

**UNIVERSIDADE FEDERAL DE VIÇOSA**

**Dosagem alélica em *Coffea arabica*: evidências de pareamento homeólogo e identificação de QTL associados à resistência a *Hemileia vastatrix***

Danúbia Rodrigues Alves  
*Doctor Scientiae*

**VIÇOSA - MINAS GERAIS  
2024**

**DANÚBIA RODRIGUES ALVES**

**Dosagem alélica em *Coffea arabica*: evidências de pareamento homeólogo e identificação de QTL associados à resistência a *Hemileia vastatrix***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

Orientadora: Eveline T. Caixeta Moura

Coorientadores: Guilherme da S. Pereira  
Dênia Pires de Almeida

**VIÇOSA - MINAS GERAIS  
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

A474d  
2024  
Alves, Danúbia Rodrigues, 1992-  
Dosagem alélica em *Coffea arabica*: evidências de  
pareamento homeólogo e identificação de OTL associados à  
resistência a *Hemileia vastatrix* / Danúbia Rodrigues Alves. –  
Viçosa, MG, 2024.

1 tese eletrônica (89 f.): il. (algumas color.).

Orientador: Eveline Teixeira Caixeta.

Tese (doutorado) - Universidade Federal de Viçosa,  
Departamento de Biologia Geral, 2024.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2024.834>

Modo de acesso: World Wide Web.

1. Café - Melhoramento genético. 2. Café - Resistência a  
doenças e pragas - Aspectos genéticos. 3. Marcadores genéticos.  
4. *Hemileia vastatrix*. I. Caixeta, Eveline Teixeira, 1972-.  
II. Universidade Federal de Viçosa. Departamento de Biologia  
Geral. Programa de Pós-Graduação em Genética e  
Melhoramento. III. Título.

CDD 22. ed. 633.732

**DANÚBIA RODRIGUES ALVES**

**Dosagem alélica em *Coffea arabica*: evidências de pareamento homeólogo e identificação de QTL associados à resistência a *Hemileia vastatrix***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 11 de setembro de 2024.

Assentimento:

---

Danúbia Rodrigues Alves  
Autora

---

Eveline Teixeira Caixeta Moura  
Orientadora

Essa tese foi assinada digitalmente pela autora em 02/01/2025 às 11:00:00 e pela orientadora em 03/01/2025 às 09:23:40. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **PN4B.PBJU.S6DQ** e clique no botão 'Validar documento'.

*A Deus, por guiar meus passos,  
Aos meus pais, Irene e Pedro,  
E ao Arthur,  
Por todo amor e dedicação,  
Dedico.*

## **AGRADECIMENTOS**

A Deus, por iluminar o meu caminho.

Aos meus pais, Irene e Pedro, por todo o amor e por acreditarem nos meus sonhos.

Ao Arthur pelo carinho, companheirismo, compreensão e incentivo. Com você ao meu lado o caminho se tornou mais fácil.

À minha família e aos amigos, que, mesmo de longe, estiveram presentes no meu coração.

À Ana Paula, pelo acolhimento e escuta atenta.

À Universidade Federal de Viçosa e ao Programa de Pós Graduação em Genética e Melhoramento, pela qualidade do ensino.

Ao Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), pela infraestrutura disponibilizada à realização do trabalho.

À minha orientadora, Dra. Eveline Teixeira Caixeta, pelos ensinamentos, confiança e incentivo.

À minha coorientadora, Dra. Dênia Pires de Almeida pelos ensinamentos transmitidos, pelas contribuições para este trabalho e pela amizade.

Ao meu coorientador, Prof. Guilherme da Silva Pereira pela disponibilidade e pelas contribuições para esse trabalho.

Ao Prof. Cosme Damião Cruz, à Prof. Iara Gonçalves dos Santos, ao Dr. Luiz Filipe Protasio Pereira e à Dra. Isabel Samila Lima Castro, pelas contribuições para a melhoria desta tese.

Ao Edson, à Bruna e ao Rafael pelas contribuições neste trabalho.

Aos amigos do Biocafé, pelo apoio, pela amizade e por tornarem os meus dias mais agradáveis.

Aos secretários do PPGGM, Marco Túlio e Odilon, pela dedicação e boa vontade em nos ajudar.

Aos órgãos de fomento: Embrapa Café, Consórcio de Pesquisa Café (CBP&C), Instituto Nacional de Ciência, Tecnologia e Inovação de Café (INCT/Café), Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A todos que, de alguma forma, contribuíram para a minha formação pessoal e profissional, meus sinceros agradecimentos.

*“É preciso sair da ilha para ver a ilha.”*  
José Saramago

## RESUMO

ALVES, Danúbia Rodrigues, D.Sc., Universidade Federal de Viçosa, setembro de 2024. **Dosagem alélica em *Coffea arabica*: evidências de pareamento homeólogo e identificação de QTL associados à resistência a *Hemileia vastatrix*.** Orientadora: Eveline Teixeira Caixeta Moura. Coorientadores: Guilherme da Silva Pereira e Dênia Pires de Almeida.

Entre as 141 espécies do gênero *Coffea*, *C. arabica* e *C. canephora* são as mais cultivadas devido à sua relevância econômica, sendo o café arábica responsável por cerca de 57% da produção mundial. A espécie *C. arabica* é um aloploide resultante da hibridação natural entre *C. eugenioides* e *C. canephora*, seguida da duplicação de todo o genoma. Evidências moleculares recentes sugerem a ocorrência de rearranjos genômicos e trocas homeólogas no genoma da espécie, destacando a complexidade genética desse poliploide. Contudo, a base genética das cultivares de *C. arabica* é estreita, devido à sua origem recente e a gargalos evolutivos, o que compromete sua resistência a estresses bióticos e abióticos. A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* é uma das doenças que mais causa prejuízo na produção do café arábica. O plantio de cultivares resistentes é a estratégia mais eficaz para controlar a doença, mas o alto potencial evolutivo do patógeno frequentemente supera a resistência das cultivares desenvolvida. Os programas de melhoramento genético de *C. arabica* têm se concentrado na seleção de progênies superiores e na transferência de genes de resistência durável, com destaque para o Híbrido de Timor (HdT). No Capítulo 1, foi apresentada uma revisão bibliográfica abordando temas relevantes que contribuiram para o embasamento e a elaboração dos capítulos subsequentes. No capítulo 2, foram investigados a dosagem alélica e possíveis pareamentos homeólogos em *C. arabica*, utilizando genotipagem em larga escala com marcadores SNP em população F2. Os resultados indicaram menores distorções de segregação dos marcadores ao utilizar a chamada de variantes tetraploides em comparação com chamadas diploides, além de evidências para um suposto pareamento não preferencial entre os dois subgenomas que constituem a espécie. O capítulo 3, teve como objetivo identificar QTL relacionados à resistência de *C. arabica* às raças I, II, XXXIII e ao patótipo 001 de *H. vastatrix*, utilizando genotipagem em larga escala com marcadores SNP e mapeamento de QTL por análise de marcas individuais. Foram identificados QTL associadas à resistência às diferentes raças de *H. vastatrix*, incluindo marcas inéditas para a resistência à raça XXXIII e novas marcas para as raças I, II e o patótipo 001. As descobertas desta pesquisa têm implicações

significativas para programas de melhoramento de *C. arabica*. Elas fornecem novas informações sobre o genoma da espécie, evidenciando a relevância de considerar a dosagem alélica e o pareamento homeólogo em estudos sobre esse alotetraploide, além de evidenciar regiões genômicas associadas à resistência a diferentes raças de *H. vastatrix*.

Palavras-chave: aloploidie; marcador molecular; melhoramento do cafeeiro; ferrugem do cafeeiro; snp

## ABSTRACT

ALVES, Danúbia Rodrigues, D.Sc., Universidade Federal de Viçosa, September, 2024. **Allele dosage in *Coffea arabica*: evidence of homeologous pairing and identification of QTL associated with resistance to *Hemileia vastatrix*.** Adviser: Eveline Teixeira Caixeta Moura. Co-advisers: Guilherme da Silva Pereira and Dênia Pires de Almeida.

Among the 141 species of the genus *Coffea*, *C. arabica* and *C. canephora* are the most cultivated due to their economic relevance, with arabica coffee accounting for approximately 57% of world production. The species *C. arabica* is an allopolyploid resulting from the natural hybridization between *C. eugenioides* and *C. canephora*, followed by duplication of the entire genome. Recent molecular evidence suggests the occurrence of genomic rearrangements and homeologous exchanges in the genome of the species, highlighting the genetic complexity of this polyploid. However, the genetic base of *C. arabica* cultivars is narrow, due to their recent origin and evolutionary bottlenecks, which compromises their resistance to biotic and abiotic stresses. Coffee rust, caused by the fungus *Hemileia vastatrix*, is one of the diseases that causes the greatest damage to arabica coffee production. Planting resistant cultivars is the most effective strategy for controlling the disease, but the pathogen's high evolutionary potential often overcomes the resistance of the cultivars developed. Genetic breeding programs for *C. arabica* have focused on the selection of superior progenies and the transfer of durable resistance genes, with emphasis on the Timor Hybrid (HdT). Chapter 1 presents a literature review covering relevant topics that contributed to the basis and preparation of subsequent chapters. Chapter 2 investigates allelic dosage and possible homeologous pairings in *C. arabica* using large-scale genotyping with SNP markers in an F2 population. The results indicate lower segregation distortions of markers when using tetraploid variant calling compared to diploid calls, in addition to evidence for a supposed non-preferential pairing between the two subgenomes that constitute the species. Chapter 3 aimed to identify QTL related to *C. arabica* resistance to races I, II, XXXIII, and pathotype 001 of *H. vastatrix*, using large-scale genotyping with SNP markers and QTL mapping by analysis of individual markers. QTL associated with resistance to different races of *H. vastatrix* were identified, including novel markers for resistance to race XXXIII and new markers for races I, II and pathotype 001. The findings of this research have significant implications for breeding programs of *C. arabica*. They provide new information about the genome of the species, highlighting the relevance of considering allelic dosage and homeologous pairing in studies on this

allotetraploid, in addition to highlighting genomic regions associated with resistance to different races of *H. vastatrix*.

Keywords: allopolyploid; molecular marker; coffee breeding; coffee rust; snp

## Sumário

<b>CAPÍTULO 1: REVISÃO BIBLIOGRÁFICA .....</b>	<b>11</b>
1.1. Gênero <i>Coffea</i> .....	12
1.2. Espécie <i>Coffea arabica</i> .....	13
1.3. Melhoramento de <i>Coffea arabica</i> visando resistência a <i>Hemileia vastatrix</i> .....	14
1.4. Mapeamento genético e identificação de QTL em <i>Coffea</i> .....	16
1.5. Mapeamento genético e identificação de QTL em poliploides .....	18
2. Referências Bibliográficas .....	20
<b>CAPÍTULO 2: DOSAGEM ALÉLICA E POTENCIAL PAREAMENTO HOMEÓLOGO EM <i>Coffea arabica</i> .....</b>	<b>32</b>
1. Introdução.....	35
2. Material e métodos .....	37
2.1. População experimental.....	37
2.2. Genotipagem .....	38
2.2.1. Extração de DNA.....	38
2.2.2. Genotipagem com marcadores SNP .....	38
2.3. Chamadas de variantes e dosagem alélica .....	39
3. Resultados .....	40
4. Discussão .....	46
5. Referências Bibliográficas .....	51
<b>CAPÍTULO 3: IDENTIFICAÇÃO DE QTL ASSOCIADOS À RESISTÊNCIA A <i>Hemileia vastatrix</i> EM <i>Coffea arabica</i> COM ANÁLISE INÉDITA PARA A RAÇA XXXIII DO PATÓGENO .....</b>	<b>59</b>
1. Introdução.....	63
2. Material e métodos .....	66
2.1. População experimental.....	66
2.2. Avaliações fenotípicas.....	66
2.3. Genotipagem .....	67
2.3.1. Extração de DNA.....	67
2.3.2. Genotipagem com marcadores SNP .....	68
2.4. Análise dos dados genômicos.....	68
2.5. Identificação de QTL .....	69
3. Resultados .....	70
4. Discussão .....	73
5. Referências Bibliográficas .....	77
<b>Material Suplementar .....</b>	<b>87</b>

## **CAPÍTULO 1**

### **REVISÃO BIBLIOGRÁFICA**

### 1.1. Gênero *Coffea*

O gênero *Coffea* (família Rubiaceae) possui 141 espécies/táxons (GUYOT et al., 2020). Apesar do grande número de espécies relatadas, apenas duas são mais cultivadas, por apresentarem maior importância econômica, *C. arabica* e *C. canephora* (SAKIYAMA, 2015; DAVIS et al., 2019). A espécie *C. arabica* ocorre naturalmente na Etiópia e no Sudão do Sul. Já as plantas selvagens da espécie *C. canephora* ocorrem em grande parte na África tropical úmida (DAVIS et al., 2006, 2019). O café arábica (*C. arabica*) corresponde em média 57% do café produzido no mundo e o café Robusta e café Conilon, variedades da espécie *C. canephora*, abrange os 43% restantes. A produção prevista para a safra 2024/2025 é de 99,9 milhões de sacas de 60kg para *C. arabica* e 76,4 milhões de sacas de 60kg para *C. canephora* (CONAB, 2024; MAPA, 2024). O café arábica é mais apreciado pelos consumidores, devido ao menor amargor em seu sabor e menor teor de cafeína (LEROY et al., 2006; WANG et al., 2020). Desse modo, o café arábica é vendido a um preço superior ao *C. canephora* (CONAB, 2024; MAPA, 2024).

Nos últimos 30 anos, uma demanda progressiva resultou na expansão da produção e da exportação de café, com um valor estimado em mais de 200 bilhões de dólares anuais (USDA, 2022). Atualmente, o café, principal commodity dos países tropicais, é cultivado principalmente como produto de exportação em várias regiões tropicais e serve como importante fonte de renda e emprego em países em desenvolvimento da América Latina, África e Ásia. A cafeicultura tem grande importância socioeconômica no Brasil, que atualmente é líder nas exportações mundiais e está entre os maiores países produtores e consumidores (CONAB, 2024; MAPA, 2024; USDA, 2024). Entretanto, por ser uma cultura perene e cultivada em grandes áreas, ela está sujeita a riscos iminentes que afetam diretamente a produtividade e, conseqüentemente, a economia do país. Assim, a cafeicultura brasileira é caracterizada por utilizar tecnologias que garantem a quantidade e qualidade do produto no mercado, contrabalançando as adversidades encontradas (TSOUROS et al., 2019; CAMPUZANO-DUQUE & BLAIR et al., 2022).

Perturbações à cadeia produtiva provocadas por mudanças climáticas, crises sanitária, instabilidade política e, atualmente, conseqüências das interrupções socioeconômicas provocada pela pandemia da covid-19, estão influenciando as tendências e padrões do consumo de café no mundo (RHINEY et al., 2021; CAMPUZANO-DUQUE & BLAIR et al., 2022). Além disso, os produtores de café enfrentam desafios provocados por pragas e doenças na lavoura, dentre elas a ferrugem do cafeeiro, uma das doenças que mais causa prejuízo na produção do café arábica (HARELIMANA et al., 2022; ZAMBOLIM et al., 2024). De tal modo,

é de suma importância a cooperação entre os setores público e privado no enfrentamento dos desafios da cafeicultura. São necessários investimentos adicionais no avanço da pesquisa cafeeira para lidar com os desafios e conseguir atender à crescente demanda por café (KRISHNAN et al., 2021; RHINEY et al., 2021; CAMPUZANO-DUQUE & BLAIR et al., 2022).

## 1.2. Espécie *Coffea arabica*

A espécie *C. arabica* é um aloploidio natural que se reproduz principalmente por autopolinização. No gênero *Coffea*, apenas *C. arabica* é poliploide, apresentando  $2n = 4x = 44$  cromossomos. As outras 140 espécies são diploides ( $2n = 2x = 22$  cromossomos) (DAVIS et al., 2021). A origem de *C. arabica* supostamente ocorreu entre 350.000 e 610.000 anos (SALOJÄRVI et al., 2024), a partir da hibridação natural entre duas espécies diploides, *C. eugenioides* ( $2n = 2x = 22$  cromossomos) e *C. canephora* ( $2n = 2x = 22$  cromossomos) (LASHERMES, et al., 1999), seguida de uma duplicação de todo o genoma, provavelmente durante um único evento de aloploidização (CLARINDO & CARVALHO, 2009; LASHERMES, et al., 2014; SCALABRIN et al., 2020; BAWIN et al., 2021).

A Etiópia é considerada o principal centro de origem e diversificação do café arábica (SYLVAIN, 1955). Cultivares que tiveram suas origens genéticas nas variedades botânicas Típica e, ou Bourbon, foram a base para o cultivo de café arábica na América do Sul e Central (KRUG et al., 1939; CARVALHO et al., 1989; SETOTAW et al., 2013). Posteriormente, os híbridos interespecíficos Híbrido de Timor (HdT) e Icatu (derivados do cruzamento natural de *C. arabica* com *C. canephora*) também contribuíram para a base genética das cultivares modernas de café arábica (SETOTAW et al., 2013; SCALABRIN et al. 2020).

Entretanto, a base genética das cultivares de *C. arabica*, utilizadas atualmente nos diversos países produtores, tem sido descrita como sendo estreita, devido à sua recente origem alotetraploide e fortes gargalos durante sua história evolutiva (SETOTAW et al., 2020; SCALABRIN et al., 2020; SALOJÄRVI et al., 2024). Calcula-se que 97,55% da base genética das cultivares brasileiras de *C. arabica* é derivada de apenas 13 linhagens ancestrais, assim, a maioria das cultivares lançadas são linhagens irmãs com histórico genético semelhante (SETOTAW et al., 2020). A baixa diversidade das cultivares da espécie tem sido reconhecida como uma limitação significativa da resistência/tolerância das variedades a estresses bióticos e abióticos (ANTHONY et al., 2002), como por exemplo a ferrugem do cafeeiro causada pelo

fungo biotrófico *H. vastatrix*, uma das doenças mais devastadoras dos cafezais mundialmente (HARELIMANA et al., 2022; SILVA et al., 2022a; ZAMBOLIM et al., 2024).

### 1.3. Melhoramento de *Coffea arabica* visando resistência a *Hemileia vastatrix*

O melhoramento genético do cafeeiro tem contribuições marcantes no desenvolvimento agrônomo e tecnológico da cafeicultura, por meio da obtenção de cultivares portadoras de elevado potencial produtivo, aliado a outras características de interesse, como qualidade do grão, qualidade da bebida e resistência/tolerância a estresses bióticos e abióticos (SAKIYAMA, 2015; MELESE & KOLECH, 2021). Diferentes métodos de melhoramento têm sido aplicados em diversos países produtores de café ao redor do mundo com o objetivo de melhorar o rendimento e a qualidade da cultura. A aplicação dos métodos varia dependendo da quantidade de variação genética disponível, das condições ecológicas e do objetivo final dos programas de melhoramento de *Coffea* (MELESE & KOLECH, 2021).

Um dos objetivos dos programas de melhoramento do cafeeiro é a busca por cultivares resistentes à principal doença da cultura, à ferrugem do cafeeiro. Essa doença é causada pelo fungo biotrófico *H. vastatrix* (ordem Uredinales, família Pucciniaceae) que infecta exclusivamente o gênero *Coffea*, acarretando perdas de produção de mais de perda de 30-50% na produção (ZAMBOLIM, 2016; ZAMBOLIM et al., 2024). A maioria das variedades de *C. arabica* é suscetível a esse fungo, enquanto que a espécie *C. canephora* apresenta variedades resistentes (CARVALHO et al., 2010; SILVA et al., 2022a).

Até o momento foram caracterizadas mais de 50 raças de *H. vastatrix*, sendo identificadas no Brasil 16 diferentes raças (I, II, III, VII, X, XIII, XV, XVI, XVII, XXI, XXII, XXIII, XXIV, XXV ou XXXI, XXXIII, XXXVII). A raça II está presente em maior frequência nas lavouras de café do Brasil e a raça XXXIII tem sido reportada como a mais complexa, contendo dois ou três genes de virulência (VARZEA et al., 2005; ZAMBOLIM et al., 2005; CAPUCHO et al., 2012; ZAMBOLIM & CAIXETA, 2021; ZAMBOLIM et al., 2024). O plantio de cultivares resistentes é a medida mais eficaz e econômica de controle da ferrugem, embora tratamentos químicos, como o uso de fungicidas, possam ser realizados de forma preventiva (WALLER et al., 2007; ZAMBOLIM, 2016; ZAMBOLIM et al., 2024).

Na busca por cultivares de café resistentes a essa doença, já foram identificadas diferentes fontes de resistência. A resistência à ferrugem do cafeeiro é governada por pelo menos nove

genes dominantes (S<sub>H1</sub> a S<sub>H9</sub>) de acordo, com a teoria gene a gene proposta por Flor (1971). Os genes S<sub>H1</sub>, S<sub>H2</sub>, S<sub>H4</sub> e S<sub>H5</sub> foram identificados, individualmente ou em combinação, em cafeeiros da espécie *C. arabica*, enquanto S<sub>H6</sub>, S<sub>H7</sub>, S<sub>H8</sub> e S<sub>H9</sub> em *C. canephora* e S<sub>H3</sub> em *C. liberica* (NORONHA-WAGNER & BETTENCOURT, 1967; BETTENCOURT & NORONHA-WAGNER, 1969; BETTENCOURT et al., 1980, 1988). Além disso, outros genes de resistência também foram identificados (gene CC-NBS-LRR e gene HdT\_LRR\_RLK2) (BARKA et al., 2020; ALMEIDA et al., 2021b).

O grande obstáculo para o melhoramento de *Coffea* buscando resistência a *H. vastatrix* é o potencial evolutivo das populações do fungo. Estes são capazes de suplantar a resistência das cultivares desenvolvidas, já que ao longo do tempo pode haver a ocorrência de novas raças do patógeno, para as quais as plantas não possuem resistência. Um exemplo disso são o S<sub>H1</sub>, S<sub>H2</sub>, S<sub>H4</sub> e S<sub>H5</sub> que se encontram em diferentes variedades comerciais. Esses genes já foram suplantados pelos patógenos que ocorrem no Brasil, representando assim, um desafio para os melhoristas na busca de resistência durável a ferrugem do cafeeiro (BETTENCOURT, 1973; VÁRZEA & MARQUES, 2005; ZAMBOLIM et al., 2005; DINIZ, et al., 2012; VAN DER VOSSSEN et al., 2015).

Por muitos anos, a seleção para resistência a *H. vastatrix* foi baseada na resistência completa, altamente específica, originada dos principais genes S<sub>H</sub> introgrididos, derivados de *C. arabica*, *C. canephora* ou *C. liberica*. No entanto, o alto grau de variabilidade e adaptabilidade do patógeno resultou na perda gradual da resistência baseada nos principais genes S<sub>H</sub>. Os genes S<sub>H</sub> de *C. arabica* já foram suplantados pelas raças de *H. vastatrix* existentes no Brasil (BETTENCOURT, 1973; DINIZ, et al., 2012). A resistência poligênica parcial e não específica a *H. vastatrix* foi identificada em *C. canephora*, em alguns genótipos de *C. arabica* e em híbridos interespecíficos (HERRERA et al., 2009; ROMERO et al., 2010). Assim, os programas de melhoramento do cafeeiro buscam desenvolver estratégias para melhorar a durabilidade da resistência à *H. vastatrix* em cultivares de café arábica, através da combinação de resistência qualitativa (completa ou raça-específica) e quantitativa (parcial ou não raça específica) (BARKA et al., 2020; ALMEIDA et al., 2021b; SERA et al., 2022; SILVA et al., 2022a; ALVES et al., 2024).

Os programas de melhoramento de *C. arabica* possuem linhas de pesquisa visando a seleção de progênies superiores derivadas de cruzamentos intra e interespecíficos para a transferência de genes de resistência à ferrugem do cafeeiro (SAKIYAMA, 2005; SERA et al., 2022). Os genes S<sub>H6</sub>, S<sub>H7</sub>, S<sub>H8</sub> e S<sub>H9</sub>, identificados em *C. canephora*, estão presentes no Híbrido de Timor

(HdT), um híbrido natural entre *C. arabica* e *C. canephora*. O HdT e as progênies derivadas do seu cruzamento com outras cultivares, vêm sendo estudadas em diversas regiões produtoras de café no mundo, devido a sua resistência a *H. vastatrix*. O seu germoplasma tem sido utilizado em programas de melhoramento que visam a resistência durável à ferrugem e outras doenças do cafeeiro (CAPUCHO et al., 2009; ALMEIDA et al., 2021a, 2021b; SERA et al., 2022; SILVA et al., 2006, 2022a; ALVES et al., 2024).

O uso eficiente dos recursos disponíveis nesses híbridos naturais resistentes é fundamental para o melhoramento contínuo das variedades de *C. arabica*, já que o HdT, além de apresentar resistência à ferrugem se cruza facilmente com as cultivares de *C. arabica*, favorecendo a transferência de sua resistência (CARVALHO et al., 1989; LASHERMES et al., 2009; SILVA et al., 2022a; ZAMBOLIM et al., 2024). Assim, os programas de melhoramento do cafeeiro têm utilizado os HdT como fonte de resistência aliada a estratégia de seleção assistida por marcadores (SAM) e piramidação de genes, para se obter resistência durável a *H. vastatrix*. Um dos métodos mais utilizados para realizar a SAM é o mapeamento genético e identificação de QTL (*Quantitative Trait Loci*) associados a característica de interesse (ALKIMIM et al., 2017; ALMEIDA et al., 2021a; MEDEIROS et al., 2021; SAAVEDRA et al., 2023; FEITOSA et al., 2024).

#### **1.4. Mapeamento genético e identificação de QTL em *Coffea***

O mapeamento genético associado ao estudo de QTL ligados a características de interesse é uma abordagem de fundamental importância para os estudos genéticos, incluindo estudo de herança genética, estudos de evolução comparativa, da organização dos cromossomos e dos processos biológicos relacionados a genes de interesse. Entre as aplicações dessas técnicas nos programas de melhoramento, se encontram a introgressão de características de interesse em cultivares comerciais, escolha de genitores superiores, seleção assistida de plantas e clonagem de genes (ALMEIDA et al., 2021a, 2021b; ARORA et al., 2021; XIONG et al., 2021; WANG et al., 2023; ZHAO et al., 2023).

A expansão das pesquisas genômicas, possibilitada pelo desenvolvimento de novas ferramentas biotecnológicas, aumentou em larga escala a quantidade de informações biológicas e moleculares disponíveis (EGAN et al., 2012; MICHNO & STUPAR, 2018). Os métodos de Sequenciamento de Nova Geração (NGS) permitiram a difusão do processo de sequenciamento, gerando milhares de sequências de DNA em um único ensaio e reduzindo o custo de *data-point*.

À medida que os custos e as capacidades dessas tecnologias continuam evoluir, novos campos de estudo continuam sendo abertos, permitindo pesquisadores a obter uma visão cada vez mais profunda da região genômica alvo (JAGANATHAN et al., 2020; WANG et., 2023; ZHAO et al., 2023).

Os avanços nas plataformas de genotipagem com marcadores SNP (*Single Nucleotide Polymorphism*) de alto rendimento consolidaram-se como ferramentas poderosas para a genotipagem em todo o genoma de organismos, graças à sua elevada eficiência e ao baixo custo por marcador. Essas plataformas demonstraram ser eficazes para análise de ligação, mapeamento de QTL e análise de associação em todo o genoma. O mapeamento de QTL com marcadores SNP é o caminho para uma caracterização molecular mais detalhada e estudos funcionais da relação entre polimorfismo e variação de características. Dessa forma, a genotipagem com marcadores SNP tornou-se altamente relevante para os programas de melhoramento genético, facilitando o estudo de QTL e integrando-o de maneira mais eficiente ao melhoramento assistido por marcadores e à seleção genômica (COCKRAM & MACKAY, 2018; YOU et al., 2019; WANG et., 2023; ZHAO et al., 2023).

Em trabalhos anteriores, mapas genéticos de *Coffea* foram construídos usando marcadores moleculares desenvolvidos para espécies diploides, como RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) e SSR (*Simple Sequence Repeat*) (LASHERMES et al., 2001; PEARL et al., 2004; PRAKASH et al., 2004; TEIXEIRA-CABRAL et al., 2004; PRIOLLI et al., 2008; ROMERO et al., 2014; PESTANA et al., 2015; FERRÃO et al., 2015; MONCADA et al., 2016). Nesses trabalhos, apesar dos mapas serem pouco saturados, foram identificados QTL associados a características de importância para os programas de melhoramento do cafeeiro. Prakash et al. (2004) desenvolveram os primeiros marcadores específicos de locus co-segregando com S<sub>H3</sub>, derivados da introgressão em *C. arabica*. No trabalho desenvolvido por Pestana et al. (2015), foi construído um mapa genético de ligação para *C. arabica*, composto por 111 marcadores (RAPD, AFLP, SSR), distribuídos em 12 grupos de ligação, cobrindo 976,8 cM do genoma. Neste mapa, foram identificadas duas regiões nos grupos de ligações 2 e 10 que correspondem a dois loci associados à resistência do cafeeiro ao patótipo 001 e às raças I e II de *H. vastatrix*.

Há dois mapas genéticos utilizando SNP para identificação de QTL associado a características de interesse para o melhoramento de *C. arabica*. Em um dos estudos, um mapa genético com 22 grupos de ligação, construído a partir 848 marcadores SNP e SSR, foi usado para identificar QTL associados à produção de café, altura da planta e tamanho do grão

(MONCADA et al., 2016). No outro estudo o mapa de ligação com 11 grupos de ligação foi constituído de 699 marcadores SNP para a análise de QTL associado a resistência à *Coffee Berry Disease* (CBD) (GIMASE et al., 2020).

Entretanto, não existem estudos que adotem abordagens que considerem as peculiaridades do genoma poliploide dessa espécie, no mapeamento genético e na identificação de QTL. De tal modo, a utilização de metodologias direcionadas para estudos genéticos de espécies diploides em *C. arabica* podem gerar resultados irrealistas, incluindo distorção de segregação de marcadores moleculares, devido a uma limitação intrínseca desses métodos que está associada à negligência do nível de ploidia (GARCIA et al., 2013; GERARD et al., 2024) e possíveis trocas homeólogas nessa espécie alotetraploide (LASHERMES et al., 2014, 2016; ORTIZ & SHARBROUGH, 2024; SALOJARVI et al. 2024). Destacando a importância de entender a complexidade do genoma dessa de *C. arabica*.

### **1.5. Mapeamento genético e identificação de QTL em poliploides**

Embora as espécies poliploides sejam de grande relevância na agricultura, o mapeamento genético e a identificação de QTL relacionados a características agrônomicas de interesse em culturas poliploides progrediram de forma mais lenta em comparação às culturas diploides. Esse atraso é amplamente atribuído à complexidade inerente à organização genômica dos poliploides (GARCIA et al., 2013; GERARD et al., 2024). Tradicionalmente, os estudos genéticos em poliploides concentravam-se em marcadores desenvolvidos para espécies diploides, como AFLP e SSR. Esses marcadores ganharam ampla popularidade em diversas espécies vegetais, incluindo poliploides, devido à sua capacidade de serem aplicados sem a necessidade de informações prévias de sequência, e se destacavam-se por sua alta reprodutibilidade em diferentes laboratórios, tornando-se ferramentas essenciais para análises genéticas em plantas (CORDEIRO et al., 2000; MCGREGOR et al., 2000; MCNEIL et al., 2011).

A genotipagem com marcadores SNP avançou de forma significativa em diversas culturas poliploides, como morango, rosa, batata, trigo e crisântemo. (BASSIL et al., 2015; KONING-BOUCOIRAN et al., 2015; VOS et al., 2015; WINFIELD et al., 2016; VAN GEEST et al., 2017). Esses marcadores moleculares têm se mostrado ferramentas valiosas para estudos de ligação, mapeamento de QTL de características importantes e análises de associação em nível de genoma completo em espécies poliploides (YOU et al., 2019; WANG et al., 2022). Ao

identificar SNP em todo o genoma e realizar a chamada de doses, é possível avaliar a frequência de cada alelo no conjunto poliploide, aumentando substancialmente a quantidade de informações disponíveis em cada locus, proporcionando uma visão mais abrangente das variações genéticas presentes no DNA dessas espécies (VOORRIPS et al., 2011; HACKETT et al., 2013; GARCIA et al., 2013; GERARD et al., 2018).

A dosagem de alelo refere-se ao número de cópias de um alelo específico em um locus gênico, comparado ao alelo de referência. Em espécies tetraploides, por exemplo, onde "A" é o alelo de referência, a dosagem pode variar desde aaaa (nulliplex), passando por Aaaa (simplex) até AAAA (quadruplex). O principal desafio ao trabalhar com poliploides é diferenciar corretamente os tipos de heterozigotos, como Aaaa, AAaa e AAAa (GARCIA et al., 2013; ANNICCHIARICO et al., 2015; BIAZZI et al., 2017). Usando sistemas de marcadores tradicionais como AFLPs ou SSRs, os genótipos são frequentemente classificados de forma binária: 0 para ausência de banda (aaaa) ou 1 para presença de banda (Aaaa, AAaa, AAAa e AAAA), sem considerar o nível de ploidia. Portanto, sistemas de marcadores que possibilitam a distinção clara de todos os alelos presentes nos genótipos são mais informativos para o estudo de genomas poliploides. No entanto, isso requer a aplicação de métodos estatísticos especializados, a fim de permitir uma análise e interpretação detalhadas dos dados nesse contexto complexo (GARCIA et al., 2013).

Além disso, pesquisas moleculares recentes sugerem que rearranjos genômicos, envolvendo trocas homeólogas, desempenharam um papel crucial na evolução de genomas aloploides, possivelmente ocorrendo nas fases iniciais do desenvolvimento dessas espécies, por meio de pareamentos multivalentes ou menos restritos durante a meiose (LASHERMES et al., 2014; SCALABRIN et al., 2020, 2024). Embora a detecção dessas variações seja desafiadora devido à complexidade das alterações genômicas, marcadores moleculares modernos têm sido valiosos para esses estudos (LASHERMES et al., 2014, 2016; XIONG et al., 2021; CHU et al., 2021; SCALABRIN et al., 2020, 2024). No entanto, assim como ocorre com os estudos de dosagem alélica, o pareamento homeólogo em *C. arabica* ainda é uma área pouco explorada nas abordagens de mapeamento de QTL, devido à complexidade do genoma poliploide da espécie (SCALABRIN et al., 2024; SALOJARVI et al., 2024), além da escassez de modelos estatísticos que integrem o pareamento preferencial em análises com aloploides (GERARD et al., 2024).

## 2. Referências Bibliográficas

- ALKIMIM, E. R., CAIXETA, E. T., SOUSA, T. V., PEREIRA, A. A., DE OLIVEIRA, A. C. B., ZAMBOLIM, L., & SAKIYAMA, N. S. (2017). Marker-assisted selection provides arabica coffee with genes from other *Coffea* species targeting on multiple resistance to rust and coffee berry disease. *Molecular Breeding*, 37, 1-10.
- ALMEIDA, D. P., CAIXETA, E. T., MOREIRA, K. F., DE OLIVEIRA, A. C. B., DE FREITAS, K. N. P., PEREIRA, A. A., ... & CRUZ, C. D. (2021a). Marker-assisted pyramiding of multiple disease resistance genes in coffee genotypes (*Coffea arabica*). *Agronomy*, 11(9), 1763.
- ALMEIDA, D. P., CASTRO, I. S. L., MENDES, T. A. D. O., ALVES, D. R., BARKA, G. D., BARREIROS, P. R. R. M., ... & CAIXETA, E. T. (2021b). Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. *Scientia Agricola*, 78.
- ALVES, D. R., ALMEIDA, D. P., DE ANDRADE SILVA, E. M., CASTRO, I. S. L., BARREIROS, P. R. R. M., DE OLIVEIRA MENDES, T. A., ... & CAIXETA, E. T. (2024). Unravelling the role of key genes involved coffee leaf rust resistance. *Current Plant Biology*, 38, 100347.
- ANNICCHIARICO, P., BARRETT, B., BRUMMER, E. C., JULIER, B., & MARSHALL, A. H. (2015). Achievements and challenges in improving temperate perennial forage legumes. *Critical Reviews in Plant Sciences*, 34(1-3), 327-380.
- ANTHONY, F., COMBES, M. C., ASTORGA, C., BERTRAND, B., GRAZIOSI, G., & LASHERMES, P. (2002). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theoretical and Applied Genetics*, 104, 894-900.
- ARORA, S., KAUR, S., DHILLON, G. S., SINGH, R., KAUR, J., SHARMA, A., & CHHUNEJA, P. (2021). Introgression and genetic mapping of leaf rust and stripe rust resistance in *Aegilops triuncialis*. *Journal of Genetics*, 100, 1-11.

BARKA, G. D., CAIXETA, E. T., FERREIRA, S. S., & ZAMBOLIM, L. (2020). In silico guided structural and functional analysis of genes with potential involvement in resistance to coffee leaf rust: A functional marker based approach. *PloS one*, 15(7), e0222747.

BAWIN, Y., RUTTINK, T., STAELENS, A., HAEGEMAN, A., STOFFELEN, P., MWANGA MWANGA, J. C. I., ... & JANSSENS, S. B. (2021). Phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica*. *Journal of Systematics and Evolution*, 59(5), 953-963.

BETTENCOURT, A. J., & NORONHA WAGNER, H. (1969). Genetic factors conditioning resistance of *Coffea arabica* L. to *Hemileia vastatrix* Berk. et Br. *Agronomia Lusitana (Portugal)* v. 31 p. 285-292.

BETTENCOURT, A. J. (1973). Considerações gerais sobre o Híbrido de Timor. Instituto Agronómico, Sao Paulo (Brasil).

BETTENCOURT, A. J.; NORONHA-WAGNER, M.; LOPES, M. (1980), Factor genético que condiciona a resistência do clone 1343/269 (Híbrido de Timor) à *Hemileia vastatrix* Berk. and Br. *Brotéria Genética*, v.I, n.LXXVI, p.53-58.

BETTENCOURT, A. J., RODRIGUES JUNIOR, C. J., CLARKE, R. J., & MACRAE, R. (1988). Principles and practice of coffee breeding for resistance to rust and other diseases Coffee, *Agronomy*, 4.

BASSIL, N. V., DAVIS, T. M., ZHANG, H., FICKLIN, S., MITTMANN, M., WEBSTER, T., ... & VAN DE WEG, E. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry *Fragaria × ananassa*. *BMC genomics*, 16, 1-30.

BIAZZI, E., NAZZICARI, N., PECETTI, L., BRUMMER, E. C., PALMONARI, A., TAVA, A., & ANNICCHIARICO, P. (2017). Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. *PloS one*, 12(1), e0169234.

- CAMPUZANO-DUQUE, L. F., & BLAIR, M. W. (2022). Strategies for Robusta Coffee (*Coffea canephora*) improvement as a new crop in Colombia. *Agriculture*, 12(10), 1576.
- CAPUCHO, A. S., CAIXETA, E. T., ZAMBOLIM, E. M., & ZAMBOLIM, L. (2009). Herança da resistência do Híbrido de Timor UFV 443-03 à ferrugem-do-cafeeiro. *Pesquisa Agropecuária Brasileira*, 44, 276-282.
- CAPUCHO, A. S., ZAMBOLIM, E. M., FREITAS, R. L., HADDAD, F., CAIXETA, E. T., & ZAMBOLIM, L. (2012). Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. *Australasian Plant Disease Notes*, 7, 189-191.
- CARVALHO, A., FAZUOLI, L. C., & COSTA, W. M. D. (1989). Melhoramento do cafeeiro: XLI. Produtividade do Híbrido de Timor, de seus derivados e de outras fontes de resistência a *Hemileia vastatrix*. *Bragantia*, 48, 73-86.
- CARVALHO, V. L., CHALFOUN, S. M., & CUNHA, R. D. (2010). Manejo de doenças do cafeeiro. Reis, PR; Cunha, RL *Café arábica: do plantio a colheita*. Lavras: EPAMIG, 1, 689-756.
- CHU, J., LI, W., PIAO, D., LIN, F., HUO, X., ZHANG, H., ... & ZHANG, C. (2021). Identification of a major QTL related to resistance to soybean mosaic virus in diverse soybean genetic populations. *Euphytica*, 217(9), 176.
- CLARINDO, W. R., & CARVALHO, C. R. (2009). Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. *Plant Cell Reports*, 28, 73-81.
- CONAB, C. N. D. A. (2024). Acompanhamento da Safra Brasileira - Café. [S. l.: s. n.],
- CORDEIRO, G. M., TAYLOR, G. O., & HENRY, R. J. (2000). Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. *Plant science*, 155(2), 161-168.
- COCKRAM, J., & MACKAY, I. (2018). Genetic mapping populations for conducting high-resolution trait mapping in plants. *Plant genetics and molecular biology*, 109-138.

DAVIS, A. P., GOVAERTS, R., BRIDSON, D. M., & STOFFELEN, P. (2006). An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Botanical Journal of the Linnean Society*, 152(4), 465-512.

DAVIS, A. P., CHADBURN, H., MOAT, J., O'SULLIVAN, R., HARGREAVES, S., & NIC LUGHADHA, E. (2019). High extinction risk for wild coffee species and implications for coffee sector sustainability. *Science advances*, 5(1), eaav3473.

DAVIS, A. P., & RAKOTONASOLO, F. (2021). Six new species of coffee (*Coffea*) from northern Madagascar. *Kew Bulletin*, 76(3), 497-511.

DINIZ, I., TALHINHAS, P., AZINHEIRA, H. G., VÁRZEA, V., MEDEIRA, C., MAIA, I., ... & DO CÉU SILVA, M. (2012). Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and nonhost resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. *European Journal of Plant Pathology*, 133, 141-157.

EGAN, J. B., SHI, C. X., TEMBE, W., CHRISTOFORIDES, A., KURDOGLU, A., SINARI, S., ... & STEWART, A. K. (2012). Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood, The Journal of the American Society of Hematology*, 120(5), 1060-1066.

FEITOSA, F. D. M., SANTOS, I. G. D., PEREIRA, A. A., DE OLIVEIRA, A. C. B., & CAIXETA, E. T. (2024). Gene pyramiding for achieving enhanced disease and insect multiple resistance in *Coffea arabica*. *Crop Science*. 64(5), 2736-2747.

FERRÃO, L. F. V., CAIXETA, E. T., PENA, G., ZAMBOLIM, E. M., CRUZ, C. D., ZAMBOLIM, L., ... & SAKIYAMA, N. S. (2015). New EST-SSR markers of *Coffea arabica*: transferability and application to studies of molecular characterization and genetic mapping. *Molecular Breeding*, 35, 1-5.

FLOR, H. H. (1971). Current status of the gene-for-gene concept. *Annual review of phytopathology*, 9(1), 275-296.

GARCIA, A. A., MOLLINARI, M., MARCONI, T. G., SERANG, O. R., SILVA, R. R., VIEIRA, M. L., ... & SOUZA, A. P. (2013). SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. *Scientific reports*, 3(1), 3399.

GERARD, D., FERRÃO, L. F. V., GARCIA, A. A. F., & STEPHENS, M. (2018). Genotyping polyploids from messy sequencing data. *Genetics*, 210(3), 789-807.

GERARD, D., THAKKAR, M., & FERRÃO, L. F. V. (2024). Tests for Segregation Distortion in Tetraploid F1 Populations. *bioRxiv*, 2024-02.

GIMASE, J. M., THAGANA, W. M., OMONDI, C., CHESEREK, J. J., GICHIMU, B. M., & GICHURU, E. K. (2020). Quantitative trait loci (QTL) mapping of resistance to coffee berry disease (*Colletotrichum kahawae* Waller & Bridge) in *Coffea arabica* L. variety Rume Sudan.

GUYOT, R., HAMON, P., COUTURON, E., RAHARIMALALA, N., RAKOTOMALALA, J. J., LAKKANNA, S., ... & BONNET, P. (2020). WCSdb: a database of wild *Coffea* species. *Database*, 2020, baaa069.

HACKETT, C. A., MCLEAN, K., & BRYAN, G. J. (2013). Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. *PloS one*, 8(5), e63939.

HARELIMANA, A., RUKAZAMBUGA, D., & HANCE, T. (2022). Pests and diseases regulation in coffee agroecosystems by management systems and resistance in changing climate conditions: A review. *Journal of Plant Diseases and Protection*, 129(5), 1041-1052.

HERRERA P, J. C., ALVARADO A, G., CORTINA G, H. A., COMBES, M. C., ROMERO G, G., & LASHERMES, P. (2009). Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk & Br.) introgressed into the cultivated *Coffea arabica* L. from the diploid *C. canephora* species. *Euphytica*, 167, 57-67.

JAGANATHAN, D., BOHRA, A., THUDI, M., & VARSHNEY, R. K. (2020). Fine mapping and gene cloning in the post-NGS era: advances and prospects. *Theoretical and Applied Genetics*, 133(5), 1791-1810.

KONING-BOUCOIRAN, C. F., ESSELINK, G. D., VUKOSAVLJEV, M., VAN'T WESTENDE, W. P., GITONGA, V. W., KRENS, F. A., ... & SMULDERS, M. J. (2015). Using RNA-Seq to assemble a rose transcriptome with more than 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for rose (*Rosa L.*). *Frontiers in plant science*, 6, 249.

KRISHNAN, S., MATSUMOTO, T., NAGAI, C., FALCONER, J., SHRINER, S., LONG, J., ... & VEGA, F. E. (2021). Vulnerability of coffee (*Coffea* spp.) genetic resources in the United States. *Genetic Resources and Crop Evolution*, 68(7), 2691-2710.

KRUG, C. A., MENDES, J. E. T., & CARVALHO, A. (1939). Taxonomia de *Coffea arabica* (Vol. 62). Instituto Agronómico do Estado em Campinas.

LASHERMES, P., COMBES, M. C., ROBERT, J., TROUSLOT, P., D'HONT, A., ANTHONY, F., & CHARRIER, A. (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics MGG*, 261(2), 259-266.

LASHERMES, P., COMBES, M. C., PRAKASH, N. S., TROUSLOT, P., LORIEUX, M., & CHARRIER, A. (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome*, 44(4), 589-595.

LASHERMES, P., BERTRAND, B., ETIENNE, H. (2009). Breeding coffee (*Coffea arabica*) for sustainable production. *Breeding plantation tree crops: Tropical species*, 525-543.

LASHERMES, P., COMBES, M. C., HUEBER, Y., SEVERAC, D., & DEREPPER, A. (2014). Genome rearrangements derived from homoeologous recombination following allopolyploidy speciation in coffee. *The Plant Journal*, 78(4), 674-685.

LASHERMES, P., HUEBER, Y., COMBES, M. C., SEVERAC, D., & DEREPPER, A. (2016). Inter-genomic DNA exchanges and homeologous gene silencing shaped the nascent allopolyploid coffee genome (*Coffea arabica* L.). *G3: Genes, Genomes, Genetics*, 6(9), 2937-2948.

LEROY, T., RIBEYRE, F., BERTRAND, B., CHARMETANT, P., DUFOUR, M., MONTAGNON, C., ... & POT, D. (2006). Genetics of coffee quality. *Brazilian journal of plant physiology*, 18, 229-242.

MAPA, 2024: Ministério da Agricultura e Pecuária.

MCGREGOR, C. E., LAMBERT, C. A., GREYLING, M. M., LOUW, J. H., & WARNICH, L. (2000). A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica*, 113, 135-144.

MCNEIL, M. D., HERMANN, S., JACKSON, P. A., & AITKEN, K. S. (2011). Conversion of AFLP markers to high-throughput markers in a complex polyploid, sugarcane. *Molecular Breeding*, 27, 395-407.

MEDEIROS, A. C., CAIXETA, E. T., OLIVEIRA, A. C. B. D., SOUSA, T. V., STOCK, V. D. M., CRUZ, C. D., ... & PEREIRA, A. A. (2021). Combining ability and molecular marker approach identified genetic resources to improve agronomic performance in *Coffea arabica* breeding. *Frontiers in Sustainable Food Systems*, 5, 705278.

MELESE, Y. Y., & KOLECH, S. A. (2021). Coffee (*Coffea arabica* L.): Methods, objectives, and future strategies of breeding in Ethiopia. *Sustainability*, 13(19), 10814.

MICHNO, J. M., & STUPAR, R. M. (2018). The importance of genotype identity, genetic heterogeneity, and bioinformatic handling for properly assessing genomic variation in transgenic plants. *BMC biotechnology*, 18, 1-9.

MONCADA, M. D. P., TOVAR, E., MONTOYA, J. C., GONZÁLEZ, A., SPINDEL, J., & MCCOUCH, S. (2016). A genetic linkage map of coffee (*Coffea arabica* L.) and QTL for yield, plant height, and bean size. *Tree genetics & genomes*, 12, 1-17.

NORONHA-WAGNER, M., & BETTENCOURT, A. J. (1967). Genetic study of the resistance of *Coffea* spp. to leaf rust: I. Identification and behavior of four factors conditioning disease reaction in *Coffea arabica* to twelve physiologic races of *Hemileia vastatrix*. *Canadian Journal of Botany*, 45(11), 2021-2031.

ORTIZ, A. J., & SHARBROUGH, J. (2024). Genome-wide patterns of homoeologous gene flow in allotetraploid coffee. *Applications in Plant Sciences*, 12(4), e11584.

PEARL, H. M., NAGAI, C. H. I. F. U. M. I., MOORE, P. H., STEIGER, D. L., OSGOOD, R. V., & MING, R. (2004). Construction of a genetic map for arabica coffee. *Theoretical and Applied Genetics*, 108, 829-835.

PESTANA, K. N., CAPUCHO, A. S., CAIXETA, E. T., DE ALMEIDA, D. P., ZAMBOLIM, E. M., CRUZ, C. D., ... & SAKIYAMA, N. S. (2015). Inheritance study and linkage mapping of resistance loci to *Hemileia vastatrix* in Híbrido de Timor UFV 443-03. *Tree Genetics & Genomes*, 11, 1-13.

PRAKASH, N. S., MARQUES, D. V., VARZEA, V. M. P., SILVA, M. C., COMBES, M. C., & LASHERMES, P. (2004). Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. *Theoretical and Applied Genetics*, 109, 1311-1317.

PRIOLLI, R. H., MAZZAFERA, P., SIQUEIRA, W. J., MÖLLER, M., ZUCCHI, M. I., RAMOS, L. C. S., ... & COLOMBO, C. A. (2008). Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). *Genetics and Molecular Biology*, 31, 498-504.

RHINEY, K., GUIDO, Z., KNUDSON, C., AVELINO, J., BACON, C. M., LECLERC, G., ... & BEBBER, D. P. (2021). Epidemics and the future of coffee production. *Proceedings of the National Academy of Sciences*, 118(27), e2023212118.

ROMERO G, G., ALVARADO A, G., CORTINA G, H., LIGARRETO M, G., GALEANO, N. F., & HERRERA P, J. C. (2010). Partial resistance to leaf rust (*Hemileia vastatrix*) in coffee (*Coffea arabica* L.): genetic analysis and molecular characterization of putative candidate genes. *Molecular breeding*, 25, 685-697.

ROMERO, G., VÁSQUEZ, L. M., LASHERMES, P., & HERRERA, J. C. (2014). Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). *Plant Breeding*, 133(1), 121-129.

SAAVEDRA, L. M., CAIXETA, E. T., BARKA, G. D., BORÉM, A., ZAMBOLIM, L., NASCIMENTO, M., ... & PEREIRA, A. A. (2023). Marker-assisted recurrent selection for pyramiding leaf rust and coffee berry disease resistance alleles in *Coffea Arabica* L. *Genes*, 14(1), 189.

SAKIYAMA, N. S. (2005). Melhoramento do Café Arábica. In: BORÉM, A. (org.). Melhoramento de espécies cultivadas. 2a ed.ed. Viçosa: Editora UFV – Universidade Federal de Viçosa, p. p.605-658.

SAKIYAMA, N. S. (2015). O Café Arábica. In: SAKIYAMA, N. S. et al. (org.). Café arábica: do plantio à colheita. Viçosa: Editora UFV - Universidade Federal de Viçosa, p. 316p.

SALAJÄRVI, J., RAMBANI, A., YU, Z., GUYOT, R., STRICKLER, S., LEPELLEY, M., ... & DESCOMBES, P. (2024). The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. *Nature genetics*, 56(4), 721-731.

SCALABRIN, S., TONIUTTI, L., DI GASPERO, G., SCAGLIONE, D., MAGRIS, G., VIDOTTO, M., ... & BERTRAND, B. (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Scientific reports*, 10(1), 1-13.

SCALABRIN, S., MAGRIS, G., LIVA, M., VITULO, N., VIDOTTO, M., SCAGLIONE, D., ... & MORGANTE, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, 15(1), 463.

SERA, G. H., DE CARVALHO, C. H. S., DE REZENDE ABRAHÃO, J. C., POZZA, E. A., MATIELLO, J. B., DE ALMEIDA, S. R., ... & DOS SANTOS BOTELHO, D. M. (2022). Coffee leaf rust in Brazil: Historical events, current situation, and control measures. *Agronomy*, 12(2), 496.

SETOTAW, T. A., CAIXETA, E. T., PEREIRA, A. A., DE OLIVEIRA, A. C. B., CRUZ, C. D., ZAMBOLIM, E. M., ... & SAKIYAMA, N. S. (2013). Coefficient of parentage in *Coffea arabica* L. cultivars grown in Brazil. *Crop Science*, 53(4), 1237-1247.

SETOTAW, T. A., CAIXETA, E. T., ZAMBOLIM, E. M., SOUSA, T. V., PEREIRA, A. A., DE OLIVEIRA, A. C. B., ... & SAKIYAMA, N. S. (2020). Genome introgression of Híbrido de Timor and its potential to develop high cup quality *C. arabica* cultivars.

SILVA, M. D. C., VÁRZEA, V., GUERRA-GUIMARÃES, L., AZINHEIRA, H. G., FERNANDEZ, D., PETITOT, A. S., ... & NICOLE, M. (2006). Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian journal of plant physiology*, 18, 119-147.

SILVA, M. D. C., GUERRA-GUIMARÃES, L., DINIZ, I., LOUREIRO, A., AZINHEIRA, H., PEREIRA, A. P., ... & VÁRZEA, V. (2022a). An overview of the mechanisms involved in *Coffee-Hemileia vastatrix* interactions: Plant and pathogen perspectives. *Agronomy*, 12(2), 326.

SYLVAIN, P. G. (1955). Some observations on *Coffea arabica* L. in Ethiopia.

TEIXEIRA-CABRAL, T. A., SAKIYAMA, N. S., ZAMBOLIM, L., PEREIRA, A. A., & SCHUSTER, I. (2004). Single-locus inheritance and partial linkage map of *Coffea arabica* L. *Crop Breeding and Applied Biotechnology*, 4(4).

TSOUROS, D. C., BIBI, S., & SARIGIANNIDIS, P. G. (2019). A review on UAV-based applications for precision agriculture. *Information*, 10(11), 349.

USDA, U. S. D. of A. A. R. S. (2022). Coffee: World Markets and Trade Coffee.

USDA, U. S. D. of A. A. R. S. (2024). Coffee: World Markets and Trade Coffee.

VAN DER VOSSSEN, H., BERTRAND, B., & CHARRIER, A. (2015). Next generation variety development for sustainable production of arabica coffee (*Coffea arabica* L.): a review. *Euphytica*, 204(2), 243-256.

VAN GEEST, G., BOURKE, P. M., VOORRIPS, R. E., MARASEK-CIOLAKOWSKA, A., LIAO, Y., POST, A., ... & ARENS, P. (2017). An ultra-dense integrated linkage map for hexaploid chrysanthemum enables multi-allelic QTL analysis. *Theoretical and Applied Genetics*, 130, 2527-2541.

VÁRZEA, V. M. P.; MARQUES, D. V. (2005). Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: ZAMBOLIM, L.; ZAMBOLIM, E. M.; VÁRZEA, V. M. P. (org.). Durable resistance to coffee leaf rust. Viçosa: Editora UFV - Universidade Federal de Viçosa. p. 53–74.

VOORRIPS, R. E., GORT, G., & VOSMAN, B. (2011). Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC bioinformatics*, 12, 1-11.

VOS, P. G., UITDEWILLIGEN, J. G., VOORRIPS, R. E., VISSER, R. G., & VAN ECK, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an insight into the breeding history. *Theoretical and Applied Genetics*, 128, 2387-2401.

XIONG, H., LI, Y., GUO, H., XIE, Y., ZHAO, L., GU, J., ... & LIU, L. (2021). Genetic mapping by integration of 55K SNP array and KASP markers reveals candidate genes for important agronomic traits in hexaploid wheat. *Frontiers in Plant Science*, 12, 628478.

ZAMBOLIM, L., ZAMBOLIM, E. M., VALE, F. D., PEREIRA, A. A., SAKIYAMA, N. S., & CAIXETA, E. T. (2005). Physiological races of *Hemileia vastatrix* Berk. et Br. *Brazil-Physiological variability, current situation and future prospect*. In Zambolim L, Zambolim EM and Várzea VMP (eds) *Durable resistance to coffee leaf rust*. UFV, Viçosa, 75-98.

ZAMBOLIM, L. (2016). Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology*, 41, 1-8.

ZAMBOLIM, L., & CAIXETA, E. T. (2021). An overview of physiological specialization of coffee leaf rust-new designation of pathotypes. *Int. J. Curr. Res*, 13, 15564-15575.

ZAMBOLIM, L., CAIXETA, E. T., GUERREIRO FILHO, O., SERA, G. H., SERA, T., PEREIRA, A. A., & DE CARVALHO, C. H. (2024). Breeding Programs Against Coffee Leaf Rust in Brazil: A Review. *Journal of Agricultural Science*, 16(5), 61-74.

ZHAO, N., XUE, D., MIAO, Y., WANG, Y., ZHOU, E., ZHOU, Y., ... & WANG, X. (2023). Construction of a high-density genetic map for faba bean (*Vicia faba* L.) and quantitative trait loci mapping of seed-related traits. *Frontiers in Plant Science*, 14, 1201103.

YOU, Q., YANG, X., PENG, Z., ISLAM, M. S., SOOD, S., LUO, Z., ... & WANG, J. (2019). Development of an Axiom Sugarcane100K SNP array for genetic map construction and QTL identification. *Theoretical and Applied Genetics*, 132, 2829-2845.

WALLER, J. M., BIGGER, M., & HILLOCKS, R. J. (Eds.). (2007). Coffee pests, diseases and their management. CABI.

WANG, X., LIM, L. T., & FU, Y. (2020). Review of analytical methods to detect adulteration in coffee. *Journal of AOAC International*, 103(2), 295-305.

WANG, L., YANG, J., ZHANG, H., TAO, Q., ZHANG, Y., DANG, Z., ... & LUO, Z. (2022). Sequence coverage required for accurate genotyping by sequencing in polyploid species. *Molecular ecology resources*, 22(4), 1417-1426.

WANG, R., XING, S., BOURKE, P. M., QI, X., LIN, M., ESSELINK, D., ... & FANG, J. (2023). Development of a 135K SNP genotyping array for *Actinidia arguta* and its applications for genetic mapping and QTL analysis in kiwifruit. *Plant Biotechnology Journal*, 21(2), 369-380.

WINFIELD, M. O., ALLEN, A. M., BURRIDGE, A. J., BARKER, G. L., BENBOW, H. R., WILKINSON, P. A., ... & EDWARDS, K. J. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant biotechnology journal*, 14(5), 1195-1206.

## **CAPÍTULO 2**

### **DOSAGEM ALÉLICA E POTENCIAL PAREAMENTO HOMEÓLOGO EM *Coffea arabica***

## Resumo

A espécie *Coffea arabica*, uma commodity de importância mundial, é um aloploiploide que proveio da hibridação natural entre duas espécies diploides, *C. eugenioides* e *C. canephora*, seguida de uma duplicação de todo o genoma, provavelmente durante um único evento de aloploidização. Inicialmente, *C. arabica* foi classificada como um alotetraploide segmentar, um tipo de aloploiploide originado da hibridação de espécies próximas o suficiente, que permite o pareamento parcial entre cromossomos homeólogos. No entanto, evidências citogenéticas e moleculares indicaram a ausência de pareamento homeólogo em *C. arabica*, sugerindo um comportamento meiótico semelhante ao de espécies diploides com formação bivalentes, sendo considerada, portanto, como alotetraploide perfeita. A princípio, os estudos genéticos em *C. arabica*, baseados em marcadores moleculares, foram realizados usando marcadores desenvolvidos para espécies diploides. Entretanto, análises moleculares mais recentes, utilizando marcadores SNP, mostraram que rearranjos genômicos, envolvendo trocas homeólogas, ocorreram em *C. arabica*. Assim, a utilização de metodologias direcionadas para estudos genéticos de espécies diploides em *C. arabica* podem gerar resultados irrealistas, devido a uma limitação intrínseca desses métodos que está associada à negligência de possíveis trocas homeólogas, além e à desconsideração do nível de ploidia. Neste trabalho, foi investigada a dosagem alélica e possíveis pareamento homeólogos em *C. arabica*, utilizando genotipagem em larga escala com marcadores SNP em população F<sub>2</sub>. Foram observadas menores distorções de segregação dos marcadores ao utilizar a chamada de variantes tetraploides em comparação com chamadas diploides, além de evidências para um suposto pareamento não preferencial entre os dois subgenomas que constituem a espécie *C. arabica*. Assim como os estudos de dosagem alélica, o pareamento homeólogo em *C. arabica* ainda é uma área pouco explorada. Dessa forma, as descobertas desta pesquisa têm implicações significativas para programas de melhoramento de *C. arabica*, fornecendo novas informações cruciais para a compreensão do genoma dessa espécie, evidenciando a importância de considerar a dosagem alélica e o pareamento homeólogo em estudos sobre esse alotetraploide.

**Palavras-chave:** Aloploiploide; Marcador molecular; Melhoramento do cafeeiro; Poliploides; SNP.

### Abstract

The species *Coffea arabica*, a commodity of global importance, is an allopolyploid that emerged from the natural hybridization between two diploid species, *C. eugenioides* and *C. canephora*, followed by a duplication of the full genome, probably during a single allopolyploidization event. Initially, *C. arabica* was classified as a segmental allotetraploid, an allopolyploid originated from the hybridization of species close enough to allow partial pairing between homeologous chromosomes. However, cytogenetic and molecular evidence indicated the absence of homeologous pairing in *C. arabica*, suggesting a meiotic behavior similar to diploid species with bivalent formation, and therefore being considered a perfect allotetraploid. Initially, genetic studies in *C. arabica*, based on molecular markers, were performed using markers developed for diploid species. However, more recent molecular analyses using SNP markers have shown that genomic rearrangements involving homeologous exchanges have occurred in *C. arabica*. Therefore, the use of methodologies aimed at genetic studies of diploid species in *C. arabica* may generate unrealistic results, due to an intrinsic limitation of these methods that is associated with the neglect of possible homeologous exchanges, in addition to the disregard of the ploidy level. This study investigated allelic dosage and possible homeologous pairings in *C. arabica* using large-scale genotyping with SNP markers in an F<sub>2</sub> population. Smaller segregation distortions of markers were observed when using tetraploid variant calling compared to diploid calls. In addition, a supposed non-preferential pairing between the two subgenomes that constitute the species *C. arabica* was shown. Homeologous pairing and allelic dosage studies in *C. arabica* are still an underexplored area. Thus, the findings of this research have significant implications for *C. arabica* breeding programs, providing new information crucial for understanding the genome of this species, highlighting the importance of considering allelic dosage and homeologous pairing in studies on this allotetraploid.

**Keywords:** Allopolyploid; Molecular marker; Coffee breeding; Polyploids; SNP.

## 1. Introdução

A duplicação completa de um genoma, evento conhecido como poliploidia, é um mecanismo importante na especiação e evolução dos genomas das plantas. Sabe-se que a poliploidia ocorreu com frequência durante a evolução das angiospermas e que todas as espécies de plantas com flores passaram por uma ou mais duplicação do genoma em sua história evolutiva (JIAO et al., 2011; SOLTIS et al., 2014). A alopoliploidização, mecanismo que envolve a hibridação interespecífica antes da duplicação completa de um genoma, também contribuiu com a adaptação das plantas a nichos mais amplos, desempenhando um papel fundamental na história evolutiva de importantes culturas agrícolas, como o trigo (*Triticum aestivum*), algodão (*Gossypium hirsutum*) e café (*Coffea arabica*) (SOLTIS et al., 2014; LASHERMES et al., 2014).

A espécie *C. arabica*, amplamente reconhecida como uma commodity de importância mundial (CONAB, 2024; MAPA, 2024; USDA, 2024), é um alopoliploide que se reproduz principalmente por autopolinização. No gênero *Coffea*, apenas *C. arabica* é poliploide ( $2n = 4x = 44$  cromossomos). Em contraste, as outras 140 espécies/táxons são diploides ( $2n = 2x = 22$  cromossomos) (GUYOT et al., 2020). *C. arabica* proveio da hibridação natural entre duas espécies diploides, *C. eugenioides* e *C. canephora* (LASHERMES, et al., 1999), seguida de uma duplicação de todo o genoma, provavelmente durante um único evento de alopoliploidização (CLARINDO & CARVALHO, 2008; LASHERMES, et al., 2014; BAWIN et al., 2021).

As espécies alopoliploides podem seguir diferentes caminhos evolutivos, o que torna difícil fazer generalizações sobre seus processos de evolução (DOYLE et al., 2008; SALOJARVI et al., 2024). Esses organismos exibem padrões de herança genética que variam de completamente disômicos, nos quais os cromossomos homólogos (originados do mesmo subgenoma) se pareiam durante a meiose, até tetrassômicos, nos quais cromossomos homólogos e homeólogos (originados de subgenomas diferentes) podem interagir e se parear (ENDRIZZI, 1962; LEAL-BERTIOLI et al., 2018). Além disso, diferentes padrões de herança podem ocorrer dentro do mesmo genoma (SCALABRIN et al., 2024). Entretanto, a compreensão das alterações genômicas pós duplicação e suas consequências ainda é limitada, mas há evidências que pode ocorrer choque genômico, em termos de ativação de elementos transponíveis, troca homeóloga, instabilidade meiótica e até variação de cariótipo (DOYLE et al., 2008; SALOJARVI et al., 2024).

A origem poliploide de *C. arabica* resultante do cruzamento entre duas espécies diploides com genomas semelhantes, é amplamente aceita. No entanto, a sua classificação em relação ao padrão de herança genética foi por muitas vezes debatida (CLARINDO & CARVALHO 2008; 2009; CENCI et al., 2012). Inicialmente, *C. arabica* foi classificada como um alotetraploide segmentar, um tipo de aloploiploide originado da hibridação de espécies próximas o suficiente, que permite o pareamento parcial entre cromossomos homeólogos (OROZCO-CASTILLO et al., 1996; PINTO-MAGLIO & CRUZ, 1998). No entanto, evidências moleculares sugeriram a ausência de pareamento homeólogo em *C. arabica*, não devido a diferenciação estrutural entre os genomas parentais, mas possivelmente pelo efeito de um ou mais genes reguladores, semelhantes ao *Ph* (supressor de pareamento homeólogo) (LASHERMES et al. 2000). A natureza alotetraploide de *C. arabica* também foi reforçada por análises citogenética, que mostraram que, apesar da estreita relação entre os dois subgenomas, *C. arabica* exibe comportamento meiótico semelhante ao de espécies diploides com formação bivalentes, podendo ser, portanto, considerada alotetraploide perfeita (CLARINDO & CARVALHO 2008; 2009).

Inicialmente, os estudos genéticos em *Coffea*, baseados em marcadores moleculares, foram realizados usando marcadores desenvolvidos para espécies diploides, como RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) e SSR (*Simple Sequence Repeat*) (LASHERMES et al., 2000, 2001; PRAKASH et al., 2004; PEARL et al., 2004; TEIXEIRA-CABRAL et al., 2004; OLVEIRA et al., 2007; PRIOLLI et al., 2008; ROMERO et al., 2014; FERRÃO et al., 2015; PESTANA et al., 2015), considerando o comportamento meiótico das espécies do gênero, incluindo *C. arabica*, semelhante ao observado em espécies diploides, caracterizado pela formação de bivalentes. Além disso, esses marcadores ganharam ampla popularidade em diversas espécies vegetais, incluindo poliploides, devido à sua capacidade de serem aplicados sem a necessidade de informações prévias de sequência, e se destacavam-se por sua alta reprodutibilidade em diferentes laboratórios, tornando-se ferramentas essenciais para análises genéticas em plantas (CORDEIRO et al., 2000; MCGREGOR et al., 2000; MCNEIL et al., 2011).

Mapas genéticos de *Coffea* foram desenvolvidos usando marcadores moleculares RAPD, AFLP e SSR (PAILLARD et al., 1996; KY et al., 2000; PEARL et al., 2004; TEIXEIRA-CABRAL et al., 2004; OLVEIRA et al., 2007; PESTANA et al., 2015). Nesses estudos, apesar dos mapas serem pouco saturados, foram identificados QTL associados a características de importância para os programas de melhoramento do cafeeiro. Entretanto,

foram relatados um alto número de marcadores com distorção de segregação (DS), o que dificulta a construção de mapas saturados.

Posteriormente, novas tecnologias de sequenciamento possibilitam a geração de dados de genotipagem de alto rendimento usando SNP (*Single Nucleotide Polymorphism*), que são mais informativos do que os marcadores usados inicialmente. Essa abordagem permite uma melhor cobertura do genoma e pode auxiliar no desenvolvimento de modelos mais realistas para estudos genéticos em poliploides (SERANG et al., 2012; LARA et al., 2019; GERARD et al., 2024). Os marcadores SNP possibilitaram a observação direta de todos os genótipos, aumentando significativamente a informação incorporada em cada locus (dosagem alélica), e oferecendo uma visão mais completa das variações presentes no DNA das espécies poliploides (VOORRIPS et al., 2011; HACKETT et al., 2013; GARCIA et al., 2013; GERARD et al., 2018).

Análises moleculares mais recentes, utilizando SNP, mostraram que rearranjos genômicos, envolvendo trocas homeólogas, ocorreram em *C. arabica* (LASHERMES et al., 2014, 2016; ORTIZ & SHARBROUGH, 2024; SALOJARVI et al 2024). Esses estudos indicaram que variações de sequência entre os subgenomas coexistem com variações alélicas dentro dos subgenomas. Assim, a utilização de metodologias direcionadas para estudos genéticos de espécies diploides em *C. arabica* podem gerar resultados irrealistas, incluindo distorção de segregação de marcadores moleculares, devido a uma limitação intrínseca desses métodos que está associada à negligência do nível de ploidia e possíveis trocas homeólogas nessa espécie alotetraploide (GARCIA et al., 2013; GERARD et al., 2024). Destacando a importância de entender a complexidade do genoma dessa espécie.

Desse modo, o objetivo desse trabalho foi investigar a dosagem alélica e possíveis pareamento homeólogos em *C. arabica*, por meio de genotipagem em larga escala com marcadores SNP, visando uma compreensão mais aprofundada da estrutura genética da espécie. Posteriormente, as informações obtidas poderão ser aplicadas nos estudos genéticos, como nos desenvolvimentos de marcadores moleculares mais eficientes, para esse aloploide, considerando suas peculiaridades.

## **2. Material e métodos**

### **2.1. População experimental**

A população utilizada nesse estudo, é uma população F<sub>2</sub> de *C. arabica*, derivada da autofecundação de uma planta F<sub>1</sub>. Esse genótipo foi originado do cruzamento entre Híbrido de Timor UFV 443-03 (genitor masculino) e Catuaí Amarelo IAC 64 (UFV 2148-57). Os cruzamentos para a obtenção da geração F<sub>1</sub> foram realizados por Pereira (1995) no Departamento de Fitopatologia, na Universidade Federal de Viçosa (Viçosa/MG) (20°45'14"S, 42°52'53"O) – UFV e nas Fazendas Experimentais da EPAMIG. A planta da geração F<sub>1</sub> (H 511-1), foi submetida à autofecundação controlada, para obtenção da geração segregante F<sub>2</sub>, com 246 plantas que se encontram em campo experimental da UFV (CAPUCHO et al., 2009).

## **2.2. Genotipagem**

### **2.2.1. Extração de DNA**

Foram coletados um par de folhas jovens de cada uma das 232 plantas da população F<sub>2</sub> segregante, da F<sub>1</sub>, H 511-1, e dos genitores da F<sub>1</sub>, UFV 443-03 e UFV 2148-57. Esse material biológico foi enviado ao Laboratório de Biotecnologia do Cafeeiro (BioCafé/UFV), onde foi realizada a extração do DNA genômico conforme a metodologia descrita por (DINIZ et al., 2005). Posteriormente, a qualidade do DNA foi avaliada em gel de agarose (1%) e quantificada com espectrofotômetro Qubit DNA BR (*Life Technologies*). A concentração final de DNA foi ajustada para 20 ng/μL.

### **2.2.2. Genotipagem com marcadores SNP**

Após a extração e análise de qualidade, as amostras de DNA foram enviadas para a empresa *Rapid Genomics*, Flórida (EUA), para a genotipagem por sequenciamento de regiões genômicas. Essas regiões sequenciadas foram inicialmente identificadas por Resende et al. (2016). Estes autores desenharam sondas com base em sequências disponíveis em bancos de dados de *Expressed Sequence Tags* (ESTs) desenvolvidos pelo Projeto Genoma Brasileiro de Café (VIEIRA et al., 2006) e também do genoma de referência de *C. canephora* (DENOEUDE et al., 2014), com o objetivo de cobrir todo o genoma de *C. arabica*. Para o presente trabalho, foram selecionadas 400 sondas polimórficas mais informativas, contendo maior número de SNP, com base em trabalhos anteriores (SOUSA et al., 2017; SILVA et al., 2018; SILVA et al., 2022b). Foram desenhados pares de primers para a amplificação destas regiões usando a tecnologia *Flex-Seq EX-L* (*Rapid Genomics*, 2021). A clonagem e o sequenciamento das sondas amplificadas foram realizados utilizando a plataforma *Illumina HiSeq*.

### 2.3. Chamadas de variantes e dosagem alélica

Com o uso do *software* bowtie2, as leituras obtidas com o sequenciamento foram mapeadas utilizando o genoma de *C. arabica* como referência, disponível no NCBI (GCF\_003713225.1), os códigos que representam os cromossomos do genoma de *C. arabica* estão disponíveis na tabela 1, e com base nas informações de alinhamento, foram feitas as identificações das variantes. Foram realizados dois mapeamentos diferentes. Inicialmente, procedeu-se com o mapeamento contra o genoma *C. arabica* com chamada de variantes considerando ploidia igual a 2 (diploides, gerando até três classes genotípicas: 0, 1, 2) e posteriormente foi feito um mapeamento com chamadas de variantes considerando ploidia igual a 4 (tetraploides, gerando até cinco classes genotípicas: 0, 1, 2, 3, 4). A ferramenta FreeBayes (GARRISON & MARTH, 2012) foi utilizada para realizar a chamada de variantes a partir dos arquivos gerados pelo alinhamento com o bowtie2. As análises posteriores dos VCF foram feitas utilizando *scripts* personalizados no *software* R (4.2.3).

Tabela 1: Relação entre os códigos que representam os cromossomos dos subgenomas (C<sup>c</sup> e C<sup>e</sup>) do genoma de *C. arabica*, disponível no NCBI.

<b>Cromossomo</b>	<b>Código</b>	<b>Comprimento (bp)</b>
1c	NC_039898.1	50.636.588
1e	NC_039899.1	48.756.970
2c	NC_039900.1	66.155.350
2e	NC_039901.1	71.633.312
3c	NC_039902.1	41.566.753
3e	NC_039903.1	37.271.464
4c	NC_039904.1	41.786.336
4e	NC_039905.1	42.507.429
5e	NC_039906.1	39.439.615
5c	NC_039907.1	45.899.693
6c	NC_039908.1	55.181.588
6e	NC_039909.1	52.240.725
7c	NC_039910.1	38.854.053
7e	NC_039911.1	35.392.230
8e	NC_039912.1	45.117.557
8c	NC_039913.1	39.008.463
9c	NC_039914.1	38.064.651
9e	NC_039915.1	35.839.895
10e	NC_039916.1	40.458.934
10c	NC_039917.1	45.429.025
11c	NC_039918.1	36.215.491
11e	NC_039919.1	42.465.768

A fim realizar a chamada de dose e estudar de forma mais precisa as distribuições dos genótipos da população em estudo, foi utilizada a ferramenta updog, para gerar probabilidades de genótipo e distribuições posteriores de genótipo para cada indivíduo em cada SNP (GERARD et al., 2018). Com base nas contagens de leitura contidas nos VCF, foram inferidas as dosagens de alelos de referência e alternativos usando modelo desenvolvido para o estudo de populações poliploides. O modelo utilizado foi o “s1pp”, o qual considera uma população tetraploide resultantes de uma única geração de autofecundação (geração S1), estimando com precisão a frequência de alelos e evidenciando classes genotípicas que antes eram confundidas (GERARD et al., 2018).

Foram obtidos gráficos de dispersão com o objetivo de visualizar a distribuição de indivíduos nas classes genotípicas (dosagens), a partir do FreeBayes e do updog resultantes de segregações esperadas em uma população  $F_2$ . Para cada marcador específico na população em estudo, o gráfico de dispersão contém as contagens de leituras do alelo de referência no eixo y e as contagens de leituras do alelo alternativo no eixo x. Nos gráficos de dispersão com os dados obtidos a partir do mapeamento com chamadas de variantes considerando ploidia igual a 4, indivíduos com quatro contagem de alelos de referência foram alocados perto da linha da vertical ( $x = 0$ ), enquanto indivíduos com uma contagem de zero alelos de referência estariam perto da linha horizontal  $y = 0$ .

### **3. Resultados**

Após realizar o mapeamento contra o genoma de *C. arabica*, com chamada de variantes diploide, o arquivo VCF resultante apresentou 2660 marcadores SNP para os 235 indivíduos, distribuídos em todos os 22 cromossomos do genoma (Figura 1).

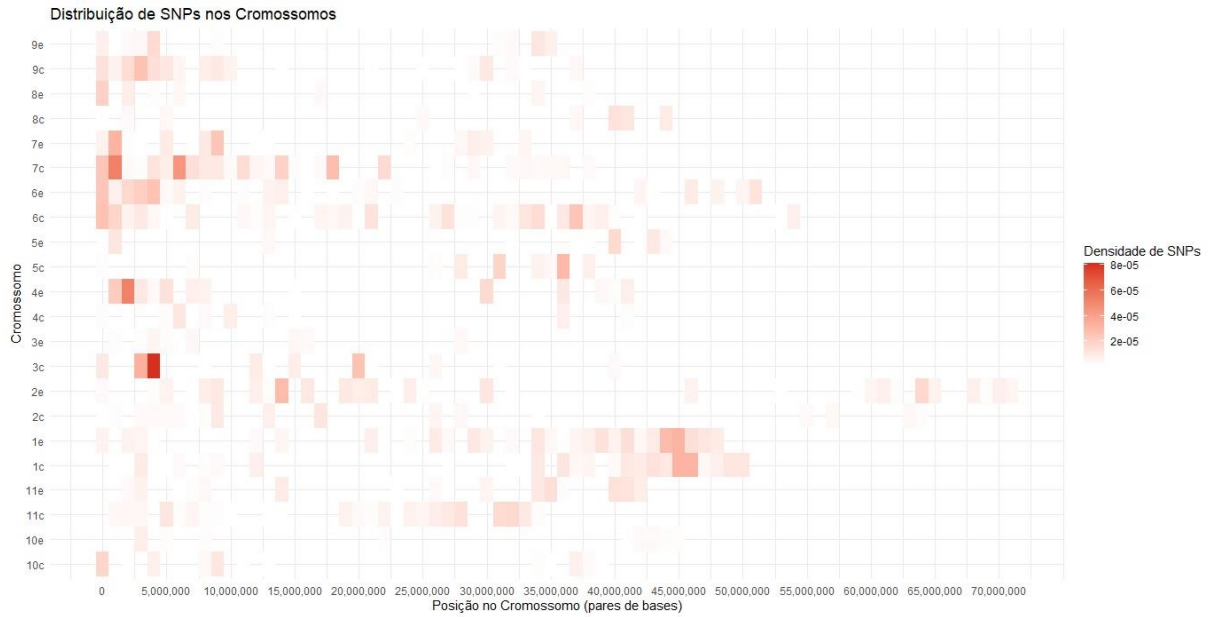


Figura 1. Distribuição dos marcadores SNP com chamada diploide nos cromossomos do genoma de *C. arabica*.

A maior parte dos gráficos de dispersão obtidos a partir desse mapeamento com o objetivo de visualizar a distribuição de indivíduos nas classes genóticas (dosagens), a partir do FreeBayes, resultantes de segregações esperadas em uma população  $F_2$  apresentaram desajuste para a maioria dos marcadores. Foi observado que a classe de genótipos heterozigotos não se comportou de acordo com a segregação esperada de 1:2:1 (1AA: 2Aa: 1aa) para uma população  $F_2$  segregante, diploide, derivada de uma autofecundação de um genótipo heterozigoto, para todos os loci (Figura 2). O teste de qui-quadrado resultou em 2510 marcadores distorcidos, restando apenas 150 marcadores segregando na proporção esperada.

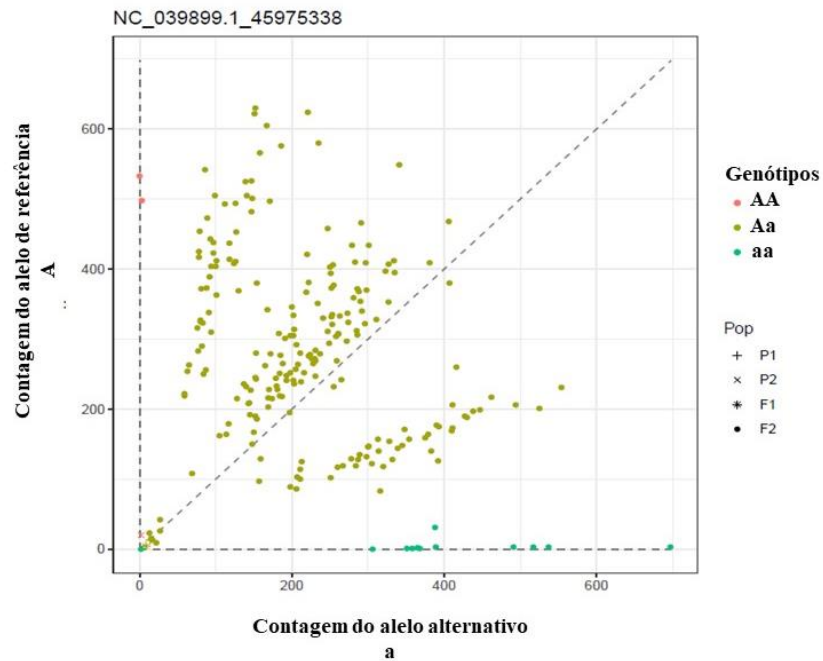


Figura 2. Distribuição dos genótipos para o marcador NC\_039899.1\_4575338 nas classes esperadas 1AA: 2Aa: 1aa.

Já o arquivo VCF resultante do mapeamento contra o genoma de *C. arabica*, com chamada de variantes tetraploide apresentou 2766 marcadores SNP para os 235 indivíduos, distribuídos por todos os 22 cromossomos do genoma de *C. arabica* (Figura 3). Os gráficos de dispersão obtidos a partir desses dados apresentaram menor desajuste para a maioria dos marcadores, em comparação com os dados de chamada de variantes diploides (Figura 4).

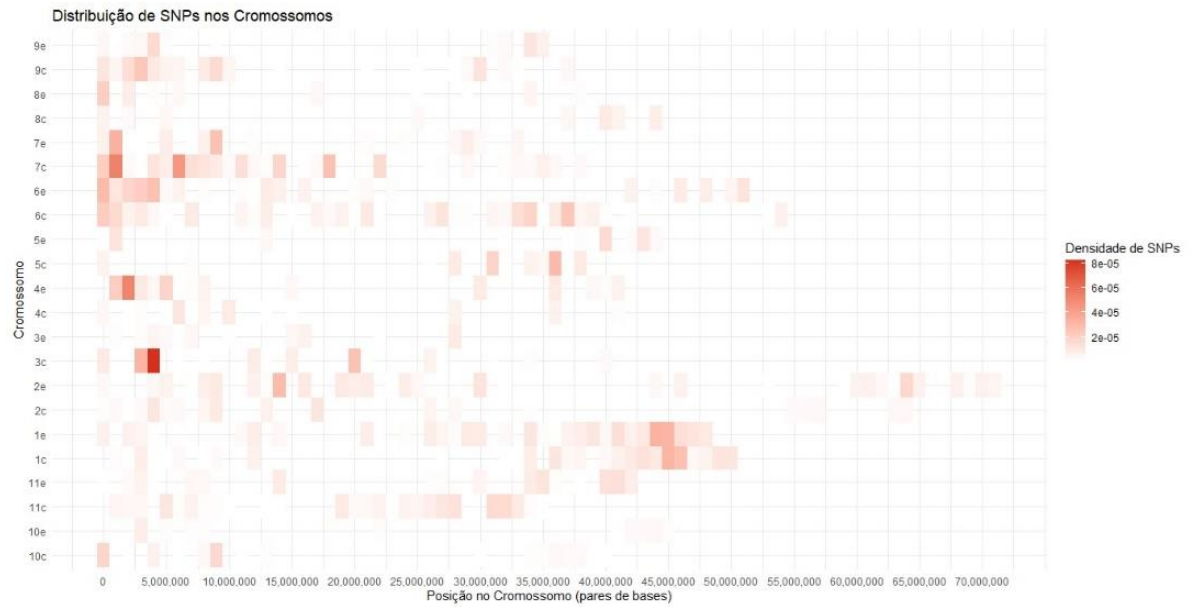


Figura 3. Distribuição dos marcadores SNP com chamada tetraploide nos cromossomos do genoma de *C. arabica*.

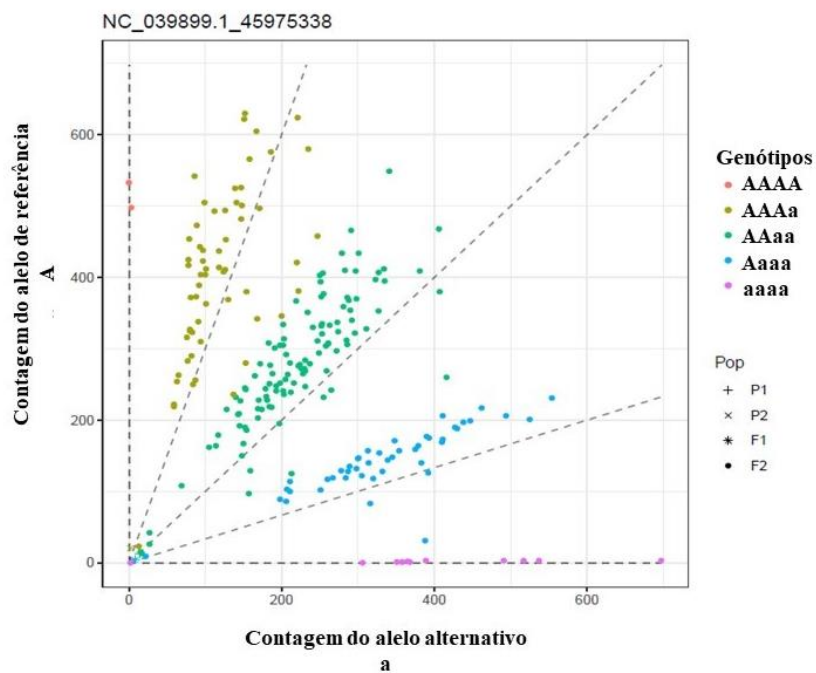


Figura 4. Distribuição dos genótipos para o marcador NC\_039899.1\_4575338 nas classes esperadas 1AAAA: 4AAAa: 6AAaa: 4Aaaa: 1aaaa.

Utilizando a ferramenta updog, modelo “s1pp” para a análise dos 2766 marcadores SNP com chamada de variantes tetraploides, foram obtidas melhores distribuição de indivíduos nas classes genóticas (dosagens) (Figura 5).

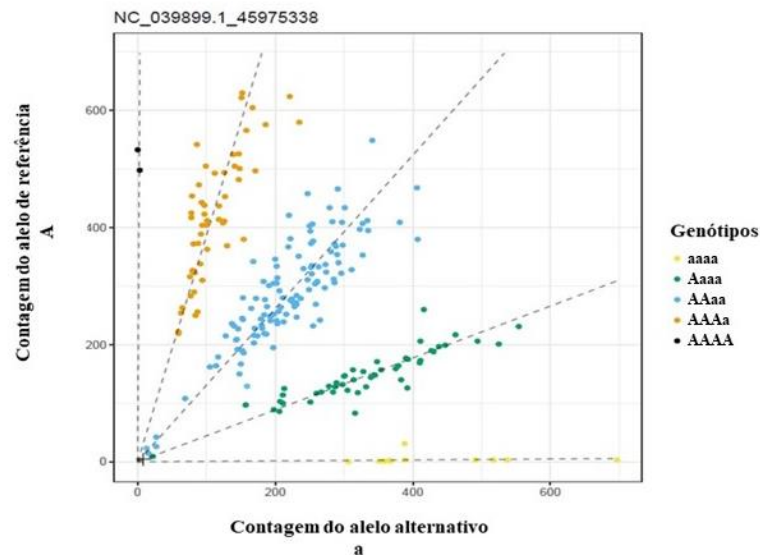


Figura 5. Distribuição dos genótipos para o marcador NC\_039899.1\_4575338, utilizando o modelo “s1pp” do updog, nas classes esperadas 1AAAA: 4AAAAa: 6AAaa: 4Aaaa: 1aaaa, ajustado aos vieses de segregação esperados.

Entretanto, foi observado que alguns marcadores não se ajustaram nas dosagens esperadas para uma população  $F_2$  segregante alotetraploide (1AAAA: 4AAAAa: 6AAaa: 4Aaaa: 1aaaa). Os dados evidenciaram que determinados marcadores apresentam segregação de 1:2:1 (1AAAA: 2AAaa: 1aaaa), distribuição esperada para uma população  $F_2$  segregante, tetraploide, derivada de uma autofecundação de um genótipo que possui loci heterozigotos para a porção de um dos subgenomas e homozigoto para outro (Figura 6).

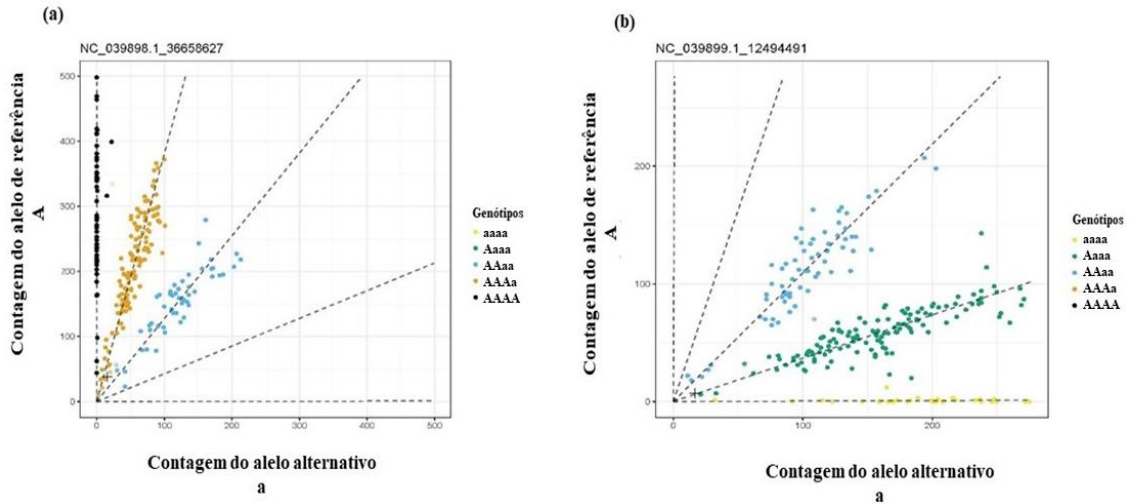


Figura 6. Distribuição dos genótipos para os marcadores NC\_039898.1\_36658627 (a) e NC\_039899.1\_12494491 (b).

Além disso, foi identificado que determinados marcadores apresentam a distribuição de doses de 1:2:1 (1AAAA: 2AAaa: 1aaa) quase perfeita. No entanto, existem alguns poucos indivíduos presentes nas classes que representam as dosagens Aaaa e AAAa. Essa observação evidencia classes resultantes de um suposto pareamento não preferencial, que pode ser definido como pareamento entre cromossomos pertencentes aos diferentes subgenomas de *C. arabica* (pareamento entre homeólogos) que não é esperado na segregação de espécie exclusivamente alotetraploide (Figura 7).

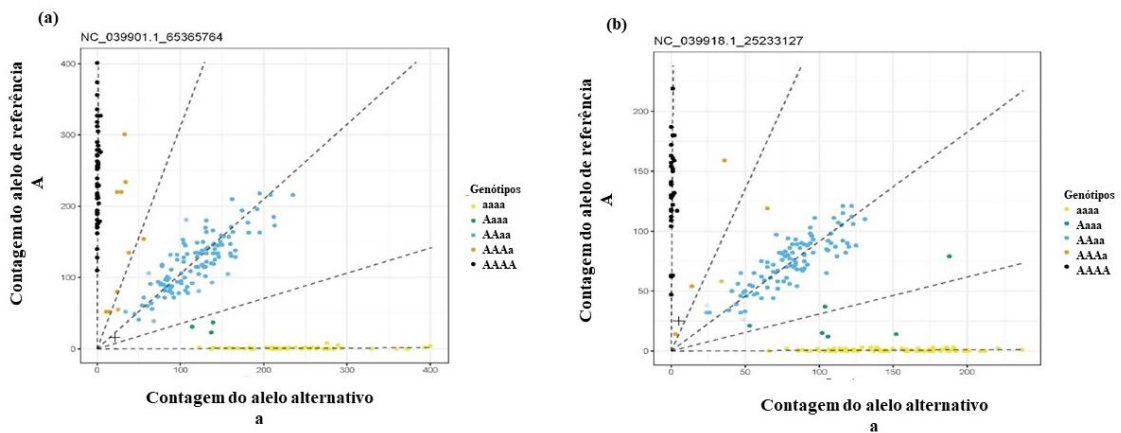


Figura 7. Gráficos apresentando evidências para um suposto pareamento entre os cromossomos homeólogos de *C. arabica* (cromossomos pertencentes aos subgenomas C<sup>c</sup> e C<sup>e</sup>) na distribuição dos marcadores NC\_039901.1\_65365764 (a) e NC\_039918.1\_25233127 (b).

#### 4. Discussão

Tradicionalmente, os estudos genéticos de poliploides focavam principalmente em marcadores moleculares desenvolvidos para espécies diploides, como AFLP e SSR. Esses marcadores se tornaram populares em muitas espécies de plantas e tinham como maior vantagem não necessitar de nenhuma informação prévia de sequência, além de serem técnicas altamente reprodutíveis em laboratórios (CORDEIRO et al., 2000; MCGREGOR et al., 2000; MCNEIL et al., 2011). Entretanto, as espécies poliploides não se beneficiavam totalmente das informações desses marcadores. Pois essas técnicas não permitem determinar o número de cópias de cada alelo (dosagem alélica) em um locus polimórfico, apresentando suposições mais simplificadas do genoma em estudo (MCGREGOR et al., 2000; MCNEIL et al., 2011; GARCIA et al., 2013).

Com o surgimento das tecnologias de sequenciamento de nova geração e o progresso em métodos genéticos e estatísticos, tornaram-se possíveis estudos mais detalhados de genomas poliploides em diferentes culturas (HACKETT et al., 2013; GARCIA et al., 2013; SCALABRIN et al., 2024). As tecnologias de genotipagem baseadas em marcadores SNP proporcionaram uma ampla cobertura do genoma e possibilitaram a identificação da dosagem alélica. Essa abordagem permite a observação direta de todos os genótipos, aumentando significativamente a informação incorporada em cada locus e oferecendo uma visão mais completa das variações presentes no DNA das espécies poliploides (VOORRIPS et al., 2011; HACKETT et al., 2013; GARCIA et al., 2013; GERARD et al., 2018).

A dosagem de alelo se refere ao número de cópias dos alelos em um locus gênico em relação ao alelo de referência. Em tetraploide, por exemplo, onde A é o alelo de referência, o genótipo varia de aaaa para nulliplex, Aaaa para simplex (dose única) e assim por diante até AAAA (quadruplex). O grande desafio em trabalhar com espécies poliploides está na distinção correta entre diferentes tipos de heterozigotos, no exemplo anterior: Aaaa, AAaa e AAAa. Nesse caso, se os genótipos fossem avaliados com um marcador AFLP, por exemplo, eles seriam pontuados como 0 (banda de gel ausente) para aaaa ou 1 (banda de gel presente) para todos os outros indivíduos (Aaaa, AAaa, AAAa e AAAA). Outro exemplo, utilizado um marcador codominante

SSR seriam diferenciados os genótipos aaaa e AAAA, mas os genótipos Aaaa, AAaa, AAAa não seriam distinguidos. Esses resultados irrealistas ocorrem devido a uma limitação intrínseca dos métodos que está associada à negligência do nível de ploidia (GARCIA et al., 2013; ANNICCHIARICO et al., 2015; BIAZZI et al., 2017).

Neste estudo, utilizando dados de genotipagem de marcadores SNP para uma população de *C. arabica*, foram observados desajustes das dosagens dos genótipos em algumas classes de marcadores. Essas distorções foram identificadas na maioria dos marcadores a partir do alinhamento contra o genoma de *C. arabica*, empregando chamadas de variantes diploides. A segregação esperada utilizando esses marcadores codominantes para uma população F<sub>2</sub> segregante, diploide, derivada de uma autofecundação de um genótipo heterozigoto, para todos os loci é 1:2:1. Entretanto, os resultados mostraram que a classe heterozigota apresentou alta variação de segregação (94% de marcadores distorcidos).

Utilizando o alinhamento contra o genoma de *C. arabica* com chamadas de variantes tetraploides, foram observadas menores distorções de segregação para a maioria dos marcadores, em comparação com as chamadas de variantes diploides. A fim de realizar a chamada de dose e estudar de forma mais precisa as distribuições dos genótipos da população em estudo, foi utilizada a ferramenta updog, que se baseia em uma modelagem probabilística que melhora a distribuição de classes confundidas. Esses resultados demonstraram uma melhor distinção dos genótipos heterozigotos, separando-os em diferentes classes, o que reforça a importância de considerar a dosagem alélica na análise do genótipo dessa espécie.

Em diferentes estudos para a construção de um mapa genético para *C. arabica*, com marcadores moleculares desenvolvidos para espécies diploides, também foram relatados marcadores com distorção de segregação. Utilizando marcadores RAPD, Oliveira et al. (2007) e Teixeira-Cabral et al. (2004) verificaram 24,85% e 5,50% de segregação com distorção, respectivamente. Para marcadores AFLP, Ky et al. (2000) e Pearl et al. (2004) verificaram distorção na ordem de 30%, em diferentes espécies de *Coffea* sp. e 25% em população de *C. arabica*. No trabalho desenvolvido por Pestana et al. (2015), utilizando populações derivadas do Híbrido de Timor (HdT), como neste trabalho, e marcadores RAPD, AFLP e SSR, os autores verificaram distorção em 21 marcadores do total de 137 (15,33% de distorção). Esses resultados corroboram a ideia de que metodologias desenvolvidas para estudos genéticos de espécies diploides podem gerar resultados imprecisos em *C. arabica*, como a identificação de distorções de segregação em marcadores moleculares que, na realidade, não apresentam essa característica. Essa limitação decorre da inadequação do uso desses métodos, que

desconsideram a dosagem alélica de espécies poliploides (GARCIA et al., 2013; GERARD et al., 2024), evidenciando a necessidade de abordagens específicas para a interpretação precisa dos dados desse aloploiploide.

Sabe-se que *C. arabica* é resultante da hibridação entre duas espécies relacionadas, *C. eugenioides* e *C. canephora*. Os dois subgenomas, das espécies parentais apresentam alta colinearidade (SCALABRIN et al., 2020; SALOJARVI et al., 2024). Assim, deve-se levar em consideração mais cópias de alelos para determinado locus, ou seja, loci que são compartilhados pelos dois subgenomas. De tal modo, os genótipos heterozigotos irão se distribuir em diferentes classes, e não apenas uma (exemplo: Aaaa, AAaa, AAAa), ao utilizar marcadores que identificam esses loci compartilhados. Isso explica a distorção de segregação de alguns marcadores desenvolvidos para espécies diploides, que consideram que os genótipos heterozigotos estão presentes apenas na classe Aa.

Entretanto, apesar do baixo nível de divergência entre os dois genomas de *C. arabica* sabe-se que variações de sequência entre subgenomas coexistem com variações alélicas dentro de subgenomas. A introgressão de características de interesse em espécies comerciais, prática realizada nos programas de melhoramento do cafeeiro, aumentou a diversidade de algumas regiões genômicas de determinados genótipos (SCALABRIN et al., 2020; SALOJARVI et al., 2024). Um exemplo é o uso de variedades originadas do HdT em cruzamentos com cultivares de *C. arabica*, visando a concentração de características agrônomicas relevantes, como a introgressão de genes associados à resistência contra uma das principais doenças dessa cultura, a ferrugem do cafeeiro (CAPUCHO et al., 2009; SETOTAW et al., 2013; ALMEIDA et al., 2021a, 2021b; SERA et al., 2022; SILVA et al., 2006, 2022a; ALVES et al., 2024).

No presente estudo, foi observado que alguns marcadores segregaram na proporção 1:2:1, segregação semelhante a que seria observada em genótipos com comportamento de tetraploides perfeitos, evidenciando alguns loci heterozigotos para a porção de um dos subgenomas e ausentes ou em homozigose para o outro genitor parental. Entretanto, as classes de genótipos heterozigóticos, quando os alelos estavam presentes nos dois subgenomas (segregando em um e em homozigose no outro) foram distribuídos em duas classes distintas, de acordo com a dosagem alélica. A proporção 1:2:1 nesse caso corresponde 1(AAAA ou aaaa):2(AAAa ou Aaaa):1(AAaa) e não 1(homozigoto): 2(heterozigoto):1(homozigoto). Nesse contexto, se fossem utilizados marcadores dominantes, o resultado seria uma proporção 3:1, juntando as classes heterozigotas, gerando conclusões equivocadas. O mesmo aconteceria se fossem

utilizados marcadores codominantes em gel. As classes de heterozigotos apresentariam pouca diferença e poderia resultar em interpretações errôneas.

Além disso, alguns marcadores analisados no presente estudo que apresentaram segregação 1:2:1 (1AAAA:2AAaa:1aaaa) exibiram alguns poucos genótipos presentes nas outras classes (AAAA e Aaaa), não esperadas na segregação dessa espécie alotetraploide, nesse contexto. No caso, para ocorrer essa segregação, 1AAAA:2AAaa:1aaaa, esperava-se que os alelos estivessem presentes apenas em um dos subgenomas e ausente no outro. A presença de alguns genótipos nas classes AAAA e Aaaa fornece evidência para um suposto pareamento não preferencial entre os dois subgenomas que constituem a espécie *C. arabica*, pareamento entre homeólogos. Devido ao baixo número de genótipos que apresentaram esse comportamento, a espécie parece apresentar um pareamento preferencial parcial, havendo maior preferência para o pareamento entre cromossomos do mesmo subgenoma.

Diferentes estudos demonstraram que a espécie *C. arabica* exibe comportamento meiótico semelhante ao diploide (herança dissômica), com formação apenas de bivalentes (CLARINDO & CARVALHO 2008; 2009), possivelmente pelo efeito de um ou mais genes supressores de pareamento homeólogo (LASHERMES et al. 2000). Entretanto, análises moleculares mais recentes mostraram que rearranjos genômicos envolvendo trocas homeólogas ocorreram em *C. arabica* e podem ter tido um papel importante na evolução desse genoma aloploiploide (LASHERMES et al., 2016; SCALABRIN et al., 2024; SALOJARVI et al., 2024). Uma possibilidade é que a maioria dessas trocas homeólogas tenham ocorrido precocemente e poderia ser o resultado de um possível pareamento multivalente (ou bivalente menos estrito) na meiose no início da evolução da espécie (LASHERMES et al., 2016). Entretanto, de acordo com Scalabrin et al. (2024) e Salojarvi et al. (2024), essas trocas não aconteceram apenas após o evento de poliploidização, mas, ao contrário do que se supõe, pareamento e troca homeólogos ocasionais ainda podem ocorrer. O que corrobora com o baixo número de genótipos que apresentaram esse comportamento no estudo, sugerindo que a espécie apresenta pareamento entre homeólogos, mas demonstra uma maior preferência pelo pareamento entre cromossomos do mesmo subgenoma.

Assim como os estudos de dosagem alélica, o pareamento homeólogo em *C. arabica* ainda é uma área pouco explorada. Dessa forma, as descobertas desta pesquisa fornecem novas informações cruciais para a compreensão do genoma de *C. arabica*, evidenciando a importância de considerar a dosagem alélica em estudos sobre esse aloploiploide. Além disso, enquanto a espécie foi classificada como alotetraploide perfeita, dados desse trabalho revelam evidências

de pareamento entre cromossomos homeólogos. Essa perspectiva destaca a importância para os programas de melhoramento do cafeeiro em entender a complexidade do genoma de *C. arabica* e utilizar esse conhecimento para desenvolver modelos mais realistas para estudos genéticos dessa espécie alotetraploide.

## 5. Referências Bibliográficas

- ALMEIDA, D. P., CAIXETA, E. T., MOREIRA, K. F., DE OLIVEIRA, A. C. B., DE FREITAS, K. N. P., PEREIRA, A. A., ... & CRUZ, C. D. (2021a). Marker-assisted pyramiding of multiple disease resistance genes in coffee genotypes (*Coffea arabica*). *Agronomy*, 11(9), 1763.
- ALMEIDA, D. P., CASTRO, I. S. L., MENDES, T. A. D. O., ALVES, D. R., BARKA, G. D., BARREIROS, P. R. R. M., ... & CAIXETA, E. T. (2021b). Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. *Scientia Agricola*, 78.
- ALVES, D. R., ALMEIDA, D. P., DE ANDRADE SILVA, E. M., CASTRO, I. S. L., BARREIROS, P. R. R. M., DE OLIVEIRA MENDES, T. A., ... & CAIXETA, E. T. (2024). Unravelling the role of key genes involved coffee leaf rust resistance. *Current Plant Biology*, 38, 100347.
- ANNICCHIARICO, P., BARRETT, B., BRUMMER, E. C., JULIER, B., & MARSHALL, A. H. (2015). Achievements and challenges in improving temperate perennial forage legumes. *Critical Reviews in Plant Sciences*, 34(1-3), 327-380.
- BAWIN, Y., RUTTINK, T., STAELENS, A., HAEGEMAN, A., STOFFELEN, P., MWANGA MWANGA, J. C. I., ... & JANSSENS, S. B. (2021). Phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica*. *Journal of Systematics and Evolution*, 59(5), 953-963.
- BIAZZI, E., NAZZICARI, N., PECETTI, L., BRUMMER, E. C., PALMONARI, A., TAVA, A., & ANNICCHIARICO, P. (2017). Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. *PloS one*, 12(1), e0169234.
- CAPUCHO, A. S., CAIXETA, E. T., ZAMBOLIM, E. M., & ZAMBOLIM, L. (2009). Herança da resistência do Híbrido de Timor UFV 443-03 à ferrugem-do-cafeeiro. *Pesquisa Agropecuária Brasileira*, 44, 276-282.
- CENCI, A., COMBES, M. C., & LASHERMES, P. (2012). Genome evolution in diploid and tetraploid *Coffea* species as revealed by comparative analysis of orthologous genome segments. *Plant molecular biology*, 78(1), 135-145.

CLARINDO, W. R., & CARVALHO, C. R. (2008). First *Coffea arabica* karyogram showing that this species is a true allotetraploid. *Plant Systematics and Evolution*, 274, 237-241.

CLARINDO, W. R., & CARVALHO, C. R. (2009). Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. *Plant Cell Reports*, 28, 73-81.

CONAB, C. N. D. A. (2024). Acompanhamento da Safra Brasileira - Café. [S. l.: s. n.],

CORDEIRO, G. M., TAYLOR, G. O., & HENRY, R. J. (2000). Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. *Plant science*, 155(2), 161-168.

DENOEUDE, F., CARRETERO-PAULET, L., DEREPPER, A., DROC, G., GUYOT, R., PIETRELLA, M., ... & LASHERMES, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*, 345(6201), 1181-1184.

DINIZ, L. E., SAKIYAMA, N. S., LASHERMES, P., CAIXETA, E. T., OLIVEIRA, A. C. B., ZAMBOLIM, E. M., ... & ZAMBOLIM, L. (2005). Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. *Crop Breeding and Applied Biotechnology*, 5(4), 387.

DOYLE, J. J., FLAGEL, L. E., PATERSON, A. H., RAPP, R. A., SOLTIS, D. E., SOLTIS, P. S., & WENDEL, J. F. (2008). Evolutionary genetics of genome merger and doubling in plants. *Annual review of genetics*, 42(1), 443-461.

ENDRIZZI, J. (1962). The diploid-like cytological behavior of tetraploid cotton. *Evolution*, 325-329.

FERRÃO, L. F. V., CAIXETA, E. T., PENA, G., ZAMBOLIM, E. M., CRUZ, C. D., ZAMBOLIM, L., ... & SAKIYAMA, N. S. (2015). New EST-SSR markers of *Coffea arabica*: transferability and application to studies of molecular characterization and genetic mapping. *Molecular Breeding*, 35, 1-5.

- GARCIA, A. A., MOLLINARI, M., MARCONI, T. G., SERANG, O. R., SILVA, R. R., VIEIRA, M. L., ... & SOUZA, A. P. (2013). SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. *Scientific reports*, 3(1), 3399.
- GARRISON, E., & MARTH, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.
- GERARD, D., FERRÃO, L. F. V., GARCIA, A. A. F., & STEPHENS, M. (2018). Genotyping polyploids from messy sequencing data. *Genetics*, 210(3), 789-807.
- GERARD, D., THAKKAR, M., & FERRÃO, L. F. V. (2024). Tests for Segregation Distortion in Tetraploid F1 Populations. *bioRxiv*, 2024-02.
- GUYOT, R., HAMON, P., COUTURON, E., RAHARIMALALA, N., RAKOTOMALALA, J. J., LAKKANNA, S., ... & BONNET, P. (2020). WCSdb: a database of wild *Coffea* species. *Database*, 2020, baaa069.
- HACKETT, C. A., MCLEAN, K., & BRYAN, G. J. (2013). Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. *PloS one*, 8(5), e63939.
- JIAO, Y., WICKETT, N. J., AYYAMPALAYAM, S., CHANDERBALI, A. S., LANDHERR, L., RALPH, P. E., ... & DEPAMPHILIS, C. W. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473(7345), 97-100.
- KY, C. L., BARRE, P., LORIEUX, M., TROUSLOT, P., AKAFFOU, S., LOUARN, J., ... & NOIROT, M. (2000). Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). *Theoretical and Applied Genetics*, 101, 669-676.
- LARA, L. A., SANTOS, M. F., JANK, L., CHIARI, L., VILELA, M. D. M., AMADEU, R. R., ... & GARCIA, A. A. F. (2019). Genomic selection with allele dosage in *Panicum maximum* jacq. *G3: Genes, Genomes, Genetics*, 9(8), 2463-2475.

LASHERMES, P., COMBES, M. C., ROBERT, J., TROUSLOT, P., D'HONT, A., ANTHONY, F., & CHARRIER, A. (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics MGG*, 261(2), 259-266.

LASHERMES, P., PACZEK, V., TROUSLOT, P., COMBES, M. C., COUTURON, E., & CHARRIER, A. (2000). Brief communication. Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific Hybrid *C. arabica* X *C. canephora*. *Journal of Heredity*, 91(1), 81-85.

LASHERMES, P., COMBES, M. C., PRAKASH, N. S., TROUSLOT, P., LORIEUX, M., & CHARRIER, A. (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome*, 44(4), 589-595.

LASHERMES, P., COMBES, M. C., HUEBER, Y., SEVERAC, D., & DEREPPER, A. (2014). Genome rearrangements derived from homoeologous recombination following allopolyploidy speciation in coffee. *The Plant Journal*, 78(4), 674-685.

LASHERMES, P., HUEBER, Y., COMBES, M. C., SEVERAC, D., & DEREPPER, A. (2016). Inter-genomic DNA exchanges and homeologous gene silencing shaped the nascent allopolyploid coffee genome (*Coffea arabica* L.). *G3: Genes, Genomes, Genetics*, 6(9), 2937-2948.

LEAL-BERTIOLI, S. C., GODOY, I. J., SANTOS, J. F., DOYLE, J. J., GUIMARÃES, P. M., ABERNATHY, B. L., ... & BERTIOLI, D. J. (2018). Segmental allopolyploidy in action: increasing diversity through polyploid hybridization and homoeologous recombination. *American journal of botany*, 105(6), 1053-1066.

MAPA, 2024: Ministério da Agricultura e Pecuária.

MCGREGOR, C. E., LAMBERT, C. A., GREYLING, M. M., LOUW, J. H., & WARNICH, L. (2000). A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica*, 113, 135-144.

MCNEIL, M. D., HERMANN, S., JACKSON, P. A., & AITKEN, K. S. (2011). Conversion of AFLP markers to high-throughput markers in a complex polyploid, sugarcane. *Molecular Breeding*, 27, 395-407.

OLIVEIRA, A. C. B. D., SAKIYAMA, N. S., CAIXETA, E. T., ZAMBOLIM, E. M., RUFINO, R. J. N., & ZAMBOLIM, L. (2007). Partial map of *Coffea arabica* L. and recovery of the recurrent parent in backcross progenies.

OROZCO-CASTILLO, C., CHALMERS, K. J., POWELL, W., & WAUGH, R. (1996). RAPD and organelle specific PCR re-affirms taxonomic relationships within the genus *Coffea*. *Plant Cell Reports*, 15, 337-341.

ORTIZ, A. J., & SHARBROUGH, J. (2024). Genome-wide patterns of homoeologous gene flow in allotetraploid coffee. *Applications in Plant Sciences*, 12(4), e11584.

PAILLARD, M., LASHERMES, P., & PÉTIARD, V. (1996). Construction of a molecular linkage map in coffee. *Theoretical and Applied Genetics*, 93, 41-47.

PEARL, H. M., NAGAI, C. H. I. F. U. M. I., MOORE, P. H., STEIGER, D. L., OSGOOD, R. V., & MING, R. (2004). Construction of a genetic map for arabica coffee. *Theoretical and Applied Genetics*, 108, 829-835.

PESTANA, K. N., CAPUCHO, A. S., CAIXETA, E. T., DE ALMEIDA, D. P., ZAMBOLIM, E. M., CRUZ, C. D., ... & SAKIYAMA, N. S. (2015). Inheritance study and linkage mapping of resistance loci to *Hemileia vastatrix* in Híbrido de Timor UFV 443-03. *Tree Genetics & Genomes*, 11, 1-13.

PINTO-MAGLIO, C. A. F., & DA CRUZ, N. D. (1998). Pachytene chromosome morphology in *Coffea* L. II. *C. arabica* L. complement. *Caryologia*, 51(1), 19-35.

PRAKASH, N. S., MARQUES, D. V., VARZEA, V. M. P., SILVA, M. C., COMBES, M. C., & LASHERMES, P. (2004). Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. *Theoretical and Applied Genetics*, 109, 1311-1317.

PRIOLLI, R. H., MAZZAFERA, P., SIQUEIRA, W. J., MÖLLER, M., ZUCCHI, M. I., RAMOS, L. C. S., ... & COLOMBO, C. A. (2008). Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). *Genetics and Molecular Biology*, 31, 498-504.

RESENDE, M., CAIXETA, E.T., & ALKIMIM, E.R. (2016). High-throughput targeted genotyping of *Coffea arabica* and *Coffea canephora* using next generation sequencing. San Diego, CA. 1.

ROMERO, G., VÁSQUEZ, L. M., LASHERMES, P., & HERRERA, J. C. (2014). Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). *Plant Breeding*, 133(1), 121-129.

SALOJÄRVI, J., RAMBANI, A., YU, Z., GUYOT, R., STRICKLER, S., LEPELLEY, M., ... & DESCOMBES, P. (2024). The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. *Nature genetics*, 56(4), 721-731.

SCALABRIN, S., TONIUTTI, L., DI GASPERO, G., SCAGLIONE, D., MAGRIS, G., VIDOTTO, M., ... & BERTRAND, B. (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Scientific reports*, 10(1), 1-13.

SCALABRIN, S., MAGRIS, G., LIVA, M., VITULO, N., VIDOTTO, M., SCAGLIONE, D., ... & MORGANTE, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, 15(1), 463.

SERA, G. H., DE CARVALHO, C. H. S., DE REZENDE ABRAHÃO, J. C., POZZA, E. A., MATIELLO, J. B., DE ALMEIDA, S. R., ... & DOS SANTOS BOTELHO, D. M. (2022). Coffee leaf rust in Brazil: Historical events, current situation, and control measures. *Agronomy*, 12(2), 496.

- SERANG, O., MOLLINARI, M., & GARCIA, A. A. F. (2012). Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. *PLoS One*, 7(2), e30906.
- SETOTAW, T. A., CAIXETA, E. T., PEREIRA, A. A., DE OLIVEIRA, A. C. B., CRUZ, C. D., ZAMBOLIM, E. M., ... & SAKIYAMA, N. S. (2013). Coefficient of parentage in *Coffea arabica* L. cultivars grown in Brazil. *Crop Science*, 53(4), 1237-1247.
- SILVA, M. D. C., VÁRZEA, V., GUERRA-GUIMARÃES, L., AZINHEIRA, H. G., FERNANDEZ, D., PETITOT, A. S., ... & NICOLE, M. (2006). Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian journal of plant physiology*, 18, 119-147.
- SILVA, M. D. C., GUERRA-GUIMARÃES, L., DINIZ, I., LOUREIRO, A., AZINHEIRA, H., PEREIRA, A. P., ... & VÁRZEA, V. (2022a). An overview of the mechanisms involved in *Coffea-Hemileia vastatrix* interactions: Plant and pathogen perspectives. *Agronomy*, 12(2), 326.
- SILVA, F.L., ALKIMIM, E. R., BARREIRO, P. R. R. M., LEICHTWEIS, B. G., SILVA, A. C. A., DA SILVA, R. A., ... & CAIXETA, E. T. (2022b). Genome-wide association study of plant architecture and diseases resistance in *Coffea canephora*. *Euphytica*, 218(7), 92.
- SILVA, R. A.; ZAMBOLIM, L.; CASTRO, I. S. L.; RODRIGUES, H. S.; CRUZ, C. D.; CAIXETA, E. T. (2018) The Híbrido de Timor germplasm: identification of molecular diversity and resistance sources to coffee berry disease and leaf rust. *Euphytica*, v. 214, n. 9, p. 153.
- SOLTIS, D. E., VISGER, C. J., & SOLTIS, P. S. (2014). The polyploidy revolution then... and now: Stebbins revisited. *American journal of botany*, 101(7), 1057-1078.
- SOUSA, T. V; CAIXETA, E. T.; ALKIMIM, E. R.; DE OLIVEIRA, A. C. B.; PEREIRA, A. A.; SAKIYAMA, N. S. JÚNIOR, M. F. R. R.; ZAMBOLIM, L. (2017). Population structure and genetic diversity of coffee progenies derived from Catuaí and Híbrido de Timor revealed by genome-wide SNP marker. *Tree genetics & genomes*, v. 13, n. 6, p. 124.
- TEIXEIRA-CABRAL, T. A., SAKIYAMA, N. S., ZAMBOLIM, L., PEREIRA, A. A., & SCHUSTER, I. (2004). Single-locus inheritance and partial linkage map of *Coffea arabica* L. *Crop Breeding and Applied Biotechnology*, 4(4).

USDA, U. S. D. of A. A. R. S. (2024). Coffee: World Markets and Trade Coffee.

VIEIRA, L. G. E., ANDRADE, A. C., COLOMBO, C. A., MORAES, A. H. D. A., METHA, Â., OLIVEIRA, A. C. D., ... & PEREIRA, G. A. G. (2006). Brazilian coffee genome project: an EST-based genomic resource. *Brazilian Journal of Plant Physiology*, 18, 95-108.

VOORRIPS, R. E., GORT, G., & VOSMAN, B. (2011). Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC bioinformatics*, 12, 1-11.

## CAPÍTULO 3

**IDENTIFICAÇÃO DE QTL ASSOCIADOS À RESISTÊNCIA A *Hemileia vastatrix* EM *Coffea arabica* COM ANÁLISE INÉDITA PARA A RAÇA XXXIII DO PATÓGENO**

## Resumo

A cafeicultura mundial enfrenta desafios significativos causados por pragas e doenças, sendo a ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix*, uma das doenças mais prejudiciais à produção de *Coffea arabica*. A ferrugem pode acarretar perdas de mais de 30-50% na produção cafeeira, sendo o plantio de cultivares de café resistentes a esse fungo a medida mais eficaz e econômica no controle dessa doença. No entanto, a emergência de novas raças de *H. vastatrix*, capazes de suplantam a resistência das cultivares existentes, representa um desafio contínuo. Os programas de melhoramento genético têm buscado estratégias para aumentar a durabilidade da resistência à *H. vastatrix* em cultivares de café arábica, através de seleção de progênies superiores derivadas de cruzamentos intra e interespecíficos, para a transferência de genes de resistência à ferrugem do cafeeiro. O Híbrido de Timor (HdT), um híbrido natural entre *C. arabica* e *C. canephora*, e as progênies derivadas do seu cruzamento com outras cultivares, tem se destacado como uma fonte valiosa de resistência à ferrugem, combinando características agronômicas desejáveis e resistência ao patógeno. Por sua alta compatibilidade com cultivares de *C. arabica*, o HdT facilita a transferência de genes de resistência e tem sido amplamente utilizado em programas de melhoramento, em combinação com abordagens como a seleção assistida por marcadores (SAM) e a piramidação de genes para obtenção de resistência durável. Um método promissor para viabilizar essa estratégia é a identificação de QTL associados a características de interesse, como à resistência. Esse estudo teve como objetivo identificar QTL em *C. arabica* relacionados à resistência às raças I, II, XXXIII e ao patótipo 001 de *H. vastatrix*, utilizando genotipagem em larga escala com marcadores SNP e mapeamento de QTL por análise de marcas individuais. A população analisada corresponde a uma F<sub>2</sub> segregante para resistência a ferrugem, proveniente de cruzamento com genótipo oriundo de HdT. Dados fenotípicos da raça XXXIII foram obtidos em experimentos conduzidos em casa de vegetação, enquanto informações sobre as raças I, II e o patótipo 001 foram extraídas de estudos prévios. A análise de herança indicou que a resistência à raça XXXIII é condicionada pela ação de três genes. E a identificação de QTL revelou regiões genômicas associadas à resistência a diferentes raças de *H. vastatrix*, incluindo marcas inéditas para a resistência à raça XXXIII e novas marcas para as raças I, II e o patótipo 001. Essas regiões genicas oferecem suporte direto à SAM, permitindo a seleção eficiente de genótipos portadores de genes de resistência e sua posterior introgressão em outras variedades. Assim, os resultados obtidos contribuem significativamente para o entendimento da herança genética de uma importante fonte de resistência a *H. vastatrix* e fortalecem as estratégias de

melhoramento de *C. arabica*, promovendo o desenvolvimento de cultivares com resistência durável às diferentes raças desse patógeno.

**Palavras-chave:** Ferrugem do cafeeiro; Genes de resistência; Marcador molecular; Melhoramento do cafeeiro; SNP.

## Abstract

Coffee production worldwide faces significant challenges caused by pests and diseases. Coffee rust, caused by the biotrophic fungus *Hemileia vastatrix*, has been one of the most damaging diseases to *Coffea arabica* production. Rust can cause losses of more than 30-50% in coffee production, and planting coffee cultivars resistant to this fungus is the most effective and economical measure to control this disease. However, the emergence of new *H. vastatrix* races capable of overcoming the resistance of existing cultivars represents an ongoing challenge. Breeding programs have sought strategies to increase the durability of resistance to *H. vastatrix* in arabica coffee cultivars by selecting superior progenies derived from intra and interspecific crosses for the transfer of coffee rust resistance genes. The Timor Hybrid (HdT), a natural hybrid between *C. arabica* and *C. canephora*, and the progenies derived from its crossing with other cultivars, has emerged as a valuable source of rust resistance, combining desirable agronomic traits and resistance to the pathogen. Due to its high compatibility with *C. arabica* cultivars, HdT facilitates the transfer of resistance genes and has been widely used in breeding programs, in combination with approaches such as marker-assisted selection (MAS) and gene pyramiding to obtain durable resistance. A promising method to make this strategy viable is the identification of QTL associated with traits of interest, such as resistance. This study aimed to identify QTL in *C. arabica* related to resistance to races I, II, XXXIII and pathotype 001 of *H. vastatrix*, using large-scale genotyping with SNP markers and QTL mapping by analysis of individual markers. The population analyzed corresponds to an F<sub>2</sub> segregating for rust resistance, resulting from a cross with a genotype originating from HdT. Phenotypic data for race XXXIII were obtained in experiments conducted in a greenhouse, while information on races I, II and pathotype 001 was extracted from previous studies. The inheritance analysis indicated that resistance to race XXXIII is conditioned by the action of three genes. The identification of QTL showed genomic regions associated with resistance to different races of *H. vastatrix*, including novel markers for resistance to race XXXIII and new markers for races I, II and pathotype 001. These genomic regions offer direct support to SAM, allowing the efficient selection of genotypes carrying resistance genes and their subsequent introgression into other varieties. Thus, the results obtained contribute significantly to the understanding of the genetic inheritance of an important source of resistance to *H. vastatrix* and strengthen the breeding strategies of *C. arabica*, promoting the development of cultivars with durable resistance to the different races of this pathogen.

**Keywords:** Coffee rust; Resistance genes; Molecular marker; Coffee breeding; SNP.

## 1. Introdução

A cafeicultura mundial enfrenta desafios significativos causados por pragas e doenças, sendo a ferrugem do cafeeiro uma das doenças mais prejudiciais à produção de *Coffea arabica* (HARELIMANA et al., 2022; ZAMBOLIM et al., 2024). Essa espécie representa, em média, 57% da produção mundial de café, enquanto *Coffea canephora* (variedades Robusta e Conilon) responde os 43% restantes (CONAB, 2024; MAPA, 2024). O fungo biotrófico *Hemileia vastatrix* (ordem Uredinales, família Pucciniaceae) é o patógeno causador da ferrugem do cafeeiro e infecta exclusivamente o gênero *Coffea*, acarretando perdas de mais de 30-50% na produção (ZAMBOLIM, 2016; ZAMBOLIM et al., 2024). Essas perdas ocorrem devido à suscetibilidade da maioria das variedades comerciais de *C. arabica* plantadas, enquanto *C. canephora* apresenta variedades resistentes a esse fungo (CARVALHO et al., 2010; SILVA et al., 2022a).

A base genética estreita de *C. arabica* tem sido reconhecida como uma limitação significativa da resistência/tolerância das variedades a estresses bióticos e abióticos (ANTHONY et al., 2002; SCALABRIN et al. 2020; SALOJÄRVI et al., 2024). Essa limitação é atribuída à sua recente origem alotetraploide, estimada entre 350.000 e 610.000 anos, seguida por gargalos de pré-domesticação que contribuíram para a redução da variabilidade genética (SALOJÄRVI et al., 2024). A origem dessa espécie supostamente ocorreu a partir da hibridação natural entre duas espécies diploides ( $2n = 2x = 22$  cromossomos), *C. eugenoides* e *C. canephora* (LASHERMES, et al., 1999), seguida de uma duplicação de todo o genoma, provavelmente durante um único evento de aloploidização ( $2n = 4x = 44$  cromossomos) (CLARINDO & CARVALHO, 2009; LASHERMES, et al., 2014; SCALABRIN et al., 2020; BAWIN et al., 2021).

O plantio de cultivares de café resistentes ao fungo causador da ferrugem do cafeeiro é a medida mais eficaz e econômica de controle dessa doença, embora tratamentos químicos, como o uso de fungicidas, possam ser realizados de forma preventiva (WALLER et al., 2007; ZAMBOLIM, 2016; ZAMBOLIM et al., 2024). Entretanto, ao longo do tempo pode haver a ocorrência de novas raças do patógeno, para as quais as plantas não possuem resistência, já que esse fungo é capaz de suplantar a resistência das cultivares desenvolvidas (ZAMBOLIM & CAIXETA, 2021; ZAMBOLIM et al., 2024). Até o momento foram caracterizadas mais de 50 raças de *H. vastatrix*, sendo identificadas no Brasil 16 delas (I, II, III, VII, X, XIII, XV, XVI, XVII, XXI, XXII, XXIII, XXIV, XXV ou XXXI, XXXIII, XXXVII). A raça

II está presente em maior frequência nas lavouras de café do Brasil e a raça XXXIII tem sido reportada como a mais complexa, contendo dois ou três genes de virulência (VARZEA et al., 2005; ZAMBOLIM et al., 2005; CAPUCHO et al., 2012; ZAMBOLIM & CAIXETA, 2021; ZAMBOLIM et al., 2024).

Por muitos anos, a seleção para resistência a *H. vastatrix* foi baseada na resistência completa, altamente específica à raça do patógeno, originada dos principais genes  $S_H$  introgrididos, derivados de *C. arabica*, *C. canephora* ou *C. liberica*. No entanto, o alto grau de variabilidade e adaptabilidade do patógeno, resultou na perda gradual da resistência baseada nos principais genes  $S_H$ . Os genes  $S_H$  de *C. arabica* já foram suplantados pelas raças de *H. vastatrix* existentes no Brasil (BETTENCOURT, 1973; DINIZ, et al., 2012; ZAMBOLIM & CAIXETA, 2021; ZAMBOLIM et al., 2024). A resistência poligênica parcial e não específica a *H. vastatrix* foi identificada em *C. canephora*, em alguns genótipos de *C. arabica* e em híbridos interespecíficos dessas espécies (HERRERA et al., 2009; ROMERO et al., 2010).

Os programas de melhoramento do cafeeiro buscam desenvolver estratégias para melhorar a durabilidade da resistência à *H. vastatrix* em cultivares de café arábica, através da combinação de diferentes genes de resistência qualitativa (completa ou raça-específica) e de quantitativa (parcial ou não raça específica) (BARKA et al., 2020; ALMEIDA et al., 2021b; SERA et al., 2022; SILVA et al., 2022a; ALVES et al., 2024). Esses programas visam à seleção de progênies superiores derivadas de cruzamentos intra e interespecíficos para a transferência de genes de resistência à ferrugem do cafeeiro (SAKIYAMA, 2005; SERA et al., 2022).

Os genes  $S_{H6}$ ,  $S_{H7}$ ,  $S_{H8}$  e  $S_{H9}$ , identificados em *C. canephora*, estão presentes no Híbrido de Timor (HdT), um híbrido natural entre *C. arabica* e *C. canephora*. O HdT e as progênies derivadas do seu cruzamento com outras cultivares, vêm sendo estudadas em diversas regiões produtoras de café no mundo, devido a sua resistência a *H. vastatrix*. O seu germoplasma tem sido utilizado em programas de melhoramento que visam a resistência durável à ferrugem e outras doenças do cafeeiro (CAPUCHO et al., 2009; ALMEIDA et al., 2021a, 2021b; SERA et al., 2022; SILVA et al., 2006, 2022a; ALVES et al., 2024). O uso eficiente dos recursos disponíveis nesses híbridos naturais resistentes é fundamental para o melhoramento contínuo das variedades de *C. arabica*, já que o HdT, além de apresentar apresentam características agronômicas relevantes, dentre elas a resistência à ferrugem, os seus genótipos se cruzam facilmente com as cultivares de *C. arabica*, favorecendo a transferência de genes (CARVALHO et al., 1989; LASHERMES et al., 2009; SILVA et al., 2022a; ZAMBOLIM et al., 2024).

Nos programas de melhoramento do cafeeiro, os HdT têm sido utilizados como fonte de resistência aliada a estratégia de seleção assistida por marcadores (SAM) e piramidação de genes, para se obter resistência durável a *H. vastatrix*. Um dos métodos mais eficazes para viabilizar a SAM é a identificação de QTL (*Quantitative Trait Loci*) associados à característica de interesse (ALKIMIM et al., 2017; ALMEIDA et al., 2021a; MEDEIROS et al., 2021; SAAVEDRA et al., 2023; FEITOSA et al., 2024). Entre as aplicações dos estudos de QTL nos programas de melhoramento, se encontram a introgressão de características de interesse em cultivares comerciais, escolha de genitores superiores e clonagem de genes (ALMEIDA et al., 2021a, 2021b; ARORA et al., 2021; XIONG et al., 2021; WANG et al., 2023; ZHAO et al., 2023).

Mapas genéticos para a identificação de QTL associados a características de importância para os programas de melhoramento do cafeeiro, foram obtidos usando marcadores moleculares desenvolvidos para espécies diploides, como RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) e SSR (*Simple Sequence Repeat*). (LASHERMES et al., 2001; PEARL et al., 2004; PRAKASH et al., 2004; TEIXEIRA-CABRAL et al., 2004; PRIOLLI et al., 2008; ROMERO et al., 2014; PESTANA et al., 2015; FERRÃO et al., 2015; MONCADA et al., 2016). Posteriormente, foram desenvolvidos dois mapas genéticos utilizando marcadores de alta densidade, SNP, para identificação de QTL associado a características de interesse para o melhoramento de *C. arabica*, um dos estudos foi realizado para identificar QTL associados à produção de café, altura da planta e tamanho do grão (MONCADA et al., 2016), no outro para a análise de QTL associado a resistência à *Coffee Berry Disease* (CBD) (GIMASE et al., 2020). Entretanto, a construção de mapas genéticos associada ao estudo de QTL que governam as características de interesse em *C. arabica* continuam limitados, devido à complexidade do genoma alotetraploide, a sua baixa diversidade genética e a presença de cromossomos homeólogos (LASHERMES et al., 2014; MÉROT-L'ANTHOËNE et al., 2019; SATTLER et al., 2022; SCALABRIN et al., 2020, 2024).

A análise de marcas individuais (SMA - *Single Marker Analysis*) é uma técnica robusta que avalia a associação entre um marcador específico e um fenótipo. Por meio da comparação das médias fenotípicas entre os diferentes genótipos do marcador, a técnica detecta diferenças significativas, permitindo a identificação de QTL (MCCOUCH & DOERGE et al., 1995; COLLARD et al., 2005). Assim, não é necessário o uso de um mapa de ligação para a busca de QTL associados as características de interesse. Desse modo, o objetivo deste trabalho foi identificar QTL em *C. arabica* associados à resistência às raças I, II, XXXIII e ao patótipo 001

de *H. vastatrix*, utilizando genotipagem baseada em marcadores SNP e mapeamento de QTL por meio da abordagem SMA, considerando a ploidia da espécie. Este estudo visa ampliar o entendimento dos genes envolvidos na resistência do cafeeiro ao patógeno e identificar marcadores moleculares que possam auxiliar de forma mais eficiente nos programas de melhoramento genético da cultura.

## 2. Material e métodos

### 2.1. População experimental

A população utilizada nesse estudo corresponde a uma F<sub>2</sub> segregante para resistência a ferrugem. A planta F, autofecundada para dar origem a população segregante foi originada do cruzamento entre Híbrido de Timor UFV 443-03 (genitor masculino), planta resistente à ferrugem do cafeeiro, e Catuaí Amarelo IAC 64 (UFV 2148-57), cultivar suscetível. Os cruzamentos foram realizados por Pereira (1995) no Departamento de Fitopatologia, na Universidade Federal de Viçosa (Viçosa-MG) (20°45'14"S, 42°52'53"O) – UFV e nas Fazendas Experimentais da EPAMIG. A planta da geração F<sub>1</sub> (H 511-1), foi submetida à autofecundação controlada, para obtenção da geração segregante F<sub>2</sub>, com 246 plantas que se encontram em campo experimental da UFV (CAPUCHO et al., 2009).

### 2.2. Avaliações fenotípicas

Foi conduzido em casa de vegetação (UFV) um experimento para avaliar a resistência e a suscetibilidade dos genótipos da população F<sub>2</sub> segregante à raça XXXIII de *H. vastatrix*. Em primeiro momento, os urediniósporos da raça XXXIII de *H. vastatrix*, pertencente à micoteca do BioCafé, foram multiplicados e posteriormente utilizados como inóculo de acordo com a metodologia descrita por Capucho et al. (2009). Foram obtidos 16 discos foliares de 1,5 cm de diâmetro para cada um dos 232 genótipos da população F<sub>2</sub> segregante e para os parentais resistente (HdT- UFV 443-03) e suscetível (Catuaí-UFV 2148-57), para constituírem as unidades amostrais. Os urediniósporos foram depositados na face abaxial de cada disco com o auxílio de um pincel. Após a inoculação, os discos foram colocados em caixas gerbox (caixa plástica 11 x 11 x 33 cm) contendo espuma, saturado em água e tela de metal. As caixas gerbox foram fechadas e mantidas no escuro por 48h em uma câmara a 22°C e posteriormente a 22°C

mantidos sob um fotoperíodo de 12h (luz/escuro). Os discos foram limpos com algodão após às 48h para evitar o aparecimento de hiperparasitas (ESKES, 2019).

A avaliação de resistência ou suscetibilidade foi iniciada quando o genótipo suscetível (Catuai-UFV 2148-57) começou a esporular, aproximadamente 25 dias após a inoculação. A avaliação seguiu a escala de Tamayo et al. (1995), que se baseia na ausência ou presença de sinais. Plantas com pontuação de 1 a 3 foram consideradas resistentes (1 = sem sintomas, 2 = pequenas lesões cloróticas e 3 = grandes lesões cloróticas sem esporulação). Plantas com pontuação de 4 a 6 foram consideradas suscetíveis, pois apresentavam urediniósporos (4 = <25% da área da lesão com urediniósporos, 5 = 25–50% da área da lesão com urediniósporos e 6 = >50% da lesão área com urediniósporos). Amostras de folhas foram coletadas e inoculadas com urediniósporos da raça XXXIII de *H. vastatrix* em dois tempos diferentes, novembro (2022) e abril (2023).

Com o intuito de verificar a segregação fenotípica para resistência/suscetibilidade das plantas (estudo de herança), foi utilizado o teste qui-quadrado, com o auxílio do *software* Genes (CRUZ, 2013). Foram testadas as proporções fenotípicas esperadas para um gene (3:1 e 1:3), dois genes (15:1, 13:3, 7:9, 9:7, 3:13 e 1:15) e três genes (63:1, 61:3, 55:9, 37:27, 27:37, 9:55, 3:61, 1:63, 57:7, 51:13, 49:15, 43:21, 25:39 e 19:45), significativos a 5% de probabilidade.

Os dados fenotípicos das raças I e II e patótipo 001 foram obtidos de Pestana et al. (2015), os quais foram coletados seguindo mesma metodologia descrita acima para a raça XXXIII.

## **2.3. Genotipagem**

### **2.3.1. Extração de DNA**

Foram coletados um par de folhas jovens de cada uma das 232 plantas da população F<sub>2</sub> segregante, da F<sub>1</sub>, H 511-1, e dos genitores da F<sub>1</sub>, UFV 443-03 e UFV 2148-57. Esse material biológico foi enviado ao Laboratório de Biotecnologia do Cafeeiro (BioCafé/UFV), onde foi realizada a extração do DNA genômico conforme a metodologia descrita por (DINIZ et al., 2005). Posteriormente, a qualidade do DNA foi avaliada em gel de agarose (1%) e quantificada com espectrofotômetro Qubit DNA BR (*Life Technologies*). A concentração final de DNA foi ajustada para 20 ng/μL.

### 2.3.2. Genotipagem com marcadores SNP

Após a extração e análise de qualidade, as amostras de DNA foram enviadas para a empresa *Rapid Genomics*, Flórida (EUA), para a genotipagem por sequenciamento de regiões genômicas. Essas regiões sequenciadas foram inicialmente identificadas por Resende et al. (2016). Estes autores desenharam sondas com base em sequências disponíveis em bancos de dados de *Expressed Sequence Tags* (ESTs) desenvolvidos pelo Projeto Genoma Brasileiro de Café (VIEIRA et al., 2006) e também do genoma de referência de *C. canephora* (DENOEU et al., 2014), com o objetivo de cobrir todo o genoma de *C. arabica*. Para o presente trabalho, foram selecionadas 400 sondas polimórficas mais informativas, contendo maior número de SNP, com base em trabalhos anteriores (SOUSA et al., 2017; SILVA et al., 2018; SILVA et al., 2022b). Foram desenhados pares de primers para a amplificação destas regiões usando a tecnologia *Flex-Seq EX-L* (*Rapid Genomics*, 2021). A clonagem e o sequenciamento das sondas amplificadas foram realizados utilizando a plataforma *Illumina HiSeq*.

### 2.4. Análise dos dados genômicos

Com o uso do *software* bowtie2, as leituras obtidas com o sequenciamento foram mapeadas utilizando o genoma de *C. arabica* como referência, disponível no NCBI (GCF\_003713225.1), e com base nas informações de alinhamento, foram feitas as identificações das variantes. Foram realizados mapeamentos separados para cada um dos subgenomas (porção *C. canephora* – C<sup>c</sup> e porção *C. eugenioides* - C<sup>e</sup>) do genoma de referência com chamadas de variantes tetraploides (0, 1, 2, 3, 4). Os códigos que representam os cromossomos dos subgenomas (C<sup>c</sup> e C<sup>e</sup>) do genoma de *C. arabica* estão disponíveis na Tabela 1. As análises posteriores dos VCF gerados foram feitas utilizando *scripts* personalizados no *software* R (4.2.3).

A fim realizar a chamada de dose e estudar de forma mais precisa as distribuições dos genótipos da população em estudo, foi utilizada a ferramenta updog, para gerar probabilidades de genótipo e distribuições posteriores de genótipo para cada indivíduo em cada SNP (GERARD et al., 2018). Com base nas contagens de leitura contidas nos VCF, foram inferidas as dosagens de alelos de referência e alternativos usando modelo desenvolvido para o estudo de populações poliploides. O modelo utilizado foi o “s1pp”, o qual considera uma população

tetraploide resultantes de uma única geração de autofecundação (geração S1), estimando com precisão a frequência de alelos e evidenciando classes genóticas que antes eram confundidas (GERARD et al., 2018).

Tabela 1: Relação entre os códigos que representam os cromossomos dos subgenomas (C<sup>c</sup> e C<sup>e</sup>) do genoma de *C. arabica*, disponível no NCBI.

<b>Cromossomo</b>	<b>Código</b>	<b>Comprimento (bp)</b>
1c	NC_039898.1	50.636.588
1e	NC_039899.1	48.756.970
2c	NC_039900.1	66.155.350
2e	NC_039901.1	71.633.312
3c	NC_039902.1	41.566.753
3e	NC_039903.1	37.271.464
4c	NC_039904.1	41.786.336
4e	NC_039905.1	42.507.429
5e	NC_039906.1	39.439.615
5c	NC_039907.1	45.899.693
6c	NC_039908.1	55.181.588
6e	NC_039909.1	52.240.725
7c	NC_039910.1	38.854.053
7e	NC_039911.1	35.392.230
8e	NC_039912.1	45.117.557
8c	NC_039913.1	39.008.463
9c	NC_039914.1	38.064.651
9e	NC_039915.1	35.839.895
10e	NC_039916.1	40.458.934
10c	NC_039917.1	45.429.025
11c	NC_039918.1	36.215.491
11e	NC_039919.1	42.465.768

## 2.5. Identificação de QTL

A identificação de QTL associados à resistência de *C. arabica* às raças I, II, patótipo 001 de *H. vastatrix* (conforme avaliado por Pestana et al., 2015) e à raça XXXIII avaliada neste estudo foi realizada através das seguintes etapas. Foram analisados os QTL para os marcadores que não apresentaram distorção de segregação, alinhados com o subgenoma C<sup>c</sup> e, separadamente, com o subgenoma C<sup>e</sup> de *C. arabica*. A metodologia empregada para a identificação dos QTL foi a análise de marcadores individuais (SMA - *Single Marker Analyses*), utilizando o teste de regressão linear simples. Esta análise foi conduzida no *software* R (versão 4.2.3) e a significância do efeito de cada marcador foi avaliada com base no valor de  $-\log_{10}$  (p-valor) superior a 3, equivalente a um *p-valor* de 0,001. Os marcadores foram então representados nos cromossomos correspondentes aos subgenomas de *C. arabica*. Adicionalmente, foi estimado o

efeito aditivo do marcador com a maior associação para cada fenótipo avaliado, identificado pelo menor p-valor.

### 3. Resultados

Os dados referentes a maior nota na avaliação da resistência/suscetibilidade dos genótipos da população F<sub>2</sub> segregante para resistência às raças I, II e XXXIII e, patótipo 001 de *H. vastatrix*, foram descritos na tabela do material suplementar 1.

Na avaliação da raça XXXIII, o parental HdT (UFV 443-03) inoculado com o patógeno se mostrou resistente, enquanto o Catuaí (UFV 2148-57) foi suscetível (Material suplementar 1). Na população F<sub>2</sub>, obteve-se dados fenotípicos para 230 plantas, sendo observadas 199 plantas resistentes (86,5%), notas variando de 1 a 3 e 31 suscetíveis (13,5%), notas variando de 4 a 6 (Figura 1 e Material suplementar 1).

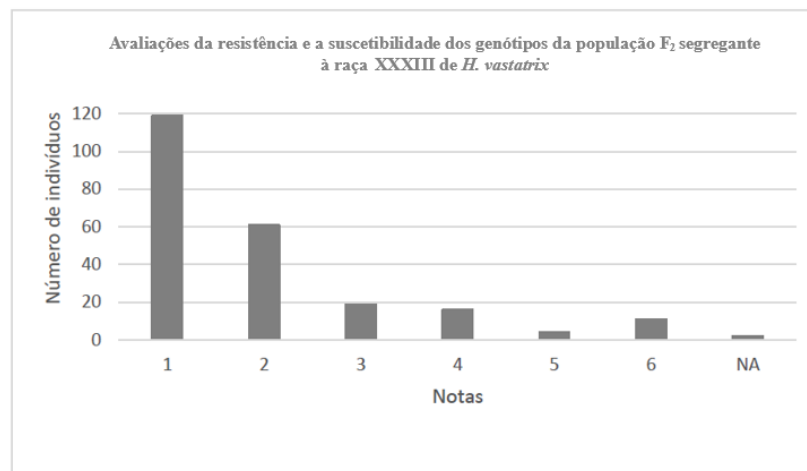


Figura 1: Número de indivíduos da população F<sub>2</sub> pertencentes as diferentes notas associadas a resistência (1 a 3) e a suscetibilidade (4 a 6) à raça XXXIII de *H. vastatrix*.

No estudo de herança da resistência, o teste de qui-quadrado apresentou duas hipóteses de segregação significativas a 5%, ambas envolvendo o efeito de três genes. Uma das hipóteses envolvendo um gene dominante e dois recessivos complementares (55:9) e a outra um gene com efeito dominante independente e dois genes dominantes complementares (57:7) (Tabela 2).

Tabela 2: Proporção esperada, observada e resultado do teste de qui-quadrado para dois e três genes responsáveis pela resistência na população F<sub>2</sub> segregante.

Proporção esperada		Número observado		Número esperado		X <sup>2</sup>	Probabilidade (%)
Resistentes	Suscetíveis	Resistentes	Suscetíveis	Resistentes	Suscetíveis		
3	1	199	31	172,5	57,5	16,284	0,0054
15	1	199	31	215,625	14,375	20,5089	0,0005
13	3	199	31	186,875	43,125	4,1957	4,0525
7	9	199	31	100,625	129,375	170,9783	0,0
9	7	199	31	129,375	100,625	85,6449	0,0
3	13	199	31	43,125	186,875	693,4265	0,0
1	15	199	31	14,375	215,625	2529,3089	0,0
63	1	199	31	226,4062	3,5937	212,3199	0,0
61	3	199	31	219,2187	10,7812	39,7822	0,0
55	9	199	31	197,6562	32,3437	0,0649	79,8817
37	27	199	31	132,9687	97,0312	77,7258	0,0
27	37	199	31	97,0312	132,9687	185,3535	0,0
9	55	199	31	32,3437	197,6562	999,2407	0,0
3	61	199	31	10,7812	219,2187	3447,5199	0,0
1	63	199	31	3,5937	226,4062	10793,6532	0,0
57	7	199	31	204,8437	25,1562	1,5242	21,6984
51	13	199	31	183,2812	46,7187	6,6367	0,9989
49	15	199	31	176,0937	53,9062	12,7131	0,0363
43	21	199	31	154,5312	75,4687	38,999	0,0
25	39	199	31	89,8437	140,1562	217,6329	0,0
19	45	199	31	68,2812	161,7187	355,9112	0,0

Para os dados genotípicos, os marcadores SNP identificados na população F<sub>2</sub> forma submetidos à análise de segregação, além de serem mapeados contra o genoma de *C. arabica*. No subgenoma C<sup>c</sup>, obteve-se 146 marcadores SNP não distorcidos e associados a resistência. Nesse subgenoma, 62 marcadores SNP foram significativamente associados à resistência à raça XXXIII, 39 à raça I, 33 à raça II e 113 ao patótipo 001. Os marcadores associados à resistência à raça XXXIII estão presentes nos cromossomos 1, 2 e 11, à raça I nos cromossomos 1, 2, 3, 4, 5, 6, 7, 9 e 11, à raça II nos cromossomos 1, 2, 6 e 7 e ao patótipo 001 nos cromossomos 1, 4, 7 e 11, referentes ao genoma de *C. arabica* (Figura 2). O efeito aditivo do marcador que apresentou maior associação à resistência à raça XXXIII, raça I, raça II e patótipo 001 estão apresentados nos cromossomos 1, 5, 7 e 1, respectivamente.

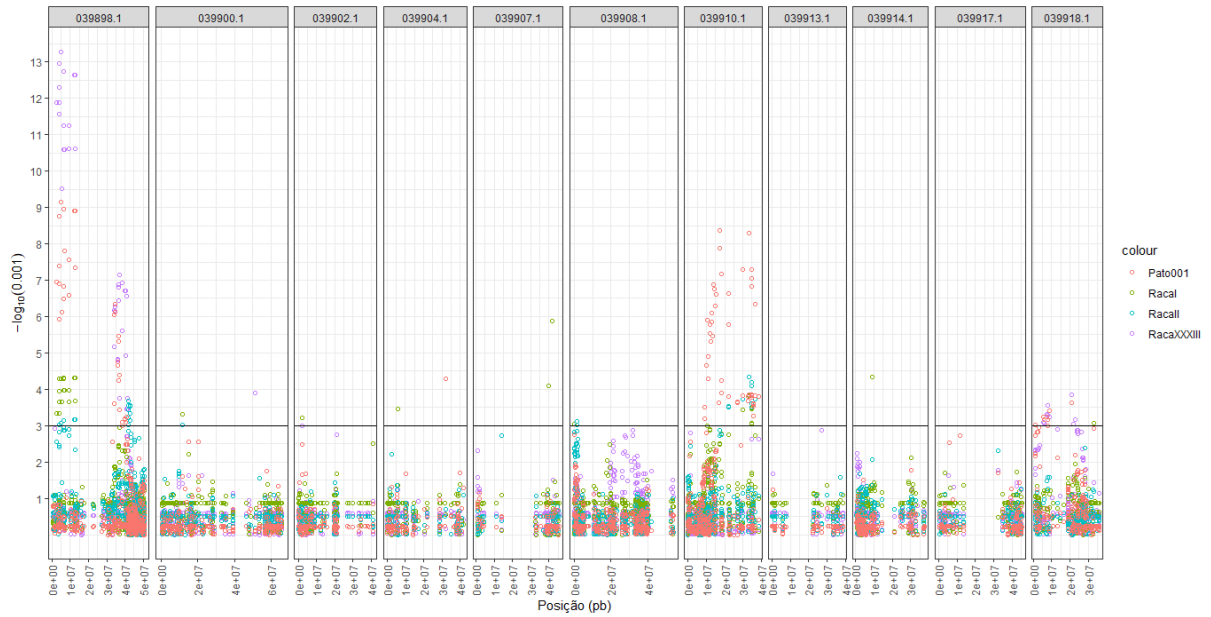


Figura 2. Distribuição dos marcadores SNP associados à resistência a três raças e ao patótipo 001 de *H. vataatrix*, nos cromossomos do subgenoma  $C^c$  de *C. arabica*.

O mapeamento contra o genoma de *C. arabica*, porção  $C^c$ , resultou em 1462 marcadores SNP não distorcidos e associados a resistência. Nesse subgenoma, 60 marcadores SNP foram significativamente associados à resistência à raça XXXIII, 43 à raça I, 42 à raça II e 134 ao patótipo 001. Os marcadores associados à resistência à raça XXXIII estão presentes nos cromossomos 1, 8 e 11, à raça I nos cromossomos 1, 3, 5 e 7, à raça II nos cromossomos 1, 6 e 7 e ao patótipo 001 nos cromossomos 1, 2, 7 e 11, referentes ao genoma de *C. arabica* (Figura 3). O efeito aditivo do marcador que apresentou maior associação à resistência à raça XXXIII, raça I, raça II e patótipo 001 estão apresentados nos cromossomos 1, 5, 7 e 1, respectivamente.

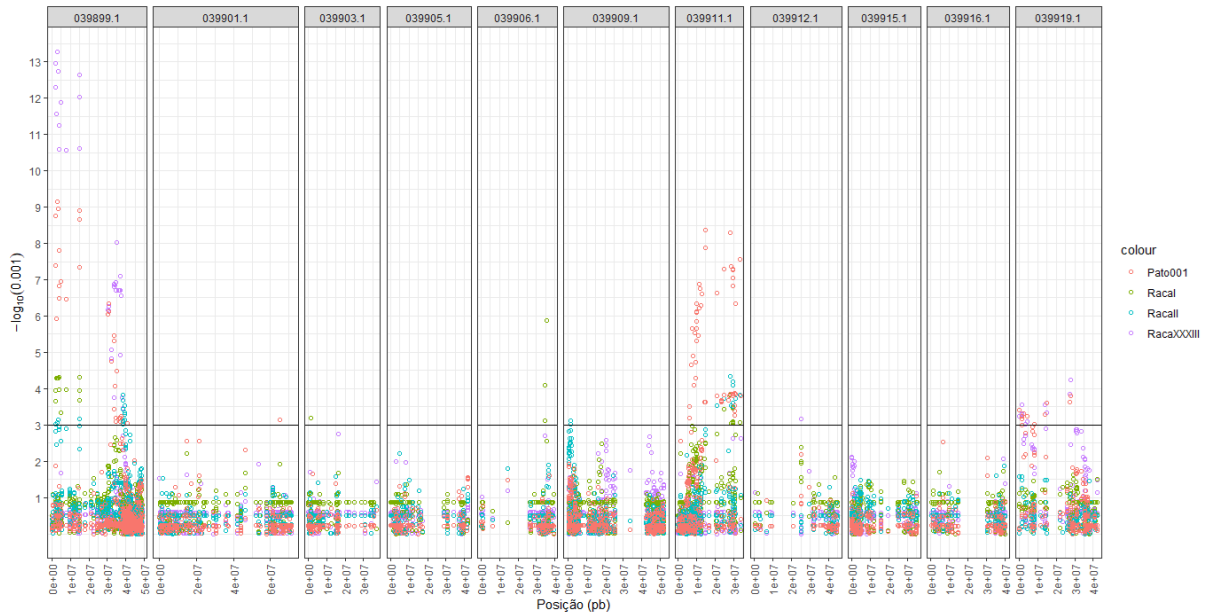


Figura 3. Distribuição dos marcadores SNP associados à resistência à raça três raças de *H. vastatrix* e ao patótipo 001, em estudo, nos cromossomos do subgenoma  $C^c$  de *C. arabica*.

#### 4. Discussão

A resistência à ferrugem do cafeeiro é governada por pelo menos 11 genes dominantes ( $S_{H1}$  a  $S_{H11}$ ) (BETTENCOURT et al., 1988; BARKA et al., 2020; ALMEIDA et al., 2021b). No HdT, estão presentes os genes  $S_{H6}$ - $S_{H11}$ , identificados em *C. canephora*, além de outros ainda não caracterizados (SETOTAW et al. 2020; BARKA et al., 2020; ALMEIDA et al., 2021b). A teoria de Flor (1971) sugere que para cada gene que condiciona a reação de resistência no hospedeiro há um gene correspondente no parasita que condiciona a patogenicidade. No patossistema *Coffea-H. vastatrix*, as diferentes raças do fungo podem expressar a combinação de diferentes genes de virulência (VÁRZEA et al., 2005; ZAMBOLIM et al., 2021). A raça XXXIII de *H. vastatrix*, utilizada neste estudo, é considerada complexa pois contém dois ou três genes de virulência (v 5,7 ou v 5,7,9) (VARZEA et al., 2005; ZAMBOLIM et al., 2005; CAPUCHO et al., 2009; ZAMBOLIM & CAIXETA, 2021; ZAMBOLIM et al., 2024).

O estudo de herança da resistência de *C. arabica* a *H. vastatrix*, realizado com raça XXXIII do patógeno, evidenciou que à resistência a essa raça é condicionada pela ação de três genes. Nesse modelo, podem estar atuando um gene dominante e dois genes recessivos, na proporção 55:9, ou um gene dominante independente e dois genes dominantes complementares, na proporção 57:7. Sugerindo ainda o envolvimento de efeito epistático. Em outros estudos com o HdT, inoculados com diferentes raças de *H. vastatrix*, os autores concluíram que os acessos desse

híbrido apresentam herança da resistência a ferrugem governada por um a três genes dominantes, como encontrado neste trabalho (PEREIRA, 1995; BRITO et al., 2010; ROMERO et al., 2014; PESTANA et al., 2015).

O estudo de herança não permitiu concluir se a resistência do HdT UFV 443-03 à raça XXXIII de *H. vastatrix* é controlada por um locus dominante e dois loci recessivos ou um locus dominante independente e dois loci dominantes complementares. Além disso, o tamanho populacional não permitiu testar se a resistência é condicionada por mais de três loci. Assim, o mapeamento de QTL foi útil para complementar o estudo de herança afim de verificar a segregação fenotípica para resistência/suscetibilidade das plantas de café.

Atualmente, não existe uma ferramenta abrangente para o desenvolvimento de mapas de ligação para estudos de QTL em genomas aloploidos, capaz de abordar de forma completa todas as peculiaridades e desafios inerentes a esses genomas (GARCIA et al., 2013; GERARD et al., 2024). No entanto, neste estudo, a análise de QTL utilizando a metodologia SMA, considerando a ploidia da espécie, apresentou resultados promissores na identificação de regiões gênicas associadas à resistência a *H. vastatrix*, incluindo resultados inéditos de marcas associadas a resistência à raça XXXIII do patógeno e novas marcas para a resistência as raças I e II e patótipo 001. Utilizando a abordagem de SMA foi avaliado se havia diferença fenotípicas entre indivíduos de classes genotípicas contrastantes, sem a utilização de um mapa genético. Assim, quando houve uma associação entre fenótipo e genótipo foi decretado um QTL putativo e posteriormente os QTL verdadeiramente associados foram obtidos a partir de  $-\log_{10}$  (p-valor) superior a 3, equivalente a um *p-valor* de 0,001 (MCCOUCH & DOERGE, 1995). Apesar de fácil implementação, essa metodologia não permite o estudo de epistasia.

No trabalho de Pestana et al. (2015), em estudo genômico da resistência as raças I, II e patótipo 001 de *H. vastatrix*, utilizando a mesma população desta pesquisa, foram identificados marcadores do tipo SSR, AFLP, RAPD e um marcador específico (CARF001) ligado a QTL associados a resistência de *Coffea*. Entretanto, neste trabalho foram identificados marcadores do tipo SNP associados a QTL, permitindo uma maior cobertura do genoma e mais informações do que os marcadores desenvolvidos para genomas diploides (SERANG et al., 2012; LARA et al., 2019; GERARD et al., 2024).

A análise de QTL para os marcadores alinhados com a porção C<sup>e</sup> do genoma de *C. arabica*, indicou que os marcadores associados às diferentes raças do patógeno estão presentes em nove cromossomos (1, 2, 3, 4, 5, 6, 7, 9 e 11). Os marcadores alinhados com a porção C<sup>e</sup> do genoma

de *C. arabica*, estão presentes em oito cromossomos (1, 2, 3, 5, 6, 7, 8 e 11). Marcadores significativos associados à raça XXXIII foram encontrados nos cromossomos 1 e 11, à raça I nos cromossomos 1, 3, 5 e 7, à raça II nos cromossomos 1, 6 e 7 e ao patótipo 001 nos cromossomos 1 e 7, nas análises de ambos subgenomas de *C. arabica* ( $C^c$  e  $C^e$ ). Sabe-se que a espécie *C. arabica* é um alotetraploide resultante da hibridação entre duas espécies relacionadas, *C. eugenioides* e *C. canephora*. Os dois subgenomas, das espécies parentais, apresentam alta colinearidade (SCALABRIN et al., 2020; SALOJARVI et al., 2024), o que explica a identificação de regiões associadas a resistência nos mesmos cromossomos na análise dos dois subgenomas.

A análise de QTL para as raças I e II e patótipo 001 realizada por Pestana et al. (2015), indicou QTL associados a resistência em diferentes grupos de ligação (GL) relacionados ao mapa desenvolvido no trabalho. Os autores concluíram que duas regiões estão envolvidas na resistência do HdT (GL2 e GL10) para as raças I, II e patótipo 001 e outras duas regiões (GL2 e GL3) para a raça I e patótipo 001. Assim, os autores sugerem que mesmos genes podem estar envolvidos na resistência à raça I e ao patótipo 001 ou a região contém genes próximos que conferem resistência a ambos. Neste trabalho, utilizando marcadores SNP, também foram encontrados marcadores associados para as raças I e II e o patótipo 001 em cromossomos semelhantes (cromossomos: 1 e 7). Corroborando com as conclusões de Pestana et al (2015) de que pode haver genes semelhantes ou próximos envolvidos na resistência de diferentes raças do patógeno.

Genes de resistência tendem a estar frequentemente agrupados no genoma, o que pode ser uma consequência de duplicações gênicas e processos de recombinação originados a partir de genes dos progenitores. A recombinação e mutações entre genes estreitamente ligados ou próximos no cromossomo podem gerar novas especificidades no reconhecimento de patógenos, desempenhando um papel crucial na coevolução planta-patógeno (KEEN, 1990; MEYERS et al., 2005; BARRAGÁN & WEIGEL, 2021). Assim, a combinação desses diferentes genes pode ser uma estratégia chave na busca de resistência durável de *Coffea* a *H. vastatrix* (BARÇA et al., 2020; ALMEIDA et al., 2021b; SERA et al., 2022; SILVA et al., 2022a; ALVES et al., 2024).

Pela primeira vez foram identificados QTL associados a resistência à raça XXXIII de *H. vastatrix*. Esses marcadores associados foram detectados nos cromossomos 1 e 11 dos dois subgenomas, sendo que o cromossomo 1 apresentou marcadores associados a todas as raças e patótipo avaliados e no cromossomo 11 também foi detectado marcador associado a raça I e ao

patótipo 001. Evidenciando mais uma vez a hipótese de haver genes semelhantes ou próximos envolvidos na resistência de diferentes raças desse patógeno (PESTANA et al., 2015). Além disso, este trabalho identificou cromossomos específicos com regiões associadas exclusivamente à resistência a determinadas raças, como a raça XXXIII, a raça I, ou ao patótipo 001. Esse resultado permite o desenvolvimento de marcadores moleculares exclusivos para algumas raças de *H. vastatrix* o que pode auxiliar no estudo da evolução desse patógeno.

Assim, este estudo identificou marcadores do tipo SNP associados a QTL relacionados à resistência às raças I, II, XXXIII e ao patótipo 001 de *H. vastatrix*, considerando a ploidia da espécie e proporcionando maior cobertura genômica com os marcadores SNP, além de ampliar a profundidade das informações geradas em comparação com estudos anteriores. A utilização dessas regiões genômicas associadas a genes de resistência, por meio da seleção assistida por marcadores, permitirá a triagem eficiente de genótipos portadores de genes de resistência, os quais poderão ser utilizados para a introgressão dessas características em outras variedades. Dessa forma, os resultados obtidos contribuem significativamente para o entendimento da herança genética de uma importante fonte de resistência a *H. vastatrix* e fortalecem as estratégias de melhoramento do cafeeiro que visam ao desenvolvimento de plantas com resistência ou tolerância durável a diferentes raças desse patógeno.

## 5. Referências Bibliográficas

ALKIMIM, E. R., CAIXETA, E. T., SOUSA, T. V., PEREIRA, A. A., DE OLIVEIRA, A. C. B., ZAMBOLIM, L., & SAKIYAMA, N. S. (2017). Marker-assisted selection provides arabica coffee with genes from other *Coffea* species targeting on multiple resistance to rust and coffee berry disease. *Molecular Breeding*, 37, 1-10.

ALMEIDA, D. P., CAIXETA, E. T., MOREIRA, K. F., DE OLIVEIRA, A. C. B., DE FREITAS, K. N. P., PEREIRA, A. A., ... & CRUZ, C. D. (2021a). Marker-assisted pyramiding of multiple disease resistance genes in coffee genotypes (*Coffea arabica*). *Agronomy*, 11(9), 1763.

ALMEIDA, D. P., CASTRO, I. S. L., MENDES, T. A. D. O., ALVES, D. R., BARKA, G. D., BARREIROS, P. R. R. M., ... & CAIXETA, E. T. (2021b). Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. *Scientia Agricola*, 78.

ALVES, D. R., ALMEIDA, D. P., DE ANDRADE SILVA, E. M., CASTRO, I. S. L., BARREIROS, P. R. R. M., DE OLIVEIRA MENDES, T. A., ... & CAIXETA, E. T. (2024). Unravelling the role of key genes involved coffee leaf rust resistance. *Current Plant Biology*, 38, 100347.

ANTHONY, F., COMBES, M. C., ASTORGA, C., BERTRAND, B., GRAZIOSI, G., & LASHERMES, P. (2002). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theoretical and Applied Genetics*, 104, 894-900.

ARORA, S., KAUR, S., DHILLON, G. S., SINGH, R., KAUR, J., SHARMA, A., & CHHUNEJA, P. (2021). Introgression and genetic mapping of leaf rust and stripe rust resistance in *Aegilops triuncialis*. *Journal of Genetics*, 100, 1-11.

BARKA, G. D., CAIXETA, E. T., FERREIRA, S. S., & ZAMBOLIM, L. (2020). In silico guided structural and functional analysis of genes with potential involvement in resistance to coffee leaf rust: A functional marker based approach. *PloS one*, 15(7), e0222747.

BARRAGAN, A. C., & WEIGEL, D. (2021). Plant NLR diversity: the known unknowns of pan-NLRomes. *The Plant Cell*, 33(4), 814-831.

BAWIN, Y., RUTTINK, T., STAELENS, A., HAEGEMAN, A., STOFFELEN, P., MWANGA MWANGA, J. C. I., ... & JANSSENS, S. B. (2021). Phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica*. *Journal of Systematics and Evolution*, 59(5), 953-963.

BETTENCOURT, A. J. (1973). Considerações gerais sobre o Híbrido de Timor. Instituto Agronômico, Sao Paulo (Brasil).

BETTENCOURT, A. J., RODRIGUES JUNIOR, C. J., CLARKE, R. J., & MACRAE, R. (1988). Principles and practice of coffee breeding for resistance to rust and other diseases Coffee, *Agronomy*, 4.

BRITO, G. G., CAIXETA, E. T., GALLINA, A. P., ZAMBOLIM, E. M., ZAMBOLIM, L., DIOLA, V., & LOUREIRO, M. E. (2010). Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. *Euphytica*, 173, 255-264.

CAPUCHO, A. S., CAIXETA, E. T., ZAMBOLIM, E. M., & ZAMBOLIM, L. (2009). Herança da resistência do Híbrido de Timor UFV 443-03 à ferrugem-do-cafeeiro. *Pesquisa Agropecuária Brasileira*, 44, 276-282.

CAPUCHO, A. S., ZAMBOLIM, E. M., FREITAS, R. L., HADDAD, F., CAIXETA, E. T., & ZAMBOLIM, L. (2012). Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. *Australasian Plant Disease Notes*, 7, 189-191.

CARVALHO, A., FAZUOLI, L. C., & COSTA, W. M. D. (1989). Melhoramento do cafeeiro: XLI. Produtividade do Híbrido de Timor, de seus derivados e de outras fontes de resistência a *Hemileia vastatrix*. *Bragantia*, 48, 73-86.

CARVALHO, V. L., CHALFOUN, S. M., & CUNHA, R. D. (2010). Manejo de doenças do cafeeiro. Reis, PR; Cunha, RL Café arábica: do plantio a colheita. Lavras: EPAMIG, 1, 689-756.

CLARINDO, W. R., & CARVALHO, C. R. (2009). Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. *Plant Cell Reports*, 28, 73-81.

COLLARD, B. C., JAHUFER, M. Z. Z., BROUWER, J. B., & PANG, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, 142, 169-196.

CONAB, C. N. D. A. (2024). Acompanhamento da Safra Brasileira - Café. [S. l.: s. n.],

CRUZ, C. D. (2013). GENES: software para análise de dados em estatística experimental e em genética quantitativa. *Acta Scientiarum. Agronomy*, 35, 271-276.

DENOEUDE, F., CARRETERO-PAULET, L., DEREPPER, A., DROC, G., GUYOT, R., PIETRELLA, M., ... & LASHERMES, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*, 345(6201), 1181-1184.

DINIZ, L. E., SAKIYAMA, N. S., LASHERMES, P., CAIXETA, E. T., OLIVEIRA, A. C. B., ZAMBOLIM, E. M., ... & ZAMBOLIM, L. (2005). Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. *Crop Breeding and Applied Biotechnology*, 5(4), 387.

DINIZ, I., TALHINHAS, P., AZINHEIRA, H. G., VÁRZEA, V., MEDEIRA, C., MAIA, I., ... & DO CÉU SILVA, M. (2012). Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and nonhost resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. *European Journal of Plant Pathology*, 133, 141-157.

ESKES, A. B. (2019). Resistance. In *Coffee Rust* (pp. 171-292). CRC Press.

FEITOSA, F. D. M., SANTOS, I. G. D., PEREIRA, A. A., DE OLIVEIRA, A. C. B., & CAIXETA, E. T. (2024). Gene pyramiding for achieving enhanced disease and insect multiple resistance in *Coffea arabica*. *Crop Science*. 64(5), 2736-2747

FERRÃO, L. F. V., CAIXETA, E. T., PENA, G., ZAMBOLIM, E. M., CRUZ, C. D., ZAMBOLIM, L., ... & SAKIYAMA, N. S. (2015). New EST-SSR markers of *Coffea arabica*: transferability and application to studies of molecular characterization and genetic mapping. *Molecular Breeding*, 35, 1-5.

- FLOR, H. H. (1971). Current status of the gene-for-gene concept. *Annual review of phytopathology*, 9(1), 275-296.
- GARCIA, A. A., MOLLINARI, M., MARCONI, T. G., SERANG, O. R., SILVA, R. R., VIEIRA, M. L., ... & SOUZA, A. P. (2013). SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. *Scientific reports*, 3(1), 3399.
- GERARD, D., FERRÃO, L. F. V., GARCIA, A. A. F., & STEPHENS, M. (2018). Genotyping polyploids from messy sequencing data. *Genetics*, 210(3), 789-807.
- GERARD, D., THAKKAR, M., & FERRÃO, L. F. V. (2024). Tests for Segregation Distortion in Tetraploid F1 Populations. *bioRxiv*, 2024-02.
- GIMASE, J. M., THAGANA, W. M., OMONDI, C., CHESEREK, J. J., GICHIMU, B. M., & GICHURU, E. K. (2020). Quantitative trait loci (QTL) mapping of resistance to coffee berry disease (*Colletotrichum kahawae* Waller & Bridge) in *Coffea arabica* L. variety Rume Sudan.
- HARELIMANA, A., RUKAZAMBUGA, D., & HANCE, T. (2022). Pests and diseases regulation in coffee agroecosystems by management systems and resistance in changing climate conditions: A review. *Journal of Plant Diseases and Protection*, 129(5), 1041-1052.
- HERRERA P, J. C., ALVARADO A, G., CORTINA G, H. A., COMBES, M. C., ROMERO G, G., & LASHERMES, P. (2009). Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk & Br.) introgressed into the cultivated *Coffea arabica* L. from the diploid *C. canephora* species. *Euphytica*, 167, 57-67.
- KEEN, N. (1990). Gene-for-gene complementarity in plant-pathogen interactions. *Annual review of genetics*, 24, 447-463.
- LARA, L. A., SANTOS, M. F., JANK, L., CHIARI, L., VILELA, M. D. M., AMADEU, R. R., ... & GARCIA, A. A. F. (2019). Genomic selection with allele dosage in *Panicum maximum* jacq. *G3: Genes, Genomes, Genetics*, 9(8), 2463-2475.

LASHERMES, P., COMBES, M. C., ROBERT, J., TROUSLOT, P., D'HONT, A., ANTHONY, F., & CHARRIER, A. (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics MGG*, 261(2), 259-266.

LASHERMES, P., COMBES, M. C., PRAKASH, N. S., TROUSLOT, P., LORIEUX, M., & CHARRIER, A. (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome*, 44(4), 589-595.

LASHERMES, P., BERTRAND, B., ETIENNE, H. (2009). Breeding coffee (*Coffea arabica*) for sustainable production. *Breeding plantation tree crops: Tropical species*, 525-543.

LASHERMES, P., COMBES, M. C., HUEBER, Y., SEVERAC, D., & DEREPPER, A. (2014). Genome rearrangements derived from homoeologous recombination following allopolyploidy speciation in coffee. *The Plant Journal*, 78(4), 674-685.

MAPA, 2024: Ministério da Agricultura e Pecuária.

MCCOUGH, S. R., & DOERGE, R. W. (1995). QTL mapping in rice. *Trends in Genetics*, 11(12), 482-487.

MEDEIROS, A. C., CAIXETA, E. T., OLIVEIRA, A. C. B. D., SOUSA, T. V., STOCK, V. D. M., CRUZ, C. D., ... & PEREIRA, A. A. (2021). Combining ability and molecular marker approach identified genetic resources to improve agronomic performance in *Coffea arabica* breeding. *Frontiers in Sustainable Food Systems*, 5, 705278.

MEROT-L'ANTHOENE, V., TOURNEBIZE, R., DARRACQ, O., RATTINA, V., LEPELLEY, M., BELLANGER, L., ... & PONCET, V. (2019). Development and evaluation of a genome-wide Coffee 8.5 K SNP array and its application for high-density genetic mapping and for investigating the origin of *Coffea arabica* L. *Plant biotechnology journal*, 17(7), 1418-1430.

MEYERS, B. C., KAUSHIK, S., & NANDETY, R. S. (2005). Evolving disease resistance genes. *Current opinion in plant biology*, 8(2), 129-134.

MONCADA, M. D. P., TOVAR, E., MONTOYA, J. C., GONZÁLEZ, A., SPINDEL, J., & MCCOUCH, S. (2016). A genetic linkage map of coffee (*Coffea arabica* L.) and QTL for yield, plant height, and bean size. *Tree genetics & genomes*, 12, 1-17.

PEARL, H. M., NAGAI, C. H. I. F. U. M. I., MOORE, P. H., STEIGER, D. L., OSGOOD, R. V., & MING, R. (2004). Construction of a genetic map for arabica coffee. *Theoretical and Applied Genetics*, 108, 829-835.

PEREIRA, A. A. (1995). Herança da resistência a *Hemileia vastatrix* Berk. et Br. em cafeeiros derivados do Híbrido de Timor.

PESTANA, K. N., CAPUCHO, A. S., CAIXETA, E. T., DE ALMEIDA, D. P., ZAMBOLIM, E. M., CRUZ, C. D., ... & SAKIYAMA, N. S. (2015). Inheritance study and linkage mapping of resistance loci to *Hemileia vastatrix* in Híbrido de Timor UFV 443-03. *Tree Genetics & Genomes*, 11, 1-13.

PRAKASH, N. S., MARQUES, D. V., VARZEA, V. M. P., SILVA, M. C., COMBES, M. C., & LASHERMES, P. (2004). Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. *Theoretical and Applied Genetics*, 109, 1311-1317.

PRIOLLI, R. H., MAZZAFERA, P., SIQUEIRA, W. J., MÖLLER, M., ZUCCHI, M. I., RAMOS, L. C. S., ... & COLOMBO, C. A. (2008). Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). *Genetics and Molecular Biology*, 31, 498-504.

RESENDE, M., CAIXETA, E.T., & ALKIMIM, E.R. (2016). High-throughput targeted genotyping of *Coffea arabica* and *Coffea canephora* using next generation sequencing. San Diego, CA. 1.

ROMERO G, G., ALVARADO A, G., CORTINA G, H., LIGARRETO M, G., GALEANO, N. F., & HERRERA P, J. C. (2010). Partial resistance to leaf rust (*Hemileia vastatrix*) in coffee (*Coffea arabica* L.): genetic analysis and molecular characterization of putative candidate genes. *Molecular breeding*, 25, 685-697.

ROMERO, G., VÁSQUEZ, L. M., LASHERMES, P., & HERRERA, J. C. (2014). Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). *Plant Breeding*, 133(1), 121-129.

SAAVEDRA, L. M., CAIXETA, E. T., BARKA, G. D., BORÉM, A., ZAMBOLIM, L., NASCIMENTO, M., ... & PEREIRA, A. A. (2023). Marker-assisted recurrent selection for pyramiding leaf rust and coffee berry disease resistance alleles in *Coffea Arabica* L. *Genes*, 14(1), 189.

SAKIYAMA, N. S. (2005). Melhoramento do Café Arábica. In: BORÉM, A. (org.). Melhoramento de espécies cultivadas. 2a ed.ed. Viçosa: Editora UFV – Universidade Federal de Viçosa, p. p.605-658.

SALOJÄRVI, J., RAMBANI, A., YU, Z., GUYOT, R., STRICKLER, S., LEPELLEY, M., ... & DESCOMBES, P. (2024). The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. *Nature genetics*, 56(4), 721-731.

SATTLER, M. C., DE OLIVEIRA, S. C., MENDONÇA, M. A. C., & CLARINDO, W. R. (2022). *Coffea* cytogenetics: from the first karyotypes to the meeting with genomics. *Planta*, 255(6), 112.

SCALABRIN, S., TONIUTTI, L., DI GASPERO, G., SCAGLIONE, D., MAGRIS, G., VIDOTTO, M., ... & BERTRAND, B. (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Scientific reports*, 10(1), 1-13.

SCALABRIN, S., MAGRIS, G., LIVA, M., VITULO, N., VIDOTTO, M., SCAGLIONE, D., ... & MORGANTE, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, 15(1), 463.

SERA, G. H., DE CARVALHO, C. H. S., DE REZENDE ABRAHÃO, J. C., POZZA, E. A., MATIELLO, J. B., DE ALMEIDA, S. R., ... & DOS SANTOS BOTELHO, D. M. (2022).

Coffee leaf rust in Brazil: Historical events, current situation, and control measures. *Agronomy*, 12(2), 496.

SERANG, O., MOLLINARI, M., & GARCIA, A. A. F. (2012). Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. *PLoS One*, 7(2), e30906.

SETOTAW, T. A., CAIXETA, E. T., ZAMBOLIM, E. M., SOUSA, T. V., PEREIRA, A. A., DE OLIVEIRA, A. C. B., ... & SAKYAMA, N. S. (2020). Genome introgression of Híbrido de Timor and its potential to develop high cup quality *C. arabica* cultivars.

SILVA, M. D. C., VÁRZEA, V., GUERRA-GUIMARÃES, L., AZINHEIRA, H. G., FERNANDEZ, D., PETITOT, A. S., ... & NICOLE, M. (2006). Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian journal of plant physiology*, 18, 119-147.

SILVA, M. D. C., GUERRA-GUIMARÃES, L., DINIZ, I., LOUREIRO, A., AZINHEIRA, H., PEREIRA, A. P., ... & VÁRZEA, V. (2022a). An overview of the mechanisms involved in *Coffea-Hemileia vastatrix* interactions: Plant and pathogen perspectives. *Agronomy*, 12(2), 326.

SILVA, F.L., ALKIMIM, E. R., BARREIRO, P. R. R. M., LEICHTWEIS, B. G., SILVA, A. C. A., DA SILVA, R. A., ... & CAIXETA, E. T. (2022b). Genome-wide association study of plant architecture and diseases resistance in *Coffea canephora*. *Euphytica*, 218(7), 92.

SILVA, R. A.; ZAMBOLIM, L.; CASTRO, I. S. L.; RODRIGUES, H. S.; CRUZ, C. D.; CAIXETA, E. T. (2018) The Híbrido de Timor germplasm: identification of molecular diversity and resistance sources to coffee berry disease and leaf rust. *Euphytica*, v. 214, n. 9, p. 153.

SOUSA, T. V; CAIXETA, E. T.; ALKIMIM, E. R.; DE OLIVEIRA, A. C. B.; PEREIRA, A. A.; SAKIYAMA, N. S. JÚNIOR, M. F. R. R.; ZAMBOLIM, L. (2017). Population structure and genetic diversity of coffee progenies derived from Catuaí and Híbrido de Timor revealed by genome-wide SNP marker. *Tree genetics & genomes*, v. 13, n. 6, p. 124.

TAMAYO, P. J. VALE, FXR do; ZAMBOLIM, L.; CHAVES, GM; PEREIRA, AA (1995). Resistência do Catimor à ferrugem e virulência de raças fisiológicas de *Hemileia vastatrix* Berk. & Br. *Fitopatologia Brasileira*, 20, 572-576.

TEIXEIRA-CABRAL, T. A., SAKIYAMA, N. S., ZAMBOLIM, L., PEREIRA, A. A., & SCHUSTER, I. (2004). Single-locus inheritance and partial linkage map of *Coffea arabica* L. *Crop Breeding and Applied Biotechnology*, 4(4).

VÁRZEA, V. M. P.; MARQUES, D. V. (2005). Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: ZAMBOLIM, L.; ZAMBOLIM, E. M.; VÁRZEA, V. M. P. (org.). Durable resistance to coffee leaf rust. Viçosa: Editora UFV - Universidade Federal de Viçosa. p. 53–74.

VIEIRA, L. G. E., ANDRADE, A. C., COLOMBO, C. A., MORAES, A. H. D. A., METHA, Â., OLIVEIRA, A. C. D., ... & PEREIRA, G. A. G. (2006). Brazilian coffee genome project: an EST-based genomic resource. *Brazilian Journal of Plant Physiology*, 18, 95-108.

XIONG, H., LI, Y., GUO, H., XIE, Y., ZHAO, L., GU, J., ... & LIU, L. (2021). Genetic mapping by integration of 55K SNP array and KASP markers reveals candidate genes for important agronomic traits in hexaploid wheat. *Frontiers in Plant Science*, 12, 628478.

ZAMBOLIM, L., ZAMBOLIM, E. M., VALE, F. D., PEREIRA, A. A., SAKIYAMA, N. S., & CAIXETA, E. T. (2005). Physiological races of *Hemileia vastatrix* Berk. et Br. *Brazil-Physiological variability, current situation and future prospect*. In Zambolim L, Zambolim EM and Várzea VMP (eds) *Durable resistance to coffee leaf rust*. UFV, Viçosa, 75-98.

ZAMBOLIM, L. (2016). Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology*, 41, 1-8.

ZAMBOLIM, L., & CAIXETA, E. T. (2021). An overview of physiological specialization of coffee leaf rust-new designation of pathotypes. *Int. J. Curr. Res*, 13, 15564-15575.

ZAMBOLIM, L., CAIXETA, E. T., GUERREIRO FILHO, O., SERA, G. H., SERA, T., PEREIRA, A. A., & DE CARVALHO, C. H. (2024). Breeding Programs Against Coffee Leaf Rust in Brazil: A Review. *Journal of Agricultural Science*, 16(5), 61-74.

ZHAO, N., XUE, D., MIAO, Y., WANG, Y., ZHOU, E., ZHOU, Y., ... & WANG, X. (2023). Construction of a high-density genetic map for faba bean (*Vicia faba* L.) and quantitative trait loci mapping of seed-related traits. *Frontiers in Plant Science*, 14, 1201103.

WALLER, J. M., BIGGER, M., & HILLOCKS, R. J. (Eds.). (2007). Coffee pests, diseases and their management. CABI.

WANG, R., XING, S., BOURKE, P. M., QI, X., LIN, M., ESSELINK, D., ... & FANG, J. (2023). Development of a 135K SNP genotyping array for *Actinidia arguta* and its applications for genetic mapping and QTL analysis in kiwifruit. *Plant Biotechnology Journal*, 21(2), 369-380.

## Material Suplementar

Material suplementar 1. Informações dos indivíduos genotipados. ID: código dos genótipos - *Rapid Genomics*; ID POP: código dos genótipos da população; XXXIII, I, II, P 001: nota da avaliação fenotípica dos genótipos para as raças XXXIII, I, II e patótipo 001, respectivamente.

ID	ID POP	XXXIII	I	II	P 001	ID	ID POP	XXXIII	I	II	P 001
EMB_144101_P002_WG07	H511/1-1	1	2	3	2	EMB_144101_P004_WA05	H511/1-131	1	3	1	2
EMB_144101_P002_WG08	H511/1-3	5	3	2	1	EMB_144101_P004_WA06	H511/1-132	1	3	3	2
EMB_144101_P002_WG09	H511/1-4	1	2	3	1	EMB_144101_P004_WA07	H511/1-133	1	3	3	1
EMB_144101_P002_WG10	H511/1-5	1	3	2	1	EMB_144101_P004_WA08	H511/1-134	6	5	2	2
EMB_144101_P002_WG11	H511/1-6	1	2	1	2	EMB_144101_P004_WA09	H511/1-135	1	3	1	1
EMB_144101_P002_WG12	H511/1-7	2	3	1	2	EMB_144101_P004_WA10	H511/1-136	5	2	3	2
EMB_144101_P002_WH01	H511/1-8	2	3	2	1	EMB_144101_P004_WA11	H511/1-137	2	3	3	2
EMB_144101_P002_WH02	H511/1-9	1	2	3	1	EMB_144101_P004_WA12	H511/1-138	3	3	3	2
EMB_144101_P002_WH03	H511/1-10	1	2	1	1	EMB_144101_P004_WB01	H511/1-139	2	2	2	2
EMB_144101_P002_WH04	H511/1-11	2	3	3	2	EMB_144101_P004_WB02	H511/1-140	2	4	3	3
EMB_144101_P002_WH05	H511/1-12	1	3	2	4	EMB_144101_P004_WB03	H511/1-141	1	3	2	1
EMB_144101_P002_WH06	H511/1-13	4	4	3	2	EMB_144101_P004_WB04	H511/1-142	1	3	3	2
EMB_144101_P002_WH07	H511/1-14	1	3	2	2	EMB_144101_P004_WB05	H511/1-143	1	3	2	3
EMB_144101_P002_WH08	H511/1-15	1	2	1	1	EMB_144101_P004_WB06	H511/1-144	2	3	1	2
EMB_144101_P002_WH09	H511/1-16	1	3	2	2	EMB_144101_P004_WB07	H511/1-145	1	3	3	2
EMB_144101_P002_WH10	H511/1-17	1	2	2	2	EMB_144101_P004_WB08	H511/1-146	1	3	3	3
EMB_144101_P002_WH11	H511/1-18	1	3	3	3	EMB_144101_P004_WB09	H511/1-147	1	3	2	3
EMB_144101_P002_WH12	H511/1-19	2	2	3	2	EMB_144101_P004_WB10	H511/1-148	1	3	3	1
EMB_144101_P003_WA01	H511/1-20	1	2	2	2	EMB_144101_P004_WB11	H511/1-149	1	2	3	3
EMB_144101_P003_WA02	H511/1-21	1	3	3	2	EMB_144101_P004_WB12	H511/1-150	6	3	1	2
EMB_144101_P003_WA03	H511/1-22	4	2	2	5	EMB_144101_P004_WC01	H511/1-151	2	3	3	3
EMB_144101_P003_WA04	H511/1-23	2	2	2	1	EMB_144101_P004_WC02	H511/1-152	1	3	3	2
EMB_144101_P003_WA05	H511/1-24	1	3	2	3	EMB_144101_P004_WC03	H511/1-153	3	3	3	2
EMB_144101_P003_WA06	H511/1-25	NA	3	1	1	EMB_144101_P004_WC04	H511/1-154	6	3	4	4
EMB_144101_P003_WA07	H511/1-26	2	2	2	2	EMB_144101_P004_WC05	H511/1-155	1	3	3	2
EMB_144101_P003_WA08	H511/1-27	6	3	3	3	EMB_144101_P004_WC06	H511/1-156	2	3	3	1
EMB_144101_P003_WA09	H511/1-28	2	3	2	2	EMB_144101_P004_WC07	H511/1-157	2	4	3	2
EMB_144101_P003_WA10	H511/1-29	1	2	3	2	EMB_144101_P004_WC08	H511/1-158	1	3	2	2
EMB_144101_P003_WA11	H511/1-30	1	2	3	2	EMB_144101_P004_WC09	H511/1-159	1	4	6	6
EMB_144101_P003_WA12	H511/1-31	1	3	3	3	EMB_144101_P004_WC10	H511/1-160	1	3	2	1
EMB_144101_P003_WB01	H511/1-32	4	3	3	2	EMB_144101_P004_WC11	H511/1-161	1	2	2	1
EMB_144101_P003_WB02	H511/1-33	3	3	2	2	EMB_144101_P004_WC12	H511/1-162	6	3	2	3
EMB_144101_P003_WB03	H511/1-35	1	2	3	2	EMB_144101_P004_WD01	H511/1-163	2	3	3	2
EMB_144101_P003_WB04	H511/1-36	4	2	1	3	EMB_144101_P004_WD02	H511/1-164	1	2	3	2
EMB_144101_P003_WB05	H511/1-38	1	9	3	2	EMB_144101_P004_WD03	H511/1-165	1	4	3	3
EMB_144101_P003_WB06	H511/1-39	6	6	6	6	EMB_144101_P004_WD04	H511/1-166	1	2	3	1
EMB_144101_P003_WB07	H511/1-40	1	3	3	2	EMB_144101_P004_WD05	H511/1-167	3	2	2	2
EMB_144101_P003_WB08	H511/1-41	2	2	3	2	EMB_144101_P004_WD06	H511/1-169	2	2	3	2
EMB_144101_P003_WB09	H511/1-42	1	2	3	2	EMB_144101_P004_WD07	H511/1-170	2	3	2	2
EMB_144101_P003_WB10	H511/1-43	2	3	1	2	EMB_144101_P004_WD08	H511/1-171	1	3	2	1
EMB_144101_P003_WB11	H511/1-44	1	3	3	1	EMB_144101_P004_WD09	H511/1-172	2	3	1	2
EMB_144101_P003_WB12	H511/1-45	1	2	3	3	EMB_144101_P004_WD10	H511/1-173	2	3	1	2
EMB_144101_P003_WC01	H511/1-46	3	2	2	1	EMB_144101_P004_WD11	H511/1-175	2	3	2	3
EMB_144101_P003_WC02	H511/1-47	2	2	3	2	EMB_144101_P004_WD12	H511/1-176	2	3	2	2
EMB_144101_P003_WC03	H511/1-48	2	2	3	3	EMB_144101_P004_WE01	H511/1-177	1	3	2	2
EMB_144101_P003_WC04	H511/1-49	1	2	3	2	EMB_144101_P004_WE02	H511/1-178	1	2	2	2
EMB_144101_P003_WC05	H511/1-50	3	3	3	2	EMB_144101_P004_WE03	H511/1-179	1	3	2	2
EMB_144101_P003_WC06	H511/1-51	5	3	3	4	EMB_144101_P004_WE04	H511/1-180	2	2	3	1
EMB_144101_P003_WC07	H511/1-52	4	3	2	2	EMB_144101_P004_WE05	H511/1-181	NA	2	2	2
EMB_144101_P003_WC08	H511/1-54	3	2	2	1	EMB_144101_P004_WE06	H511/1-185	1	3	3	3
EMB_144101_P003_WC09	H511/1-55	2	3	3	3	EMB_144101_P004_WE07	H511/1-186	1	2	2	2
EMB_144101_P003_WC10	H511/1-56	4	3	6	6	EMB_144101_P004_WE08	H511/1-187	1	2	3	2
EMB_144101_P003_WC11	H511/1-57	2	2	2	2	EMB_144101_P004_WE09	H511/1-188	3	3	3	2

EMB_144101_P003_WC12	H511/1-58	2	3	3	2	EMB_144101_P004_WE10	H511/1-189	1	3	3	2
EMB_144101_P003_WD01	H511/1-59	1	3	3	2	EMB_144101_P004_WE11	H511/1-190	3	3	3	2
EMB_144101_P003_WD02	H511/1-64	2	3	3	2	EMB_144101_P004_WE12	H511/1-191	3	2	3	2
EMB_144101_P003_WD03	H511/1-65	2	4	2	2	EMB_144101_P004_WF01	H511/1-192	4	3	3	3
EMB_144101_P003_WD04	H511/1-66	1	4	5	4	EMB_144101_P004_WF02	H511/1-193	1	2	3	2
EMB_144101_P003_WD05	H511/1-67	1	2	2	2	EMB_144101_P004_WF03	H511/1-194	2	3	3	2
EMB_144101_P003_WD06	H511/1-68	1	2	3	2	EMB_144101_P004_WF04	H511/1-195	1	2	2	1
EMB_144101_P003_WD07	H511/1-69	2	2	3	3	EMB_144101_P004_WF05	H511/1-196	1	3	2	3
EMB_144101_P003_WD08	H511/1-70	1	3	2	3	EMB_144101_P004_WF06	H511/1-197	3	3	2	3
EMB_144101_P003_WD09	H511/1-71	2	2	2	2	EMB_144101_P004_WF07	H511/1-198	1	3	3	2
EMB_144101_P003_WD10	H511/1-72	1	2	3	2	EMB_144101_P004_WF08	H511/1-199	4	4	2	4
EMB_144101_P003_WD11	H511/1-73	1	3	3	3	EMB_144101_P004_WF09	H511/1-200	1	3	3	3
EMB_144101_P003_WD12	H511/1-74	1	2	2	2	EMB_144101_P004_WF10	H511/1-201	1	3	3	2
EMB_144101_P003_WE01	H511/1-75	1	3	1	2	EMB_144101_P004_WF11	H511/1-202	2	3	3	2
EMB_144101_P003_WE02	H511/1-76	1	3	2	2	EMB_144101_P004_WF12	H511/1-203	1	3	3	3
EMB_144101_P003_WE03	H511/1-77	1	3	3	2	EMB_144101_P004_WG01	H511/1-204	1	3	3	2
EMB_144101_P003_WE04	H511/1-78	1	2	3	2	EMB_144101_P004_WG02	H511/1-205	1	3	2	3
EMB_144101_P003_WE05	H511/1-81	1	2	2	2	EMB_144101_P004_WG03	H511/1-206	2	3	1	2
EMB_144101_P003_WE06	H511/1-82	1	3	3	2	EMB_144101_P004_WG04	H511/1-207	1	2	1	2
EMB_144101_P003_WE07	H511/1-83	1	2	2	2	EMB_144101_P004_WG05	H511/1-208	1	3	2	2
EMB_144101_P003_WE08	H511/1-84	1	2	3	1	EMB_144101_P004_WG06	H511/1-209	1	3	2	2
EMB_144101_P003_WE09	H511/1-85	1	3	3	2	EMB_144101_P004_WG07	H511/1-210	2	4	3	2
EMB_144101_P003_WE10	H511/1-86	1	3	2	3	EMB_144101_P004_WG08	H511/1-211	4	5	6	6
EMB_144101_P003_WE11	H511/1-87	2	2	2	2	EMB_144101_P004_WG09	H511/1-212	6	3	3	2
EMB_144101_P003_WE12	H511/1-88	3	3	3	1	EMB_144101_P004_WG10	H511/1-213	2	2	3	2
EMB_144101_P003_WF01	H511/1-89	3	3	1	2	EMB_144101_P004_WG11	H511/1-214	1	2	3	3
EMB_144101_P003_WF02	H511/1-90	2	2	3	2	EMB_144101_P004_WG12	H511/1-215	2	3	3	2
EMB_144101_P003_WF03	H511/1-91	1	5	3	5	EMB_144101_P004_WH01	H511/1-216	2	3	3	2
EMB_144101_P003_WF04	H511/1-92	1	3	3	2	EMB_144101_P004_WH02	H511/1-217	4	3	3	3
EMB_144101_P003_WF05	H511/1-93	2	3	3	3	EMB_144101_P004_WH03	H511/1-218	1	3	3	2
EMB_144101_P003_WF06	H511/1-95	1	2	1	1	EMB_144101_P004_WH04	H511/1-219	2	3	2	2
EMB_144101_P003_WF07	H511/1-96	1	3	1	1	EMB_144101_P004_WH05	H511/1-220	1	3	3	2
EMB_144101_P003_WF08	H511/1-97	2	4	2	3	EMB_144101_P004_WH06	H511/1-221	1	3	2	2
EMB_144101_P003_WF09	H511/1-98	5	5	4	5	EMB_144101_P004_WH07	H511/1-222	1	3	2	2
EMB_144101_P003_WF10	H511/1-99	2	3	3	2	EMB_144101_P004_WH08	H511/1-223	1	3	3	3
EMB_144101_P003_WF11	H511/1-100	3	3	3	2	EMB_144101_P004_WH09	H511/1-224	4	3	3	3
EMB_144101_P003_WF12	H511/1-101	3	3	3	2	EMB_144101_P004_WH10	H511/1-225	1	3	3	2
EMB_144101_P003_WG01	H511/1-102	6	3	2	3	EMB_144101_P004_WH11	H511/1-226	1	3	3	2
EMB_144101_P003_WG02	H511/1-103	1	2	2	2	EMB_144101_P004_WH12	H511/1-227	1	3	3	2
EMB_144101_P003_WG03	H511/1-104	2	3	2	3	EMB_144101_P005_WA01	H511/1-228	3	3	1	2
EMB_144101_P003_WG04	H511/1-105	4	4	4	2	EMB_144101_P005_WA02	H511/1-229	4	2	2	2
EMB_144101_P003_WG05	H511/1-106	1	2	2	2	EMB_144101_P005_WA03	H511/1-230	1	3	3	2
EMB_144101_P003_WG06	H511/1-107	2	3	3	2	EMB_144101_P005_WA04	H511/1-231	3	3	3	5
EMB_144101_P003_WG07	H511/1-108	2	3	2	2	EMB_144101_P005_WA05	H511/1-232	3	3	2	1
EMB_144101_P003_WG08	H511/1-109	1	2	2	2	EMB_144101_P005_WA06	H511/1-233	1	3	2	2
EMB_144101_P003_WG09	H511/1-110	1	3	3	2	EMB_144101_P005_WA07	H511/1-234	2	3	1	2
EMB_144101_P003_WG10	H511/1-111	1	2	3	1	EMB_144101_P005_WA08	H511/1-235	2	2	2	2
EMB_144101_P003_WG11	H511/1-112	2	2	2	1	EMB_144101_P005_WA09	H511/1-236	1	3	1	2
EMB_144101_P003_WG12	H511/1-113	2	2	2	1	EMB_144101_P005_WA10	H511/1-237	1	3	3	2
EMB_144101_P003_WH01	H511/1-114	2	3	3	2	EMB_144101_P005_WA11	H511/1-238	2	3	3	1
EMB_144101_P003_WH02	H511/1-115	1	2	3	3	EMB_144101_P005_WA12	H511/1-239	2	3	3	2
EMB_144101_P003_WH03	H511/1-116	4	2	3	2	EMB_144101_P005_WB01	H511/1-240	3	2	3	1
EMB_144101_P003_WH04	H511/1-117	1	3	3	2	EMB_144101_P005_WB02	H511/1-241	2	3	3	2
EMB_144101_P003_WH05	H511/1-118	6	3	3	2	EMB_144101_P005_WB03	H511/1-242	1	2	2	2
EMB_144101_P003_WH06	H511/1-120	6	6	4	6	EMB_144101_P005_WB04	H511/1-243	1	3	3	2
EMB_144101_P003_WH07	H511/1-121	4	3	1	2	EMB_144102_P002_WG04	H511/1-244	2	2	2	2
EMB_144101_P003_WH08	H511/1-122	1	2	1	1	EMB_144102_P002_WG05	H511/1-245	1	2	2	2
EMB_144101_P003_WH09	H511/1-123	1	3	2	2	EMB_144102_P002_WG06	H511/1-246	2	2	3	2
EMB_144101_P003_WH10	H511/1-124	1	3	2	3	EMB_144102_P002_WG07	H511/1-247	1	3	1	2
EMB_144101_P003_WH11	H511/1-125	1	3	3	1	EMB_144102_P002_WG08	H511/1-248	1	3	2	3
EMB_144101_P003_WH12	H511/1-126	2	3	2	1	EMB_144102_P002_WG09	H511/1-249	1	2	3	1
EMB_144101_P004_WA01	H511/1-127	4	3	2	3	EMB_144102_P002_WG10	2148-57	6	5	6	5

EMB_144101_P004_WA02	H511/1-128	2	3	2	3	EMB_144102_P002_WG11	HT443-3	2	2	2	2
EMB_144101_P004_WA03	H511/1-129	6	4	2	3	EMB_144102_P002_WG12	H511/1 (F1)	NA	2	2	2
EMB_144101_P004_WA04	H511/1-130	1	3	3	2						

---