

## IDENTIFICAÇÃO E SEQUENCIAMENTO DE UM *RESISTANCE GENE ANALOG* (RGA) EM HÍBRIDO DE TIMOR CIFIC 832/2

Isabel Samila Lima Castro<sup>2</sup>; Dênia Pires de Almeida<sup>3</sup>; Danúbia Rodrigues Alves<sup>4</sup>; Edson Mario de Andrade<sup>5</sup>; Brunna de Figueiredo Duarte<sup>6</sup>; Eveline Teixeira Caixeta<sup>7</sup>; Tiago Antônio de Oliveira Mendes<sup>8</sup>; Laércio Zambolim<sup>9</sup>

<sup>1</sup>Trabalho financiado pelo Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café, CNPq, Fapemig, INCT-Café

<sup>2</sup>Estudante, MSc, Universidade Federal de Viçosa, Viçosa - MG, isabel.castro@ufv.br

<sup>3</sup>Estudante, DSc, Universidade Federal de Viçosa, Viçosa - MG, denia\_pires@hotmail.com

<sup>4</sup>Estudante, MSc, Universidade Federal de Viçosa, Viçosa - MG, danubia.rodriguesalves@gmail.com

<sup>5</sup>Estudante, MSc, Universidade Federal de Minas Gerais, Belo Horizonte - MG, mariodeandrade@gmail.com

<sup>6</sup>Estudante, BS, Universidade Federal de Viçosa, Viçosa - MG, brunnafigueiredo09@gmail.com

<sup>7</sup>Pesquisadora, DSc, Embrapa Café, Brasília - DF, eveline.caixeta@embrapa.br

<sup>8</sup>Professor, DSc, Universidade Federal de Viçosa, Viçosa - MG, tiagomgmendes@yahoo.com.br

<sup>9</sup>Professor, PhD, Universidade Federal de Viçosa, Viçosa - MG, zambolim@ufv.br

**RESUMO:** O café, uma das *commodities* agrícolas mais importante no mundo, possui papel de destaque na economia de vários países. A ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix* é a principal doença da cultura. Seu controle exige gastos e, conseqüentemente, aumenta os custos de produção. Dessa forma, o uso de variedades resistentes a ferrugem é o mais indicado, pois além de eficaz, é econômico e sustentável. Nesse trabalho, selecionou-se um RGA (*Resistance Gene Analog*), a partir de dados do transcriptoma da interação cafeeiro-*H. vastatrix*, com potencial envolvimento no mecanismo de defesa da planta. A análise de expressão gênica do RGA mostrou diferença significativa entre as variedades resistente e suscetível, sendo o gene *up* regulado na variedade resistente às 12 e 24 horas após a inoculação (hai), sugerindo o seu envolvimento na resposta de defesa do hospedeiro. Foi realizado o rastreamento do RGA em uma biblioteca de clones BAC (*Bacterial Artificial Chromosome*) obtida do Híbrido de Timor (HdT) CIFIC 832/2. O sequenciamento do clone 101-2P, que possui o fragmento de interesse, revelou a presença de três RGAs que poderão fornecer informações importantes sobre o mecanismo de resistência da planta ao patógeno, assim como serem utilizados para a identificação e seleção de variedades resistentes nos programas de melhoramento do cafeeiro. A análise funcional e a caracterização desse RGA estão sendo realizadas.

**PALAVRAS-CHAVE:** Genômica, interação planta-patógeno, sistema imune da planta e genes de resistência

## IDENTIFICATION AND SEQUENCING OF A RESISTANCE GENE ANALOG (RGA) IN HYBRID TIMOR CIFIC 832/2

**ABSTRACT:** Coffee, one of the most important agricultural commodities in the world, has an important role in the economy of many countries. Coffee rust, caused by biotrophic fungus *Hemileia vastatrix*, is the main disease of the crop. Their control requires spendings and, consequently, increases production costs. Thus, the use of rust resistant varieties is more suitable, because, besides being effective, it is more economical and sustainable. In this work, an RGA (*Resistance Gene Analog*) was selected from transcriptome data of *coffee-H. vastatrix* interaction, with potential involvement in the plant's defense mechanism. RGA gene expression analysis showed significant difference between resistant and susceptible varieties, with the gene being up regulated in the resistant variety at 12 and 24 hours after inoculation (hai), suggesting its involvement in the host defense response. RGA screening was performed on a BAC (*Bacterial Artificial Chromosome*) clone library obtained from the Híbrido de Timor (HdT) CIFIC 832/2. The sequencing of clone 101-2P, which has the fragment of interest, revealed the presence of three RGAs that may provide important information about the plant's resistance mechanism to the pathogen, as well as for the identification and selection of resistant varieties in the coffee breeding programs. Functional analysis and characterization of this RGA are being performed.

**KEY WORDS:** Genomics, plant-pathogen interaction, plant immune system and resistance genes

## INTRODUÇÃO

O café, uma das *commodities* agrícolas mais importante no mundo, possui papel de destaque na economia de vários países. *Coffea arabica* e *Coffea canephora* são as espécies mais importantes economicamente (Davis et al., 2011). A ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix* é a principal doença da cultura. A doença pode atingir com gravidade toda a lavoura sendo responsável por perdas da ordem de 35 a 50% da produção se não forem implantadas medidas de controle eficiente. Seu controle exige gastos e, conseqüentemente, aumenta os custos de produção (Zambolim, 2016). Dessa forma, o uso de variedades resistentes a ferrugem é o mais indicado, pois além de eficaz, é econômico e sustentável. O modo de interação entre o patógeno e o hospedeiro envolve sofisticados

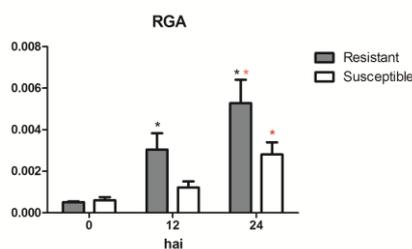
mecanismos de supressão do sistema imunológico e a indução de genes específicos na planta para o estabelecimento da colonização pelo fungo (Schulze-Lefert e Panstruga, 2003; Guimarães et al., 2015). Dessa forma, a identificação de genes envolvidos no mecanismo de defesa da planta é essencial para a identificação e seleção de plantas resistentes a ferrugem. Derivados do Híbrido de Timor (HdT), híbrido natural interespecífico de *C. arabica* e *C. canephora*, tem sido utilizados como a principal fonte de genes de resistência à *H. vastatrix*. Tendo em vista sua importância, Cação et al. (2013) construíram uma biblioteca de cromossomo artificial bacteriano (BAC) do HdT CIFC 832/2, genótipo parental de vários cultivares modernas do cafeeiro, com o objetivo de disponibilizar uma ferramenta útil para a identificação de genes de interesse. A maioria dos genes de resistência (genes *R*) clonados e sequenciados até o momento contêm um sítio de ligação a nucleotídeo (NBS) e um domínio de repetição rica em leucina (LRR) onde são encontrados motivos altamente conservados (Rommens & Kishore, 2000). Os *RGAs* (*Resistance genes analogs*), assim como os genes *R*, possuem domínios conservados, principalmente NBR, ligados a resistência. Dessa forma, a identificação e caracterização de *RGAs* são fundamentais para o desenvolvimentos de estratégias para os programas de melhoramento que visam a obtenção de variedades resistentes. O objetivo desse trabalho foi identificar e sequenciar um *RGA* com potencial envolvimento no mecanismo de defesa do cafeeiro à *H. vastatrix*.

## MATERIAL E MÉTODOS

A partir de dados do transcriptoma da interação entre cafeeiros resistentes (HdT CIFC 832/1) / suscetíveis (Caturra 19/01) e *H. vastatrix* raça XXXIII (Florez et al., 2017), foi selecionado um *RGA* potencialmente envolvido no mecanismo de resistência da planta. Esse *RGA* foi identificado somente na biblioteca de transcritos da interação incompatível. Para a avaliação da expressão gênica, foi montado um experimento inteiramente casualizado com três réplicas biológicas para cada tempo de coleta após a inoculação de urediniósporos de *H. vastatrix* raça XXXIII (0, 12 e 24 horas após a inoculação) em plantas resistentes (HdT CIFC 832/1) e suscetíveis (Caturra 19/01). As folhas coletadas em cada tempo, previamente congeladas, foram maceradas em nitrogênio líquido e aproximadamente 100 mg foram destinadas a extração de RNA utilizando-se o Concert® (Invitrogen) como extrator. A partir da sequência do *RGA* selecionado, obtida pela montagem por referência do transcriptoma da interação, foram desenhados *primers* que a amplifica. A síntese de cDNA foi realizada utilizando o kit ImProm-II™ *Reverse Transcription System Protocol RT-PCR* (Promega), segundo as orientações do fabricante. Para a realização da PCR quantitativa em tempo real foi utilizado o sistema de detecção de fluorescência SYBR Green I. O nível de expressão dos genes foi calculado utilizando os valores médios de Ct resultante de três réplicas biológicas e três réplicas técnicas por meio de curva padrão relativa. As análises estatísticas foram realizadas usando o *GraphPad Prism*. O *RGA* selecionado foi rastreado a partir de uma biblioteca de 56.832 clones BAC oriunda de um clone de Híbrido de Timor CIFC 832/2, resistente a ferrugem, construída por Cação et al. (2013) pelo método de decomposição de *pool* de clones BAC. O DNA da BAC identificada foi extraído utilizando o *Kit Wizard® SV Plus Minipreps DNA Purification System* (Promega), seguindo o protocolo do fabricante. O clone contendo o fragmento correspondente do gene de interesse foi sequenciado por meio da plataforma Illumina MiSeq, utilizando *MiSeq Reagent Kit v2*, seguindo as recomendações do fabricante. A qualidade das *reads* foi avaliada utilizando o *software* FastQC 0.11.5. O *software* trimmomatic foi utilizado, com valor de Phred=30, para a seleção das *reads* de boa qualidade e eliminação dos adaptadores. A montagem da sequência do clone foi realizada utilizando a estratégia *de novo* por meio do *software* SPAdes. A predição de genes foi feita por meio do *software* AUGUSTUS. *Solanum lycopersicum* foi utilizado como genoma de referência por ser a espécie mais próxima filogeneticamente disponível nesse *software* até a presente data. A anotação gênica foi feita por meio da ferramenta tBLASTx com base nos genomas das espécies *C. arabica*, *C. canephora* e *Coffea eugenioides*, disponíveis no *National Center for Biotechnology Information* (NCBI).

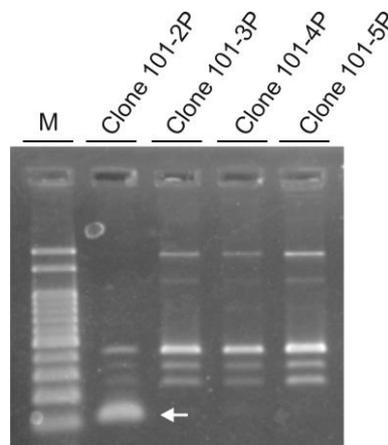
## RESULTADOS E DISCUSSÃO

A identificação e caracterização de *RGAs* é extremamente importante para o desenvolvimento de estratégias de melhoramento de plantas visando à resistência a patógenos. A análise de expressão gênica do *RGA* selecionado a partir de dados do transcriptoma mostrou expressão significativa diferenciada entre o genótipo resistente e suscetível em 12 e 24 hai, sendo esse gene *up* regulado no genótipo resistente (Figura 1). O padrão de expressão desse gene se mostrou crescente durante as horas após a infecção nos dois genótipos, porém com uma maior e mais rápida expressão no genótipo resistente na fase inicial (12-24 hai) da infecção. Esse padrão de expressão reforça a hipótese de que a resistência a patógenos em plantas pode depender da taxa e extensão da síntese de uma ou mais enzimas induzidas no hospedeiro pelo patógeno (Diniz et al. 2012). Diniz et al. (2012) sugeriram que a resposta rápida de resistência limita a formação de haustórios, estrutura especializada responsável pela absorção de nutrientes a partir do hospedeiro, além da indução de mudanças estruturais na célula hospedeira (Ramiro et al., 2009), e isso pode ser a base para o sucesso da resistência da planta à infecção pelo patógeno.



**Figura 1:** Validação do gene *RGA* por qPCR. O eixo y representa o nível de expressão em relação aos genes endógenos em três momentos, 0, 12 e 24 horas após a inoculação (hai) por *H. vastatrix*. Barras de erros = SEM, n = 3 réplicas biológicas independentes. Asterisco (\*) indica diferença significativa do nível de expressão na mesma hai entre as interações resistente e suscetível. Asterisco (\*\*) indica diferença significativa do nível de expressão entre as hai dentro da mesma interação.

Com o objetivo de obter a sequência completa do *RGA*, este foi rastreado em uma biblioteca de clones BAC e identificado no clone BAC 101-2P (placa 101, coluna 2 e linha P) (Figura 2). A partir do sequenciamento desse clone, foi obtido um total de 81352 *reads pair* com boa qualidade utilizadas para a montagem. Apesar dos genomas de referência de *C. canephora* e *C. arabica* estarem disponíveis, optou-se pela estratégia de montagem *de novo*, pois o loco sequenciado pertence a um híbrido interespecífico, o Híbrido de Timor. Dessa forma, mudanças estruturais nesse loco poderiam não ser identificadas por meio da montagem utilizando um genoma de referência.



**Figura 2:** Gel de agarose 1,5 %. M: marcador de peso molecular 100 pb. Seta: indica a banda de interesse presente no clone 101-2P. Demais bandas foram consideradas inespecíficas.

A montagem realizada com o *software* SPAdes gerou 2600 *contigs* e 2597 *scaffolds* com valor de N50=1615. O tamanho total da montagem foi de 221871pb sendo o maior *scaffold* com 56551 pb (Tabela 1). A partir dos *scaffolds* foram preditos 1645 ORFs (*Open Reading Frame*) por meio do *software* AUGUSTUS. Foram filtrados somente as ORFs que apresentaram tamanho acima de 100 pares de bases (pb), totalizando 1636 ORFs. Esses possíveis genes foram então anotados utilizando a ferramenta tBLASTx e os genomas de *C. arabica*, *C. canephora* e *C. eugenioides*. Um total de 297 genes foram anotados (dados não mostrados). Vinte genes (~7%) foram anotados como proteínas não caracterizadas, enquanto que a grande maioria foi anotada como proteína que participa do metabolismo primário da planta. Três genes (*scaffold1g1*, *scaffold1g3* e *scaffold3g5*) foram anotados como proteínas RGAs, confirmando que a montagem foi eficiente, apesar de relativa fragmentação observada. As sequências desses três genes foram submetidas ao *software* Pfam para a análise de domínios conservados em proteínas. O resultado indicou que as proteínas codificadas pelos genes identificados possuem os domínios Rx-N e NB-ARC, ambos encontrados em proteínas envolvidas na resistência de plantas à patógenos (Rommens & Kishore, 2000). A maioria das proteínas de resistência nas plantas possuem sítio de ligação de nucleotídeos e repetição rica em leucina (NLR, também conhecidos como superfamília NB-LRR) (Lee and Yeom, 2015). Tais proteínas possuem domínios NB-ARC na região central e LRR na região C-terminal. Já foi demonstrado que genes codificadores de NB sem LRRs também pode funcionar na imunidade da planta (Nandety et al., 2013).

**Tabela 1:** Dados quantitativos da montagem do clone 101-2P utilizando o *software* SPAdes.

Características	Números
<i>Reads</i> de alta qualidade	81352
Tamanho total	221871pb*
Número de <i>scaffolds</i>	2600
Maior <i>scaffold</i>	56551pb
N50	1615
L50	11
Tamanho total (>= 1000 pb)	121403pb
Tamanho total (>= 5000 pb)	94573pb
Tamanho total (>= 10000 pb)	85323pb
Tamanho total (>= 25000 pb)	56551pb
Tamanho total (>= 50000 pb)	56551pb
CG (%)	43,53

\*pb: Pares de bases

## CONCLUSÕES

1 - As informações obtidas nesse trabalho foram úteis para o entendimento inicial da estrutura do loco *RGA* presente no cafeeiro HdT 832/2, fornecendo informações importantes que poderão ser utilizadas para a compreensão da origem do loco *RGA* e da função do mesmo durante a resposta de defesa da planta.

2 - Observou-se que a região sequenciada em análise apresenta três diferentes *RGAs*. Estudos funcionais e de caracterização desses genes estão sendo realizados.

3 - A união dessas informações poderá auxiliar o desenvolvimento de estratégias que visam o controle da doença por meio da identificação e seleção de variedades resistentes.

4 - Dessa forma, os três *RGAs* identificados constituem alvos importantes para os programas de melhoramento do cafeeiro que visam a resistência à *H. vastatrix*.

## REFERÊNCIAS BIBLIOGRÁFICAS

- CAÇÃO SMB, SILVA NV, DOMINGUES DS et al (2013) Construction and Characterization of a BAC Library From the *Coffea arabica* Genotype Timor Hybrid C1FC 832/2. *Genetica*. doi: 10.1007/S10709-013-9720-Y
- DAVIS AP, TOSH J, RUCH N, FAY MF (2011) Growing coffee: *Psilanthus (Rubiaceae)* subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot J Linn Soc* 167:357–377. doi:10.1111/j.1095-8339.2011.01177.x.
- DINIZ I, TALHINHAS P, AZINHEIRA HG et al (2012) Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and nonhost resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. *European Journal Plant Pathology*, 133, 141-157.
- FLOREZ JC, MOFATTO LS, LOPES RLF, FERREIRA SS, ZAMBOLIM EM, CARAZZOLLE MF, ZAMBOLIM L, CAIXETA ET (2017) High throughput transcriptome analysis of coffee reveals prehaustorial resistance in response to *Hemileia vastatrix* infection. *Plant Mol Biol*; 95:607–623 DOI 10.1007/s11103-017-0676-7
- GUIMARÃES LG, TENENTE R, PINHEIRO C, CHAVES I, SILVA MC, CARDOSO FNH et al (2015) Proteomic analysis of apoplastic fluid of *Coffea arabica* leaves highlights novel biomarkers for resistance against *Hemileia vastatrix*. *Frontiers In Plant Science*. Volume 6.
- LEE HA, AND YEOM SI (2015) Plant NB-LRR proteins: tightly regulated sensors in a complex manner. *Brief Funct Genomics* 14, 233–242. doi: 10.1093/bfgp/elv012
- NANDETY RS, CAPLAN JL, CAVANAUGH K, PERROUD B, WROBLEWSKI T, MICHELMORE RW et al (2013) The role of TIR-NBS and TIR-X proteins in plant basal defense responses. *Plant Physiol*. 162, 1459–1472. doi: 10.1104/pp.113.219162
- ROMMENS CM, KISHORE GM (2000) Exploiting the full potential of disease-resistance genes for agricultural use. *Current Opinion Biotechnology*, London, v.11, n.2, p.120-25.
- SCHULZE-LEFERT P, PANSTRUGA R (2003) Establishment of biotrophy by parasitic fungi and reprogramming of host cells for disease resistance. *Annu Rev Phytopathol* 41:641–667. doi: 10.1146/annurev.phyto.41.061002.083300
- ZAMBOLIM L (2016) Current Status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology*, 41(1), 1-8. doi 10.1007/S40858-016-0065-9