AVALIAÇÃO DE METODOLOGIAS DE EXTRAÇÃO DE DNA DE ALTA QUALIDADE PARA *Coffea Canephora*

Isabella Sobreira Alexandre Assis[1], Matheus Alves Silva[1], Maria Clara Bolari Ramos[1], Maíra Silva Cattem[1], Adésio Ferreira[1], Marcia Flores da Silva Ferreira[1].

[1]Universidade Federal do Espírito Santo (UFES)/Centro de Ciências Agrárias e Engenharias (CCAE), Laboratório de Genética, Alto Universitário, s/n, Guararema - 29500-000 - Alegre-ES, Brasil, isabella.assis@edu.ufes.br, alvesmatheuss21@gmail.com, mariabolari1@hotmail.com, maira_cattem@hotmail.com, adesioferreira@gmail.com, mfloressf@gmail.com.

Resumo

A obtenção de DNA genômico de alta qualidade é uma exigência para o sequenciamento de nova geração (NGS), sobretudo em espécies com elevado teor de polissacarídeos e compostos fenólicos, como o cafeeiro. Os protocolos de extração descritos na literatura são variáveis quanto aos componentes e procedimentos, o que dificulta a escolha do método mais adequado para cada espécie. Neste trabalho objetivou-se comparar três protocolos de extração de DNA baseados em brometo de cetiltrimetilamônio (CTAB) — Doyle e Doyle modificado; um protocolo adaptado pela UFSC para sequenciamento por nanoporos e o protocolo institucional CIMMYT_B/GRP_PR0002 — e um kit comercial de coluna de sílica (NucleoSpin Plant II), quanto ao rendimento e à integridade do DNA de amostras foliares de *Coffea canephora*. O DNA foi quantificado por espectrofotometria (NanoDrop) e sua integridade analisada por eletroforese em gel de agarose a 0,8%. Embora o protocolo Doyle e Doyle-IAC tenha apresentado a maior concentração média (1.460,5 ng/µL), os perfis eletroforéticos revelaram retenção de material nos poços e arraste (*smear*), indicando presença de contaminações residuais. Em contrapartida, o protocolo adaptado pela UFSC proporcionou o melhor desempenho qualitativo apesar da menor concentração média (269,0 ng/µL), produziu bandas bem definidas de alto peso molecular, sem retenção nos poços e com elevadas razões de pureza (A260/280 = 1,91). O método CIMMYT e o kit comercial apresentaram baixa pureza e rendimento insuficiente, respectivamente. Conclui-se que o protocolo adaptado para sequenciamento por nanoporos pela UFSC foi o mais adequado para a extração de DNA de alta qualidade em *C. canephora*.

Palavras-chave: Extração de DNA. CTAB. *Coffea canephora*. Sequenciamento de nova geração. Qualidade de DNA.

Área do Conhecimento: Ciências Agrárias - Agronomia

Introdução

O avanço dos estudos genômicos e genéticos em café, impulsionado pelo sequenciamento do genoma de referência de *C. canephora* (Denoeud et al., 2014; Salojärvi et al., 2024), tem ampliado as ferramentas disponíveis para o melhoramento genético da espécie, o que torna a obtenção de DNA genômico de boa qualidade uma etapa cada vez mais relevante para a genética vegetal.

A alta qualidade do DNA é uma exigência para o sequenciamento de nova geração (NGS): as bibliotecas de sequenciamento demandam DNA de alto peso molecular, íntegro e livre de contaminantes como proteínas, RNA, polissacarídeos e compostos fenólicos, que podem inibir as enzimas utilizadas no preparo das bibliotecas (Healey et al., 2014). Em espécies vegetais ricas em polissacarídeos e polifenóis, como o cafeeiro, essas substâncias precipitam com o DNA durante sua extração, reduzindo o rendimento e a pureza do material obtido (Silva, 2010; Aboul-Maaty; Oraby, 2019).

Prevenir a precipitação desses contaminantes com o DNA continua sendo um desafio durante as extrações de DNA. As abordagens de sequenciamento de nova geração, como o sequenciamento por síntese da Illumina, utilizam polimerases (Chen, 2014), que são inibidas por polissacarídeos.

Os protocolos de extração de DNA descritos na literatura são variáveis quanto aos componentes químicos e procedimentos empregados. A maioria deles deriva do método de extração à base de CTAB descrito por Doyle e Doyle (1987), posteriormente modificado por diferentes grupos de pesquisa e instituições para adequação a espécies e finalidades específicas, ao passo que kits comerciais baseados em colunas de sílica oferecem alternativas mais padronizadas, ainda que nem sempre igualmente eficientes para tecidos ricos em metabólitos secundários (Habibi et al., 2022). Portanto, é necessário testar e comparar protocolos antes de sua adoção, considerando que a escolha pode depender da técnica de sequenciamento e tipo de material e espécie.

Diante disso, este trabalho teve como objetivo avaliar três protocolos de extração de DNA baseados em CTAB — todos fundamentados no método de Doyle e Doyle (1987), com modificações — e um kit comercial de coluna de sílica, comparando o rendimento e a integridade do DNA extraído de folhas de *C. canephora*, a fim de identificar o protocolo mais adequado à obtenção de DNA de alta qualidade visando sua utilização em sequenciamento de nova geração.

Metodologia

Foram utilizadas amostras de folhas jovens liofilizadas de dez genótipos de *Coffea canephora*. De cada amostra 100 mg de tecido vegetal foram macerados no MagNALyser previamente a etapa de lise. As extrações de DNA foram realizadas no Laboratório de Genética do Centro de Ciências Agrárias e Engenharias (CCAE) da Universidade Federal do Espírito Santo (UFES), campus de Alegre-ES.
Os quatro protocolos testados foram:

i) *Doyle e Doyle-IAC modificado*: baseado no método de Doyle e Doyle (1987), com modificações propostas pelo Instituto Agronômico de Campinas (IAC) para plantas com alta concentração de polissacarídeos, utilizando tampão de extração à base de CTAB 2%, PVP-40 1%, NaCl 1,4 M, Tris-HCl 100 mM, EDTA 20 mM, β-mercaptoetanol 0,2% e proteinase K, seguido de purificação com clorofórmio:álcool isoamílico (24:1) e precipitação do DNA com isopropanol e acetato de amônio.

(ii) *UFSC*: protocolo adaptado para preparo de amostras destinadas ao sequenciamento por nanoporos, também baseado em tampão CTAB 2% (NaCl 1,4 M, Tris-HCl 100 mM, EDTA 20 mM, PVP-40 1%, β-mercaptoetanol 0,2% e proteinase K), com etapas adicionais de purificação com clorofórmio:álcool isoamílico, sendo a última após a adição de RNase, e precipitação com isopropanol gelado e acetato de amônio *overnight*.

(iii) *CIMMYT* (protocolo institucional CIMMYT_B/GRP_PR0002, do International Maize and Wheat Improvement Center): tampão CTAB 1% com NaCl 700 mM, Tris-HCl 100 mM, EDTA 50 mM, PVP 1% e β-mercaptoetanol 140 mM, purificação com clorofórmio:octanol (24:1) e precipitação com isopropanol gelado.

(iv) *Kit comercial*: NucleoSpin Plant II, Mini kit for DNA from Plants (Macherey-Nagel), baseado em coluna de sílica, utilizado conforme instruções do fabricante (Macherey-Nagel, 2014) nas variantes de tampão de lise PL1 (com e sem adição de 1% de PVP) e PL2/3, aplicado como ensaio piloto em duas das dez amostras (3 e 8).

A concentração e a pureza do DNA extraído por cada protocolo foram determinadas por espectrofotometria em NanoDrop, calculando-se as razões de absorbância A260/280, indicadora de contaminação por proteínas, e A260/230, indicadora de contaminação por compostos fenólicos, polissacarídeos e sais residuais do tampão de extração (Thermo Fisher Scientific, 2007). A integridade do DNA genômico foi avaliada por eletroforese em gel de agarose a 0,8%.

Os valores de concentração e das razões de pureza obtidos para as dez amostras em cada um dos três protocolos baseados em CTAB foram organizados e comparados por meio de estatística descritiva (média e desvio padrão). Os resultados do kit comercial, obtidos para apenas duas amostras em caráter de ensaio piloto, foram descritos individualmente.

Resultados

As análises por espectrofotometria demonstraram variações marcantes entre os métodos baseados em CTAB (Tabela 1).

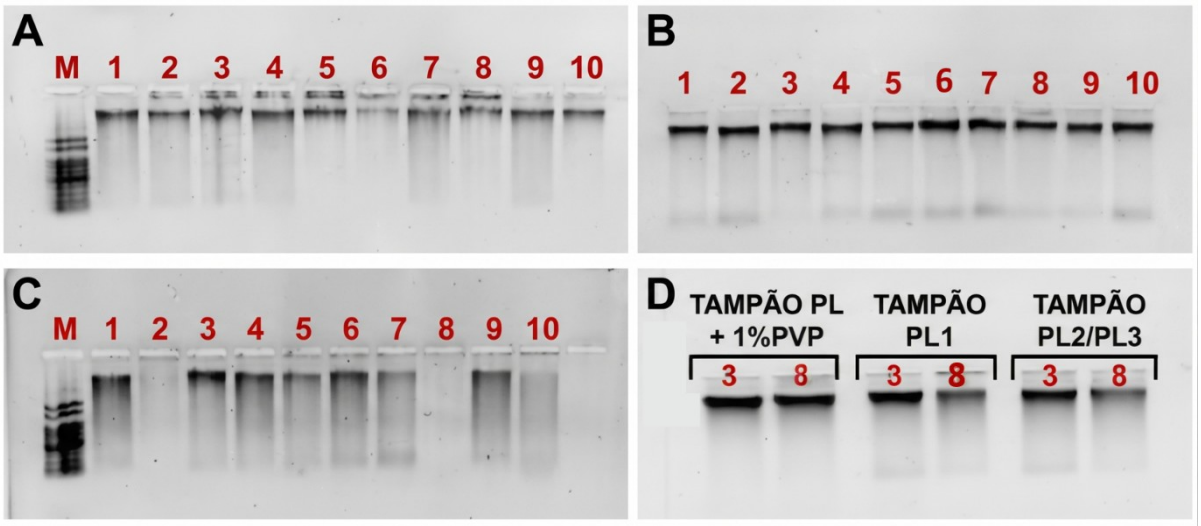
Tabela 1 - Concentração (ng/µL) e razões de pureza A260/280 e A260/230 do DNA genômico de dez amostras de *C. canephora*, obtidas por espectrofotometria (NanoDrop) para os três protocolos de extração baseados em CTAB.

Amostra	Doyle e Doyle-IAC			UFSC			CIMMYT		
	Conc. (ng/µL)	A260/ 280	A260/ 230	Conc. (ng/µL)	A260/ 280	A260/ 230	Conc. (ng/µL)	A260/ 280	A260/ 230
1	1379,6	2,00	2,10	174,5	1,90	1,90	648,6	1,60	0,80
2	1588,3	2,10	2,20	170,1	1,90	2,10	590,1	1,70	0,80
3	2249,8	2,00	2,00	282,7	1,80	1,30	680,7	1,50	0,60
4	927,0	1,90	2,10	323,8	1,90	2,00	845,5	0,80	0,80
5	2399,9	2,10	2,30	188,3	1,90	2,10	953,2	1,60	0,70
6	842,3	1,90	1,90	398,8	1,90	1,70	887,8	1,60	0,70
7	1033,5	2,10	2,30	292,3	2,00	2,00	931,2	1,70	0,90
8	1438,6	2,10	2,30	372,5	1,90	1,80	468,6	1,30	0,44
9	1871,8	2,10	2,30	297,6	2,00	2,00	1384,6	1,80	1,10
10	874,0	1,90	2,10	189,1	1,90	1,90	570,0	1,40	0,50
Média±DP	**1460±567**	**2,02±0,09**	**2,16±0,14**	**269±84**	**1,91±0,06**	**1,88±0,24**	**796±266**	**1,50±0,29**	**0,73±0,19**

Fonte: dados da pesquisa (2026).

O protocolo Doyle e Doyle-IAC forneceu as maiores concentrações de DNA (média de 1.460,5 ng/µL). Contudo, a análise do perfil eletroforético (Figura 1A) revelou forte retenção de material no poço de aplicação e arraste (*smear*), evidenciando a extração indesejada de compostos celulares juntamente com o DNA. O protocolo CIMMYT (Figura 1C) apresentou degradação e baixas razões de pureza (A260/280 = 1,50; A260/230 = 0,73). O kit comercial (Figura 1D) forneceu rendimentos muito baixos (10,4 a 28,0 ng/µL), embora o uso do Tampão PL1 + 1% PVP tenha mantido melhor integridade como observado no gel de agarose 0,8%.

Figura 1. Perfil eletroforético em gel de agarose (0,8%) do DNA genômico de *C. canephora* submetido a diferentes protocolos de extração. **(A)** Doyle e Doyle modificado pelo IAC (amostras 1 a 10); **(B)** Protocolo UFSC adaptado para nanoporos (amostras 1 a 10), demonstrando bandas íntegras sem retenção nos poços; **(C)** Protocolo CIMMYT (amostras 1 a 10); **(D)** Ensaio piloto com o kit comercial NucleoSpin Plant II nas amostras 3 e 8 testando variações de tampão de lise (PL + 1% PVP, PL1, PL2/PL3). **M:** marcador de peso molecular



Por outro lado, o protocolo CTAB adaptado para sequenciamento por nanoporos (UFSC) (Figura 1B) destacou-se como o melhor método. Apesar de resultarem em menor quantidade total de DNA (269,0 ng/µL em média), as amostras apresentaram excelente grau de pureza (A260/280 = 1,91), ausência de retenção residual nos poços e bandas de alto peso molecular bem definidas e íntegras, sem migração retardada por poluentes. A concentração mais baixa é compensada pela qualidade superior do DNA, parâmetro necessário para o sucesso no sequenciamento por nanoporos e demais plataformas NGS.

Discussão

Embora os três protocolos avaliados sejam derivados do método de Doyle e Doyle (1987) e compartilhem os componentes básicos do tampão de extração: CTAB, NaCl, Tris-HCl, EDTA e β-mercaptoetanol, diferem quanto à concentração desses reagentes, ao número de etapas de purificação com solvente orgânico, velocidade de centrifugação, ao momento de aplicação da RNase e ao tempo de precipitação do DNA. Essas diferenças, mais do que a composição química do tampão, permitem explicar o diferente desempenho dos protocolos quanto ao rendimento e à pureza do DNA obtido. Revisões recentes sobre o método CTAB apontam que praticamente todas as etapas do protocolo original foram alvo de modificações por diferentes grupos de pesquisa, e que a falta de detalhamento dessas alterações nas publicações contribui para a inconsistência de resultados relatada na literatura (Schenk et al., 2023).

Prevenir a coprecipitação de polissacarídeos e compostos fenólicos ainda é um dos maiores e mais importantes desafios durante as extrações de DNA para o NGS. As abordagens de sequenciamento de nova geração, como o sequenciamento por síntese da Illumina (San Diego, Califórnia, EUA), utilizam polimerases (Chen, 2014), que são inibidas por polissacarídeos.

Valores de A260/280 entre 1,8 e 2,0 e de A260/230 entre 2,0 e 2,2 são considerados indicativos de DNA puro, livre de contaminação relevante por proteínas ou por compostos residuais do processo de extração, como fenol, sais e polissacarídeos (Thermo Fisher Scientific, 2007). Assim sendo, os protocolos Doyle e Doyle (IAC) e Modificado para sequenciamento por nanoporos (UFSC) foram os que mais se aproximaram do padrão de qualidade recomendado entre os protocolos avaliados, demonstrando boas razões de pureza.

O protocolo Doyle-Doyle (IAC) apresentou um rendimento elevado, o que é desejável para o preparo de bibliotecas de sequenciamento de nova geração, que costuma exigir quantidades da ordem de microgramas de DNA de alto peso molecular (Healey et al., 2014), entretanto o gel de agarose das amostras obtidas por esse método apresentaram arraste e retenção no poço de aplicação, indicativos de contaminação e degradação do material.

No protocolo Doyle-Doyle (IAC), a RNase é adicionada apenas na etapa final de ressuspensão do DNA, e não entre as rodadas de purificação com clorofórmio:álcool isoamílico, de modo que o RNA permanece em solução durante toda a etapa de precipitação com isopropanol, podendo coprecipitar com o DNA e contribuir para o arraste observado no gel. A proteinase K, por sua vez, atua durante a lise inicial, degradando histonas e outras proteínas ligadas à cromatina (Schenk et al., 2023), mas a remoção efetiva dos fragmentos proteicos gerados depende da eficiência das extrações subsequentes com solvente orgânico, o que parece não ter sido suficiente neste protocolo para evitar a retenção de material no poço de aplicação.

O protocolo UFSC, embora tenha apresentado boa razão A260/280, resultou em rendimento menor. Essa diferença pode estar relacionada à coleta de quantidade menor de sobrenadante para evitar a interfase, adaptado originalmente para preparo de bibliotecas de sequenciamento por nanoporos, que costuma priorizar a pureza e a integridade das moléculas em detrimento do rendimento bruto. O protocolo do CIMMYT, apesar de ter fornecido concentrações intermediárias, apresentou as menores razões de pureza entre os três protocolos, sobretudo A260/230, o que sugere remoção incompleta de polissacarídeos, compostos fenólicos ou resíduos do tampão de extração, problema já relatado para protocolos CTAB aplicados a tecidos vegetais ricos nesses compostos, quando as etapas de purificação são insuficientes (Silva, 2010; Aboul-Maaty; Oraby, 2019).

A maior pureza obtida com o protocolo UFSC pode ser atribuída a um conjunto de ajustes incorporados em relação ao Doyle e Doyle-IAC. Primeiro, a coleta de um volume fixo e reduzido de sobrenadante (três alíquotas de 180 µL) evita a interface entre as fases aquosa e orgânica, um dos passos mais críticos da extração à base de CTAB, já que a perturbação dessa interface é apontada como uma das principais causas de carreamento de proteínas desnaturadas e polissacarídeos para o extrato final (Schenk et al., 2023), o que também explica, a menor concentração de DNA recuperada, uma vez que no momento da coleta, descarta-se uma quantidade considerável de sobrenadante. Segundo, a digestão com RNase é realizada entre duas rodadas de purificação com clorofórmio:álcool isoamílico, e não apenas ao final da extração como no protocolo IAC, permitindo a remoção do RNA antes da etapa de precipitação do DNA. Terceiro, a precipitação em isopropanol gelado ocorre por período prolongado (overnight, a −20 °C), seguida de lavagens sucessivas com etanol 70% e depois 95%, o que favorece a remoção mais completa de sais e de CTAB residual, um composto capaz de

inflar artificialmente a leitura de A260 quando carreado para a amostra final (Schenk et al., 2023). Esse conjunto de ajustes é compatível com a lógica de protocolos otimizados para sequenciamento de terceira geração, nos quais a integridade de moléculas de alto peso molecular é priorizada mesmo às custas do rendimento total.

As menores razões de pureza obtidas com o protocolo CIMMYT também podem ser relacionadas a características de seu desenho metodológico, nesse protocolo, a concentração de CTAB no tampão corresponde à metade da utilizada nos demais protocolos (1% x 2%), é realizada apenas uma etapa de purificação com clorofórmio:octanol, sendo uma segunda etapa apenas opcional, a centrifugação ocorre a uma velocidade inferior (3.700 rpm) e não há emprego de proteinase K em nenhuma etapa. Concentrações mais baixas de CTAB tendem a ser menos eficientes na remoção de polissacarídeos (Schenk et al., 2023; Lade et al., 2014), e a ausência de proteinase K, combinada à apenas duas extrações orgânicas e a uma menor velocidade de centrifugação, reduz a eficiência de remoção de proteínas e outros contaminantes, fatores que, em conjunto, ajudam a explicar as baixas razões A260/280 e, sobretudo, A260/230 observadas nesse protocolo.

As bandas de baixo peso molecular observadas na parte inferior do gel de agarose são compatíveis com a presença de RNA residual, o que não compromete necessariamente a utilização do DNA em sequenciamento, mas indica a necessidade de uma etapa adicional ou mais eficiente de tratamento com ribonuclease A (RNase A) antes do preparo de bibliotecas, de modo a evitar interferência na quantificação e na normalização das amostras.

O kit comercial testado em caráter piloto (NucleoSpin Plant II) forneceu concentrações de DNA muito inferiores às obtidas pelos protocolos baseados em CTAB, mesmo na variante com adição de PVP, recomendada pelo fabricante para tecidos ricos em compostos fenólicos (Macherey-Nagel, 2014). Resultado semelhante foi relatado por Habibi et al. (2022) ao comparar protocolos CTAB e kits comerciais de coluna de sílica em espécie lenhosa rica em polissacarídeos e polifenóis, na qual o método CTAB também apresentou o melhor desempenho para sequenciamento de nova geração. Ainda assim, o baixo rendimento do kit neste estudo pode estar relacionado à pequena massa de tecido inicial recomendada pelo protocolo do fabricante, e não necessariamente à sua eficácia intrínseca, o que sugere a necessidade de testes adicionais antes de descartar essa alternativa.

Conclusão

O protocolo adaptado para nanoporos (UFSC) apresentou o melhor equilíbrio entre rendimento e pureza do DNA genômico extraído de folhas de *C. canephora*, sendo recomendado para a obtenção de DNA de alta qualidade visando aplicações em sequenciamento de nova geração.

Referências

ABOUL-MAATY, N. A. F.; ORABY, H. A. S. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. **Bulletin of the National Research Centre**, v. 43, n. 25, 2019.

CHEN, C.-Y. DNA polymerases drive DNA sequencing-by-synthesis technologies: Both past and present. **Frontiers in Microbiology** 5: 305, 2014. doi: 10.3389/fmicb.2014.00305.

DENOEUD, F. et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, v. 345, n. 6201, p. 1181-1184, 2014.

DOYLE, J. J.; DOYLE, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochemical Bulletin**, v. 19, n. 1, p. 11-15, 1987.

HABIBI, N. et al. Comparison and optimization of DNA isolation protocols for high throughput genomic studies of *Acacia pachyceras* Schwartz. **MethodsX**, v. 9, 101799, 2022.

HEALEY, A.; FURTADO, A.; COOPER, T.; HENRY, R. J. Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. **Plant Methods**, v. 10, n. 21, 2014.

LADE, B. D.; PATIL, A. S.; PAIKRAO, H. M. Efficient genomic DNA extraction protocol from medicinal rich *Passiflora foetida* containing high level of polysaccharide and polyphenol. **SpringerPlus**, v. 3, n. 457, 2014.

MACHEREY-NAGEL. **NucleoSpin Plant II, Genomic DNA Purification: user manual.** Düren: Macherey-Nagel, 2014.

SALOJÄRVI, J. et al. The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. **Nature Genetics**, v. 56, n. 4, p. 721–731, 2024. DOI: https://doi.org/10.1038/s41588-024-01695-w. Acesso em: 01 set. 2026.

SCHENK, J. J.; BECKLUND, L. E.; CAREY, S. J.; FABRE, P. P. What is the “modified” CTAB protocol? Characterizing modifications to the CTAB DNA extraction protocol. **Applications in Plant Sciences**, v. 11, n. 3, e11517, 2023.

SILVA, M. N. da. Extração de DNA genômico de tecidos foliares maduros de espécies nativas do cerrado. **Revista Árvore**, v. 34, n. 6, p. 973-978, 2010.

THERMO FISHER SCIENTIFIC. **260/280 and 260/230 ratios: NanoDrop ND-1000 and ND-8000 8-sample spectrophotometers.** Wilmington: NanoDrop Technologies, 2007.

Agradecimentos

Os autores agradecem ao Laboratório de Genética do Centro de Ciências Agrárias e Engenharias (CCAE) da Universidade Federal do Espírito Santo (UFES), campus de Alegre-ES, pela infraestrutura disponibilizada para a realização deste estudo, à Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (Fapes) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo financiamento da pesquisa.