

IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES RELACIONADOS À BIOSSÍNTESE DE ÁCIDOS CLOROGÊNICOS EM *Coffea arabica* L.

Suzana Tiemi Ivamoto¹; Leonardo Murai Sakuray²; Tiago Benedito dos Santos³; Luiz Filipe Protasio Pereira⁴

¹ Doutoranda em Genética e Biologia Molecular, UEL, Londrina-PR, suzanatiemi@yahoo.com.br

² Graduando em Engenharia Agrônômica, UEL, Londrina-PR, leonardomurai@gmail.com

³ Pós doutorando, Instituto Agronômico do Paraná, Londrina-PR, tiagobio02@yahoo.com.br

⁴ Pesquisador, PhD, Embrapa Café, Londrina-PR, filipe.pereira@embrapa.br

RESUMO: O Brasil é o maior produtor e exportador de café no cenário mundial, além de ser o segundo maior mercado consumidor da bebida, atrás apenas dos Estados Unidos. Pesquisas estão em desenvolvimento para aumentar o conhecimento genético de cafeeiros relacionada com a qualidade do produto. A qualidade da bebida é diretamente influenciada pelos compostos químicos presentes no grão de café, como por exemplo, ácidos clorogênicos (CGAs) que são relacionados à adstringência da bebida. Visando uma maior compreensão dos mecanismos moleculares e genéticos associados a produção dos CGAs, os objetivos deste trabalho são: i) identificar dos genes da via de biossíntese dos CGAs; ii) caracterizar o perfil transcricional *in silico* para alguns genes candidatos. Neste estudo foi utilizada uma montagem *de Novo* do transcriptoma de *C. arabica* e os seus respectivos 65,480 unigenes obtidos através de RNA-Seq de folhas, flores e de perisperma de frutos em diferentes fases de desenvolvimento. Foram identificados 45 genes candidatos para a via de biossíntese dos CGAs, os quais incluem todos os principais genes descritos para esta rota metabólica (fenilalanina amonialiase, cinamato 4-hidroxilase, 4 couamato CoA ligase, 4-couamato 3-hidroxilase, hidroxicinamoil quinato transferase e cafeoil-CoA O-metiltransferase).

PALAVRAS-CHAVE: RNA-Seq, coffee, atividade transcricional, ácidos clorogênicos e qualidade da bebida

IDENTIFICATION AND CHARACTERIZATION OF GENES RELATED TO CHLOROGENIC ACIDS BIOSYNTHESIS IN *Coffea arabica* L.

ABSTRACT: Brazil is the largest producer and exporter of coffee worldwide. It is also the second largest coffee consumer market. Several studies are underway in order to increase the genetic knowledge related to coffee quality. Cup quality is directly influenced by biochemical compounds present in coffee beans, such as chlorogenic acids (CGAs) that infer astringency properties to the beverage. In order to better understand the molecular and genetics mechanisms involved in the CGAs production, the main goals of this work are to: i) identify CGAs candidate genes; ii) characterize the transcriptional profile of these candidate genes. In this study, we used a panel of 65,480 unigenes generated by an *de Novo* Assembly of *C. arabica* transcriptome data (RNA-Seq) obtained from leaves, flowers and fruit perisperm in several development stages. 45 candidate genes were identified as being related to CGAs biosynthesis, which includes all genes already describe in the literature for this metabolic pathway (phenylalanine ammoniase, cinnamate 4-hydroxylase, 4-coumarate CoA ligase, hydroxycinnamoyl-Coenzyme A shikimate/quinate, p-coumaroyl shikimate/quinate 3'-hydroxylase and Caffeoyl-CoA O-methyltransferase).

KEYWORDS: RNA-Seq, coffee, transcriptional activity, chlorogenic acids and beverage quality

INTRODUÇÃO

O estudo molecular de genes envolvidos em rotas metabólicas dos principais componentes químicos do café, visa identificar genes candidatos e aumentar o conhecimento sobre as suas influências na qualidade da bebida. Esse estudos podem gerar marcadores seja para características relacionadas a bebida ou para rastreabilidade do produto. Uma maior qualidade ou mesmo um diferencial rotas metabólicas podem agregar valor ao produto final, visando beneficiar o produtor e o consumidor (Shi et al., 2011).

Os ácidos clorogênicos (CGAs) são precursores da síntese de lignina e estão relacionados com a proteção da planta contra estresses bióticos e abióticos (Kochko et al., 2003; McCarthy et al., 2007). Eles estão também relacionados com a saúde humana devido as suas propriedades antioxidantes (Bakuradze et al., 2011).

Para identificar genes de interesse agrônômico e caracterizar os seus respectivos perfis de expressão, uma das estratégias utilizadas é acessar as informações obtidas por um sequenciamento de alto desempenho de RNAs (Dudareva et al., 2013). Esta tecnologia gera um grande volume de informações e pode ser obtido a um baixo custo e em um curto período de tempo. O objetivo desse trabalho é a identificação e caracterização do perfil transcricional *in silico* dos genes relacionados aos CGAs em *C. arabica*.

MATERIAL E MÉTODOS

Foram utilizados um painel de 65,480 unigenes gerados pela montagem *de Novo* de um transcriptoma de *C. arabica* (Pereira & Ivamoto et al., 2015).

A busca de genes candidatos para as vias de biossíntese dos ácidos clorogênicos foi realizada através de comparações nucleotídicas e proteicas através da ferramenta BlastN e BlastX, respectivamente. Foram utilizadas como parâmetros de seleção os seguintes parâmetros: i) e-value menor que $1e-30$; ii) pontuação de *score* maior do que 400; iii) Tamanho da sequência proteica maior do que 300 aminoácidos. Os perfis transcricionais foram determinados utilizando valores normalizados de RPKM, e as imagens foram obtidas com o software Genesis.

RESULTADOS E DISCUSSÃO

Foram identificados 45 genes candidatos para a via de biossíntese dos ácidos clorogênicos (tabela 1). Os resultados incluem os genes: fenilalanina amonialiase (PAL), cinamato 4-hidroxilase (C4H), 4 couramato CoA ligase (4CL), 4-couramato 3-hidroxilase (C3'H), hidroxicinamoil quinato transferase (CQT) e cafeoil-CoA O-metiltransferase (CCoAMOT).

Tabela 1: Genes Candidatos identificados para a biossíntese de CGAs

Nome do Gene	Número de Genes	Descrição da Proteína	Número EC
<i>CaPAL</i>	3	Fenilalanina amonialiase	4.3.1.5
<i>CaC4H</i>	3	Cinamato 4-hidroxilase	1.14.13.11
<i>Ca4CL</i>	5	4 couramato CoA ligase	6.2.1.12
<i>CaCQT</i>	14	Hidroxicinamoil quinato transferase	2.3.1.133
<i>C3'H</i>	17	4-couramato 3-hidroxilase	1.14.13.36
<i>CaCCoAMOT</i>	3	Cafeoil-CoA O-metiltransferase	2.1.1.104

Os genes candidatos identificados possuem funções importantes para as etapas de biossíntese dos CGAs, as quais podem ser observadas na figura 1.

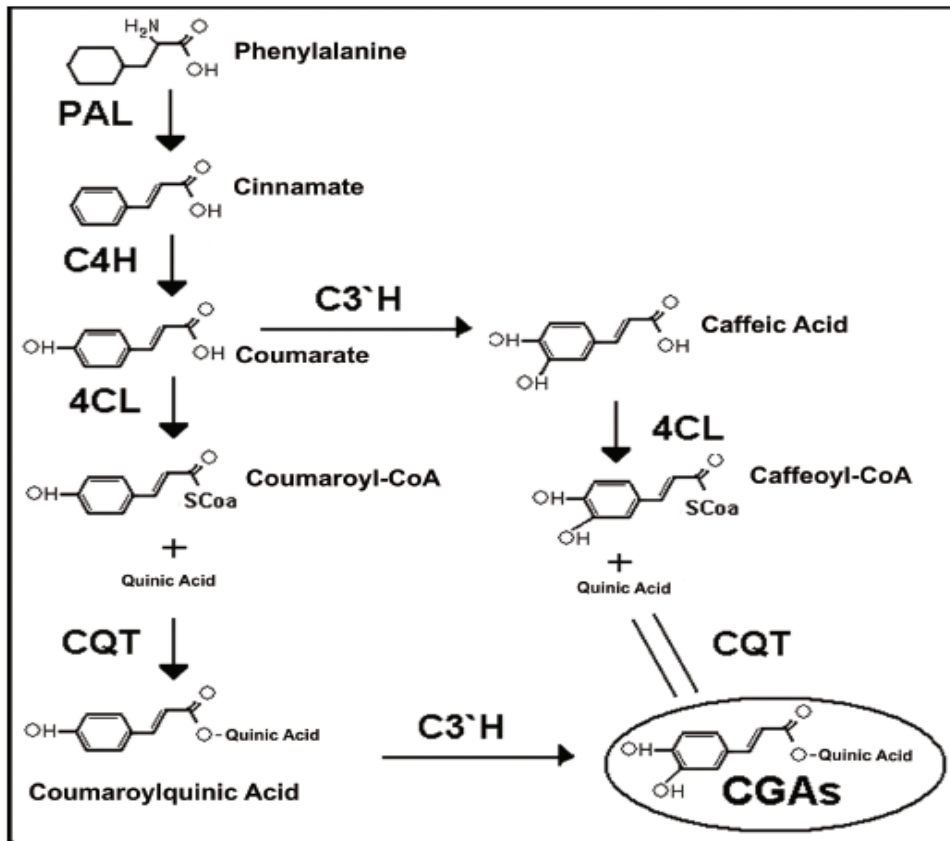


Figura 1: Biossíntese dos ácidos clorogênicos (CGAs). Fenilalanina amonialiase (PAL), cinamato 4-hidroxilase (C4H), 4 couramato CoA ligase (4CL) 4-couramato 3-hidroxilase (C3'H) and hidroxicinamoil quinato transferase (CQT). Ivamoto et al., 2013.

O padrão de atividade transcricional encontrada para os 45 genes candidatos ao longo do desenvolvimento dos frutos de café podem ser observados na figura 2. Os órgãos analisados por este trabalho foram flores, folha e perisperma de frutos nos cinco primeiros estádios de desenvolvimento (30, 60, 90, 120 e 150 dias após a floradas-DAF).

O perfil de expressão gênica observado para todos genes de PAL e 2 C4H, em flores, folhas e frutos nas etapas iniciais de desenvolvimento foram elevados. Além destes, os genes *CaC4H2* e *CaC3H14* apresentaram um padrão de expressão estável e alto em todos os órgãos/tecidos analisados. *Ca4CL1* e *CaC3H13* apresentaram um padrão interessante de expressão, eles foram expressos em folhas e perisperma, porém, apresentaram uma baixa atividade transcricional em flores.

Por outro lado alguns dos genes apresentaram maior atividade de acordo com o órgão analisado, por exemplo, para os genes *CaC3H9*, *CaC3H11*, *CaCQT3* e *CaCQT10* foram detectados um numero maior de transcritos em flores. Os genes *CaC3H2*, *CaC3H3*, *CaC3H6*, *CaC3H15* e *CaCQT5* foram transcritos principalmente em folhas. E para o gene *CaC3H17* foi observado uma maior atividade transcricional na fase inicial de desenvolvimento do perisperma dos frutos .

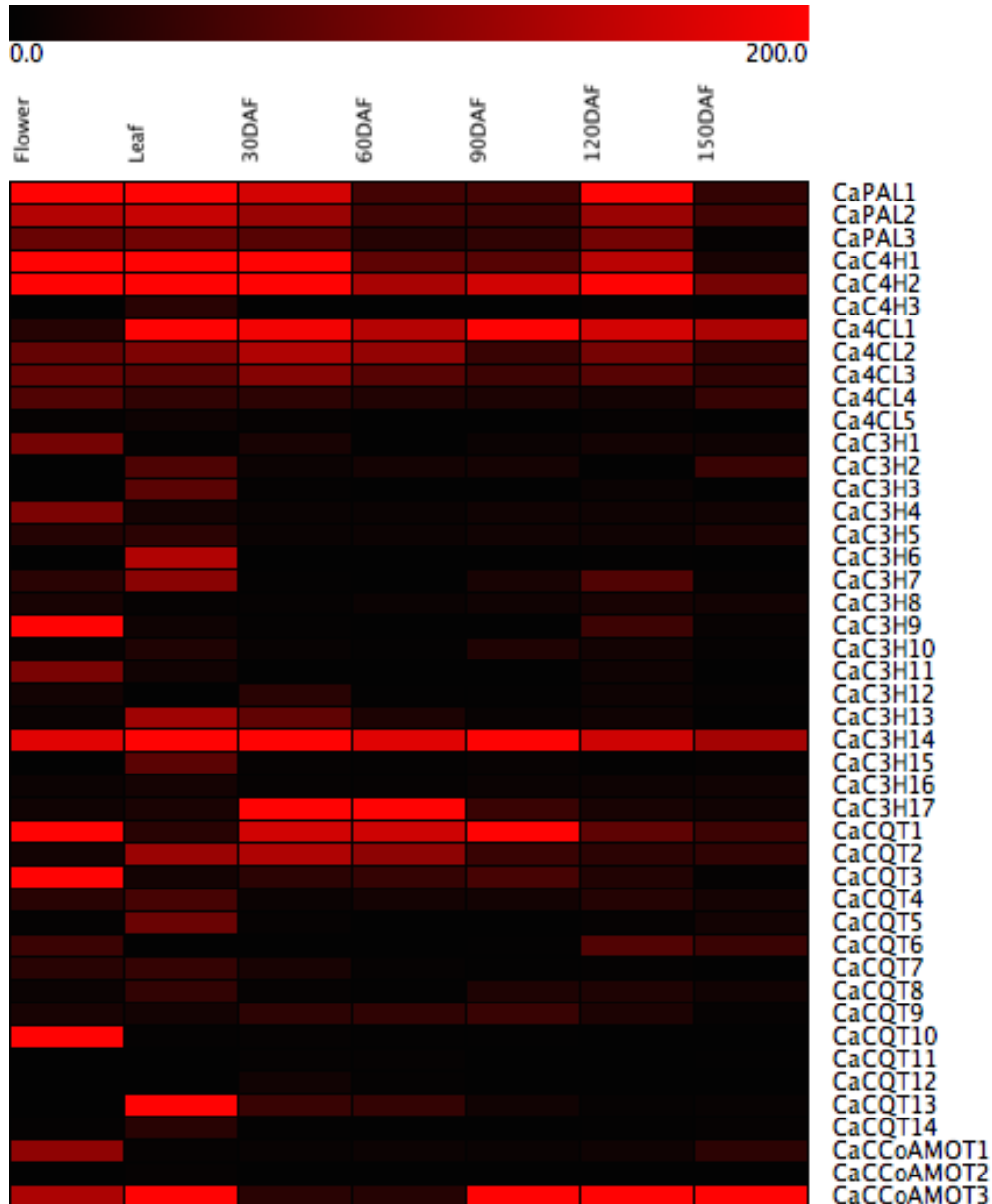


Figura 2: Padrão de atividade transcricional dos 45 genes candidatos identificados para a via de biossíntese dos CGAs. A cor vermelha indica que o gene estava transcricionalmente ativo, e a cor preta indica que o gene estava transcricionalmente inativo neste determinado órgão e/ou tecido analisado.

CONCLUSÃO

Foram identificados 45 genes relacionados com a formação de CGAs em flores, folhas e frutos. O padrão transcricional permitiu identificar genes com atividades específicas assim como constitutiva.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bakuradge, T. et al. (2011) Antioxidant-rich coffee reduces DNA damage, elevates glutathione status and contributes to weight control: results from intervention study. *Molecular Nutrition and Food Research*, 55: 793-797.
- Carazzolle, M.F., Rabello, F.R., Martins, N.F., de Souza, A.A., do Amaral, A.M., Freitas-Astua, J., ... & Mehta, A. (2011). Identification of defence-related genes expressed in coffee and citrus during infection by *Xylella fastidiosa*. *European journal of plant pathology*, 130(4), 529-540.
- Chang, S, Puryear, J., Cairney, J.A. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology*, 11, 113-116.
- Chevreur, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E., Wetter, T., & Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome research*, 14(6), 1147-1159.
- Dias, R.C.E., Campanha, F.G., Vieira, L.G.E., Ferreira, L.P., Pot, D., Marraccini, P., Benassi, M.D.T. (2010). Evaluation of kahweol and cafestol in coffee tissues and roasted coffee by a new high-performance liquid chromatography methodology. *Journal of agricultural and food chemistry*, 58(1), 88-93.
- Dudareva, N., Klemplen, A., Muhlemann, J.K., & Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist*.
- Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., ... & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research*, 36(10), 3420-3435.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., ... & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644-652.
- Huang, X., & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome research*, 9(9), 868-877.
- Kitzberger, C.S.G., Scholz, M.B.D.S., Pereira, L.F.P., Vieira, L.G.E., Sera, T., Silva, J.B.G.D., Benassi, M.D.T. (2013). Diterpenes in green and roasted coffee of *Coffea arabica* cultivars growing in the same edapho-climatic conditions. *Journal of Food Composition and Analysis*.
- Kochko, A. de. et al. (2003) Genetic mapping of caffeoylcoenzyme A 3-O-methyltransferase gene in coffee trees: impact in chlorogenic acid content. *Theoretical and Applied Genetics*, 107 (4):751-756.
- Ivamoto, S.T., et al. (2013) Diversidade nucleotídica de genes envolvidos na biossíntese de ácidos clorogênicos de cafeeiros. *Coffee Science*, 8(2): 148-156.
- Ivamoto, S.T. et al. (2015). Identification of the transcriptionally active cytochrome P450 repertoire in *Coffea arabica*. *Genetics and Molecular research*, 14(1): 2399-2412.
- Lulin, H., Xiao, Y., Pei, S., Wen, T., & Shangqin, H. (2012). The first Illumina-based de novo transcriptome sequencing and analysis of safflower flowers. *PloS one*, 7(6), e38653.
- McCarthy, J. et al. (2007) Chlorogenic acid synthesis in coffee: an analysis of CGA content and real-time RTPCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. *Plant Science*, 172: 861-1060.
- Mondego, J.M., Vidal, R.O., Carazzolle, M.F., Tokuda, E.K., Parizzi, L.P., Costa, G.G., ... & Pereira, G.A. (2011). An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. *BMC plant biology*, 11(1), 30.
- Pereira, L.F.P & Ivamoto, S.T. (2015). Characterization of coffee genes involved in isoprenoid and diterpenes metabolic pathway. In: Preedy, V.R. (Ed.). *Coffee in Health and Disease Prevention*. Academic Press.
- Shi, C.Y., Yang, H., Wei, C.L., Yu, O., Zhang, Z.Z., Jiang, C.J., ... & Wan, X.C. (2011). Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. *BMC genomics*, 12(1), 131.
- Ziemann, M., Kamboj, A., Hove, R.M., Loveridge, S., El-Osta, A., & Bhawe, M. (2013). Analysis of the barley leaf transcriptome under salinity stress using mRNA-Seq. *Acta Physiologiae Plantarum*, 1-10