

SELEÇÃO DE MARCADORES MICROSSATÉLITES PARA IDENTIFICAÇÃO GENÉTICA DE MUDAS DE *Coffea canephora* NO ESPÍRITO SANTO

Viviane Alexia Correia Silva, Aline dos Santos Bergamin, Stéffany de Freitas Badaró Rodrigues, Nayara Gonçalves Vallory, Adésio Ferreira, Marcia Flores da Silva Ferreira.

Universidade Federal do Espírito Santo/Programa de Pós-Graduação em Genética e Melhoramento, Alto Universitário, S/N, Guararema – 29500-000 – Alegre-ES, Brasil. vivianeac.correia@gmail.com, alinebergamin258@hotmail.com, steffany.fbr@gmail.com, nayaravallory@gmail.com, adesioferreira@gmail.com, mfloressf@gmail.com

Resumo

A espécie *Coffea canephora* possui grande importância econômica no Espírito Santo, sendo propagada a. O DNA *fingerprinting*, utilizando marcadores microssatélites (SSRs), permite a obtenção de perfis genéticos únicos, auxiliando na identificação e certificação dessas mudas, sendo uma importante ferramenta para prevenção de fraudes. Este estudo avaliou um conjunto de SSRs visando selecionar aqueles mais polimórficos, informativos e consistentes para aplicação em *fingerprinting*. Para isso, foram utilizados um conjunto de 46 SSRs em 12 indivíduos de *C. Canephora*. Dos 46 SSRs amplificados, 29 foram passíveis de genotipagem. Foram identificados 159 alelos, com média de 5,49 e número de alelos variando de 2 a 8. O conteúdo de informação polimórfica (PIC) médio foi de 0,717, indicando que os SSR avaliados, em sua maioria, apresentam alto poder discriminatório. Ao final, selecionou-se um conjunto de 23 marcadores SSR que apresentaram alta qualidade de amplificação, elevado polimorfismo e altos valores de PIC, possuindo potencial para serem utilizados na identificação genética e certificação de mudas clonais.

Palavras-chave: Café. Marcadores SSRs. DNA *Fingerprinting*.

Área do Conhecimento: Ciências Agrárias – Agronomia.

Introdução

A cafeicultura desempenha um papel fundamental na economia nacional e o Espírito Santo ocupa a posição de maior produtor de *C. canephora* do Brasil, produzindo cerca de 13.8 milhões de sacas beneficiadas, representando cerca de 69% da produção nacional (CONAB, 2025). Em escala comercial, as mudas de *Coffea canephora* são obtidas predominantemente por propagação vegetativa, especialmente via estaquia. Esse método assegura a formação de clones que preservam a constituição genética da planta matriz (Ferrão et al., 2019). Com o aumento da demanda por mudas de alta qualidade e o crescimento da comercialização desse material, torna-se essencial garantir a identificação genética precisa e confiável dos genótipos.

O DNA *fingerprinting* é uma metodologia que permite a identificação genética, auxiliando na distinção entre cultivares, variedades e linhagens a partir de um perfil genético único (Gazendam, Mojapelo e Bairu, 2022; Silva et al., 2024). A determinação do perfil genético de uma variedade de planta exige um conjunto de marcadores moleculares, que são que são ferramentas imparciais para caracterização de genótipos, já que não sofrem com interferências ambientais e nem dependem de fase de desenvolvimento das plantas (Jamali, Cockram, Hickey, 2018) e um método para detectá-los, como a eletroforese capilar ou em gel. Dentre eles, destacam-se os marcadores microssatélites ou SSRs (*Simple Sequence Repeats*) por seu alto polimorfismo, codominância, multialelismo e reprodutibilidade, sendo amplamente utilizados em estudos de diversidade genética, mapeamento genético, genômica funcional e identificação de cultivares (Marwal e Gaur, 2020).

Os SSRs são sequências de DNA curtas de 2 a 6 pares de bases repetidas em tandem (Marwal; Gaur, 2020), e variam, portanto, no número de nucleotídeos na unidade de repetição. De acordo com Testolin et al. (2022), são considerados critérios para seleção de SSRs para *fingerprinting*: motivos trinucleotídeos ou mais longos, quando possível; *locus* único; pelo menos dois loci não sobrepostos;

A Ciência do NANO e seu impacto transformador no MACRO

alto polimorfismo; baixo *stuttering*; reprodutibilidade consistente; *loci* com segregação independente distribuídos pelo genoma. SSRs com motivos longos (tri, tetra, penta e hexanucleotídeos) são preferidos para fingerprinting, pois a maior distância entre alelos facilita sua distinção (Wang et al., 2016).

Portanto, o objetivo deste trabalho foi testar um conjunto de 46 SSRs visando a seleção dos marcadores mais polimórficos, informativos e com boa qualidade de amplificação para serem utilizados em estudos de identificação de genótipos de *C. canephora*.

Metodologia

Foram selecionados 12 genótipos de *C. canephora* previamente caracterizados ampla variabilidade genotípica, conforme descrito previamente pelo Laboratório de Genética e Melhoramento Vegetal da Universidade Federal do Espírito Santo (UFES) (dados não publicados), sugerindo a possibilidade de serem genótipos híbridos. Devido a essa característica, esses materiais foram considerados apropriados para a etapa inicial de triagem dos marcadores microsatélites, por aumentarem a probabilidade de detecção de maior número de polimorfismos. Esses materiais foram obtidos na Fazenda Experimental de Marilândia (FEM), pertencente ao Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (Incaper).

Tabela 1 – Identificação dos 12 genótipos de *C. canephora* utilizados.

Identificação	Clone	Cultivar
1	BAG558	Não pertence
2	V03	Vitória – Incaper 8142
3	V07	Vitória – Incaper 8142
4	BAG110	Não pertence
5	BAG138	Não pertence
6	BAG328	Não pertence
7	BAG380	Não pertence
8	203	Jequitibá – ES 8122
9	A1	Andina, Plena, Tributun e Diamante ES8112
10	R2	Robustão Capixaba – EMCAPA 8141
11	R5	Robustão Capixaba – EMCAPA 8141
12	207	Jequitibá – ES 8122

As folhas coletadas tiveram seu DNA extraído segundo protocolo CTAB (brometo de cetil-trimetilamônio), descrita por Doyle & Doyle (1990), com modificações do IAC para otimização em café. As amostras extraídas foram diluídas para a concentração final de 20 ng.µL⁻¹ de DNA e armazenadas em freezer a -20 °C.

Foram testados 46 SSRs descritos na literatura que foram selecionados considerando as seguintes características: amplificação em estudos anteriores, altos valores de PIC (maior que 0,6) e tipo de motivo. As PCRs seguiram o protocolo da Sinapse Biotecnologia com volume final de 20 µL, contendo na reação: 40 ng/µL de DNA, 1X PCR Buffer Mg²⁺ plus, 1 U de Taq DNA Polimerase (Sinapse Biotecnologia®), 0,2 µM dos iniciadores *forward/reverse* e 0,2 mM de dNTP. As amplificações foram feitas no termociclador Veriti™ 96-Well Thermal Cycler (Thermo Fisher Scientific®) sob as seguintes condições: 1 ciclo de três minutos a 94°C; 35 ciclos de 40 segundos a 94 °C, um minuto a temperatura de anelamento entre 51 e 61°C (a depender da exigência do primer), um minuto a 72 °C; seguido de extensão final de 72 °C por 10 minutos. Os amplicons foram avaliados em gel de poliacrilamida a 10% corado com GelRed®. Foi feito dois controles negativos com água para todos os SSRs.

Os indivíduos para cada marcador foram genotipados, com codificação dos alelos por números inteiros consecutivos (ex: 1, 2, 3...). Cada genótipo foi representado por um par de números,

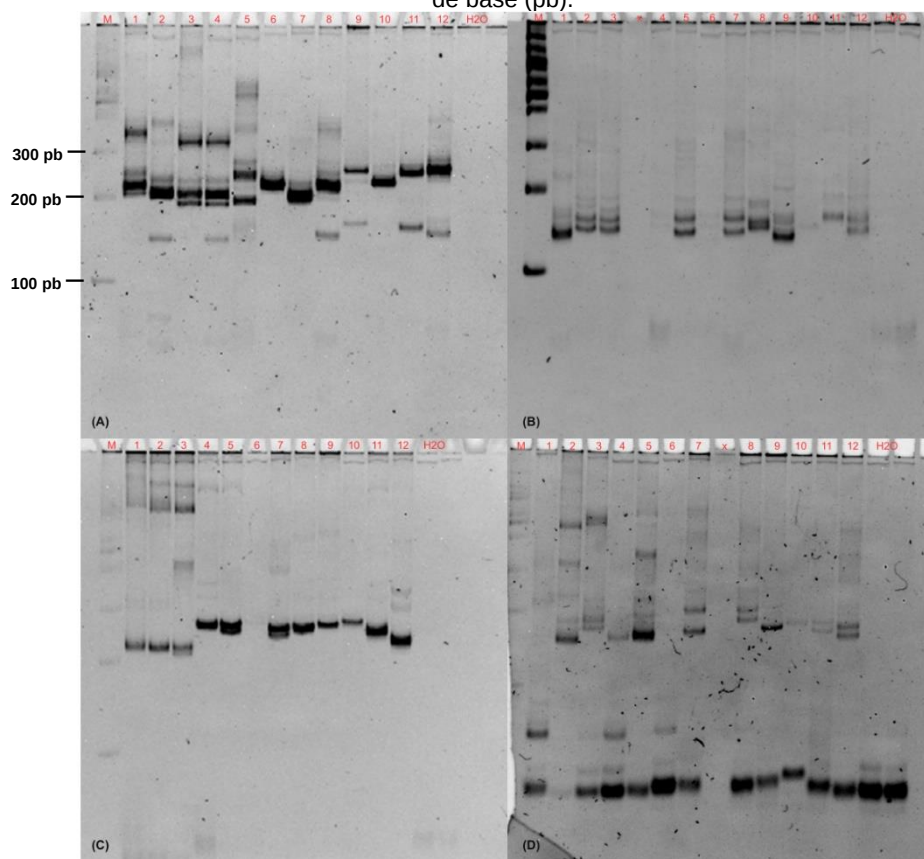
considerando a natureza diploide de *C. canephora*, sendo os pares idênticos indicativos de homozigose (ex: 11, 22) e os diferentes, de heterozigose (ex: 12, 23).

A análise foi realizada no RStudio (Posit Team, 2025), utilizando os seguintes pacotes: readxl (Wickham; Bryam, 2025), tidyverse (Wickham et al., 2019), pegas (Paradis, 2010) e adegenet (Jombart, 2008). Foram estimados os seguintes parâmetros para cada SSR: frequência alélica, número total de alelos (N), heterozigosidade observada (H_o), heterozigosidade esperada (H_e), conteúdo de informação polimórfica (PIC) e coeficiente de endogamia (F).

Resultados

Dos 46 SSRs analisados, 17 apresentaram bandas arrastadas ou ausência completa de amplificação, o que inviabilizou sua genotipagem. Os 29 marcadores restantes apresentaram resultados consistentes e passíveis de análise, dos quais 23 foram selecionados para o conjunto final e seis foram descartados (Tabela 2). A seleção considerou critérios como bandas nítidas, alto polimorfismo, baixa amplificação inespecífica, motivo do SSR e amplificação consistente na maioria das amostras. A Figura 1 ilustra quatro dos SSRs escolhidos, que atendem a esses parâmetros, destacando-se pelo elevado polimorfismo e perfis claros. Os demais marcadores selecionados apresentaram comportamento semelhante.

Figura 1 – Géis de eletroforese em gel de poliacrilamida a 10% de quatro SSRs: (A) CAM32 – (TA)₁₂; (B) CCRM14 – (CT)₂₂; (C) SSR46 – (GAT)₈; (D) CC1P11 – (AGG)₁₄. M = Marcador de peso molecular 100 pares de base (pb).



Fonte: Viviane Correia.

Na população analisada, foram identificados 159 alelos utilizando 29 marcadores SSRs, com média de 5,49 alelos, variando de 2 (SSR119) a 8 alelos (CaEST-002 e CAM22). A heterozigosidade observada variou de 0,083 (CAM18) a 1 (SSR84), com média de 0,480. A heterozigosidade esperada

A Ciência do NANO e seu impacto transformador no MACRO

apresentou valores entre 0,444 (SSR119) e 0,846 (CC1P11), com média de 0,737. O conteúdo de informação polimórfica (PIC) variou de 0,395 (SSR119) a 0,826 (CC1P11), com média de 0,717. O coeficiente de endogamia (F) mostrou ampla variação, oscilando entre -0,412 (SSR106) e 0,899 (CAM18), com média de 0,349.

Tabela 2 – Número de alelos (N), heterozigosidade observada (Ho), heterozigosidade esperada (He), conteúdo de informação polimórfica (PIC), coeficiente de endogamia (F) para 29 *locus* SSRs.

Locus	Motivo	N	Ho	He	PIC	F
CaEST-002 ¹	-	8	0,727	0,826	0,815	0,120
CAM22 ²	(AT)15	8	0,667	0,809	0,798	0,176
CC1P11 ³	(AGG)14	7	0,667	0,846	0,836	0,212
CAM27 ²	(TA)7 (GT)14	7	0,500	0,845	0,835	0,408
CAM03 ²	(AC)11	7	0,750	0,823	0,811	0,089
SSR46 ⁵	(GAT)8	7	0,182	0,818	0,806	0,778
CCRM40 ³	(CT)29	7	0,500	0,802	0,789	0,377
SSR34 ¹	(AC)10 (CA)6	7	0,500	0,800	0,788	0,375
CAM16 ²	(TC)11	7	0,333	0,799	0,786	0,583
CCRM16 ³	(TC)26	7	0,500	0,747	0,731	0,330
CAM18 ²	(TC)9	6	0,083	0,823	0,810	0,899
SSR84 ⁵	-	6	1	0,781	0,765	-0,280
CC2P4 ⁴	(AT)22	6	0,727	0,777	0,760	0,064
SSR35 ¹	(CA)9	6	0,400	0,750	0,731	0,467
CCRM17 ³	(CT)17	6	0,455	0,744	0,727	0,389
CAM32 ²	(TA)12	5	0,250	0,767	0,749	0,674
CAM55 ²	(GA)4 (G)4 (A)27	5	0,714	0,765	0,748	0,067
*CAM12 ²	(CAG)10	5	0,600	0,765	0,747	0,216
CCRM19 ³	(GA)26	5	0,333	0,747	0,728	0,554
*Mg_368 ⁶	(TG)13	5	0,800	0,745	0,723	-0,074
CAM33 ²	(A)13 -5bp (AG)18	5	0,364	0,678	0,649	0,463
CCRM14 ³	(CT)22	4	0,667	0,685	0,659	0,027
*CAM24 ²	(CCT) 5-87bp (CTG) 6	4	0,167	0,688	0,656	0,758
CCRM33 ³	(CT)12	4	0,545	0,678	0,646	0,195
*CAM41 ²	(TAAA)5	4	0,125	0,664	0,640	0,812
CC4P5 ⁴	(ACT)10	3	0,182	0,645	0,608	0,718
*CCRM42 ³	(GA)13	3	0,222	0,593	0,560	0,625
SSR106 ⁵	(CA)18	3	0,750	0,531	0,500	-0,412
*SSR119 ¹	(CAT)7	2	0,222	0,444	0,395	0,500
Média		5,483	0,480	0,737	0,717	0,349

*Marcadores não selecionados.

¹Silva et al. (2024); ²Hendre (2008); ³Hendre e Aggarwal (2014); ⁴Ogutu et al. (2016); ⁵Sousa et al. (2022);

⁶Alekcevetch (2013).

Discussão

Dos 29 SSRs genotipados, 23 apresentaram desempenho satisfatório e seis não foram selecionados (Mg_368, CAM12, CAM24, CAM41, CCRM42 e SSR119). Esse resultado evidencia que

nem todos os primers, mesmo aqueles desenvolvidos especificamente para *C. canephora*, apresentam boa performance. A eletroforese capilar, método de detecção amplamente utilizado em análises de *fingerprinting* (Wang et al., 2016), requer marcadores robustos e de fácil interpretação (Testolin et al., 2022), além de eficientes e adequados para multiplexação, que é uma estratégia que possibilita a análise simultânea de múltiplos loci e é essencial para reduzir custos e tempo de processamento (Marwal & Gaur, 2020). Assim, destaca-se a importância da triagem inicial, que garante que apenas os melhores marcadores sejam utilizados.

A baixa ou ausência de amplificação, bem como a presença de amplificação inespecífica e bandas arrastadas, podem estar relacionadas à quantidade, integridade e pureza da sequência-alvo e às condições de reação. Esses problemas também podem decorrer do desenho dos primers ou de fatores ligados à PCR, como composição da reação, temperatura de anelamento e número de ciclos utilizados na programação.

O número elevado de alelos totais (159) e a média de 5,48 alelos por *locus* evidenciam a presença de ampla variabilidade genética nos 12 indivíduos analisados. Entre os 29 marcadores avaliados, 21 apresentaram número de alelos igual ou superior a 5, reforçando a alta diversidade detectada. Esse valor médio é similar ao reportado por Silva et al. (2024), que encontraram 5,19 alelos por *locus* ao analisar 12 SSRs em 96 genótipos de *C. canephora*, e por Sousa et al. (2022), que observaram 2,9 alelos por *locus* em 44 genótipos utilizando 21 SSRs. Assim, mesmo em uma população de tamanho reduzido, os marcadores empregados mostraram-se altamente polimórficos e podem ser adequados para elaboração de perfis genéticos e identificação de genótipos. Apesar de apresentar apenas 3 alelos, o marcador SSR106 foi selecionado devido a qualidade da sua amplificação.

A heterozigosidade observada (H_o) corresponde à proporção de indivíduos heterozigotos em uma população e é um indicador da variação genética (Sousa et al., 2022). Entre os marcadores analisados, os SSR84, Mg_368 e SSR106 apresentaram H_o maior que H_e , configurando um excesso de heterozigotos nesse *locus*. Para os demais *loci*, o padrão foi inverso, com H_o consistentemente inferior à heterozigosidade esperada (H_e). De modo geral, a média de H_o (0,49) foi menor que a de H_e (0,738), revelando um déficit de heterozigotos.

O valor médio do PIC foi de 0,718, indicando que a maioria dos marcadores utilizados é altamente informativa e, portanto, adequada para análises de discriminação de genótipos. Segundo a classificação proposta por Botstein et al. (1980), valores de PIC superiores a 0,50 caracterizam *loci* altamente informativos; valores entre 0,25 e 0,50 indicam *loci* moderadamente informativos; e valores inferiores a 0,25 correspondem a *loci* pouco informativos. Nesse estudo, apenas o marcador SSR119 (PIC = 0,395) foi enquadrado como moderadamente informativo, enquanto todos os demais apresentaram alto poder discriminatório.

O coeficiente de endogamia (F) reflete o equilíbrio entre homozigotos e heterozigotos na população (Sousa et al., 2022). Alguns marcadores, como SSR84, Mg_368 e SSR106, apresentaram F negativo e H_o maior que H_e , evidenciando excesso de heterozigotos nesses loci e sugerindo que os alelos não estão sendo fixados pela endogamia (Sousa et al., 2023). Em contraste, os marcadores CAM18, SSR46, CAM41 e CC4P5 exibiram altos valores de F , indicando predominância de homozigotos.

Conclusão

O conjunto final de 23 marcadores SSR selecionados apresentaram alta qualidade de amplificação, elevado polimorfismo e valores de PIC altos, tornando-os adequados para a identificação genética e certificação de mudas clonais de *C. canephora*. A triagem realizada neste estudo foi essencial para eliminar marcadores pouco informativos, reduzindo tempo e custos em análises futuras. Estes marcadores têm potencial para serem aplicados em estudos por eletroforese capilar, o que permitirá a investigação da eficiência e o impacto do motivo SSR na detecção de polimorfismos.

Referências

ALEKCEVETCH, J. C. et al. Estudo da diversidade genética de uma população de *Coffea canephora* var. conilon por meio de marcadores moleculares do tipo SSR. 2013.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da safra brasileira de café, Brasília, DF, v.12, n. 3, terceiro levantamento, setembro 2025.

FERRÃO, R. G. et al. Conilon Coffee: The *Coffea canephora* produced in Brasil. 3rd Updated and expanded. Vitória, ES: 2019, 974 p.

GAZENDAM, I.; MOJAPPELO, P.; BAIRU, M. W. Potato cultivar identification in south africa using a custom SNP panel. **Plants**, v. 11, n. 12, p. 1546, 2022.

HENDRE, P. S.; AGGARWAL, R. K. Development of genic and genomic SSR markers of robusta coffee (*Coffea canephora* Pierre Ex A. Froehner). **PloS one**, v. 9, n. 12, p. e113661, 2014.

HENDRE, P. S. et al. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. **BMC plant biology**, v. 8, n. 1, p. 51, 2008.

JAMALI, S. H.; COCKRAM, J.; HICKEY, L. T. Insights into deployment of DNA markers in plant variety protection and registration. **Theoretical and Applied Genetics**, v. 132, p. 1911-1929, 2019.

JOMBART, T. adegenet: a R package for the multivariate analysis of genetic markers. **Bioinformatics**, v. 24, n. 11, p. 1403-1405, 2008.

MARWAL, A.; GAUR, R. K. Molecular markers: tool for genetic analysis. In: Animal biotechnology. **Academic Press**, 2020. p. 353-372.

OGUTU, C. et al. Characterization and utilization of microsatellites in the *Coffea canephora* genome to assess genetic association between wild species in Kenya and cultivated coffee. **Tree Genetics & Genomes**, v. 12, n. 3, p. 54, 2016.

PARADIS, E. pegas: an R package for population genetics with an integrated–modular approach. **Bioinformatics**, v. 26, n. 3, p. 419-420, 2010.

POSIT TEAM. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA, 2025. Disponível em: <<http://www.posit.co/>>.

SILVA, L. de F et al. Fingerprinting Amazonian coffees: assessing diversity through molecular markers. **Euphytica**, v. 220, n. 2, p. 28, 2024.

SOUSA, P. et al. *Coffea canephora*: Heterotic crosses indicated by molecular approach. **Plants**, v. 11, n. 22, p. 3023, 2022.

TESTOLIN, R. et al. SSR-based DNA fingerprinting of fruit crops. **Crop Science**, v. 63, n. 2, p. 390-459, 2023.

WANG, R. J.; GAO, X. F.; KONG, X. R.; YANG, J. An efficient identification strategy of clonal tea cultivars using long-core motif SSR markers. **SpringerPlus**, v. 5, n. 1, p. 1152, 2016.

WICKHAM, H. et al. Package 'readxl'. Version, v. 13, n. 1, 2019.

WICKHAM H. et al. Welcome to the Tidyverse. **Journal of open source software**, v. 4, n. 43, p. 1686, 2019.

Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES) – projeto número 290/2024, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Secretaria de Estado da Agricultura, Abastecimento, Aquicultura e Pesca (SEAG), à Universidade Federal do Espírito Santo (UFES) pelo apoio financeiro a este estudo e ao Laboratório de Genética e Melhoramento Vegetal e Laboratório de Biometria da UFES pela estrutura.