

UNIVERSIDADE FEDERAL DE VIÇOSA

Desenvolvimento e aplicação de marcadores moleculares de polimorfismos de nucleotídeo único em *Coffea*

Bruna Lopes Mariz
Doctor Scientiae

**VIÇOSA - MINAS GERAIS
2025**

BRUNA LOPES MARIZ

Desenvolvimento e aplicação de marcadores moleculares de polimorfismos de nucleotídeo único em *Coffea*

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: Dênia Pires de Almeida
Iara G. dos Santos
Marcos D. V. de Resende

**VIÇOSA - MINAS GERAIS
2025**

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

T

M343d
2025
Mariz, Bruna Lopes, 1994-
Desenvolvimento e aplicação de marcadores moleculares de
polimorfismos de nucleotídeo único em *Coffea* / Bruna Lopes
Mariz. – Viçosa, MG, 2025.
1 tese eletrônica (92 f.): il. (algumas color.).

Orientador: Eveline Teixeira Caixeta.
Tese (doutorado) - Universidade Federal de Viçosa,
Biologia Geral e Agronomia, 2025.
Inclui bibliografia.
DOI: <https://doi.org/10.47328/ufvbbt.2026.074>
Modo de acesso: World Wide Web.

1. Café - Melhoramento genético. 2. Marcadores genéticos.
I. Caixeta, Eveline Teixeira, 1972-. II. Universidade Federal de
Viçosa. Biologia Geral e Agronomia. Programa de
Pós-Graduação em Genética e Melhoramento. III. Título.

CDD 22. ed. 633.732

BRUNA LOPES MARIZ

Desenvolvimento e aplicação de marcadores moleculares de polimorfismos de nucleotídeo único em *Coffea*

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: 15 de dezembro de 2025.

Assentimento:

Bruna Lopes Mariz
Autora

Eveline Teixeira Caixeta Moura
Orientadora

Essa tese foi assinada digitalmente pela autora em 18/03/2026 às 09:27:46 e pela orientadora em 19/03/2026 às 09:56:26. 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 **6Z4B.I4D3.N1FR** e clique no botão 'Validar documento'.

AGRADECIMENTOS

A Deus por todas as oportunidades e inspirações.

Aos meus pais, Helenice e Hamilton, à minha vó Nadir, ao meu irmão Breno e ao meu querido Bruno, pela humildade dos muitos ensinamentos, pelas inúmeras orações, pela confiança e o imensurável amor.

À Eveline Caixeta pelo incentivo ao aprendizado constante, aos desafios propostos e infinita generosidade.

Aos pesquisadores Antônio Carlos Baião de Oliveira, Carolina Gonçalves Santos, Danúbia Rodrigues Alves, Dênia Pires de Almeida, Iara Gonçalves dos Santos e Marcos Deon Vilela de Resende por acompanharem esse trabalho com entusiasmo, dedicação e muito carinho.

A toda equipe Biocafé, gratidão pelos trabalhos realizados e por todos os bons momentos.

À Universidade Federal de Viçosa e ao Programa de Genética e Melhoramento pelo ensino de excelência, aos grandes profissionais atuantes na educação e formação profissional.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

"Nada na vida deve ser temido, somente compreendido. Agora é hora de compreender mais, para temer menos."

Marie Curie

RESUMO

MARIZ, Bruna Lopes, D.Sc., Universidade Federal de Viçosa, dezembro de 2025. **Desenvolvimento e aplicação de marcadores moleculares de polimorfismos de nucleotídeo único em *Coffea***. Orientadora: Eveline Teixeira Caixeta Moura. Coorientadores: Dênia Pires de Almeida, Iara Gonçalves dos Santos e Marcos Deon Vilela de Resende.

Os polimorfismos de nucleotídeo único (SNPs) se consolidaram como ferramentas genômicas fundamentais para o melhoramento genético vegetal moderno, destacando-se por sua elevada sensibilidade analítica, eficiência, custo-benefício e versatilidade de aplicação em diferentes contextos laboratoriais. A integração estratégica desses marcadores moleculares com métodos convencionais de melhoramento representa um paradigma inovador para acelerar o desenvolvimento de cultivares de *Coffea* spp. com características agrônômicas superiores. Este estudo teve como objetivos: (i) conduzir uma análise bibliométrica abrangente sobre a aplicação de SNPs no melhoramento genético do cafeeiro e (ii) desenvolver e validar marcadores SNPs associados ao controle da altura de plantas em *C. arabica*. Os resultados das análises bibliométricas reportaram aplicações promissoras para o melhoramento assistido por marcadores SNP na redução do tempo de desenvolvimento de novas cultivares e no aumento da acurácia seletiva. Embora as abordagens de genotipagem de alto rendimento (high throughput) tenham gerado um extenso catálogo de SNPs em *Coffea* spp., persiste um desafio crítico: a transição entre a descoberta e a validação funcional desses marcadores para aplicação prática, particularmente em metodologias de baixo rendimento (low throughput). Para superar essa lacuna, foram desenvolvidos e validados quatro marcadores SNPs pela genotipagem High-Resolution Melting (HRM) e dois marcadores por Tetra-Primer Amplification Refractory Mutation System (Tetra-Primer ARMS), para estudos de altura de plantas em *C. arabica*. Estes marcadores ampliaram o conhecimento nos aspectos da complexidade genética da altura da planta e abrem perspectivas para a seleção de plantas com porte previsível. Essa previsibilidade permite otimizar o espaçamento de plantio, além de facilitar a colheita mecânica e tornar o manejo cultural mais eficiente. A abordagem integrada entre genômica e melhoramento convencional representa um avanço significativo no desenvolvimento de cultivares de café, alinhando pesquisa científica com demandas práticas da cadeia produtiva.

Palavras-chave: marcadores SNPs; tecnologias de genotipagem; melhoramento assistido por marcadores; seleção genotípica

ABSTRACT

MARIZ, Bruna Lopes, D.Sc., Universidade Federal de Viçosa, December, 2025. **Development and application of molecular markers of single nucleotide polymorphisms in *Coffea***. Adviser: Eveline Teixeira Caixeta Moura. Co-advisers: Dênia Pires de Almeida, Iara Gonçalves dos Santos and Marcos Deon Vilela de Resende.

Single nucleotide polymorphisms (SNPs) have become established as fundamental genomic tools in modern plant breeding, owing to their high analytical sensitivity, efficiency, cost-effectiveness, and versatility across diverse laboratory applications. The strategic integration of these molecular markers with conventional breeding methods represents an innovative paradigm for accelerating the development of *Coffea* spp. cultivars with superior agronomic traits. This study aimed to: (i) conduct a comprehensive bibliometric analysis of the application of SNPs in coffee breeding and (ii) develop and validate SNP markers associated with the control of plant height in *C. arabica*. The bibliometric analyses revealed promising applications of SNP-based marker-assisted selection, particularly in reducing the time required for the development of new cultivars and increasing selection accuracy. Although high-throughput genotyping approaches have generated an extensive catalog of SNPs in *Coffea* spp., a critical challenge remains: bridging the gap between marker discovery and functional validation for practical application, especially in low-throughput methodologies. To address this limitation, four SNP markers were developed and validated using High-Resolution Melting (HRM) genotyping, and two additional markers were validated using the Tetra-Primer Amplification Refractory Mutation System (Tetra-Primer ARMS) for plant height studies in *C. arabica*. These markers enhance the understanding of the genetic complexity underlying plant height and open new perspectives for the selection of plants with predictable growth habits. Such predictability enables optimized planting density, facilitates mechanical harvesting, and improves overall crop management efficiency. The integrated approach combining genomics and conventional breeding represents a significant advance in coffee cultivar development, aligning scientific research with the practical demands of the production chain.

Keywords: SNP markers; genotyping technologies; marker-assisted selection; genotypic selection

SUMÁRIO

Introdução geral.....	9
Referências	12

CAPÍTULO 1: ANÁLISE BIBLIOMÉTRICA DAS APLICAÇÕES DE MARCADORES SNP EM MELHORAMENTO GENÉTICO DO CAFEIEIRO

1	Introdução.....	18
1.1	Revisão bibliográfica.....	20
1.1.1	Como os SNPs se tornaram os marcadores mais versáteis para estudos genômicos em plantas cultivadas nos últimos anos?.....	20
1.1.2	Identificação de SNPs em genotipagens de <i>high throughput</i> e <i>low throughput</i> ...	21
1.1.3	GWS e GWAS como ferramentas genômicas para identificação de SNPs aplicáveis ao melhoramento genético vegetal.....	26
1.2	Material e métodos	27
1.3	Resultados.....	28
1.4	Discussão	35
1.5	Conclusão	39
1.6	Referências	40

CAPÍTULO 2: DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES MOLECULARES SNP PARA ALTURA DE PLANTAS EM *Coffea arabica*

2	Introdução.....	55
2.1	Material e Métodos.....	58
2.1.1	Material vegetal	58
2.1.2	Extração de DNA.....	60
2.1.3	Seleção de SNPs associados à altura de planta	60
2.1.4	Desenvolvimento de marcadores SNP <i>Low-throughput</i>	62

2.1.5	Modelagem Longitudinal da Altura das Plantas e Análise de Associação Genética	64
2.2	Resultados.....	66
2.2.1	Desenvolvimento dos marcadores moleculares SNP para altura de cafeeiros	66
2.2.2	Genotipagem da população de testagem utilizando marcadores HRM e Tetra <i>Primer</i> ARMS desenvolvidos para altura de plantas	71
2.2.3	Análise da segregação dos marcadores desenvolvidos.....	76
2.2.4	Modelagem longitudinal da altura de plantas da população de testagem.....	77
2.2.5	Associação fenotípica e genotípica dos marcadores moleculares validados	78
2.2.6	Análise descritiva das médias fenotípicas em relação aos marcadores moleculares validados para a altura de plantas.....	80
2.2.7	Identificação de genes e anotação funcional da região gênica entre V16099 e V1190	81
2.4	Discussão	83
2.5	Conclusão	86
2.6	Referências	87

Introdução geral

O café é uma das bebidas mais consumidas no mundo, com estimativas de consumo anual de aproximadamente 10 milhões de toneladas de grãos, o que equivale a cerca de 500 bilhões de xícaras por ano. O Brasil destaca-se como o maior produtor e exportador mundial, com previsões de receita bruta de cerca de 6 bilhões de dólares e é responsável por cerca de 37% do total global de café exportado, seguido por países como Vietnã e Colômbia (Cecafé, 2025; Conab, 2025; FAO, 2025; ICO, 2025b). Este mercado de café é marcado por uma forte dinâmica de importação, com os Estados Unidos, a União Europeia e o Japão representando os maiores consumidores, responsáveis por 60% do total de importações. O aumento da demanda por cafés especiais e sustentáveis, aliado à expansão do consumo em mercados emergentes, como Ásia e África, tem impulsionado o crescimento contínuo das exportações (ICO, 2023; Amrouk et al., 2025).

O gênero *Coffea* compreende 141 espécies descritas (Guyot et al., 2020). Do ponto de vista comercial, *C. arabica* (arábica) e *C. canephora* (robusta e conilon) dominam o mercado global, respondendo por 64% e 36% da produção mundial de café, respectivamente (ICO, 2025a). A espécie *C. liberica*, variedade Liberica (café liberica) e variedade Dewevrei (café excelsa), embora de importância econômica marginal, apresentam relevância regional em mercados específicos do Sudeste Asiático (Davis et al., 2022; Melia et al., 2025).

Enquanto a maioria das espécies do gênero são diploides, autoincompatíveis e de fecundação cruzada, *C. arabica* se distingue por ser alotetraploide e apresentar uma taxa de autofecundação superior a 90% (Clarindo & Carvalho, 2008; Lashermes et al., 2016; Salojärvi et al., 2024; Sattler et al., 2022; Scalabrin et al., 2020). Essa característica, aliada à sua base genética estreita, uma vez que as cultivares disponíveis têm origem nas formas botânicas Typica e Bourbon, contribui para que o melhoramento genético dessa espécie seja um processo desafiador e demorado, com até 30 anos para lançamento de cultivares (Anthony et al., 2002; Bawin et al., 2021; Combes et al., 2000). A falta de diversidade genética nas linhagens de melhoramento de café arábica tem sido reconhecida como uma limitação significativa da tolerância varietal a estresses bióticos e abióticos (Huang et al., 2020; Millet, Delahaie, et al., 2024; Montagnon et al., 2021; Scalabrin et al., 2024).

Devido à sua importância econômica global, o melhoramento genético de cafés tem sido amplamente realizado para atender às demandas do mercado produtor e consumidor. A introgressão de genes de interesse agrônômico é realizada por meio de cruzamentos inter e intraespecíficos, produzindo progênies segregantes que, submetidas a ciclos de seleção

genealógica e recorrente, dão origem a cultivares elite (Barbosa et al., 2020; Feitosa et al., 2024; Mishra, 2020; Saavedra et al., 2023; Prakash et al., 2016). O café arábica teve um incremento de produtividade média cerca de 50% nas últimas duas décadas principalmente devido a melhoramento genético e em manejo, representando 80% das exportações globais de café, acompanhado pela crescente demanda de café canéfora (Carvalho, 1981; FAO, 2025). Contudo, a necessidade por novas cultivares é um processo contínuo e cada vez mais desafiador, dada a crescente exigência por maiores níveis de produtividade, resistência a doenças e adaptação a condições ambientais adversas (Grand View Research, 2024; ICO, 2025b; Malta et al., 2021).

Os programas de melhoramento genético buscam cultivares com: (i) caracteres morfoagronômicos que contribuam para alta produtividade e vigor vegetativo; (ii) resistência vertical e horizontal à agentes bióticos, que comprometam o desenvolvimento da planta, como fungos, bactérias, nematoides e insetos predadores; (iii) uniformidade de maturação; e (iv) qualidade superior da bebida (Ariyoshi, Freitas, et al., 2022; Barbosa et al., 2020; Feitosa et al., 2024; Koutouleas et al., 2023; Mishra, 2020; Santos et al., 2022). Complementarmente, atributos como adaptação à colheita mecanizada, adensamento, estabilidade fenotípica em diferentes regiões e redução de custos de implantação são determinantes para a adoção de novas cultivares (Van der Vossen et al., 2015).

A introgressão simultânea dessas características – que visam produtividade, resiliência a estresses e qualidade – demanda aplicação de métodos de melhoramento como seleção genealógica, seleção recorrente e retrocruzamentos, cuja eficiência está condicionada à variabilidade genética disponível e aos objetivos do programa (Ariyoshi, Freitas, et al., 2022; Malta et al., 2021; Pereira et al., 2021). Em uma estratégia integrada e multitemporal, prioriza-se, a longo prazo, a introgressão assistida de alelos oriundos de espécies silvestres relacionadas (como *C. canephora* e *C. eugenioides*), visando ampliar a base genética. Paralelamente, a médio e curto prazo, recomenda-se a busca de variações genéticas dentro da espécie *C. arabica*, explorando acessos tradicionais e germoplasma subutilizado para incorporação mais rápida de alelos favoráveis. Contudo, o desenvolvimento de novas cultivares de café pode exigir mais de três décadas na ausência de ferramentas biotecnológicas, devido à biologia perene da espécie, à autogamia e à necessidade de cinco a sete ciclos de autofecundação para a obtenção de cultivares (Ariyoshi et al., 2025; Morais & Melo, 2011; Surya Prakash et al., 2016).

Neste contexto, a biotecnologia surge como uma ferramenta estratégica capaz de otimizar tempo e custos no desenvolvimento de cultivares de café (Ariyoshi et al., 2025; Huang et al., 2020; Koutouleas et al., 2023). Com o advento de tecnologias moleculares e de sequenciamento de nova geração, tornou-se possível a genotipagem de milhares de polimorfismos genéticos,

principalmente os polimorfismos de nucleotídeo único (SNP), que constituem marcadores moleculares altamente informativos para a análise da diversidade genética e para o entendimento da genômica quantitativa, funcional e evolutiva (Melia et al., 2025; Morais & Melo, 2011; Santos et al., 2022; Scalabrin et al., 2024; Tran et al., 2016).

Os marcadores SNP podem ser associados a locos de interesses agrônômicos e utilizados na Seleção Assistida por Marcadores Moleculares (SAM) em programas de melhoramento genético (Luis et al., 2022; Tran et al., 2018; Zhang et al., 2021). Essa abordagem permite: seleção precoce e acurada de genótipos portadores de alelos desejáveis em estágios de plântula, independentemente das condições ambientais; redução de ciclos de melhoramento pela eliminação de indivíduos não promissores em estágios iniciais; e piramidação eficiente de múltiplos genes de resistência/tolerância (Zargar et al., 2016). Esses marcadores demonstram versatilidade, sendo eficazes na discriminação de características tanto monogênicas quanto poligênicas (Koutouleas et al., 2023; Mishra, 2020; Morais & Melo, 2011; Tran et al., 2016).

Esses avanços têm impulsionado significativamente as áreas de genômica e bioinformática, viabilizando a associação entre dados genotípicos e fenótipos agronomicamente relevantes, contribuindo para o desenvolvimento de cultivares mais produtivas, resilientes e adaptadas às demandas do mercado global (Dmitriev et al., 2022; Filippis, 2017; Jian & Li, 2021; Scalabrin et al., 2020). Nesse contexto, a presente tese relata aplicação de marcadores moleculares de polimorfismos de nucleotídeo único no gênero *Coffea*.

No Capítulo I, é apresentada uma revisão bibliométrica abordando a cultura do cafeeiro associados as temáticas de estudos de seleção e associação genômica ampla, e as principais metodologias de genotipagem de SNP de baixo rendimento (*low-throughput*), que contribuiu para o embasamento e a elaboração do capítulo subsequente. No Capítulo II são apresentados os resultados do desenvolvimento de marcadores moleculares SNP identificados em loco de característica quantitativa (QTL- *Quantitative Trait loci*) associado à redução da altura de plantas para estudos em *C. arabica*. Estes estudos justificam-se pela necessidade de transpor a lacuna entre a descoberta de marcadores e sua aplicação prática, bem como elucidar aspectos da complexidade genética da característica altura de plantas e projetar perspectivas para a seleção de plantas com porte previsível. Os resultados destes estudos trazem implicações relevantes para programas de melhoramento genético, pois permitem a aplicação de marcadores na seleção assistida no gênero *Coffea*.

Referências

- Amrouk, E. M., Palmeri, F., & Magrini, E. (2025). Global coffee market and recent price developments. *FAO - Food and Agriculture Organization of the United Nations*. <https://www.fao.org/faostat/en/#data>
- 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*(5), 894–900. <https://doi.org/10.1007/S00122-001-0798-8>
- Ariyoshi, C., De Oliveira, F. F., Ferreira, R. V., Sera, G. H., & Pereira, L. F. P. (2025). Genomic research on coffee and its impact on molecular breeding. *Advances in Botanical Research*, *114*, 455–484. <https://doi.org/10.1016/bs.abr.2024.05.003>
- Ariyoshi, C., Freitas, F., Shigueoka, L. H., da Silva, A. G., Arias, A. G., Villalobos, J. V., Ivamoto-Suzuki, S. T., Sera, G. H., Caixeta, E. T., da Rocha, M. R., Moraes, W. B., Luiz Partelli, F., & Protasio Pereira, L. F. (2022). Current challenges and genomic advances toward the development of coffee genotypes resistant to biotic stress. *Genomic Designing for Biotic Stress Resistant Technical Crops*, 159–189. https://doi.org/10.1007/978-3-031-09293-0_4
- Barbosa, I. de P., de Oliveira, A. C. B., Rosado, R. D. S., Sakiyama, N. S., Cruz, C. D., & Pereira, A. A. (2020). Sensory analysis of arabica coffee: cultivars of rust resistance with potential for the specialty coffee market. *Euphytica*, *216*(10), 1–12. <https://doi.org/10.1007/S10681-020-02704-9>
- Bawin, Y., Ruttink, T., Staelens, A., Haegeman, A., Stoffelen, P., Mwanga Mwanga, J. C. I., Roldán-Ruiz, I., Honnay, O., & Janssens, S. B. (2021). Phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica*. *Journal of Systematics and Evolution*, *59*(5), 953–963. <https://doi.org/10.1111/jse.12694>
- Carvalho, A. de (1981). Café: Novas variedades mais produtivas. *Agricultura de Hoje*, Rio de Janeiro, 6:32–34. <http://www.sbicafe.ufv.br:80/handle/123456789/7770>
- Cecafé. (2025). *CECAFE - Relatório Mensal Março 2025*. <https://docs.google.com/viewerng/viewer?url=https://www.cecafe.com.br/site/wp-content/uploads/graficos/cecafe-relatorio-mensal-marco-2025.pdf&hl=en>
- Clarindo, W. R., & Carvalho, C. R. (2008). First *Coffea arabica* karyogram showing that this species is a true allotetraploid. *Plant Systematics and Evolution*, *274*(3–4), 237–241. <https://doi.org/10.1007/s00606-008-0050-y>
- Combes, M. C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G., & Lashermes, P. (2000). Characterization of microsatellite loci in *Coffea arabica* and related coffee species. *Molecular Ecology*, *9*(8), 1178–1180. <https://doi.org/10.1046/j.1365-294X.2000.00954-5.x>
- Conab. (2025). Série histórica de café total (Arábica e Conilon). *Companhia Nacional de Abastecimento*. <https://www.gov.br/conab/pt-br/atuacao/informacoes-agropecuarias/safras/series-historicas/cafe/cafetotalseriehist.xls/view>

- Davis, A. P., Kiwuka, C., Faruk, A., Walubiri, M. J., & Kalema, J. (2022). The re-emergence of Liberica coffee as a major crop plant. *Nature Plants*, 8(12), 1322–1328. <https://doi.org/10.1038/S41477-022-01309-5>
- Dmitriev, A. A., Pushkova, E. N., & Melnikova, N. V. (2022). Plant genome sequencing: modern technologies and novel opportunities for breeding. *Molecular biology 2022* 56:4, 56(4), 495–507. <https://doi.org/10.1134/S0026893322040045>
- FAO. (2025). *FAOSTAT*. Food and Agriculture Organization of the United Nations. <https://www.fao.org/faostat/en/#data/qcl>
- Feitosa, F. de M., Santos, I. G. dos, 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. <https://doi.org/10.1002/csc2.21303>
- Filippis, L. F. De. (2017). Plant bioinformatics: next generation sequencing approaches. *Plant Bioinformatics: Decoding the Phyta*, 1–106. https://doi.org/10.1007/978-3-319-67156-7_1
- Grand View Research. (2024). *Specialty Coffee Market Size, Share & Trends Analysis Report By Age Group (18-24 Years, 25-39 Years, 40-59 Years, Above 60), By Distribution Channel (Retail, Away From Home), By Region, And Segment Forecasts, 2025 - 2030*. <https://www.grandviewresearch.com/industry-analysis/specialty-coffee-market-report#>
- Guyot, R., Hamon, P., Couturon, E., Raharimalala, N., Rakotomalala, J.-J., Lakkanna, S., Sabatier, S., Affouard, A., & Bonnet, P. (2020). WCSdb: a database of wild *Coffea* species. *Database*, 2020. <https://doi.org/10.1093/database/baaa069>
- Huang, L., Wang, X., Dong, Y., Long, Y., Hao, C., Yan, L., & Shi, T. (2020). Resequencing 93 accessions of coffee unveils independent and parallel selection during *Coffea* species divergence. *Plant Molecular Biology*, 103(1–2), 51–61. <https://doi.org/10.1007/s11103-020-00974-4>
- ICO. (2023). Coffee development report. *International Coffee Organization*, 1, 1–118. <https://www.icocoffee.org/documents/cy2024-25/coffee-development-report-2022-23.pdf>
- ICO. (2025a). Coffee market report - March 2025. *International Coffee Organization*, 1, 1–14. <https://ico.org/resources/coffee-market-report-statistics-section/>
- ICO. (2025b). I-CIP in a pocket of downturn amid increased uncertainty. *International Coffee Organization*, 01–14. <https://www.ico.org/documents/cy2024-25/cmr-0325-e.pdf>
- Jian, Y., & Li, M. (2021). A narrative review of single-nucleotide polymorphism detection methods and their application in studies of *Staphylococcus aureus*. *Journal of Bio-X Research*, 4(1), 1–9. <https://doi.org/10.1097/jbr.0000000000000071>
- Koutouleas, A., Collinge, D. B., & Ræbild, A. (2023). Alternative plant protection strategies for tomorrow's coffee. *Plant Pathology*, 72(3), 409–429. <https://doi.org/10.1111/ppa.13676>
- Lashermes, P., Hueber, Y., Combes, M.-C., Severac, D., & Dereeper, 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. <https://doi.org/10.1534/g3.116.030858>
- Luis, S. C. J., Paulino, P. R., Bello-Bello, J. J., Esteban, E. P., Heber, A. R. V., Tarsicio, C. T., Gabino, G. de los S., & Victorino, M. R. (2022). SNP markers identification by genome wide association study for chemical quality traits of coffee (*Coffea* spp.) Germplasm. *Molecular Biology Reports*, 49(6), 4849–4859. <https://doi.org/10.1007/S11033-022-07339-8>
- Malta, M. R., Oliveira, A. C. B., Liska, G. R., Carvalho, G. R., Pereira, A. A., Silva, A. D., Alvaro, L. N., & Mota, D. M. (2021). Selection of elite genotypes of *Coffea arabica* L. to produce specialty coffees. *Frontiers in Sustainable Food Systems*, 5, 715385. <https://doi.org/10.3389/fsufs.2021.715385>
- Melia, T., Fatayat, Wahibah, N. N., Fatonah, S., Roslim, D. I., & Adnan, A. (2025). Genome-wide DNA polymorphisms in two peatland adapted *Coffea liberica* varieties. *BMC Genomic Data*, 26(1), 11. <https://doi.org/10.1186/S12863-025-01305-6>
- Millet, C. P., Delahaie, B., Georget, F., Allinne, C., Solano-Sánchez, W., Zhang, D., Jeune, W., Toniutti, L., & Poncet, V. (2024). Guadeloupe and Haiti’s coffee genetic resources reflect the crop’s regional and global history. *Plants People Planet*, 7(1), 245–262. <https://doi.org/10.1002/ppp3.10584>
- Mishra, M. K. (2020). Genetic resources and breeding of coffee (*Coffea* spp.). *Advances in Plant Breeding Strategies: Nut and Beverage Crops*, 4, 475–515. https://doi.org/10.1007/978-3-030-23112-5_12
- Montagnon, C., Mahyoub, A., Solano, W., & Sheibani, F. (2021). Unveiling a unique genetic diversity of cultivated *Coffea arabica* L. in its main domestication center: Yemen. *Genetic Resources and Crop Evolution*, 68(6), 2411–2422. <https://doi.org/10.1007/s10722-021-01139-y>
- Morais, T. P. de, & Melo, B. de. (2011). Biotecnologia aplicada ao melhoramento genético do cafeeiro. *Ciência Rural*, 41(5), 753–760. <https://doi.org/10.1590/S0103-84782011005000056>
- Pereira, D. R., Nadaleti, D. H. S., Rodrigues, E. C., da Silva, A. D., Malta, M. R., de Carvalho, S. P., & Carvalho, G. R. (2021). Genetic and chemical control of coffee rust (*Hemileia vastatrix* Berk et Br.): impacts on coffee (*Coffea arabica* L.) quality. *Journal of the Science of Food and Agriculture*, 101(7), 2836–2845. <https://doi.org/10.1002/jsfa.10914>
- Prakash, N. S., Devasia, J., Raghuramulu, Y., & Aggarwal, R. K. (2016). Genetic diversity and coffee improvement in India. Volume 2 (pp. 231–268). *Springer International Publishing*. https://doi.org/10.1007/978-3-319-27090-6_11
- Salojärvi, J., Rambani, A., Yu, Z., Guyot, R., Strickler, S., Lepelley, M., Wang, C., Rajaraman, S., Rastas, P., Zheng, C., Muñoz, D. S., Meidanis, J., Paschoal, A. R., Bawin, Y., Krabbenhoft, T. J., Wang, Z. Q., Fleck, S. J., Aussel, R., Bellanger, L., ... 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. <https://doi.org/10.1038/s41588-024-01695-w>

- Santos, T. B. dos, Silva Ferreira, M. F. da, Marques, I., Oliveira, S. C., Zaidan, I. R., Oliveira, M. G., Rodrigues, W. P., Ribas, A. F., Guyot, R., Ramalho, J. C., Ribeiro-Barros, A. I., Pereira, L. F. P., & Partelli, F. L. (2022). Current challenges and genomic advances towards the development resilient coffee genotypes to abiotic stresses. *Genomic Designing for Abiotic Stress Resistant Technical Crops*, 41–69. https://doi.org/10.1007/978-3-031-05706-9_3
- 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. *Plant*, 255(6), 1–16. <https://doi.org/10.1007/s00425-022-03898-z>
- Scalabrin, S., Magris, G., Liva, M., Vitulo, N., Vidotto, M., Scaglione, D., Del Terra, L., Ruosi, M. R., Navarini, L., Pellegrino, G., Berny Mier y Teran, J. C., Toniutti, L., Suggi Liverani, F., Cerutti, M., Di Gaspero, G., & Morgante, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, 15(1), 463. <https://doi.org/10.1038/s41467-023-44449-8>
- Scalabrin, S., Toniutti, L., Di Gaspero, G., Scaglione, D., Magris, G., Vidotto, M., Pinosio, S., Cattonaro, F., Magni, F., Jurman, I., Cerutti, M., Suggi Liverani, F., Navarini, L., Del Terra, L., Pellegrino, G., Ruosi, M. R., Vitulo, N., Valle, G., Pallavicini, A., ... 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), 4642. <https://doi.org/10.1038/s41598-020-61216-7>
- Tran, H. T., Lee, L. S., Furtado, A., Smyth, H., & Henry, R. J. (2016). Advances in genomics for the improvement of quality in coffee. *Journal of the Science of Food and Agriculture*, 96(10), 3300–3312. <https://doi.org/10.1002/jsfa.7692>
- Tran, H. T. M., Furtado, A., Vargas, C. A. C., Smyth, H., Slade Lee, L., & Henry, R. (2018). SNP in the *Coffea arabica* genome associated with coffee quality. *Tree Genetics and Genomes*, 14(5), 1–15. <https://doi.org/10.1007/s11295-018-1282-9>
- Van der Vossen, 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. <https://doi.org/10.1007/s10681-015-1398-z>
- Zargar, S. M., Raatz, B., Sonah, H., Muslimanazir, Bhat, J. A., Dar, Z. A., Agrawal, G. K., & Rakwal, R. (2016). Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. *Journal of Crop Science and Biotechnology 2015 18:5*, 18(5), 293–308. <https://doi.org/10.1007/S12892-015-0037-5>
- Zhang, D., Vega, F. E., Solano, W., Su, F., Infante, F., & Meinhardt, L. W. (2021). Selecting a core set of nuclear SNP markers for molecular characterization of arabica coffee (*Coffea arabica* L.) genetic resources. *Conservation Genetics Resources*, 13(3), 329–335. <https://doi.org/10.1007/s12686-021-01201-y>

CAPÍTULO 1

ANÁLISE BIBLIOMÉTRICA DAS APLICAÇÕES DE MARCADORES SNP EM MELHORAMENTO GENÉTICO DO CAFEIRO

DEZEMBRO / 2025

VIÇOSA – MG

RESUMO

A revolução genômica tem transformado o melhoramento genético de plantas ao longo das últimas décadas, permitindo a identificação precisa de locos associados a características de interesse por meio de marcadores moleculares. Dentre esses, os polimorfismos de nucleotídeo único (SNP) emergiram por sua alta densidade genômica, estabilidade e facilidade de automação, consolidando-se como ferramentas importantes para abordagens modernas de Seleção Genômica Ampla, Associação Genômica Ampla (GWAS) e Seleção Assistida por Marcadores. No gênero *Coffea*, que apresenta singular complexidade genômica devido ao hábito perene, variações de ploidia e base genética relativamente estreita, os SNPs representam uma ferramenta estratégica para superar limitações do melhoramento convencional. Esta lacuna motivou a presente análise bibliométrica, que tem como objetivo mapear sistematicamente as pesquisas sobre aplicações de SNPs em genotipagem de alto (*High throughput genotyping*) e baixo (*Low throughput genotyping*) rendimento no melhoramento do cafeeiro. A literatura revela que a aplicação de SNP no café é relativamente recente, com apenas 31 manuscritos publicados, envolvendo colaboração entre pesquisadores de 15 países. Nos estudos com genotipagem de alto rendimento, identificou-se 83 termos agrupados em 9 clusters temáticos. As genotipagem de baixo rendimento concentraram-se na validação de SNPs para características de qualidade da bebida, resistência a patógenos e rastreabilidade, com 94 termos distribuídos em 7 clusters. As metodologias de baixo rendimento aplicadas foram: *High-Resolution Melting* (HRM), *Kompetitive Allele-Specific PCR* (KASP), *TaqMan*, *Tetra-Primer Amplification Refractory Mutation System* (Tetra-Primer ARMS) e *Allele-Specific PCR* (AS-PCR). Embora as abordagens de varredura genômica ampla tenha possibilitado a identificação de um grande número de SNPs, observa-se ainda uma importante defasagem entre a fase de descoberta e a validação funcional desses marcadores para aplicação prática, particularmente quando se consideram metodologias de baixo rendimento.

Palavras-chave: *Coffea*. Painel de genotipagem. Validação de marcadores SNPs. HRM. KASP. *TaqMan*. *Tetra-Primer* ARMS. AS-PCR.

1 Introdução

O uso de marcadores moleculares revolucionou o ritmo e a precisão da análise genética de plantas, o que, por sua vez, facilitou a implementação do melhoramento molecular de culturas. As últimas quatro décadas testemunharam enormes avanços na evolução dos sistemas de marcadores e das respectivas tecnologias de detecção (Dmitriev et al., 2022; Kumar et al., 2024; Nadeem et al., 2018). Paralelamente aos avanços da genômica, os marcadores moleculares vêm se sofisticando progressivamente, impulsionados pela expansão exponencial de informações genômicas disponíveis (Mishra, 2020; Morais & Melo, 2011; Zargar et al., 2016).

Os marcadores de DNA podem ser identificados com base em técnicas como Reação em Cadeia da Polimerase (PCR), hibridização e sequenciamento, que permitem detectar diferenças entre os genomas dos indivíduos (Grover & Sharma, 2016). Além disso, podem ser categorizados de acordo com sua escala de processamento em alto rendimento de genotipagem (*High throughput genotyping*) e baixo rendimento de genotipagem (*Low throughput genotyping*), que se refere à quantidade de informação que um método pode gerar por unidade de tempo ou custo. Nesse contexto, o rendimento indica quantos locos (localizações no genoma) podem ser genotipados (identificados) simultaneamente (Filippis, 2017; Kockum et al., 2023; Lu et al., 2024; Merot-L'anthoene et al., 2019; Sahoo et al., 2025).

Os marcadores de alto rendimento permitem a genotipagem de milhares de locos simultaneamente, e são ideais para análises em larga escala. Entre os mais utilizados destacam-se os SNPs, *Single Feature Polymorphisms* (SFPs), *Diversity Arrays Technology* (DArt), *Reduced-Representation Libraries* (RRLs), *Restriction Site-associated DNA* (RAD-seq), *Complexity Reduction of Polymorphic Sequences* (CRoPS), *Genotyping by Sequencing* (GBS) e *Shotgun Genotyping* (MSG). As principais diferenças entre essas plataformas estão associadas às metodologias de detecção, que variam quanto ao princípio de captura dos polimorfismos, à necessidade de sequenciamento, ao grau de multiplexação e à profundidade de cobertura genômica (Schatz et al., 2012; Singh et al., 2022; Yadav et al., 2016).

Os marcadores de baixo rendimento permitem genotipar um número limitado de loco por vez, esse processo de genotipagem permite uma análise aprofundada de regiões específicas do genoma. Dentre os marcadores de baixa escala estão *Restriction Fragment Length Polymorphism* (RFLP), *Random Amplified Polymorphic DNA* (RAPD), *Sequence Characterized Amplified Regions* (SCAR), *Amplified Fragment Length Polymorphism* (AFLP), *Simple Sequence Repeat* (Microssatélites ou SSRs), e também os SNPs (Mishra, 2020; Morais & Melo, 2011; Sahoo et al., 2025).

Esses diferentes tipos de marcadores moleculares são adotados no melhoramento do cafeeiro para a identificação de alelos que conferem resistência múltipla a doenças (Alkimim et al., 2017; Almeida et al., 2021; Alvarenga et al., 2011; Brenes et al., 2025; Feitosa et al., 2024; Saavedra et al., 2023), conservação de bancos de germoplasma (Zhang et al., 2021), caracterização de genes de interesse (Almeida, et al., 2021; Alves, et al., 2022; Combes et al., 2000; Gichuru et al., 2008; Mahé et al., 2008; Prakash et al., 2004), análise de diversidade genética (Bikila, 2015; Ferrão et al., 2013; Montagnon et al., 2021; Nong et al., 2025; Surya Prakash et al., 2016), escolha de genitores para formação de populações (Medeiros et al., 2021; Setotaw et al., 2020; Sousa et al., 2019), clonagem (Mishra et al., 2023), monitoramento da pureza genética (Bolívar-González et al., 2022; Silva et al., 2023), estudos da organização e mapeamento genômico (Lashermes et al., 2016; Merot-L'anthoene et al., 2019; Salojärvi et al., 2024; Scalabrin et al., 2024), análise de locos quantitativos (Moncada et al., 2016; Romero et al., 2014; R. A. Silva et al., 2024), estudos de filogenia e evolução (Bawin et al., 2021; Davis et al., 2011), entre muitas outras aplicações e perspectivas.

Os marcadores moleculares do tipo SNPs emergiram como a mais recente geração de marcadores moleculares, destacando-se como ferramentas multifuncionais na genômica contemporânea (Breyer & Smith, 2020; Filippis, 2017; Singh et al., 2022; Yadav et al., 2016). Esses marcadores consistem em variações pontuais de um único nucleotídeo em posições específicas do genoma, decorrentes de substituições das bases nitrogenadas, podendo estas serem transições (purina-purina ou pirimidina-pirimidina) e transversões (purina-pirimidina) (Udoh et al., 2021; Zargar et al., 2016). Eles podem estar localizados tanto em regiões codificantes (*exons*) quanto em regiões não codificantes (*introns* e regiões intergênicas). Essa ubiquidade genômica, combinada com sua elevada densidade nos genomas vegetais, confere aos SNPs uma aplicabilidade eficiente em estratégias de genotipagem, desde metodologias de baixo rendimento a plataformas de alto rendimento (Mammadov et al., 2012; Morgil et al., 2020).

A aplicação de marcadores moleculares SNPs configura-se como uma abordagem genômica inovadora para superar limitações intrínsecas ao melhoramento convencional do cafeeiro, tais como o longo ciclo reprodutivo da espécie, a baixa taxa de recombinação e a dificuldade de fenotipagem de características quantitativas altamente influenciadas pelo ambiente. Os SNPs viabilizam estratégias como Análise de Associação Genômica Ampla (*Genome-Wide Association Studies – GWAS*) (Ariyoshi, et al., 2022; Silva et al., 2022; Ferrão et al., 2023; Gimase et al., 2021; Gonzales et al., 2025; Sant'Ana et al., 2018; Silva et al., 2024; Suela et al., 2023; Zaidan et al., 2025), Seleção Genômica Ampla (*Genome-Wide Selection – GWS*)

(Alkimim et al., 2020; Arcanjo et al., 2024; Sousa et al., 2019); mapeamento de alta resolução e identificação de locos de características quantitativas (*Quantitative Trait Loci* – QTLs) (Merot-L'anthoene et al., 2019), e, conseqüentemente, a Seleção Assistida por Marcadores (SAM) (Akpertey et al., 2021; Bolívar-González et al., 2022; Tao et al., 2013; Zhou et al., 2016).

Diante desta perspectiva, o presente estudo emprega análise bibliométrica sistemática para: (i) mapear a produção científica sobre a utilização de SNPs em GWS, GWAS e plataformas de genotipagem de baixo rendimento; (ii) identificar padrões temporais e geográficos na pesquisa; e (iii) delinear lacunas tecnológicas e oportunidades para o desenvolvimento de cultivares com maior produtividade, resistência a patógenos e qualidade sensorial - atributos essenciais para atender às exigências do mercado mundial de café.

1.1 Revisão bibliográfica

1.1.1 Como os SNPs se tornaram os marcadores mais versáteis para estudos genômicos em plantas cultivadas nos últimos anos?

A superioridade analítica dos SNPs reside em sua alta densidade e distribuição genômica, considerados a forma mais frequente de variação genética natural, com resolução na ordem de kilobases (kb). Estima-se que os SNPs ocorrem uma vez a cada 1.000 bases nitrogenadas ou mais (Mammadov et al., 2012; Marth et al., 1999), o que proporciona cerca de 1% de abundância considerando uma população resultado da baixa taxa de mutação espontânea do DNA. Essa abundância representa mais de 90% de todas as diferenças entre eucariotos, portanto, eles são considerados o melhor recurso de variação genética para estudos populacionais e mapeamento do genoma (Bawin et al., 2021; Huang et al., 2020; Mekbib et al., 2022; Udoh et al., 2021).

O impacto de uma variação genética SNP depende de sua localização no genoma e pode ser consideradas variantes codificadoras e não codificadoras, com possibilidade de efeitos benéficos, neutros ou negativos para o genótipo (Bolívar-González et al., 2022; Mammadov et al., 2012; Marth et al., 1999). As variantes codificadoras ocorrem em éxons e podem alterar a sequência de aminoácidos, potencialmente modificando a estrutura e a função da proteína resultante (Morgil et al., 2020). As variantes não codificadoras englobam diferentes tipos de alterações genéticas como: (i) mutações em íntrons, que podem influenciar o *splicing* do RNA mensageiro, determinando quais éxons serão incluídos na tradução e, conseqüentemente, afetando propriedades essenciais da proteína, como sua solubilidade ou associação com membranas celulares; (ii) variações em regiões regulatórias, como promotores e

intensificadores, que podem modificar os níveis de expressão gênica, impactando não apenas o gene diretamente afetado, mas também genes vizinhos no mesmo cromossomo; (iii) alterações em pequenos RNAs não codificantes, como microRNAs, podem interferir na estabilidade e tradução de RNAs mensageiros, regulando indiretamente a síntese proteica (Kockum et al., 2023).

No entanto, uma limitação dos SNPs é sua natureza bialélica predominante, o que resulta em menor conteúdo informativo por marcador em comparação com marcadores multialélicos, como microssatélites. Estima-se que aproximadamente cinco marcadores SNPs sejam necessários para fornecer informações equivalentes a um único marcador microssatélite. Apesar dessa limitação, a combinação de alta densidade genômica, estabilidade, aplicabilidade em larga e pequena escala, torna os SNPs ferramentas promissoras (Merot-L'anthoene et al., 2019; Zhang et al., 2021).

1.1.2 Identificação de SNPs em genotipagens de *high throughput* e *low throughput*

Os SNPs são detectados quando um nucleotídeo em uma leitura difere daquele presente na posição correspondente do genoma de referência. Na ausência de um genoma de referência, essa comparação é realizada mediante comparação de leituras entre diferentes genótipos utilizando métodos de “montagem de novo” (Kumar et al., 2024).

O processo de chamada de SNPs emprega arquivos de alinhamento gerados por *softwares* de mapeamento, aplicando critérios rigorosos que incluem profundidade de leitura adequada, qualidade das bases e proporções de nucleotídeos no consenso. A eficiência na detecção de SNPs é significativamente ampliada quando se analisam múltiplos genótipos representativos da diversidade da espécie, permitindo uma caracterização mais abrangente da variabilidade genética populacional. (Delomas et al., 2021; Lu et al., 2024; Scheben et al., 2017; Singh et al., 2022).

As plataformas de genotipagem e sequenciamento de SNPs classificam-se em três categorias principais conforme escala e aplicabilidade: *Microarrays*, Sequenciamento de Nova Geração (NGS) e Reação em Cadeia da Polimerase (PCR) (Tabela 1). Os *Microarrays* permitem a análise de densidade extremamente alta (100.000 a 5 milhões de SNPs) com capacidade moderada (centenas a milhares de amostras), sendo adequados para estudos de GWAS (Merot-L'anthoene et al., 2019; Zhang et al., 2021). Derivações de NGS proporcionam densidade que varia de moderada a alta (1.000–100.000 SNPs), dependendo do método, da cobertura, da profundidade de leitura, da espécie e da complexidade do genoma, com rendimento comparável, configurando-se como alternativas versáteis para estudos que exigem

equilíbrio entre custo e resolução genômica (Dmitriev et al., 2022; Filippis, 2017; Yadav et al., 2016).

As plataformas de genotipagem baseadas em PCR são consideradas sistemas de baixo rendimento por analisarem tipicamente de 1 a 100 SNPs por reação, o que as torna particularmente adequadas para análise de marcadores SNPs previamente identificados como promissores para determinada finalidade (Lawrie & Massey, 2023; Marth et al., 1999; Schatz et al., 2012). A alta especificidade na detecção de variantes em locos pré-determinados e fácil análise são vantagens destas metodologias em relação a genotipagem em larga escala. Entre as principais aplicações destacam-se: validação de SNPs candidatos em GWAS; a genotipagem rotineira em programas de melhoramento genético e a implementação em laboratórios de diagnóstico genômico. Consolidando seu papel como metodologia versátil e acessível para análise genômica direcionada (Jian & Li, 2021).

A escolha entre as diversas plataformas de genotipagem por SNPs deve considerar fatores interdependentes, como: (i) a densidade de marcadores requerida; (ii) o tamanho e estrutura da população de estudo; (iii) a infraestrutura computacional disponível; e (iv) as limitações orçamentárias. A seleção criteriosa da tecnologia é essencial para equilibrar custo, resolutividade e viabilidade técnica (Breyer & Smith, 2020).

Tabela 1: Exemplos da diversidade de marcadores moleculares de alto e baixo rendimento e suas características de processamento.

Marcadores	Rendimento	Quantidade de DNA na reação	Enzimas de Restrição	PCR	Informações de sequência	Multiplexação	Tamanho amostral
RFLPs	Baixo	2-10 µg	Sim	Não	Não	Difícil	<50-100
RAPDs	Médio	5-10 µg	Não	Sim	Sim	Difícil	<100
AFLPs	Médio	~1 µg	Sim	Sim	Sim	Possível	<100
SSRs	Médio	10-20 ng	Não	Sim	Sim	Possível	48-384
SNPs	Alto	> 5ng	Não	Sim	Sim	Possível	Variável
SFPs	Alto	50-100 ng	Sim	Não	Não	Não	100-500
DArT	Alto	50-100 ng	Sim	Não	Não	Não	100-500
RRLs	Muito alto	25 ng	Sim	Sim	Sim	Possível	>1000
RAD-seq	Muito alto	300 ng	Sim	Sim	Sim	Possível	>1000
CRoPs	Muito alto	300 ng	Sim	Sim	Sim	Possível	>1000
GBS	Muito alto	100 ng	Sim	Sim	Sim	Possível	>1000
MSG	Muito alto	10 ng	Sim	Sim	Sim	Possível	>1000

RFLPs: *Restriction Fragment Length Polymorphisms*; RAPDs: *Random Amplified Polymorphic DNA*; SCAR: *Sequence Characterized Amplified Region*; AFLPs: *Amplified Fragment Length Polymorphisms*; SSRs: *Simple Sequence Repeats* (ou *Microsatelites*); SNPs: *Single Nucleotide Polymorphisms*; SFPs: *Single Feature Polymorphisms*; DArT: *Diversity Arrays Technology*; RRLs: *Reduced Representation Libraries*; RAD:seq: *Restriction:site Associated DNA Sequencing*; CRoPs: *Cleaved Randomly Amplified Polymorphic Sequences*; GBS: *Genotyping-by:Sequencing*; MSG: *Multiplexed Shotgun Genotyping*. Adaptado de Zargari et al., 2016.

Para a genotipagem *low throughput* de SNPs na cultura do cafeeiro, as metodologias reportadas na literatura incluem técnicas baseadas em PCR em tempo real, com análises por curvas de dissociação (*High Resolution Melting* - HRM), PCR alelo-específico competitivo (*Kompetitive Allele-Specific PCR* – KASP), TaqMan (baseado em sondas); além de abordagens que podem ser executadas em gel de agarose, como *Tetra-Primer ARMS-PCR* (*Tetra-Primer Amplification Refractory Mutation System*) e alelo-específico (*Allele-Specific PCR* – AS-PCR).

1.1.2.1 Genotipagem *High Resolution Melting* (HRM)

A genotipagem HRM é uma técnica pós-PCR que permite a detecção de variantes genéticas com base no comportamento de dissociação térmica do DNA de fita dupla (dsDNA), realizada em um sistema de tubo fechado. Após a amplificação por PCR da região contendo os SNPs de interesse, o DNA é submetido a um gradiente de temperatura, enquanto a fluorescência dos corantes intercalantes específicos é monitorada continuamente (Bolívar-González et al., 2022; Gatica-Arias et al., 2023; Kim et al., 2023; E. M. de A. Silva et al., 2023).

O princípio de detecção baseia-se na ligação do corante fluorescente ao DNA de dupla fita em baixas temperaturas, quando a fluorescência é máxima. À medida que a temperatura aumenta, ocorre a desnaturação do DNA e liberação do corante, causando uma queda abrupta na fluorescência (cerca de 1000 vezes). Os SNPs alteram a estabilidade térmica do DNA, gerando perfis de dissociação distintos que podem ser discriminados pela Temperatura de *Melting* (T_m) (Combes et al., 2018; Wojdacz et al., 2008).

A forma e o T_m das curvas de fusão são influenciados por propriedades intrínsecas do *amplicon*, como teor de GC, comprimento e sequência. Os SNPs que envolvem trocas C/T, G/A, C/A ou G/T geralmente promovem variações térmicas mais detectáveis ($\Delta T_m > 0,5-1,0^\circ\text{C}$), enquanto trocas C/G (ΔT_m de $0,2-0,5^\circ\text{C}$) e A/T ($\Delta T_m < 0,2^\circ\text{C}$) podem ser mais difíceis de distinguir. Os dados gerados na análise de HRM são usualmente representados por diferentes formas gráficas de fluorescência em função da temperatura. A maioria dos *softwares* de análise interpreta de modo automatizado os resultados (Kim et al., 2023).

1.1.2.2 Genotipagem *Kompetitive Allele-Specific PCR* (KASP)

A tecnologia KASP utiliza dois *primers* específicos para os SNPs, cada um contendo uma sequência complementar à região flanqueadora do polimorfismo, mas com um nucleotídeo diferente na extremidade 3', correspondente ao alelo específico. Eles são marcados com fluoróforos distintos (por exemplo, FAM e HEX), permitindo a detecção diferenciada dos alelos

durante a PCR. Um *primer reverse* é utilizado para amplificar a região de interesse que é complementar a ambos *primers* específicos (Akpertey et al., 2021; Semagn et al., 2014).

A detecção é realizada após a amplificação, utilizando um leitor de placas calibrado com filtros específicos para os comprimentos de onda dos fluoróforos (Merot-L'anthoene et al., 2019). A intensidade da fluorescência é analisada para determinar o genótipo da amostra. Por exemplo, a presença de fluorescência apenas no comprimento de onda associado ao fluoróforo FAM indica homozigose para o alelo 1, enquanto a fluorescência apenas no comprimento de onda associado ao fluoróforo HEX indica homozigose para o alelo 2. A presença de fluorescência em ambos os comprimentos de onda indica heterozigose (Millet, Allinne, et al., 2024; Millet, Delahaie, et al., 2024; Semagn et al., 2014).

1.1.2.3 Genotipagem TaqMan

A genotipagem de SNPs pela tecnologia TaqMan também baseia-se no uso *primers* marcados com sondas. Para os SNPs, duas sondas são projetadas: uma específica para o alelo de referência e outra para o alelo alternativo. Cada sonda é marcada com fluoróforos diferentes que emitem comprimentos de onda distintos (por exemplo, FAM e VIC) (Sahoo et al., 2025; Tao et al., 2013).

Cada sonda consiste em uma sequência oligonucleotídica complementar à região flangeadora dos SNPs, marcada com um fluoróforo em sua extremidade 5' e um *quencher* (bloqueador) na extremidade 3'. O *quencher* suprime a fluorescência do fluoróforo enquanto a sonda está intacta. Durante os ciclos de PCR, na etapa extensão, a DNA polimerase sintetiza a nova fita de DNA a partir dos *primers*, quando a polimerase encontra a sonda TaqMan, sua atividade exonuclease 5'→3' cliva a sonda, liberando o fluoróforo do *quencher* e permitindo a emissão de fluorescência (Campos et al., 2023; Hicks et al., 2023; Tao et al., 2013).

Cada fluoróforo liberado emite luz em um comprimento de onda específico, correspondendo ao alelo presente. Ao final da reação, a intensidade da fluorescência é analisada para determinar o genótipo da amostra. Por exemplo, a presença de fluorescência apenas no comprimento de onda associado ao corante FAM indica homozigose para o alelo 1, enquanto a fluorescência apenas no comprimento de onda associado ao corante VIC indica homozigose para o alelo 2. A presença de fluorescência em ambos os comprimentos de onda indica heterozigose. A tecnologia TaqMan oferece possibilidade de genotipagem de múltiplos SNPs em uma única reação (multiplexação) (Sahoo et al., 2025).

1.1.2.4 Genotipagem *Tetra-Primer Amplification Refractory Mutation System (Tetra-Primer ARMS)*

A genotipagem *Tetra-Primer ARMS* é uma abordagem baseada em PCR multiplex que permite a detecção específica de SNPs por meio da amplificação seletiva de alelos. Esse método fundamenta-se na utilização de dois *primers* internos (*Inner Forward/Reverse*) que são projetados com uma base discriminatória na extremidade 3', permitindo ligação preferencial aos alelos-alvo da troca dos SNPs (Ye et al., 2001). Essa base não-complementar é inserida na penúltima posição 3' para aumentar a especificidade, inibindo a extensão pela DNA polimerase em casos de incompatibilidade. Dois *primers* externos (*Outer Forward/Reverse*) amplificam um fragmento longo de DNA como controle (Tanha et al., 2015; Ye et al., 2001).

A combinação desses quatro *primers* em uma única reação de PCR permite a amplificação simultânea, gerando *amplicons* de tamanhos distintos que podem ser facilmente diferenciados por eletroforese em gel de agarose (Alyethodi et al., 2018). Observam-se três tipos de padrões de bandas: um fragmento maior, que serve como controle interno da reação; e dois fragmentos menores, cada um correspondendo a um alelo específico dos SNPs analisados. Esses padrões permitem a discriminação dos genótipos em indivíduos heterozigotos exibem as três bandas (controle + dois alelos específicos), enquanto homozigotos apresentam apenas duas bandas (controle + um alelo, variando entre dominante ou recessivo) (Medrano & Oliveira, 2014; Schincariol, 2011).

1.1.2.5 Genotipagem *Allele-Specific PCR (AS-PCR)*

A genotipagem AS-PCR proporciona distinção de variantes alélicas de um até cinco nucleotídeos. Esta técnica baseia-se na amplificação diferencial de alelos através de *primers* específicos, possibilitando a identificação de SNPs em uma única reação realizada sob condições convencionais de PCR (Gaudet et al., 2009).

O sistema utiliza três oligonucleotídeos: um *primer reverso* comum e dois *primers forwards* alelo-específicos, estes últimos contendo caudas não homólogas de distintos comprimentos. Esta configuração gera produtos de amplificação com tamanhos diferenciados para cada alelo, os quais podem ser discriminados por eletroforese em gel de agarose (Gaudet et al., 2009). A especificidade da reação é significativamente aumentada pela introdução estratégica de um *mismatch* desestabilizador na penúltima posição da extremidade 3' dos *primers* alelo-específicos, o que inibe eficientemente a amplificação de alelos não-alvo (Kalendar, Shustov, et al., 2022; Liu et al., 2012).

Como principais vantagens, destaca-se o baixo custo operacional em comparação com todos os métodos que necessitam de qPCR. É limitada por sua otimização trabalhosa dos parâmetros de reação e avaliação de resultados delongada. Por isso, algumas metodologias de modificação vêm sendo propostas para otimização de resultados e uso em qPCR (Amangeldiyeva et al., 2023; He et al., 2020, 2022; Kalendar, Baidyussen, et al., 2022; Lefever et al., 2019; Liu et al., 2012).

1.1.3 GWS e GWAS como ferramentas genômicas para identificação de SNPs aplicáveis ao melhoramento genético vegetal

As ferramentas de Seleção Genômica Ampla (GWS) e Estudos de Associação Genômica Ampla (GWAS) representam pilares fundamentais para a identificação e aplicação de SNPs no melhoramento genético vegetal de precisão. Ambas se fundamentam na análise de dados genômicos, representados por milhares de marcadores SNP distribuídos ao longo do genoma, obtidos pelo sequenciamento do DNA de populações de plantas.

A GWS é uma metodologia de predição que utiliza os SNPs para estimar os Valores Genéticos Estimados de Melhoramento (GEBVs), e assim, permitir a seleção de indivíduos com base no seu potencial genético, mesmo antes da expressão fenotípica. Em contrapartida, o GWAS é uma abordagem que identifica SNPs individualmente associados ao fenótipo de características agronomicamente relevantes, o que fornece *insights* sobre a arquitetura genética do caráter e viabiliza a seleção assistida para genes de efeito maior (Korte & Farlow, 2013; Mekbib et al., 2022, Merot-L'anthoene et al., 2019; Meuwissen & Goddard, 2001; Santos et al., 2022; Spence et al., 2025).

O princípio dessas metodologias é o desequilíbrio de ligação, no qual SNPs próximos no mesmo cromossomo apresentam maior associação entre si do que aqueles mais distantes, devido à menor probabilidade de serem separados por eventos de recombinação genética (Silva et al., 2022; Silva et al., 2024; Zargar et al., 2016). Como a taxa de recombinação é geralmente proporcional à distância entre os locos, SNPs adjacentes tendem a ser co-herdados. Variações nas taxas de recombinação ao longo do genoma levam à formação de blocos de desequilíbrio de ligação (ou blocos de haplótipos), regiões com forte associação alélica entre múltiplos SNPs, que são essenciais para inferir relações genótipo-fenótipo (Arcanjo et al., 2024; Kockum et al., 2023; Meuwissen & Goddard, 2001).

Uma vez identificados os SNPs candidatos com correlação com o fenótipo-alvo, estes podem ser convertidos em marcadores moleculares específicos utilizando metodologias de genotipagem de baixo rendimento, permitindo a genotipagem direcionada exclusivamente para

o polimorfismo de interesse. Esses marcadores desenvolvidos e validados, ao explorarem eficientemente a variação genética natural presente em populações diversas, tornam-se ferramentas cruciais para a seleção assistida por marcadores, otimizando a identificação de genótipos superiores e acelerando o desenvolvimento de cultivares melhoradas com características agrônomicas desejáveis (Arcanjo et al., 2024; Uffelman et al., 2021).

1.2 Material e métodos

Para análise da produção científica relacionada ao gênero *Coffea* e suas aplicações na genotipagem de SNPs, adotou-se uma abordagem exploratória baseada em bibliometria. A busca bibliográfica foi conduzida prioritariamente na base *Scopus* (*Elsevier*), selecionada por sua ampla cobertura multidisciplinar e consistência na indexação de periódicos relevantes para a área. Embora consultas exploratórias tenham sido realizadas também nas bases *Web of Science* e *PubMed*, observou-se que estas retornaram um número significativamente menor de artigos diretamente vinculados aos termos de busca utilizados, reforçando a adequação da *Scopus* como fonte primária para o escopo deste estudo.

A revisão sistemática foi estruturada em dois eixos metodológicos complementares, visando abranger desde a identificação até a aplicação de marcadores SNPs funcionais em *Coffea* spp. O primeiro eixo abrangeu estudos genômicos com ênfase na identificação de marcadores SNPs, agronomicamente promissores ao melhoramento genético, por meio de abordagens de GWS e GWAS. Para essa finalidade, utilizou-se a *string* de busca: "*Coffea AND (GWS OR GWAS OR SNP-MARKER)*". O segundo eixo metodológico foi direcionado à validação e aplicação de marcadores SNPs previamente identificados abrangendo todas as técnicas de genotipagem de baixo rendimento documentadas na literatura cafeeira. Para essa abordagem, empregou-se *string* de busca: "*Coffea AND (HRM OR KASP OR TaqMan OR rhAmp OR AS-PCR OR ARMS OR MTPA)*".

A busca na base de dados foi realizada considerando a ocorrência dos termos especificados, nos campos de título, resumo e palavras-chave dos registros. Todas as publicações recuperadas foram sistematicamente analisadas, a fim de verificar sua real aderência aos critérios temáticos definidos nas *strings* de busca. Não foram aplicadas restrições quanto ao período de publicação, garantindo a inclusão abrangente de todos os estudos relevantes disponíveis na base consultada.

Os dados bibliográficos foram processados no *software VOSviewer*, especializado em análise de redes científicas (Eck & Waltman, 2010). O algoritmo *VOS* (*Visualization of Similarities*) agrupa termos por similaridade, minimizando distâncias entre elementos

correlacionados, e mapeia coocorrências em que o teor de associação entre termos é calculada com base na frequência de aparição conjunta no texto. A partir disso, são gerados *clusters* temáticos em que cada *cluster* representa um conjunto de termos fortemente interligados. Nos *clusters* o tamanho reflete a frequência dos termos ou o volume de publicações associadas, e a proximidade espacial entre *clusters* indica relações temáticas.

Para a análise da base de dados bibliográfica, foram aplicadas técnicas de agrupamento baseadas em logaritmo de modularidade (*Log-likelihood/Modularity clustering*), visando identificar padrões de associação entre termos e colaborações científicas. Foram analisadas a coocorrência de palavras-chave, extraídas dos registros bibliográficos, considerando suas frequências e relações. Também foi feita análise da rede de coautoria entre países, em que foram mapeadas as colaborações científicas com base na afiliação dos autores.

1.3 Resultados

A plataforma retornou 39 publicações, das quais 20 estudos de fato relacionavam o gênero *Coffea* com GWS, GWAS e/ou marcadores SNPs na busca "*Coffea AND (GWS OR GWAS OR SNP-MARKER)*" (Figura 1). Os primeiros estudos foram publicados em 2016, abordando qualidade de bebida, caracteres agrônômicos, estresses bióticos e abióticos. Dentre os trabalhos analisados, a maioria tinha relatos de cruzamentos interespecíficos para melhoramento genético populacional, em que 12 manuscritos estavam diretamente ligados a *C. arabica*, seis a *C. canephora* e dois incluíam ambas as espécies.

Quanto à produção acadêmica, foram encontrados 108 autores que publicaram sobre o tema. No entanto, apenas 24 (27,58%) haviam publicado mais de dois trabalhos. A produção científica está concentrada em 10 países, com 12 manuscritos com cooperação entre países (Tabela 2). As áreas temáticas das publicações ficaram concentradas em ciências agrárias e biológicas (53%), e bioquímica, genética e biologia molecular (41%).

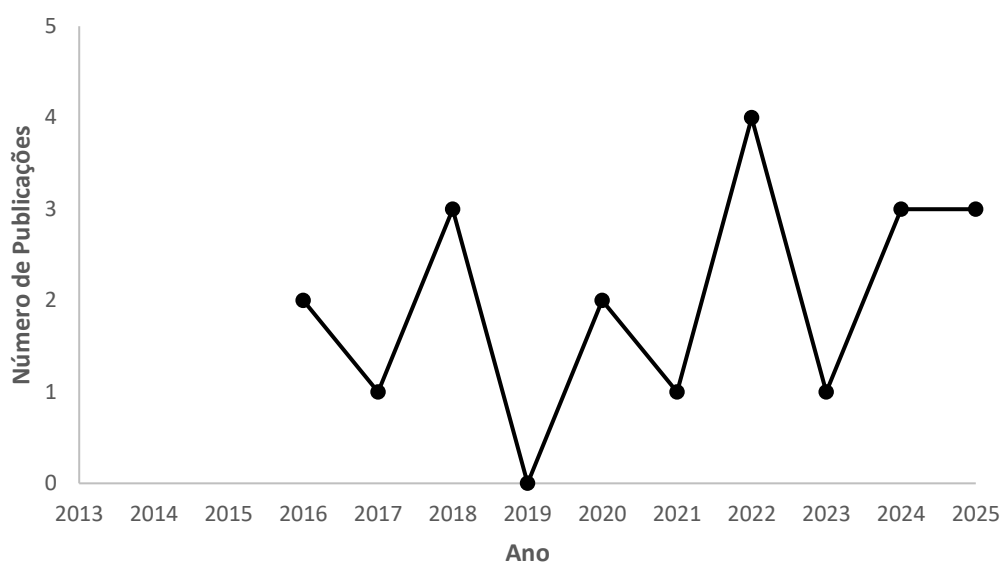


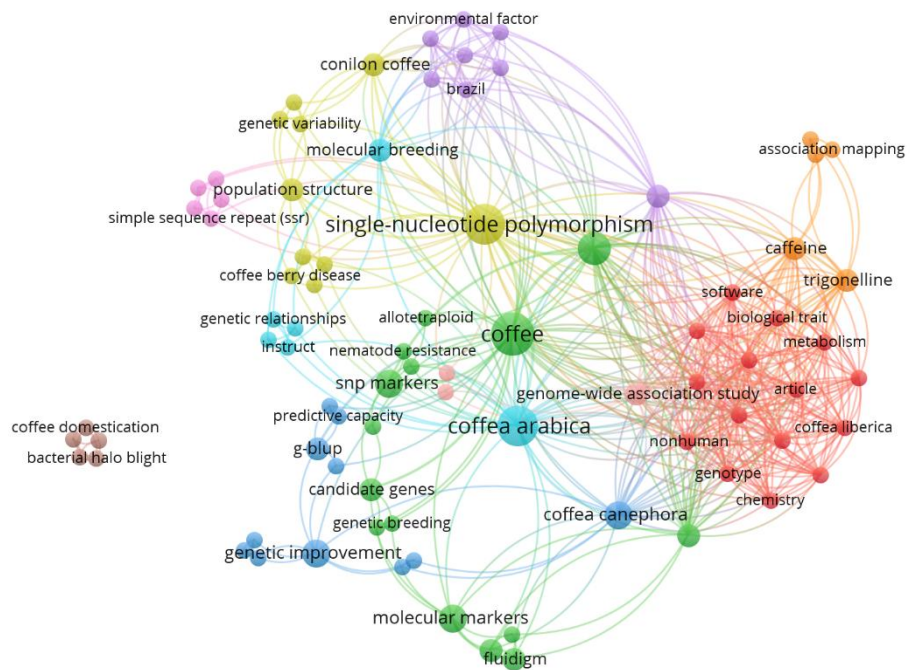
Figura 1: Revisão sistemática do número de publicações científicas de GWS e GWAS aplicados a *Coffea* spp.

Tabela 2: Revisão sistemática da distribuição geográfica de publicações sobre GWS e GWAS aplicados a *Coffea* spp.

País	Documentos	Citações
Estados Unidos	7	207
Costa Rica	2	47
Brasil	14	173
China	2	45
México	2	22
Austrália	1	30
França	1	61
Vietnã	1	30
Colômbia	1	38
Quênia	1	28

Na construção do mapa bibliométrico, foram identificados 83 termos, organizados em nove *clusters* interconectados e um *cluster* isolado, relacionado à doença mancha aureolada (*Pseudomonas syringae*) (Figura 2A). O gráfico de evolução temporal revelou tendências recentes em palavras-chave, como: algoritmos genéticos, fatores ambientais, Brasil, melhoramento genético e marcadores SNPs (Figura 2B).

A



B

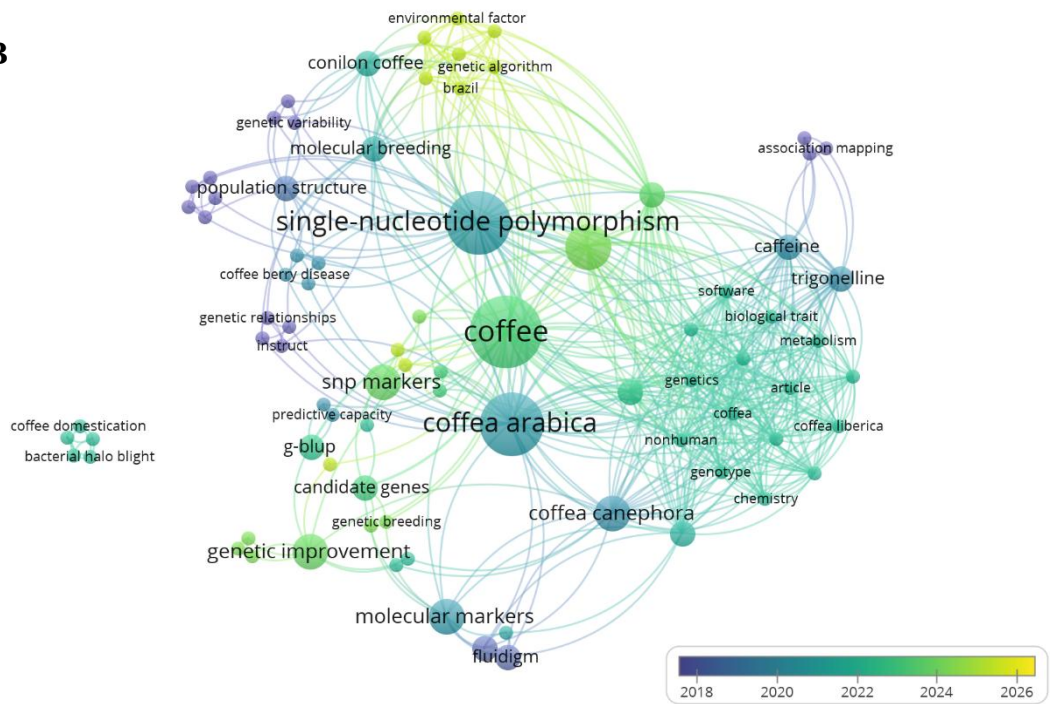


Figura 2: Mapeamento bibliométrico de estudos sobre GWS e GWAS aplicados a *Coffea spp.*

A: Rede de coocorrência. **B:** Evolução temporal.

Os resultados de busca para métodos de genotipagem de baixo rendimento para SNPs (PCR ou qPCR), inicialmente retornou 30 publicações, das quais 11 de fato estavam relacionadas a temática proposta (Figura 3). O primeiro estudo foi publicado em 2013, e posteriormente uma produção científica esparsa no tema. As publicações mais recentes abordam temas como agrobiodiversidade e variabilidade genética, refletindo uma tendência crescente em aplicações de rastreabilidade e de melhoramento genético.

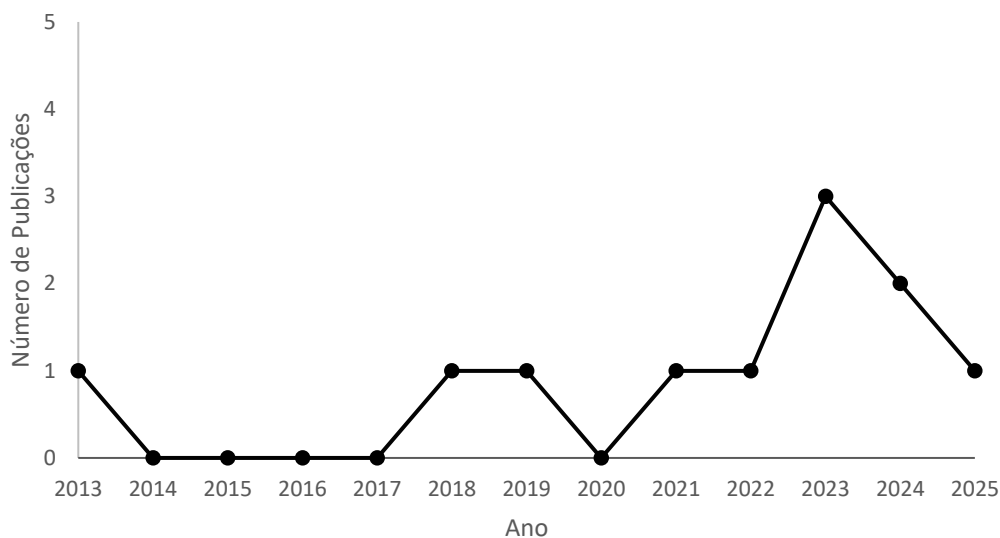


Figura 3: Revisão sistemática do número de publicações científicas sobre metodologias de genotipagem SNPs de baixo rendimento aplicados a *Coffea* spp.

Foram identificados 58 autores associados às publicações, distribuídos em 11 países, com 11 manuscritos com colaboração entre países (Tabela 3). Desses, nove autores possuíam mais de 10 citações, indicando um núcleo de pesquisadores ativos, porém limitado. As publicações classificaram-se em seis categorias principais: Ciências Agrárias e Biológicas; Bioquímica (53%); Genética e Biologia Molecular (29%); Engenharia (6%); Imunologia e Microbiologia (6%); Multidisciplinar (6%).

Tabela 3: Revisão sistemática da distribuição geográfica de publicações sobre metodologias de genotipagem de baixo rendimento para SNPs aplicados a *Coffea* spp.

País	Documentos	Citações
Costa Rica	4	11
Estados Unidos	3	30
França	3	40
Haiti	2	2
Bélgica	1	2
Vietnã	1	2
Brasil	4	14
China	1	24
Gana	1	28
Tailândia	1	24
Holanda	1	21

Na análise de redes, foram catalogadas 94 palavras-chave, organizadas em sete *clusters*: *cluster* 1 com 29 itens, *cluster* 2 com 19 itens, *cluster* 3 com 12 itens, *cluster* 4 com 11 itens, *cluster* 5 com 10 itens, *cluster* 6 com 7 itens e *cluster* 7 com 6 itens (Figura 4A). Os termos com maior força de conexão e publicações mais recentes foram *Coffea*, HRM, genética, *Coffea arabica*, doenças de planta e SNPs (Figura 4B).

Na análise bibliométrica conjunta de todos os manuscritos, os "Avanços em Genética e Melhoramento de Plantas de Café e Detecção de Patógenos" emergiu como conceito central que unifica as diversas linhas de pesquisa (Figura 5). Este núcleo temático se ramifica em sete vertentes principais de investigação: estudos comparativos entre espécies de *Coffea*, desenvolvimento e aplicação de marcadores moleculares, uso de GWAS, estratégias de melhoramento para resistência a doenças, gestão de bancos de germoplasma, análises de diversidade genética e investigações sobre os processos históricos de domesticação. O mapeamento sistemático identificou e organizou 71 conceitos-chave em um fluxograma temático abrangente, revelando as interconexões entre pesquisa básica e aplicada.

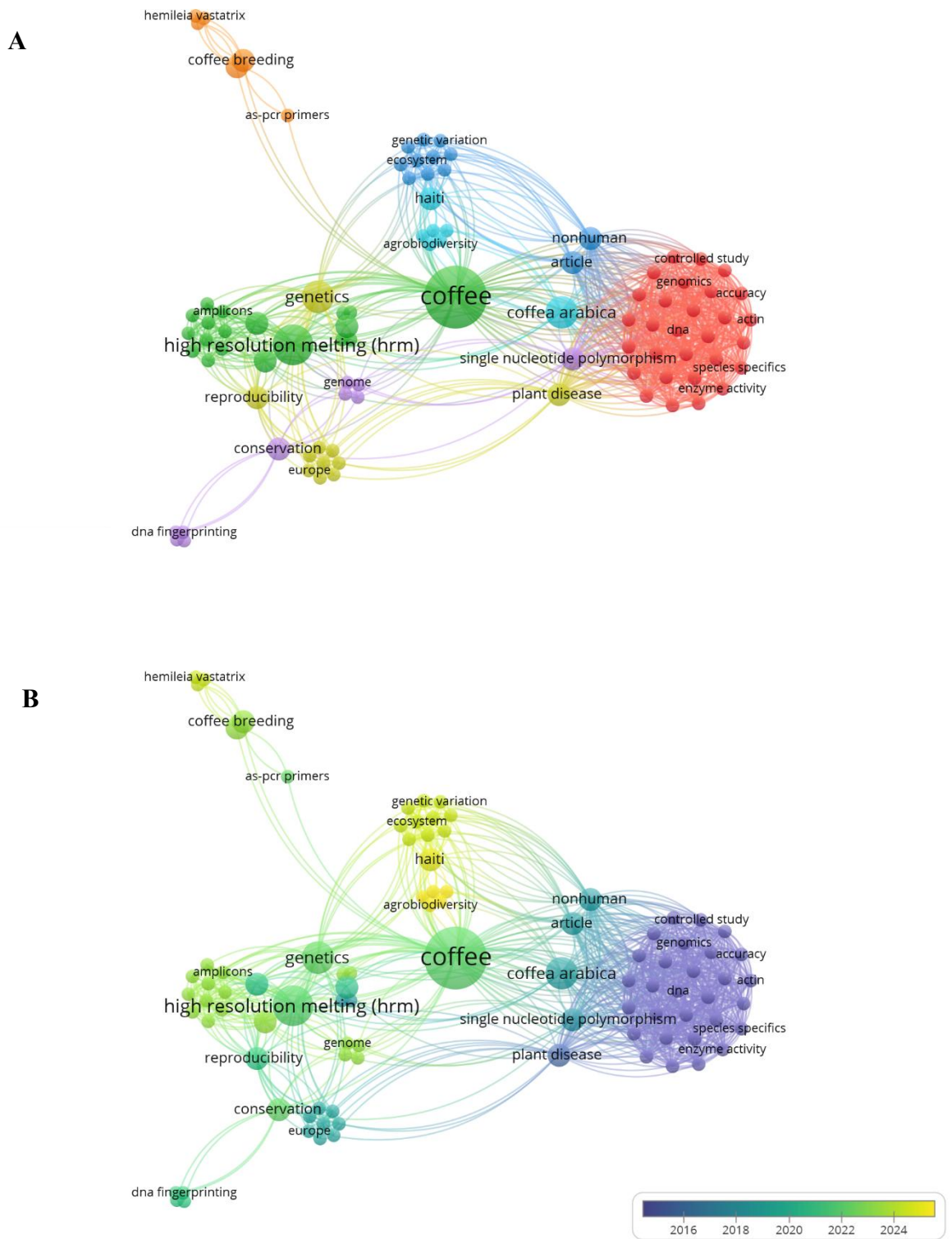


Figura 4: Mapeamento bibliométrico de estudos aplicados a genotipagem de baixo rendimento para SNPs em *Coffea*. **A:** Rede de coocorrência. **B:** Evolução temporal.

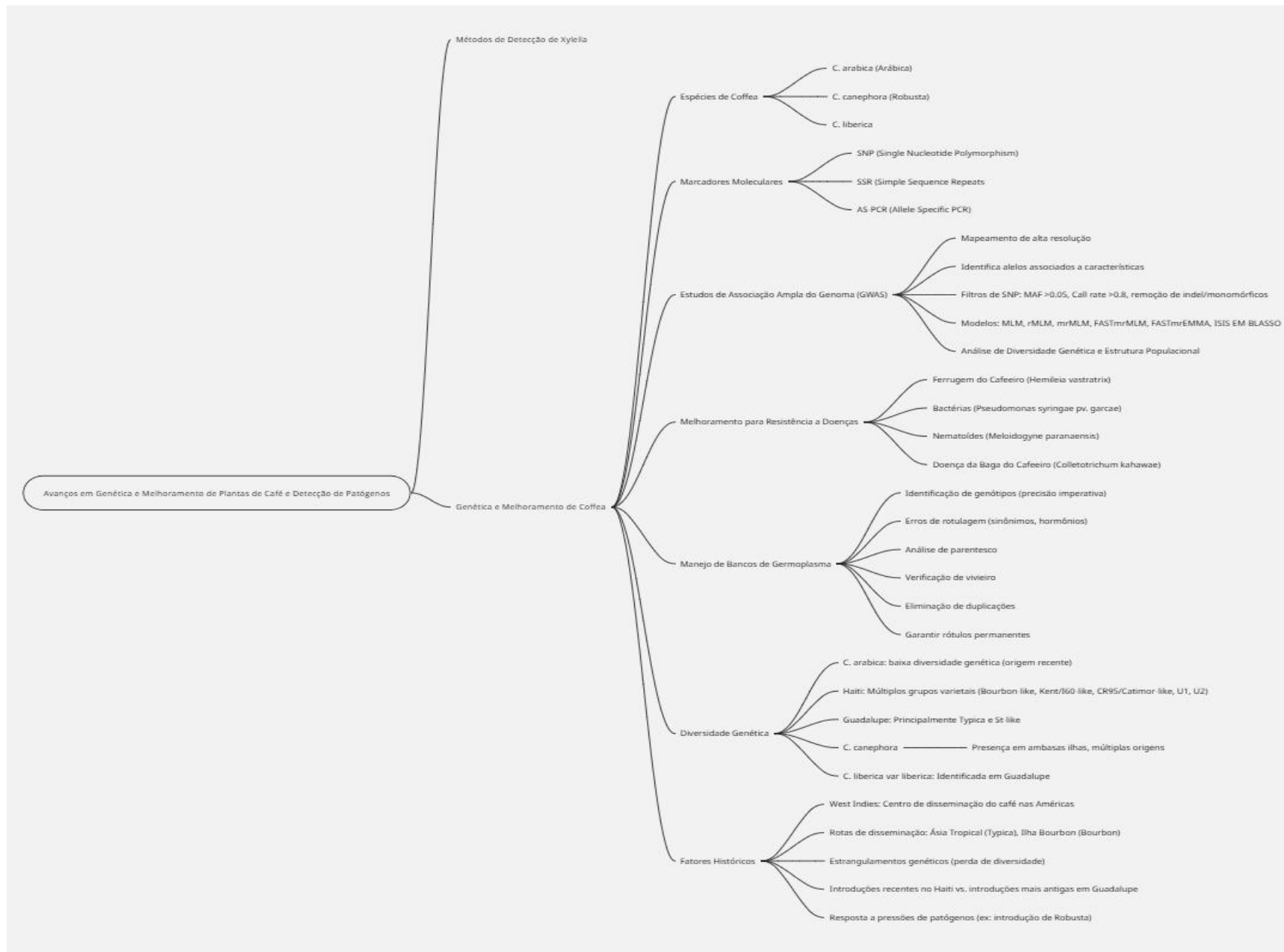


Figura 5: Análise sistemática em *clusters* dos temas dos manuscritos avaliados nos estudos genômicos aplicados ao gênero *Coffea* envolvendo GWS, GWAS e métodos de genotipagem de SNPs de baixo rendimento.

1.4 Discussão

Este estudo bibliométrico evidencia um cenário científico incipiente, porém promissor, no emprego de GWS, GWAS e validação de marcadores SNPs no melhoramento genético de *Coffea* spp. A identificação de apenas 31 manuscritos nas temáticas, desde 2013, sugere que essas abordagens genômicas são relativamente recentes na pesquisa cafeeira. No entanto, em comparação com culturas anuais de amplo estudo genômico, como milho (*Zea mays*), soja (*Glycine max*) e trigo (*Triticum aestivum*), as pesquisas em *Coffea* spp. ainda apresentam lacunas significativas, tanto em número de publicações quanto em aplicações práticas consolidadas (Luis et al., 2022; Sant’Ana et al., 2018; Suela et al., 2023; Uffelmann et al., 2021; Zaidan et al., 2025).

Essa defasagem decorre, em grande parte, da biologia perene do cafeeiro, cujo ciclo longo demanda extensas áreas experimentais, elevados custos operacionais e longos períodos de condução dos experimentos de campo (Nonato et al., 2021; Scalabrin et al., 2020, 2024). A esse cenário somam-se os elevados custos da genotipagem em larga escala, contrastando com o baixo investimento financeiro destinado à cultura, uma vez que os programas de melhoramento dependem majoritariamente de instituições públicas (Clarindo & Carvalho, 2008, 2009; Lashermes et al., 2016; Salojärvi et al., 2024). Ademais, a padronização de protocolos específicos para o poliploide *C. arabica* ainda constitui um desafio, demandando investigações aprofundadas sobre a interação entre subgenomas (Clevenger et al., 2015; Lu et al., 2024).

Os estudos disponíveis de GWS e GWAS concentram-se na identificação de marcadores moleculares associados a características agrônômicas com influências na produtividade e qualidade do cultivo. Dentre estas estão resistência a patógenos como ferrugem (*Hemileia vastatrix*) (Silva et al., 2022; Silva et al., 2024; Sousa et al., 2024; Zaidan et al., 2025), cercosporiose (*Cercospora cafeicola*) (Silva et al., 2022), mancha aureolada (*Pseudomonas syringae* pv. *Garcae*) (Ariyoshi, et al., 2022), antracnose dos frutos - CBD (*Colletotrichum kahawae*) (Gimase et al., 2020), nematoides (*Meloidogyne paranaenses*) (Gonzales et al., 2025) e bicho mineiro (*Leucoptera coffeella*) (Alkimim et al., 2017; Sousa et al., 2019).

Além dos atributos de resistência, há também um número crescente de estudos voltados à qualidade química dos grãos, com destaque para teores de cafeína, trigonelina e compostos voláteis (Luis et al., 2022; Muniz et al., 2024; H. T. M. Tran et al., 2018), bem como lipídios e diterpenos (Muniz et al., 2024; Sant’Ana et al., 2018). Também são relatados para caracteres

morfoagronômicos (Alkimim et al., 2018, 2020; Silva et al., 2022; Silva et al., 2024; Sousa et al., 2024; Suela et al., 2023) e diversidade genética (Sousa et al., 2017; Zhou et al., 2016).

As metodologias de baixo rendimento de genotipagem de SNPs em café mostram-se em fase de consolidação, com 58% das publicações concentradas nos últimos três anos. Este crescimento recente pressupõe uma fase anterior de descoberta, na qual abordagens de alto rendimento identificaram genes e QTLs de interesse, criando assim a base necessária para o desenvolvimento e aplicação de ferramentas genotípicas mais direcionadas e acessíveis. Simultaneamente isso revela um potencial promissor para estudos longitudinais, especialmente para popularização de metodologias para validação, e estudos comparativos entre plataformas de genotipagem, quanto a acurácia, custo e aplicabilidade em programas de melhoramento.

A genotipagem por HRM foi aplicada em três estudos para desenvolvimento de marcadores SNPs em café. Silva et al. (2023) e Combes et al., (2018) validaram marcadores SNPs via HRM detecção de adulterações entre *C. arabica* com *C. canephora*. Bolívar-González et al. (2022), identificaram 15 marcadores SNPs interespecíficos e 4 intraespecíficos em *C. arabica*, em genes relacionados ao acúmulo de sacarose nas sementes (*SUS2*, *CaKO1* e *CcOAOMT*) e na região não codificadora *ITS2*, comprovando sua eficácia para caracterização de germoplasma, rastreabilidade e autenticação na cadeia exportadora de café.

Os estudos que empregaram a genotipagem KASP em café concentraram-se predominantemente em análises de diversidade genética, tanto entre espécies quanto intraespecíficas. Merot-L'anthoene et al. (2019) sintetizaram e validaram ensaios KASP para 42 SNPs, selecionados a partir de um *microarray* de 8,5 mil SNPs, com o objetivo de discriminar geneticamente *C. arabica* e *C. canephora*. Posteriormente, Akpertey et al. (2021) utilizaram 120 marcadores SNPs KASP para examinar erros de parentesco, rotulagem incorreta, diversidade genética e estrutura populacional em 400 acessos de *C. canephora* procedentes de diferentes países e mantidos em um banco de germoplasma em Gana. Mais recentemente, Millet et al. (2024a) identificaram fluxo gênico entre variedades tradicionais de *C. arabica* (Typica, Bourbon e Catimor) por meio de 96 SNPs genotipados via KASP, em sistemas agroflorestais no Haiti. No mesmo ano, Millet et al. (2024b) empregaram 80 SNPs em ensaios KASP para caracterizar a diversidade genética e varietal dos recursos cafeeiros nas ilhas de Guadalupe e Haiti.

Na literatura científica, o único relato de genotipagem TaqMan em café foi no estudo de Tao et al. (2013), que desenvolveu um marcador SNPs baseado no gene *GAPDH* para detecção específica de *C. kahawae*, agente causal da antracnose do cafeeiro, demonstrando alta

especificidade para diagnóstico e monitoramento da doença. Também é reportada uma única aplicação de genotipagem Tetra-Primer ARMS-PCR em café, validando 10 SNPs para características de qualidade de bebida (açúcares, cafeína, ácidos clorogênicos), tolerância hídrica e produção de grãos (Schincariol, 2011).

Pesquisas empregando genotipagem por AS-PCR desenvolveram marcadores SNPs para resistência a doenças. Almeida et al. (2024) desenvolveram o marcador CaRHv10-AS, demonstrando mais de 97% de eficácia na identificação de genótipos de *C. arabica* resistentes a múltiplas raças de *H. vastatrix*, agente causal da ferrugem. Ariyoshi, Sera, et al. (2022) validaram em populações F₂ e de retrocruzamento o marcador SNPs Psg_QL1 para resistência a *Pseudomonas syringae* pv. *Garcae*, causador da mancha aureolada.

Considerando os eixos metodológicos abordados nesta pesquisa, observa-se uma disparidade crítica entre a identificação de SNPs em estudos de GWS/GWAS e sua subsequente validação funcional. Enquanto centenas de SNPs são reportados como associados a características agronômicas de interesse, apenas uma fração mínima é efetivamente convertida em marcadores funcionais aplicáveis à seleção assistida por marcadores (SAM) por meio de metodologias de baixo rendimento (como PCR/qPCR), conforme evidenciado pelas 11 publicações identificadas nessa etapa de validação. Essa lacuna técnica resulta de desafios intrínsecos ao processo de validação: muitos marcadores candidatos apresentam comportamento monomórfico quando testados experimentalmente, o que inviabiliza sua aplicação prática. Além disso, mesmo os marcadores que se mostram polimórficos demandam extensiva validação estatística em distintas populações para confirmar sua associação fenotípica e a porcentagem de variação explicada, acrescentando complexidade e tempo ao processo de desenvolvimento de ferramentas moleculares robustas.

A predominância de estudos em *C. arabica* (representando mais de 70% dos manuscritos analisados) sobre *C. canephora* está diretamente associada à sua maior relevância econômica global, responsável por aproximadamente 60% da produção mundial de café (ICO, 2023, 2025a). Além disso, a maior sensibilidade de *C. arabica* a estresses bióticos e abióticos justifica o foco da pesquisa em estratégias genômicas para mitigar essas vulnerabilidades. Contudo, *C. canephora* apresenta características estratégicas para o futuro da cafeicultura ligados a rusticidade frente a diferentes ambientes, diversidade genética e um mercado em expansão impulsionado pela indústria de solúveis e *blends*. Espécies silvestres como *C. liberica* e *C. eugenoides* também representam oportunidades adicionais para estudos comparativos de genômica funcional, hibridação interespecífica e a introgressão de locos de interesse.

No panorama geopolítico da pesquisa, a liderança dos Estados Unidos e Brasil na produção científica reflete papéis estratégicos no mercado global do café. Os Estados Unidos, como maior importador mundial e o Brasil, como maior produtor e exportador, investem em pesquisas para garantia de qualidade e segurança alimentar. Essa sinergia é estratégica, mas requer maior integração com países produtores da África e América Central, onde a diversidade silvestre de *Coffea* oferece recursos genéticos ainda inexplorados (Davis et al., 2011, 2022; Davis & Rakotonasolo, 2021).

O sequenciamento do genoma referência de *C. arabica* e *C. canephora* foi concluído, mas novas montagens e anotações continuam refinando esses genomas (Denoeud et al., 2014; Salojärvi et al., 2024). Essas informações proporcionam avanços significativos na genômica do cafeeiro, permitindo a identificação mais precisa de genes envolvidos em diversas características agronômicas de interesse. Esses dados genômicos viabilizaram o desenvolvimento de painéis de SNPs de alta densidade, a construção de mapas genéticos mais robustos, e a implementação de estratégias de seleção assistida por marcadores e predição genômica. (Combes et al., 2013, 2023; Denoeud et al., 2014; Herrera et al., 2012; Lashermes et al., 2016; Salojärvi et al., 2024). Paralelamente, o desenvolvimento de algoritmos computacionais e modelos estatísticos avançados tem aumentado a acurácia das análises preditivas, ampliando a aplicabilidade de GWS e GWAS no melhoramento do cafeeiro (Arcanjo et al., 2024; Paixão et al., 2022; Sousa et al., 2024; Suela et al., 2023).

1.5 Conclusão

Este estudo bibliométrico identificou um elevado número de SNPs descritos em 20 manuscritos por meio de abordagens de varredura genômica ampla em cafeeiros. No entanto, observa-se uma defasagem significativa entre a descoberta inicial desses marcadores e sua validação funcional para aplicação prática por meio de metodologias *low throughput*, conforme relatado em apenas 11 publicações. O avanço das plataformas de genotipagem baseadas em PCR (HRM, KASP, TaqMan, AS-PCR) tem contribuído para reduzir essa lacuna, ainda que de forma incipiente.

As aplicações *low throughput* reportadas concentram-se principalmente em cruzamentos interespecíficos envolvendo *C. arabica* e *C. canephora*, voltados ao melhoramento genético populacional. Paralelamente, nota-se a necessidade de aprimoramento na fenotipagem de precisão, em redes de ensaios multiambientais, e no desenvolvimento de modelos estatísticos e ferramentas bioinformáticas específicas para a genética das espécies de café.

Um viés geopolítico também é evidente: a produção científica concentra-se predominantemente nos maiores países produtores e consumidores de café, enquanto os países africanos – reconhecidos como centro de origem e principal reservatório de diversidade genética do gênero *Coffea* – permanecem sub-representados, mesmo diante de seu potencial estratégico para o desenvolvimento de marcadores moleculares.

Diante desse cenário, esforços futuros devem priorizar a integração de pipelines que unam descoberta, validação e aplicação de SNPs, além do fortalecimento de colaborações internacionais – especialmente com países detentores de ampla diversidade genética do gênero. A consolidação dessas estratégias poderá acelerar o desenvolvimento de cultivares mais produtivas, resilientes e alinhadas às demandas do mercado global de café.

1.6 Referências

- Akpertey, A., Padi, F. K., Meinhardt, L., & Zhang, D. (2021). Effectiveness of single nucleotide polymorphism markers in genotyping germplasm collections of *Coffea canephora* using KASP assay. *Frontiers in Plant Science*, *11*, 612593. <https://doi.org/10.3389/fpls.2020.612593>
- Alkimim, E. R., Caixeta, E. T., Sousa, T. V., Da Silva, F. L., Sakiyama, N. S., & Zambolim, L. (2018). High-throughput targeted genotyping using next-generation sequencing applied in *Coffea canephora* breeding. *Euphytica*, *214*(3), 1–18. <https://doi.org/10.1007/s10681-018-2126-2>
- 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), 6. <https://doi.org/10.1007/s11032-016-0609-1>
- Alkimim, E. R., Caixeta, E. T., Sousa, T. V., Resende, M. D. V., da Silva, F. L., Sakiyama, N. S., & Zambolim, L. (2020). Selective efficiency of genome-wide selection in *Coffea canephora* breeding. *Tree Genetics & Genomes*, *16*(3), 41. <https://doi.org/10.1007/s11295-020-01433-3>
- Almeida, D. P., Caixeta, E. T., Moreira, K. F., Oliveira, A. C. B. de, Freitas, K. N. P. de, Pereira, A. A., Rosado, R. D. S., Zambolim, L., & Cruz, C. D. (2021). Marker-assisted pyramiding of multiple disease resistance genes in coffee genotypes (*Coffea arabica*). *Agronomy* *2021*, Vol. 11, Page 1763, *11*(9), 1763. <https://doi.org/10.3390/agronomy11091763>
- Almeida, D. P. de, Carneiro, D. G., Caixeta, E. T., Alves, D. R., & Zambolim, L. (2024). New and efficient AS-PCR molecular marker for selection of coffee resistant to coffee leaf rust. *Crop Breeding and Applied Biotechnology*, *24*(4), e498424410. <https://doi.org/10.1590/1984-70332024v24n4a47>
- Almeida, D. P. de, Castro, I. S. L., Mendes, T. A. de O., Alves, D. R., Barka, G. D., Barreiros, P. R. R. M., Zambolim, L., Sakiyama, N. S., & Caixeta, E. T. (2021). Receptor-like kinase (Rlk) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. *Scientia Agricola*, *78*(6), 1–9. <https://doi.org/10.1590/1678-992x-2020-0023>
- Alvarenga, S. M., Caixeta, E. T., Hufnagel, B., Thiebaut, F., Maciel-Zambolim, E., Zambolim, L., & Sakiyama, N. S. (2011). Molecular markers from coffee genome expressed sequences potentially involved in resistance to rust. *Pesquisa Agropecuária Brasileira*, *46*(8), 890–898. <https://doi.org/10.1590/s0100-204x2011000800015>
- Alyethodi, R. R., Singh, U., Kumar, S., Alex, R., Deb, R., Sengar, G. S., Raja, T. V., & Prakash, B. (2018). T-ARMS PCR genotyping of SNP rs445709131 using thermostable strand displacement polymerase. *BMC Research Notes*, *11*(1), 1–5. <https://doi.org/10.1186/S13104-018-3236-6>

- Amangeldiyeva, A., Baidyussen, A., Kuzbakova, M., Yerzhebayeva, R., Jatayev, S., & Shavrukov, Y. (2023). Modified allele-specific qPCR (ASQ) genotyping. *Methods in Molecular Biology*, 2638, 231–247. https://doi.org/10.1007/978-1-0716-3024-2_16
- Arcanjo, E. S., Nascimento, M., Azevedo, C. F., Caixeta, E. T., de Oliveira, A. C. B., Pereira, A. A., & Nascimento, A. C. C. (2024). Low-density marker panels for genomic prediction in *Coffea arabica* L. *Acta Scientiarum. Agronomy*, 47(1), e69698. <https://doi.org/10.4025/actasciagron.v47i1.69698>
- Ariyoshi, C., Sant’ana, G. C., Felicio, M. S., Sera, G. H., Nogueira, L. M., Rodrigues, L. M. R., Ferreira, R. V., Silva, B. S. R. da, Resende, M. L. V. de, Destéfano, S. A. L., Domingues, D. S., & Pereira, L. F. P. (2022). Genome-wide association study for resistance to *Pseudomonas syringae* pv. *garcae* in *Coffea arabica*. *Frontiers in Plant Science*, 13, 989847. <https://doi.org/10.3389/fpls.2022.989847>
- Ariyoshi, C., Sera, G. H., Rodrigues, L. M. R., Carvalho, F. G., Shigueoka, L. H., Mendonça, A. E. S., Pereira, C. T. M., Destéfano, S. A. L., & Pereira, L. F. P. (2022). Development and validation of an allele-specific marker for resistance to bacterial halo blight in *Coffea arabica*. *Agronomy* 2022, Vol. 12, Page 3178, 12(12), 3178. <https://doi.org/10.3390/agronomy12123178>
- Bawin, Y., Ruttink, T., Staelens, A., Haegeman, A., Stoffelen, P., Mwanga Mwanga, J. C. I., Roldán-Ruiz, I., Honnay, O., & Janssens, S. B. (2021). Phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica*. *Journal of Systematics and Evolution*, 59(5), 953–963. <https://doi.org/10.1111/JSE.12694>
- Bikila, B. A. (2015). *Estimation of genetic parameters and SNPs based molecular diversity of Coffea canephora* [Universidade Federal de Viçosa]. <http://www.sbicafe.ufv.br:80/handle/123456789/8395>
- Bolívar-González, A., Molina-Bravo, R., Solano-Sánchez, W., Araya-Valverde, E., Ivamoto-Suzuki, S. T., Pereira, L. F. P., & Gatica-Arias, A. (2022). SNP markers found in non-coding regions can distinguish among low-variant genotypes of arabica and other coffee species. *Genetic Resources and Crop Evolution*, 70(4), 1215–1228. <https://doi.org/10.1007/S10722-022-01498-0>
- Brenes, E. G., Zambolim, L., Almeida, D. P. de, Ribeiro, P. M., Mariz, B. L., & Caixeta, E. T. (2025). Assisted stacking of fungal disease resistance genes in central american coffee cultivars. *Agronomy*, 15(1), 230. <https://doi.org/10.3390/agronomy15010230/s1>
- Breyer, J. P., & Smith, J. R. (2020). Practical genotyping by single-nucleotide primer extension. *Biology Methods and Protocols*, 5(1). <https://doi.org/10.1093/biomethods/bpaa002>
- Campos, M. D., Campos, C., & Cardoso, H. (2023). TaqMan probes for plant species identification and quantification in food and feed traceability. *Methods in Molecular Biology*, 2638, 301–314. https://doi.org/10.1007/978-1-0716-3024-2_21

- Clarindo, W. R., & Carvalho, C. R. (2008). First *Coffea arabica* karyogram showing that this species is a true allotetraploid. *Plant Systematics and Evolution*, 274(3–4), 237–241. <https://doi.org/10.1007/s00606-008-0050-y>
- 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(1), 73–81. <https://doi.org/10.1007/s00299-008-0621-y>
- Clarindo, W. R., Carvalho, C. R., Caixeta, E. T., & Koehler, A. D. (2013). Following the track of “Híbrido de Timor” origin by cytogenetic and flow cytometry approaches. *Genetic Resources and Crop Evolution*, 60(8), 2253–2259. <https://doi.org/10.1007/s10722-013-9990-3>
- Combes, M. C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G., & Lashermes, P. (2000). Characterization of microsatellite loci in *Coffea arabica* and related coffee species. *Molecular Ecology*, 9(8), 1178–1180. <https://doi.org/10.1046/j.1365-294x.2000.00954-5.x>
- Combes, M. C., Dereeper, A., Severac, D., Bertrand, B., & Lashermes, P. (2013). Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid *Coffea arabica* grown at contrasted temperatures. *New Phytologist*, 200(1), 251–260. <https://doi.org/10.1111/nph.12371>
- Combes, M. C., Joët, T., & Lashermes, P. (2018). Development of a rapid and efficient DNA-based method to detect and quantify adulterations in coffee (Arabica versus Robusta). *Food Control*, 88, 198–206. <https://doi.org/10.1016/j.foodcont.2018.01.014>
- Combes, M. C., Joët, T., Stavriniades, A. K., & Lashermes, P. (2023). New cup out of old coffee: contribution of parental gene expression legacy to phenotypic novelty in coffee beans of the allopolyploid *Coffea arabica* L. *Annals of Botany*, 131(1), 157–170. <https://doi.org/10.1093/aob/mcac041>
- Davis, A. P., Kiwuka, C., Faruk, A., Walubiri, M. J., & Kalema, J. (2022). The re-emergence of Liberica coffee as a major crop plant. *Nature Plants*, 8(12), 1322–1328. <https://doi.org/10.1038/S41477-022-01309-5>
- Davis, A. P., & Rakotonasolo, F. (2021). Six new species of coffee (*Coffea*) from northern Madagascar. *Kew Bulletin*, 76(3), 497–511. <https://doi.org/10.1007/s12225-021-09952-5>
- Davis, A. P., Tosh, J., Ruch, N., & Fay, M. F. (2011). Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Botanical Journal of the Linnean Society*, 167(4), 357–377. <https://doi.org/10.1111/j.1095-8339.2011.01177.x>
- Delomas, T. A., Willis, S. C., Parker, B. L., Miller, D., Anders, P., Schreier, A., & Narum, S. (2021). Genotyping single nucleotide polymorphisms and inferring ploidy by amplicon sequencing for polyploid, ploidy-variable organisms. *Molecular Ecology Resources*, 21(7), 2288–2298. <https://doi.org/10.1111/1755-0998.13431>

- Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., Zheng, C., Alberti, A., Anthony, F., Aprea, G., Aury, J. M., Bento, P., Bernard, M., Bocs, S., Campa, C., Cenci, A., Combes, M. C., Cruzillat, D., Da Silva, C., ... Lashermes, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*, *345*(6201), 1181–1184. <https://doi.org/10.1126/science.1255274>
- Dmitriev, A. A., Pushkova, E. N., & Melnikova, N. V. (2022). Plant genome sequencing: modern technologies and novel opportunities for breeding. *Molecular Biology 2022* *56*:4, *56*(4), 495–507. <https://doi.org/10.1134/S0026893322040045>
- Eck, N. J. Van, & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. <https://doi.org/10.1007/S11192-009-0146-3>
- Feitosa, F. de M., Santos, I. G. dos, 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. <https://doi.org/10.1002/CSC2.21303>
- Ferrão, M. A. G., da Fonseca, A. F. A., Volpi, P. S., de Souza, L. C., Comério, M., Filho, A. C. V., Riva-Souza, E. M., Munoz, P. R., Ferrão, R. G., & Ferrão, L. F. V. (2023). Genomic-assisted breeding for climate-smart coffee. *Plant Genome*, *17*(1), e20321. <https://doi.org/10.1002/TPG2.20321>
- Ferrão, L. F. V., Caixeta, E. T., Souza, F. de F., Zambolim, E. M., Cruz, C. D., Zambolim, L., & Sakiyama, N. S. (2013). Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. *Plant Systematics and Evolution*, *299*(1), 225–238. <https://doi.org/10.1007/S00606-012-0717-2>
- Filippis, L. F. De. (2017). Plant Bioinformatics: Next generation sequencing approaches. *plant bioinformatics: decoding the Phyta*, 1–106. https://doi.org/10.1007/978-3-319-67156-7_1
- Gatica-Arias, A., Bolívar-González, A., Sánchez-Barrantes, E., Araya-Valverde, E., & Molina-Bravo, R. (2023). High Resolution Melt (HRM) Genotyping for detection of induced mutations in coffee (*Coffea arabica* L. var. Catuai). In *Mutation Breeding in Coffee with Special Reference to Leaf Rust: Protocols* (pp. 275–290). Springer Berlin Heidelberg Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-67273-0_20
- Gaudet, M., Fara, A. G., Beritognolo, I., & Sabatti, M. (2009). Allele-specific PCR in SNP genotyping. *Methods in Molecular Biology (Clifton, N.J.)*, *578*, 415–424. https://doi.org/10.1007/978-1-60327-411-1_26
- Gichuru, E. K., Agwanda, C. O., Combes, M. C., Mutitu, E. W., Ngugi, E. C. K., Bertrand, B., & Lashermes, P. (2008). Identification of molecular markers linked to a gene conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica*. *Plant Pathology*, *57*(6), 1117–1124. <https://doi.org/10.1111/j.1365-3059.2008.01846.x>
- Gimase, J. M., Thagana, W. M., Omondi, C. O., Cheserek, J. J., Gichimu, B. M., Gichuru, E. K., Ziyomo, C., & Sneller, C. H. (2020). Genome-wide association study identify the

- genetic loci conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica* var. Rume Sudan. *Euphytica*, 216(6), 1–17. <https://doi.org/10.1007/s10681-020-02621-x>
- Gonzales, R. V., Caixeta, E. T., Salgado, S. M. de L., Alves, D. R., dos Reis Fatobene, B. J., Silva, E. M. de A., Correia, A. M., Nascimento, M., Oliveira, A. C. B. de, & Pereira, A. A. (2025). Genome-wide association studies of Amphillo coffee resistance to *Meloidogyne paranaensis*. *European Journal of Plant Pathology* 2025, 1–22. <https://doi.org/10.1007/S10658-025-03025-0>
- Grover, A., & Sharma, P. C. (2016). Development and use of molecular markers: Past and present. *Critical Reviews in Biotechnology*, 36(2), 290–302. <https://doi.org/10.3109/07388551.2014.959891>
- He, Q., Chen, M., Lin, X., & Chen, Z. (2020). Allele-specific PCR with a novel data processing method based on difference value for single nucleotide polymorphism genotyping of ALDH2 gene. *Talanta*, 220. <https://doi.org/10.1016/j.talanta.2020.121432>
- He, Q., Hu, O., Chen, M., Liang, Z., Liang, L., & Chen, Z. (2022). A novel and cost-efficient allele-specific PCR method for multiple SNP genotyping in a single run. *Analytica Chimica Acta*, 1229. <https://doi.org/10.1016/j.aca.2022.340366>
- Herrera, J. C., Romero, J. V., Camayo, G. C., Caetano, C. M., & Cortina, H. A. (2012). Evidence of intergenomic relationships in triploid hybrids of coffee (*Coffea* sp.) as revealed by meiotic behavior and genomic in situ hybridization. *Tropical Plant Biology*, 5(3), 207–217. <https://doi.org/10.1007/s12042-012-9105-x>
- Hicks, L. K., Fili, S., & Walker, E. L. (2023). Rapid genotyping of plant samples using TaqMan™ assays. *Research Square*, 1, 1-15. <https://doi.org/10.21203/rs.3.rs-2670154/v1>
- Huang, L., Wang, X., Dong, Y., Long, Y., Hao, C., Yan, L., & Shi, T. (2020). Resequencing 93 accessions of coffee unveils independent and parallel selection during *Coffea* species divergence. *Plant Molecular Biology*, 103(1–2), 51–61. <https://doi.org/10.1007/s11103-020-00974-4>
- ICO. (2023). Coffee development report. *International Coffee Organization*, 1, 1–118. <https://www.icocoffee.org/documents/cy2024-25/coffee-development-report-2022-3.pdf>
- ICO. (2025). Coffee market report - March 2025. *International Coffee Organization*, 1, 1–14. <https://ico.org/resources/coffee-market-report-statistics-section/>
- Jian, Y., & Li, M. (2021). A narrative review of single-nucleotide polymorphism detection methods and their application in studies of *Staphylococcus aureus*. *Journal of Bio-X Research*, 4(1), 1–9. <https://doi.org/10.1097/jbr.0000000000000071>
- Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, 9(1), 29. <https://doi.org/10.1186/1746-4811-9-29>

- Luis, S. C. J., Paulino, P. R., Bello-Bello, J. J., Esteban, E. P., Heber, A. R. V., Tarsicio, C. T., Gabino, G. de los S., & Victorino, M. R. (2022). SNP markers identification by genome wide association study for chemical quality traits of coffee (*Coffea* spp.) Germplasm. *Molecular Biology Reports*, *49*(6), 4849–4859. <https://doi.org/10.1007/S11033-022-07339-8>
- Kalendar, R., Baidyussen, A., Serikbay, D., Zotova, L., Khassanova, G., Kuzbakova, M., Jatayev, S., Hu, Y. G., Schramm, C., Anderson, P. A., Jenkins, C. L. D., Soole, K. L., & Shavrukov, Y. (2022). Modified “Allele-Specific qPCR” method for SNP genotyping based on FRET. *Frontiers in Plant Science*, *12*. <https://doi.org/10.3389/FPLS.2021.747886>
- Kalendar, R., Shustov, A. V., Akhmetollayev, I., & Kairov, U. (2022). Designing allele-specific competitive-extension PCR-based assays for high-throughput genotyping and gene characterization. *Frontiers in Molecular Biosciences*, *9*. <https://doi.org/10.3389/FMOLB.2022.773956>
- Kim, N., Kwon, J. S., Kang, W. H., & Yeom, S. I. (2023). High-Resolution Melting (HRM) genotyping. *Methods in Molecular Biology (Clifton, N.J.)*, *2638*, 337–349. https://doi.org/10.1007/978-1-0716-3024-2_24
- Kockum, I., Huang, J., & Stridh, P. (2023). Overview of genotyping technologies and methods. *Current Protocols*, *3*(4), e727. <https://doi.org/10.1002/cpz1.727>
- Kumar, R., Das, S. P., Choudhury, B. U., Kumar, A., Prakash, N. R., Verma, R., Chakraborti, M., Devi, A. G., Bhattacharjee, B., Das, R., Das, B., Devi, H. L., Das, B., Rawat, S., & Mishra, V. K. (2024). Advances in genomic tools for plant breeding: harnessing DNA molecular markers, genomic selection, and genome editing. *Biological Research 2024 57:1*, *57*(1), 1–23. <https://doi.org/10.1186/S40659-024-00562-6>
- Lashermes, P., Hueber, Y., Combes, M.-C., Severac, D., & Dereeper, 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. <https://doi.org/10.1534/g3.116.030858>
- Lawrie, R. D., & Massey, S. E. (2023). Agrigenomic diversity unleashed: current single nucleotide polymorphism genotyping methods for the agricultural sciences. *Applied Biosciences 2023*, *Vol. 2*, *Pages 565-585*, *2*(4), 565–585. <https://doi.org/10.3390/applbiosci2040036>
- Lefever, S., Rihani, A., Van der Meulen, J., Pattyn, F., Van Maerken, T., Van Dorpe, J., Hellemans, J., & Vandesompele, J. (2019). Cost-effective and robust genotyping using double-mismatch allele-specific quantitative PCR. *Scientific Reports*, *9*(1), 1–11. <https://doi.org/10.1038/s41598-019-38581-z>

- Liu, J., Huang, S., Sun, M., Liu, S., Liu, Y., Wang, W., Zhang, X., Wang, H., & Hua, W. (2012). An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods*, 8(1), 1–9. <https://doi.org/10.1186/1746-4811-8-34>
- Lu, M., Li, J., Sun, X., Zhao, D., Zong, H., Tang, C., Li, K., Zhou, Y., & Xiao, J. (2024). Genotyping single nucleotide polymorphisms in homologous regions using multiplex kb level amplicon capture sequencing. *Molecular Genetics and Genomics*, 299(1), 1–12. <https://doi.org/10.1007/S00438-024-02192-9>
- Mahé, L., Combes, M. C., Várzea, V. M. P., Guilhaumon, C., & Lashermes, P. (2008). Development of sequence characterized DNA markers linked to leaf rust (*Hemileia vastatrix*) resistance in coffee (*Coffea arabica* L.). *Molecular Breeding*, 21(1), 105–113. <https://doi.org/10.1007/s11032-007-9112-z>
- Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, 2012, 728398. <https://doi.org/10.1155/2012/728398>
- Marth, G. T., Korf, I., Yandell, M. D., Yeh, R. T., Gu, Z., Zakeri, H., Stitzel, N. O., Hillier, L. D., Kwok, P. Y., & Gish, W. R. (1999). A general approach to single-nucleotide polymorphism discovery. *Nature Genetics*, 23(4), 452–456. <https://doi.org/10.1038/70570>
- Medeiros, A. C., Caixeta, E. T., Oliveira, A. C. B. de, Sousa, T. V., Stock, V. de M., Cruz, C. D., Zambolim, L., & 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. <https://doi.org/10.3389/fsufs.2021.705278>
- Medrano, R. F. V., & Oliveira, C. A. (2014). Guidelines for the Tetra-Primer ARMS-PCR technique development. *Molecular Biotechnology*. <https://doi.org/10.1007/s12033-014-9734-4>
- Mekbib, Y., Tesfaye, K., Dong, X., Saina, J. K., Hu, G.-W., & Wang, Q.-F. (2022). Whole-genome resequencing of *Coffea arabica* L. (Rubiaceae) genotypes identify SNP and unravels distinct groups showing a strong geographical pattern. *BMC Plant Biology*, 22(1), 69. <https://doi.org/10.1186/s12870-022-03449-4>
- Merot-L'anthoene, V., Tournebize, R., Darracq, O., Rattina, V., Lepelley, M., Bellanger, L., Tranchant-Dubreuil, C., Coulée, M., Pégard, M., Metairon, S., Fournier, C., Stoffelen, P., Janssens, S. B., Kiwuka, C., Musoli, P., Sumirat, U., Legnaté, H., Kambale, J., Ferreira da Costa Neto, J., ... Poncet, V. (2019). Development and evaluation of a genome-wide Coffee 8.5K 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. <https://doi.org/10.1111/pbi.13066>

- Meuwissen, T. H., Hayes, B. J., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819-1829. <https://doi.org/10.1093/genetics/157.4.1819>
- Millet, C. P., Allinne, C., Vi, T., Marraccini, P., Verleysen, L., Couderc, M., Ruttink, T., Zhang, D., Sánchez, W. S., Tranchant-Dubreuil, C., Jeune, W., & Poncet, V. (2024). Haitian coffee agroforestry systems harbor complex arabica variety mixtures and under-recognized genetic diversity. *Plos One*, *19*(4), e0299493. <https://doi.org/10.1371/journal.pone.0299493> (a)
- Millet, C. P., Delahaie, B., Georget, F., Allinne, C., Solano-Sánchez, W., Zhang, D., Jeune, W., Toniutti, L., & Poncet, V. (2024). Guadeloupe and Haiti's coffee genetic resources reflect the crop's regional and global history. *Plants People Planet*, *7*(1), 245–262. <https://doi.org/10.1002/ppp3.10584> (b)
- Mishra, M. K. (2020). Genetic resources and breeding of coffee (*Coffea* spp.). *Advances in plant breeding strategies: Nut and beverage crops*, *4*, 475–515. https://doi.org/10.1007/978-3-030-23112-5_12
- Mishra, M. K., Jingade, P., Huded, A. K. C., & Muniswamy, B. (2023). Assessment of genetic homogeneity of somatic embryo-derived plants and seed-derived plants of a robusta coffee cultivar using molecular markers and functional genes sequencing. *Plant Cell, Tissue and Organ Culture*, *153*(2), 319–332. <https://doi.org/10.1007/S11240-023-02468-9>
- 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), 5. <https://doi.org/10.1007/s11295-015-0927-1>
- Montagnon, C., Mahyoub, A., Solano, W., & Sheibani, F. (2021). Unveiling a unique genetic diversity of cultivated *Coffea arabica* L. in its main domestication center: Yemen. *Genetic Resources and Crop Evolution*, *68*(6), 2411–2422. <https://doi.org/10.1007/s10722-021-01139-y>
- Morais, T. P. de, & Melo, B. de. (2011). Biotecnologia aplicada ao melhoramento genético do cafeeiro. *Ciencia Rural*, *41*(5), 753–760. <https://doi.org/10.1590/s0103-84782011005000056>
- Morgil, H., Gercek, Y. C., Tulum, I., Morgil, H., Gercek, Y. C., & Tulum, I. (2020). Single nucleotide polymorphisms (SNPs) in plant genetics and breeding. *The Recent Topics in Genetic Polymorphisms*. <https://doi.org/10.5772/intechopen.91886>
- Muniz, H. V. L., Ariyoshi, C., Ferreira, R. V., Felicio, M. S., Filipe, L., Pereira, P., Filho, M., Celso, A., & Zanin, W. (2024). A new set of quantitative trait loci linked to lipid content in *Coffea arabica*. *Crop Breeding and Applied Biotechnology*, *24*(2), e478824212. <https://doi.org/10.1590/1984-70332024v24n2a25>

- Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doğan, Y., Comertpay, G., Yıldız, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N., Özkan, H., Chung, G., & Baloch, F. S. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology and Biotechnological Equipment*, *32*(2), 261–285. <https://doi.org/10.1080/13102818.2017.1400401>
- Nonato, J. V. A., Carvalho, H. F., Borges, K. L. R., Padilha, L., Maluf, M. P., Fritsche-Neto, R., & Guerreiro Filho, O. (2021). Association mapping reveals genomic regions associated with bienniality and resistance to biotic stresses in arabica coffee. *Euphytica*, *217*(10), 190. <https://doi.org/10.1007/s10681-021-02922-9>
- Nong, T. H., Phan, T. H. T., Nguyen, T. M., Dinh, T. T. O., Nguyen, T. H., Pham, T. T. H., Nguyen, T. T. L., & Ninh, T. T. (2025). Genetic diversity of *Coffea arabica* accessions using RAPD and DAMD molecular markers. *New Zealand Journal of Crop and Horticultural Science*. <https://doi.org/10.1080/01140671.2025.2454606>
- Ortiz, A. J., & Sharbrough, J. (2024). Genome-wide patterns of homoeologous gene flow in allotetraploid coffee. *Applications in Plant Sciences*, *12*(4), e11584. <https://doi.org/10.1002/APS3.11584>
- Paixão, P. T. M., Nascimento, A. C. C., Nascimento, M., Azevedo, C. F., Oliveira, G. F., da Silva, F. L., & Caixeta, E. T. (2022). Factor analysis applied in genomic selection studies in the breeding of *Coffea canephora*. *Euphytica*, *218*(4), 1–9. <https://doi.org/10.1007/S10681-022-02998-x>
- 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*(6), 1311–1317. <https://doi.org/10.1007/s00122-004-1748-z>
- 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. <https://doi.org/10.1111/pbr.12127>
- Saavedra, L. M., Caixeta, E. T., Barka, G. D., Borém, A., Zambolim, L., Nascimento, M., Cruz, C. D., Oliveira, A. C. B. de, & 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. <https://doi.org/10.3390/genes14010189>
- Sahoo, J., Mishra, R., & Joshi, R. K. (2025). PCR-based single nucleotide polymorphism (SNP) genotyping for crop improvement-current status and future prospects. *Discover Plants 2025 2:1*, *2*(1), 1–19. <https://doi.org/10.1007/S44372-025-00262-9>
- Salojärvi, J., Rambani, A., Yu, Z., Guyot, R., Strickler, S., Lepelley, M., Wang, C., Rajaraman, S., Rastas, P., Zheng, C., Muñoz, D. S., Meidanis, J., Paschoal, A. R., Bawin, Y.,

- Krabbenhoft, T. J., Wang, Z. Q., Fleck, S. J., Aussel, R., Bellanger, L., ... 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. <https://doi.org/10.1038/s41588-024-01695-w>
- Sant’Ana, G. C., Pereira, L. F. P., Pot, D., Ivamoto, S. T., Domingues, D. S., Ferreira, R. V., Pagiatto, N. F., Da Silva, B. S. R., Nogueira, L. M., Kitzberger, C. S. G., Scholz, M. B. S., De Oliveira, F. F., Sera, G. H., Padilha, L., Labouisse, J. P., Guyot, R., Charmetant, P., & Leroy, T. (2018). Genome-wide association study reveals candidate genes influencing lipids and diterpenes contents in *Coffea arabica* L. *Scientific Reports*, *8*(1), 1–12. <https://doi.org/10.1038/S41598-017-18800-1>
- Santos, T. B. dos, Silva Ferreira, M. F. da, Marques, I., Oliveira, S. C., Zaidan, I. R., Oliveira, M. G., Rodrigues, W. P., Ribas, A. F., Guyot, R., Ramalho, J. C., Ribeiro-Barros, A. I., Pereira, L. F. P., & Partelli, F. L. (2022). Current challenges and genomic advances towards the development resilient coffee genotypes to abiotic stresses. *Genomic Designing for Abiotic Stress Resistant Technical Crops*, 41–69. https://doi.org/10.1007/978-3-031-05706-9_3
- 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. *Plants*, *255*(6), 1–16. <https://doi.org/10.1007/S00425-022-03898-Z>
- Scalabrin, S., Magris, G., Liva, M., Vitulo, N., Vidotto, M., Scaglione, D., Del Terra, L., Ruosi, M. R., Navarini, L., Pellegrino, G., Berny Mier y Teran, J. C., Toniutti, L., Suggi Liverani, F., Cerutti, M., Di Gaspero, G., & Morgante, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, *15*(1), 463. <https://doi.org/10.1038/s41467-023-44449-8>
- Scalabrin, S., Toniutti, L., Di Gaspero, G., Scaglione, D., Magris, G., Vidotto, M., Pinosio, S., Cattonaro, F., Magni, F., Jurman, I., Cerutti, M., Suggi Liverani, F., Navarini, L., Del Terra, L., Pellegrino, G., Ruosi, M. R., Vitulo, N., Valle, G., Pallavicini, A., ... 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), 4642. <https://doi.org/10.1038/s41598-020-61216-7>
- Schatz, M. C., Witkowski, J., & McCombie, W. R. (2012). Current challenges in de novo plant genome sequencing and assembly. *Genome Biology*, *13*(4), 1–7. <https://doi.org/10.1186/gb-2012-13-4-243>
- Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. *Plant Biotechnology Journal*, *15*(2), 149–161. <https://doi.org/10.1111/pbi.12645>
- Schincariol, L. P. (2011). *Identificação de marcadores snp pelas técnicas de PCR-RFLP e TETRA-PRIMER ARMS-PCR e suas associações com qualidade de bebida em café*

[DISSERTAÇÃO]. Instituto Agronômico de Campinas.
https://www.iac.sp.gov.br/areadoinstituto/posgraduacao/repositorio/storage/teses_dissertacoes/pb1213309.pdf

- Semagn, K., Babu, R., Hearne, S., & Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. *Molecular Breeding*, 33(1), 1–14. <https://doi.org/10.1007/s11032-013-9917-x>
- Setotaw, T. A., Caixeta, E. T., Zambolim, E. M., Sousa, T. V., Pereira, A. A., Baião, A. C., Cruz, C. D., Zambolim, L., & Sakiyama, N. S. (2020). Genome Introgression of Híbrido de Timor and its potential to develop high cup quality *C. arabica* cultivars. *Journal of Agricultural Science*, 12(4), 64. <https://doi.org/10.5539/jas.v12n4p64>
- Silva, E. M. de A., Castro, I. S. L., Aguilar, A. P., Caixeta, E. T., & Mendes, T. A. de O. (2023). New genetic markers for 100% arabica coffee demonstrate high discriminatory potential for InDel-HRM-based coffee authentication. *Food Research International*, 173, 113424. <https://doi.org/10.1016/j.foodres.2023.113424>
- Silva, L., Alkimim, E. R., Barreiro, P. R. R. M., Leichtweis, B. G., Silva, A. C. A., da Silva, R. A., Sousa, T. V., Nascimento, M., & Caixeta, E. T. (2022). Genome-wide association study of plant architecture and diseases resistance in *Coffea canephora*. *Euphytica*, 218(7), 92. <https://doi.org/10.1007/s10681-022-03042-8>
- Silva, R. A., Caixeta, E. T., Silva, L. de F., Sousa, T. V., Barreiros, P. R. R. M., Oliveira, A. C. B. de, Pereira, A. A., Barreto, C. A. V., & Nascimento, M. (2024). Identification of SNP markers and candidate genes associated with major agronomic traits in *Coffea arabica*. *Plants*, 13(13), 1876. <https://doi.org/10.3390/plants13131876>
- Singh, A., Ramakrishna, G., Kaila, T., Saxena, S., Sharma, S., Gaikwad, A. B., Abdin, M. Z., & Gaikwad, K. (2022). Next-generation sequencing technologies: approaches and applications for crop improvement. *Genomics of Cereal Crops* (pp. 31-94). https://doi.org/10.1007/978-1-0716-2533-0_3
- Sousa, I. C. de, Barreto, C. A. V., Caixeta, E. T., Nascimento, A. C. C., Azevedo, C. F., Alkimim, E. R., & Nascimento, M. (2024). The trade-off between density marker panels size and predictive ability of genomic prediction for agronomic traits in *Coffea canephora*. *Euphytica*, 220(4), 1–11. <https://doi.org/10.1007/S10681-024-03303-8>
- Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B., Pereira, A. A., Sakiyama, N. S., Zambolim, L., & Resende, M. D. V. (2019). Early selection enabled by the implementation of genomic selection in *Coffea arabica* breeding. *Frontiers in Plant Science*, 9, 1934. <https://doi.org/10.3389/fpls.2018.01934>
- Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B. de, Pereira, A. A., Sakiyama, N. S., Resende Júnior, M. F. R. de, & 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 and Genomes*, 13(6), 1–16. <https://doi.org/10.1007/s11295-017-1208-y>
- Spence, J. P., Mostafavi, H., Ota, M., Milind, N., Gjorgjieva, T., Smith, C. J., ... & Pritchard, J. K. (2025). Specificity, length and luck drive gene rankings in association studies. *Nature*, 649, 918–925. <https://doi.org/10.1038/s41586-025-09703-7>
- Suela, M. M., Azevedo, C. F., Nascimento, A. C. C., Momen, M., Oliveira, A. C. B. de, Caixeta, E. T., Morota, G., & Nascimento, M. (2023). Genome-wide association study for morphological, physiological, and productive traits in *Coffea arabica* using structural equation models. *Tree Genetics and Genomes*, 19(3), 1–17. <https://doi.org/10.1007/S11295-023-01597-8>
- Surya Prakash, N., Devasia, J., Raghuramulu, Y., & Aggarwal, R. K. (2016). Genetic diversity and coffee improvement in India. *Molecular Breeding for Sustainable Crop Improvement*, 2 (pp. 231-268). Cham: Springer International Publishing. 231–268. https://doi.org/10.1007/978-3-319-27090-6_11
- Tanha, H. M., Naeini, M. M., Rahgozar, S., Rasa, S. M. M., & Vallian, S. (2015). Modified tetra-primer ARMS PCR as a single-nucleotide polymorphism genotyping tool. *Genetic Testing and Molecular Biomarkers*, 19(3), 156–161. <https://doi.org/10.1089/GTMB.2014.0289>
- Tao, G., Hyde, K. D., & Cai, L. (2013). Species-specific real-time PCR detection of *Colletotrichum kahawae*. *Journal of Applied Microbiology*, 114(3), 828–835. <https://doi.org/10.1111/JAM.12068>
- Tran, H. T. M., Furtado, A., Vargas, C. A. C., Smyth, H., Slade Lee, L., & Henry, R. (2018). SNP in the *Coffea arabica* genome associated with coffee quality. *Tree Genetics and Genomes*, 14(5), 1–15. <https://doi.org/10.1007/S11295-018-1282-9>
- Udoh, L. I., Obaseoji, W. P., & Uzoabo, C. (2021). Single nucleotide polymorphisms: A modern tool to screen plants for desirable traits. *Plant Breeding - Current and Future Views*. <https://doi.org/10.5772/intechopen.94935>
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1), 1–21. <https://doi.org/10.1038/S43586-021-00056-9>
- Wojdacz, T. K., Dobrovic, A., & Hansen, L. L. (2008). Methylation-sensitive high-resolution melting. *Nature Protocols*, 3(12), 1903–1908. <https://doi.org/10.1038/nprot.2008.191>
- Yadav, P., Vaidya, E., Rani, R., Yadav, N. K., Singh, B. K., Rai, P. K., & Singh, D. (2016). Recent perspective of next generation sequencing: applications in molecular plant biology and crop improvement. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences* 2016 88:2, 88(2), 435–449. <https://doi.org/10.1007/S40011-016-0770-7>

- Ye, S., Dhillon, S., Ke, X., Collins, A. R., & Day, I. N. M. (2001). An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Research*, *29*(17), e88–e88. <https://doi.org/10.1093/nar/29.17.e88>
- Zaidan, I. R., Ferreira, M. F. da S., Pereira do Couto, D., Santos, J. G., Silva, M. A., Canal, G. B., Bernardes, C. de O., Azevedo, C. F., & Ferreira, A. (2025). Genome-wide association analysis of traits related to development, abiotic and biotic stress resistance in *Coffea canephora*. *Scientia Horticulturae*, *341*, 114004. <https://doi.org/10.1016/j.scienta.2025.114004>
- Zargar, S. M., Raatz, B., Sonah, H., Muslimanazir, Bhat, J. A., Dar, Z. A., Agrawal, G. K., & Rakwal, R. (2016). Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. *Journal of Crop Science and Biotechnology 2015 18:5*, *18*(5), 293–308. <https://doi.org/10.1007/s12892-015-0037-5>
- Zhang, D., Vega, F. E., Solano, W., Su, F., Infante, F., & Meinhardt, L. W. (2021). Selecting a core set of nuclear SNP markers for molecular characterization of Arabica coffee (*Coffea arabica* L.) genetic resources. *Conservation Genetics Resources*, *13*(3), 329–335. <https://doi.org/10.1007/S12686-021-01201-y>
- Zhou, L., Vega, F. E., Tan, H., Lluch, A. E. R., Meinhardt, L. W., Fang, W., Mischke, S., Irish, B., & Zhang, D. (2016). Developing single nucleotide polymorphism (SNP) markers for the identification of coffee germplasm. *Tropical Plant Biology*, *9*(2), 82–95. <https://doi.org/10.1007/S12042-016-9167-2>

CAPÍTULO 2

DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES MOLECULARES SNP PARA ALTURA DE PLANTAS EM *Coffea arabica*

DEZEMBRO / 2025

VIÇOSA – MG

RESUMO

A seleção precoce de cafeeiros com porte definido é estratégica em programas de melhoramento genético para otimizar a colheita, o manejo, o adensamento e a eficiência no uso de insumos na lavoura. Este estudo objetivou desenvolver marcadores moleculares de polimorfismo de nucleotídeo único (SNP), previamente identificados por Estudo de Associação Genômica Ampla (GWAS), para uso na seleção assistida e para melhor compreensão da característica altura de plantas em *Coffea arabica*. Para isso, utilizou-se uma população de mapeamento para desenvolvimento de marcadores e uma população de testagem para os marcadores desenvolvidos. Foram empregadas as metodologias de genotipagem *low-throughput* dos SNPs: *High Resolution Melting* (HRM), *RNase H-dependent Amplification* (rhAmp), *Tetra-Primer Amplification Refractory Mutation System* (Tetra Primer ARMS) e *Modified Tetra-Primer Amplification Refractory Mutation System* (MTPA). Foram obtidos seis marcadores moleculares utilizando as metodologias HRM e Tetra-Primer ARMS para os SNPs V16099 (HRM-V16099-2, 4P.ARMS-V16099-1), V1190 (HRM-V1190-3, 4P.ARMS-V1190-1), V42 (HRM-V42-2) e V16476 (HRM-V16476-1), demonstrando seu potencial para aplicação em seleção assistida. As demais metodologias testadas, rhAmp e MTPA, não foram eficientes no desenvolvimento de marcadores. Na população de testagem, confirmou-se a capacidade desses marcadores em diferenciar os alelos de interesse; no entanto, a análise de associação de marca simples não detectou efeitos significativos isoladamente sobre a altura das plantas. Esta divergência pode ser atribuída à mudança de *background* genético entre as populações, possíveis alterações no padrão de desequilíbrio de ligação (LD) ou à modulação por efeitos epistáticos, corroborando a natureza poligênica e complexa do caráter. A região genômica dos marcadores V16099 e V1190 revelou 13 genes candidatos envolvidos em processos de morfogênese e controle hormonal (auxina, ácido jasmônico e ácido abscísico), sugerindo uma rede gênica determinante para a arquitetura da planta. Este trabalho reporta os primeiros marcadores moleculares para discriminação de altura em cafeeiros, oferecendo ferramentas potenciais para a seleção assistida em programas de melhoramento genético.

Palavras – Chave: Melhoramento genético; Porte de café; Seleção Assistida por Marcadores Moleculares; Genotipagem HRM; Tetra *Primer* ARMS.

2 Introdução

Os complexos genomas do gênero *Coffea* representam, ao mesmo tempo, desafios e oportunidades para os programas de melhoramento genético. Elucidar sua origem, evolução e diversidade genômica é essencial para explorar seu potencial no desenvolvimento de cultivares de café melhoradas (Huang et al., 2020; Merot-L'anthoene et al., 2019; Salojärvi et al., 2024; Scalabrin et al., 2020, 2024). Dentre as 141 espécies descritas (Guyot et al., 2020), a mais cultivada é *C. arabica* ($2n = 4x = 44$), uma espécie alotetraploide originada de uma hibridação natural entre os ancestrais diploides *C. canephora* (subgenoma C^C) e *C. eugenioides* (subgenoma C^E) (Medrano et al. 2025; Bawin et al. 2021; Lashermes et al. 1999, 2016).

Atualmente, o Brasil possui 127 cultivares de *C. arabica* de domínio público registradas (Brasil, 2025). A maioria das cultivares de *C. arabica* resistentes a doenças apresenta introgressões de *C. canephora* provenientes do cafeeiro Híbrido de Timor (HdT) (Bettencourt et al., 1992; Bettencourt & Rodrigues, 1988; Carvalho et al., 1989; Huang et al., 2020; Mekbib et al., 2022; Montagnon et al., 2021; Setotaw et al., 2020; Silvarolla et al., 1997). Os cafeeiros HdT, originados de cruzamento espontâneo entre *C. canephora* e *C. arabica*, constitui a principal fonte de resistência múltipla a fatores bióticos e abióticos incorporada em progênies de *C. arabica* em escala mundial (Bettencourt, 1973; Bettencourt et al., 1992; Clarindo et al., 2013; Clarindo & Carvalho, 2009; Lashermes et al., 2016; Sattler et al., 2022). Entre os híbridos derivados do HdT utilizados no melhoramento genético, destacam-se as populações oriundas dos cruzamentos de Caturra × HdT e Vila Sarchi × HdT, conhecidas na literatura como Catimor e Sarchimor, respectivamente (Bertrand et al., 1997; Bettencourt & Fazuoli, 2008; Chaves & Zambolin, 1976; Severino et al., 2008; Várzea et al., 2008).

O uso desses híbridos interespecíficos enfrenta um obstáculo: o HdT possui porte elevado de plantas (com relatos de até 9 metros), característica que pode ser indesejável para manejo fitotécnico da lavoura e para realização das colheitas manual, semimecanizada e mecanizada (Bettencourt, 1973; Bonomo et al., 2004; Silvarolla et al., 1997). Quando cruzados com cultivares de *C. arabica* consideradas de porte baixo, os descendentes exibem significativa segregação para a altura das plantas. Como resultado, surgem populações nas quais apenas uma fração dos indivíduos combina a estatura desejada com o perfil de resistências almejado, exigindo assim uma seleção rigorosa ao longo de várias gerações (Brasil, 2024; Lashermes et al., 2016).

A altura da planta, componente central de sua arquitetura, é um caráter de grande relevância agrônômica, com interação entre as características genéticas e o ambiente de cultivo. Ela está

intrinsecamente relacionada a fatores decisivos para a produtividade, como o grau de dominância apical, a alocação de biomassa e a eficiência da captação luminosa (Busov et al., 2008; Ju et al., 2018; Lanfear et al., 2013). Além disso, a altura é um parâmetro fundamental para o planejamento do manejo, influenciando diretamente a definição do espaçamento de plantio, a viabilidade do adensamento e a eficiência da colheita. Nesse contexto, a capacidade de prever a altura das plantas em programas de melhoramento emerge como uma ferramenta estratégica para o desenvolvimento de cultivares adaptadas a diferentes sistemas produtivos na cafeicultura (Carvalho et al., 1979; Carvalho & Monaco, 1972; Dias et al., 2020; Silvarolla et al., 1997).

As principais pesquisas sobre a herança genética da característica altura em café arábica são da década de 1980, desenvolvidas por meio de cruzamentos contrastantes entre cultivares consideradas baixas e altas (Carvalho et al., 1979, 1984b, 1984a; Carvalho & Monaco, 1972; Silvarolla et al., 1997). Carvalho et al. (1984b) sugeriram que o porte reduzido é governado pelos locos *Sb*, *Sr* e *Ct*, presentes nas cultivares São Bernardo, San Ramón e Caturra, respectivamente; e levantaram hipóteses de que outros locos responsáveis pelo porte baixo poderiam estar presentes em cafés como Villa Sarchi, Pacas e Vila Lobos.

Os locos *Sb*, *Sr* e *Ct* segregam independentemente e podem ter combinações diíbridas, com possibilidade de os alelos estarem presentes em ambos os subgenomas (C^C e C^E), resultando em até quatro cópias de um mesmo alelo nos cromossomos homeólogos. A característica porte reduzido é condicionada por genes dominantes em relação ao porte normal. Em cruzamentos entre plantas com alelos para dois locos diferentes, ocorre uma interação epistática do tipo duplo dominante, na qual a presença de um alelo dominante condiciona porte reduzido e máscara o efeito no outro (proporção 15:1 em F_2 , e 3:1 em cruzamento teste) (Carvalho et al., 1984b, 1984a).

Acredita-se que a mutação *Ct* de Caturra tenha surgido no Bourbon Vermelho, dando origem a cultivar Caturra Vermelho (Bonomo et al., 2004; Carvalho & Monaco, 1972). Seu principal efeito fenotípico é a redução do comprimento médio dos internódios dos ramos ortotrópicos e plagiotrópicos, que diminui a altura das plantas, conferindo-lhes aspecto compacto. Linhagens portadoras do gene *Ct*, tendem a ter menor diâmetro de copa em relação às linhagens não portadoras (Carvalho et al., 1984a). As características altura da planta, diâmetro da copa e comprimento dos internódios são diretamente relacionados à expressão desse gene, e guardam importantes correlações entre si (Aguiar et al., 2004; A. Carvalho et al., 1979).

Trabalho mais recente de Carvalho et al. (2022) sugere que a expressão da altura do cafeeiro é modulada pela interação dos alelos *Ct* ou *ct* com genes de efeito secundário, justificando a variação fenotípica dentro das categorias de porte baixo e alto. Sob a perspectiva da caracterização por descritores mínimos, a altura de planta pode ser classificada em *Ct Ct* ou *Ct ct* (muito baixa, baixa e média) e *ct ct* (alta e muito alta). Por essa classificação, as cultivares dos grupos Vila Lobos são consideradas muito baixas; Caturra e Villa Sarchi, baixas (< 2,0 m); Catuaí Amarelo, médias (~2,0 m), Acaiaí, Bourbon e Mundo Novo, altas (> 3,30m) (Aguiar et al., 2004; Brasil. Ministério da Agricultura, 2000; Carvalho et al., 1984a; Pereira et al., 2025).

Entretanto, para fins práticos no melhoramento genético, prevalece uma classificação binária simplificada em porte baixo e porte alto. Nesse sistema, o padrão de porte baixo é representado por cultivares com altura semelhante à do 'Catuaí', derivadas da introgressão de alelos de 'Villa Sarchi' (como Obatã, Tupi, IAPAR 59, IPR 98) ou 'Caturra' (como Catuaí Vermelho, Catuaí Amarelo, Rubi e Topázio). Nas condições do sudeste brasileiro, essas plantas geralmente atingem de 1,8 a 2,2 m de altura quando adultas. O padrão de porte alto, por sua vez, inclui cultivares como 'Mundo Novo', 'Acaiaí' e 'Icatu', que atingem de 2,2 a 3,0 m ou mais. Nesta classificação, as cultivares Bourbon Amarelo e Bourbon Vermelho também são consideradas de porte alto, por serem portadoras dos alelos *ct ct*, embora sejam classificadas como de porte médio na metodologia de descritores mínimos (Carvalho et al., 2022; Fazuoli et al., 2008; Julca-Otiniano et al., 2023; Valarini et al., 2005)

Um recente Estudo de Associação Genômica Ampla (GWAS) conduzido por Silva et al. (2024) identificou um Loco de Característica Quantitativa (QTL) associado à altura de plantas no cromossomo 6 de *C. arabica*. Entre os 49 SNPs significativos associados ao caráter, 42 localizavam-se no subgenoma C^C , enquanto 11 no subgenoma C^E . Os efeitos fenotípicos desses SNPs na redução da altura variaram de 4,61% a 21,21%. Esses SNPs encontram-se em desequilíbrio de ligação e abrangem uma porção extensa do cromossomo, o que sugere que múltiplos genes possam estar envolvidos no controle dessa característica. A hipótese de que a altura é regulada por uma rede complexa de genes de pequeno efeito é reforçada por estudos de Seleção Genômica Ampla (GWS) realizados na mesma população, os quais reportaram baixa herdabilidade genotípica e identificaram 202 QTLs putativos associados à altura das plantas (Sousa et al., 2019).

Com base nessa descoberta, o QTL identificado por Silva et al. (2024) constitui um alvo genômico importante para entendimento dessa característica e para o avanço do melhoramento genético por meio de marcadores moleculares. Diante disso, o presente estudo teve como

objetivo aprofundar o conhecimento sobre a característica altura de plantas em *C. arabica*, mediante ao desenvolvimento e validação molecular de marcadores SNPs localizados neste QTL. O propósito final consistiu em converter esses marcadores moleculares em ferramentas aplicáveis à seleção assistida, para otimizar o tempo necessário no desenvolvimento de novas cultivares de café com porte definido.

2.1 Material e Métodos

2.1.1 Material vegetal

Nesse trabalho utilizou-se uma população de mapeamento para desenvolvimento dos marcadores e uma população de testagem para avaliar os marcadores desenvolvidos.

A população de mapeamento utilizada no GWAS foi empregada neste estudo para identificar marcadores polimórficos e avaliar a precisão na discriminação alélica, uma vez que seus genótipos haviam sido previamente sequenciados, assegurando assim a reprodutibilidade dos resultados de genotipagem. Essa população foi desenvolvida a partir de cruzamentos contrastantes entre genitores de *C. arabica* de porte reduzido do grupo Catuaí Amarelo (UFV 2143-236, UFV 2148-57 e UFV 2154-344) e plantas de porte alto derivadas de Híbrido de Timor (UFV 440-10, UFV 445-46 e CIFC 832/2).

Além dos genitores, incluíram-se nas análises genóticas os híbridos F_1 oriundos desses cruzamentos controlados (H419-1, H419-10, H513-5, H514-7 e H514-8), bem como retrocruzamentos de diferentes combinações (1_1, 3_4, 1_5, 1_3, 9_1 e 2_2). Também foram avaliados três indivíduos segregantes da geração F_2 (10_4, 10_1 e 13_1), totalizando 20 genótipos utilizados. A descrição completa dessa população de mapeamento encontra-se em Sousa et al. (2017).

Para avaliação e validação dos marcadores *low-throughput* em um *background* genético distinto, os SNPs foram analisados em uma população segregante F_2 de 150 genótipos, derivada do cruzamento entre o acesso Híbrido de Timor MG0357 e a linhagem Tupi Amarelo IAC 5162 (Figura 1), aqui designada como população de testagem. Estudos preliminares confirmaram elevada capacidade de combinação deste cruzamento, que resultou em duas plantas F_1 (C12P-8-B20-E5 e C12P-22-B20-E5) (Medeiros et al., 2021). Estas plantas F_1 foram subsequentemente autofecundadas em Viçosa, MG, Brasil (20°44'28,4"S 42°50'53,9"W), gerando a população F_2 utilizada neste estudo. Este *pedigree* populacional combinou genitores com alturas contrastantes, viabilizando a análise da segregação do caráter na progênie.

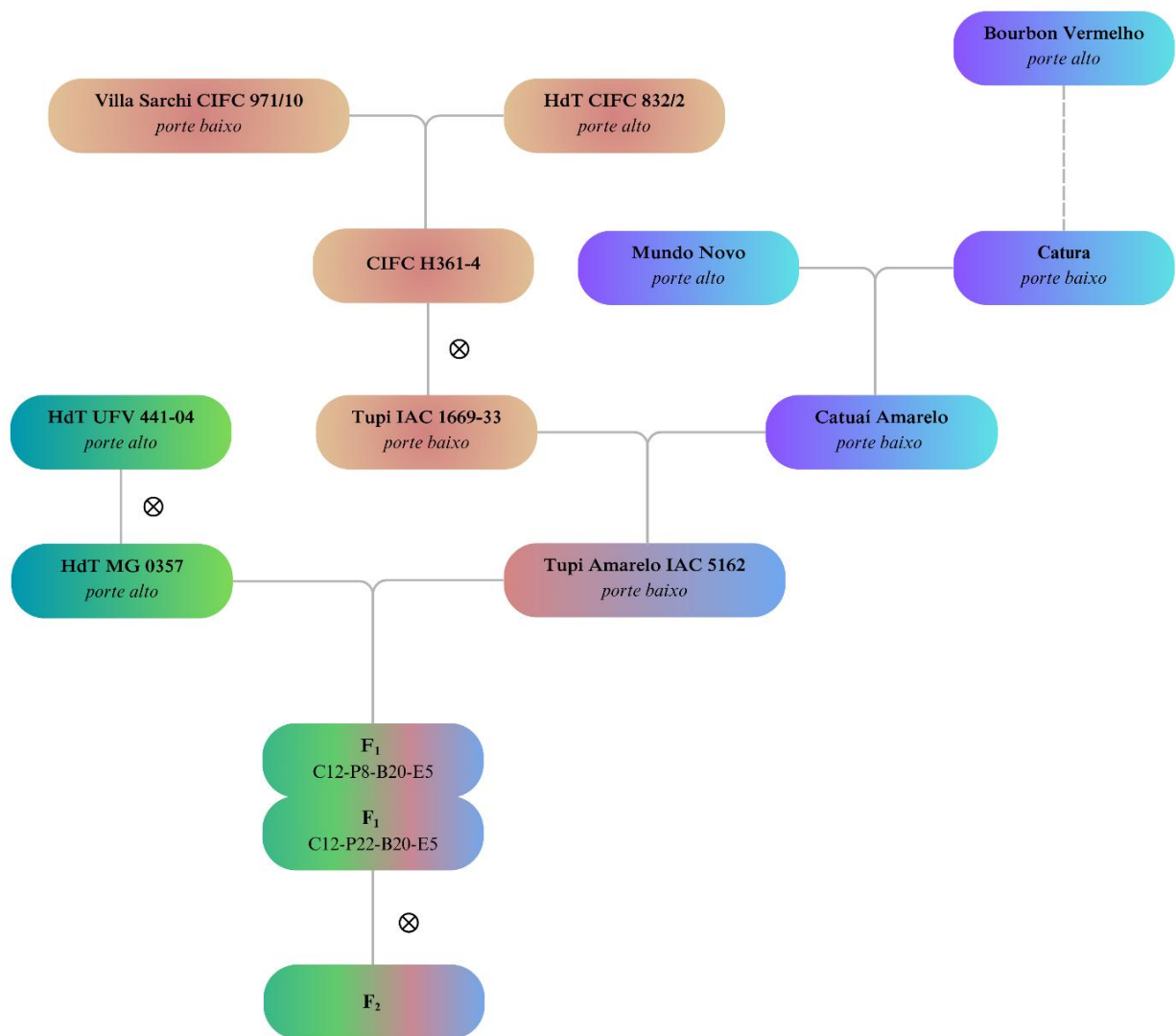


Figura 1: Estrutura do Programa de Melhoramento Genético de *Café arabica* direcionado à redução de altura de planta, resultante da integração de dados moleculares e fenotípicos.

O HdT MG0357, pertencente ao Banco de Germoplasma da EPAMIG, é derivado do acesso HdT UFV 441-04, originalmente introduzido no Brasil pelo Centro de Investigação das Ferrugens do Cafeeiro (CIFIC). Em contraste, o cafeeiro Tupi Amarelo IAC 5162, apresenta porte baixo, frutos graúdos, alta resistência a patógenos e elevado potencial produtivo. Foi desenvolvido pelo Programa de Melhoramento do Instituto Agrônomo de Campinas (IAC), a partir da identificação de uma planta com frutos amarelos em uma lavoura da cultivar Tupi IAC 1669-33. Sua origem remonta ao germoplasma Sarchimor, sendo que a Tupi IAC 1669-33 é derivada do híbrido CIFIC H361-4 – este, por sua vez, originado do cruzamento entre Villa Sarchi CIFIC 971/10 e HdT CIFIC 832/2. Acredita-se que o Tupi Amarelo IAC 5162 seja resultado de um cruzamento natural entre a Tupi IAC 1669-33 e um cafeeiro da cultivar Catuaí

Amarelo. Como resultado dessa genealogia, o Tupi Amarelo IAC 5162 tem fenótipo de porte baixo, frutos graúdos, alta resistência a patógenos e elevado potencial produtivo (Brasil, 2025; Carvalho et al., 2022; Fazuoli et al., 2008).

A população F₂ foi plantada em 2016 em delineamento de blocos aumentados, com quatro repetições de 50 plantas cada, espaçadas em 3,0 m entre linhas e 0,80 m entre plantas. Após avaliação da sobrevivência, a população efetiva para análise de altura foi estabelecida em 135 plantas. Foram incluídas como testemunhas as cultivares Paraíso MG H419-1 e Catuaí Vermelho IAC 144, com três plantas de cada por bloco, selecionadas por suas características de baixo porte e alto vigor vegetativo.

A fenotipagem da altura foi realizada medindo-se o ramo ortotrópico principal da superfície do solo até seu ápice, em centímetros. As mensurações foram conduzidas durante os períodos produtivos de 2018, 2020, 2021, 2022, 2023 e 2024, proporcionando uma série temporal abrangente para análise da característica.

2.1.2 Extração de DNA

O DNA genômico da população de mapeamento e da população de testagem foi extraído a partir de folhas jovens, seguindo a metodologia estabelecida por (Diniz et al., 2005). A quantificação foi realizada utilizando os sistemas NanoDrop™ 2000/2000c (*NanoDrop Technologies, Wilmington, EUA*) e Qubit 2.0 (*Invitrogen*). A qualidade foi avaliada mediante eletroforese em gel de agarose a 1%.

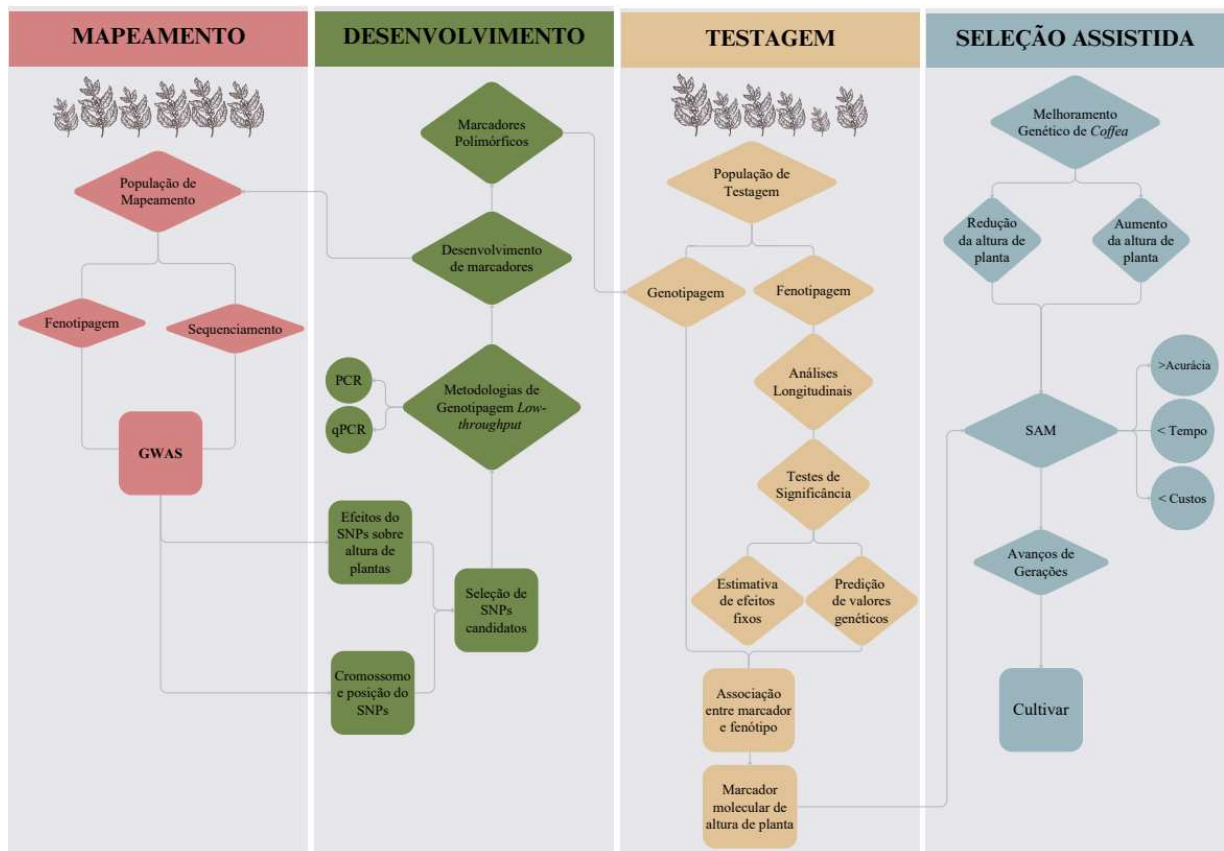
2.1.3 Seleção de SNPs associados à altura de planta

O desenvolvimento de marcadores SNPs usados nesse trabalho teve início com a construção de 40.000 sondas de DNA (Sousa et al., 2017, Resende et al., 2016). A partir deste conjunto, foram selecionadas 10.000 sondas polimórficas, de alta eficiência e distribuídas ao longo do genoma (Sousa et al., 2019). Usando essas sondas na genotipagem por *Capture-Seq (RAPiD Genomics)*, Silva et al. (2024) identificaram 21.211 marcadores SNP e, por meio de análise de GWAS, mapearam QTL no cromossomo 6, contendo 49 SNPs significativamente associados à altura da planta ($p < 0,05$) (Tabela 1). Estes 49 SNPs mapeados no QTL do cromossomo 6 foram empregados para avançar na compreensão da característica altura de plantas e no desenvolvimento de ferramentas genômicas para *C. arabica* (Figura 2).

Tabela 1. Marcadores SNPs associados à altura de plantas identificados por Silva et al. (2024).

SNPs	Subgenoma	Posição	<i>p</i> -value	R ²	Alelo	Alelo Favorável	Efeito (%)
V15959	<i>C. canephora</i>	18653343	1.02E-06	0.13359	T/G	T	7.84281
V15966	<i>C. canephora</i>	18683698	5.12E-08	0.1642	G/A	A	9.32675
V15967	<i>C. canephora</i>	18683725	1.15E-06	0.13684	C/T	T	8.03744
V16012	<i>C. canephora</i>	19668205	1.56E-06	0.13291	G/T	T	7.58232
V16014	<i>C. canephora</i>	19668268	6.28E-07	0.14141	T/C	C	8.56697
V16027	<i>C. canephora</i>	19711172	2.52E-06	0.12814	A/G	G	7.79241
V16029	<i>C. canephora</i>	19711182	1.75E-06	0.13192	A/T	T	7.81282
V16050	<i>C. canephora</i>	20516595	2.37E-07	0.14913	A/T	A	8.43056
V16099	<i>C. canephora</i>	21356008	1.25E-06	0.1079	C/T	T	19.56234
V16297	<i>C. canephora</i>	26782635	1.93E-06	0.12073	C/A	C	5.29129
V16298	<i>C. canephora</i>	26782664	1.83E-06	0.11905	C/T	C	5.1215
V16293	<i>C. canephora</i>	26937789	3.99E-07	0.14288	C/T	T	7.50054
V16294	<i>C. canephora</i>	26937799	1.68E-07	0.15273	G/A	G	7.62358
V16295	<i>C. canephora</i>	26937808	1.22E-06	0.13603	A/G	A	7.40663
V16296	<i>C. canephora</i>	26937817	2.99E-07	0.14881	G/A	G	7.88765
V15785	<i>C. canephora</i>	30136750	9.53E-07	0.10301	T/C	C	4.60664
V15786	<i>C. canephora</i>	30136754	1.60E-07	0.11814	A/C	C	4.98635
V1190	<i>C. canephora</i>	30457537	1.53E-06	0.13143	T/C	C	21.21391
V15787	<i>C. canephora</i>	30968458	7.64E-07	0.13943	A/G	A	7.66001
V1094	<i>C. canephora</i>	31352169	2.34E-06	0.13012	C/T	T	7.18741
V16180	<i>C. canephora</i>	32858769	2.93E-07	0.14951	G/A	A	8.48874
V16206	<i>C. canephora</i>	33299570	1.57E-07	0.15237	A/G	G	8.46431
V1666	<i>C. canephora</i>	33759015	1.21E-07	0.15923	T/G	G	7.90736
V1504	<i>C. canephora</i>	34372993	1.04E-06	0.13284	G/T	G	6.92785
V16508	<i>C. canephora</i>	34559139	1.39E-06	0.10804	T/C	T	4.87099
V16509	<i>C. canephora</i>	34559142	1.79E-06	0.10844	T/C	T	4.88017
V16482	<i>C. canephora</i>	34836377	2.32E-06	0.12723	G/A	G	6.92015
V16481	<i>C. canephora</i>	34852348	3.98E-07	0.14382	A/G	A	8.26294
V16476	<i>C. canephora</i>	34852929	3.19E-07	0.13838	T/C	T	6.63111
V16479	<i>C. canephora</i>	34852983	8.28E-08	0.1506	A/G	A	7.22469
V16480	<i>C. canephora</i>	34852993	4.04E-07	0.13818	A/G	A	6.83245
V1690	<i>C. canephora</i>	35004072	2.07E-06	0.10846	G/A	G	7.98351
V16589	<i>C. canephora</i>	37176686	1.31E-06	0.12377	G/A	G	6.12801
V2101	<i>C. canephora</i>	37696025	1.98E-07	0.1507	A/C	C	7.83073
V40	<i>C. canephora</i>	37767529	1.75E-06	0.1264	T/A	T	7.68267
V41	<i>C. canephora</i>	37767535	1.79E-06	0.12654	T/C	T	7.70794
V42	<i>C. canephora</i>	37767560	2.05E-06	0.12198	T/C	T	7.51621
V16336	<i>C. canephora</i>	54173849	9.68E-07	0.13528	A/G	G	7.7757
V15899	<i>C. eugenioides</i>	16766939	8.84E-07	0.11037	T/C	T	5.81278
V15914	<i>C. eugenioides</i>	17010018	4.41E-07	0.11965	A/G	G	6.11464
V16102	<i>C. eugenioides</i>	20768840	1.93E-06	0.13382	A/T	A/T	4.70157
V16391	<i>C. eugenioides</i>	43620472	5.09E-07	0.14393	T/A	A	8.11656
V16392	<i>C. eugenioides</i>	43620494	2.50E-07	0.15093	G/A	A	8.17788
V16395	<i>C. eugenioides</i>	43620540	9.61E-07	0.13135	C/T	T	7.41266
V16418	<i>C. eugenioides</i>	47341413	1.37E-06	0.12705	T/C	C	7.32441
V16762	<i>C. eugenioides</i>	51038354	8.60E-07	0.12894	C/A	A	7.62583
V16744	<i>C. eugenioides</i>	51085545	1.93E-07	0.14412	G/A	A	7.80832
V16709	<i>C. eugenioides</i>	51312825	2.21E-06	0.10674	G/C	C	5.6536
V16630	<i>C. eugenioides</i>	51939316	1.50E-06	0.12542	A/G	A	7.61825

Figura 2. Fluxograma dos processos para o desenvolvimento e obtenção de marcadores para altura de planta, destinados à seleção assistida em café.



2.1.4 Desenvolvimento de marcadores SNP *Low-throughput*

A metodologia de genotipagem de SNP em baixa escala utilizada foi *High Resolution Melting* (HRM), justificada por sua comprovada eficiência na detecção alélica em alotetraploides, robustez em ensaios preliminares e utilização em outros estudos com o gênero *Coffea* (Bolívar-González et al., 2022; Combes et al., 2018; Gatica-Arias et al., 2023; Silva et al., 2023).

Para o desenvolvimento dos marcadores HRM, as sequências flanqueadoras de cada SNP — correspondendo a 300 pares de bases a montante e 300 pares de bases a jusante da posição polimórfica — foram inicialmente utilizadas como base para o desenho dos pares de *primers*. Nos SNPs V40, V41, V42, V16293, V16294, V16295, V16296, V16476, V16479, V16480, V16392, V16395, V15967 e V15966, foi necessário desenvolver *amplicons* que combinassem dois ou mais SNPs, devido à proximidade física no genoma. As regiões foram analisadas quanto à qualidade, ausência de repetições e temperatura de *melting* (*T_m*) adequada, assegurando amplificação específica e fragmentos compatíveis com a técnica HRM.

Em seguida, cada sequência de *amplicon* prevista (incluindo o par de *primers* e a região interna do SNP) foi submetida a alinhamento contra o genoma de referência de *C. arabica* cv. Caturra Vermelho CCC135-36, utilizando a ferramenta *Basic Local Alignment Search Tool* (BLAST) (*National Center for Biotechnology Information* – NCBI, 2024). Essa etapa teve como finalidade confirmar a especificidade dos *primers* no genoma e assegurar que cada *amplicon* fosse único, sem alinhamentos secundários indesejados.

Os ensaios de HRM foram realizados conforme protocolo padronizado QIAGEN© (Qiagen, 2009) e as diretrizes estabelecidas por Słomka et al. (2017). Para cada SNP alvo, foram customizados dois ou três conjuntos de *primers*, com posterior seleção do par mais eficiente por testes *in vitro*. Os *primers* foram desenhados para gerar produtos de amplificação entre 132-345 pb, com *Tm* aproximada de 60°C. As reações de qPCR foram conduzidas em triplicata técnica, em volume final de 10-25 µL, contendo 1-50 ng de DNA genômico, 1X HRM PCR Master Mix QIAGEN©, 0,7 µM de cada *primer* (*forward* e *reverse*). O perfil de amplificação consistiu em: ativação enzimática a 95°C por 5 minutos, seguida de 40 ciclos de desnaturação a 95°C por 10 segundos e anelamento/extensão a 55°C por 30 segundos. A etapa de dissociação foi realizada com gradiente de temperatura de 65°C a 95°C, com incremento de 0,1°C a cada 2 segundos. A amplificação e análise dos dados foram efetuadas no termociclador *Rotor-Gene Q Series* (QIAGEN, Alemanha).

Os marcadores V16099 e V1190, selecionados por seus elevados efeitos fenotípicos, foram submetidos a metodologias complementares de genotipagem. Além do HRM, empregou-se a genotipagem *RNase H-dependent Amplification* (rhAmp), baseada em PCR em tempo real (qPCR), bem como métodos em gel de agarose, incluindo o *Tetra-Primer Amplification Refractory Mutation System* (Tetra Primer ARMS) e *Modified Tetra-Primer Amplification Refractory Mutation System* (MTPA). A aplicação estratégica de múltiplas metodologias *low-throughput* fundamentadas em princípios distintos (fusão térmica e tamanho de fragmento) visou aumentar as probabilidades de desenvolvimento bem-sucedido de marcadores funcionais e propor abordagens aos diferentes cenários operacionais de um programa de melhoramento.

Os ensaios de genotipagem rhAmp foram realizados de acordo com o protocolo padronizado IDT – *Integrated DNA Technologies*© e as diretrizes estabelecidas por (Dobosy et al., (2011). Para cada SNP alvo, a empresa desenvolveu 10 conjuntos de *primers*, dos quais foi selecionado o par mais eficiente para validação *in vitro*. A discriminação alélica foi realizada mediante detecção dos fluoróforos *Yakima Yellow*® (canal VIC, emissão em 551 nm - região verde do espectro) e *Fluorescein*® (canal FAM, emissão em 517 nm - região azul). Os alelos

de referência foram marcados com o FAM e os alternativos com VIC. As reações continham 3-10 ng de DNA genômico, *rhAmp Genotyping Master Mix (2X)*, *rhAmp Reporter Mix (40X)* e Ensaio SNP *rhAmp (20X)*, em volume final de 10 µL. O perfil de amplificação consistiu em: ativação enzimática a 95°C por 10 minutos; seguida de 40 ciclos de desnaturação a 95°C por 10 segundos, anelamento 60°C por 30 segundos, extensão a 68°C por 20 segundos. A amplificação e análise dos dados foram efetuadas no termociclador *7500 Real-Time PCR System (Thermo Fisher Scientific - EUA)*.

A genotipagem por *Tetra-Primer ARMS* foi realizada conforme o protocolo de Ye et al., (2001) com as adaptações propostas por Medrano & Oliveira (2014). Os *primers* foram desenvolvidos utilizando o *software* específico para *Tetra-Primer ARMS* (Ye et al., 2001). As reações de PCR foram otimizadas em volume final de 25 µL, contendo 250 ng de DNA genômico, 1X tampão de reação, 3-4 mM de MgCl₂, 400 µM de dNTPs, 0,12 µM de cada *primer* (externo e interno) e 2,5 U/µL da enzima Taq DNA polimerase *Platinum (Invitrogen, Thermo Fisher Scientific)*. O programa de amplificação consistiu em: desnaturação inicial a 94°C por 5 minutos; 5 ciclos *touchdown* (desnaturação a 94°C por 30 segundos, anelamento de 60°C a 55°C com redução de 1°C por ciclo por 30 s, extensão a 72°C por 30 s); seguidos por 30 ciclos padrão (94°C por 30 s, 57°C por 30 s, 72°C por 40 s); e extensão final a 72°C por 10 min. A separação dos produtos de PCR foi efetuada por eletroforese em gel de agarose 2,5% e gel de poliacrilamida 6%, sob condições de 70V, 300mA e 90W, permitindo a discriminação dos genótipos mediante análise do perfil de bandas.

A genotipagem por *MTPA* foi conduzida conforme protocolo estabelecido por Tanha et al., (2015), mantendo as diretrizes fundamentais do método *ARMS* (Ye et al., 2001) para otimização das reações de PCR. O programa de amplificação no termociclador consistiu em: desnaturação inicial a 94°C por 5 minutos; 30 ciclos de desnaturação a 94°C por 30 segundos, anelamento a 54-60°C por 30 segundos e extensão a 72°C por 40 segundos; seguidos de extensão final a 72°C por 10 minutos. Para análise dos produtos, alíquotas de 5 µL de cada reação foram misturadas com 2 µL de corante de carga e submetidas à eletroforese em gel de agarose a 2% e gel de poliacrilamida 6%, sob condições de 70 V, 300 mA e 90 W.

2.1.5 Modelagem Longitudinal da Altura das Plantas e Análise de Associação Genética

Para a análise da genotipagem da população de testagem, os resultados foram submetidos a Testes de Qui-Quadrado a 5% de significância. A fenotipagem do crescimento das plantas da população de testagem ao longo dos seis anos de avaliação (2018, 2020, 2021, 2022, 2023 e

2024) foi modelada considerando o caráter longitudinal dos dados. Isso permite capturar a dinâmica temporal do crescimento e discriminar a contribuição de efeitos genéticos e ambientais em diferentes fases do desenvolvimento da planta perene, superando a limitação de avaliações fenotípicas pontuais.

Modelos mistos com diferentes estruturas de covariância residual foram testados. A estrutura de covariância dos resíduos foi selecionada mediante comparação de modelos pelo Critério de Informação Bayesiano (BIC), testando-se as estruturas: Simetria Composta (Uniforme), Exponencial, Exponencial Heterogênea, Antedependência de Primeira Ordem e Não Estruturada. A modelagem mista foi realizada utilizando o pacote ASReml-R v. 4.1 (Gilmour et al. 2015) conforme modelo a seguir:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \mathbf{W}\mathbf{i} + \boldsymbol{\varepsilon}_{\xi}$$

Onde \mathbf{y} é o vetor de dados fenotípicos tomados entre 2018 e 2024; $\boldsymbol{\beta}$ é o vetor de efeitos fixos de anos; $\boldsymbol{\alpha}$ é o vetor aleatório de genótipos, $\boldsymbol{\alpha} \sim \mathbf{N}(\mathbf{0}, \mathbf{I}\sigma_g^2)$, em que \mathbf{I} é uma matriz identidade e σ_g^2 é a variância genotípica; \mathbf{i} é o vetor aleatório da interação genótipos \times anos, $\mathbf{i} \sim \mathbf{N}(\mathbf{0}, \mathbf{I}\sigma_i^2)$, em que σ_i^2 é a variância da interação genótipos \times anos; e $\boldsymbol{\varepsilon}_{\xi}$ é o vetor aleatório dos resíduos modelado utilizando uma matriz uniforme, ou exponencial, ou exponencial heterogênea, ou de antedependência de primeira ordem ou não-estruturada. As letras maiúsculas representam as matrizes de incidência para os efeitos referidos.

Os componentes de variância foram obtidos por meio do *Restricted Maximum Likelihood* (REML) (Patterson & Thompson, 1971). Além disso, foram obtidas estimativas dos efeitos fixos via *Best Linear Unbiased Estimates* (BLUEs) e previsões de valores genéticos via *Best Linear Unbiased Predictions* (BLUPs) (Henderson 1975; Thompson 1973). O teste de Wald foi utilizado para verificar a significância do efeito fixo de ano e o *Likelihood Ratio Test* (LRT) foi utilizado para testar a significância dos componentes de variância. A associação entre marcadores e altura das plantas foi realizada pela análise de marca simples (*Single-Marker Analysis*), utilizando o pacote R/qtl (Broman et al., 2003).

Para caracterizar a região genômica compreendida entre os marcadores V16099 e V1190, foram realizados o mapeamento e a anotação funcional de genes candidatos associados ao controle da altura em *C. arabica*. A identificação dos genes foi efetuada mediante consulta ao genoma de referência da espécie, disponível na plataforma *NCBI - National Center for Biotechnology Information*, (2025).

2.2 Resultados

2.2.1 Desenvolvimento dos marcadores moleculares SNP para altura de cafeeiros

A análise dos SNPs obtidos por sequenciamento na população de mapeamento (Sousa et al., 2017) revelou que os genitores derivados de Híbrido de Timor (HdT) apresentaram genótipos associados ao porte alto, enquanto os genitores Catuaí exibiram alelos favoráveis à redução de altura (Tabela 2). Uma exceção foi o acesso HdT CIFIC 832/2, que, por ser um material geneticamente mais rústico e não ter sido submetido a ciclos de autofecundação, manteve heterozigose em todos os locos analisados, embora também apresente fenótipo de porte alto. Esse padrão refletiu-se na progênie H513-5 (UFV 2148-57 × CIFIC 832/2), que exibiu homozigose na maioria dos loci devido à segregação alélica específica desse parental. As demais progênies (F₁, F₂ e retrocruzamentos) apresentaram segregação genotípica compatível com a herança mendeliana.

Para a obtenção de marcadores SNP *low-throughput* associados à altura de plantas, as sequências contendo os SNPs identificados para essa característica via GWAS (Silva et al., 2024) foram analisadas no genoma de *C. arabica* cv Caturra disponível no NCBI. Os alinhamentos *in silico* revelaram alta similaridade entre os subgenomas C^C e C^E no cromossomo 6. As sequências do subgenoma C^C apresentaram identidade superior a 80% com a sequência do subgenoma C^E. Foram selecionados 16 SNPs exclusivos do subgenoma C^C, sendo eles V16099, V1190, V40, V41, V42, V16293, V16294, V16295, V16296, V16476, V16479, V16480, V16392, V16395, V15966 e V15967. Foram desenvolvidos 16 marcadores HRM, 2 *rhAmp*, 4 Tetra-Primer ARMS, 2 MTPA e 2 para sequenciamento (Tabela 3).

A genotipagem HRM permitiu a discriminação alélica eficiente nos marcadores: HRM-V16099-2, HRM-V1190-3, HRM-V42-2, HRM-V16476-1 e HRM-V15967.V15966-2. Estes exibiram um padrão codominante, com amplificação e curvas de *melting*/dissociação compatíveis com os protocolos estabelecidos. Os marcadores projetados com *amplicons* contendo múltiplos SNPs (com exceção de HRM-V15967.V15966-2) não exibiu uma fase de amplificação exponencial clara e apresentaram ruídos de leitura em ciclos tardios. Os dois marcadores com *amplicons* mais complexos, contendo quatro SNPs (HRM-V16293.V16294.V16295.V16296-1 e -2), mostraram evidências de formação de dímeros de *primer* nos ciclos iniciais, resultando em perfis de amplificação inconsistentes e perfis de *melting* não interpretáveis.

Tabela 2. Sequenciamento da população de mapeamento utilizado para desenvolvimento e validação de marcadores HRM associados à altura em cafeeiros.

		<i>SNPs</i>															
		<i>C/T</i>	<i>T/C</i>	<i>T/A</i>	<i>T/C</i>	<i>T/C</i>	<i>C/T</i>	<i>G/A</i>	<i>A/G</i>	<i>G/A</i>	<i>T/C</i>	<i>A/G</i>	<i>A/G</i>	<i>G/A</i>	<i>C/T</i>	<i>G/A</i>	<i>C/T</i>
		<i>Alelo favorável a redução de altura</i>															
Genótipo	Genealogia	V16099	V1190	V40	V41	V42	V16293	V16294	V16295	V16296	V16476	V16479	V16480	V16392	V16395	V15966	V15967
UFV 440-10	Genitor - Híbrido de Timor CIFC 4192B	CT	TT	AA	CC	CC	CC	AA	GG	AA	CC	GG	GG	GG	CC	GG	CC
UFV 445-46	Genitor - Híbrido de Timor CIFC 4193C	CT	TT	AA	CC	CC	CC	AA	GG	AA	CC	GG	GG	GG	CC	GG	CC
CIFC 832/2	Genitor - Híbrido de Timor CIFC 832/2	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
UFV 2143-236	Genitor - Catuai amarelo IAC 30	TT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	TT	AA	TT
UFV 2148-57	Genitor - Catuai amarelo IAC 64	TT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	CT	AA	TT
UFV 2154-344	Genitor - Catuai amarelo IAC 86	TT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	CT	AA	TT
H419-1	F ₁ (UFV2143-235 x UFV445-46)	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
H419-10	F ₁ (UFV2143-235 x UFV445-46)	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
H513-5	F ₁ (UFV 2148-57 X CIFC 832/2)	CT	TC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	CT	AA	TT
H514-7	F ₁ (UFV 2154-344 x UFV 440-10)	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
H514-8	F ₁ (UFV 2154-344 x UFV 440-10)	CT	TC	TA	TC	TC	CT	GA	AG	GG	TC	AG	AG	GA	CT	GA	CT
1_1	Retrocruzamento de H419-1 x (UFV 445-46)	CT	TT	AA	CC	CC	CC	AA	GG	AA	CC	GG	GG	GG	CC	GG	CC
3_4	Retrocruzamento de H514-8 x (UFV 440-10)	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
1_5	Retrocruzamento de H419-1 x (UFV 445-46)	CT	TC	TA	TC	TC	CT	GA	AG	GA	TC	AG	AG	GA	CT	GA	CT
1_3	Retrocruzamento de H419-1 x (UFV 445-46)	TT	TT	AA	CC	CC	CC	AA	GG	AA	CC	GG	GG	GG	CC	GG	CC
9_1	Retrocruzamento de UFV 2148-57 x (H513-5)	CT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	TT	AA	TT
2_2	Retrocruzamento de H419-1 x (UFV 2143-235)	TT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	CT	AA	TT
10_4	F ₂ de H514-8	CT	TT	AA	CC	CC	CC	AA	GG	AA	CC	GG	GG	GG	CC	GG	CC
10_1	F ₂ de H514-8	CT	TC	TT	TT	TT	CT	GA	AG	GA	TC	AG	AG	GA	CT	GG	CC
13_1	F ₂ de H513-5	CT	CC	TT	TT	TT	TT	GG	AA	GG	TT	AA	AA	AA	CT	AA	TT

A genotipagem Tetra *Primer ARMS* permitiu o desenvolvimento dos marcadores 4P.ARMS-V16099-1 (leitura em agarose) e 4P.ARMS-V1190-1 (leitura em poliacrilamida). Esta metodologia demonstrou dependência crítica da qualidade do DNA, com amostras contendo razão 260/230 inferior a 1,8 exibindo amplificação deficiente (bandas fracas ou ausentes). As bandas controle (*Outer Forward* e *Outer Reverse*) apresentaram-se bem definidas, enquanto as bandas alelo-específicas exibiram menor intensidade, requerendo análise criteriosa, como já descrito pela metodologia.

A genotipagem por rhAmp mostrou-se ineficaz para a discriminação alélica dos SNPs V16099 e V1190. Na análise de dispersão, todos os genótipos (controles, genitores, F₁ e F₂) agruparam-se próximos, indicando comportamento monomórfico. Os eletroferogramas revelaram sinais elevados para o fluoróforo VIC (alelos alternativos: V16099-T e V1190-C) e sinais próximos ao *baseline* para FAM (alelos de referência). Adicionalmente, observou-se alta proporção de chamadas indeterminadas pela análise de *Ct* (*Cycle threshold*), com redução da amplificação em concentrações mais elevadas de DNA.

A genotipagem por *MTPA* não permitiu o desenvolvimento de marcadores para V16099, mesmo com otimizações de parâmetros. Para V1190, ambos os *primers* testados falharam na validação, com distinção alélica limitada apenas às bandas de controle e ao alelo C em gel de agarose. A análise em gel de poliacrilamida também não permitiu discriminação genotípica adequada.

Tabela 3. *Primers* customizados para os SNPs do QTL associados à altura de cafeeiros para as genotipagens *HRM*, *rhAmp*, Tetra Primer *ARMS*, *MTPA* e sequenciamento.

Genotipagem	SNPs	Primers	Identificação	Forward (5' - 3')*	Reverse (5' - 3')*	Amplicon
<i>HRM</i>	V16099	1	HRM-V16099-1	TCTTAGCTCGTGATAGTGATCGAAA	ATCTCCGGGATGAAATTTAGGCTT	241
		2	HRM-V16099-2	GATCGAAATCAAATCCATCCAAACT	CGGGATGAAATTTAGGCTTTCAG	219
		3	HRM-V16099-3	TCTTAGCTCGTGATAGTGATCGAAAT	TCTCCGGGATGAAATTTAGGCTTT	240
	V1190	4	HRM-V1190-1	CGTACTAATGATCTGGTTTGCTCAC	TTTCGACTCCATTCAACTCTAACAC	246
		5	HRM-V1190-2	TACTAATGATCTGGTTTGCTCACGA	TTTCGACTCCATTCAACTCTAACACG	243
		6	HRM-V1190-3	TAATGATCTGGTTTGCTCACGAATG	GTTTCGACTCCATTCAACTCTAACAC	242
	V40, V41, V42	7	HRM-V40.V41.V42-1	TCCGACACCTCGAAAGAGTT	TAGGCTGTCTGGACCATTACC	192
	V42	8	HRM-V42-2	ACTCTTGTTGTAATCCAG	TAGGCTGTCTGGACCATTACC	138
	V16293, V16294, V16295, V16296	9	HRM-V16293.V16294.V16295.V16296-1	AGAGCTTCCCTTGAACCTTGC	AGTGTGAGAGCATTCTAACCTA	194
	V16293, V16294, V16295, V16296	10	HRM-V16293.V16294.V16295.V16296-2	AGAGCTTCCCTTGAACCTTGCAT	TTCTCGAAAGAGTTTATGTGGTTGA	323
	V16476	11	HRM-V16476-1	GACAAGCAGCGACGACATTG	TTAGCCGTCCTCGTCACTTG	95
	V16476, V16479, V16480	12	HRM-V16476.V16479.V16480-2	CAAGCAGCGACGACATTGAT	GCCATTGTCCGTCTCTACCA	205
	V16395	13	HRM-V16395-1	TCAAGGGTACAAGTTCAGCCA	GGTCATTGGCACTTTTCTAGGG	183
	V16392, V16395	14	HRM-V16392.V16395-2	AGTGAAACAGTCGCTCGCTT	GGTCATTGGCACTTTTCTAGGG	298
	V15967	15	HRM- V15967-1	AGCTGCTACAGGATCAATCAACAT	ATAGGGAGACGTAGCCAAGCAT	107
	V15967, V15966	16	HRM- V15967.V15966-2	CATTATGTAATTTGAC <u>ACC</u> AT	ATAGGGAGACGTAGCCAAGCAT	133
<i>rhAmp</i>	V16099	17	rhAmp-V16099	AATGCTAACCTGCAAATAAAATTCTAC <u>C</u> GAATGCTAACCTGCAAATAAAATTCTAT <u>T</u>	GGTTTGAGTTTACCAGGAACATAGC	185
	V1190	18	rhAmp-V1190	ACAGCCCTTCTACCA <u>A</u> ACAGCCCTTCTACCC <u>G</u>	GCGACAAGGTTTAGGACACTTGG	135

... continuação

Genotipagem	SNPs	Primers	Identificação	Forward (5' - 3')*	Reverse (5' - 3')*	Amplicon	
Tetra Primer ARMS	V16099	19	4P.ARMS-V16099-1	<i>Outer:</i> GAGCATGTCTTAGCTCGTGATAGTG <i>Inner:</i> TGCTAACCTGCAAATAAAATTCTCT <u>C</u>	GCTTCCAAACATGTTATGTTCCCTG GAAATTTAGGCTTTCAGCATGTTTAA <u>A</u>	383 51	
		20	4P.ARMS-V16099-2	<i>Outer:</i> AGCATGTCTTAGCTCGTGATAGTGAT <i>Inner:</i> TGCTAACCTGCAAATAAAATTCTCT <u>C</u>	GCTTCCAAACATGTTATGTTCCCTG GAAATTTAGGCTTTCAGCATGTTTAA <u>A</u>	382 51	
	V1190	21	4P.ARMS-V1190-1	<i>Outer:</i> GGCATGAGCGTACTAATGATCTGGTTT <i>Inner:</i> AACGCTTTGACCGTGCCTCT <u>G</u>	ACAACATCGTACAGAAGTCCGTGTGA GTCGATCTCTACAGCCCTTCCTCC <u>T</u>	353 45	
		22	4P.ARMS-V1190-2	<i>Outer:</i> CTAATGATCTGGTTTGCTCACGAATGA <i>Inner:</i> AACGCTTTGACCGTGCCTCT <u>G</u>	ACAACATCGTACAGAAGTCCGTGTGA GTCGATCTCTACAGCCCTTCCTCC <u>T</u>	341 45	
	V16099	-	-	-	-	-	-
	MTPA	V1190	23	4P.MTPA-V1190-1	<i>Outer:</i> TCTGGTTTGCTCACGAATGAGAGTT <i>Inner:</i> ACGCTTTGACCGTGCCTCT <u>G</u>	TCCGTGTGATTTTTGGTTGTGCGC CGATCTCTACAGCCCTTCCTCC <u>T</u>	317 42
24			4P.MTPA-V1190-2	<i>Outer:</i> TTGAGTACGAGGCATGAGCGTAATA <i>Inner:</i> AACGCTTTGACCGTGCCTCT <u>G</u>	CTGGCCCAACCTACAACATCGTCCA GTCGATCTCTACAGCCCTTCCTCC <u>T</u>	375 45	

* Nucleotídeos sublinhados representam os SNPs.

2.2.2 Genotipagem da população de testagem utilizando marcadores HRM e Tetra Primer ARMS desenvolvidos para altura de plantas

Os marcadores desenvolvidos na população de mapeamento (HRM-V16099-2, 4P.ARMS-V16099-2, HRM-V1190-3, 4P.ARMS-V1190-1, HRM-V42-2, HRM-V16476-1 e HRM-V15967.V15966-2) foram analisados com a genotipagem da população de testagem.

2.2.2.1 Marcadores moleculares desenvolvidos para o SNP V16099

O marcador HRM-V16099-2 demonstrou um perfil codominante, com curvas de dissociação diferenciadas para os genótipos homozigotos (diferença de 0,2 °C) e uma curva distinta para os heterozigotos (Figura 3). A genotipagem da população alcançou acurácia superior a 85% na discriminação alélica, estimada pelo *software* de análise. O ponto de dissociação (T_m) dos produtos de amplificação foi de 77,5 °C.

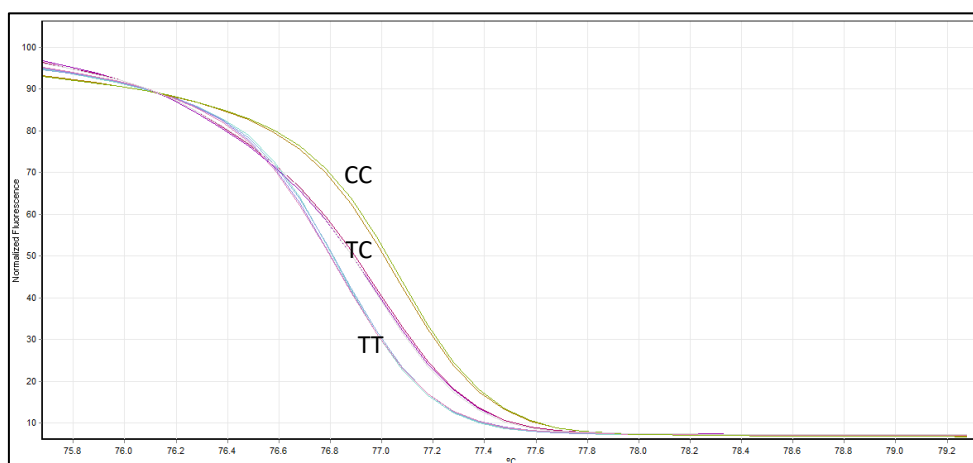


Figura 3: Perfil de dissociação e leitura do marcador molecular HRM -V16099-2. dF/dT : derivada do negativo da taxa de mudança no sinal de fluorescência (dF) por derivada do aumento na temperatura (dT)

Paralelamente, o marcador 4P.ARMS-V16099-1 foi desenvolvido exibindo um padrão de herança dominante. Neste sistema, os genótipos homocigotos CC são claramente identificados, enquanto os genótipos TT e CT compartilham o mesmo padrão de bandas (Figura 4). Este fenômeno é atribuído a uma incompatibilidade intencional (*-2 mismatch*) no terminal 3' do *primer Inner*, que privilegia a amplificação pelos *primers Outer*, resultando em bandas de controle de alta intensidade e bandas alelo-específicas de menor intensidade.

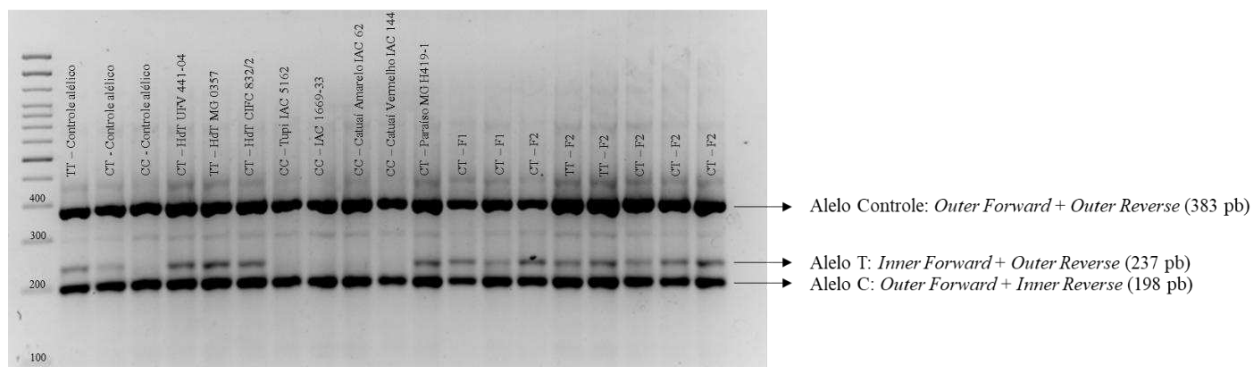


Figura 4: Perfil eletroforético do marcador molecular 4P.ARMS-V16099-1, em gel de agarose 2,5%.

Os dados de genotipagem por HRM e Tetra *Primer ARMS* confirmaram a constituição genotípica dos parentais: os genitores femininos oriundos de HdT são homocigotos TT, os genitores do parental masculino são homocigotos CC, e a progênie F₁ é heterocigota CT, com a população F₂ exibindo segregação mendeliana. De acordo com o estudo de GWAS, o alelo T está associado à redução da altura da planta, enquanto o alelo C correlaciona-se com o aumento da estatura, corroborando com a genotipagem realizada e a fenotipagem descrita na literatura.

2.2.2.2 Marcadores moleculares desenvolvidos para o SNP V1190

O marcador HRM-V1190-3 foi selecionado para a análise da população devido à sua maior amplitude de leitura, com uma diferença de *Tm* de 0,6 °C entre os homocigotos, permitindo uma clara distinção entre os três genótipos (Figura 5).

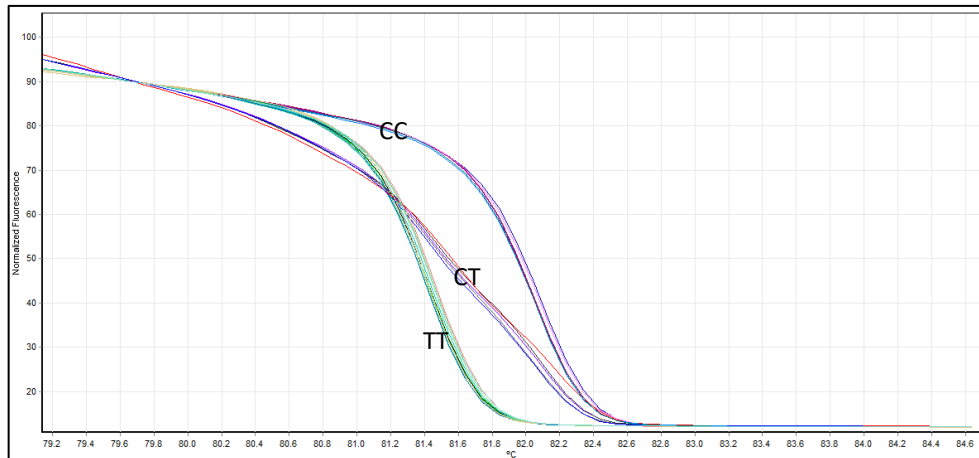


Figura 5: Perfil de dissociação e leitura do marcador molecular HRM-V1190-3. dF/dT : derivada do negativo da taxa de mudança no sinal de fluorescência (dF) por derivada do aumento na temperatura (dT).

Na técnica *Tetra Primer ARMS*, os sistemas de *primers* inicialmente produziram quatro bandas monomórficas em gel de agarose. No entanto, o marcador 4P.ARMS-V1190-1 permitiu a distinção entre homozigotos e heterozigotos quando analisado em gel de poliacrilamida (Figura 6). Neste sistema, os alelos específicos são diferenciados por uma diferença de 55 pb, visualizando-se na posição de 600 pb no gel.

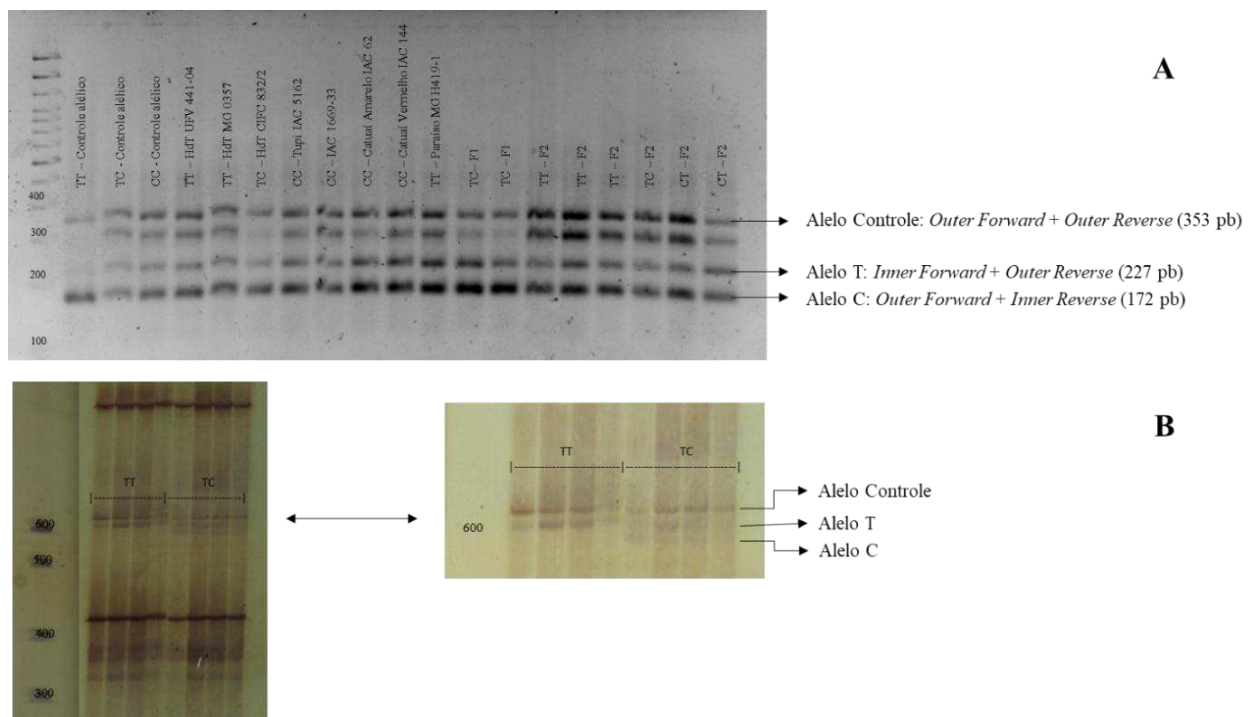


Figura 6: Perfil eletroforético do marcador molecular 4P.ARMS-V1190. (A) Bandas de amplificação monomórficas em gel de agarose. (B) Polimorfismo dos alelos de forma codominante em gel de poliacrilamida.

Ambas as metodologias de genotipagem (*HRM* e *Tetra Primer ARMS*) confirmaram a constituição genotípica esperada: o genitor de porte alto como homozigoto TT e o genitor de porte baixo como homozigoto CC. Estes resultados validam as associações identificadas pelo estudo de GWAS, nas quais o alelo C se mostrou favorável à redução da altura das plantas.

2.2.2.3 Marcador molecular desenvolvido para o SNP V42

O marcador HRM-V42-2 apresentou padrão codominante com diferença de temperatura de dissociação de 0,5 °C entre os homozigotos, permitindo distinção clara dos três genótipos na faixa de fluorescência de 15 a 80 unidades (Figura 7). O genitor de porte alto apresentou genótipo CC ($T_m \approx 77,4$ °C), enquanto o de porte baixo mostrou genótipo TT ($T_m \approx 78,2$ °C), validando a associação identificada pelo GWAS de que o alelo T está associado à redução da altura das plantas para este locus.

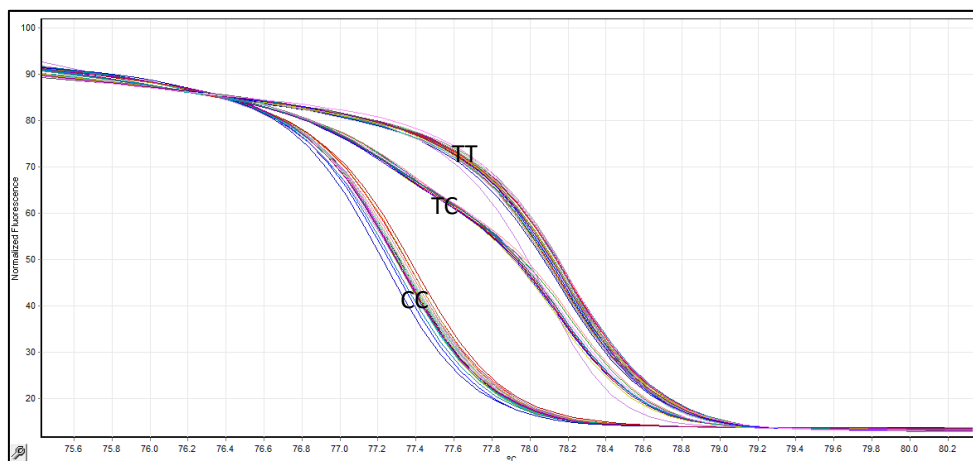


Figura 7: Perfil de dissociação HRM do marcador molecular V42-2 com SNPs T/C. dF/dT : derivada do negativo da taxa de mudança no sinal de fluorescência (dF) por derivada do aumento na temperatura (dT)

2.2.2.4 Marcador molecular desenvolvido para o SNP V16476

Para os SNPs V16476, o marcador HRM-V16476-1 demonstrou diferença de dissociação de 0,3 °C entre homozigotos, com leitura na faixa de 79,8 a 80,8 °C e fluorescência entre 60-70 unidades (Figura 8). O genitor alto apresentou genótipo CC ($T_m \approx 81,2$ °C) e o baixo, genótipo TT ($T_m \approx 80,8$ °C), confirmando que o alelo T contribui para a redução neste SNP.

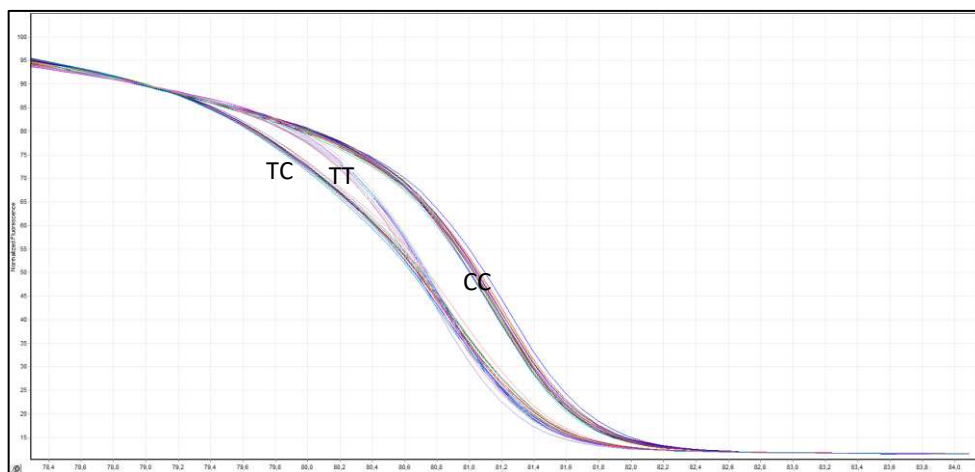


Figura 8: Perfil de dissociação HRM do marcador molecular HRM-V16476-1 com SNPs T/C. dF/dT : derivada do negativo da taxa de mudança no sinal de fluorescência (dF) por derivada do aumento na temperatura (dT)

2.2.2.5 Marcador molecular desenvolvido para os SNPs V15967 e V15966

O marcador HRM-V15967.V15966-2 foi desenvolvido com dois SNPs estratégicos: um no *primer forward* e outro na região central do *amplicon*. A análise revelou diferença de 0,4 °C entre homozigotos, porém os heterozigotos diferiram em apenas 0,1 °C do homozigoto AA, resultando em faixas estreitas de leitura (75,9-76,2 °C; fluorescência 70-76) (Figura 9).

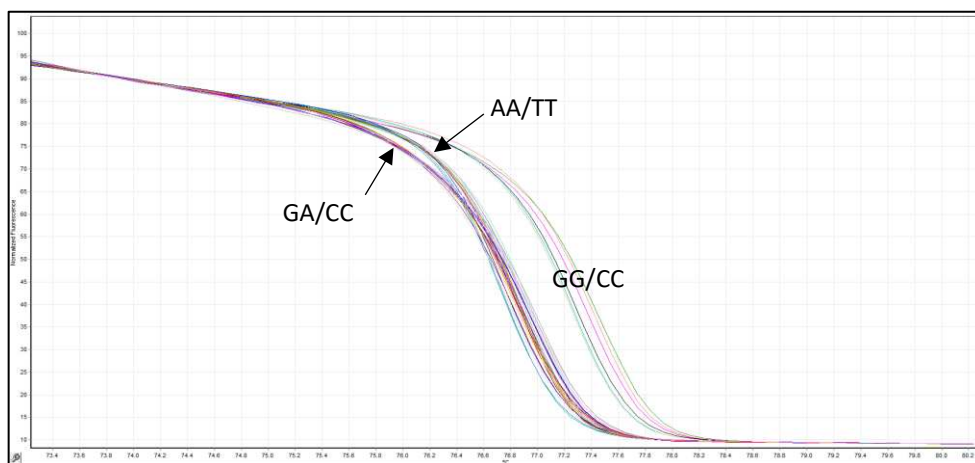


Figura 9: Perfil de dissociação HRM do marcador molecular HRM- V15967.V15966-2 com SNPs C/T e G/A, respectivamente. dF/dT : derivada do negativo da taxa de mudança no sinal de fluorescência (dF) por derivada do aumento na temperatura (dT)

A análise genotípica validou as associações previamente identificadas pelo GWAS: o genitor de porte alto apresentou haplótipo CC/GG (V15967/V15966; $T_m \approx 77,5$ °C), enquanto o genitor de porte baixo exibiu haplótipo TT/AA ($T_m \approx 77,1$ °C), confirmando que os alelos T

(V15967) e A (V15966) estão associados à redução da altura das plantas. No entanto, as diferenças térmicas reduzidas entre genótipos GA/CC e AA/TT (0,1°C) comprometem a confiabilidade da discriminação alélica, conforme documentado na literatura para sistemas HRM com baixa variação de *Tm*. Diante desta limitação técnica, o marcador HRM-V15967.V15966-2 foi excluído das análises subsequentes na população de testagem.

2.2.3 Análise da segregação dos marcadores desenvolvidos

A genotipagem da população de testagem com os marcadores desenvolvidos para os SNPs V16099 (HRM-V16099-2 e 4P.ARMS-V16099-1), V1190 (HRM-V1190-3 e 4P.ARMS-V1190-1), V42 (HRM-V42-2) e V16476 (HRM-V16476-1) apresentou padrões de segregação codominantes consistentes como o esperado, conforme confirmado pelo teste de Qui-quadrado (Tabela 4). Para um grau de liberdade igual a 2, a hipótese de segregação 1:2:1 foi aceita para todos os marcadores ($p > 0,05$), com probabilidades variando entre 56% e 99%. As análises confirmaram os genótipos parentais dominantes para TT e CC, e segregação mendeliana esperada na população F₂. As testemunhas apresentaram genótipos condizentes com os esperados para porte baixo.

Tabela 4: Análise de segregação dos marcadores desenvolvidos para os SNPs V16099, V1190, V41 e V16476 por meio do teste de qui-quadrado (χ^2).

Genealogia		V16099	V1190	V41	V16476
Testemunha	Paraíso MG H419-1	TT	TT	CC	CC
Testemunha	Catuai Vermelho IAC 144	CC	CC	TT	TT
Progenitor Feminino	HT UFV 441-04	TT	TT	CC	CC
Progenitor Masculino	Tupi IAC 1669-33	CC	CC	TT	TT
Genitor Feminino	HT MG 0357	TT	TT	CC	CC
Genitor Masculino	Tupi Amarelo IAC 5162	CC	CC	TT	TT
F ₁	C12 P8 B20 E5	CT	CT	CT	CT
	C12 P22 B20 E5	CT	CT	CT	CT
F ₂		31 genótipos - TT	31 genótipos - TT	37 genótipos - TT	43 genótipos - TT
		36 genótipos - CC	34 genótipos - CC	30 genótipos - CC	40 genótipos - CC
		75 genótipos - CT	77 genótipos - CT	75 genótipos - CT	77 genótipos - CT
Teste χ^2 *		X ² calculado 0,80	X ² calculado 1,14	X ² calculado 1,14	X ² calculado 0,58
		Grau de Liberdade 2 Probabilidade 66,94%	Grau de Liberdade 2 Probabilidade 56,53%	Grau de Liberdade 2 Probabilidade 56,53%	Grau de Liberdade 2 Probabilidade 74,92%

* χ^2 a 5% de significância.

2.2.4 Modelagem longitudinal da altura de plantas da população de testagem

Para a análise fenotípica da população de testagem (F₂) considerando modelagem longitudinal, o modelo que melhor descreveu os dados foi o Modelo de Antedependência de Ordem 1, selecionado por apresentar o menor valor de BIC entre os modelos avaliados (Tabela 5). Os modelos Uniforme, Exponencial heterogêneo e Não estruturado, não convergiram as matrizes de dados, portanto os resultados de BIC não puderam ser calculados. O efeito fixo Ano foi significativo, indicando a existência de variação entre as médias de altura da população nos diferentes anos.

O teste de razão de verossimilhança (LRT) foi significativo a 1% para o componente de variância de genótipos considerando o modelo adotado ($p\text{-value} = 1.657e-10$ ***). De forma geral, verificou-se que a maioria dos genótipos apresentou incremento progressivo nos valores genéticos preditos ao longo dos anos, compatível com o padrão biológico de crescimento da planta (Figura 10). Observou-se um aumento acumulado de altura de aproximadamente 57% para a testemunha 1, 61% para a testemunha 2 e 54% para F₂. A altura média das plantas apresentou um incremento de 45% no período de 2018 a 2024.

Tabela 5: Modelagem longitudinal da altura de plantas da população de validação.

Modelos		BIC	
Uniforme		-	
Exponencial		5430.66	
Exponencial Heterogêneo		-	
Antedependência de Ordem 1		5345.29	
Não Estruturado		-	
Teste Wald	GL	Wald	Pr (Chisq)
Ano	5	2042.5	< 2.2e-16 ***
Anos avaliados	Estimativa (cm)	Erro Padrão (cm)	
2018	86.34	1.43	
2020	139.38	1.89	
2021	147.23	2.16	
2022	163.10	2.44	
2023	179.84	3.00	
2024	189.83	3.09	

GL: Graus de Liberdade; Probabilidade (Pr) sob a distribuição Qui-quadrado (Chisq).

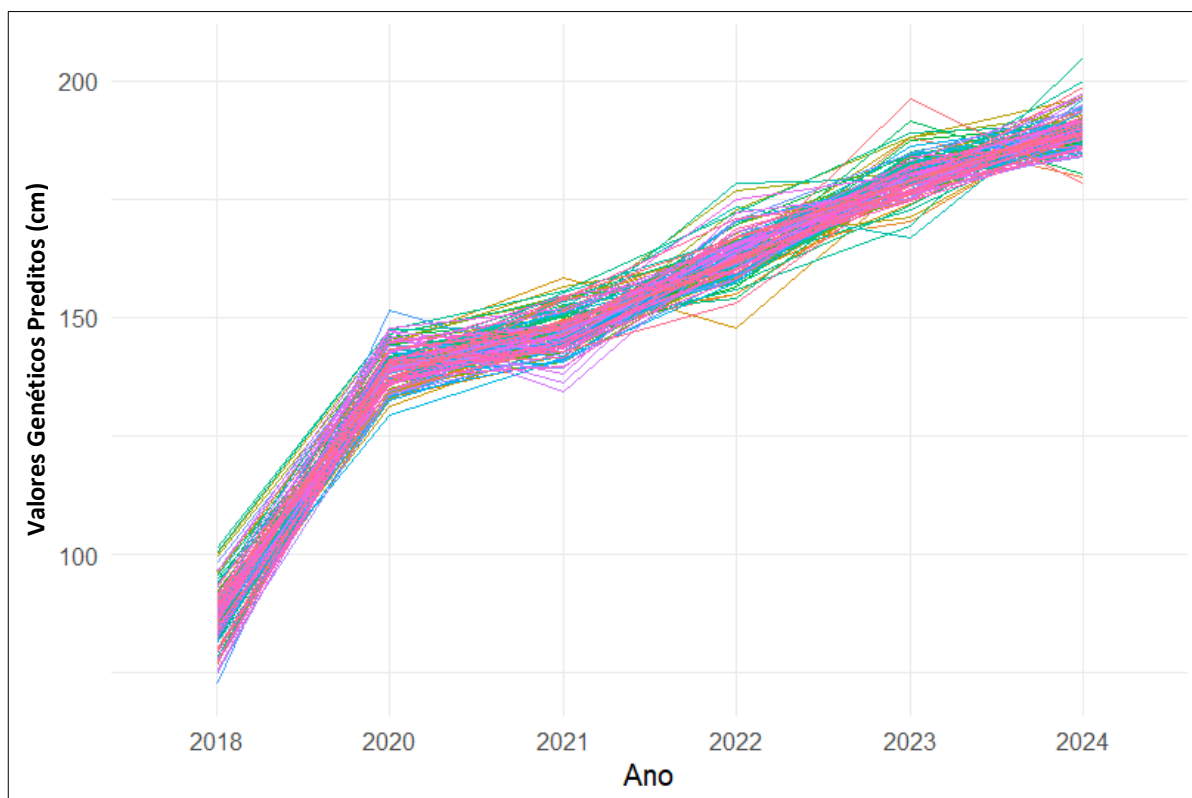


Figura 10: Valores genéticos preditos (BLUP) para a progênie segregante ao longo das avaliações fenotípicas em centímetros.

2.2.5 Associação fenotípica e genotípica dos marcadores moleculares validados

A análise de distribuição fenotípica por genótipo (PxG) mostrou acentuada sobreposição entre os genótipos para todos os marcadores, indicando ausência de diferenças fenotípicas consistentes entre os alelos ao longo dos anos (Figura 11). Complementarmente, a análise de marcador individual resultou em valores de LOD entre 0,03 e 1,66, abaixo do limiar de significância estabelecido ($LOD \geq 3$) (Tabela 6). Estes resultados indicam que, embora o desenvolvimento dos marcadores V16099, V1190, V42 e V16476 tenha sido bem-sucedido na população de mapeamento, a análise de associação fenótipo-genótipo na população de testagem não revelou efeitos significativos desses marcadores sobre a altura das plantas.

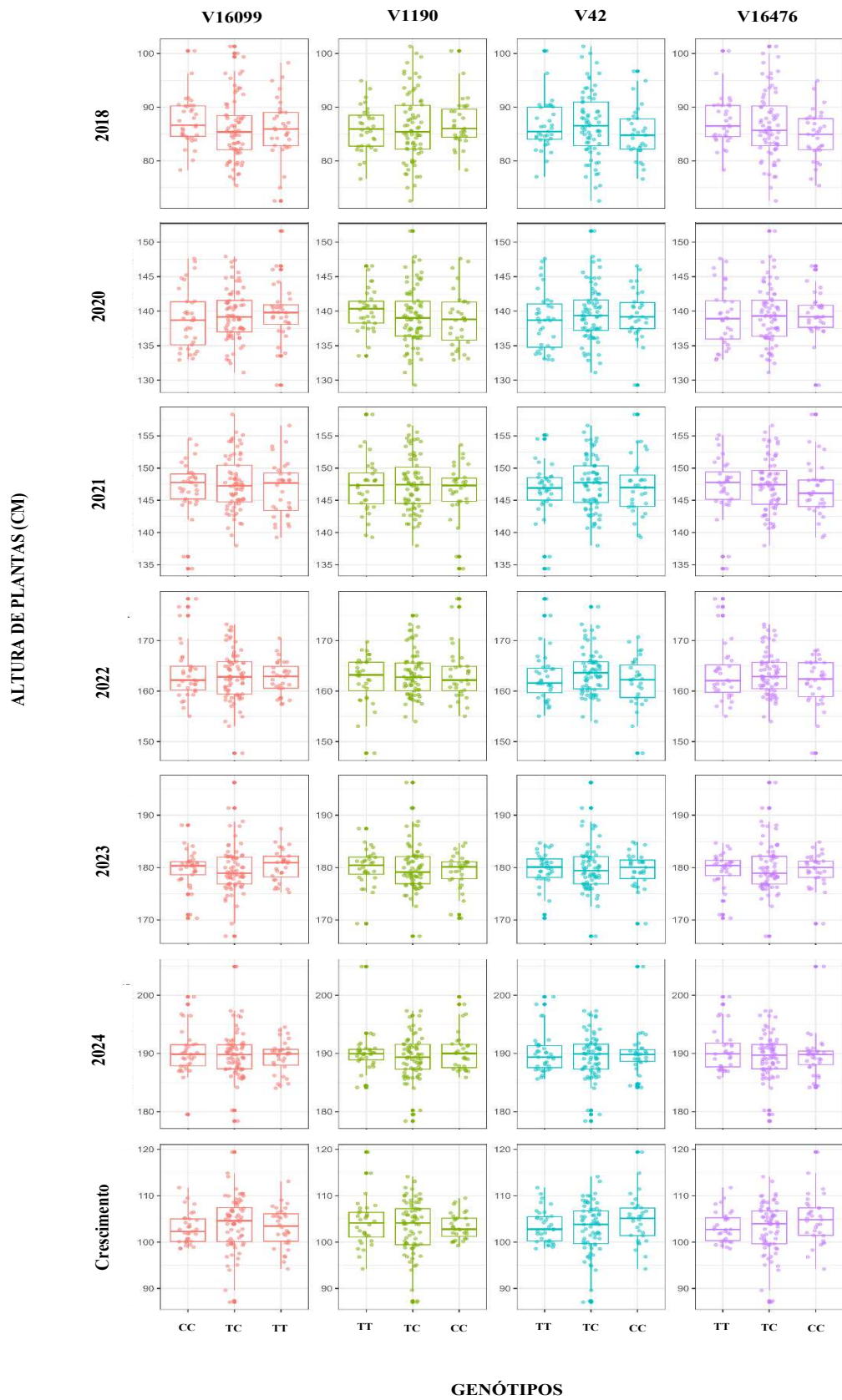


Figura 11: Distribuição fenotípica do incremento de crescimento da altura das plantas (em centímetros), por genótipo, para os marcadores V16099, V1190, V42 e V16476, ao longo dos anos avaliados.

Tabela 6: Análises de Marcador Individual por regressão com valores de LOD dos marcadores V16099, V1190, V42 e V16476 individualizado por ano e crescimento.

Marcador	Posição	2018	2020	2021	2022	2023	2024	Crescimento
V16099	0.00	0.32	0.17	0.21	0.16	0.15	0.23	0.19
V1190	518.08	0.15	0.39	0.27	0.18	0.15	0.36	0.15
V42	1036.16	0.46	0.32	0.24	0.85	0.06	0.06	0.74
V16476	1045.06	0.90	0.04	0.15	0.62	0.08	0.42	0.75

2.2.6 Análise descritiva das médias fenotípicas em relação aos marcadores moleculares validados para a altura de plantas

Considerando os efeitos estimados pelo estudo de GWAS - redução de 19,56% para o alelo T em V16099 e 21,21% para o alelo C em V1190 (Silva et al., 2024) - esta população apresentou reduções médias de 2,6% para genótipos TT (V16099) e 3,1% para CC (V1190) (Tabela 7). A distribuição genotípica na população de testagem revelou predominância de heterozigotos duplos (CTCT; n=60), seguidos por homozigotos CCCC (n=26) e TTTT (n=23). Notavelmente, não foram identificados indivíduos duplo homozigotos (TTCC) na população. As menores médias de altura em todos os anos foram observadas nas combinações TT (V16099) + CT (V1190) e CT (V16099) + CC (V1190). Em 2024, os genótipos TTCT e CTCC apresentaram alturas 9,5% e 6,5% inferiores à média geral, respectivamente. Estes resultados sugerem que a combinação de alelos homozigotos favoráveis à redução (TT-V16099 e CC-V1190) com heterozigotos (CT-V16099 e CT-V1190) resulta em plantas de menor porte, considerando somente esses SNPs. As maiores alturas foram registradas em genótipos CC (V16099) + CT (V1190), que superaram a média geral em 9,4% (2021), 7,7% (2022), 9,3% (2023) e 5,1% (2024).

A associação entre V16099 e V15967.V15966 nos haplótipos TTCT e CTTT promoveu redução de altura superior a 3 cm em relação à média da população de testagem. Extrapolando esses padrões, deduz-se que as combinações homozigóticas TTTT (para redução de altura) e CCCC (para aumento de altura) representariam os extremos fenotípicos ideais. Contudo, estes genótipos não foram detectados na população de testagem. Os marcadores V42-2 e V15967.V15966 apresentaram efeitos teóricos sobre a altura das plantas de aproximadamente 8%, conforme indicado pelo GWAS; e na população de validação, a análise das médias revelou baixas diferenças (~1 cm) entre os genótipos homozigotos dominantes.

Foram realizadas análises considerando as médias dos BLUPs em 20% da progênie F₂ classificados como de maior e de menor altura. No entanto, não foram observadas diferenças

que associassem os genótipos identificados aos fenótipos de altura de plantas, o que evidencia o forte efeito ambiental na expressão da característica.

Tabela 7: Associação da genotipagem dos marcadores validados (HRM-V16099-2, HRM-V1190-3, HRM-V42-2 e HRM-V16476-1) com as médias fenotípicas da altura da população cafeeira em centímetros.

V16099	V1190	V42	V15967.V15966	Nº	2018	2020	2021	2022	2023	2024	Médias	
				TT	31	86.53	139.61	143.45	158.87	177.42	185.48	148.56
				CT	72	86.67	140.17	149.10	164.21	180.19	190.42	151.79
				CC	32	90.37	138.16	147.06	164.81	181.31	192.81	152.42
				TT	31	87.40	143.10	147.90	162.55	182.42	193.06	152.74
				CT	74	86.90	139.32	148.92	164.57	180.26	188.72	151.45
				CC	30	89.00	136.50	142.77	160.17	176.07	189.33	148.97
				TT	34	88.59	136.29	144.26	161.15	179.29	191.03	150.10
				CT	71	88.30	140.69	148.63	164.73	179.01	187.54	151.48
				CC	30	84.50	140.60	147.67	161.57	182.33	194.00	151.78
				TT	31	90.77	139.52	147.68	147.68	180.39	193.23	149.88
				CT	72	88.02	139.32	147.71	163.68	179.01	187.57	150.88
				CC	32	83.56	140.16	146.09	160.81	181.09	191.72	150.57
				Médias gerais	135	87,44	139,56	147,32	163,13	179,82	189,85	

2.2.7 Identificação de genes e anotação funcional da região gênica entre V16099 e V1190

Na região genômica delimitada pelos marcadores V16099 e V1190, marcadores de maiores efeitos no estudo de Silva et al. (2024), foram identificados 225 genes, dos quais mais de 75% codificam proteínas conservadas como ortólogos em outras espécies. Deste total, 16 genes apresentaram duplicação e dois genes em triplicação, com funções atribuídas a processos como regulação do ciclo celular, reparo de DNA e catabolismo proteico. Aproximadamente 31% dos genes codificantes permanecem não caracterizados, representando candidatos potenciais para envolvimento nos mecanismos de controle da altura da planta.

Para análise funcional direcionada, foram selecionados 13 genes associados a famílias gênicas envolvidas em morfogênese e controle hormonal com potencial regulação da altura (Tabela 8). Entre estes, destacam-se genes relacionados à sinalização de auxina, ácido jasmônico e ácido salicílico. Adicionalmente, os pseudogenes LOC113694124 e LOC113693263 mostraram associação a vias regulatórias do ciclo de divisão celular.

Tabela 8: Anotação funcional de genes com potencial envolvimento na expressão e regulação da altura de plantas na região entre os marcadores moleculares para V16099 e V1190.

Gene	Descrição	Função
LOC113693172	Bidirectional sugar transporter SWEET7b-like	Permite a atividade do transportador transmembrana de açúcar. Envolvido no transporte de carboidratos. Localizado na membrana (Ji <i>et al.</i> , 2022).
LOC113693177	Protein enhanced pseudomonas susceptibilty 1-like	Envolvido na regulação da resposta de defesa às bactérias e ao ácido jasmônico. Atua no processo biossintético do ácido salicílico (LI <i>et al.</i> , 2023).
LOC113693179	Protein RALF-like 27	Super expressão de Fator de Alcalinização Rápida (RALF) inibem o alongamento da raiz e sensibilidade aumentada ao estresse salino (Sharma <i>et al.</i> , 2016).
LOC113694124	Anaphase-promoting complex subunit 8-like	Envolvido em vários aspectos do desenvolvimento, controlando a divisão e o alongamento celular (Xu <i>et al.</i> , 2019).
LOC113693567	Polyubiquitin	Envolvido em processo catabólico de proteína dependente de modificação e ubiquitinação de proteína. Desempenha um papel essencial na sinalização de auxina. Localizado no citoplasma e núcleo (Doroodian & Hua, 2021).
LOC113693205	Senescence specific cysteine protease SAG39-like	Habilita atividade endopeptidase do tipo cisteína. Envolvida na proteólise e no processo catabólico de proteínas. Localizada no espaço extracelular e no lisossomo (WANG <i>et al.</i> , 2022).
LOC113691287	Abscisic acid-insensitive 5-like protein 2	Habilita a atividade do fator de transcrição de ligação ao DNA. Envolvido na regulação positiva da transcrição modelada do DNA. Localizado no núcleo (MARQUES <i>et al.</i> , 2024).
LOC113691881	Protein plant cadmium resistance 10-like	Cádmio, elemento não essencial que reduz as taxas de germinação das sementes e causa deficiências nutricionais e estresse oxidativo nas plantas, o que pode levar ao retardo do crescimento e até mesmo à morte das plantas (GUAN <i>et al.</i> , 2023).
LOC113693234	Heavy metal associated isoprenylated plant protein 45-like	Permite a ligação de íons metálicos, como ferro, manganês, magnésio, zinco. (BARR; WERNER; TILSNER, 2023).
LOC113692013	Probable lysophospholipase bodyguard 4	Envolvido no desenvolvimento da cutícula e morfogênese. Relações diretas com ácido abscísico (JAKOBSON <i>et al.</i> , 2016).
LOC113691904	Myb-related protein 306-like	Envolvido na resposta à auxina (MARQUES <i>et al.</i> , 2024).
LOC113691882	Plant intracellular Ras-group-related LRR protein 6-like	Permite a ligação às proteínas. Modula a morfogênese celular regulando a formação e montagem da parede celular e/ou a polarização do crescimento (T.G. Ajith & Shah, 2024).
LOC113693263	Cell division cycle protein 48 homolog pseudogene	Atua antes ou durante a divisão celular, regulação negativa da resposta de defesa, germinação do pólen, crescimento do tubo polínico, desestabilização de proteínas (INÈS <i>et al.</i> , 2024).

2.4 Discussão

O presente estudo estabeleceu uma base metodológica para a genotipagem de SNPs associados à altura em *C. arabica* por meio do desenvolvimento de marcadores SNP *low-throughput* usando as técnicas HRM (HRM-V16099-2, HRM-V1190-3, HRM-V42-2 e HRM-V16476-1) e Tetra-Primer ARMS (4P.ARMS-V16099-1 e 4P.ARMS-V1190-1). Estes marcadores representam um avanço para a implementação de seleção assistida em programas de melhoramento, permitindo a discriminação precoce de alelos associados a porte alto ou baixo de forma acessível e confiável.

Dos 49 marcadores SNP identificados por Silva et al. (2024), 16 atenderam aos critérios para customização de *primers*, e destes, seis resultaram em marcadores SNP *low-throughput*. Foram utilizadas diferentes metodologias de genotipagem para os SNP-alvos de maior efeito no GWAS (V16099 e V1190), o que foi estratégico para aumentar a probabilidade de sucesso na conversão de marcadores teóricos (identificados por GWAS) em marcadores para seleção assistida. Estas abordagens diversificadas também proporcionam diferentes modos de detecção, aumentando a aplicabilidade dos marcadores em programas de seleção assistida com diferentes investimentos tecnológicos (Zargar et al., 2016).

A genotipagem HRM demonstrou elevada eficiência, sensibilidade e escalabilidade, apresentando perfis de dissociação reprodutíveis e compatíveis com estudos anteriores em café (Bolívar-González et al., 2022; Combes et al., 2018; Gatica-Arias et al., 2023). Sua capacidade de *multiplex* em tempo real e o fato de não requerer sondas fluorogênicas tornam essa abordagem ideal para triagem inicial de grandes populações em plataformas de qPCR.

Por sua vez, a genotipagem por Tetra-Primer ARMS, embora enfrente desafios na confecção simultânea de quatro *primers* com parâmetros distintos em uma mesma reação de tubo fechado (Medrano & Oliveira, 2014), mostrou-se uma alternativa viável de baixo custo e visualização direta em géis de agarose (Schincariol, 2011). Em contraste, as metodologias rhAmp e MTPA não produziram resultados satisfatórios para os SNPs-alvo (V16099 e V1190), possivelmente devido a particularidades das técnicas e das regiões flanqueadoras.

A disponibilização desses marcadores *low-throughput* customizados representa uma alternativa estratégica às abordagens *high-throughput*, como SNP arrays e GBS. Embora essas tecnologias sejam eficazes para estudos de descoberta em larga escala (GWS, GWAS e mapeamento de QTL), demandam consideráveis investimentos financeiros e infraestrutura especializada (Arcanjo et al., 2024; Singh et al., 2022; Uffelman et al., 2021). Em programas de melhoramento que empregam a seleção assistida por marcadores para características

específicas, marcadores *low-throughput* proporcionam custos substancialmente menores, maior acessibilidade técnica e aplicação direcionada. Esses atributos tornam esse tipo de marcador especialmente útil para populações grandes, seleção de genótipos para avanço de geração e na identificação de piramidação de alelos. Nesses contextos, a necessidade de alta escalabilidade típica das plataformas de sequenciamento é substituída pela prioridade em precisão, confiabilidade e custo-efetividade da genotipagem (Grover & Sharma, 2016; Sahoo et al., 2025; Zargar et al., 2016).

Os marcadores desenvolvidos na população de mapeamento foram aplicados a uma outra população em melhoramento, aqui denominada população de testagem, com o objetivo de avaliar sua eficiência discriminatória em população geneticamente distinta. Nessa população de testagem observou-se a discriminação alélica clara e a determinação dos genótipos, confirmando a capacidade dos marcadores diferenciar os alelos de interesse mesmo em um novo contexto genômico.

Para investigar a associação entre marcadores e a altura das plantas na população de testagem foi realizada análise de marca simples. Os efeitos detectados foram inferiores aos observados na população original e não reproduziram os efeitos significativos previamente reportados por Silva et al. (2024). Essa ausência de replicação e os testes estatísticos empregados para os dados fenotípico sugerem que a população de testagem não apresenta contraste fenotípico suficiente para revelar a contribuição real dos marcadores. Tal resultado está alinhado a natureza complexa e poligênica da altura de plantas, com forte influência ambiental e interação entre múltiplos genes de pequeno efeito (Busov et al., 2008; Wang et al., 2017; Salgotra & Stewart Jr., 2020).

Além do pouco contraste da característica na população, essa divergência pode ser atribuída, em parte, às diferenças entre a população de testagem (Híbrido de Timor × Tupi Amarelo) e a população do estudo GWAS original (Catuaí × Híbrido de Timor). A mudança de *background* genético altera padrões de desequilíbrio de ligação e epistasia, modificando a expressão dos alelos favoráveis. Além disso, os genitores do cruzamento de testagem pertencem a grupos genéticos distintos, porém com distância genética moderada (Medeiros et al., 2021), o que pode ter contribuído para recombinações que diluíram ou alteraram os efeitos epistáticos originalmente observados.

Também reforça essa divergência a ancestralidade do Tupi Amarelo, portador dos loco *Ct* (Caturra) e de fatores genéticos de Villa Sarchi relacionados ao porte reduzido (Carvalho et al., 1984a; Carvalho & Monaco, 1972). Interações epistáticas de tipo duplo dominante, sugeridas

na literatura, podem ter sido atenuadas ao longo de sucessivas gerações de melhoramento, resultando em menor expressão fenotípica do porte reduzido mesmo na presença de alelos teoricamente favoráveis (Carvalho et al., 1984b). De maneira semelhante, linhagens Catuaí Vermelho e Amarelo tendem a ser mais altas que as descendências de Caturra Vermelho x HdT e Villa Sarchi x HdT, apesar de compartilharem o mesmo alelo *Ct* que confere porte reduzido às plantas ((Carvalho et al., 1984a; Silvarolla et al., 1997), reforçando a importância desse cenário genético na modulação da altura de plantas.

O GWAS estimou em aproximadamente 20% a contribuição dos SNPs V16099 e V1190 para a redução da altura em cafeeiros (SILVA et al., 2024). Considerando a análise descritiva por médias da F₂, revelou-se que genótipos homocigotos favoráveis proporcionaram redução média de 3% na altura (Tabela 6). Quando considerados em conjunto, arranjos genotípicos que combinam homocigotos favoráveis (TT - V16099 e CC - V1190) com heterocigotos (CT - V16099 e CT - V1190) resultaram em reduções de aproximadamente 11% na altura das plantas. A proximidade física desses marcadores no QTL (cerca de 9 Mb), associada às distribuições genotípicas observadas, sugere possível ligação gênica entre estes locos. Embora, a confirmação desta hipótese requer investigações adicionais em outras estruturas populacionais.

Do ponto de vista funcional, a região entre V16099 e V1190 contém genes envolvidos em sinalização hormonal, incluindo auxina, ácido jasmônico e ácido salicílico, que são essenciais para o desenvolvimento das plantas e, conseqüentemente, na expressão do fenótipo altura (Doroodian & Hua, 2021; Guan et al., 2023; Sharma et al., 2016; C. Wang et al., 2022). Além disso, os marcadores HRM-V15967.V15966-2 e HRM-V42-2 localizam-se em genes com funções biológicas relevantes - uma histidina-tRNA ligase organelar e uma proteína-quinase do tipo receptor rico em leucina, respectivamente - representando alvos promissores para estudos de associação (NCBI - National Center for Biotechnology Information, 2025).

As análises *in silico* revelaram-se fundamentais considerando a elevada similaridade entre os subgenomas de *C. arabica*, onde a homologia sequencial pode levar à amplificação não específica por pareamento inespecífico de *primers* (Guyot et al., 2020; Scalabrin et al., 2020, 2024). Embora regiões gênicas conservadas representem sítios preferenciais para o desenvolvimento de marcadores funcionais devido à sua estabilidade evolutiva, polimorfismos em regiões não codificantes apresentam igual relevância para aplicações em melhoramento genético. A eficácia de um marcador molecular é determinada primordialmente por sua capacidade de detectar variações genéticas entre indivíduos, e não por sua localização genômica (Hasan et al., 2021; Park et al., 2025; Salgotra & Stewart Jr, 2020).

Em programas de melhoramento, a seleção demonstra maior eficiência em estágios iniciais como F₂, nos quais existe altas taxas de segregação e diversidade alélica. A conversão dos SNPs identificados em marcadores diagnósticos viabiliza a seleção precoce baseada em genótipo, reduzindo significativamente o tempo necessário para identificação de linhagens superiores e maximizando o ganho por unidade de tempo (Salgotra & Stewart Jr, 2020).

2.5 Conclusão

Foram desenvolvidos e validados seis marcadores moleculares *low-throughput* (HRM-V16099-2, 4P.ARMS-V16099-1, HRM-V1190-3, 4P.ARMS-V1190-1, HRM-V42-2 e HRM-V16476-1) para quatro SNPs pertencentes ao QTL associado à altura, localizado no cromossomo 6 de *Coffea arabica* (V16099, V1190, V16476 e V42). Esses marcadores demonstraram potencial para aplicação em seleção assistida, sobretudo em abordagens que considerem múltiplos locos de forma integrada.

A análise de associação de marca simples em população diferente da que o marcador foi identificado e desenvolvido não evidenciou efeitos significativos isolados sobre a altura. Esses resultados sugerem que a expressão fenotípica do QTL pode ser modulada pelo *background* genético populacional, possivelmente mediada por alterações no desequilíbrio de ligação ou por interações epistáticas. Esses dados reforçam o caráter quantitativo e de arquitetura genética complexa da altura em café arábica.

A região genômica de aproximadamente 9 Mb entre os marcadores V16099 e V1190 revelou a presença de genes candidatos envolvidos em morfogênese e vias hormonais, sugerindo uma regulação em rede. Paralelamente, a alta influência ambiental observada corrobora que a altura é governada por múltiplos genes de efeito menor, sensíveis à interação com o ambiente.

A conversão desses SNPs em ferramentas de seleção assistida representa uma estratégia viável para acelerar o desenvolvimento de cultivares com porte definido. No entanto, sua aplicação deve considerar a validação prévia em diferentes contextos genéticos e a integração com dados fenotípicos multiambientais. Dessa forma, este estudo oferece tanto ferramentas moleculares aplicáveis quanto um aprofundamento da compreensão da base genética da altura, contribuindo para o melhoramento direcionado da arquitetura da planta em *C. arabica*.

2.6 Referências

- Aguiar, A. T. da E., Guerreiro-Filho, O., Maluf, M. P., Gallo, P. B., & Fazuoli, L. C. (2004). Caracterização de cultivares de *Coffea arabica* mediante utilização de descritores mínimos. *Bragantia*, 63(2), 179–192. <https://doi.org/10.1590/S0006-87052004000200003>
- Arcanjo, E. S., Nascimento, M., Azevedo, C. F., Caixeta, E. T., de Oliveira, A. C. B., Pereira, A. A., & Nascimento, A. C. C. (2024). Low-density marker panels for genomic prediction in *Coffea arabica* L. *Acta Scientiarum. Agronomy*, 47(1), e69698. <https://doi.org/10.4025/actasciagron.v47i1.69698>
- Bolívar-González, A., Molina-Bravo, R., Solano-Sánchez, W., Araya-Valverde, E., Ivamoto-Suzuki, S. T., Pereira, L. F. P., & Gatica-Arias, A. (2022). SNP markers found in non-coding regions can distinguish among low-variant genotypes of arabica and other coffee species. *Genetic Resources and Crop Evolution*, 70(4), 1215–1228. <https://doi.org/10.1007/S10722-022-01498-0>
- Brasil. (2024). *Ministério da Agricultura, Pecuária e Abastecimento*. Registro Nacional de Cultivares - RNC. http://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares_registradas.php
- Brasil. Ministério da Agricultura, Pecuária e Abastecimento. (2000). Instruções para execução dos ensaios de distinguibilidade, homogeneidade e estabilidade de cultivares de café (*Coffea* spp.). *Serviço Nacional de Proteção de Cultivares, Diário Oficial da União, Poder Executivo*, Brasília, DF. http://www.agricultura.gov.br/arq_editor/file/vegetal/registroautorizacoes/formularios%20prote%c3%a7%c3%a3o%20cultivares/cafe_us.doc
- Broman, K. W., Wu, H., Sen, S., Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19(7), 889–890. <https://doi.org/10.1093/bioinformatics/btg112>
- Busov, V. B., Brunner, A. M., & Strauss, S. H. (2008). Genes for control of plant stature and form. *New Phytologist*, 177(3), 589–607. <https://doi.org/10.1111/j.1469-8137.2007.02324.x>
- Carvalho, A., Fazuoli, L. C., & Costa, W. M. da. (1989). Melhoramento do cafeeiro: XLI. Produtividade do híbrido de timor, de seus derivados e de outras fontes de resistência a *Hemileia vastatrix*. *Bragantia*, 48(1), 73–86. <https://doi.org/10.1590/S0006-87051989000100007>
- Carvalho, A., Medina Filho, H. P., Fazuoli, L. C., & Costa, W. M. da. (1984a). Genética de *Coffea*: XXVI. Hereditariedade do porte reduzido do cultivar Caturra. *Bragantia*, 43, 443–458. <https://doi.org/10.1590/S0006-87051984000200015>
- Carvalho, A., Medina Filho, H. P., Fazuoli, L. C., & Costa, W. M. da. (1984b). Número de locos e ação gênica de fatores para porte pequeno em *Coffea arabica* L. *Bragantia*, 43(2), 425–442. <https://doi.org/10.1590/S0006-87051984000200014>

- Carvalho, A., & Monaco, L. C. (1972). Transferência do fator caturra para o cultivar Mundo Novo de *Coffea arabica*. *Bragantia*, 31, 379–399. <https://doi.org/10.1590/S0006-87051972000100031>
- Carvalho, A., Mônico, L. C., & Fazuoli, L. C. (1979). Melhoramento do cafeeiro: XL-estudos de progênies e híbridos de café catuaí. *Bragantia*, 38, 203–216. <https://doi.org/10.1590/S0006-8705197900010002>
- 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(1), 73–81. <https://doi.org/10.1007/s00299-008-0621-y>
- Clarindo, W. R., Carvalho, C. R., Caixeta, E. T., & Koehler, A. D. (2013). Following the track of “Híbrido de Timor” origin by cytogenetic and flow cytometry approaches. *Genetic Resources and Crop Evolution*, 60(8), 2253–2259. <https://doi.org/10.1007/s10722-013-9990-3>
- Combes, M. C., Joët, T., & Lashermes, P. (2018). Development of a rapid and efficient DNA-based method to detect and quantify adulterations in coffee (Arabica versus Robusta). *Food Control*, 88, 198–206. <https://doi.org/10.1016/j.foodcont.2018.01.014>
- Conab. Companhia Nacional de Abastecimento (2020). *Série Histórica de produção de café arábica no Brasil*. <https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras>
- Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., Zheng, C., Alberti, A., Anthony, F., Aprea, G., Aury, J. M., Bento, P., Bernard, M., Bocs, S., Campa, C., Cenci, A., Combes, M. C., Crouzillat, D., Da Silva, C., ... Lashermes, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*, 345(6201), 1181–1184. <https://doi.org/10.1126/science.1255274>
- Dias, R. E. B. A., Dias, R. A. A., Botelho, C. E., Abrahão, J. C. de R., Rezende, T. T., & Carvalho, G. R. (2020). Genetic determination of characteristics related to semi-mechanized coffee harvests. *Crop Breeding and Applied Biotechnology*, 20(1), 2020. <https://doi.org/10.1590/1984-70332020v20n1n16>
- Diniz, L. E. C., Sakiyama, N. S., Lashermes, P., Caixeta, E. T., Oliveira, A. C. B., Loureiro, M. E., Pereira, A. A., & Zambolim, L. (2005). Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. *Crop Breeding and Applied Biotechnology*, 5, 387–393. <https://doi.org/10.12702/1984-7033.v05n04a03>
- Doroodian, P., & Hua, Z. (2021). The ubiquitin switch in plant stress response. *Plants*, 10(2), 246. <https://doi.org/10.3390/plants10020246>
- Gatica-Arias, A., Bolívar-González, A., Sánchez-Barrantes, E., Araya-Valverde, E., & Molina-Bravo, R. (2023). High Resolution Melt (HRM) Genotyping for detection of induced mutations in coffee (*Coffea arabica* L. var. Catuaí). In *Mutation Breeding in Coffee with Special Reference to Leaf Rust: Protocols* (pp. 275–290). Springer Berlin Heidelberg Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-67273-0_20

- Grover, A., & Sharma, P. C. (2016). Development and use of molecular markers: Past and present. *Critical Reviews in Biotechnology*, 36(2), 290–302. <https://doi.org/10.3109/07388551.2014.959891>
- Guan, J., Yang, Y., Shan, Q., Zhang, H., Zhou, A., Gong, S., Chai, T., & Qiao, K. (2023). Plant cadmium resistance 10 enhances tolerance to toxic heavy metals in poplar. *Plant Physiology and Biochemistry*, 203, 108043. <https://doi.org/10.1016/j.plaphy.2023.108043>
- Guyot, R., Hamon, P., Couturon, E., Raharimalala, N., Rakotomalala, J.-J., Lakkanna, S., Sabatier, S., Affouard, A., & Bonnet, P. (2020). WCSdb: a database of wild *Coffea* species. *Database*, 2020. <https://doi.org/10.1093/database/baaa069>
- Henderson, C. R. (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423-449. <https://doi.org/10.2307/2529430>
- Huang, L., Wang, X., Dong, Y., Long, Y., Hao, C., Yan, L., & Shi, T. (2020). Resequencing 93 accessions of coffee unveils independent and parallel selection during *Coffea* species divergence. *Plant Molecular Biology*, 103(1–2), 51–61. <https://doi.org/10.1007/s11103-020-00974-4>
- Ju, Y., Feng, L., Wu, J., Ye, Y., Zheng, T., Cai, M., Cheng, T., Wang, J., Zhang, Q., & Pan, H. (2018). Transcriptome analysis of the genes regulating phytohormone and cellular patterning in Lagerstroemia plant architecture. *Scientific Reports*, 8(1), 15162. <https://doi.org/10.1038/s41598-018-33506-8>
- Lanfear, R., Ho, S. Y. W., Jonathan Davies, T., Moles, A. T., Aarssen, L., Swenson, N. G., Warman, L., Zanne, A. E., & Allen, A. P. (2013). Taller plants have lower rates of molecular evolution. *Nature Communications*, 4(1), 1879. <https://doi.org/10.1038/ncomms2836>
- Lashermes, P., Hueber, Y., Combes, M.-C., Severac, D., & Dereeper, 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. <https://doi.org/10.1534/g3.116.030858>
- Marques, I., Fernandes, I., Paulo, O. S., Batista, D., Lidon, F. C., Rodrigues, A. P., Partelli, F. L., DaMatta, F. M., Ribeiro-Barros, A. I., & Ramalho, J. C. (2024). Transcriptomic analyses reveal that *Coffea arabica* and *Coffea canephora* have more complex responses under combined heat and drought than under individual stressors. *International Journal of Molecular Sciences*, 25(14), 7995. <https://doi.org/10.3390/ijms25147995>
- Medeiros, A. C., Caixeta, E. T., Oliveira, A. C. B. de, Sousa, T. V., Stock, V. de M., Cruz, C. D., Zambolim, L., & 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. <https://doi.org/10.3389/fsufs.2021.705278>

- Medrano, R. F. V., & Oliveira, C. A. (2014). Guidelines for the Tetra-Primer ARMS-PCR Technique Development. *Molecular Biotechnology*. <https://doi.org/10.1007/s12033-014-9734-4>
- Mekbib, Y., Tesfaye, K., Dong, X., Saina, J. K., Hu, G.-W., & Wang, Q.-F. (2022). Whole-genome resequencing of *Coffea arabica* L. (Rubiaceae) genotypes identify SNP and unravels distinct groups showing a strong geographical pattern. *BMC Plant Biology*, *22*(1), 69. <https://doi.org/10.1186/s12870-022-03449-4>
- Montagnon, C., Mahyoub, A., Solano, W., & Sheibani, F. (2021). Unveiling a unique genetic diversity of cultivated *Coffea arabica* L. in its main domestication center: Yemen. *Genetic Resources and Crop Evolution*, *68*(6), 2411–2422. <https://doi.org/10.1007/s10722-021-01139-y>
- National Center for Biotechnology Information- NCBI. (2024, September 24). *Coffea arabica genome assembly (GCF_003713225.1)*. https://www.ncbi.nlm.nih.gov/datasets/genome/gcf_003713225.1
- Patterson, H. D., & Thompson, R. (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* *58*, 545-554. <https://doi.org/10.2307/2334389>
- R Core Team. (2018). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, 16. <https://www.R-project.org/>
- Resende, D. M. V., & Alves, R. S. (2022). Statistical significance, selection accuracy, and experimental precision in plant breeding. *Crop Breeding and Applied Biotechnology*, *22*(3), e42712238. <https://doi.org/10.1590/1984-70332022v22n3a31>
- Resende, M. D. V. de. (2016). Software Selegen-REML/BLUP: A useful tool for plant breeding. *Crop Breeding and Applied Biotechnology*, *16*(4), 330–339. <https://doi.org/10.1590/1984-70332016v16n4a49>
- Resende, M., Caixeta, E., & Alkimim, E. R. (2016). High-throughput targeted genotyping of *Coffea arabica* and *Coffea canephora* using next generation sequencing. *San Diego, CA, I*.
- Sahoo, J., Mishra, R., & Joshi, R. K. (2025). PCR-based single nucleotide polymorphism (SNP) genotyping for crop improvement-current status and future prospects. *Discover Plants 2025 2:1*, *2*(1), 1–19. <https://doi.org/10.1007/S44372-025-00262-9>
- Salgotra, R. K., & Stewart Jr, C. N. (2020). Functional markers for precision plant breeding. *International journal of molecular sciences*, *21*(13), 4792. <https://doi.org/10.3390/ijms21134792>
- Salojärvi, J., Rambani, A., Yu, Z., Guyot, R., Strickler, S., Lepelley, M., Wang, C., Rajaraman, S., Rastas, P., Zheng, C., Muñoz, D. S., Meidanis, J., Paschoal, A. R., Bawin, Y., Krabbenhoft, T. J., Wang, Z. Q., Fleck, S. J., Aussel, R., Bellanger, L., ... 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. <https://doi.org/10.1038/s41588-024-01695-w>
- 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. *Plants*, 255(6), 1–16. <https://doi.org/10.1007/s00425-022-03898-z>
- Scalabrin, S., Magris, G., Liva, M., Vitulo, N., Vidotto, M., Scaglione, D., Del Terra, L., Ruosi, M. R., Navarini, L., Pellegrino, G., Berny Mier y Teran, J. C., Toniutti, L., Suggi Liverani, F., Cerutti, M., Di Gaspero, G., & Morgante, M. (2024). A chromosome-scale assembly reveals chromosomal aberrations and exchanges generating genetic diversity in *Coffea arabica* germplasm. *Nature Communications*, 15(1), 463. <https://doi.org/10.1038/s41467-023-44449-8>
- Scalabrin, S., Toniutti, L., Di Gaspero, G., Scaglione, D., Magris, G., Vidotto, M., Pinosio, S., Cattonaro, F., Magni, F., Jurman, I., Cerutti, M., Suggi Liverani, F., Navarini, L., Del Terra, L., Pellegrino, G., Ruosi, M. R., Vitulo, N., Valle, G., Pallavicini, A., ... 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), 4642. <https://doi.org/10.1038/s41598-020-61216-7>
- Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. *Plant Biotechnology Journal*, 15(2), 149–161. <https://doi.org/10.1111/pbi.12645>
- Schincariol, L. P. (2011). *Identificação de marcadores snp pelas técnicas de pcr-rflp e tetra-primer arms-pcr e suas associações com qualidade de bebida em café* [DISSERTAÇÃO]. Instituto Agronômico de Campinas. https://www.iac.sp.gov.br/areadoinstituto/posgraduacao/repositorio/storage/teses_dissertacoes/pb1213309.pdf
- Setotaw, T. A., Caixeta, E. T., Zambolim, E. M., Sousa, T. V., Pereira, A. A., Baião, A. C., Cruz, C. D., Zambolim, L., & Sakiyama, N. S. (2020). Genome Introgression of Híbrido de Timor and its potential to develop high cup quality *C. arabica* cultivars. *Journal of Agricultural Science*, 12(4), 64. <https://doi.org/10.5539/jas.v12n4p64>
- Sharma, A., Hussain, A., Mun, B.-G., Imran, Q. M., Falak, N., Lee, S.-U., Kim, J. Y., Hong, J. K., Loake, G. J., Ali, A., & Yun, B.-W. (2016). Comprehensive analysis of plant rapid alkalization factor (RALF) genes. *Plant Physiology and Biochemistry*, 106, 82–90. <https://doi.org/10.1016/j.plaphy.2016.03.037>
- Silva, R. A., Caixeta, E. T., Silva, L. de F., Sousa, T. V., Barreiros, P. R. R. M., Oliveira, A. C. B. de, Pereira, A. A., Barreto, C. A. V., & Nascimento, M. (2024). Identification of SNP markers and candidate genes associated with major agronomic traits in *Coffea arabica*. *Plants*, 13(13), 1876. <https://doi.org/10.3390/plants13131876>

- Silvarolla, M. B., Guerreiro-Filho, O., Lima, M. M. A. & Fazuoli, L. C. (1997). Avaliação de progênies derivadas do Híbrido de Timor com resistência ao agente da ferrugem. *Bragantia*, 56(1), 47-58. <https://doi.org/10.1590/s0006-87051997000100005>
- Singh, A., Ramakrishna, G., Kaila, T., Saxena, S., Sharma, S., Gaikwad, A. B., Abdin, M. Z., & Gaikwad, K. (2022). Next-generation sequencing technologies: approaches and applications for crop improvement. *Genomics of Cereal Crops* (pp. 31-94). https://doi.org/10.1007/978-1-0716-2533-0_3
- Słomka, M., Sobalska-Kwapis, M., Wachulec, M., Bartosz, G., & Strapagiel, D. (2017). High resolution melting (HRM) for high-throughput genotyping—limitations and caveats in practical case studies. *International Journal of Molecular Sciences* 2017, Vol. 18, 2316, 18(11), 2316. <https://doi.org/10.3390/IJMS18112316>
- Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B. de, Pereira, A. A., Sakiyama, N. S., Resende Júnior, M. F. R. de, & 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 and Genomes*, 13(6), 1–16. <https://doi.org/10.1007/s11295-017-1208-y>
- Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B., Pereira, A. A., Sakiyama, N. S., Zambolim, L., & Resende, M. D. V. (2019). Early selection enabled by the implementation of genomic selection in *Coffea arabica* breeding. *Frontiers in Plant Science*, 9, 1934. <https://doi.org/10.3389/fpls.2018.01934>
- Thompson, R (1973) The estimation of variance and covariance components when records are subject to culling. *Biometrics* 29, 527-550. <https://doi.org/10.2307/2529174>
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1), 1–21. <https://doi.org/10.1038/s43586-021-00056-9>
- Wang, C., Chen, N., Liu, J., Jiao, P., Liu, S., Qu, J., Guan, S., & Ma, Y. (2022). Overexpression of ZmSAG39 in maize accelerates leaf senescence in *Arabidopsis thaliana*. *Plant Growth Regulation*, 98(3), 451–463. <https://doi.org/10.1007/s10725-022-00874-1>
- Wang, Y., Zhao, J., Lu, W., & Deng, D. (2017). Gibberellin in plant height control: old player, new story. *Plant Cell Reports*, 36(3), 391–398. <https://doi.org/10.1007/s00299-017-2104-5>
- Ye, S., Dhillon, S., Ke, X., Collins, A. R., & Day, I. N. M. (2001). An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Research*, 29(17), e88–e88. <https://doi.org/10.1093/nar/29.17.e88>
- Zargar, S. M., Raatz, B., Sonah, H., Muslimanazir, Bhat, J. A., Dar, Z. A., Agrawal, G. K., & Rakwal, R. (2016). Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. *Journal of Crop Science and Biotechnology* 2015 18:5, 18(5), 293–308. <https://doi.org/10.1007/S12892-015-0037-5>