

**VERÔNICA DA COSTA SARAIVA**

**MAPEAMENTO CITOGÊNÉTICO COMPARATIVO DOS GENES *DXMT1* E *DXMT2*  
DA VIA DE BIOSÍNTESE DA CAFEÍNA EM *Coffea* L.**

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

Orientador: Wellington Ronildo Clarindo

**VIÇOSA – MINAS GERAIS  
2022**

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

T

S243m  
2022 Saraiva, Verônica da Costa, 1992-  
Mapeamento citogenético comparativo dos genes DXMT1 e  
DXMT2 da via de biossíntese da cafeína em Coffea L. /  
Verônica da Costa Saraiva. – Viçosa, MG, 2022.  
1 dissertação eletrônica (33 f.): il. (algumas color.).

Orientador: Wellington Ronildo Clarindo.  
Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Biologia Geral, 2022.  
Referências bibliográficas: f. 29-33.  
DOI: <https://doi.org/10.47328/ufvbbt.2022.747>  
Modo de acesso: World Wide Web.

1. Café - Melhoramento genético. 2. Café - Citologia.  
3. Mapeamento genômico vegetal. I. Clarindo, Wellington  
Ronildo, 1979-. II. Universidade Federal de Viçosa.  
Departamento de Biologia Geral. Programa de Pós-Graduação  
em Genética e Melhoramento. III. Título.

CDD 22. ed. 633.732

Bibliotecário(a) responsável: Bruna Silva CRB-6/2552

VERÔNICA DA COSTA SARAIVA

**MAPEAMENTO CITOGENÉTICO COMPARATIVO DOS GENES *DXMT1* E *DXMT2*  
DA VIA DE BIOSÍNTESE DA CAFEÍNA EM *Coffea* L.**

Dissertação apresentada à universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para a obtenção do título de *Magister Scientiae*.

APROVADA: 5 de agosto de 2022.

Assentimento:



---

Verônica da Costa Saraiva  
Autora



---

Wellington Ronaldo Clarindo  
Orientador

*Aos meus pais, Francisca Almeida e João Barbosa, pelos ensinamentos e valores. Mostrando sempre que a conquista dos nossos objetivos é fruto de muito esforço, força de trabalho, humildade e honestidade. E aos meus irmãos Wanderson Saraiva e Mônica Saraiva.*

**Dedico.**

## **AGRADECIMENTOS**

A Deus pelo dom da vida e por iluminar todos os meus passos e confortar meu coração diante das dificuldades.

À minha família que sempre me apoiou em todas as minhas conquistas e que esteve sempre ao meu lado nos bons e maus momentos. Minha mãe, Francisca Almeida Costa, minha maior inspiração e meu exemplo de pessoa. Ela que sempre me motivou a acreditar que meus sonhos e objetivos eram possíveis. Ao meu pai João Barbosa. Meus irmãos, Wanderson Saraiva e Monica Saraiva.

À Universidade Federal de Viçosa (UFV) pela oportunidade e conhecimentos adquiridos.

O presente trabalho foi realizado com o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG pela concessão da bolsa de estudo.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

Ao meu amigo Maurício pelo apoio, incentivo e pela nossa grande amizade.

Ao Prof. Dr. Wellington Ronildo Clarindo, por todas as trocas de saberes que contribuíram para o meu desenvolvimento científico e profissional.

À Dra. Jéssica Silva pela paciência, acompanhamento e orientações durante a minha pesquisa.

À Dra. Mariana Sattler pelos ensinamentos e contribuições nas técnicas de citogenética molecular.

Aos membros da banca examinadora, pelo interesse, disponibilidade e contribuições que irão enriquecer o meu estudo.

Aos meus amigos do Laboratório de Citogenética e Citometria Vegetal da UFV, em especial ao João pelo carinho e amizade.

## BIOGRAFIA

Verônica da Costa Saraiva, filha de Francisca Almeida Costa e João Barbosa Neto, natural de Floriano – PI, nasceu em 20 de maio de 1992, em Floriano, Piauí. Em janeiro de 2012, ingressou no curso de Licenciatura em Ciências Biológicas no Instituto Federal do Piauí (IFPI), *Campus* Floriano, Piauí. Concluiu o curso de graduação em março de 2017, obtendo o título de Licenciado em Ciências Biológicas em julho de 2017, durante a cerimônia de colação de grau.

Em agosto de 2021, concluiu o curso de pós-graduação *Lato Sensu* em Projetos e Licenciamento Ambiental, pela Faculdade Única de Ipatinga, Minas Gerais.

Em julho de 2020 ingressou no curso de Mestrado do Programa de Pós-Graduação em Genética e Melhoramento da Universidade Federal de Viçosa - MG, na linha de pesquisa de Biologia Evolutiva e Citogenética, sob orientação do professor Dr. Wellington Ronildo Clarindo. Atuou em atividades de pesquisa na área de Cultura de Tecidos Vegetais e Citogenética, submetendo-se à defesa de dissertação em 5 de agosto de 2022.

## RESUMO

SARAIVA, Verônica da Costa, M.Sc., Universidade Federal de Viçosa, agosto de 2022. **Mapeamento citogenético comparativo dos genes *DXMT1* e *DXMT2* da via de biossíntese da cafeína em *Coffea* L.** Orientador: Wellington Ronildo Clarindo.

A rota biossintética da cafeína é regulada por diferentes genes da família *N-metiltransferases* (*NMTs*). Dentre eles, os genes *DXMT1* e *DXMT2* codificam enzimas que atuam na última reação da via, convertendo a teobromina em cafeína. Vinte e três genes *NMTs* são encontrados em espécies de *Coffea*. No entanto, estudos citogenômicos baseados no mapeamento físico desses genes ainda não foram realizados em *Coffea* ou em outras espécies vegetais. Esses estudos podem fornecer informações sobre a organização genômica e evolução das espécies, como *Coffea eugenioides*, *Coffea canephora* e *Coffea arabica*. Desse modo, nosso objetivo foi mapear os genes *DXMT1* e *DXMT2*, utilizando uma única sonda, nos cromossomo mitóticos de *Coffea* diploides (*C. eugenioides* e *C. canephora*) e do alotetraploide verdadeiro (*C. arabica*). A sequência utilizada como sonda para mapear os genes foi sequenciada. Com base nos dados de sequenciamento, a similaridade em relação às sequências dos genes *DXMT1* e *DXMT2* foi confirmada. Os cariótipos dos *Coffea* diploides apresentaram  $2n = 2x = 22$  cromossomos e o alotetraploide  $2n = 4x = 44$  cromossomos. Por meio da sonda de 400 pb, nós mapeamos simultaneamente as sequências dos genes *DXMT1* e *DXMT2*. Dois sinais de hibridização foram identificados em *C. eugenioides* para o gene *DXMT1* e dois sinais para o gene *DXMT2* em *C. canephora*. O cariótipo de *C. arabica* apresentou quatro sinais de hibridização, dois sinais para o gene *DXMT1* possivelmente herdado de *C. eugenioides*, e dois para o gene *DXMT2* proveniente de *C. canephora*. Todos os sinais de hibridização estavam localizados em cromossomos submetacêntricos na região intersticial do braço longo. A presença de uma cópia do gene *DXMT1* em *C. eugenioides*, uma do gene *DXMT2* em *C. canephora* e duas cópias em *C. arabica*, sendo uma do gene *DXMT1* e outra cópia do gene *DXMT2* fornece evidências adicionais para a origem de *C. arabica*. Portanto, os resultados fornecem a base para a realização de outras abordagens genômicas e evolutivas no gênero *Coffea*. Além disso, nosso estudo contribui para a caracterização citogenômica de *Coffea* e também para outras espécies que

possuem os genes *NMTs* envolvidos na biossíntese da cafeína.

**Palavras-chave:** Mapeamento físico. Café. Citogenômica comparativa. metabolismo secundário. Sequenciamento. Mapa genético.

## ABSTRACT

SARAIVA, Verônica da Costa, M.Sc., Universidade Federal de Viçosa, August 2022. **Comparative cytogenetic mapping of the *DXMT1* and *DXMT2* genes of the caffeine biosynthesis pathway in *Coffea* L.** Advisor: Wellington Ronildo Clarindo.

Caffeine biosynthesis is regulated by genes of the N-methyltransferases (NMTs) family. Among them, the *DXMT1* and *DXMT2* genes encode enzymes that act in the last reaction of the pathway, converting theobromine into caffeine. Twenty-three NMTs genes are found in the genome of *Coffea* species. However, cytogenomic studies based on the physical mapping of these genes have not yet been carried out in *Coffea* or other plant species. These studies can provide information about the genomic organization and evolution of species, such as *Coffea eugenioides*, *Coffea canephora* and *Coffea arabica*. Thus, we aimed to map the *DXMT1* and *DXMT2* genes, using a single probe, on the mitotic chromosomes of diploids (*C. eugenioides* and *C. canephora*) and true allotetraploid (*C. arabica*) *Coffea*. The sequence used as probe to map the genes was sequenced. Based on the sequencing data, the similarity was confirmed in relation to the *DXMT1* and *DXMT2* genes. The karyotypes of the diploid *Coffea* showed  $2n = 2x = 22$  chromosomes, and the allotetraploid  $2n = 4x = 44$  chromosomes. Using the 400 bp probe, we simultaneously mapped the *DXMT1* and *DXMT2* gene. Two hybridization signals for the *DXMT1* gene were identified in *C. eugenioides* and two signals for the *DXMT2* gene in *C. canephora*. *C. arabica* karyotype showed four hybridization signals, two signals for the *DXMT1* gene possibly inherited from *C. eugenioides*, and two for the *DXMT2* gene from *C. canephora*. All hybridization signals were mapped in the interstitial region of the long arm of a submetacentric chromosome. The presence of one copy of the *DXMT1* gene in *C. eugenioides*, one copy of the *DXMT2* gene in *C. canephora*, and two copies in *C. arabica*, one of the *DXMT1* gene and the other copy of the *DXMT2* gene, provides additional evidence for the true allotetraploid origin of *C. arabica*. Therefore, the results provide the basis for further genomic and evolutionary approaches in the *Coffea* genus. In addition, our study contributes to the cytogenomic characterization of *Coffea*, as well as other species that have NMTs genes involved in caffeine biosynthesis.

**Keywords:** Physical mapping. Coffee. Comparative cytogenomics. Secondary metabolism. Sequencing. Genetic map.

## SUMÁRIO

|                                |    |
|--------------------------------|----|
| INTRODUÇÃO .....               | 11 |
| MATERIAL E MÉTODOS.....        | 14 |
| RESULTADOS.....                | 18 |
| DISCUSSÃO .....                | 22 |
| CONCLUSÕES .....               | 26 |
| INFORMAÇÕES SUPLEMENTARES..... | 27 |
| REFERÊNCIAS.....               | 29 |

## INTRODUÇÃO

*Coffea* L. é um gênero com grande representatividade dentro da família Rubiaceae, compreendendo cerca de 124 espécies nativas da África e ilhas do Oceano Índico, estendendo-se ao sul e sudeste da Ásia e Australásia. O gênero *Coffea* tem origem recente de 7,87 Mya, sendo a região central Africana o principal centro de diversidade de espécies de *Coffea* (Tosh et al. 2013). As espécies do gênero *Coffea* apresentam número básico de cromossomos  $x = 11$ , sendo quase todas diploides com  $2n = 2x = 22$  cromossomos. *Coffea arabica* L. é o único alotetraploide do gênero, com  $2n = 4x = 44$  cromossomos (Hamon et al. 2017). O evento de hibridação natural que originou *C. arabica* ocorreu possivelmente entre *Coffea canephora* Pierre ex A. Froehner e *Coffea eugenioides* S. Moore (Lashermes et al. 1999; Clarindo e Carvalho 2009).

Nas últimas décadas, *Coffea* tornou-se um gênero de grande interesse econômico, sendo utilizado principalmente para produção de café. O café é uma das bebidas mais importantes do mundo, com aproximadamente 2,25 bilhões de xícaras consumidas diariamente (Denoeud et al. 2014). O interesse econômico mundial está baseado principalmente em duas espécies: *Coffea arabica* que corresponde a 70% da produção, e *C. canephora* que representa 30%. Atualmente, as espécies de *Coffea* são encontradas em mais de 70 países de regiões tropicais e subtropicais distribuídas pela América do Sul e Central, Ásia e África (Vieira et al. 2019). A diferença na maior produção de *C. arabica* em relação a *C. canephora* se dá principalmente pelo fato de *C. arabica* ser uma espécie que gera uma bebida de alta qualidade, aroma marcante, menor amargor, suave e com menor teor de cafeína, sendo dessa forma mais apreciado pelos consumidores. A bebida de *C. canephora* é menos aromática, densa, sabor mais amargo e com elevado teor de cafeína (Huded et al. 2020).

No gênero *Coffea*, a presença da cafeína é de grande interesse, principalmente nos grãos de café, em virtude do impacto na qualidade da bebida (Kato et al. 2009). A cafeína (1,3,7-trimetilxantina) é membro de um grupo de compostos conhecidos coletivamente como alcaloides purínicos, derivados da xantosina (Leibrock et al. 2022). A cafeína é um metabólito secundário encontrado em quase 100 espécies de plantas pertencentes aos gêneros *Camellia*, *Coffea*, *Theobroma* e *Paullinia* (Kato e Nakayama 2017). A principal via de biossíntese da

cafeína ocorre a partir da xantosina (Denoeud et al. 2014). A via envolve três etapas de metilação catalisadas por N-metiltransferases (NMTs) dependentes de S-adenosil-1-metionina (SAM) e uma reação de nucleosidase. A primeira etapa na via biossintética da cafeína é a conversão da xantosina em 7-metilxantosina por uma N-metiltransferase dependente de SAM. A 7- metilxantosina é hidrolisada por uma nucleosidase em 7-metilxantina, a qual é metilada no 3-N em teobromina (3,7-dimetilxantina). Em seguida, a 3,7-dimetilxantina é metilada no 1-N formando a cafeína (Ashihara et al. 2008; Perrois et al. 2015).

As enzimas NMTs envolvidas na biossíntese da cafeína são codificadas por diferentes genes *NMTs*. No genoma de *Coffea*, os genes NMTs pertencem a uma única família que agrupa 23 genes que codificam enzimas que atuam na via biossintética da cafeína (Denoeud et al. 2014). Os diferentes genes envolvidos na biossíntese da cafeína possuem sequências genômicas que variam de 1.829 a 2.063 bp, bem como o mesmo número de éxons e íntrons, sendo 4 e 3 respectivamente (Denoeud et al. 2014). Ademais, suas sequências exibem alta similaridade (>80%) entre si (Perrois et al. 2015). A análise filogenética realizada com diferentes N-metiltransferases envolvidas na síntese da cafeína evidenciou que os *NMTs* estão agrupados em três clusters que correspondem, respectivamente, a *XMT* (xantosinometiltransferase), *MXMT* (7- metilxantina methyltransferase) e *DXMT* (3,7-dimetilxantina methyltransferase) (Perrois et al. 2015).

Em espécies do gênero *Coffea*, como *C. canephora*, três genes para N-metiltransferases foram identificados, com base em dados de sequenciamento, e localizados duas regiões distintas: o *DXMT2* no cromossomo 1 e os genes *XMT1* e o *MXMT1* no cromossomo 9 (Denoeud et al. 2014). No genoma de *C. arabica* foram identificados seis genes relacionados à via de biossíntese da cafeína, sendo: *XMT*, *XMT2*, *MXMT1*, *MXMT2*, *DXMT1* e *DXMT2*. Os genes ortólogos *XMT1*, *MXMT1* e *DXMT2* encontrados em *C. arabica* foram herdados de *C. canephora*, enquanto os *XMT2*, *MXMT2* e *DXMT1* foram provavelmente herdados do progenitor *C. eugenioides* (Perrois et al. 2015).

Genes de cópia única, sequências com poucas cópias no genoma e sequências repetitivas podem ser identificadas pela construção de mapas citogenéticos. Com base nos mapas citogenéticos, a evolução dos genomas de plantas vem sendo investigada e caracterizada. Em *Coffea*, o mapeamento citogenético foi conduzido para as sequências repetitivas (repeatoma) de rDNA 5S e

45S (18S e 26S), sendo importante para discriminação de cromossomos e entendimento da organização evolutiva do gênero (Hamon et al. 2009). O mapeamento de gene rDNA 5S realizado nos diploides *C. canephora* e *C. eugenioides* exibiu a presença de um loci na porção intersticial do braço longo do cromossomo 4 para *C. canephora*, e na porção pericentromérica do braço longo do cromossomo 8 em *C. eugenioides* (Sattler et al. 2022).

A distribuição cromossômica de outras sequências do repeatoma, dois transposons (*MuDR* e *Tip100*) e um LTR-retroelemento (*Del 1*), foi avaliada em *C. canephora*, *C. eugenioides* e *C. arabica*. Os sinais de hibridização dos elementos móveis foram mapeados predominantemente nas porções teloméricas em *C. eugenioides* e *C. canephora*, e porções intersticiais, em maior número, em *C. arabica*. *Del 1* apresentou sinais de hibridização mais dispersos nas três espécies (Lopes et al. 2013). O mapeamento citogenético de retrotransposons LTR *Ciganos* comuns nos centrômeros dos cromossomos de plantas também foi realizado em *Coffea*. O mapeamento trouxe um direcionamento sobre a composição das regiões centroméricas no gênero *Coffea*, e também evidências genômicas acerca da relação de parentesco de *C. arabica* com seus dois progenitores (Nunes et al. 2018).

Em outras espécies de plantas, a citogenômica vem avançando no mapeamento de genes de cópia única. A primeira tentativa de mapeamento de genes de cópia única em plantas foi realizada em *Petroselinum crispum* L. (Huang et al. 1988). O mapeamento citogenético de sequências específicas foi conduzido para *Allium cepa* L. e *Allium fistulosum* L., envolvendo a localização de genes membros da família multigênica da alinase (Khrustaleva et al. 2019). Em *Triticum aestivum* L. e *Triticum turgidum* L., genes que conferem resistência à toxicidade do alumínio foram mapeados (Oliveira e Maglio 2020). Os genes *Glh*, *Bph-3* e *Xa-5* identificados em *Oryza sativa* L. foram mapeados em *Oryza granulata* Nees et Arn. Ex Watt.

Dessa forma, considerando a importância do mapeamento citogenético na genômica, o nosso objetivo foi construir um mapa citogenético do gene *DXMT1* em *C. eugenioides*, do *DXMT2* em *C. canephora*, e dos genes *DXMT1* e *DXMT2* em *C. arabica*.

## MATERIAL E MÉTODOS

### *Material biológico*

Folhas de indivíduos das espécies *C. canephora*, *C. arabica* e *C. eugenioides* (doadores de explantes) foram coletadas de plantas cultivadas em casa de vegetação na Universidade Federal de Viçosa (UFV), Minas Gerais, Brasil. As plantas foram desinfestadas com solução contendo 10,0 mL L<sup>-1</sup> Assist (BASF<sup>®</sup>, Brasil), 1,0 g L<sup>-1</sup> Curathane (Dow AgroSciences<sup>®</sup>, Brasil), 2,0 g L<sup>-1</sup> Agri-micina PM (Pfizer<sup>®</sup>, Brasil) e 0,5 mL L<sup>-1</sup> Ethion 500 (Bayer Crop- Science<sup>®</sup>, Brasil). Após 24 h da desinfestação, as folhas foram coletadas e lavadas superficialmente com detergente líquido e mantidas em água corrente por 2 h. Em seguida, as folhas foram desinfestadas em câmara de fluxo laminar com etanol 70% (Merck<sup>®</sup>) por 1 min e solução contendo 500 mL de hipoclorito de sódio (Merck<sup>®</sup>), 500 mL de dH<sub>2</sub>O estéril e 10 gotas L<sup>-1</sup> de Tween 20 (Sigma<sup>®</sup>) por 20 min. Finalizando, as folhas foram lavadas por três a quatro vezes em dH<sub>2</sub>O estéril (Clarindo et al. 2012).

### *Estabelecimento das suspensões de agregados celulares*

Fragmentos das folhas desinfestadas (~2 cm<sup>2</sup>) de cada *Coffea* doadora de explante foram inoculados em placas de Petri, 60 × 15mm (Prolab<sup>®</sup>, Brasil) contendo o meio de indução de calo friável. O meio de indução de calo friável utilizado foi idêntico ao proposto por Sanglard et al. (2019). Trinta placas de Petri para cada *Coffea* com cinco fragmentos foram mantidas no escuro a 22 ± 2°C por 90 dias para indução de calos friáveis. Decorrido os 90 dias, 0,5 g de calos friáveis foi transferido para Erlenmeyers contendo 30 mL de meio líquido. A composição do meio líquido foi semelhante ao meio de indução, diferindo na ausência do Phytigel (Sigma<sup>®</sup>). Os Erlenmeyers foram mantidos no escuro a 25 ± 2°C em agitador orbital a 100 rpm, ocorrendo seis subcultivos a cada 15 dias para estabelecimento das suspensões de agregados celulares (SAC), que foram utilizadas para obtenção de núcleos e cromossomos (Sanglard et al. 2019).

Sete dias após o sexto subcultivo, as SAC foram tratadas com 4 µM do agente anti-tubulínico amiprofos-metil e mantidas no escuro, em agitador orbital a 100 rpm

e  $25 \pm 2^\circ\text{C}$  por 8 h. As SAC foram fixadas em solução de metanol: ácido acético (3:1) com três trocas de 10 min cada e mantidas a  $-20^\circ\text{C}$ . Posteriormente, as SAC foram novamente lavadas em  $\text{dH}_2\text{O}$  por três vezes 10 min cada e maceradas por 2 h a  $36^\circ\text{C}$  em pool enzimático (0,5% celulase, 0,5% hemicelulase Sigma<sup>®</sup>, 0,5% macerozyme Orozuka R10 Yakult) diluído em  $\text{dH}_2\text{O}$  na proporção 1:10 (pool enzimático:  $\text{dH}_2\text{O}$ ). Após a maceração, as SAC foram lavados em  $\text{dH}_2\text{O}$  e fixados em 3:1 metanol:ácido acético e armazenado a  $-20^\circ\text{C}$ . A partir das SAC maceradas as lâminas foram preparadas por meio das técnicas de dissociação celular e secagem ao ar (Clarindo e Carvalho 2006).

#### *Amplificação genômica para caracterização dos genes relacionados à cafeína*

O DNA genômico de *C. eugenoides*, *C. canephora* e *C. arabica* foi extraído utilizando o kit GenElute<sup>™</sup> Plant Genomic DNA Miniprep (Sigma<sup>®</sup>). A concentração e pureza do DNA foram avaliadas por espectrofotometria usando NanoDrop (Invitrogen), e a integridade do DNA foi verificada por eletroforese em gel de agarose a 1,5%. Os primers DXMT-C *sense* 5'-CAGGGACATCTGGAGGAAGA-3' e *antisense*: 5'-GCTTTTTGGCGAGAGAAATG-3', foram desenhados a partir de uma região consenso obtida do alinhamento das sequências completas dos genes *DXMT1* e *DXMT2* (3,7- dimetilxantina methyltransferase) disponíveis no National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>).

As sequências foram amplificadas em reações de PCR utilizando DNA genômico de *C. eugenoides* gerando uma sonda de 400 pb. A mistura da reação de amplificação foi composta por 0,5  $\mu\text{M}$  de cada primer, 100 ng de DNA genômico, 200  $\mu\text{M}$  de cada dNTP (Promega), tampão de reação 1x (Invitrogen), 2 mM  $\text{MgSO}_4$  (Invitrogen) e 2,5 U Platinum<sup>®</sup> Taq DNA Polimerase High Fidelity (Sigma<sup>®</sup>) para um volume final de 30  $\mu\text{L}$ . As condições da PCR no termociclador foram: desnaturação inicial a  $95^\circ\text{C}$  por 2 min; 30 ciclos de desnaturação a  $95^\circ\text{C}$  durante 1 min; anelamento a  $45^\circ\text{C}$  por 45 seg; extensão a  $68^\circ\text{C}$  por 1 min; e extensão final a  $72^\circ\text{C}$  por 2 min. Os produtos de amplificação foram analisados por eletroforese em gel de agarose 1,5%, purificados por meio de E-gel<sup>®</sup> SizeSelect (Invitrogen) e quantificados pelo Qubit<sup>™</sup> (Invitrogen).

A sequência alvo de DNA foi amplificada via PCR para gerar a sonda *DXMT* usando os primers *DXMT-C*. A reação de produção da sonda consistiu em 1,25 U de

Platinum® Taq DNA Polymerase High Fidelity (Invitrogen); 1x do tampão de reação da enzima (Invitrogen); 1,8 mM de MgSO<sub>4</sub> (Invitrogen), 0,2 mM de dATP, dCTP, dGTP e 0,15 mM de dTTP; 0,05 mM de Chromatide® Alexa Fluor® 488-5-dUTP; 0,5 µM de cada primer e 100 ng do DNA molde amplificado, um volume final de 20 µL para cada reação. A reação de marcação foi conduzida na mesma condição de reação de PCR estabelecida para a amplificação.

Lâminas contendo cromossomos metafásicos e núcleos foram lavadas em tampão fosfato salino (PBS 1X) por 5 min, fixadas em solução de formalina 4% por 10 min, lavadas novamente com PBS 1X por 5 min em temperatura ambiente e desidratadas em série alcoólica (70%, 85% e 100%) por 5 min cada a -20°C. A desnaturação dos cromossomos e núcleos foi realizada em tampão de formamida 70% em solução salina de citrato de sódio (SSC 2X) a 67°C por 3 min, seguida novamente de série alcoólica gelada (70%, 85% e 100%) por 3 min cada. A mistura de hibridização consistiu em formamida 50%, SSC 2x e 200 ng da sonda marcada. A desnaturação a 85°C por 5 min em termociclador MJ e imediatamente colocada no gelo.

As lâminas foram incubadas com 35 µL da mistura de hibridização, coberta com lamínulas de plástico HybriSlip™ (Sigma®) e seladas com cola Rubber Cement (Elmer's®). A hibridização foi conduzida no equipamento ThermoBrite™ (StatSpin®) a 37°C por 24 h. Posteriormente, as lâminas foram submetidas à lavagem de estringência realizada a 37°C em solução SSC 2X por 5 min. As metáfases e os núcleos foram contracoradas com 40% glicerol + 4',6-diamidino-2-fenilindol (Soares et al. 2020).

Os núcleos e metáfases foram capturados com uma câmera digital 12-bit CCD (Olympus®) acoplada a um fotomicroscópio Olympus BX-60 equipado com iluminação de campo claro, epifluorescência e objetiva de imersão (100×, NA = 1,4). As imagens foram capturadas na exposição de 1.8 e processadas no software Image-Pro Plus 6,1 (Media Cybernetics, Inc).

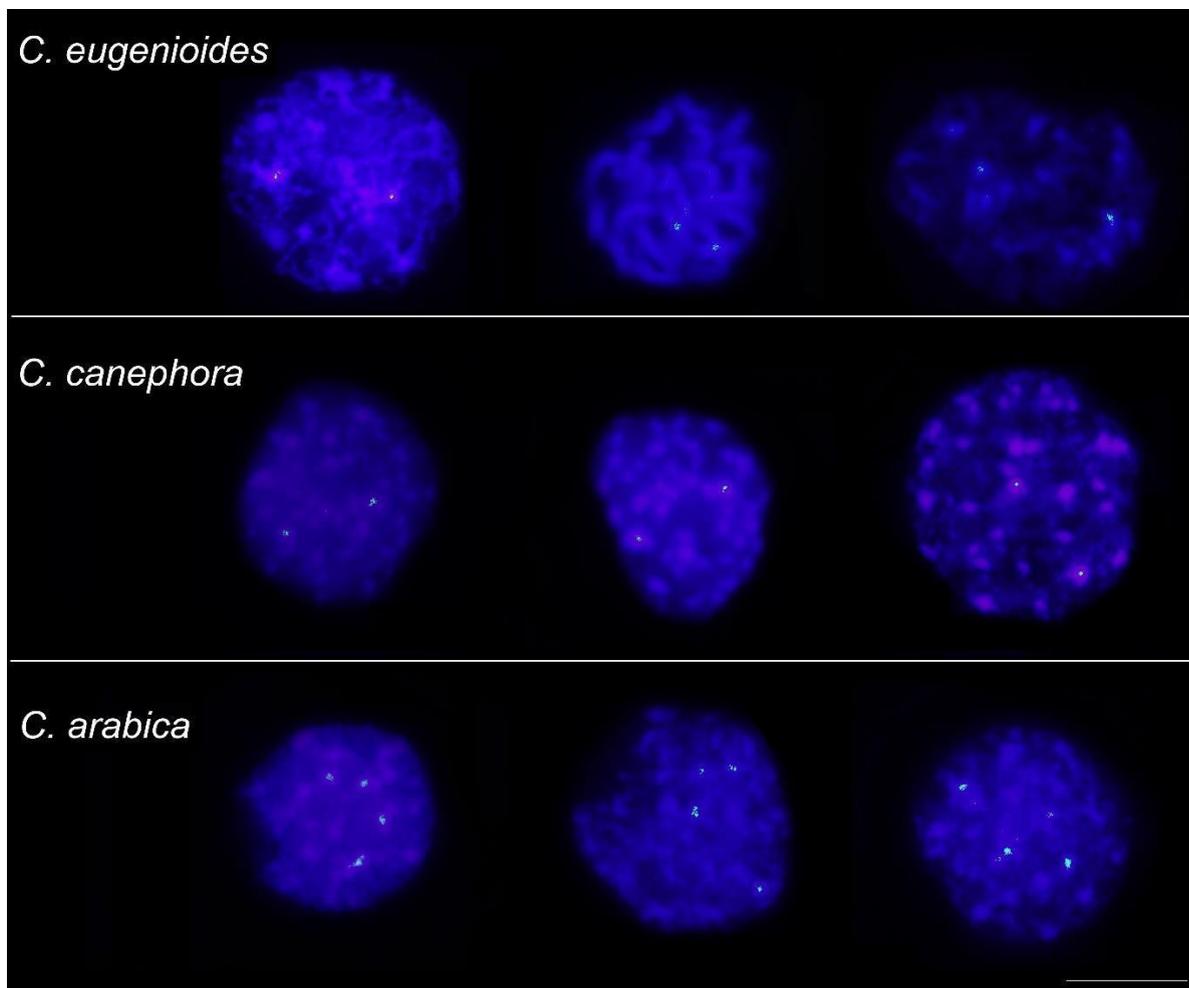
### *Sequenciamento dos produtos de amplificação*

As reações de PCR para sequenciamento foram feitas para um volume total

de 50  $\mu\text{L}$  e consistiram em 0,5  $\mu\text{M}$  de cada primer, 200 ng de DNA amplificado e purificado, 200  $\mu\text{M}$  de cada dNTP (Promega<sup>®</sup>), 1,3 mM de  $\text{MgCl}_2$ , 1X de tampão Colorless GoTaq<sup>®</sup> Flexi (Promega<sup>®</sup>) e 1 U da enzima de Platinum<sup>®</sup> Taq DNA Polymerase High Fidelity (Invitrogen). As condições da PCR foram as mesmas utilizadas para a amplificação das sequências e produção das sondas. A purificação das amostras foi realizada com o kit Wizard<sup>®</sup> SV Gel and PCR Clean-Up System (Promega<sup>®</sup>). A concentração e pureza do produto de PCR foram determinadas por espectrofotômetro NanoDrop 1000 (Thermo Fisher Scientific<sup>®</sup>), e sua integridade foi avaliada por eletroforese em gel de agarose 1,5%. Os produtos de PCR purificados foram enviados para sequenciamento pelo método de Sanger na empresa ACTGene Análises Molecular, Brasil. A similaridade da sequência foi realizada usando o software Molecular Evolutionary Genetics Analysis (MEGAX - versão 11, 64-bit) e pelo banco de dados NCBI GenBank.

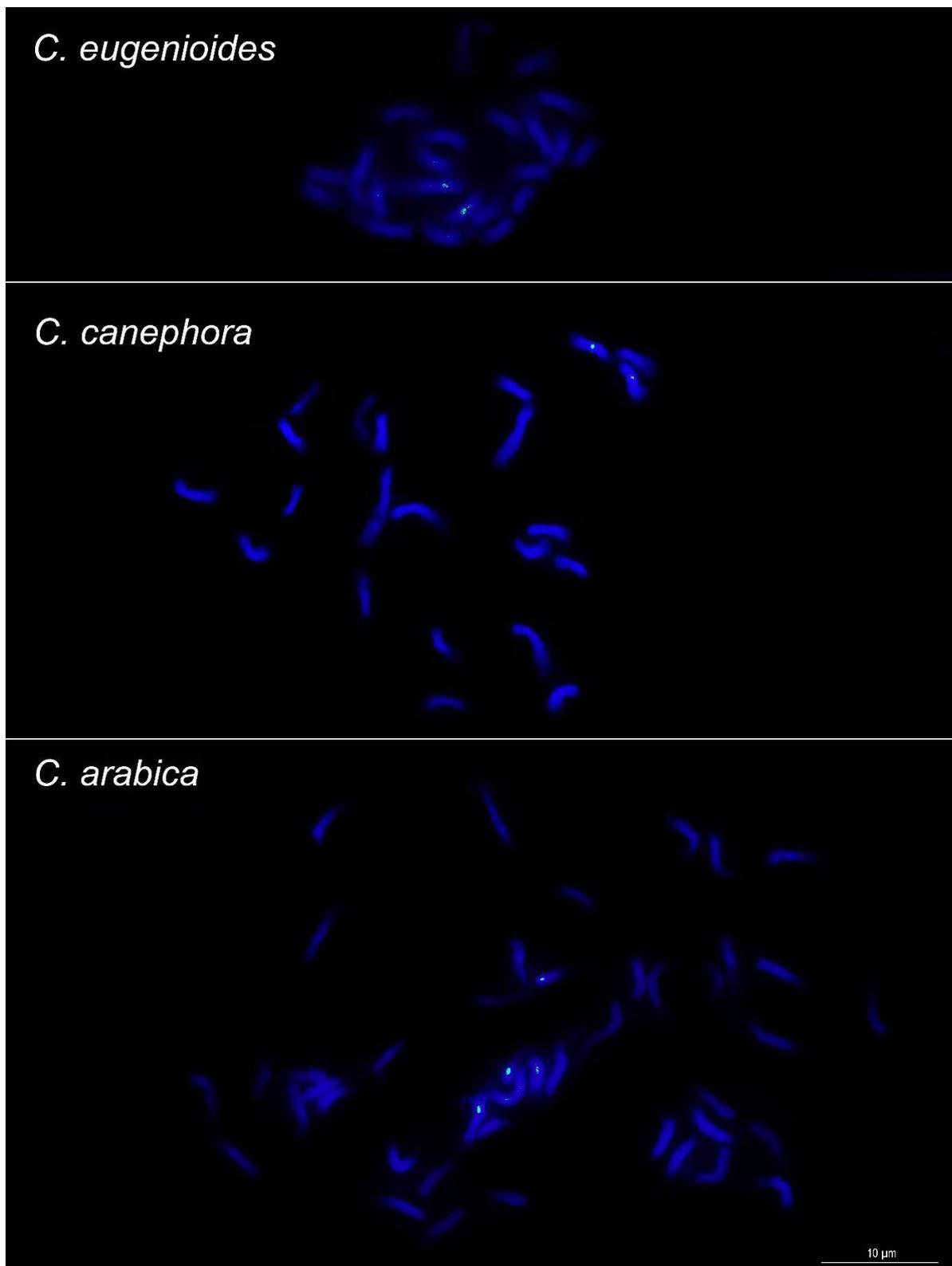
## RESULTADOS

Uma única sonda *DXMT* construída a partir do primer *DXMT-C*, com 400 pb, foi utilizada na hibridização *in situ* fluorescente em núcleos interfásicos e metáfases de *Coffea*. Os núcleos interfásicos foram utilizados para identificar o número de cópias dos genes *DXMT1* e *DXMT2* em *Coffea*. Dessa forma, dois sinais de hibridização foram identificados em quinze núcleos interfásicos de *C. eugenioides* e em 8 de *C. canephora*. Os sinais identificados mostram que *C. eugenioides* possui uma cópia do gene *DXMT1*, assim como *C. canephora* apresenta uma cópia do gene *DXMT2*. Em *C. arabica*, quatro sinais de hibridização foram identificados em 8 núcleos interfásicos, indicando que essa espécie possui uma cópia do *DXMT1* e outra do *DXMT2*. A hibridização *in situ* realizada nos núcleos interfásicos das três espécies confirmaram o número de cópias dos genes *DXMT1* e *DXMT2*, demonstrando também a reprodutividade e especificidade da sonda utilizada para os genes *DXMT1* e *DXMT2* (Figura 1).



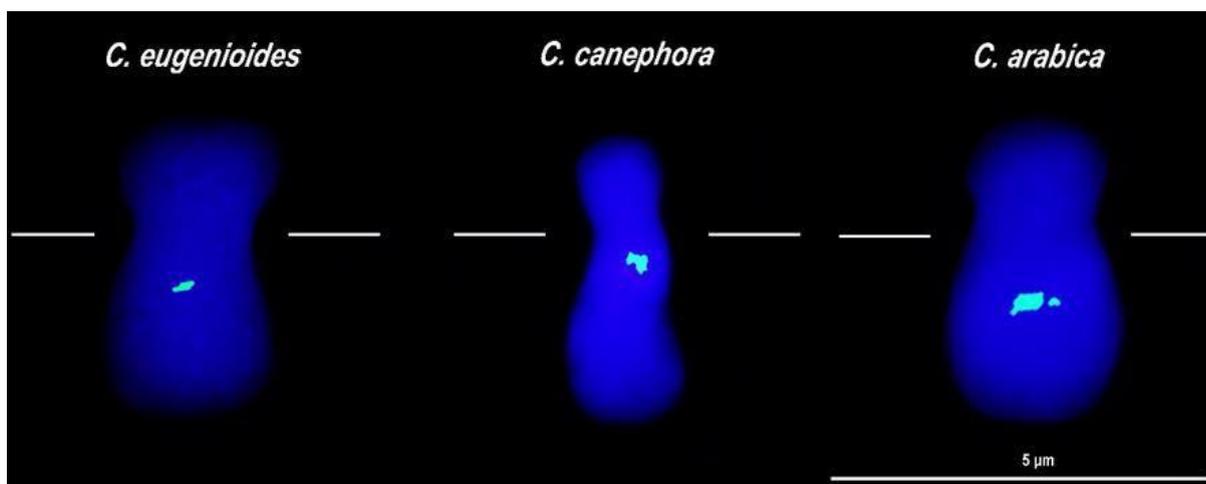
**Figura 1:** Núcleos interfásicos de *Coffea* exibindo sinais de hibridização para o genes da cafeína (*DXMT1* e *DXMT2*). *C. eugenioides* com dois sinais para o gene *DXMT1*, *C. canephora* com dois sinais do gene *DXMT2* e *C. arabica* com quatros sinais, sendo dois para o gene *DXMT1* e dois para *DXMT2*. Barra: 10  $\mu$ m.

As metáfases dos diploides *C. eugenioides* e *C. canephora* apresentaram  $2n = 2x = 22$  cromossomos, e as do alotetraploide *C. arabica* ( $2n = 4x = 44$  cromossomos), confirmando o nível de ploidia das três espécies. Sequências dos genes *DXMT1* e *DXMT2* foram mapeadas, pela primeira vez, em metáfases de *Coffea*. Dois sinais de hibridização foram obtidos para o gene *DXMT1* em metáfases de *C. eugenioides* e dois sinais para o gene *DXMT2* em *C. canephora*. Quatro sinais de hibridização foram observados no cariótipo de *C. arabica*, sendo dois sinais para o gene *DXMT1* e dois sinais para *DXMT2* em diferentes pares cromossômicos. (Figura 2).



**Figura 2:** Mapeamento de sequências dos genes *DXMT1* e *DXMT2* em cromossomos metafásicos de *Coffea*. *C. eugenioides* exibindo dois sinais do gene *DXMT1*, *C. canephora* dois sinais para o gene *DXMT2* e *C. arabica* quatro sinais, sendo dois sinais para o gene *DXMT1* e dois sinais para *DXMT2*. Barra 10 μm.

Nós observamos que os cromossomos de *C. eugenoides*, *C. canephora* e *C. arabica* exibiram sinais de hibridização para a sonda na região intersticial do braço longo de um par cromossômico de *C. eugenoides*, um par em *C. canephora* e em dois pares cromossômicos de *C. arabica* (Figura 3).



**Figura 3.** Cromossomos metafásicos com sinais de hibridização das sequências do gene *DXMT1* em *C. eugenoides*, do gene *DXMT2* em *C. canephora* e em *C. arabica* para *DXMT1* e *DXMT2*. Os sinais estão localizados na região intersticial do braço longo. Barra: 5 µm.

A sonda gerada a partir do DNA genômico de *Coffea* foi sequenciada e sua similaridade verificada no MEGAX e pelo banco de dados NCBI GenBank, usando a função de Busca de Alinhamento Local Básico (BLAST). O alinhamento realizado no MEGAX demonstrou similaridade de 95,75% e 96,69% para as sequências sense e antisense, respectivamente, com os primers utilizados (Tabela Suplementar 2). O percentual de similaridade entre os dados do sequenciamento com as sequências disponíveis no NCBI foi equivalente a 97% para a fita sense e 97% para antisense do gene *DXMT1* em *C. eugenoides* e 96% na fita sense e 97% na antisense do gene *DXMT2* em *C. canephora*. Enquanto, *C. arabica* apresentou 97% para fita sense e 96% na antisense para os dois genes *DXMT1* e *DXMT2*.

## DISCUSSÃO

Nós realizamos, pela primeira vez, o mapeamento físico de uma sequência referente aos genes *DXMT1* e *DXMT2* em cromossomos metafásicos do alotetraploide *C. arabica* e dos seus dois progenitores *C. eugenioides* e *C. canephora*. Os genes *DXMT1* e *DXMT2* são responsáveis por codificar uma das cafeína sintase (*DXMT*) que converte a teobromina em cafeína. Os genes *XMT*, *MXMT* e *DXMT* envolvidos na via de biossíntese da cafeína apresentam alta similaridade (>80%) entre si (Perrois et al. 2015), fato que dificulta o desenho de primers específicos para cada um desses genes envolvidos. O alinhamento de múltiplas sequências dos genes N-metiltransferases apresentaram alta similaridade entre os diferentes genes agrupados (Leibrock et al. 2022). O percentual de similaridade que esses genes apresentam entre si nos permitiu construir uma única sonda, que mapeou simultaneamente as sequências dos genes *DXMT1* e *DXMT2* e determinou o número cópias desses dois genes. A similaridade da sonda utilizada com os genes *DXMT1* e *DXMT2* foi confirmada por meio do sequenciamento da sequência, pelo banco de dados NCBI GenBank. Perrois et al. (2015) projetaram primers “universais” para identificar os diferentes genes *XMT*, *MXMT* e *DXMT* presentes em *C. arabica* e *C. canephora*, explorando a alta similaridade entre as sequências.

A ploidia dos diploides *C. eugenioides* e *C. canephora*, e do alotetraploide *C. arabica* foi confirmada. *C. arabica*, inicialmente classificado como um alotetraploide segmental natural (Pinto-Maglio e Cruz 1998), originou-se de um evento hibridação entre duas espécies diploides com genomas semelhantes (Stebbins 1949; Sattler, 2022). O alotetraploide *C. arabica* tem como seus progenitores potenciais *C. eugenioides* e *C. canephora* (Lashermes et al. 1999; Ruas et al. 2003; Clarindo e Carvalho 2009). Considerando a origem evolutiva de *C. arabica*, a sonda construída mapeou dois genes diferentes relacionados à via de biossíntese de cafeína, tanto o *DXMT1* encontrado no genoma de *C. eugenioides* quanto o gene *DXMT2* em *C. canephora*. Dessa forma, a sonda gerou sinais de hibridização para esses dois genes no genoma de *C. arabica*. O mapeamento dessas sequências nos permite sugerir que não ocorreram eventos evolutivos envolvendo a perda desses genes no genoma de *C. arabica*, o que afetaria a biossíntese da cafeína nessa espécie. Uma condição

diferente do que ocorre em *Coffea humblotiana* Baill, uma espécie caracterizada pela ausência de cafeína nas folhas e sementes. A ausência da cafeína em *C. humblotiana* se deve a perda do loci *DXMT*, uma vez que a análise microssintética revelou a ausência de um segmento de 76 kb neste loci (Raharimalala et al. 2021). Alterações no número de cópias de um gene resulta em alterações no nível de expressão gênica, o que leva a existência de diferentes fenótipos (Hastings et al. 2009).

Genomas poliploides, como o de *C. arabica*, estão sujeitos a alterações genéticas, envolvendo mutações e alterações cromossômicas estruturais (Chen 2007; Lopes et al. 2013). O sequenciamento dos dois acessos de *C. arabica* 'AR41' e 'Caturra' evidenciou a perda de 179 genes nos onze cromossomos de referência ao subgenoma de *C. canephora* (Lashermes et al. 2016). A perda de sequências parentais também foi observada em outros aloploides como *Tragopogon miscellus* Ownbey (Shaked et al. 2001; Tate et al. 2006). Indivíduos alotetraploides de *T. miscellus* apresentaram uma média de ~20% menos, equivalente a cerca de 70 loci em um dos cromossomos homeólogos, de um dos alelos dos loci investigados (Buggs et al. 2012). Em poliploides naturais e sintéticos de *Arabidopsis suecica* Fries, a ocorrência de alterações cromossômicas estruturais envolveu a perda de um loci de rDNA parental (Pontes et al. 2003; Pontes et al. 2004). O mapeamento físico realizado para o gene *11A* nas espécies *Brachiaria ruziziensis* Germain e Evrard, *Brachiaria brizantha* Stapf. cv. Marandu e *Brachiaria decumbens* Stapf identificou a ocorrência de deleções em virtude da falta de um sinal de hibridização em um dos cromossomos homólogos do par 7 em *B. brizantha* (Nani et al. 2018).

*C. arabica* e seus dois progenitores *C. eugenioides* e *C. canephora* apresentam diferenças no nível de cafeína em folhas e grãos, e nos diferentes estágios de maturação dos grãos (Perrois et al. 2015). Os sinais de hibridização dos núcleos interfásicos e metáfases confirmaram o número de cópias dos genes *DXMT1* em *C. eugenioides*, do gene *DXMT2* em *C. canephora* e dos genes *DXMT1* e *DXMT2* em *C. arabica*. Dessa forma, inferimos que a diferença no teor de cafeína nessas espécies possivelmente está relacionada ao genótipo e/ou à expressão gênica desses genes, e não na perda do loci dos genes *DXMT1* e *DXMT2*. Segundo Ky et al. (2001), os valores médios do conteúdo de cafeína expressos em base percentual da matéria seca (% dmb) nas espécies *Coffea* variam de 0% (dmb) em *Coffea pseudozanguebariae* Bridson a 2,5% (dmb) em *C. canephora*. O conteúdo

de cafeína expresso em grãos verdes de *C. eugenoides* equivale a 0,5% (dmb) e em *C. arabica* a 1,6% (dmb) (Campa et al. 2005). Os genes *DXMT1* e *DXMT2* estão diretamente relacionados ao acúmulo de cafeína em folhas e grãos de *Coffea* (Raharimalala et al. 2021). Os níveis de expressão gênica dos genes envolvidos na via de biossíntese da cafeína em *C. arabica* são diferentes em relação à *C. eugenoides* e *C. canephora*, o que leva a uma produção intermediária de cafeína em suas folhas e grãos (Perrois et al. 2015). Em *C. arabica*, diferentes padrões de expressão foram encontrados para genes homeólogos de *C. eugenoides* e *C. canephora*. A variação na expressão gênica contribui para a variação no teor de cafeína nos grãos de diferentes espécies e variedades de plantas (Leibrock et al. 2022). Segundo Tran et al. (2018), 40 SPNs potencialmente associados à via de biossíntese de cafeína foram identificados e podem levar às diferenças fenotípicas no teor de cafeína *Coffea*.

O mapeamento da sonda confeccionada revelou que os sinais de hibridização em prometáfases e metáfases de *C. eugenoides*, *C. canephora* e *C. arabica* não estão localizadas no cromossomo 1, divergindo dos resultados do sequenciamento de *C. canephora* que aponta que o gene *DXMT2* está localizado no cromossomo 1 (Denoëud et al. 2014). O par cromossômico com sinal de hibridização de *C. canephora* foi comparado com o cariótipo da espécie estabelecido por Clarindo e Carvalho (2009), indicando ser um dos pares cromossômicos (2, 3, 5, 7 ou 8). Os cromossomos com sinais de hibridização para sequência do *DXMT1* em *C. eugenoides* apresentaram semelhança morfológica com os cromossomos 8 e 9 do cariógrama reportado por Sattler (2022). Portanto, nós acreditamos que o gene *DXMT1* em *C. eugenoides* está localizado em um desses cromossomos.

Segundo Denoëud et al (2014), os genes *XMT*, *MXMT* e *DXMT* em *Coffea* evoluíram como resultado de duplicações em tandem no cromossomo 1 ancestral, e supostamente os genes *XMT* e *MXMT* foram translocados para o cromossomo 9 e o gene *DXMT* permaneceu no cromossomo 1. O agrupamento de genes *NMTs* de *Coffea* é sintético a *NMTs* de outros gêneros, como *Gardenia* (Rubiaceae), não produtores de cafeína. No entanto, uma análise comparativa do genoma *Gardenia* e *Coffea* mostrou que o cluster de genes *NMTs* envolvidos na biossíntese de cafeína em *C. canephora* mostra sintenia com uma região em *Gardenia jasminoides* Ellis, mas que são anteriores a divergência entre *Coffea* e *Gardenia* (Xu et al. 2020). Esses dados sugerem que as duplicações em tandem dos genes *NMTs* em *Coffea*

ocorreram após a diversificação do gênero *Coffea* em relação aos gêneros *Theobroma*, *Paullinia*, *Citrus*, *Camellia*, *Gardenia* (Deneud et al. 2014; Huang et al., 2016). As duplicações em tandem trazem novas funções metabólicas e papéis adicionais na evolução do genoma das plantas (Salojarvi et al. 2017).

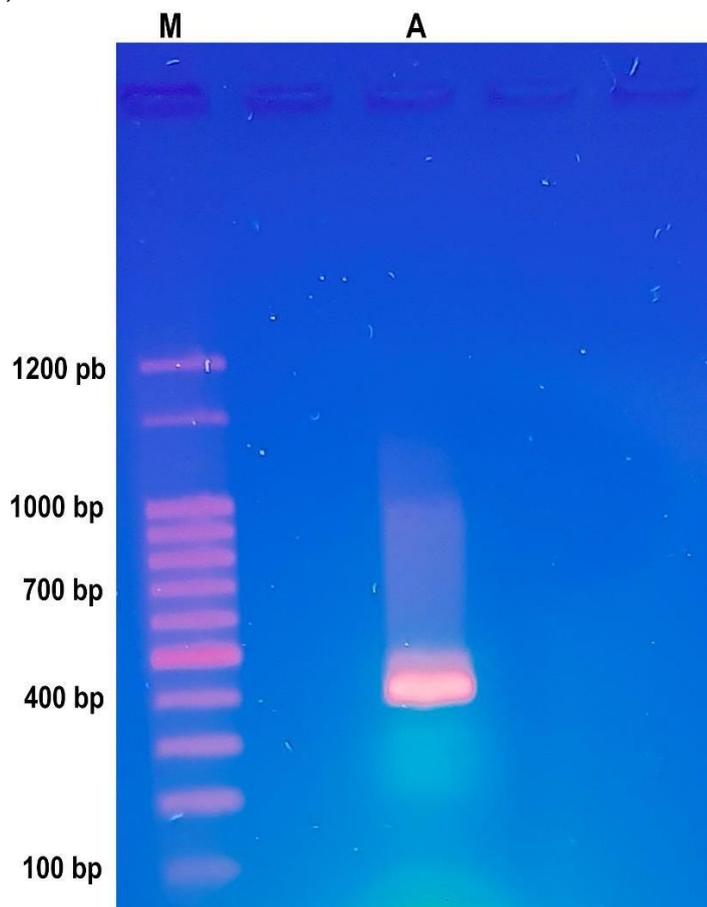
O número de genes *NMTs* é distinto entre as espécies desses diferentes gêneros. Por exemplo, *Camellia sinensis* L. possui 13 genes *NMTs*, *Theobroma cacao* Linn. 21 e *C. canephora* 23 genes relacionados à biossíntese da cafeína (Xia et al. 2017). No entanto, a localização cromossômica dessas sequências por mapeamento citogenético não foi realizada em nenhuma dessas espécies. O mapeamento citogenético tem possibilitado a identificação de genes de cópia única e sequências repetitivas em diversas espécies de plantas (Jiang 2019). Os mapas citogenéticos são valiosos apenas por integrar as informações sobre estrutura, organização e evolução dos cromossomos individuais, e também de genomas inteiros (Wang 2006). A localização simultânea de diferentes sequências de DNA utilizando 15 BACs de *Phaseolus vulgaris* L. foi usada para construir um mapa citogenético comparativo para *Phaseolus lunatus* L. O mapa construído possibilitou a identificação de todos os cromossomos de *P. lunatus*, permitindo uma análise de sintonia entre os genomas das duas espécies (Bonifácio et al. 2012). Em *Zea mays* L., um mapa citogenético foi construído para oito genes, incluindo os genes *DEK1*, *SERK2*, *RF2*, *ACC1*, *ACC2* e *MYO1* (Lamb et al. 2007). Entre as espécies vegetais, *Z. mays* teve seu mapeamento físico explorado tanto para sequências de cópia única quanto para sequências do repeatoma (Yu et al. 2007; Lamb et al. 2007), agregando informações com aplicações diretas no melhoramento de plantas (Nani et al. 2018).

## CONCLUSÕES

Nós, pela primeira vez, determinamos o número de cópias e mapeamos sequências dos genes *DXMT1* e *DXMT2* relacionados a via de biossíntese da cafeína em três espécies *Coffea*. *C. eugenioides* possui uma cópia do gene *DXMT1*, *C. canephora* uma cópia do gene *DXMT2* e *C. arabica* duas cópias gênicas, sendo uma do *DXMT1* proveniente de *C. eugenioides* e uma do gene *DXMT2* herdada de *C. canephora*. Com isso, inferimos que a diferença no teor de cafeína presente nessas três espécies está relacionado à diferença da expressão gênica desses genes ou dos diferentes genes relacionados à via de biossíntese da cafeína e não ao número de cópia dos genes *DXMT1* e *DXMT2*. Além disso, nossos dados contribuem para aumentar as análises citogenômicas e o entendimento da evolução e estrutura do genoma de *Coffea*.

## INFORMAÇÕES SUPLEMENTARES

**Figura Suplementar 1:** Produto de PCR resultante da amplificação do DNA genômico de *C. eugenioides* utilizando o primer para os genes *DXMT1* e *DXMT2*, em gel de agarose 1,5%



M: marcador, A: amplicon, bp: pares de base.

**Tabela Suplementar 1:** Sequências utilizadas para alinhamento no MEGAX.

```

>DXMT_sense
CCCCCCTTTCMTGTTTCGTCTATGCACCTTCAGTAGAAGAAGTAAAGTGCATAGTTGAGGAGGAAGGTT
CTTTTGAAATTTTATACCTGGAGACTTTTAAGGTCCCTTATGATGCTGGCTTCTCTATTGATGATGATTAC
CAAGGAAGATCCCATTCCCAGTATCCTGCGATGAACATGCTAGAGCAGCGCATGTGGCATCTGTCGTT
AGATCAGTTTACGAACCCATCCTCGCAAGTCATTTTGGAGAAGCTATTATACCTGACTTATCCCACAGGA
TTGCGAAGAATGCAGCAAAGGTTCTCCGCTCGGGCAAAGGCTTCTATGATAATCTTATCATTCTCTCGC
CAAAAAGC
>DXMT_antisense
CAGTCTSGAAGCCTTTGCCCGGCGGAGACCTTTGCTGCATTCTTCGCAATCCTGTGGAATAAGTCAGT
ACAATAGCTTCTCCAAAATGACTTGCAGGATGGGTTTCGTAAACTGATCTAACGACAGATGCCACATGC
GCTGCTCTAGCATGTTTCATCGCAGGATACTGGGGAAATGGGATCTTCCTTGGTAATCATCATCAATAGAG
AAGCCAGCATCATAAGGGGCTTAAAAGTCTCCAGGTATAAAATTTCAAAGAACCTTCTCCTCAACTA
TGCACTTTACTTCTTCTGCTGAAGGTGCATAGACTGGAACATTGAAACTGTCCAATTTTCTCCTCCAGA
TGTCCTGA
>ORIGEM
CAGGGACATCTGGAGGAAGAAAAATTGGACAGTTTCAATGTTCCAATCTATGCACCTTCAACAGAAGAAG
TAAAGTGCATAGTTGAGGAGGAAGGTTCTTTTGAATTTTATACCTGGAGACTTTTAAGGTCCCTTATGAT
GCTGGCTTCTCTATTGATGATGATTACCAAGGAAGATCCCATTCCCAGTATCCTGCGATGAACATGCTA
GAGCAGCGCATGTGGCATCTGTCGTTAGATCAATTTTGAACCCATCGTCGCAAGTCATTTTGGAGAAGC
TATCATGCCTGACTTATCCCACAGGATTGCGAAGAATGCAGCAAAGGTTCTTCGCTCCGGCAAAGGCTTC
TATGATAGTCTTATCATTCTCTCGCCAAAAAGC

```

**Tabela Suplementar 2:** Similaridade das sequências *Sense/Antisense* (sequenciamento) com a sequência “única” de origem para os genes *DXMT1* e *DXMT2*.

|                         | <i>DXMT_sense</i> | <i>DXMT_antisense</i> |
|-------------------------|-------------------|-----------------------|
| <i>DXMT_sense</i>       | 0,0363            |                       |
| <i>DXMT_antisense</i>   | 0,0363            |                       |
| <i>Origem</i>           | <b>0,0425</b>     | <b>0,0331</b>         |
| <i>Similaridade (%)</i> | <b>95,75496</b>   | <b>96,692396</b>      |

## REFERÊNCIAS

- Ashihara H, Santo H, Crozier A (2008) Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* 69:841-856. <https://doi.org/10.1016/j.phytochem.2007.10.029>
- Bonifácio E M, Fonsêca A, Almeida C, Santos K G B, Haranda-Pedrosa (2012) A Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor Appl Genet* 124, 1513–152. [doi.org/10.1007/s00122-012-1806-x](https://doi.org/10.1007/s00122-012-1806-x)
- Buggs R J A, Chamala S, Wu w, Tate J A, Schnable P S, Soltis D E, Soltis P S, Barbazuk W B (2012) Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin. *Current Biology* 22:248-252. <https://doi.org/10.1016/j.cub.2011.12.027>
- Campa C, Doulebeau S, Dussert S, Hamom S, Noirot M (2005) Diversity in bean caffeine content among wild *Coffea* species: evidence of a discontinuous distribution. *Food Chemistry* 91: 633–63. [doi.org/10.1016/j.foodchem.2004.06.032](https://doi.org/10.1016/j.foodchem.2004.06.032).
- Chen Z J (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu Rev Plant Biol* 58:377-406. doi: 10.1146/annurev.arplant.58.032806.103835
- Clarindo W, Carvalho C (2006) A high quality chromosome preparation from cell suspension aggregates culture of *Coffea canephora*. *Cytologia* 71: 243- 249. <https://doi.org/10.1508/cytologia.71.243>
- Clarindo W R, Carvalho C R (2009) Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. *Plant Cell Rep* 28:73-81. <https://doi.org/10.1007/s00299-008-0621-y>
- Clarindo W R, Carvalho C R, Mendonça M A C (2012) Cytogenetic and flow cytometry data expand knowledge of genome evolution in three *Coffea* species. *Plant Syst Evol* 298:835-844. <https://doi.org/10.1007/s00606-012-0595-7>
- Denoeud F, Carretero-Paulet L, Dereeper A et al (2014) The *Coffee* genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345: 1181-1184. <https://doi.org/10.1126/science.1255274>
- Hamon P, Siljak-Yokvlev S, Srisuwan S, Robin O, Poncet O, Hamon S et al (2009) Physical mapping of rDNA and heterochromatin in chromosomes of 16 *Coffea* species: a revised view of species differentiation. *Chromosome* 17: 291–304. <https://doi.org/10.1007/s10577-009-9033-2>
- Huang P L, Hahlbrock K, Somssich I E (1988) Detection of a single-copy gene on plant chromosomes by in situ hybridization. *Mol Gen Genet* 211:143–147. <https://doi.org/10.1007/BF00338405>

Hamon P, Grover C E, Davis A P et al (2017) Genotyping-by-sequencing provides the first well-resolved phylogeny for coffee (*Coffea*) and insights into the evolution of caffeine content in its species. *Mol Phylogenet Evol* 109:351–361. <https://doi.org/10.1016/j.ympev.2017.02.009>

Huded A K C, Jingade P, Bychappa M, Mishra M K (2020) Genetic diversity and population structure analysis of *Coffea* (*Coffea canephora*) germplasm collections in indian gene bank employing SRAP and SCoT. *International Journal of Fruit Science* 20:757-784. <https://doi.org/10.1080/15538362.2020.1768618>

Hastings PJ, Lupski JR, Rosenberg SM, Ira G (2009) Mechanisms of change in gene copy number. *Nature Reviews Genetics* 10:551-564. <https://doi.org/10.1038/nrg2593>

Huang R, O'Donnell A, Barboline J, Barkman T (2016) Convergent evolution of caffeine in plants by co-option of exapted ancestral enzymes. *PNAS* 113:10613-10618. <https://doi.org/10.1073/pnas.1602575113>

Jiang J (2019) Fluorescence in situ hybridization in plants: recent developments and future applications. *Chromosome Res* 27:153–165 (2019). <https://doi.org/10.1007/s10577-019-09607-z>

Khrustaleva L, Kudryavtseva N, Romanov D, Ermolaev A, Kirov I (2019) Comparative tyramide-FISH mapping of the genes controlling flavor and bulb color in *Allium* species revealed an altered gene order. *Sci Rep* 9:12007. <https://doi.org/10.1038/s41598-019-48564-9>.

Kato M, Noda M, Inoue M, Kadowaki T, Tsugane S (2009) Psychological factors, *Coffee* and risk of diabetes mellitus among middle-aged japanese: a population-based prospective study in the JPHC study cohort. *Endocrine Journal* 56:459. <https://doi.org/10.1507/endocrj.K09E-003>

Kato M, Nakayama F (2017). Biotechnology of natural products. *Springer*. 131-143. [https://doi.org/10.1007/978-3-319-67903-7\\_6](https://doi.org/10.1007/978-3-319-67903-7_6)

Ky C L, J. Louarn J, Dussert S, Guyot B, Hamon S, Noirot M (2001) Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions. *Food Chemistry* 75: 223-230.

Lashermes P, Combes M C C, Robert J, Trouslot P, D'Hont A, Charrier A (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics* 261:259–266. <https://doi.org/10.1007/s004380050965>

Lashermes P, Hueber Y, Combes M C, Dany Severac, Dereeper A (2016) Inter-genomic DNA exchanges and homeologous gene silencing shaped the nascent allopolyploid *Coffea* genome (*Coffea arabica* L.), *G3 Genes/Genomes/Genetics* 6: 2937–2948. <https://doi.org/10.1534/g3.116.030858>

Lopes F R, Jjingo D, Silva C R M, Andrade A C, Marraccini P, Texeira M F et al (2013) Transcriptional activity, chromosomal distribution and expression effects of transposable elements in *Coffea* genomes. *Plos One* 11.

<https://doi.org/10.1371/journal.pone.0078931>

Lamb J C, Danilova T, Bauer M J, Meyer J M, Holland J J, Jensen M D, Birchler J A (2007) Single-gene detection and karyotyping using small-target fluorescence in situ hybridization on maize somatic chromosomes. *Genetics* 175: 1047-58. doi: 10.1534/genetics.106.065573.

Leibroek N V, Santegoets J, Mooijman P J W, Yusuf F, Zuijdgheest X C L et al (2022) The biological feasibility and social context of gene-edited, caffeine-free coffee. *Food Sci Biotechnol* 31:635–655. <https://doi.org/10.1007/s10068-022-01082-3>

Nunes R C N, Orozco-Arias S, Crouzillat D, Mueller L A, Strickler S R, Descombes P et al (2018) Structure and distribution of centromeric retrotransposons at diploid and allotetraploid *Coffea* centromeric and pericentromeric regions. *Front. Plant Sci* 9:175. <https://doi.org/10.3389/fpls.2018.00175>

Nani T F, Schnable J C, Washburn J D, albert P, Pereira W A, Sobrinho F R et al (2018) Location of low copy genes in chromosomes of *Brachiaria* spp. *Mol Biol* 45:109– 118. <https://doi.org/10.1007/s11033-018-4144-5>

Oliveira E C, Pinto-Maglio C A F (2020) Cytogenetic mapping of the ALMT (aluminum-activated malate transporter) gene in wheat genotypes. *Scientia Agricola* 77.doi: <http://dx.doi.org/10.1590/1678-992X-2019-0012>

Perrois C, Strickler S R, Mathieu G, Lepelley M, Bedon L, Michaux S et al (2015) Differential regulation of caffeine metabolism in *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta). *Planta* 241:179–191. <https://doi.org/10.1007/s00425-014-2170-7>

Pinto-Maglio C, Cruz N (1998) Pachytene chromosome morphology in *Coffea* L. II. *C.arabica* L. complement. *Caryologia* 51:19-35 <https://doi.org/10.1080/00087114.1998.10589117>.

Pontes O, Lawrence R J, Neves N, Silva M, Lee J H, et al. (2003) Natural variation in nucleolar dominance reveals the relationship between nucleolus organizer chromatin topology and rRNA gene transcription in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 100: 11418–11423. <https://doi.org/10.1073/pnas.1932522100>

Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Viegas L C W, Pikaard C S (2004) Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proceedings of the National Academy of Sciences of the United States of America* 101: 18240–18245. <https://doi.org/10.1073/pnas.0407258102>

Raharimalala N, Rombauts S, McCarthy A, Garavito A, Orozco-Arias S, Bellanger, L et al (2021) The absence of the caffeine synthase gene is involved in the naturally decaffeinated status of *Coffea humblotiana*, a wild species from comoro archipelago. *Sci Rep* 11:8119. <https://doi.org/10.1038/s41598-021-87419-0>

Ruas PM, Ruas CF, Rampim L et al (2003) Genetic relationship in *Coffea* species and parentage determination of interspecific hybrids using ISSR (Inter-Simple Sequence Repeat) markers. *Genet Mol Biol* 26:319–327. <https://doi.org/10.1590/S1415-47572003000300017>

Soares F A F, Carvalho C R, Sattler M C, Silva J C, Pinto D E E, Passamani P Z, Clarindo W R (2020) Plant chromosome-specific probes by microdissection of a single chromosome: Is that a reality?. *Frontiers in plant science* 11: 334. <https://doi.org/10.3389/fpls.2020.00334>

Sattler M C, de Oliveira S C, Mendonça M A C et al (2022) *Coffea* cytogenetics: from the first karyotypes to the meeting with genomics. *Planta* 255: 112. <https://doi.org/10.1007/s00425-022-03898-z>

Sanglard N A, Amaral-Silva P M, Sattler M C et al (2019) Indirect somatic embryogenesis in *Coffea* with different ploidy levels: a revisiting and updating study. *Plant cell Tissue Organ Cull PCTOC* 136:255-267. <https://doi.org/10.1007/s11240-018-1511-9>

Shaked H, Kashkush K, Ozkan H, Feldman M, Levy A A (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13:1749–59. <https://doi.org/10.1105/TPC.010083>

Salojarvi J, Smolander O P, Nieminen K, Rajaraman S, Safronov O, Safdari P, Lamminmaki A, Immanen J, Lan T, Tanskanen J, et al (2017) Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. *Nat Genet* 49:904–12. <https://doi.org/10.1038/ng.3862>

Stebbins G L (1949) Variation and evolution in plants. Columbia University Press, New York.

Tate J A, Ni Z, Scheen A C, Koh J, Gilbert C A, et al (2006) Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* 173:1599–611. <https://doi.org/10.1534/genetics.106.057646>

Tosh J, Dessein S, Buerki S et al (2013) Evolutionary history of the Afro-Madagascan *Ixora* species (Rubiaceae): species diversification and distribution of key morphological traits inferred from dated molecular phylogenetic trees. *Ann Bot* 112:1723–1742. <https://doi.org/10.1093/aob/mct222>

Tran H T M, Furtado A, Vargas C A C, Smyth H, Lee L S, Henry R (2018) SNP in the *Coffea arabica* genome associated with coffee quality. *Tree Genetics & Genomes* 14: 72. <https://doi.org/10.1007/s11295-018-1282-9>

Vieira N G, Ferrari I F, Rezende J C, Mayer J L S, Mondego J M C (2019) Homeologous regulation of frigidallike genes provides insights on reproductive development and somatic embryogenesis in the allotetraploid *Coffea arabica*. *Scientific Reports* 9:8446. DOI: <https://doi.org/10.1038/s41598-019-44666-6>.

Wang C J R, Harper L, Cande W Z (2006) High-resolution single-copy gene fluorescence in situ hybridization and its use in the construction of a cytogenetic map of maize chromosome 9. *The Plant Cell* 18:529-544. <https://doi.org/10.1105/tpc.105.037838>

Xiong Z, Tan G, You A, He G, She C, Li L et al (2004) Comparative physical mapping of rice BAC clones linked to resistance genes *Glh*, *Bph-3* and *xa-5* in *Oryza sativa* L. and *O. granulata* nees et Arn. ex watt. *Chin.Sci.Bull* 49:591–596. <https://doi.org/10.1360/03wc0533>

Xia E, Zhang H B, Sheng J, Li K, Zhang Q J, Kim C et al (2017) The Tea Tree Genome Provides Insights into Tea Flavor and Independent Evolution of Caffeine Biosynthesis. *Molecular Plant* 10:866877. [doi.org/10.1016/j.molp.2017.04.002](https://doi.org/10.1016/j.molp.2017.04.002)

Xu Z, Pu X, Gao R, Demurtas O C, Fleck S J et al (2020) Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biol* 18:63. <https://doi.org/10.1186/s12915-020-00795-3>

Yu W, Lamb JC, Han F, Birchler JA (2007) Cytological visualization of transposable elements and their transposition pattern in somatic cells of maize. *Genetics* 175:31–39. <https://doi.org/10.1534/genetics.106.064238>