

EDIÇÃO GÊNICA DE CAFEIEIRO VIA CRISPR-CAS9: CLONAGEM DE sgRNAS DE GENES DA BIOSÍNTESE DA CAFEÍNA¹

Fernanda Freitas de Oliveira², Caroline Aryoshi³, Larissa Giroto⁴, Tiago Benedito dos Santos⁵, Juarez Pires Tomaz⁶, Luiz Filipe Protasio Pereira⁷

¹Trabalho financiado pela Capes

²Doutoranda, MS, Universidade Estadual de Londrina, IAPAR, Londrina-PR, fernandafreitasdeoliveira92@gmail.com

³Doutoranda, MS, Universidade Estadual de Londrina, IAPAR, Londrina-PR, carolinearyoshi@hotmail.com

⁴Doutoranda, MS, Universidade Estadual de Londrina, IAPAR, Londrina-PR, larissagirotto@yahoo.com.br

⁵Professor, PhD, Universidade do Oeste Paulista, Presidente Prudente-SP, tiagobio02@yahoo.com.br

⁶Pesquisador, PhD, IAPAR, Londrina-PR, tomaz@iapar.br

⁷Pesquisador, PhD, Embrapa Café, Londrina-PR, filipe.pereira@embrapa.br

RESUMO: O café descafeinado apresenta uma demanda crescente do mercado consumidor, sendo portanto uma alternativa para produtores e indústria. Uma das alternativas de produção de cafés descafeinados é o silenciamento gênico via CRISPR-Cas9. O CRISPR-Cas9 é uma das ferramentas mais potentes e precisas de edição gênica, que consiste na utilização de um RNA guia (sgRNA) e de uma proteína Cas9 que reconhece regiões homólogas ao sgRNA e cliva o DNA, possibilitando o silenciamento do gene alvo. Visando iniciar trabalhos para produção de plantas de café descafeinado, neste trabalho apresentamos a estratégia de clonagem e produção de vetores CRISPR-Cas9 com sgRNAs para os genes da biossíntese da cafeína *CaMXMT* e *CaDXMT*. Os sgRNAs foram desenhados na plataforma CRISPRdirect[®] utilizando o *C. canephora* como genoma de referência. Dos 276 sgRNAs de *MXMT* e 317 de *DXMT*, foram selecionados os com menor número de off-targets e posicionados próximos a região terminal do genes, sendo três sgRNAs de cada gene: sgMXMT₁, sgMXMT₂, sgMXMT₃, sgDXMT₁, sgDXMT₂, sgDXMT₃. Para validar se os sgRNAs possuíam homologia com o genoma de *C. arabica*, foi realizado blast com o software BioEdit[®]. Foi possível analisar através do alinhamento que os três sgRNAs de *CaMXMT* alinharam-se nos cromossomos “I” dos subgenoma de *C. canephora* e que sgMXMT₁ e sgMXMT₃ alinharam-se também no cromossomo “I” do subgenoma de *C. eugenioides*. Já os sgRNAs desenhados para *CaDXMT*, todos se alinharam ao cromossomo “A” no subgenoma de *C. canephora*. Apesar de não haver alinhamento no subgenoma de *C. eugenioides*, houve alinhamento no Scaffold 405, que contém a região gênica de *DXMT* de *C. eugenioides*, não sendo então um off-target. Os sgRNAs sgMXMT₁ e sgDXMT₁ foram ligados nos vetores pHAtc e pBAtc através de eletroporação. Foram positivadas por PCR 4 colônias para a ligação pHAtc/sgMXMT₁, 3 para pHAtc/sgDXMT₁ e 2 para pBAtc/sgMXMT₁. As colônias positivas na PCR foram também validadas por sequenciamento e detectou-se a inserção correta dos sgRNAs na posição esperada.

PALAVRAS-CHAVE: Edição gênica, *Coffea* spp, CRISPR-Cas9

COFFEE GENOME EDITING VIA CRISPR-CAS9: CAFFEINE GENES BIOSYNTHESIS sgRNAs CLONING

ABSTRACT: Decaffeinated coffee presents an increasing demand of the consumer market, being therefore an alternative for producers and industry. One of the alternatives to obtain decaffeinated coffee production is gene silencing via CRISPR-Cas9. The CRISPR-Cas9 is one of the most powerful and precise tools for gene editing, which consists in the use of a guide RNA (sgRNA) and a Cas9 protein that recognizes homologous regions to sgRNA and cleaves the DNA, allowing the silencing of the target gene. This work presents a strategy of cloning and production of CRISPR-Cas9 vectors with sgRNAs for *CaMXMT* and *CaDXMT* caffeine biosynthesis genes. The sgRNAs were designed on the CRISPRdirect[®] platform using *C. canephora* as the reference genome. From 276 MXMT and 317 DXMT sgRNAs, we selected the ones with the lowest number of off-targets and positioned near the gene terminal region, totalizing three sgRNAs from each gene: sgMXMT₁, sgMXMT₂, sgMXMT₃, sgDXMT₁, sgDXMT₂, sgDXMT₃. To validate if the sgRNAs had homology to the *C. arabica* genome, blast was performed using BioEdit[®] software. It was possible to analyze by alignment that the three *CaMXMT* sgRNAs aligned on the "I" chromosomes of the *C. canephora* subgenome and that sgMXMT₁ and sgMXMT₃ also aligned on the "I" chromosome of the *C. eugenioides* subgenome. In the same way, the sgRNAs designed for *CaDXMT*, all aligned to chromosome "A" in the *C. canephora* subgenome. Although there was no alignment in the *C. eugenioides* subgenome, there was alignment in Scaffold 405, which contains the *DXMT* gene region of *C. eugenioides*, and is not an off-target. The sgRNAs sgMXMT₁ and sgDXMT₁ were ligated into the vectors pHAtc and pBAtc by electroporation. Four colonies were positive for the pHAtc/sgMXMT₁, 3 binding to pHAtc/sgDXMT₁ and 2 to pBAtc/sgMXMT₁. PCR positive colonies were also validated by sequencing and the correct insertion of the sgRNAs was detected in the expected position.

KEY WORDS: Genic edition, *Coffea* spp, CRISPR-Cas9

INTRODUÇÃO

Devido aos malefícios que o consumo em excesso de café causa, existe um mercado consumidor em expansão que prefere café descafeinado. O mercado de café descafeinado movimentava 2 milhões de dólares por ano (BORRELL, 2012), sendo portanto a obtenção de café descafeinado um nicho em potencial tanto para produtores como para a indústria. Nessa linha de raciocínio, Ogita et al. (2002, 2004) demonstrou através da técnica de RNAi que é possível reduzir a produção de cafeína silenciando o gene *MXMT*. Igualmente Silvarolla; Mazzafera; Fazuoli (2004) encontraram uma cultivar de *C. arabica* descafeinada, esses autores sugerem que a ausência de cafeína se deve a uma mutação no gene *DXMT*. Entretanto, desde então, não houveram maiores progressos no avanço da obtenção de cafés descafeinados. Com o advento CRISPR-Cas9, através da edição gênica, há a possibilidade de obter cafés descafeinados e, portanto, suprir as demandas do mercado consumidor. CRISPR/Cas9 é uma das ferramentas mais potentes e precisas de edição genômica, visto que a alteração é específica e guiada por um pequeno RNA de 20 nucleotídeos (RNAsg) que se liga ao DNA alvo (CHARPENTIER; DOUDNA, 2013). Desta forma pode-se reduzir os efeitos pleiotrópicos causados eventualmente por processos tradicionais de transformação genética.

A fim de produzir cafés descafeinados, os genes das enzimas 7-Metil-xantina-metil-transferase (*MXMT*) e 3,7-Dimetil-xantina-metil-transferase (*DXMT*), que participam da principal via metabólica da cafeína, são potenciais targets para edição genômica. Para tanto, deve-se desenhar sgRNAs específicos, cloná-los e inserir-los em vetor de expressão. Posteriormente os vetores serão inseridos em *Agrobacterium tumefaciens*, para transformação e silenciamento genico em *C. arabica* e *C. canephora*.

MATERIAL E MÉTODOS

DESENHO DOS sgRNA

Para desenho dos sgRNA, utilizou-se os CDS de *MXMT* e *DXMT* de *C. arabica*. O desenho foi feito na plataforma CRISPRdirect (NAITO et al., 2015). Como genoma de referência foi utilizado o de *C. canephora* (DENOEUDE et al., 2014) disponível na plataforma.

ALINHAMENTO DOS sgRNA NO GENOMA DE *C. arabica*

Através do software BioEdit® foi feito alinhamento dos sgRNAs no genoma de *C. arabica* (Arabica Coffee Genome Consortium) para identificar em quais subgenomas os sgRNAs se alinhariam e também possíveis *off-targets*. Para que o alinhamento fosse considerado significativo, foram estipuladas cobertura e identidade de 100%.

ESCOLHA DOS VETORES

Foram testados 2 diferentes vetores pBatC, pHAtC (KIM et al., 2016), contendo a sequência correspondente a proteína Cas9, com o CaMV 35S como promotor e NOS como terminador, já para os sgRNA há o promotor U6 de *Arabidopsis*. O vetor pHAtC tem como marcador de seleção o gene de resistência a higromicina, enquanto que o vetor pBatC apresenta o gene de resistência a glufosinato de amônio.

MONTAGEM DOS VETORES DE EXPRESSÃO

A princípio, para validar a metodologia de ligação dos sgRNAs aos vetores, apenas um sgRNA de cada gene foi utilizado. Para inserção do sgRNA no vetor de expressão, primeiramente foi feita digestão do vetor com a endonuclease de restrição *AarI* a 37°C por aproximadamente 16 h, conforme informações do fabricante.

A fim de haver complementariedade do vetor digerido com os sgRNAs, foi adicionada a uma de suas fitas a sequência “GATT” e na outra fita foi adicionada a “AAAC”, ambas as sequências foram adicionadas a ponta 5', havendo assim, complementariedade com o corte do vetor realizado pela enzima *AarI*. O anelamento dos sgRNAs bem como sua ligação aos vetores foram feitos como estabelecido no protocolo de KIM et al. (2016).

Para validar se a ligação foi feita corretamente, os vetores foram eletroporados em células competentes de *E. coli* (BRASILEIRO; CARNEIRO, 2015). As células eletroporadas foram plaqueadas em meio seletivo contendo espectinomomicina na concentração de 100mg/L e mantidas a 37°C por aproximadamente 16h.

EXTRAÇÃO DE DNA, PCR E SEQUENCIAMENTO DAS CÉLULAS TRANSFORMADAS

As colônias individuais selecionadas foram replicadas para extração de DNA plasmidial. Foi feita PCR dos DNA plasmidial utilizando um primer que se anela ao sgRNA e um primer que se anela ao vetor. A PCR foi conduzida durante 30 ciclos com uma temperatura de anelamento de 63°C (Tabela 1). Os plasmídeos que apresentaram amplificação esperada via PCR foram sequenciados.

Tabela 1. Condições de termociclagem da PCR utilizada para validar a inserção dos sgRNAs no plasmídeo

Etapa	Temperatura	Tempo
Pré-ciclo	95°C	5min
	95°C	1min
Ciclo (30x)	63°C	0:45min
	72°C	0:45min
Final	72°C	5min
	4°C	∞

RESULTADOS E DISCUSSÕES

DESENHOS DOS sgRNA

Para desenho dos sgRNAs de *CaDXMT* e *CaMXMT*, foram filtrados apenas os sgRNA com alta especificidade e com poucos *off-targets* no genoma. Assim, de 276 sgRNA de *MXMT* sobraram 11 targets, e de 317 sgRNA de *DXMT* sobraram 40. Foram selecionados os 3 sgRNA de cada gene, dando preferência aos posicionados a região terminal dos genes (Tabela 1).

Tabela 2. Características dos sgRNAs de *CaMXMT* e *CaDXMT*

Gene	Nome	Start – End	Sequência Target	Strand	Tm	GC%
<i>CaMXMT</i>	sgMXMT ₁	1176 -1198	GGACTGGTATGTGTAACAGCTGG	+	72°C	50%
<i>CaMXMT</i>	sgMXMT ₂	2424 - 2446	CTCGTGCTTCAAATATCTGCAGG	+	67.9°C	45%
<i>CaMXMT</i>	sgMXMT ₃	2724 - 2746	CCACAGGCTTGC GAAGCATGCAG	-	76.7°C	60%
<i>CaDXMT</i>	sgDXMT ₁	2056 - 2078	CAAGCAAGTCCATGGAATTCGGG	+	68.3 °C	45%
<i>CaDXMT</i>	sgDXMT ₂	2059 - 2081	GCAAGTCCATGGAATTCGGGTGG	+	73.8 °C	55%
<i>CaDXMT</i>	sgDXMT ₃	2065 - 2087	CCATGGAATTCGGGTGGTCAAT	-	72.8 °C	50%

ALINHAMENTO DOS sgRNA NO GENOMA DE *C. arabica*

Foi possível analisar através do alinhamento que dois sgRNAs de *CaMXMT* (sgMXMT₁ e sgMXMT₃) alinharam-se nos cromossomos “I” dos subgenomas de *C. eugenioides* e *C. canephora*. Entretanto, o sgMXMT₂ alinhou-se apenas no cromossomo “I” do subgenoma de *C. canephora*.

Já os sgRNAs desenhados para *CaDXMT*, todos se alinharam ao cromossomo “A” no subgenoma de *C. canephora*, no entanto não foi possível observar alinhamento no subgenoma de *C. eugenioides*. Houve também um outro alinhamento para esses sgRNAs no Scaffold 405, foi realizado então o blast desse scaffold com o gene *CaDXMT* e observou-se que a sequência gênica encontra-se nesse scaffold, não sendo então um off-target, já que os sgRNAs se encontram na região gênica de *CaDXMT* que está inserida no Scaffold 405.

Todos os alinhamentos foram considerados significativos com cobertura/identidade de 100%.

Tabela 3. Alinhamento do sgRNAs no genoma de *C. arabica*

Gene	Nome	Start – End	Alinhamento	Cobertura/Identidade
<i>CaMXMT</i>	sgMXMT ₁	1176 -1198	chr_I_sg_E e chr_I_sg_C	100%
<i>CaMXMT</i>	sgMXMT ₂	2424 - 2446	chr_I_sg_C	100%
<i>CaMXMT</i>	sgMXMT ₃	2724 - 2746	chr_I_sg_E e chr_I_sg_C	100%
<i>CaDXMT</i>	sgDXMT ₁	2056 - 2078	chr_A_sg_C e Scaffold_405; HRSCAF=406	100%
<i>CaDXMT</i>	sgDXMT ₂	2059 - 2081	chr_A_sg_C e Scaffold_405; HRSCAF=406	100%
<i>CaDXMT</i>	sgDXMT ₃	2065 - 2087	chr_A_sg_C e Scaffold_405; HRSCAF=406	100%

CÉLULAS TRANSFORMADAS

A ligação entre os vetores pHA_tc e os sgMXMT₁, sgDXMT₁ apresentaram crescimento de 8 e de 9 colônias, respectivamente. Já para o vetor pBA_tc, apenas o sgMXMT₁ apresentou crescimento de 7 colônias, não sendo observado crescimento da ligação entre esse vetor e sgDXMT₁.

PCR

No vetor pHAtc com o inserto de sgMXMT₁ das 8 colônias que cresceram na placa com antibiótico, 4 foram positivas para PCR (Figura 1), já para o inserto sgDXMT₁, das 9 colônias, 3 foram positivas na PCR (Figura 1).

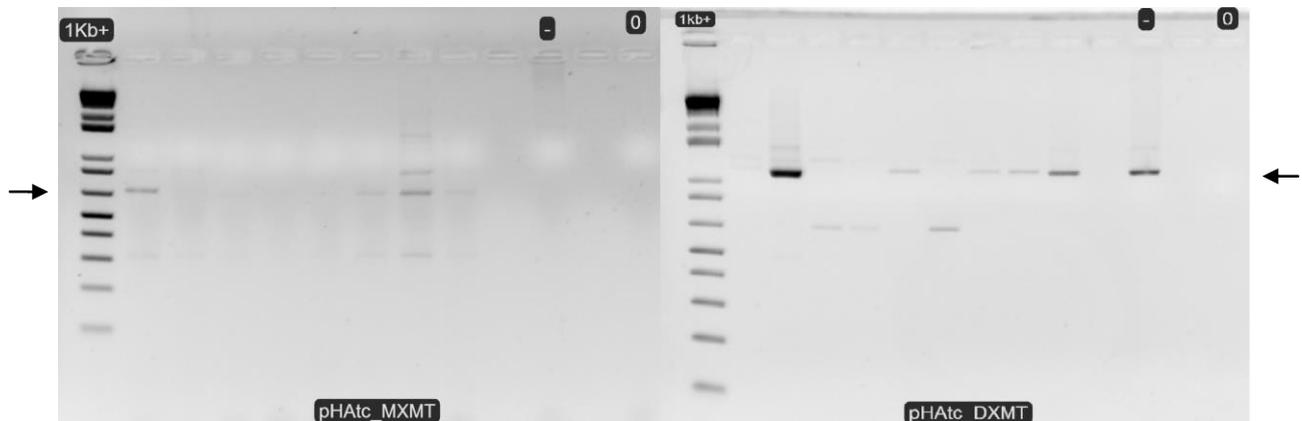


Figura 1. Eletroforese da PCR do vetor e pHAtc transformados para inserção de sgMXMT₁, sgDXMT₁. Flecha indica o Amplicon esperado de 600 pb. Amostra negativa (-) é o vetor sem transformar e o 0 é o branco

No vetor pBAtc com o inserto de sgMXMT₁ das 7 colônias que cresceram na placa com antibiótico, 2 foram positivas para PCR (Figura 2).



Figura 2. Eletroforese da PCR do vetor pBAtc transformados para inserção do sgRNA do gene sgDXMT₁. Flecha indica o Amplicon esperado de 600 pb. Amostra negativa (-) é o vetor sem transformar e o 0 é o branco.

SEQUENCIAMENTO

O sequenciamento das amostras comprovou que a inserção dos sgRNAs foi feita corretamente entre os sítios de clivagem da enzima *Aar*I e adjacente da sequência promotora de U6 (Figura 3).

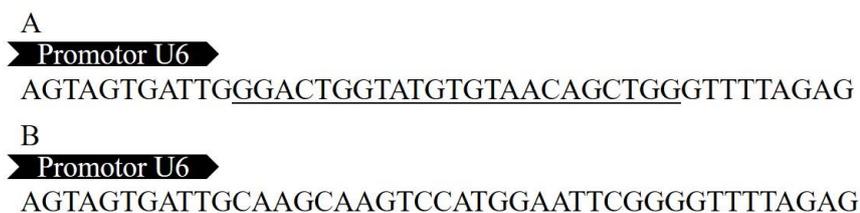


Figura 3. Sequenciamento dos sgRNAs inseridos depois do promotor U6. Fragmentos inseridos forma sublinhados: A) inserção do sgMXMT₁; B) inserção dosgDXMT₁

CONCLUSÃO

1. Foi possível desenhar os sgRNAs na plataforma CRISPRdirect utilizando o genoma de *C. canephora* como referência
2. Na análise *in silico*, os sgRNAs desenhados não apresentaram off-targets no genoma de *C. arabica*, sendo específicos para as regiões gênicas de *CaMXMT* e *CaDXMT*
3. A metodologia de inserção dos sgRNAs nos vetores pBAtc e pHAtc foi eficiente para sgMXMT₁ e sgDXMT₁, sendo confirmada por PCR e sequenciamento

AGRADECIMENTOS

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

REFERÊNCIAS

- BORRELL, B. Plant biotechnology: Make it a decaf. **Nature News**, [s. l.], v. 483, p. 264–266, 2012. Disponível em: <<http://ra.ocls.ca/ra/login.aspx?inst=conestoga&url=http://search.ebscohost.com/login.aspx?direct=true&db=awh&AN=79677545&site=eds-live>>
- BRASILEIRO, A. C. M.; CARNEIRO, V. T. de C. **Manual de transformação genética de plantas**. 2. ed. Brasília: Embrapa, 2015.
- CHARPENTIER, E.; DOUDNA, J. a. Rewriting a genome. **Nature**, [s. l.], v. 495, p. 50–51, 2013.
- DENOEUDE, F. et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science (New York, N.Y.)**, [s. l.], v. 345, n. 6201, p. 1181–4, 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/25190796>>
- KIM, H. et al. A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. **Journal of Integrative Plant Biology**, [s. l.], v. 58, n. 8, p. 705–712, 2016.
- NAITO, Y. et al. Genome analysis CRISPRdirect : software for designing CRISPR / Cas guide RNA with reduced off-target sites. [s. l.], v. 31, n. November 2014, p. 1120–1123, 2015.
- OGITA, S. et al. Genetic Modification of Coffee Plants. **Journal of Plant Biotechnology**, [s. l.], v. 4, n. 3, p. 91–94, 2002.
- OGITA, S. et al. Application of RNAi to confirm theobromine as the major intermediate for caffeine biosynthesis in coffee plants with potential for construction of decaffeinated varieties. **Plant Molecular Biology**, [s. l.], v. 54, n. 6, p. 931–941, 2004.
- SILVAROLLA, M. B.; MAZZAFERA, P.; FAZUOLI, L. C. Plant biochemistry: a naturally decaffeinated arabica coffee. **Nature**, [s. l.], v. 429, n. June, p. 826, 2004.