



BRENDA NEVES PORTO

**SEQUENCIAMENTO DO GENOMA E
IDENTIFICAÇÃO DE CANDIDATOS A
EFETORES DE *Hemileia vastatrix***

LAVRAS - MG

2016

BRENDA NEVES PORTO

**SEQUENCIAMENTO DO GENOMA E IDENTIFICAÇÃO DE
CANDIDATOS A EFETORES DE *Hemileia vastatrix***

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Doutora.

Orientador

PhD. Mário Lúcio Vilela de Resende

LAVRAS - MG

2016

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Porto, Brenda Neves.

Sequenciamento do genoma e identificação de candidatos a
efetores de *Hemileia vastatrix* / Brenda Neves Porto. – Lavras :
UFLA, 2016.

119 p. : il.

Tese(doutorado)–Universidade Federal de Lavras, 2016.

Orientador: Mário Lúcio Vilela de Resende.

Bibliografia.

1. Ferrugem do cafeeiro. 2. Interação planta-patógeno. 3. NGS
- Illumina e PacBio. 4. Proteoma I. Universidade Federal de
Lavras. II. Título.

BRENDA NEVES PORTO

**SEQUENCIAMENTO DO GENOMA E IDENTIFICAÇÃO DE
CANDIDATOS A EFETORES DE *Hemileia vastatrix***

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Doutora.

APROVADA em 25 de Fevereiro de 2016.

Dra. Eveline Teixeira Caixeta EMBRAPA Café

Dr. Ronaldo José Durigan Dalio Instituto Agrônômico de Campinas - IAC

Dr. Jorge Teodoro de Souza UFLA

Dr. Luciano Vilela Paiva UFLA

PhD. Mário Lúcio Vilela de Resende
Orientador

**LAVRAS - MG
2016**

À Izabel Oliveira Neves, minha mãe, que sempre me incentivou e apoiou em todos os momentos da minha vida.

DEDICO

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me iluminar e me dar forças para superar os obstáculos, não desistir e finalizar este trabalho.

À Universidade Federal de Lavras, em especial ao Programa de Pós-Graduação em Biotecnologia vegetal, pela oportunidade de realização do curso de doutorado.

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

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por intermédio do Instituto Nacional de Ciências e Tecnologia do Café (INCT-Café), pela concessão da bolsa de doutorado sanduíche no exterior e apoio financeiro ao nosso projeto.

Ao Consórcio de Pesquisa Café (CBP&D/Café) pelo apoio financeiro.

Ao Delaware Biotechnology Institute (DBI), onde foram realizadas as atividades de sequenciamento e montagem do genoma do fungo *Hemileia vastatrix* e à University of Delaware pela bolsa concedida por três meses.

Ao Laboratório de Biotecnologia do Cafeeiro (BioCafé), localizado na Universidade Federal de Viçosa (UFV), onde foram desenvolvidas as atividades laboratoriais.

Ao Núcleo de Análise de Biomoléculas (NuBioMol), localizado na Universidade Federal de Viçosa (UFV), onde foram desenvolvidas as atividades de bioinformática.

Ao prof. Dr. Mário Lúcio Vilela de Resende, pela oportunidade de desenvolvimento do trabalho, pela confiança depositada e orientação.

À Dra. Eveline Teixeira Caixeta, por sua coorientação, sua atenção, ensinamentos e auxílio sempre que precisei.

À Dra. Nicole Donofrio e ao Dr. Shalwn Polson pelos ensinamentos compartilhados.

Aos membros da banca pela disponibilidade e ensinamentos.

Aos professores do Programa de Pós-Graduação em Biotecnologia Vegetal pelos conhecimentos transmitidos em suas aulas teóricas e práticas.

Aos colegas e amigos do Programa de Pós-Graduação em Biotecnologia Vegetal e do Laboratório BioCafé, pela amizade, convivência e apoio técnico.

Aos amigos de diferentes partes do mundo, que fiz nos Estados Unidos, pelo aprendizado cultural e troca de ideias profissionais.

Aos brasileiros que conheci em Delaware e tornaram-se parte da minha família.

Às amigas de longas datas, feitas nos diferentes lugares por onde passei (Santa Maria, Lavras e Viçosa) que, mesmo distantes, deram-me força e torceram por mim.

À minha família, em especial minha mãe, que sempre me deu força e apoio em todos os momentos da minha vida.

A todos que colaboraram direta ou indiretamente para a realização deste trabalho, meus mais sinceros agradecimentos.

MUITO OBRIGADA!!!

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”.

Marthin Luther King.

RESUMO GERAL

A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* (Berkely & Broome), é responsável pela principal doença foliar que afeta a produção do café arábica (*Coffea arabica* L.). Quando não controlada, essa doença pode ocasionar perdas de até 50% da produção. O uso de cultivares de café resistentes à *H. vastatrix*, é a estratégia mais eficiente de controle dessa doença. Entretanto, a obtenção de genótipos resistentes tem sido um desafio devido ao alto potencial adaptativo do fungo e, conseqüentemente, o surgimento de novas raças fisiológicas do patógeno que suplantam as cultivares resistente. Durante a interação com o cafeeiro, o fungo secreta proteínas efetoras que modificam a estrutura e função da célula hospedeira, permitindo o estabelecimento da colonização parasitária. Estudos genômicos estão ajudando no entendimento dos mecanismos moleculares envolvidos no processo de interação entre planta-patógeno e no desenvolvimento de técnicas moleculares para a identificação de isolados individuais. Nosso objetivo foi sequenciar o genoma desse fungo e identificar genes que possam contribuir para a patogenicidade. Usando uma estratégia de montagem híbrida e duas plataformas de sequenciamento de nova geração, PacBio RS II e Illumina – HiSeq 2500, obteve-se um genoma parcial do isolado HV-02 (raça XXXIII) de Hv com tamanho de 576 Mb. Verificou-se que 96,37% dos genes conservados de eucariotos estavam presentes no genoma anotado de Hv, indicando um elevado nível de integridade durante o processo de montagem. Sendo que 13.034 codificam proteínas. A análise da semelhança entre as proteínas previstas no genoma e as sequências proteicas de Hv com outros fungos, mostrou que 74% apresentaram hit dentro da ordem Pucciniales, especialmente com *Puccinia. graminis* f. sp. *tritici* e *Melampsora larici-populina* e 12% não apresentaram similaridade (no-hit) com qualquer proteína descrita nos bancos de dados analisados, sendo consideradas exclusivas de Hv. Com o proteoma deduzido foi predito o secretoma. Foram identificadas 615 proteínas contendo peptídeo sinal localizado na via de secreção e sem domínios transmembranares. A partir do secretoma obtido, 111 proteínas foram consideradas candidatas a efetores específicos de Hv. Foram selecionados 17 genes (*EHv33*) para serem validados por PCR em tempo real. As análises da expressão temporal desses genes (*EHv33*) mostraram que, a maior parte deles foram mais expressos, significativamente, depois da formação dos haustórios, na interação compatível, podendo ser considerados genes candidatos a efetores translocados via haustório.

Palavras-chave: Ferrugem do cafeeiro. Genoma. Proteoma. Secretoma. Expressão gênica. Interação cafeeiro - *H. vastatrix*.

GENERAL ABSTRACT

Coffee rust caused by the fungus *Hemileia vastatrix* (Berkely & Broome) is responsible for the main leaf disease that affects the production of *Coffea arabica* L. When chemical control is not used, this disease can cause losses of up to 50% of the production. The use of coffee cultivars resistant to *H. vastatrix* (*Hv*) is the most efficient strategy for controlling this disease. However, obtaining resistant genotypes has been a challenge due to the high adaptation potential of the fungus and, consequently, the emergence of new physiological strains of the pathogen, which overcome cultivar resistance. During the interaction with coffee, the fungus secret a lot of effector proteins that modify the structure and function of the host cell, allowing or not the establishment of infection, depending on the host genotype. Functional genomics studies are enlightening the molecular mechanisms involved in the plant-pathogen interaction and the development of molecular techniques for identification of individual isolates should be pursued. The objective of this work was to sequence the whole genome of this fungus and identify genes that may contribute for microbial pathogenicity or host resistance. Using a strategy of hybrid assembly and two new generation sequencing platforms, PacBio RS II and Illumina – HiSeq 2500, we obtained a partial genome of the isolate HV-02 (strain XXXIII of *Hv*) with the size of 576 Mb. We verified that 96.37% of the conserved eukaryotes genes were present in the annotated *Hv* genome, indicating an elevated level of integrity during the assembly process, being coded 13,034 proteins. The similarity analysis of the proteins predicted in the genome and the protein sequences between *Hv* and other fungi showed that 74% hit within the Pucciniales order, especially with *Puccinia graminis* f. sp. *tritici* and *Melampsora larici-populina*, and 12% presented no similarity with any protein described in the analyzed databanks, being considered exclusive to *Hv*. With the deduced proteome, we predicted the functional secretome. We identified 615 signal peptide containing proteins located in the secretion pathway and with no transmembrane domains. Within the obtained secretome, 111 proteins were considered candidates for specific *Hv* effectors. We selected 17 putative genes (*EHv33*) to be validated by real-time PCR. The analyses of temporal expression of these genes (*EHv33*) showed that most of them were significantly up regulated after the formation of haustoria, in the compatible interaction, and can be considered candidate genes for effectors translocated via haustoria.

Keywords: Coffee rust. Genome. Proteome. Secretome. Gene expression. Coffee-*H. vastatrix* interaction.

LISTAS DE FIGURAS

CAPÍTULO 2

- Figura 1 Representação esquemática das três diferentes estratégias utilizadas na montagem do genoma do fungo *H.vastatrix* raça XXXIII. Retângulos azuis representam a primeira estratégia de montagem, a qual, utilizou somente os *reads* gerados pela plataforma PacBio. Retângulos verdes representam a segunda melhor estratégia de montagem, a qual utilizou somente os *reads* gerados pela plataforma Illumina. Retângulos vermelhos representam a terceira e mais completa estratégia de montagem do genoma de *H. vastatrix*, a qual utilizou os *reads* gerados pelas duas plataformas de sequenciamento de nova geração.61
- Figura 2 Resultado da busca por similaridade entre as sequências de proteínas presentes no genoma de *H. vastatrix* raça XXXIII contra os bancos de dados públicos de sequências de proteínas não redundantes (NCBI e UNIPROT), por meio do algoritmo BLASTp. No eixo y estão as espécies encontradas com maior frequência na busca por similaridade. No eixo x estão os números de proteínas encontradas para cada espécie64
- Figura 3 Resultado da análise de categorização funcional das proteínas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Função Molecular, no nível hierárquico 03. No eixo x estão

	os números de proteínas encontradas para cada termo do GO nessa categoria65
Figura 4	Resultado da análise de categorização funcional das proteínas encontradas no genoma de <i>H. vastatrix</i> raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o <i>software</i> BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Processo Biológico, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria66
Figura 5	Resultado da análise de categorização funcional das proteínas encontradas no genoma de <i>H. vastatrix</i> raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o <i>software</i> BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Componente Celular, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria67
Figura 6	Resultado da análise de categorização funcional das proteínas encontradas no genoma de <i>H. vastatrix</i> raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o <i>software</i> BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Componente Celular, nos níveis hierárquicos 06 e 08. No eixo

x estão os números de proteínas encontradas para cada termo do GO nessa categoria68

CAPÍTULO 3

- Figura 1 Resultado da busca por similaridade entre as sequências de proteínas secretadas presentes no genoma de *H. vastatrix* raça XXXIII contra os bancos de dados públicos de sequências de proteínas não redundantes (NCBI e UNIPROT), por meio do algoritmo BLASTp.....91
- Figura 2 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa (*e-value* < 10^{-5}) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software BLAST2GO*93
- Figura 3 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa (*e-value* < 10^{-5}) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software BLAST2GO*94
- Figura 4 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa (*e-value* < 10^{-5}) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software BLAST2GO*95

Figura 5	Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de <i>H. vastatrix</i> raça XXXIII, que apresentaram assinaturas específicas nas categorias do KOG.....	96
Figura 6	<i>Heatmap</i> resultante da análise do CD-HIT. Vermelho = proteínas secretadas exclusivas de cada fungo.....	97
Figura 7	Análise de expressão por PCR em tempo real (qRT-PCR) de seis genes <i>EHv33</i> , que codificam possíveis proteínas secretadas candidatas a efetores de <i>H. vastatrix</i> raça XXXIII	103
Figura 8	Análise de expressão por PCR em tempo real (qRT-PCR) de quatro genes <i>EHv33</i> , que codificam possíveis proteínas secretadas candidatas a efetores de <i>H. vastatrix</i> raça XXXIII	104
Figura 9	Análise de expressão por PCR em tempo real (qRT-PCR) de cinco genes <i>EHv33</i> , que codificam possíveis proteínas secretadas candidatas a efetores de <i>H. vastatrix</i> raça XXXIII	106
Figura 10	Análise de expressão por PCR em tempo real (qRT-PCR) de dois genes <i>EHv33</i> , que codificam possíveis proteínas secretadas candidatas a efetores de <i>H. vastatrix</i> raça XXXIII	107

LISTAS DE TABELAS

CAPÍTULO 2

Tabela 1	Informações estatísticas sobre a montagem do genoma do fungo <i>Hemileia vastatrix</i> raça XXXIII.....	62
Tabela 2	Resultados da análise de anotação estrutural do genoma do fungo <i>Hemileia vastatrix</i> raça XXXIII.....	63

CAPÍTULO 3

Tabela 1	Sequência dos <i>primers</i> desenhados no programa Primer Express 2.0 para análise de qRT-PCR.....	88
Tabela 2	Resultado da busca por proteínas nos <i>scaffolds</i> de <i>Hemileia vastatrix</i> raça XXXIII.....	98
Tabela 3	Domínios conservados presentes nas 11 proteínas de <i>Hemileia vastatrix</i> raça XXXIII, codificadas pelos 72 <i>scaffolds</i> presentes no genoma montado	99
Tabela 4	<i>Primers</i> escolhidos com a sequência de 13 genes que codificam possíveis proteínas candidatas a efetores de <i>H. vastatrix</i> , com base nos resultados das comparações entre os diferentes bancos de dados utilizados.....	101

SUMÁRIO

CAPÍTULO 1 Introdução geral	17
1 INTRODUÇÃO	17
2 REFERENCIAL TEÓRICO	21
2.1 Ferrugem alaranjada do cafeeiro – <i>Hemileia vastatrix</i>	21
2.2 Interações cafeeiro x <i>Hemileia vastatrix</i>	25
2.3 Proteínas efetoras de fungo, identificação e caracterização funcional.....	29
2.4 Sequenciamento de fungos fitopatogênicos	34
REFERÊNCIAS	40
CAPÍTULO 2 SEQUENCIAMENTO, MONTAGEM E ANOTAÇÃO DO GENOMA DO FUNGO <i>Hemileia vastatrix</i>	51
1 INTRODUÇÃO	53
2 MATERIAL E MÉTODOS.....	55
2.1 Multiplicação dos esporos de <i>H. vastatrix</i>	55
2.2 Germinação dos esporos de <i>H. vastatrix</i>	56
2.3 Extração de DNA de <i>H. vastatrix</i>	56
2.4 Sequenciamento e Montagem do genoma de <i>H. vastatrix</i>	57
2.5 Anotação estrutural do genoma de <i>H. vastatrix</i> e predição de genes	58
2.6 Anotação funcional das proteínas do genoma de <i>H. vastatrix</i>	59
3 RESULTADOS.....	60
4 DISCUSSÃO.....	69
5 CONCLUSÃO	73
REFERÊNCIAS	74
CAPÍTULO 3 PREDIÇÃO DO SECRETOMA, IDENTIFICAÇÃO E VALIDAÇÃO DE EFETORES DE <i>Hemileia vastatrix</i>	77
1 INTRODUÇÃO	79
2 MATERIAL E MÉTODOS.....	82
2.1 Multiplicação dos esporos de <i>H. vastatrix</i>	82
2.2 Teste de viabilidade dos esporos de <i>H. vastatrix</i>	83
2.3 Material vegetal e inoculação dos esporos de <i>H. vastatrix</i>	83
2.4 Extração de RNA e síntese de cDNA	84
2.5 Predição do secretoma de <i>H. vastatrix</i>	85
2.6 Anotação das proteínas secretadas de <i>H. vastatrix</i> com características de proteínas efetoras	85
2.7 Desenho de <i>primers</i> para RT-qPCR e teste de eficiência	87
2.8 PCR em tempo real (RT-qPCR).....	89
3 RESULTADOS.....	91

4	DISCUSSÃO	108
5	CONCLUSÃO	115
	REFERÊNCIAS	116

CAPÍTULO 1 Introdução geral

1 INTRODUÇÃO

O Brasil destaca-se por ser o maior produtor, consumidor e exportador de café do mundo. As duas espécies de café exploradas comercialmente são *Coffea arabica* (café arábica) e *C. canephora* (café conilon), representando, respectivamente, 72 e 28% da produção nacional. O estado de Minas Gerais é o maior produtor, sendo responsável por mais de 50% da produção nacional com uma área de 1.194,1 mil hectares plantados (BRASIL, 2015).

A espécie *C. arabica* é a mais cultivada e comercializada no mundo em razão da superioridade da qualidade de bebida. Essa espécie produz grãos de melhor qualidade com baixos teores de cafeína, no entanto, é altamente suscetível a doenças. Dentre as doenças, destaca-se a ferrugem alaranjada, causada pelo fungo *Hemileia vastatrix* (Berkely & Broome) responsável pela queda precoce das folhas. Como consequência, a produção do próximo ano é afetada com perdas de até 50%, dependendo das condições climáticas e do manejo realizado (ZAMBOLIM; VALE; ZAMBOLIM, 2005).

Atualmente, o controle da doença é realizado, principalmente, pelo uso de fungicidas (POZZA et al., 2010), contudo, apresenta muitas desvantagens, como por exemplo, o alto custo da produção (ABRAHÃO et al., 2009; AMARAL et al., 2008). Em função disso, a melhor opção de controle é o uso de cultivares resistentes, por ser mais eficiente, econômico e não causar impactos ambientais (ZAMBOLIM; VALE; ZAMBOLIM, 2005). Entretanto, em decorrência da alta variabilidade genética de *H. vastatrix*, o desenvolvimento de cultivares de café com resistência durável representa um desafio, pois favorece o aparecimento de novas raças fisiológicas do patógeno que ao longo do tempo,

pode levar à suplantação da resistência das cultivares melhoradas (CABRAL et al., 2009; MAIA et al., 2013; VÁRZEA; MARQUES, 2005).

Esse fungo está presente na maioria das regiões cafeeiras do mundo e já foram identificadas 49 raças fisiológicas de *H. vastatrix* (GICHURU et al., 2012), sendo 16 delas detectadas no Brasil (I, II, III, VII, X, XIII, XV, XVI, XVII, XXI, XXII, XXIII, XXIV, XXV ou XXXI, XXXIII e XXXVII) (CABRAL et al., 2009; CAPUCHO et al., 2012; ZAMBOLIM; VALE; ZAMBOLIM, 2005). A raça II é encontrada com maior frequência no Brasil (ZAMBOLIM; VALE; ZAMBOLIM, 2005) e apresenta apenas o gene de virulência v5. Já, a raça XXXIII, identificada recentemente (CAPUCHO et al., 2012), possui dois ou três genes de virulência (v5,7 ou v5,7,9), e tem sido capaz de suplantar a resistência de algumas cultivares de café derivadas de Híbrido de Timor, considerada a principal fonte de resistência utilizada nos programas de melhoramento (GICHURU et al., 2012).

Estudos realizados no Centro de Investigações das Ferrugens do Cafeeiro (CIFC), em Portugal, têm mostrado que a interação do cafeeiro-*H. vastatrix* segue a teoria de coevolução proposta por Flor (NORONHA-WAGNER; BETTENCOURT, 1967). Dessa forma, a resistência à ferrugem é condicionada por pelo menos nove genes dominantes (S_{H1} a S_{H9}), podendo-se inferir a existência de pelo menos nove genes de avirulência em *H. vastatrix*. Durante a interação com o cafeeiro, esse fungo, como as demais ferrugens, secreta proteínas efectoras que melhoram seu potencial de infecção, uma vez que são capazes de alterar a fisiologia do hospedeiro, suprimindo as respostas de defesa. Com esse mecanismo, o fungo consegue estabelecer a sua colonização parasítica (HOGENHOUT et al., 2009).

Conhecer os mecanismos moleculares de defesa do cafeeiro contra *H. vastatrix* e as diversas estratégias utilizadas pelo fungo para infectar seu hospedeiro são fundamentais para compreender a suplantação da resistência

pelas novas raças fisiológicas de *H. vastatrix* e, também, para o desenvolvimento de formas alternativas do controle da doença. Desta forma, as técnicas moleculares contribuem para elucidar os complexos mecanismos bioquímicos e moleculares envolvidos nessa interação planta-patógeno. Uma das metodologias moleculares largamente utilizadas atualmente é a análise do genoma dos organismos por meio de estudos nas áreas de genômica, transcriptômica e efetorômica.

A genômica foi, em grande parte, impulsionada pelo desejo de compreender a organização e função do genoma de fungos. Ao longo dos últimos 10 anos, com o desenvolvimento das tecnologias de sequenciamento de nova geração (NGS – *Next Generation Sequencing*), aumentou-se o número de genomas sequenciados de todas as classes de seres vivos, principalmente, dos fungos fitopatogênicos que causam as principais doenças agrícolas. Até o presente momento, foram sequenciados 28 genomas de patógenos de plantas. Desses, cinco são ferrugens, incluindo o genoma parcial de *H. vastatrix* (CRISTANCHO et al., 2014; KERSEY et al., 2014),

Apesar de muitos anos de pesquisa neste patossistema, os genes Avr complementares de *H. vastatrix* foram pouco estudados. Até o momento, alguns poucos efetores foram identificados ou caracterizados por métodos bioquímicos ou em nível molecular, por meio do sequenciamento de sequências expressas (ESTs), considerando urediniósporos germinados, apressório e haustório. Essas informações são ainda insuficientes, uma vez que conseguem identificar apenas uma pequena fração do secretoma total presente no genoma desse fungo (DUPLESSIS; JOLY; DODDS, 2012; FERNANDEZ; TALHINHAS; DUPLESSIS, 2013; TALHINHAS et al., 2014).

Um dos aspectos mais importantes na busca de informações biológicas em um contexto genômico é a associação das sequências de DNA com as suas respectivas funções biológicas. Visando entender melhor a interação do cafeeiro-

H. vastatrix, objetivou-se, neste trabalho, predizer o secretoma de *H. vastatrix* raça XXXIII e identificar possíveis genes candidatos a efetores, por meio do proteoma obtido do sequenciamento do genoma desse fungo.

2 REFERENCIAL TEÓRICO

2.1 Ferrugem alaranjada do cafeeiro – *Hemileia vastatrix*

Essa doença foi primeiramente constatada em 1861, nos cafeeiros silvestres da região do lago Victoria-Nyanza, no Quênia, na África. Somente oito anos mais tarde ela foi descrita por Berkeley e Broome no Ceilão, atual Sri Lanka (SILVA et al., 2006). No Brasil, a primeira observação da ferrugem alaranjada do café foi feita em 1970, no estado da Bahia (CHAVES et al., 1970). A partir daí ela disseminou-se para os demais estados brasileiros produtores de café e para outros países da América Latina, também, produtores de café (SCHIEBER; ZENTMYER, 1984).

A ferrugem alaranjada do cafeeiro é causada pelo fungo *H. vastatrix* Berk. & Br., que pertence ao filo *Basidiomycota*, à classe *Pucciniomycetes*, à ordem *Pucciniales* e família *Incertae sedis* (www.indexfungorum.org). O gênero *Hemileia* foi classificado desta forma devido à morfologia dos uredinósporos, que apresentam uma face lisa (côncava) e outra equinulada (convexa) coberta de espínulas (RODRIGUES, 1990).

Esse fungo é um parasita biotrófico, ou seja, depende das células vivas do seu hospedeiro para se desenvolver e se reproduzir e é exclusivo do gênero *Coffea*. A doença causada por ele resulta em queda precoce das folhas devido à alta produção de etileno (CARVALHO; SOUZA, 1998). Como consequência, ocorre redução da área fotossinteticamente ativa, ocasionando a morte dos ramos laterais, interferindo no desenvolvimento dos botões florais, no pegamento dos frutos e na produção do ano seguinte (VÁRZEA et al., 2008).

A intensidade da doença está associada ao ambiente, ao patógeno, ao hospedeiro e ao manejo da cultura (ZAMBOLIM et al., 1997). Os prejuízos

econômicos causados são variáveis, podendo causar redução de mais de 50% da produção dos cafezais (ZAMBOLIM; VALE ZAMBOLIM, 2005).

O ciclo de vida dos fungos causadores de ferrugens é muito complexo. No entanto, a ferrugem do cafeeiro evoluiu o ciclo de vida, no sentido de realizar ciclos mais curtos (microcíclico) e no mesmo hospedeiro (autoico) (KOLMER; ORDONEZ; GROTH, 2009). O ciclo biológico de *H. vastatrix* apresenta apenas três tipos de estruturas de frutificação conhecidos (urediniósporos, teliósporos e basidiósporos) e duas fases distintas (dicariótica e monocariótica) (AGRIOS, 2005; COUTINHO; RIJKENBERG; VAN ASCH, 1995; FERNANDES; EVANS; BARRETO, 2009). Na fase assexuada do ciclo, ocorre a formação de urediniósporos dicarióticos que são responsáveis pelo incremento da doença em condições de ambiente favoráveis (AGRIOS, 2005). Já os teliósporos são formados ocasionalmente e aparecem excepcionalmente em condições de chuva moderada e baixas temperaturas, infectando outros cafeeiros, devido à ação do vento e de outros agentes (FERNANDES; EVANS; BARRETO, 2009; KOLMER; ORDONEZ; GROTH, 2009).

Os sintomas da ferrugem alaranjada do café são observados na face inferior das folhas, onde aparecem pequenas manchas amarelo-pálidas, as quais evoluem para lesões maiores de aspecto pulverulento (urediniósporos) e coloração amarelo-alaranjada. Já, na fase superior, observam-se manchas cloróticas amareladas que correspondem aos limites da pústula na face inferior, a qual necrosa e causa a queda da folha infectada (FAZUOLI et al., 2007; GODOY; BERGAMIN FILHO; SALGADO, 1997; ZAMBOLIM et al., 1999).

No estudo realizado por Krohling et al. (2010), concluiu-se que a doença se encontra mais abundante nas folhas presentes na base do cafeeiro do que nas folhas que ficam no ápice, sendo 60% e 40% de folhas infectadas, respectivamente.

O controle da doença é realizado por meio do controle químico pela aplicação de fungicidas com ação protetora e fungicidas sistêmicos (VENTURA et al., 2007), sendo realizadas pulverizações foliares e aplicações via solo. Embora eficiente, o controle químico da ferrugem alaranjada do cafeeiro onera o custo de produção, aumenta os riscos de contaminação do ambiente e dos trabalhadores e é de difícil aplicação em lavouras com áreas em declive. A utilização indiscriminada de fungicidas contribui para o surgimento de novas raças fisiológicas resistentes do patógeno, que diminui a efetividade das aplicações e, conseqüentemente, a sustentabilidade da atividade (AGRIOS, 2005).

A melhor opção de controle dessa doença envolve a escolha de um local favorável ao desenvolvimento de plantas de café e pouco favorável ao desenvolvimento do patógeno e sempre que possível utilizar cultivares resistentes, por ser econômico e não causar impactos ambientais (ZAMBOLIM; VALE; ZAMBOLIM, 2005). Entretanto, o desenvolvimento de cultivares de café com resistência durável representa um desafio, em razão do aparecimento de novas raças fisiológicas do patógeno que levam à suplantação da resistência das cultivares melhoradas, como consequência do alto potencial evolutivo da população de *H. vastatrix* (CABRAL et al., 2009; MAIA et al., 2013; VÁRZEA; MARQUES, 2005).

A resistência do café a esse fungo é condicionada por pelo menos nove genes dominantes (S_H1 a S_H9). Os genes S_H1 , S_H2 , S_H4 e S_H5 foram caracterizados em *C. arabica* (BETTENCOURT; NORONHA-WAGNER, 1971; NORONHA-WAGNER; BETTENCOURT, 1967). O gene S_H3 foi caracterizado em *C. liberica* (BETTENCOURT; CARVALHO, 1968; BETTENCOURT; NORONHA-WAGNER, 1971; BETTENCOURT; RODRIGUES, 1988). Os outros genes S_H6 , S_H7 , S_H8 e S_H9 foram identificados

em *C. canephora* (BETTENCOURT, 1973; BETTENCOURT; LOPES; PALMA, 1992; BETTENCOURT; NORONHA-WAGNER; LOPES, 1980).

Atualmente, a ferrugem alaranjada do café está presente na maioria das regiões cafeeiras do mundo e já foram identificadas 49 raças fisiológicas de *H. vastatrix* (GICHURU et al., 2012) e 16 delas já foram detectadas no Brasil (I, II, III, VII, X, XIII, XV, XVI, XVII, XXI, XXII, XXIII, XXIV, XXV ou XXXI, XXXIII e XXXVII) (CABRAL et al., 2009; CAPUCHO et al., 2012; ZAMBOLIM; VALE; ZAMBOLIM, 2005). A grande variabilidade genética das populações de *H. vastatrix*, comum à maioria dos patógenos biotróficos, tem sido determinada pelo padrão de infecção observado em uma série de 18 variedades diferenciadoras de cafeeiros (ZAMBOLIM; VALE; ZAMBOLIM, 2005). A formação de novas raças fisiológicas de *H. vastatrix* está relacionada com a pressão de seleção exercida pelos genes de resistência do hospedeiro, principalmente, nas áreas onde se faz o uso de cultivares portadores de genes de resistência.

A raça II é encontrada com maior frequência no Brasil (ZAMBOLIM; VALE; ZAMBOLIM, 2005) e apresenta apenas o gene de virulência v5. Já, a raça XXXIII, recentemente identificada, possui dois ou três genes de virulência (v5,7 ou v5,7,9). Essa raça é capaz de suplantar a resistência de algumas cultivares de café que haviam sido liberadas como cultivares resistentes à ferrugem, incluindo a cultivar Oeiras MG 6851, derivada do cruzamento entre *C. arabica* cv. Caturra Vermelho CIFC 19/1 e Híbrido de Timor CIFC 832/1, uma das principais fontes de resistência nos programas de melhoramento de cafeeiro (CAPUCHO et al., 2012). Por isso, é de suma importância conhecer e entender os mecanismos moleculares envolvidos na interação planta-patógeno, pois eles ajudarão a compreender a suplantação da resistência da planta por novas raças fisiológicas do fungo.

2.2 Interações cafeeiro x *Hemileia vastatrix*

As interações entre a planta e o patógeno envolvem uma comunicação de duas vias, ou seja, não só a planta deve ser capaz de reconhecê-lo e se defender do ataque, mas o patógeno deve também ser capaz de manipular o sistema de defesa de seu hospedeiro para propiciar um ambiente adequado para seu crescimento e reprodução. Logo, milhões de anos de convívio, em um mesmo ambiente, favoreceu para que ambos evoluíssem um conjunto de genes que permitem essa comunicação e coevolução (BOYD et al., 2013). As plantas, por estarem constantemente expostas a mudanças climáticas, ataques de pragas e doenças, desenvolveram um sistema de defesa simples, porém eficiente, o qual é ativado sempre após o reconhecimento de sinais da presença de microrganismos. Por sua vez, os patógenos apresentam a habilidade de reconhecer sinais das plantas hospedeiras e iniciar a infecção (DALIO et al., 2014).

Fitopatógenos atacam plantas, porque necessitam dos nutrientes produzidos por elas, para desempenhar suas atividades vegetativas e reprodutivas. Como muitos desses nutrientes estão contidos no interior do protoplasma das células vegetais, os patógenos, para acessá-los, precisam vencer as barreiras estruturais externas das plantas, formadas pela cutícula e parede celular, bem como promover a colonização interna dos tecidos (PASCHOLATI, 2011). A maioria dos patógenos tem acesso ao interior dos hospedeiros, via penetração direta da superfície foliar, por meio das aberturas naturais ou por ferimentos. Os fungos, além de penetrarem de maneira direta, seja pela superfície foliar, pelas aberturas naturais e ou pelos ferimentos, secretam enzimas hidrolíticas, toxinas e hormônios que vão auxiliar nesse processo (PASCHOLATI et al., 1998). Depois da penetração, os patógenos bem

sucedidos colonizam o tecido do hospedeiro por meio da degradação celular e pela utilização de nutrientes. A colonização resulta em alterações morfológicas e metabólicas da planta, o que leva ao aparecimento dos sintomas da doença (PASCHOLATI, 2011).

Para que o processo infeccioso de *H. vastatrix* inicie, a germinação dos urediniósporos deve ocorrer, na face abaxial das folhas de café, um processo que demanda condições favoráveis de baixa luminosidade, água e temperatura (18° e 25°C) (SILVA et al., 2002). Em seguida, para que ocorra a penetração e colonização na planta, *H. vastatrix* diferencia várias estruturas de infecção especializadas, como tubo germinativo, apressório, hifa de penetração e haustórios. Inicialmente, os urediniósporos, ao germinar, formam o tubo germinativo, o qual se desenvolve sobre a folha até encontrar um estômato, onde se diferencia em apressório. O apressório, por sua vez, diferencia-se em hifa de penetração, a qual penetra pelos estômatos e alcança a câmara subestomática. Em seguida, ramifica-se em duas hifas laterais, formando uma estrutura em forma de âncora em cada um dos ramos laterais, chamada de célula mãe do haustório. Essas estruturas invadem as células subsidiárias do estômato, formando os haustórios primários, que se diferenciam e formam os haustórios secundários nas células do mesófilo. Assim, as hifas intercelulares crescem e vão formando novos haustórios, o que resulta na formação de um micélio denso que toma todo o mesófilo foliar (RAMIRO et al., 2009). Os haustórios são estruturas determinantes para a fase biotrófica da infecção, pois são responsáveis pela absorção de nutrientes das células do hospedeiro pelo patógeno e pela produção de proteínas efetoras que são secretadas na matriz extra-haustorial e translocadas para o interior da célula vegetal, induzindo mudanças estruturais nas células da planta (GUZZO, 2004).

As plantas, para escapar da infecção de potenciais patógenos, desenvolveram a capacidade de reconhecer moléculas derivadas desses

patógenos, como, por exemplo, a flagelina, a quitina, a glucana, entre outras. Esse processo de reconhecimento é realizado pelo seu sistema de defesa que é dividido em duas frentes principais de defesa. A primeira fornece a defesa basal contra vários patógenos porque reconhece as moléculas conservadas (PAMPs – *pathogen associated molecular patterns*) que são essenciais para o desenvolvimento do patógeno e é denominada PTI (*PAMP-triggered immunity*). A segunda é baseada no reconhecimento de proteínas efetoras produzidas pelos patógenos, denominada ETI (*effector triggered immunity*) (DODDS; RATHJEN, 2010).

As respostas de defesa da planta são ativadas, primeiramente, pelo reconhecimento dos padrões moleculares do patógeno (PAMPs) com proteínas transmembranas receptoras de sinais (PRRs – *pattern recognition receptors*), ricas em repetições de leucina (LRRs) ou motivo de lisina (Lysm) no seu domínio extracelular (BECK et al., 2012; BOYD et al., 2013). Uma vez desencadeada PTI, por ligação física entre o domínio extracelular do receptor (PRR) com o epítipo do ligante (PAMP), uma cascata de sinalização citoplasmática é ativada pelo domínio intracelular do receptor, baseada em mudanças nos níveis de cálcio, rápida produção de espécies reativas de oxigênio e cascata de sinalização da via de MAP-quinases (MONAGHAN; ZIPFEL, 2012).

Embora eficiente para prevenir a infecção, alguns patógenos adaptados adquiriram habilidade de suprimir PTI pela produção e secreção de proteínas efetoras, as quais contribuem para a virulência do patógeno, resultando na suscetibilidade desencadeada por efetores (WIN et al., 2012). Logo, uma segunda frente de defesa (ETI) foi desenvolvida pela planta, a qual utiliza receptores intracelulares (proteínas R) com domínios NB (*nucleotide-binding*) e LRR (*leucine-rich repeat*), para reconhecer efetores (proteínas Avr) do patógeno (BOYD et al., 2013). Esse reconhecimento dos efetores pelas proteínas NB-LRR

ocorre diretamente pela associação física ou indiretamente por meio de uma proteína acessória. O reconhecimento direto, também, denominado modelo ligante-receptor, baseia-se na Teoria Gene-a-Gene de Flor. Nessa teoria, quando os genes R e Avr estão presentes na planta e no patógeno, respectivamente e ocorre o reconhecimento, origina-se uma interação incompatível não havendo doença. No entanto, quando esse reconhecimento não ocorre, o patógeno é virulento e a interação é compatível dando origem à doença (FLOR, 1971). O reconhecimento indireto é conhecido como modelo guarda, em que as proteínas R reconhecem efetores ligados a proteínas da planta chamadas de ‘alvo de efetores de patógenos’ (STERGIOPOULOS; DE WIT, 2009). Dessa forma, quando ETI é ativada, a planta desencadeia rapidamente as vias de sinalização celular que acarretam em morte celular programada no sítio de infecção conhecida como reação de hipersensibilidade (HR – *hypersensitive response*). A reação conduz ao colapso do tecido e à inibição completa da proliferação do patógeno (WIN et al., 2012).

Assim como ocorre em PTI, outras respostas de defesa são ativadas em ETI associadas à HR, como, por exemplo, alterações nos níveis de cálcio no citoplasma, produção de espécies reativas de oxigênio e a cascata de sinalização das vias quinases. A diferença é que PTI permite uma resposta de defesa muito rápida ao ataque de uma ampla gama de patógenos, enquanto a ETI apresenta uma maior amplitude de resposta com uma sinalização mais forte (TSUDA; KATAGIRI, 2010). Contudo, esse tipo de resistência (ETI) pode ser superada rapidamente pelo patógeno pelas mutações e aquisição de novos efetores, seja pela perda ou pela diversificação do seu sítio de reconhecimento pelo gene R. Na coevolução, as plantas desenvolvem novas proteínas R (NB-LRR), para reconhecer esses novos efetores, sugerindo uma constante “corrida armamentista” na interação planta-patógeno (JONES; DANGL, 2006).

Em genótipo de cafeeiros suscetíveis à *H. vastatrix*, o processo de infecção culmina na esporulação sobre a superfície externa da folha, via estômatos em torno de 20 dias, após a infecção (SILVA et al., 1999). Já, em genótipos de cafeeiros resistentes à doença, o crescimento do fungo é impedido nas primeiras fases do processo infeccioso devido à reação de HR (SILVA et al., 2002, 2006).

Embora estes eventos sejam bem descritos no patossistema cafeeiro-*H.vastatrix*, diferenças temporais no desenvolvimento do fungo e nas respostas de defesa da planta podem ocorrer, dependendo do genótipo de cafeeiro e da raça do patógeno envolvidos. Observações citológicas, realizadas no estudo de Lopes (2015), mostraram que a resistência do cafeeiro Híbrido de Timor (CIFC 832/1) à *H.vastatrix* raça XXXIII é uma resistência pré-haustorial, ao contrário da resistência pós-haustorial que é geralmente descrita para interações cafeeiro - *H.vastatrix* (SILVA et al., 2002).

Por isso, a compreensão dos mecanismos chaves envolvidos tanto na defesa das plantas quanto no ataque dos patógenos pode ajudar no desenvolvimento de estratégias novas no controle de doenças na agricultura.

2.3 Proteínas efetoras de fungo, identificação e caracterização funcional

O termo “efetor” surgiu há 15 anos e tornou-se amplamente utilizado no campo das interações planta-patógeno, porque era necessário um termo neutro, que abordasse as limitações conceituais de avirulência e elicitor (HOGENHOUT et al., 2009).

Segundo Hogenhout et al. (2009), efetores são moléculas secretadas por organismos associados às plantas que modificam a estrutura e função da célula hospedeira. Porém, nem todos os efetores são secretados e nem sempre o alvo é um hospedeiro. Dessa forma, segundo Dalio et al. (2014), a definição mais

correta e atual para efetores é: efetores são moléculas liberadas ou associadas a um organismo que modificam a fisiologia de outro organismo.

De acordo com o local de atuação, os efetores podem ser divididos em dois grupos: efetores extracelulares e efetores citoplasmáticos. Os efetores extracelulares, também, conhecidos como efetores apoplásticos são secretados pelos patógenos, durante a infecção e acumulam-se na superfície exposta da célula ou nos espaços intercelulares. Os efetores citoplasmáticos são os que podem se translocar através da membrana celular e têm contato direto com compartimentos subcelulares ou organelas (DE WIT et al., 2009).

As ferrugens, assim como os outros patógenos biotróficos, produzem haustórios no mesófilo das células de seus hospedeiros durante o processo de infecção. Essas estruturas especializadas não servem apenas para a absorção de nutrientes com as células de seus hospedeiros, mas têm papel importante na secreção de proteínas efectoras para dentro da célula. Os haustórios facilitam a entrega dos efetores na matriz extra-haustorial de onde muitos serão translocados para outros compartimentos subcelulares (WIN et al., 2012). No entanto, pouco se sabe sobre como ocorre esse processo de translocação, sendo necessários métodos que permitam a detecção e visualização desse tráfego dos efetores até o local de atuação (PETRE; KAMOUN, 2014).

Como mencionado anteriormente, uma vez translocados para o citoplasma da planta, os efetores podem trafegar para diferentes compartimentos subcelulares. Um número grande de efetores se encontra no núcleo das plantas, como exemplo, TAL de *Xanthomonas*, SAP11 de fitoplasma e Crinklers (CRNs) de *Phytophthora* (BAI et al., 2009; SCHORNACK et al., 2010; VAN DEN ACKERVEKEN, 1996). Entretanto, também, encontram-se efetores em mitocôndrias, cloroplastos, citoesqueleto e complexo de golgi, como, por exemplo, HopG1, HopI1, HopA1a e HopM1, respectivamente, todos de *Pseudomonas syringae* (BLOCK et al., 2010; JELENSKA et al., 2007; LEE et

al., 2012). E em cada um desses compartimentos, esses efetores desempenham funções específicas. O efector TAL (*Xanthomonas*) ativa genes do hospedeiro envolvidos em diferentes processos, como, por exemplo, o tamanho das células, transporte de açúcares e processos epigenéticos (CHEN et al., 2010; SOUZA et al., 2012). O SAP (fitoplasma) age diretamente no desenvolvimento da planta, aumentando a massa de tecido verde, o que, por sua vez, aumenta a transmissão do vetor pelo inseto (MACLEAN et al., 2011; SUGIO et al., 2011). Em *Phytophthora*, os efetores estão associados à inibição enzimática, como, por exemplo, hidrolases (DAMASCENO et al., 2008; TIAN; BENEDETTI; KAMOUN, 2005; TIAN et al., 2004, 2007). E os efetores de *P. syringae* (HopG1, HopI1, HopA1a e HopM1) são responsáveis pela supressão da biossíntese de ácido salicílico, rompimento da estrutura de tilacoides e microtúbulos (JELENSKA et al., 2007; LEE et al., 2012).

Diante do exposto, confirma-se que há muitas outras funções distintas atribuídas aos efetores e não somente à função de imunossupressores, como uma tendência mostrada na literatura. Algumas atividades comuns dos efetores na inibição de defesa ou na inibição do reconhecimento da infecção pela planta incluem: inibição de enzimas (hidrolases) no apoplasto, inibição de ativação de proteínas PRRs, inibição de cascatas de sinalização mediadas por MAP-quinases, modificação do transcriptoma relacionado a respostas de defesa, reprogramação transcricional, degradação de componentes e proteínas relacionadas à defesa e interferência na secreção de proteínas e tráfego vesicular (GÖHRE; ROBATZEK, 2008):

Uma característica comum dos efetores é que são pequenas proteínas, geralmente secretadas, ricas em resíduos de cisteína e possuem motivos específicos conservados na região N-terminal. Na região C-terminal, encontra-se o domínio do efector que está relacionado com a sua função. Esses resíduos de cisteína podem estar envolvidos na formação de pontes dissulfeto, as quais são

importantes para a estabilidade e atividade da proteína (STERGIOPOULOS; DE WIT, 2009). Os efetores apresentam baixo grau de conservação, ou seja, não apresentam similaridade com proteínas de função conhecidas, o que dificulta inferir o possível papel biológico e, por isso, estudos, determinando suas atividades bioquímicas, ainda, são escassos (DUPLESSIS et al., 2011). Dado o grande número de efetores que ainda permanecem descaracterizados, nosso conhecimento sobre a sua identidade permanece incompleto.

A elevada taxa de evolução, observada nos efetores, é responsável pela grande diversificação da sequência dos seus genes, o que explica a baixa presença de homologia entre as proteínas efetoras já descritas e as suas diferentes especificidades nos hospedeiros. Isso resulta em mutações e seleções, durante o processo de coevolução planta-patógeno. Muitos genes que codificam proteínas efetoras estão inseridos em regiões altamente dinâmicas do genoma dos patógenos, como, por exemplo, extremidades dos cromossomos ou em meio aos elementos transponíveis, permitindo maior flexibilidade para suplantar a resistência mediada pelas proteínas R (KHANG et al., 2008; ORBACH et al., 2000; RIDOUT et al., 2006).

Inicialmente, achava-se que os efetores eram específicos, ou seja, modificavam uma proteína do hospedeiro individualmente. Hoje estudos mostram que os efetores podem ser promíscuos, isto é, um único efector pode associar-se com múltiplas proteínas da planta e afetar processos distintos no hospedeiro (WIN et al., 2012). Isso porque os genes R estão presentes em clusters multigênicos e podem ocorrer naturalmente como verdadeiros alelos através de variantes da genética original (DANGL; HORVATH; STASKAWICZ, 2013). A ativação da função de cada uma das proteínas R é feita pelo produto dos genes efetores de virulência de um patógeno específico (FLOR, 1971). Vários efetores podem ser expressos por um único patógeno e a

diversidade de efetores, em uma população de qualquer espécie patogênica, pode ser muito grande (BALTRUS et al., 2011).

As análises genéticas e bioquímicas são as duas estratégias mais populares utilizadas na identificação de efetores. No entanto, sua taxa de sucesso é relativamente baixa (WIN et al., 2012). Recentemente, com o avanço das técnicas moleculares e sequenciamento de nova geração, muita informação tem sido disponibilizada. As bases genômicas têm sido utilizadas, para identificar efetores, por meio do sequenciamento total do genoma de fungos fitopatogênicos. E defini-los funcionalmente como novos alelos R. Tem sido possível, também, predizer genes candidatos a efetores em muitos patógenos, assim como suas atividades, as quais são usadas como estratégias promissoras para a pesquisa e a obtenção de plantas resistentes a doenças (DALIO et al., 2014).

Uma estratégia utilizada para a obtenção de plantas com resistência durável é o uso de efetores alvos por meio do isolamento de genes R (PERIYANNAN et al., 2013; SAINTENAC et al., 2013). Com as tecnologias rápidas e de baixo custo de sequenciamento de DNA, um novo patógeno pode ser isolado e seu genoma gerar informações com impactos sobre as estratégias de melhoramento para um controle mais durável de doenças em plantas (VLEESHOUWERS et al., 2011).

O emprego de receptores de sistema imune, também, é uma estratégia usada na agricultura, por meio da identificação funcional das PRRs (DANGL; HORVATH; STASKAWICZ, 2013). Assim como a estratégia de empilhamento, a qual explora respostas imunes em contextos em que vários genes NLR são implantados simultaneamente. As cultivares geradas por qualquer tipo de melhoramento molecular de DNA assistido ou por transferência gênica devem exibir resistência mais durável a doenças, visto que a evasão do patógeno exigirá mutações em múltiplos genes efetores. O acesso a uma enorme diversidade

genética é permitido pelos avanços em sequenciamento de DNA das principais culturas e seus parentais para uma grande quantidade de genes NLR funcionalmente dirigidos contra diferentes efetores fundamentais (DANGL; HORVATH; STASKAWICZ, 2013).

Entretanto, para que essas abordagens obtivessem sucesso, junto com o surgimento das tecnologias de sequenciamento de nova geração, ferramentas de bioinformática foram desenvolvidas. Por meio dessas ferramentas, é possível analisar o secretoma dos patógenos e prever possíveis proteínas secretadas com dados provenientes do genoma. As ferramentas desenvolvidas, para a predição de proteínas efectoras, são baseadas nas características do peptídeo sinal das proteínas. Esses peptídeos sinais correspondem a uma pequena sequência de aminoácidos (entre 15 e 60), localizada na porção N-terminal da proteína, com a função de marcar as proteínas que serão exportadas pelo aparato de secreção (TORTO et al., 2003). Os peptídeos sinais podem ser reconhecidos por ferramentas computacionais, como, por exemplo, os algoritmos SignalP e TargetP. Posteriormente, essas proteínas são analisadas, quanto à presença de domínios transmembrânicos, utilizando o algoritmo computacional TMHMM e, finalmente, tem-se o conjunto de proteínas secretadas do genoma de um organismo (PETERSEN et al., 2011)

2.4 Sequenciamento de fungos fitopatogênicos

O sequenciamento do genoma de um organismo é um processo que determina a ordem dos nucleotídeos presentes em uma molécula de DNA desse organismo, baseando-se em métodos bioquímicos (BRANDÃO, 2009).

A molécula de ácido desoxirribonucleico (DNA) contém as informações necessárias para o desenvolvimento e funcionamento de todas as células em um organismo vivo. Para entender os mecanismos que atuam nos processos

biológicos complexos de um indivíduo e para ampliar o conhecimento das diferenças genéticas entre os diferentes organismos, é necessário explorar a composição da molécula de DNA, por meio do processo de sequenciamento do genoma (GRABHERR; MAUCELI, 2011).

O estudo da sequência de todo o genoma é essencial para os ramos da pesquisa biológica. O uso do sequenciamento, na obtenção do genoma completo de microrganismos, ocorreu há 20 anos, utilizando-se a metodologia de Sanger. Alguns anos depois, com o advento das tecnologias de sequenciamento de nova geração, houve um crescimento exponencial de novas sequências de genomas de todas as classes de seres vivos. Isso gerou uma grande quantidade de dados e possibilitou a busca de novos genes candidatos a efetores, em diversas espécies, nas quais destacam-se os fungos fitopatogênicos, agentes causadores das principais doenças agrícolas (SCHMIDT; PANSTRUGA, 2011). Essas tecnologias de sequenciamento de nova geração têm causado um impacto grande no campo das pesquisas genômicas, pelo fato de oferecerem um serviço de relativo baixo custo, com alto rendimento. Além disso, permitem uma análise mais abrangente da estrutura e conteúdo dos genomas de microrganismos do que era possível anteriormente (FORDE; O'TOOLE, 2013).

As plataformas de sequenciamento de nova geração mais usadas atualmente para estes estudos são: Illumina (HiSeq 2500) e Pacific Biosciences (PacBio RS II). Cada uma destas tecnologias proporciona novas abordagens para o sequenciamento de DNA e tem vantagens e desvantagens com relação a aplicações específicas. A tecnologia Illumina tem se destacado por oferecer protocolos de preparação de bibliotecas compatíveis com todos os sistemas, mais alta taxa de transferência de dados entre todas as plataformas e menor custo por base.

A abordagem de sequenciamento da plataforma Illumina baseia-se na química de sequenciamento-por-síntese (SBS), que utiliza nucleotídeos

terminadores reversíveis marcados por fluorescência. O resultado final é o sequenciamento base a base, permitindo dados mais precisos, sendo as leituras realizadas pela intensidade do sinal medido, durante cada ciclo, o que também reduz a taxa de erro (DIAZ-SANCHEZ et al., 2013).

Essa tecnologia foi introduzida comercialmente em 2007 e foi rapidamente adquirida por pesquisadores pelo fato de gerar uma maior quantidade de dados (30 milhões de *reads*) de maneira mais eficaz em termos de custos. Entretanto, o tamanho dos *reads* gerados era muito pequeno (35bp), o que dificultava os estudos de montagem de genoma (CAPORASO et al., 2012; LANGE et al., 2014). Com o passar dos anos, por meio da modernização das máquinas e da melhoria dos químicos e dos *softwares*, a tecnologia Illumina passou a gerar *reads* de várias centenas de pares de bases (300bp) e, assim, passou a ser utilizada, também, nos estudos de montagem *De novo* de genomas e metagenômica (UTTURKAR et al., 2014). Atualmente, esses *long reads*, produzidos pela plataforma Illumina, podem ser encontrados em MiSeq e HiSeq que produzem *reads* em *paired-end* com tamanho de 300bp e 150bp, respectivamente. No entanto, a plataforma HiSeq 2500 é a que apresenta maior rendimento, gerando 4 bilhões de *reads* em uma única corrida (HODKINSON; GRICE, 2015).

Outra plataforma muito usada, PacBio, gera *reads* longos (20kb), fazendo dessa tecnologia uma ferramenta ideal para terminar ou melhorar montagem de projetos de genomas em um tempo de execução de algumas horas. Por outro lado, seu custo é alto (US\$ 17.02 por Mb), apresenta as mais altas taxas de erro (14%) e a menores taxa de transferência de dados em comparação com as outras plataformas de sequenciamento (HODKINSON; GRICE, 2015).

A empresa Pacific Biosciences lançou, em 2010, a plataforma PacBio RS, que utiliza uma tecnologia de sequenciamento chamada *SMRT Analysis*, a qual baseia-se na detecção natural da síntese de DNA por meio de uma única

DNA polimerase. O processo de síntese começa com a introdução de nucleotídeos marcados com diferentes fluoróforos coloridos. A incorporação de fosfato marcado é base-específico. Quando uma base é detectada, a fluorescência é emitida como um pulso de luz e detectada em tempo real. A reação de sequenciamento só termina quando ocorre a dissociação entre a DNA polimerase e a fita molde (ROBERTS; CARNEIRO; SCHATZ, 2013).

Esse equipamento gera vários milhões de kilobases e *long reads* (20kb) por corrida, sendo uma plataforma ideal para a montagem de genomas. Para melhorar a qualidade da leitura das sequências e diminuir a taxa de erro, a Pacific Biosciences desenvolveu uma estratégia de sequenciamento consenso circular (CCS), que se baseia na estrutura circular das moléculas das bibliotecas da PacBio. A estrutura circular permite um sequenciamento contínuo e repetido de cadeia no sentido *sense* e *antisense*, que pode ser usado para gerar um único *read* consenso com precisão muito elevada (> 99%). Essa alta precisão é útil em metagenômica e, especialmente, em projetos de sequenciamento *De novo*, uma vez que impede a superestimativa da diversidade biológica em razão de erros de sequenciamento. PacBio RS II, atualmente, utiliza a química C3/P5 para realizar o sequenciamento. Isso proporciona os maiores comprimentos de leitura (em torno de 8.500 bases) e o maior número de leituras por corrida (cerca de 16kb) obtidas com boa qualidade (KNIEF, 2014).

Utilizando diferentes plataformas, foram desenvolvidos, segundo o *US Department of Energy-JGI*, 7.632 projetos de sequenciamento, gerando muitos dados de sequências de nucleotídeo de diversos organismos. JGI apresentou 65.717 organismos sequenciados, sendo 11.072 referentes a eucariotos e apenas 1.597 são basidiomicetos (REDDY et al., 2015).

O primeiro fungo (e eucarioto) a ter seu genoma completo sequenciado foi a levedura *Saccharomyces cerevisiae* em 1996 (GOFFEAU et al., 1996), seguida pelo sequenciamento do fungo modelo *Neurospora crassa* em 2003

(GALAGAN et al., 2003). Até o presente momento, foram sequenciados cerca de 74 genomas de fungos e oomicetos, dos quais destacam-se os patógenos de plantas economicamente importantes: *Alternaria arborescens*, *Aspergillus flavus*, *Blumeria graminis*, *Botrytis cinérea*, *Colletotrichum* spp., *Fusarium graminearum*, *Fusarium oxysporum*, *Magnaporthe oryzae*, *Melampsora lini*, *Melampsora larcini-populina*, *Mycosphaerella graminicola*, *Phytophthora infestans*, *Puccinia graminis*, *Puccinia triticina*, *Ustilago maydis* (DEAN et al., 2012; KERSEY et al., 2014; SHARMA, 2015).

Poucas espécies de ferrugem foram sequenciadas. Esses fungos são parasitas obrigatórios pertencentes ao filo basidiomicota e à ordem puccinales, compreendendo um grupo muito diverso composto por 120 gêneros e 6000 espécies (CANTU et al., 2013). A conclusão do sequenciamento do genoma de algumas ferrugens revelou a ocorrência dos maiores genomas de fungos já relatados (*Gymnosporangium confusum* -893,2Mbp e *Puccinia chrysanthemi* - 806.5Mbp) e com uma variação de até 10 vezes: O tamanho médio do genoma de basidiomicetos é de 70,4Mbp e da ordem Puccinales é de 305,5Mbp (TAVARES et al., 2014).

H. vastatrix, também, encontra-se entre os grandes genomas de ferrugens. Carvalho et al. (2013), por meio de citometria de fluxo, mediram o tamanho do genoma nuclear desse fungo, revelando um tamanho de 733,5 Mb.

Pelo descrito acima, podemos observar que os genomas de fungos variam consideravelmente em tamanho e, também, em organização. As diferenças na arquitetura do genoma entre fungos refletem-se nos diferentes processos de interação planta-patógeno e nos diferentes processos de mutação sofridos por eles. Estudos de genômica comparativa elucidaram os diferentes mecanismos de evolução que ocorrem nos patossistemas e proporcionam uma visão em traços genômicos específicos de cada espécie (STUKENBROCK; CROLL, 2014). Muitos estudos de interação planta-patógeno, obtidos por

sequenciamento do genoma e do transcriptoma, visam encontrar e caracterizar proteínas efetoras e as respostas sobre a adaptação do hospedeiro com essas proteínas, elucidando, assim, sua coevolução e o desenvolvimento de novas estratégias para o controle de doenças.

Embora o transcriptoma da interação cafeeiro - *Hemileia* tenha sido parcialmente caracterizado (FERNANDEZ et al., 2012; TALHINHAS et al., 2014), ainda, são inexistentes estudos visando identificar e caracterizar proteínas efetoras de *H. vastatrix*. baseando-se no secretoma total. presente no genoma desse fungo.

Sendo assim, diante do exposto, objetivou-se. neste trabalho, sequenciar o genoma do fungo *H. vastatrix* e realizar a anotação funcional desse genoma, buscando por proteínas efetoras totais exclusivas da espécie e, posteriormente, analisar o seu perfil de expressão durante a infecção por meio de PCR quantitativo.

REFERÊNCIAS

ABRAHÃO, A. A. et al. Influência de safras agrícolas e tratamentos fungicidas no café cereja descascado e bóia. **Ciência e Agrotecnologia**, Lavras, v. 33, p.1919-1925, 2009. Número especial.

AGRIOS, G. N. Plant diseases caused by fungi. In: _____. **Plant pathology**. 5th ed. Amsterdam: Elsevier, 2005. chap. 11, p. 562-580.

AMARAL, D. R. et al. Silicato de potássio na proteção do cafeeiro contra *Cercospora coffeicola*. **Tropical Plant Pathology**, Brasília, v. 33, n. 6, p. 425-431, dez. 2008.

BAI, X. et al. AY-WB phytoplasma secretes a protein that targets plant cell nuclei. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 22, n. 1, p. 18-30, Jan. 2009.

BALTRUS, D. A. et al. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. **PLoS Pathogens**, San Francisco, v. 7, n. 7, p. 1002132, July 2011.

BECK, M. et al. The INs and OUTs of pattern recognition receptors at the cell surface. **Current Opinion in Plant Biology**, London, v.15, n. 4, p. 367-374, Aug. 2012.

BETTENCOURT, A. J.; CARVALHO, A. Melhoramento visando à resistência do cafeeiro à ferrugem. **Bragantia**, Campinas, v. 27, n. 4, p. 35-68, fev.1968.

BETTENCOURT, A. J. **Considerações gerais sobre o ‘Híbrido de Timor’**. Campinas: Instituto Agrônômico, 1973. 20 p. (Circular técnica, 31).

BETTENCOURT, A. J.; LOPES, J.; PALMA, S. Fatores genéticos que condicionam a resistência às raças de *Hemileia vastatrix* Berk. et Br. dos clones-tipo dos grupos 1, 2 e 3 de derivados de Híbrido de Timor. **Brotéria Genética**, Lisboa, v.13, n. 80 p.185-194, 1992.

BETTENCOURT, A. J.; NORONHA-WAGNER, M. Genetic factors conditioning resistance of *Coffea arabica* L. to *Hemileia vastatrix* Berk. & Br. **Agronomia Lusitana**, Oeiras, v. 31, n. 4, p. 285-292, Apr. 1971.

BETTENCOURT, A. J.; NORONHA-WAGNER, M.; LOPES, M. Factor genético que condiciona a resistência do clone 1343/269 (Híbrido de Timor) à *Hemileia vastatrix* Berk. and Br. **Brotéria Genética**, Lisboa, v. 1, n. 76, p. 183-194, 1980.

BETTENCOURT, A. J.; RODRIGUES, J. R. C. J. Principles and practice of coffee breeding for resistance to rust and other diseases. In: CLARKE, R. J.; MACRAE, R. (Ed.). **Coffee**. London: Elsevier Applied Science, 1988. v. 4, p. 199-234.

BLOCK, A. et al. The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. **Cellular Microbiology**, Oxford, v. 12, n. 3, p. 318-330, Mar. 2010.

BOYD, L. A. et al. Plant–pathogen interactions: disease resistance in modern agriculture. **Trends in Genetics**, Cambridge, v. 29, n. 4, p. 233-240, Apr. 2013.

BRANDÃO, K. L. S. **Sequenciamento e alinhamento de DNA**. Piracicaba: [s.n.], 2009. 32 p.

BRASIL. Ministério da Agricultura pecuária e abastecimento. Companhia Nacional de Abastecimento. **Acompanhamento da safra brasileira de café, primeiro levantamento**. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/15_01_14_11_57_33bol_etim_cafe_janeiro_2015.pdf>. Acesso em: 10 jan. 2015.

CABRAL, P. G. C. et al. Identification of a new race of *Hemileia vastatrix* in Brazil. **Australasian Plant Disease Note**, Melbourne, v. 4, n. 1, p. 129-130, Dec. 2009.

CANTU, D. et al. Genome analyses of the wheat yellow (stripe) rust pathogen *Puccinia striiformis* f. sp. *tritici* reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. **BMC Genomics**, London, v. 14, p. 270-271, Apr. 2013.

CAPORASO, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. **The ISME Journal**, London, v. 6, n. 8, p. 1621-1624, Aug. 2012.

CAPUCHO, A. S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Note**, Melbourne, v. 7, p. 189-191, Oct. 2012.

- CARVALHO, G. M. A. et al. Coffee rust genome measured using flow cytometry: does size matter? **Plant Pathology**, London, v. 63, n. 5, p. 1022-1026, Oct. 2013.
- CARVALHO, V. L. de; SOUZA, S. M. C. de. **Ferrugem**: doença mais importante do cafeeiro. Lavras: EPAMIG, 1998. 20 p. (Circular técnica, 87).
- CHAVES, G. M. et al. A ferrugem do cafeeiro (*Hemileia vastatrix* Berk. & Br.). **Seiva**, Viçosa, MG, v. 30, p. 1-76, 1970. Edição especial.
- CHEN, L. Q. et al. Sugar transporters for intercellular exchange and nutrition of pathogens. **Nature**, London, v. 468, n. 7323, p. 527-532, Apr. 2010.
- COUTINHO, T. A.; RIJKENBERG, F. H. J.; VAN ASCH, M. A. J. Teliospores of *Hemileia vastatrix*. **Mycological Research**, Cambridge, v. 99, n. 8, p. 932-934, Aug. 1995.
- CRISTANCHO, M. A. et al. Annotation of a hybrid partial genome of the coffee rust (*Hemileia vastatrix*) contributes to the gene repertoire catalog of the Pucciniales. **Frontiers in Plant Science**, Lausanne, v. 5, p. 594-595, Oct. 2014.
- DALIO, R. J. D. et al. Efetores na interação planta-patógeno. **Revisão Anual de Patologia de Plantas**, Passo Fundo, v. 22, p. 25-68, 2014.
- DAMASCENO, C. M. B. et al. Structure of the glucanase inhibitor protein (GIP) family from phytophthora species suggests coevolution with plant endo-beta-1,3-glucanases. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 21, n. 6, p. 820-830, June 2008.
- DANGL, J. L.; HORVATH, D. M.; STASKAWICZ, B. J. Pivoting the plant immune system from dissection to deployment. **Science**, New York, v. 341, n. 6147, p. 746-751, Aug. 2013.
- DEAN, R. et al. The top 10 fungal pathogens in molecular plant pathology. **Molecular Plant Pathology**, Oxford, v. 13, n. 4, p. 414-430, May 2012.
- DE WIT, P. et al. Fungal effector proteins: past, present and future. **Molecular Plant Pathology**, Oxford, v. 10, n. 6, p. 735-747, Nov. 2009.
- DIAZ-SANCHEZ, S. et al. The future of molecular genetics in poultry production and food safety. **Poultry Science**, Champaign, v. 92, n. 2, p. 562-572, Feb. 2013.

DODDS, P. N.; RATHJEN, J. P. Plant immunity: towards an integrated view of plant pathogen interactions. **Nature Reviews Genetics**, London, v. 11, n. 8, p. 539-548, Aug. 2010.

DUPLESSIS, S. et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proceeding of the National Academy of Science**, Washington, v. 108, n. 22, p. 9166-9171, May 2011.

DUPLESSIS, S.; JOLY, D. L.; DODDS, P. N. Rust effectors. In: MARTIN, F.; KAMOUN, S. (Ed.). **Effectors in plant-microbe interactions**. London: Wiley-Blackwell, 2012. chap. 7, p. 155-193.

FAZUOLI, L. C. et al. Café robusta: uma nova opção para a cafeicultura paulista. **O Agrônômico**, Campinas, v. 59, n. 1, p. 71-74, 2007.

FERNANDES, R. C.; EVANS, H. C.; BARRETO, R. W. Confirmation of the occurrence of teliospores of *Hemileia vastatrix* in Brazil with observations on their mode of germination. **Tropical Plant Pathology**, Brasília, v. 34, n. 2, p. 108-113, Apr. 2009.

FERNANDEZ, D. et al. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. **Molecular Plant Pathology**, Oxford, v. 13, n. 1, p. 17-37, Jan. 2012.

FERNANDEZ, D.; TALHINHAS, P.; DUPLESSIS, S. Rust fungi: achievements and future challenges on genomics and host-parasite interactions. KEMPKEN, F. (Ed.). **Agricultural Applications**. 2nd ed. Berlin: Springer-Verlag, 2013. chap. 12, p. 345-341.

FLOR, H. H. Current status of gene-for-gene concept. **Annual Review of Phytopathology**, Palo Alto, v. 9, p. 275-296, Sept. 1971.

FORDE, B. M.; O'TOOLE, P. W. Next-generation sequencing technologies and their impact on microbial genomics. **Briefings in Functional Genomics**, Oxford, v. 12, n. 5, p. 440-453, Jan. 2013.

GALAGAN, J. E. et al. The genome sequence of the filamentous fungus *Neurospora crassa*. **Nature**, London, v. 422, n. 6934, p. 859-868, Apr. 2003.

GICHURU, E. K. et al. Additional physiological races of coffee leaf rust (*Hemileia vastatrix*) identified in Kenya. **Tropical Plant Pathology**, Brasília, v. 37, n. 6, p. 424-427, Nov./Dec. 2012.

GODOY, C. V.; BERGAMIN FILHO, A.; SALGADO, C. L. Doenças do cafeeiro (*Coffea arabica* L.). In: KIMATI, H. **Manual de fitopatologia: doenças de plantas e seu controle**. 3. ed. São Paulo: Agronômica Ceres, 1997. v. 2, p. 184-200.

GOFFEAU, A. et al. Life with 6000 genes. **Science**, Washington, v. 274, n. 5287, p. 546-567, Oct. 1996.

GÖHRE, V.; ROBATZEK, S. Breaking the barriers: microbial effector molecules subvert plant immunity. **Annual Review Phytopathology**, Palo Alto, v. 46, p. 189-215, 2008.

GRABHERR, M. G.; MAUCELI, E.; MA, L. J. Genome sequencing and assembly. **Fungal Genomics**, Totowa, v. 722, p. 1-9, 2011.

GUZZO, S. D. **Aspectos bioquímicos e moleculares da resistência sistêmica adquirida em cafeeiro contra *Hemileia vastatrix***. 2004. 236 p. Tese (Doutorado em Engenharia) - Centro de Engenharia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2004.

HODKINSON, B. P.; GRICE, E. A. Next-generation sequencing: a review of technologies and tools for wound microbiome research. **Advances in Wound Care**, New Rochelle, v. 4, n. 1, p. 50-58, Jan. 2015.

HOGENHOUT, S. A. et al. Emerging concepts in effector biology of plant-associated organisms. **Molecular Plant Microbe Interactions**, Saint Paul, v. 22, n. 2, p. 115-122, Feb. 2009.

JELENSKA, J. et al. A J-domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. **Current Biology**, London, v.17, n. 6, p. 499-508, Mar. 2007.

JONES, J. D. G.; DANGL, J. L. The plant immune system. **Nature**, London, v. 444, n. 7117, p. 323-329, Nov. 2006.

KERSEY, P. J. et al. Ensembl Genomes 2013: scaling up access to genome-wide data. **Nucleic Acids Research**, Oxford, v. 42, n. 2, p. 546-552, Feb. 2014.

KHANG, C. H. et al. Genome organization and evolution of the AVR-Pita avirulence gene family in the *Magnaporthe grisea* species complex. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 1, n. 5, p. 658-670, May 2008.

KNIEF, C. Analysis of plant microbe interactions in the era of next generation sequencing technologies. **Frontiers in Plant Science**, Lausanne, v. 5, p. 216-218, May 2014.

KOLMER, J. A.; ORDONEZ, M. E.; GROTH, J. V. The rust fungi. In: **ENCYCLOPEDIA of Life Sciences**. London: Wiley-Blackwell, 2009. p. 1-8. Disponível em: <http://www.ars.usda.gov/SP2UserFiles/person/3094/rust_fungi.pdf>. Acesso em: 25 out. 2015.

KROHLING, C. A. et al. Uma avaliação do ataque da ferrugem do cafeeiro, *Hemileia vastatrix* Berk. & Br., no sub-bosque da Reserva Biológica de Duas Bocas, Cariacica, Espírito Santo. **Natureza on line**, Santa Teresa, v. 8, n. 2, p. 63-66, 2010. Disponível em: <http://www.naturezaonline.com.br/natureza/conteudo/pdf/03_KrohlingCAetal_6366.pdf>. Acesso em: 24 out. 2015.

LANGE, V. et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. **BMC Genomics**, London, v. 15, n. 1, p. 15-63, Jan. 2014.

LEE, A. H. et al. A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. **PLoS Pathogens**, San Francisco, v. 8, n. 2, p. 1002523, Feb. 2012.

LOPES, R. L. F. **Análise citológica e perfil de expresso gênica de *Hemileia vastatrix* (raça XXXIII) na interação com o cafeeiro**. 2015. 64 p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa, MG, 2015.

MACLEAN, A. M. et al. Phytoplasma effector SAP54 induces indeterminate leaf-like flower development in Arabidopsis plant. **Plant Physiology**, Washington, v. 157, n. 2, p. 831-841, Oct. 2011.

MAIA, T. A. et al. The population structure of *Hemileia vastatrix* in Brazil inferred from AFLP. **Australasian Plant Pathology**, Melbourne, v. 42, n. 5, p. 533-542, Sept. 2013.

MONAGHAN, J.; ZIPFEL, C. Plant pattern recognition receptor complexes at the plasma membrane. **Current Opinion in Plant Biology**, London, v. 15, n. 4, p. 349-357, Aug. 2012.

NORONHA-WAGNER, M.; BETTENCOURT, A. J. Genetic study of resistance of *Coffea* sp. to leaf rust: I, identification and behavior of four factors conditioning disease reaction in *Coffea arabica* to twelve physiologic races of *Hemileia vastatrix*. **Canadian Journal of Botany**, Ottawa, v. 45, n. 11, p. 2021-2031, 1967.

ORBACH, M. J. et al. A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. **The Plant Cell**, Rockville, v. 12, n. 11, p. 2019-2032, Nov. 2000.

PASCHOLATI, S. F. et al. Mecanismos de patogenicidade em fungos. **Revisão Anual de Patologia de Plantas**, Passo Fundo, v. 6, p. 1-47, jan. 1998.

PASCHOLATI, S. F. Fisiologia do parasitismo: como os patógenos atacam as plantas. In: AMORIM, L.; REZENDE, J. A.; BERGAMIN FILHO, A. (Org.). **Manual de Fitopatologia: princípios e conceitos**. 4. ed. São Paulo: Agronômica Ceres, 2011. v. 2, cap. 34, p. 543-591.

PERIYANNAN, S. et al. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. **Science**, New York, v. 341, n. 6147, p. 786-788, Aug. 2013.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, New York, v. 8, n. 10, p. 785-786, Sept. 2011.

PETRE, B.; KAMOON, S. How do filamentous pathogens deliver effector proteins into plant cells. **PLoS Biology**, San Francisco, v. 12, n. 2, p. 1001801, Feb. 2014.

POZZA, E. A. et al. **Semiologia do cafeeiro**: sintomas de desordens nutricionais, fitossanitárias e fisiológicas. Lavras: UFLA, 2010. p. 69-101.

RAMIRO, D. A. et al. Biphasic haustorial differentiation of coffee rust (*Hemileia vastatrix* race II) 26 associated with defence responses in resistant and susceptible coffee cultivars. **Plant Pathology**, London, v. 58, n. 5, p. 944-955, Oct. 2009.

REDDY, T. B. et al. The genomes Online Database (GOLD) v.5: a metadata management system based on a four level (meta) genome project classification. **Nucleic Acids Research**, London, v. 43, Oct. 2014. Disponível em: <<http://nar.oxfordjournals.org/content/early/2014/10/27/nar.gku950.full>>. Acesso em: 25 out. 2015.

RIDOUT, C. J. et al. Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. **The Plant Cell**, Rockville, v. 18, n. 9, p. 2402-2414, Sept. 2006.

ROBERTS, R. J.; CARNEIRO, M. O.; SCHATZ, M. C. The advantages of SMRT sequencing. **Genome Biology**, London, v. 14, n. 7, p. 405, July 2013.

RODRIGUES, C. J. Coffee rusts: history, taxonomy, morphology, distribution and 4 host resistances. **Fitopatologia Brasileira**, Brasília, v. 15, p. 5-9, Mar. 1990.

SAINTENAC, C. et al. Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. **Science**, New York, v. 341, n. 6147, p. 783-786, Aug. 2013.

SCHIEBER, E.; ZENTMYER, G. A. Coffee rust in the Western Hemisphere. **Plant Disease**, Saint Paul, v. 68, n. 2, p. 89-93, Feb. 1984.

SCHMIDT, S. M.; PANSTRUGA, R. Pathogenomics of fungal plant parasites: what have we learnt about pathogenesis? **Current Opinion in Plant Biology**, London, v. 14, n. 4, p. 392-399, Aug. 2011.

SCHORNACK, S. et al. Ancient class of translocated oomycete effectors targets the host nucleus. **Proceeding of the National Academy of Science**, Washington, v. 107, n. 40, p. 17421-17426, Oct. 2010.

SHARMA, K. K. Fungal genome sequencing: basic biology to biotechnology. **Critical reviews in Biotechnology**, Boca Raton, v. 35, n. 1, p. 1-17, Sept. 2015.

SILVA, M. C. et al. Coffee resistance to the main diseases: leaf rust and coffee berry disease. **Brazilian Journal Plant Physiology**, Londrina, v. 18, n. 1, p. 119-147, Jan./Mar. 2006.

SILVA, M. C. et al. Cytochemical Aspects of the Plant–Rust Fungus Interface during the Compatible Interaction *Coffea arabica* (cv. Caturra)–*Hemileia vastatrix* (race III). **International Journal of Plant Sciences**, Chicago, v. 160, n. 1, p. 79-91, Jan. 1999.

SILVA, M. C. et al. Hypersensitive cell death and post-haustorial defense responses arrest the Orange rust (*Hemileia vastatrix*) growth in resistant coffee leaves. **Physiological and Molecular Plant Pathology**, London, v. 60, n. 4, p. 169-183, Apr. 2002.

SOUZA, T. A. et al. The TAL effector PthA4 interacts with nuclear factors involved in RNA-dependent processes including a HMG protein that selectively binds poly (U) RNA. **PLoS One**, San Francisco, v. 7, n. 2, p. 32305, Feb. 2012.

STERGIOPOULOS, I.; DE WIT, P. J. Fungal effector proteins. **Annual Review of Phytopathology**, Palo Alto, v. 47, p. 233-263, 2009.

STUKENBROCK, E. H.; CROLL, D. The evolving fungal genome. **Fungal Biology Reviews**, Amsterdam, v. 28, n. 1, p. 1-12, Feb. 2014.

SUGIO, A. et al. *Phytoplasma* protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. **Proceedings of the National Academy of Sciences**, Washington, v. 108, n. 48, p. 1254-1263, Nov. 2011.

TALHINHAS, P. et al. Overview of the functional virulent genome of the coffee leaf rust pathogen *Hemileia vastatrix* with an emphasis on early stages of infection. **Frontiers in Plant Science**, Lausanne, v. 5, p. 88-91, Mar. 2014.

TAVARES, S. et al. Genome size analyses of Pucciniales reveal the largest fungal genomes. **Frontiers in Plant Science**, Lausanne, v. 5, p. 11-112, Mar. 2014.

TIAN, M.; BENEDETTI, B.; KAMOUN, S. A second Kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato. **Plant Physiology**, Washington, v. 138, n. 3, p. 1785-1793, July 2005.

TIAN, M. et al. A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. **Journal of Biological Chemistry**, Baltimore, v. 279, n. 25, p. 26370-26377, June 2004.

TIAN, M. et al. A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. **Plant Physiology**, Washington, v. 143, n. 1, p. 364-377, Jan. 2007.

TORTO, T. A. et al. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. **Genome Research**, Cold Spring Harbor, v. 13, n. 7, p. 1675-1685, July 2003.

TSUDA, K.; KATAGIRI, F. Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. **Current opinion in Plant Biology**, London, v. 13, n. 4, p. 459-465, Aug. 2010.

UTTURKAR, S. M. et al. Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences. **Bioinformatics**, Oxford, v. 30, n. 19, p. 2709-2716, Oct. 2014.

VAN DEN ACKERVEKEN, G. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. **Cell**, Cambridge, v. 87, n. 7, p. 1307-1316, Dec. 1996.

VÁRZEA, V. M. P. et al. Resistência do cafeeiro a *Hemileia vastatrix*. In: ZAMBOLIM, L. (Ed.). **O estado da arte de tecnologia na produção de café**. Viçosa, MG: UFV, 2008. v. 8, p. 297-320.

VÁRZEA, V. M. P.; MARQUES, D. V. Population variability of *Hemileia vastatrix* vs *coffee* durable resistance. In: ZAMBOLIM, L.; ZAMBOLIM, E. M.; VÁRZEA, V. M. P. (Ed.). **Durable resistance to coffee leaf rust**. Viçosa, MG, 2005. p. 53-74.

VENTURA, J. A. et al. Diagnóstico e manejo das doenças do cafeeiro conilon. In: FERRÃO, R. G. et al. (Ed.). **Café conilon**. Vitória: INCAPER, 2007. p. 481-483.

VLEESHOUWERS, V. G. A. A. et al. Understanding and exploiting late blight resistance in the age of effectors. **Annual Review Phytopathology**, Palo Alto, v. 49, p. 507-531, 2011.

WIN, J. et al. Effector biology of plant-associated organisms: concepts and perspectives. **Cold Spring Harbor Symposia on Quantitative Biology**, Cold Spring Harbor, v. 77, p. 235-247, 2012.

ZAMBOLIM, L. et al. Café (*Coffea arabica* L.), controle de doenças causadas por fungos, bactérias e vírus. In: VALE, F. X. R.; ZAMBOLIM, L. (Ed).

Controle de doenças de plantas: grandes culturas. Visconde do Rio Branco: Suprema, 1997. p. 83-179.

ZAMBOLIM, L. et al. Manejo integrado das doenças do cafeeiro. In:

ENCONTRO SOBRE PRODUÇÃO DE CAFÉ COM QUALIDADE, 1999, Viçosa, MG. **Anais...** Viçosa, MG: UFV, 1999. v. 1, p. 134-215.

ZAMBOLIM, L.; VALE, F. X. R.; ZAMBOLIM, E. M. Doenças do cafeeiro (*Coffea arabica* e *C. canephora*). In: KIMATI, M. et al. (Ed.). **Manual de fitopatologia:** doenças das plantas cultivadas. 4. ed. São Paulo: Agronômica Ceres, 2005. v. 2, p. 165-180.

CAPÍTULO 2 SEQUENCIAMENTO, MONTAGEM E ANOTAÇÃO DO GENOMA DO FUNGO *Hemileia vastatrix*

RESUMO

O estudo da sequência de todo o genoma é essencial para os ramos da pesquisa biológica. A existência de um genoma de referência é de extrema importância para os estudos de interação planta-patógeno. Apesar de muitos anos de pesquisa com o patossistema cafeeiro-*hemileia*, existem poucas sequências genômicas disponíveis. Recentemente, a raça XXXIII do fungo *Hemileia vastatrix* foi identificada no Brasil. Essa raça, por possuir dois ou três genes de virulência (v5,7 ou v5,7,9), está sendo capaz de suplantar a resistência de algumas cultivares de cafeeiro que haviam sido liberadas como resistentes à ferrugem no Brasil, trazendo grandes preocupações para os melhoristas e produtores. Como até o momento, não está disponível nenhuma informação sobre o genoma de *H. vastatrix* raça XXXIII, objetivou-se, neste trabalho, realizar o sequenciamento, montagem e anotação (estrutural e funcional) do genoma desse fungo. Utilizando as plataformas de sequenciamento de nova geração (NGS), Pacific Biosciences (PacBio RS II) e Illumina (HiSeq 2500) e uma estratégia de montagem híbrida, foi possível obter 576 Mb do genoma do isolado Hv-02 (raça XXXIII) do fungo *H. vastatrix*. Observou-se que 96,37% dos genes conservados de eucariotos estão presentes no genoma anotado de *H. vastatrix*, indicando um elevado nível de integridade durante o processo de montagem desse genoma. Também foi possível encontrar 13.364 genes que codificam 13.034 proteínas. As análises de similaridade entre as proteínas preditas no genoma de *H. vastatrix* e as sequências de proteínas de outros fungos, depositadas nos bancos de dados públicos, mostraram 74% de *hit* com fungos da ordem Puccinales, principalmente, com *Puccinia graminis* f. sp. *tritici* e *Melampsora larici-populina*. E 12% não apresentaram similaridade (no-*hits*) com nenhuma proteína descrita nos bancos de dados analisados. Portanto, essas proteínas foram consideradas como exclusivas de *H. vastatrix*. Segundo a análise de categorização funcional, o genoma de *H. vastatrix* raça XXXIII mostrou um grande número de metaloproteínas presentes no genoma desse fungo. Os resultados de sequenciamento e montagem do genoma fungo *H. vastatrix* raça XXXIII disponibiliza dados que servirão de referência para montagens do genoma de outras raças desse fungo, bem como para estudar e compreender os mecanismos envolvidos no processo de interação do patossistema cafeeiro - *H. vastatrix*. Esse conhecimento será útil, também, no desenvolvimento de marcadores moleculares para diferenciar raças, auxiliando nas dificuldades encontradas no uso de plantas diferenciadoras.

Palavras-chave: Sequenciamento de nova geração. Genoma. Proteoma. Categorização funcional. *Blast2GO*.

ABSTRACT

The study of the sequence of the entire genome is essential for biological researches. The existence of a reference genome is of utmost importance for the plant-pathogen interaction studies. Despite many years of research with *coffee-Hemileia* pathosystem, there are few genomic sequences available. Recently, the race XXXIII fungus *Hemileia vastatrix* was identified in Brazil. This race, because it has two or three virulence genes (v5,7 or v5,7,9), being able to overcome the resistance of some coffee cultivars that had been released as resistant to rust in Brazil, bringing great concerns to breeders and producers. As no information is available about *H. vastatrix* race XXXIII genome, the objective of this work is to sequence, assemble and annotate (structurally and functionally) the genome of this fungus. Using next generation sequencing platform (NGS), Pacific Biosciences (PacBio RS II) and Illumina (HiSeq 2500) and a hybrid assembly strategy, isolate Hv-02 (race XXXIII) of fungus *H. vastatrix* has a genome size of 576 Mb. It was found that 96.37% of conserved eukaryotic genes are present in the annotated *H. vastatrix* genome, indicating a high level of integrity during the assembly process. Also, there was 13,364 13,034 genes that encode proteins. The analysis of the similarity between the predicted proteins in the genome of *H. vastatrix* and protein sequences of other fungi deposited in public databases showed 74% hit with fungi order Puccinales, especially with *Puccinia graminis* f. sp. *tritici* and *Melampsora larici-Populina*. And 12% showed no similarity (no-hit) with any protein described in databanks analyzed. Therefore, they were considered unique to *H. vastatrix*. According to the analysis of functional categorization, the genome of *H. vastatrix* race XXXIII, showed a large number of proteins involved in "binding ions". These results suggest that the sequencing and assembly of the fungus *H. vastatrix* race XXXIII genome, provides data that serve as a reference for assembly of the genome of other races of this fungus as well as to study and understand the mechanisms involved in the interaction process of *H. vastatrix* - *coffee* pathosystem. This knowledge will also be useful in the development of molecular markers to differentiate races, helping the difficulties encountered in the use of differential plants.

Keywords: Next generation sequencing. Genome. Proteome. Functional categorization. Blast2GO.

1 INTRODUÇÃO

O desenvolvimento do sequenciamento de DNA há mais de 30 anos impactou profundamente a pesquisa biológica. Hoje as tecnologias de sequenciamento evoluíram, trazendo um número crescente de novos métodos de preparação de amostras e novas ferramentas de análise de dados, gerando uma imensa diversidade de aplicações científicas. Essas tecnologias de sequenciamento de nova geração (*NGS – Next Generation Sequencing*) oferecem um serviço de baixo custo com alto rendimento, além de permitirem uma análise mais abrangente da estrutura e conteúdo dos genomas de microrganismos (FORDE; O'TOOLE, 2013).

A aplicação das *NGSs* em fitopatologia e nos estudos de interação planta-patógeno promete reduzir o tempo da obtenção da informação genética, necessária para compreender as diferentes bases moleculares presentes em cada patossistema. O entendimento da evolução da patogenicidade e dos fatores de virulência é fundamental para o desenvolvimento de estratégias mais eficazes de melhoramento para se obter uma resistência mais durável (CANTU et al., 2011).

No entanto, o progresso nessa área é dificultado pela falta de informações de sequências de genoma. A existência de um genoma de referência é de extrema importância para os estudos de interação planta-patógeno. Recentemente foram publicadas sequências do genoma completo de quatro ferrugens de importância econômica: *Puccinia graminis* f.sp.*tritici* (88.6 Mb), *Melampsora larici-populina*, (101.1 Mb), *Puccinia striiformis* f.sp.*tritici* (64.8 Mb) e *Melampsora lini* (189 Mb) (CANTU et al., 2011; DUPLESSIS et al., 2011; NEMRI et al., 2014). Apesar de muitos anos de pesquisa com o patossistema cafeeiro-*Hemileia vastatrix*, existem poucas sequências genômicas disponíveis: algumas ESTs resultantes de estudos do transcriptoma realizados por Fernandez et al. (2012) e por Talinhas et al. (2014) com a raça XIV e,

recentemente, sequências parciais do genoma híbrido, trabalho realizado por Cristancho et al. (2014).

Até o momento, não está disponível nenhuma informação sobre o genoma de *H. vastatrix* raça XXXIII. Diante do exposto, objetivou-se neste estudo realizar o sequenciamento, montagem e anotação (estrutural e funcional) do seu genoma. Os resultados obtidos serão disponibilizados em bancos de dados para servirem como referência para outros estudos de interação usando esse patossistema.

2 MATERIAL E MÉTODOS

2.1 Multiplicação dos esporos de *H. vastatrix*

Para a realização deste trabalho, foi utilizado o isolado monopostular caracterizado como raça XXXIII (isolado Hv-02). Esse isolado foi caracterizado biologicamente por Capucho et al. (2012) e desde então está sendo mantido na micoteca do Laboratório de Biotecnologia do Cafeeiro (BIOCAFÉ), sediado no Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), na Universidade Federal de Viçosa (UFV), Viçosa, MG.

A multiplicação do isolado foi realizada, de acordo com a metodologia descrita por Zambolim e Chaves (1974), em mudas de cafeeiro de *Coffea arabica* cv. Caturra (CIFC 19/1). A metodologia consiste na deposição dos urediniósporos, na face abaxial das folhas jovens e completamente desenvolvidas, com o auxílio de um pincel. Em seguida, as folhas foram aspergidas com água destilada, utilizando um borrifador, até a obtenção de um leve molhamento superficial. As plantas foram cobertas com saco plástico, a fim de manter o ambiente úmido necessário para a germinação do fungo. Posteriormente, foram mantidas no escuro por 48 horas a 22°C. Após este período, elas foram levadas para câmara de crescimento, com temperatura de 22°C e fotoperíodo de 12 horas, até o aparecimento dos urediniósporos, em torno de 30 dias após a inoculação (dai). Os urediniósporos foram coletados e acondicionados em cápsulas de gelatina dentro de um dessecador, a 4°C e 50% de umidade. Sucessivas inoculações foram realizadas, para garantir inóculos viáveis, com alta porcentagem de germinação e quantidade suficiente para que o trabalho pudesse ser realizado.

2.2 Germinação dos esporos de *H. vastatrix*

Dois gramas de urediniósporos frescos de *H. vastatrix* raça XXXIII (isolado Hv-02) foram suspensos em água destilada estéril contendo 0,05% de Tween 80 (Sigma). Após agitação em vórtex por 20 minutos, os urediniósporos foram recuperados por centrifugação a 2.500 g, por 10 minutos. Este procedimento foi realizado três vezes, para remover impurezas. Os urediniósporos lavados foram distribuídos de forma homogênea, utilizando-se uma espátula estéril, no fundo de placas de poliestireno, na presença de uma fina camada de água. As placas foram fechadas com parafilm® e envolvidas com papel de alumínio para criar um ambiente escuro e úmido necessário à germinação dos esporos. As placas foram incubadas a 22°C por 16 horas. Após esse período avaliou-se a germinação dos esporos utilizando microscópio óptico. Os urediniósporos foram raspados das placas, com auxílio de lamínulas estéreis e, imediatamente, triturados em nitrogênio líquido para extração do DNA.

2.3 Extração de DNA de *H. vastatrix*

O DNA genômico do isolado Hv-02 de *H. vastatrix* foi extraído de 2g de urediniósporos frescos, conforme instruções do kit da Mobio PowerMicrobial® Maxi DNA Isolation. A concentração e a qualidade do DNA obtidas foram verificadas em NanoDrop (Therm Fisher Scientific) e em gel de agarose 1%, respectivamente. O DNA genômico de alta qualidade foi enviado para o *Delaware Biotechnology Institute* (DBI), na *University of Delaware*, em Newark, DE, nos Estados Unidos.

2.4 Sequenciamento e Montagem do genoma de *H. vastatrix*

Essa etapa foi realizada no *Delaware Biotechnology Institute* (DBI), sediado na *University of Delaware*, em Newark, DE nos Estados Unidos.

O sequenciamento do DNA genômico da raça XXXIII (isolado Hv-02) do fungo *H. vastatrix* foi realizado utilizando duas plataformas: Pacific Biosciences (PacBio RS II) e Illumina (HiSeq 2500). Um total de 10 µg de amostra foi usado para construir duas bibliotecas single-end, sequenciadas por PacBio RS II, conforme instruções do kit PacBio Library Preparation[®] com tamanho de 20 kb. A primeira biblioteca foi sequenciada em 15 SMRT células (*Single Molecule RealTime*), usando combinações de dois químicos: XL-C2 (11 SMRT) e P4-C2 (4 SMRT). A segunda biblioteca foi sequenciada em 17 SMRT células usando apenas o químico P5-C3. Para a plataforma Illumina HiSeq 2500, a mesma quantidade de amostra foi usada para construir quatro bibliotecas paired-end com tamanho de 150 pb.

Para avaliar a qualidade dos *reads* gerados pelo sequenciamento na plataforma Illumina, foi utilizado o *software* FastQC, em que *reads* ambíguos, de baixa qualidade e com menos de 35 nucleotídeos foram descartados. Para a limpeza e filtragem dos *reads* duplicados e sobrepostos, foi usado o *software* CLC Genomics Server 6.0.4 (CLC bio, Aarhus, Denmark). Os *reads* de qualidade foram utilizados, para montagem do genoma, usando os *softwares* CLC Genomics Workbench 7.5 (CLC bio, Aarhus, Denmark) e SOAPdenovo v2r215 (LUO et al., 2015). Os *reads* gerados pelo sequenciamento na plataforma PacBio foram avaliados quanto à qualidade, limpos, filtrados e agrupados usando as configurações *default* do *software* PacBio SMRT Portal v2.2.0.p2 (SMRT ANALYSIS SOFTWARE, 2013).

Pelo fato de não existir um genoma de referência para *H. vastatrix*, foi realizada a montagem *De novo* com três estratégias: montagem do genoma,

utilizando somente os *reads* gerados pela plataforma PacBio; montagem do genoma, utilizando somente os *reads* gerados pela plataforma Illumina; e montagem híbrida, utilizando os *reads* gerados em ambas as plataformas de sequenciamento (PacBio e Illumina). Essas estratégias foram usadas na tentativa de fechar os *gaps* e diminuir os erros presentes em cada plataforma de sequenciamento.

A montagem híbrida foi realizada com o *software* CLC Genomic Workbench (CLC Bio, versão 6.04).

2.5 Anotação estrutural do genoma de *H. vastatrix* e predição de genes

Essa etapa, também, foi conduzida no *Delaware Biotechnology Institute* (DBI), sediado na *University of Delaware*, em Newark, DE, nos Estados Unidos.

Foi utilizado um *pipeline* de fácil condução chamado MAKER (CANTAREL et al., 2008), contendo *softwares* responsáveis pela anotação e predição de genes (SNAP, *RepeatMasker*, *Augustus*, *GeneMark* e FGENESH).

Inicialmente, foram identificados elementos repetitivos executando o *software RepeatMasker v4.0.1* (SMIT; HUBLEY; GREEN, 1996). Para realizar a predição de genes, foi utilizado o *software Augustus v2.5.5* (STANKE et al., 2006), por meio de alinhamento com as *ESTs* (*Expressed Sequence Tags*) de *Ustilago maydis* como espécie relacionada. Com essa estratégia, foram identificadas as possíveis sequências codificadoras (*Open Reading Frames* – ORFs). Análise UTRs (*Untranslated Region*) e transcritos alternativos foram realizados, para verificar a qualidade de montagem do gene no espaço do genoma, por meio do *software CEGMA* (PARRA; BRADNAM; KORF, 2007).

2.6 Anotação funcional das proteínas do genoma de *H. vastatrix*

Essa etapa foi realizada juntamente com o grupo de bioinformática da Universidade Federal de Viçosa (UFV), localizado no Núcleo de Análise de Biomoléculas (NuBioMol).

Com o algoritmo BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL et al., 1990), foi realizada uma busca por similaridade entre todas as sequências de proteínas encontradas no genoma de *H. vastatrix* contra o banco de dados de proteínas não redundante (nr) do NCBI (*National Center for Biotechnology Information*) e do UNIPROT (*Universal Protein Resource*). O BLAST foi executado em servidores, por linha de comando, usando scripts desenhados em linguagem *Perl*.

As sequências de proteínas do fungo com similaridade significativa com as sequências de proteínas dos bancos de dados foram analisadas quanto à sua categorização funcional (processo biológico, função molecular e componente celular) por meio do *software Blast2GO* (CONESA et al., 2005). Por outro lado, as sequências de proteínas de *H. vastatrix* que não apresentaram *hit* com as sequências de proteínas depositadas nos bancos foram consideradas possíveis proteínas exclusivas desse fungo, as quais ainda não foram descritas.

Além disso, foi realizada uma busca de similaridade contra o banco de dados *PHI* (*Pathogen-Host Interaction*), que contém proteínas relacionadas à virulência/patogenicidade.

3 RESULTADOS

Após multiplicação massiva dos uredinióporos de *H. vastatrix* (raça XXXIII) por inoculações em mudas de cafeeiro *C. arabica* cv. Caturra (CIFC 19/1), foi realizada a incubação desses esporos frescos em placas de poliestireno por 16 horas no escuro a 22°C. Observou-se cerca de 80% de germinação dos esporos, os quais tiveram o DNA genômico extraído.

O DNA extraído foi sequenciado em duas plataformas. No sequenciamento com a plataforma PacBio RS II, utilizaram-se 32 SMRT células gerando 8,39 Gpb de dados. Após a filtragem, esse número reduziu para 5,73 Gpb de dados compostos de 920.326 *long-reads*, com tamanho médio de 7,000 pb. A sobreposição dos *reads* resultou em *contigs* com tamanho médio de 186 Kb. Com essa plataforma, obteve-se um valor de N50 de 8,6 Kb e uma cobertura de sequenciamento de 20X (Figura 1).

Com a plataforma Illumina HiSeq, foram gerados um total de 264.721.502 de *short-reads*. A sobreposição desses *reads* gerou *scaffolds* com tamanhos superiores a 100 Kb, com uma cobertura de 100X o tamanho do genoma esperado de *H. vastatrix* (733.5Mb), segundo Carvalho et al. (2013). Foram utilizados vários *softwares* de montagem de genomas com os *reads* resultantes da plataforma Illumina HiSeq, entretanto, o melhor desempenho foi obtido pelo *software* SOAP *de novo* (Figura 1).

Na montagem híbrida do genoma, foram testadas múltiplas estratégias, com base nos dados de sequenciamento gerados pelas duas plataformas PacBio RS II e Illumina HiSeq (Figura 1). O objetivo foi maximizar as vantagens e minimizar as desvantagens de cada uma delas e, dessa forma, completar os *gaps* dos *scaffolds*, garantindo melhor montagem do genoma de *H. vastatrix*.

A estratégia que apresentou melhores resultados foi aquela que utilizou os *long-reads* filtrados da plataforma PacBio RS II (HGAP3) e agrupou com os

short-reads filtrados da plataforma Illumina HiSeq (CLC *de novo* com HGAP3). Isso resultou em *scaffolds* de tamanhos maiores que 750 pb. Com o *software* PBJelly, esses *scaffolds* resultantes foram revisados, verificando a qualidade da montagem (Figura 1). Com essa estratégia, foi possível montar 576 Mb do genoma de *H. vastatrix*.

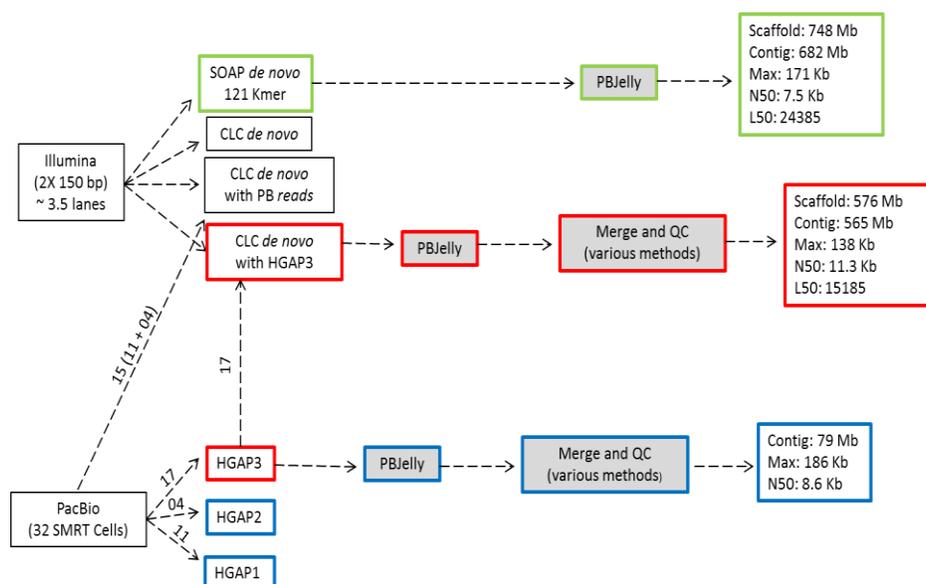


Figura 1 Representação esquemática das três diferentes estratégias utilizadas na montagem do genoma do fungo *H.vastatrix* raça XXXIII. Retângulos azuis representam a primeira estratégia de montagem, a qual, utilizou somente os *reads* gerados pela plataforma PacBio. Retângulos verdes representam a segunda melhor estratégia de montagem, a qual utilizou somente os *reads* gerados pela plataforma Illumina. Retângulos vermelhos representam a terceira e mais completa estratégia de montagem do genoma de *H. vastatrix*, a qual utilizou os *reads* gerados pelas duas plataformas de sequenciamento de nova geração.

As informações finais desse sequenciamento e montagem híbrida podem ser observadas na Tabela 1.

Tabela 1 Informações estatísticas sobre a montagem do genoma do fungo *Hemileia vastatrix* raça XXXIII

Número de <i>Scaffolds</i>	118162
Tamanho do maior <i>Scaffold</i>	137364 pb
Tamanho do menor <i>Scaffold</i>	143 pb
N50	9965 pb
L50	16308
Número de <i>Scaffolds</i> > 500 nt	118151
Número de <i>Scaffolds</i> > 1K nt	92276
Número de <i>Scaffolds</i> > 10K nt	16232
Número de <i>Scaffolds</i> > 100K nt	2
Número de <i>Scaffolds</i> > 1M nt	0
%AT	66
%GC	34
%N	0

Com base no genoma montado de *H. vastatrix*, foram realizadas anotações e predições de genes. Nos 118.162 *scaffolds*, foram identificados elementos repetitivos, utilizamos o *software RepeatMasker*. Essa análise revelou que 82% desses *scaffolds* montados são compostos por elementos repetitivos. Os elementos transponíveis representaram 35%, sendo o elemento *Gypsy* encontrado em maior número (Tabela 2).

Tabela 2 Resultados da análise de anotação estrutural do genoma do fungo *Hemileia vastatrix* raça XXXIII

Tipo de Elementos Repetitivos	Número de Elementos Repetitivos
Repetição Simples	39808
Baixa complexidade	32461
<i>Gypsy</i>	41269
<i>Copia</i>	7054
<i>EnSpm</i>	2971
<i>Harbinger</i>	2954
<i>Mariner</i>	837
<i>LINEs</i>	426
<i>Helitron</i>	181
<i>DIRs</i>	81
<i>MuDR</i>	16
<i>hAT</i>	10

Para identificar as ORFs presentes no genoma desse fungo, foi realizada uma predição de genes, utilizando as *ESTs* de *Ustilago maydis*. Espécie presente nos banco de dados do *software Augustus v2.5.5* mais relacionada com relacionada com *H. vastatrix*.

Foram identificados 13.364 genes em *H. vastatrix*. Para validar o processo de anotação gênica, usou-se o *software CEGMA* (PARRA; BRANDNAM; KORF, 2007), o qual faz uma busca por genes conservados dentro de um conjunto de 248 genes eucariotos. Observou-se que 239 (96,37%) desses genes conservados de eucariotos estão presentes, no genoma anotado de *H. vastatrix*, indicando um elevado nível de integridade durante o processo de montagem desse genoma.

Utilizando o *software Augustus*, observou-se que, dos 13.364 genes encontrados, 13.034 codificam proteínas. Essas proteínas foram analisadas quanto à similaridade contra os bancos de dados de sequências de proteínas não redundantes do NCBI e UNIPROT, por meio do algoritmo BLASTp. Com os resultados, observou-se que 1.568 sequências não apresentaram *hit* com as

sequências de proteínas, depositadas nos bancos de dados consultados, sendo então consideradas como possíveis proteínas exclusivas de *H. vastatrix*. Por outro lado, 6.944 sequências apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com sequências de proteínas de *Puccinia graminis*, 2.600 com sequências de proteínas de *Melampsora larici-pupulina* e 75 com *H. vastatrix* (Figura 2).

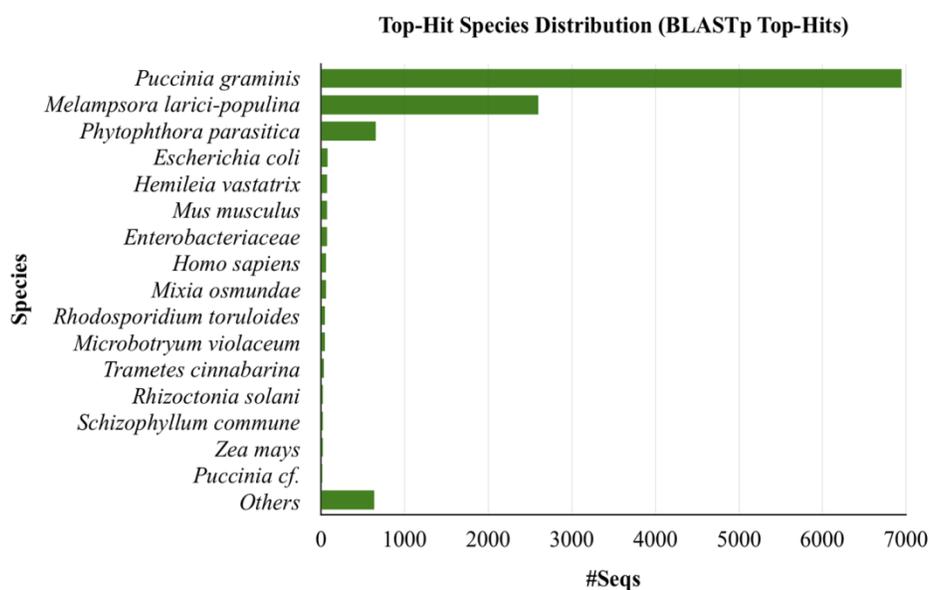


Figura 2 Resultado da busca por similaridade entre as sequências de proteínas presentes no genoma de *H. vastatrix* raça XXXIII contra os bancos de dados públicos de sequências de proteínas não redundantes (NCBI e UNIPROT), por meio do algoritmo BLASTp. No eixo y estão as espécies encontradas com maior frequência na busca por similaridade. No eixo x estão os números de proteínas encontradas para cada espécie

Utilizando-se as 11.466 sequências de proteínas que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos banco de dados (NCBI/UNIPROT), foi realizada uma análise de categorização funcional, com o *software* BLAST2GO. Destas, 7.139 proteínas

foram caracterizadas. No total, foram obtidas 12.501 anotações distribuídas em três categorias (Componente Celular, Função Molecular e Processos Biológicos) e em diferentes níveis hierárquicos do *GO*. Algumas proteínas foram anotadas em duas ou mais categorias, sendo a Função Molecular a categoria mais representada com 6.331 anotações, seguida da categoria Processo Biológico com 3.887 anotações e, por último, a categoria menos representada, Componente Celular, com 2.283 anotações.

Para a categoria Função Molecular, considerando o nível hierárquico 03, os termos “ligação a íons” e “atividade hidrolase” foram os mais representados com 889 e 538 proteínas, respectivamente (Figura 3).

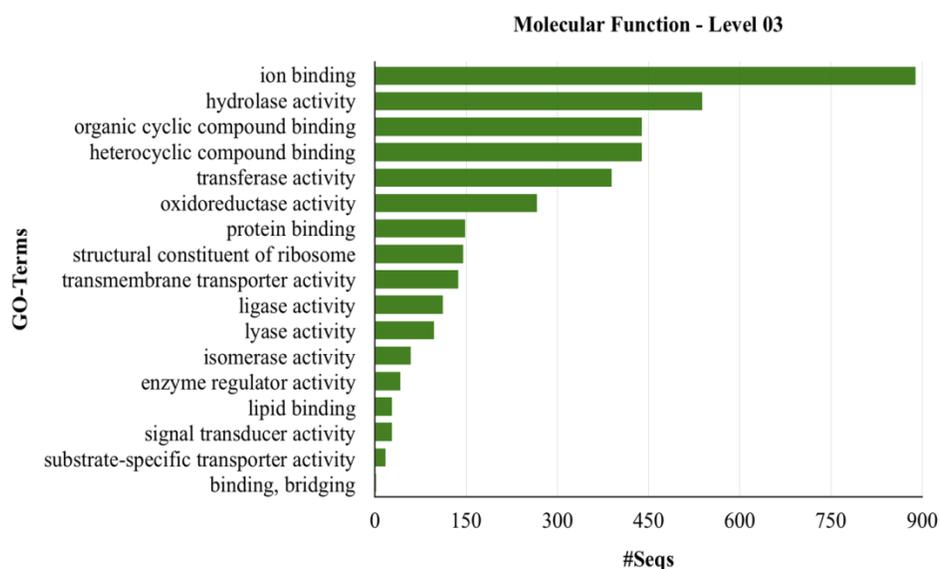


Figura 3 Resultado da análise de categorização funcional das proteínas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Função Molecular, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria

Para a categoria Processo Biológico, também, levando em consideração o nível hierárquico 03, “processo metabólico celular”, “processo metabólico primário” e “processo metabólico de substâncias orgânicas”, foram os termos mais representados com 1.243, 1.086 e 1.086 proteínas, respectivamente (Figura 4).

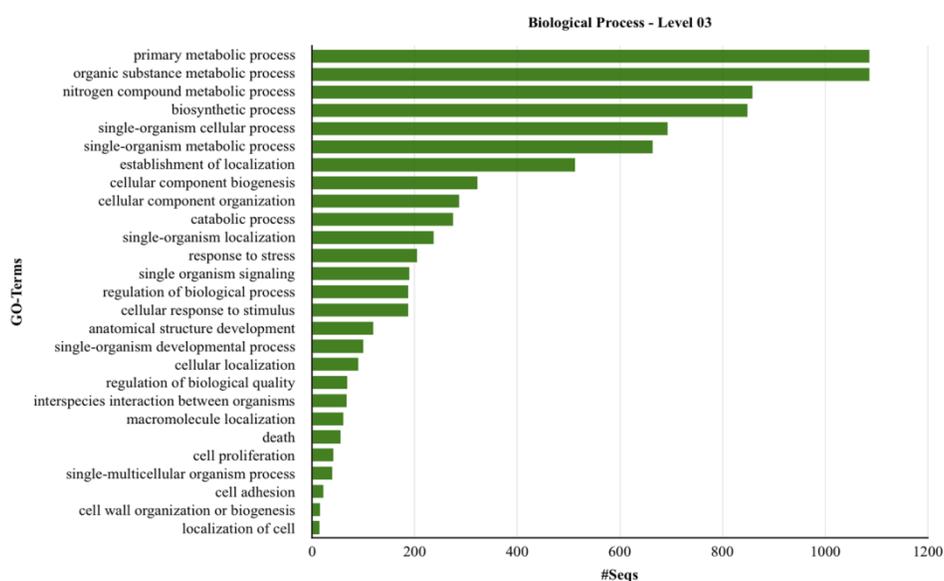


Figura 4 Resultado da análise de categorização funcional das proteínas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Processo Biológico, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria

E, na categoria Componente Celular, os termos mais representados, com o nível hierárquico 03, foram “parte celular”, “organela membranosa”,

“complexo de proteínas” e “organela não-membranosa” com 1.326, 669, 511 e 464 proteínas, respectivamente (Figura 5).

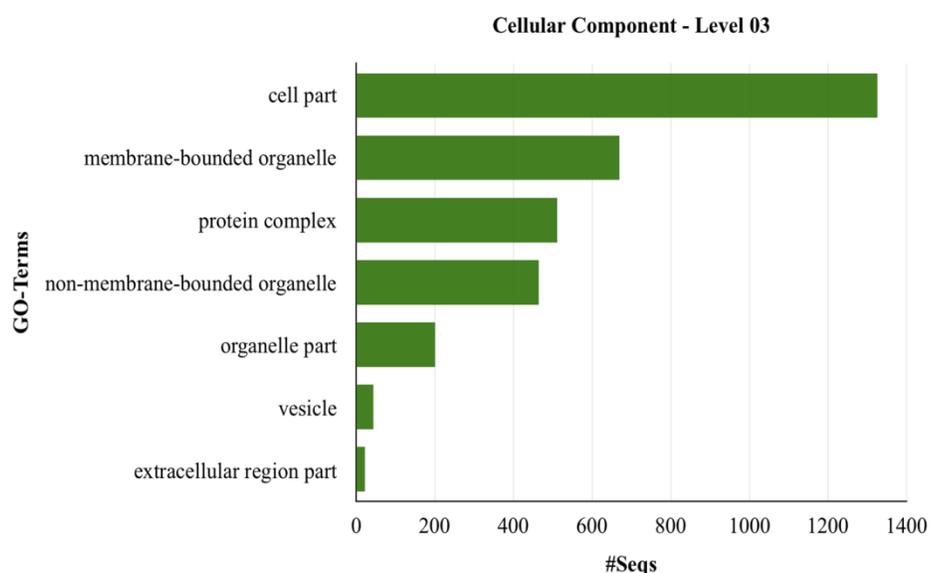


Figura 5 Resultado da análise de categorização funcional das proteínas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Componente Celular, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria

Este resultado mostra que a atuação dos genes em questão não está restrita a determinado componente celular, mas abrange toda a célula. Isso foi confirmado pela análise de outros níveis hierárquicos (06 e 08) a qual revelou que os genes estão associados, principalmente, ao núcleo, citosol, mitocôndria, ribossomo, complexo de Golgi e retículo endoplasmático (Figura 6).

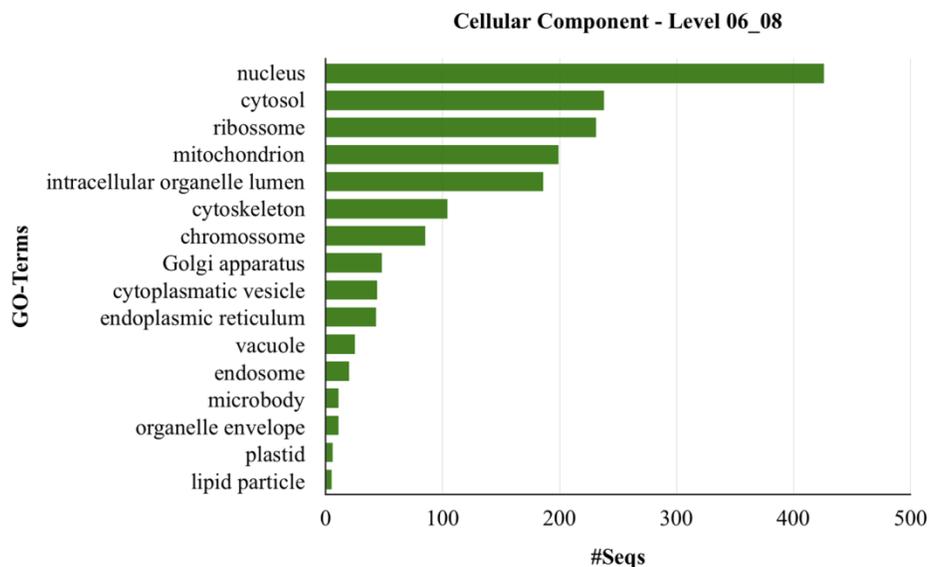


Figura 6 Resultado da análise de categorização funcional das proteínas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO. No eixo y estão os termos do GO descritos para a categoria Componente Celular, nos níveis hierárquicos 06 e 08. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria

Nestes componentes ocorrem os mais variados processos biológicos, como os mencionados anteriormente, o que justifica sua maior representação.

Observou-se, também, que 357 proteínas apresentaram alta similaridade (> 50%), quando analisadas com as proteínas associadas com a interação planta-patógeno, depositadas no banco de dados *PHI*.

4 DISCUSSÃO

Por meio da utilização de duas plataformas de sequenciamento de nova geração, PacBio RS II e Illumina HiSeq, foi gerada uma montagem híbrida e *de novo* (ou *ab initio*) para o genoma do fungo *H. vastatrix* raça XXXIII com um tamanho de 576 Mb. O tamanho do genoma obtido na montagem é inferior ao relatado nos estudos realizados por Carvalho et al. (2013) e Tavares et al. (2014), que estimaram, por meio de citometria de fluxo, tamanho de 733,5 Mb e 796,8 Mb, respectivamente. Essa divergência pode ser em razão do uso de diferentes estratégias que levam a estimativas distintas do genoma. Por outro lado, Cristancho et al. (2014), usando o sequenciamento e montagem *denovo* do genoma, obtiveram um tamanho de apenas 333 Mb. A grande quantidade de regiões repetitivas do genoma de *H. vastatrix* dificulta o processo de sequenciamento e de montagem, gerando essas divergências quanto ao tamanho obtido. Além disso, o tipo de plataforma utilizada, para realizar o sequenciamento e os *softwares* usados na montagem, podem afetar os resultados. Outro fator que pode ter influenciado para essa diferença de tamanho na montagem do genoma *H. vastatrix* foi o fato de que Cristancho et al. (2014) utilizaram uma mistura de raças para fazer o sequenciamento, o que pode levar à maior dificuldade na montagem do genoma. No presente trabalho, foi sequenciado e montado o genoma de um único isolado, o Hv-02 (raça XXXIII).

H. vastatrix possui genoma considerado grande para fungo. Em comparação com o tamanho do genoma de outras ferrugens, apenas *Gymnosporangium. confusum* com 893,2Mb e *Puccinia. chrysanthemi* com 806,5Mb (TAVARES et al., 2014) possuem genomas maiores. As demais ferrugens de grande importância econômica, como *P. graminis* f.sp. *tritici* (88,6 Mb), *P. striiformis* f. sp. *tritici* (64,8 Mb), *M. larici-populina* (101,1 Mb) e *M.*

lini (189 Mb), apresentam genomas relativamente pequenos (CANTU et al., 2011; DUPLESSIS et al., 2011; NEMRI et al., 2014).

Em geral, as ferrugens têm genomas maiores que outros fungos. Outras características que interferem nessa diferença de tamanho dos genomas são: estilo de vida, variabilidade intraespecífica, pressão de seleção e deriva genética. Segundo Spanu (2012), o biotrofismo, um estilo de vida compartilhado entre a ordem Puccinales, é responsável pelo aumento do tamanho do genoma desses fitopatógenos, comparado aos fungos não biotróficos. E, por isso, genomas grandes são característicos dessa ordem. Entretanto, essa variação pode ter menos correlação com a filogenia e, sim, com a evolução que ocorre ao longo do tempo entre patógeno e hospedeiro. Segundo Stukenbrock e Croll (2014), a variabilidade do genoma do patógeno é adaptável ao seu hospedeiro, para que possa superar suas defesas.

No entanto, verifica-se que possuir genomas grandes não significa ter maior número de genes, uma vez que as expansões no tamanho dos genomas podem ocorrer por aumento de regiões repetitivas e proliferação de elementos transponíveis (ET) (SPANU, 2012). Nesse estudo, encontrou-se alta porcentagem de elementos repetitivos (82%), dos quais 35% representam elementos transponíveis (35%). Em estudos realizados em outras ferrugens, observou-se que *P. striiformis* f. sp. *tritici*, *P. graminis* f. sp. *tritici*, *M. laricipopulina*, *M.lini* e *H. vastatrix* apresentaram 17,8%, 42,5%, 40,5%, 31,67% e 48% de ET, respectivamente (CANTU et al., 2011; CRISTANCHO et al., 2014; DUPLESSIS et al., 2011; NEMRI et al., 2014). Essa diferença na quantidade de ET, encontrada em outros estudos, pode ocorrer por dois motivos: tamanho muito grande do genoma que pode conter uma grande proporção de regiões repetitivas e pela montagem híbrida que pode superestimar o número real de regiões repetitivas. Os ET são importantes fontes de variabilidade, que ocorrem por meio de mutações durante o processo evolutivo. Eles, também, são

responsáveis pela expansão do tamanho dos genomas, uma vez que aumentam as regiões repetitivas e causam a duplicação do genoma. Pelo fato do patógeno possuir um ciclo de vida mais curto do que seu hospedeiro, ele se reproduz mais vezes que a planta, evoluindo mais rápido. Por esse motivo, que *H. vastatrix* durante a coevolução com o cafeeiro apresenta maior capacidade de adaptação e alta taxa de variabilidade genética.

A porcentagem de GC, presente no genoma de *H. vastatrix* neste estudo, foi relativamente igual à encontrada para o mesmo organismo nos estudos de Cristancho et al. (2014) com 33,74%. Por outro lado, quando comparada com *P. graminis* f. sp. *tritici* (43.3%) e *M. larici-populina* (41%) foi mais baixa. Essa redução ocorre pela subrepresentação de sequências GC em sequências repetitivas na formação de *contigs* (CRISTANCHO et al., 2014).

Ao contrário do que foi mostrado na montagem do genoma de *H. vastatrix*, realizada por Cristancho et al. (2014), que se apresentava altamente fragmentada e com um valor de N50 muito baixo, no presente trabalho, foi obtida uma montagem com elevado nível de integridade com valor de N50 igual a 11,3 Kb. Isso foi validado com as análises realizadas no CEGMA, onde se observou a presença de 96,37% de genes conservados de eucariotos no genoma.

A grande porcentagem de similaridade (74% de *hit*) entre as proteínas preditas no genoma de *H. vastatrix* e as sequências de proteínas de *P. graminis* f. sp. *tritici* e *M. larici-populina*. pode ser explicada por pertencerem à mesma ordem de *H. vastatrix*, podendo, portanto, apresentarem conteúdo gênico similar e necessário para a patogenicidade. Além disso, *P. graminis* f. sp. *tritici* e *M. larici-populina* apresentam seu genoma completo, disponibilizado com um grande número de sequências depositadas nos bancos de dados (DUPLESSIS et al., 2011). Observou-se, também, que essas proteínas apresentaram alta similaridade com as proteínas associadas com virulência\patogenicidade depositadas no banco de dados *PHI*. Dessa forma, representam uma fonte de

recurso para estudar sua relação de interação com seu hospedeiro e sua virulência.

Do total de proteínas obtidas, 12% não apresentaram similaridade (no-*hits*) com nenhuma proteína descrita nos bancos de dados analisados. Portanto, foram consideradas como exclusivas de *H. vastatrix*, pois, ainda, não foram descritas.

Na análise de categorização funcional, o genoma de *H. vastatrix* mostrou um grande número de metaloproteínas, destacando-se as proteínas ligadas ao zinco. Esse elemento é um componente estrutural das proteínas de eucariotos. Em fungos patogênicos, esses elementos são importantes, durante o processo infeccioso, pois são responsáveis por ativar a expressão das vias de absorção de zinco, atuando assim na virulência do fungo (STAATS et al., 2013). Outras proteínas abundantes foram as envolvidas na “atividade hidrolase”, que atuam na remodelagem da parede celular do fungo ou na degradação da parede celular do hospedeiro. Essas proteínas podem atuar, também, na conversão de quitina em quitosana, para impedir o reconhecimento pelos receptores de quitina, presentes no hospedeiro, favorecendo a infecção. Por esta razão, sua maior expressão deve ocorrer nos estágios iniciais da infecção, como verificado em esporos germinados de *H. vastatrix* (VIEIRA et al., 2012), *Puccinia graminis* f. sp. *tritici* e *M. larici-populina* (DUPLESSIS et al., 2011).

Relacionadas, muitas vezes, com os estágios iniciais da infecção encontram-se, também, as proteínas envolvidas nos “processos metabólicos” que foram encontradas em grande número no genoma de *H. vastatrix*.

5 CONCLUSÃO

O sequenciamento do genoma do fungo *H. vastatrix* raça XXXIII, por meio das plataformas de sequenciamento Pacific Biosciences (PacBio RS II) e Illumina (HiSeq 2500), disponibiliza dados que servirão de referência para montagens do genoma de outras raças desse fungo, bem como para estudar e compreender os mecanismos envolvidos no processo de interação desse patossistema cafeeiro - *H. vastatrix*.

Neste trabalho, obteve-se o maior, mais completo e mais íntegro genoma de *H. vastatrix* sequenciado até hoje, fornecendo dados para estudar a evolução e compreender os mecanismos moleculares envolvidos no processo de interação com a planta. Esse conhecimento será útil no desenvolvimento de marcadores moleculares para diferenciar raças, auxiliando nas dificuldades encontradas no uso de plantas diferenciadoras. Diferentes raças fisiológicas podem apresentar diferentes genes de virulência, os quais conferem a capacidade de infectar ou não determinado genótipo. Como também será essencial para desenvolver cultivares de cafeeiro com resistência mais durável, como alternativa para o controle da doença.

REFERÊNCIAS

- ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, London, v. 215, n. 3, p. 403-410, Oct. 1990.
- CANTAREL, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. **Genome Research**, Cold Spring Harbor, v. 18, n. 1, p. 188-196, Jan. 2008.
- CANTU, D. et al. Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. **PLoS One**, San Francisco, v. 6, n. 8, p. 24230, 2011.
- CAPUCHO, A. S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Note**, Melbourne, v. 7, p. 189-191, Oct. 2012.
- CARVALHO, G. M. A. et al. Coffee rust genome measured using flow cytometry: does size matter? **Plant Pathology**, London, v. 63, n. 5, p. 1022-1026, Oct. 2013.
- CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, Oxford, v. 21, n. 18, p. 3674-3676, Sept. 2005.
- CRISTANCHO, M. A. et al. Annotation of a hybrid partial genome of the coffee rust (*Hemileia vastatrix*) contributes to the gene repertoire catalog of the Pucciniales. **Frontiers in Plant Science**, Lausanne, v. 5, p. 594-595, Oct. 2014.
- DUPLESSIS, S. et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proceedings of the National Academy of Sciences of the States of America**, Washington, v. 108, n. 22, p. 9166-9171, May 2011.
- FERNANDEZ, D. et al. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. **Molecular Plant Pathology**, Oxford, v. 13, n. 1, p.17-37, Jan. 2012.
- FORDE, B. M.; O'TOOLE, P. W. Next-generation sequencing technologies and their impact on microbial genomics. **Briefings in Functional Genomics**, Oxford, v. 12, n. 5, p. 440-453, Jan. 2013.

LUO, R. et al. Erratum: SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. **Giga Science**, London, v. 4, n. 1, p. 1-18, July 2015.

NEMRI, A. et al. The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. **Frontiers in Plant Science**, Lausanne, v. 5, p. 1-14, Mar. 2014.

PARRA, G.; BRADNAM, K.; KORF, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. **Bioinformatics**, Oxford, v. 23, n. 9, p.1061-1067, May 2007.

SMIT, A.; HUBLEY, R.; GREEN, P. **RepeatMaskerOpen-3.0**. Seattle, 1996. Disponível em: <<http://www.repeatmasker.org/faq.html>>. Acesso em: 15 jul. 2014.

SMRT ANALYSIS SOFTWARE. Disponível em: <<http://www.pacb.com/products-and-services/analytical-software/devnet/>>. Acesso em: 02 set. 2013.

SPANU, P. D. The genomics of obligate (and nonobligate) biotrophs. **Annual Review Phytopathology**, Palo Alto, v. 50, p. 91-109, 2012.

STAATS, C. C. et al. Fungal zinc metabolism and its connections to virulence. **Frontiers in Cellular and Infection Microbiology**, Lausanne, n. 3, p. 65, Oct. 2013.

STANKE, M. et al. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. **BMC Bioinformatics**, Oxford, v. 7, n. 1, p. 62, Feb. 2006.

STUKENBROCK, E. H.; CROLL, D. The evolving fungal genome. **Fungal Biology Reviews**, Amsterdam, v. 28, n. 1, p. 1-12, Feb. 2014.

TALHINHAS, P. et al. Overview of the functional virulent genome of the coffee leaf rust pathogen *Hemileia vastatrix* with an emphasis on early stages of infection. **Frontiers in Plant Science**, Lausanne, v. 5, p. 88-91, Mar. 2014.

TAVARES, S. et al. Genome size analyses of Pucciniales reveal the largest fungal genomes. **Frontiers in Plant Science**, Lausanne, v. 5, p. 11-112, Mar. 2014.

VIEIRA, A. et al. Expression profiling of genes involved in the biotrophic colonisation of *Coffea arabica* leaves by *Hemileia vastatrix*. **European Journal of Plant Pathology**, Dordrecht, v. 133, n. 1, p. 261-277, 2012.

ZAMBOLIM, L.; CHAVES, G. M. Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredosporos de *Hemileia vastatrix* Berk. et Br. e *Uromyces phaseolityca* Arth. **Experientiae**, Viçosa, MG, v. 17, p. 151-184, 1974.

CAPÍTULO 3 PREDIÇÃO DO SECRETOMA, IDENTIFICAÇÃO E VALIDAÇÃO DE EFETORES DE *Hemileia vastatrix*

RESUMO

Hemileia vastatrix, assim como os demais fungos fitopatogênicos, possui a capacidade de produzir proteínas que melhoram seu potencial para infectar seu hospedeiro. Essas proteínas são conhecidas como efetores, as quais são capazes de alterar a fisiologia do hospedeiro, suprimindo as respostas de defesa. Desse modo, facilitam a infecção e conseguem estabelecer a colonização do patógeno no hospedeiro. Algumas dessas proteínas efetoras são reconhecidas por proteínas codificadas por genes de resistência o que desencadeia uma resposta de defesa da planta contra a infecção do patógeno. Uma abordagem que pode ser realizada, com a finalidade de detectar essas proteínas, é a análise do secretoma do fungo e a utilização de ferramentas de bioinformática para identificar genes candidatos a efetores de *Hemileia vastatrix*, seguida da validação por meio de PCR em tempo real. Apesar de muitos anos de pesquisa neste patossistema (*H. vastatrix* – cafeeiro), até o momento, alguns poucos efetores desse fungo foram identificados. Para ampliar o conhecimento das proteínas efetoras de *H. vastatrix*, objetivou-se, neste trabalho, predizer o secretoma, com base no proteoma obtido do sequenciamento da raça XXXIII desse fungo, visando identificar e caracterizar o perfil de expressão desses genes candidatos a efetores, por meio de PCR em tempo real. Das 13.034 proteínas preditas encontradas no genoma da raça XXXIII, foram identificadas 615, contendo peptídeo, sinal localizado na via de secreção e sem domínios transmembrânicos. Desse secretoma obtido, 111 proteínas foram consideradas candidatas a efetores específicos de *H. vastatrix*. Com base nessas análises, foram selecionados 17 genes (*EHv33*) para serem caracterizados por PCR em tempo real. As análises da expressão temporal desses genes (*EHv33*) mostraram que, com exceção dos genes *EHv33_3* e *EHv33_14*, todos os outros 15 genes (*EHv33*) apresentaram uma expressão variada, ao longo dos diferentes tempos (24, 48 e 72 h.a.i) e nas diferentes interações (incompatível e compatível). Pôde-se observar que 5 desses genes foram mais expressos, significativamente, durante os tempos iniciais da interação incompatível, sendo considerados genes candidatos a efetores pré-haustoriais. Outros 9 genes *EHv33* foram mais expressos, significativamente, depois da formação dos haustórios, na interação compatível, podendo ser considerados genes candidatos a efetores translocados via haustório. Logo, esse conhecimento será útil no desenvolvimento de marcadores moleculares para diferenciar raças, auxiliando nas dificuldades encontradas no uso de plantas diferenciadoras.

Palavras-chave: Ferrugem do cafeeiro. Interação planta-patógeno. Efetores. Caracterização de proteínas efetoras. Expressão gênica.

ABSTRACT

Hemileia vastatrix, as well as other pathogenic fungi, have the capability to produce proteins that enhance its potential to infect their host. These proteins are known as effectors, which are able to change the physiology of the host, suppressing defense responses. Thus, to facilitate infection and then establish colonization of the pathogen in the host. Some of these effector proteins are recognized by proteins encoded by resistance genes, which triggers a plant defense response against infection of the pathogen. One approach that may be performed in order to detect these proteins is the analysis of the fungus and secretome using bioinformatic tools to identify candidate *Hemileia vastatrix* effector genes, then validate by real time qPCR (RT-qPCR). Despite many years of research in this pathosystem (*H. vastatrix* - coffee), so far, few effectors of this fungus were identified. Thus, given the above, to increase knowledge about the effector proteins of *H. vastatrix*, the objective of this work is; to predict secretome based on the obtained race XXXIII proteome sequence of this fungus, to identify and characterized the expression profile of candidate genes effectors by means of real time PCR. Of the 13,034 predicted proteins found in the genome of the race XXXIII, 615 were identified containing the signal peptide located in the secretion pathway and with no transmembrane domains. From the obtained secretome, 111 proteins were considered candidates for specific effectors of *H. vastatrix*. Based on this analysis, we selected 17 genes (EHv33) to be validated by real time. Analyses of the expression of these genes (EHv33) showed that, with the exception of EHv33_3 and EHv33_14 genes, all the other 15 genes (EHv33) showed mixed expression over the different times (24, 48 and 72 hai) and different interactions (incompatible and compatible). It was observed that 5 genes were more expressed significantly during the early days of the incompatible interaction, are considered candidate genes pre-haustoriais effectors. Other 9 EHv33 genes were more expressed significantly after the formation of haustoria in the compatible interaction and can be considered candidate genes translocated effectors via haustorium. Thus, this knowledge will be useful in the development of molecular markers to differentiate races, helping the difficulties encountered in the use of differential plants.

Keywords: Coffee rust. Plant-pathogen interaction. Effectors. Validation effector proteins. Gene expression.

1 INTRODUÇÃO

A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* (Berkely & Broome), é responsável pela principal doença foliar que afeta a produção do café arábica (*Coffea arabica* L.) (SILVA et al., 2006). Esse fungo, como os demais agentes causadores de ferrugens, são parasitas biotróficos, ou seja, dependem das células vivas do seu hospedeiro para se desenvolver e se reproduzir. Durante o processo infeccioso, eles formam diversas estruturas especializadas que irão contribuir para o seu crescimento e desenvolvimento no interior de seu hospedeiro, como, por exemplo, os haustórios. Essa estrutura tem papel tanto na nutrição, uma vez que é responsável pela absorção de nutrientes das células de seu hospedeiro, quanto na secreção de proteínas efetoras para o interior das células dos hospedeiros, facilitando, assim, a entrega dos efetores na matriz extra-haustorial de onde muitos serão translocados para outros compartimentos subcelulares (WIN et al., 2012).

Sabe-se que os efetores, secretados pelos fitopatógenos, representam uma das diversas estratégias de ataque ao hospedeiro e que inúmeros efetores têm potencial para manipular a defesa das plantas. Eles são essenciais para a sua virulência, pois promovem a sua aptidão reprodutiva por meio da manipulação e supressão das respostas fisiológicas e do sistema de defesa inato do hospedeiro durante a infecção (DUPLESSIS et al., 2011; PETRE; JOLY; DUPLESSIS, 2014).

No entanto, isso contrasta com o que se vê na natureza, onde resistência é a regra, e a suscetibilidade, a exceção. Em outras palavras, uma dada espécie de planta pode ser resistente para a maioria dos patógenos e suscetível a apenas alguns. Isto se dá em função de relações incompatíveis, sendo chamada de resistência de não hospedeiro (non-host resistance). Este estudo, no entanto, está focado em identificar genes candidatos a efetores e caracterizar seu perfil de

expressão tanto nas interações incompatíveis quanto nas interações compatíveis, visando desenvolver novos métodos de controle dessa doença.

Como mencionado anteriormente, efetores são fatores de virulência, mas cada um contribuindo, parcialmente, para a virulência final. Efetores independentes podem atuar de forma redundante, podendo alterar uma mesma via de sinalização do hospedeiro. Portanto, os genes efetores podem ser perdidos sem causar impacto significativo na virulência do patógeno. Com exceção, os “efetores fundamentais”, que estão amplamente distribuídos na população de um determinado patógeno, contribuindo para a sua virulência (DANGL; HORVATH; STASKAWICZ, 2013).

Os genes de fungos são considerados candidatos a efetores, quando apresentarem, basicamente, peptídeo sinal, estiverem localizadas nas vias de secreção e não apresentarem domínios transmembranas (SAUNDERS et al., 2012).

Com o avanço das técnicas moleculares e sequenciamento de nova geração, muita informação tem sido disponibilizada. Por meio do sequenciamento total do genoma de fungos fitopatogênicos, tem sido possível prever genes candidatos a efetores em alguns patógenos (DALIO et al., 2014). Com isso, várias estratégias têm sido adotadas para se obter plantas com resistência mais durável a doenças, por meio do uso de efetores, como, por exemplo, o isolamento do gene R (PERIYANNAN et al., 2013; SAINTENAC et al., 2013). Essa estratégia pode ser realizada por ensaios de coexpressão transiente seguido por melhoramento de marcador assistido ou desenvolvimento de transgênico (GAJ; GERBACH; BARBAS, 2014; JINEK et al., 2013; SCHORNACK et al., 2013).

Outra estratégia utilizada é o emprego de receptores do sistema imune, em que cultivares geradas, seja por melhoramento molecular assistido ou por transferência gênica, devem exibir resistência mais durável a doenças, visto que

a evasão do patógeno exigirá mutações em múltiplos genes candidatos a efetores (DANGL; HORVATH; STASKAWICZ, 2013).

Apesar de muitos anos de pesquisa no patossistema cafeeiro - *H. vastatrix*, existem poucos estudos relatados com o secretoma desse fungo. Até o momento, tem-se, por meio do sequenciamento de *ESTs* de urediniósporos germinados, apressório e haustório e do recente sequenciamento parcial do genoma híbrido de *H. vastatrix*, alguns efetores identificados (CRISTANCHO et al., 2014; FERNANDEZ et al., 2012; TALHINHAS et al., 2014). Entretanto, essas informações são, ainda, insuficientes, uma vez que conseguem identificar apenas uma pequena fração do secretoma total presente no genoma completo desse fungo.

Sendo assim, objetivou-se, neste trabalho, predizer o secretoma de *H. vastatrix* raça XXXIII, visando identificar genes candidatos a efetores, por meio do proteoma obtido do sequenciamento e, posteriormente, caracterizar o perfil de expressão de alguns desses efetores encontrados utilizando PCR em tempo real.

2 MATERIAL E MÉTODOS

2.1 Multiplicação dos esporos de *H. vastatrix*

Para a realização deste trabalho, foi utilizado o isolado monopostular Hv-02 (raça XXXIII). Esse isolado foi caracterizado biologicamente por Capucho et al. (2012) e desde então está sendo mantido na micoteca do Laboratório de Biotecnologia do Cafeeiro (BIOCAFÉ), sediado no Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), na Universidade Federal de Viçosa (UFV), Viçosa, MG.

A multiplicação do isolado foi realizada, de acordo com a metodologia descrita por Zambolim e Chaves (1974), em mudas de cafeeiro *Coffea arabica* cv. Caturra (CIFC 19/1), a qual consiste na deposição dos urediniósporos na face abaxial das folhas jovens e completamente desenvolvidas com o auxílio de um pincel. Em seguida, as folhas foram aspergidas com água destilada, utilizando um borrifador, até a obtenção de um leve molhamento superficial e cobertas com saco plástico, a fim de manter o ambiente úmido necessário para a germinação do fungo. Posteriormente, as plantas foram mantidas no escuro por 48 horas a 22°C. Após este período, as plantas foram levadas para câmara de crescimento, com temperatura de 22°C e fotoperíodo de 12 horas, até o aparecimento dos urediniósporos, em torno de 30 dias após a inoculação (dai), os quais foram coletados e acondicionados em cápsulas de gelatina dentro de um dessecador, a 4°C e 50% de umidade. Sucessivas inoculações foram realizadas, para garantir um inóculo viável, ou seja, um inóculo com alta porcentagem de germinação para se obter uma quantidade suficiente de esporos para que o trabalho pudesse ser realizado.

2.2 Teste de viabilidade dos esporos de *H. vastatrix*

A viabilidade dos urediniósporos de *H. vastatrix* foi avaliada pelo teste de germinação em meio ágar-água 2%, utilizando o método descrito por Zambolim e Chaves (1974), antes da inoculação nas plantas. Somente os urediniósporos com viabilidade superior a 30% foram considerados adequados para iniciar o trabalho

Essa metodologia consiste em realizar uma suspensão de urediniósporos a uma concentração de 2mg/mL em água destilada, acrescentando-se *tween* (0,01%). Espalham-se 300µL dessa suspensão, na superfície das placas de petri, as quais são incubadas por 24 horas no escuro a 22°C. Após esse período, procede-se à contagem dos urediniósporos com auxílio do microscópio de luz Carl Zeiss. A porcentagem de germinação é dada pelo número de esporos germinados em relação ao número total de esporos presentes na placa.

2.3 Material vegetal e inoculação dos esporos de *H. vastatrix*

Foram utilizadas mudas de cafeeiro resistentes, Híbrido de Timor (CIFC 832/1), que possui os genes de resistência S_{H5} , S_{H6} , S_{H7} , S_{H8} , S_{H9} , $S_H?$ e mudas de cafeeiro suscetíveis, *C. arabica* cv. Caturra (CIFC 19/1), que possui somente o gene de resistência S_{H5} . As plantas foram mantidas em estufa e transferidas para uma câmara de crescimento com temperatura e fotoperíodo controlados (22°C de temperatura e 12 horas de luz) 24 horas antes da inoculação. Essas plantas foram inoculadas com esporos viáveis de *H. vastatrix* raça XXXIII, estabelecendo interações incompatíveis (plantas resistentes) e interações compatíveis (plantas suscetíveis). A inoculação ocorreu como descrito no item 2.1, sendo inoculado um par de folhas jovens completamente desenvolvidas em diferentes tempos: 12, 24, 48 e 72 horas. Folhas de três plantas por tratamento (4

tempos e duas cultivares) foram coletadas, imediatamente congeladas por imersão em nitrogênio líquido e armazenadas a -80°C até a extração de RNA.

2.4 Extração de RNA e síntese de cDNA

O RNA total foi extraído das amostras congeladas de folhas de cafeeiro (resistentes e suscetíveis) inoculadas com *H. vastatrix* raça XXXIII, utilizando o kit RNeasy Plant Mini kit (Qiagen, Germany), seguindo as instruções do fabricante. Para verificar a presença/ausência de contaminantes, foi medida a absorbância das amostras de RNA extraído, usando espectrofotômetro NanoDropTM Spectrophotometer ND-1000 (Thermo Fisher Scientific, New Hampshire, USA). Apenas as amostras com comprimento de onda entre 260 nm/280 nm no valor de 2.0 e com comprimento de onda entre 260 nm/230 nm no valor de 1.8 foram usadas para fazer a síntese de cDNA. A concentração de RNA total extraído foi quantificada com o fluorômetro Qubit[®] 2.0 (Invitrogen), utilizando o kit Qubit[®] RNA HS Assay, seguindo as instruções do fabricante. Somente as amostras com concentração maior do que 300 ng/ μl foram selecionadas para a síntese do cDNA. Também, verificou-se a integridade das amostras de RNA extraídas, por meio de eletroforese em gel de agarose desnaturante 1,2%, em que amostras degradadas foram excluídas.

O cDNA foi sintetizado com 1 μg de RNA total em 22 μl de volume final, utilizando o kit Impron II reverse transcriptase (Promega, Madison, WI, USA), de acordo com as instruções do fabricante. As amostras foram armazenadas a -20°C . A eficiência da síntese de cDNA foi avaliada por PCR, utilizando *os primers* endógenos beta-tubulina (*β -tubulina*), citocromo c oxidase subunidade III (*CytIII*), de *H. vastatrix*. Cada reação de RT-PCR foi constituída de uma desnaturação a 94°C por 2 minutos, seguida de 35 ciclos de 94°C por 1 minuto, 62°C por 1 minuto e 72°C por 1 minuto, seguida da extensão final a

72°C por 5 minutos. A integridade dos amplicons foi visualizada em gel de agarose desnaturante 1,2%.

2.5 Predição do secretoma de *H. vastatrix*

A predição do secretoma foi realizada, com base nas 13.034 proteínas preditas do genoma de *H. vastatrix* raça XXXIII, dados obtidos do genoma de referência (Capítulo 2). Inicialmente essas proteínas foram analisadas quanto à presença de peptídeo sinal por meio do algoritmo SignalP v4.0 (PETERSEN et al., 2011). As proteínas com D-Score = Y (SignalP) e com ponto de corte igual a 0,45 foram selecionadas para serem analisadas quanto à localização do peptídeo sinal, por meio do algoritmo TargetP v1.0 (EMANUELSSON et al., 2000). As proteínas que apresentaram o peptídeo sinal localizado na via de secreção foram selecionadas e analisadas quanto à presença de domínios transmembrânicos (TM), utilizando o algoritmo TMHMM v.2.0 (KROGH et al., 2001). As proteínas que apresentaram TM = 0 foram selecionadas.

2.6 Anotação das proteínas secretadas de *H. vastatrix* com características de proteínas efectoras

Para a anotação funcional das proteínas secretadas, primeiramente, elas foram comparadas com as sequências de proteínas não redundantes (nr) depositadas nos bancos NCBI (*National Center for Biotechnology Information*) e UNIPROT (*Universal Protein Resource*), por meio do algoritmo BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL et al., 1990). O BLASTp ($e\text{-value} = 10^{-5}$) foi executado em servidores, por linha de comando, por meio de *scripts* desenhados em linguagem *Perl*, buscando por similaridade entre essas proteínas. As sequências de proteínas que apresentaram similaridade

significativa com as sequências de proteínas presentes nos bancos de dados foram analisadas quanto sua categorização funcional (processo biológico, função molecular e componente celular) por meio do *software Blast2GO* (CONESA et al., 2005). As sequências de proteínas de *H. vastatrix* que não apresentaram *hit* foram comparadas por meio de BLASTp ($e\text{-value} = 10^{-5}$) e, também, por BLASTx ($e\text{-value} = 10^{-10}$), da mesma maneira descrita acima, com proteínas secretadas de *H. vastatrix* descritas por Fernandez et al. (2012) e Talhinhos et al. (2014). Essas sequências foram comparadas, também, com os dados de RNA-Seq obtidos por Lopes (2015). Esses dados de RNA-Seq foram gerados por meio do sequenciamento das bibliotecas da interação. Essas bibliotecas foram criadas, utilizando plantas de *C. arabica* cv. Caturra Vermelho (CIFC 19/1) e Híbrido de Timor (CIFC 832/1), inoculadas com esporos frescos de *H. vastatrix* raça XXXIII para estabelecer uma interação compatível e incompatível, respectivamente. Para realizar a comparação, neste trabalho, utilizaram-se, apenas, as bibliotecas geradas pela interação incompatível nos tempos de 12 e 24 horas após a inoculação.

Para obter a similaridade entre as categorias funcionais do KOG (EuKaryotic Orthologous Groups), banco de dados que possui um grupo de genes ortólogos de eucariotos, foi realizado psi-BLAST reverso. Além disso, foi também realizada uma busca de similaridade contra o banco de dados *PHI* (*Pathogen-Host Interaction*) que contém proteínas relacionadas à virulência/patogenicidade. Para a realização de cada uma dessas buscas, contra um banco de dados, foram considerados somente os melhores *hits*.

Para identificar as proteínas secretadas exclusivas de *H. vastatrix*, o secretoma de outras quatro ferrugens (*P. graminis* f. sp. *tritici*, *P. striiformis* f. sp. *tritici*, *M. larici-populina* e *M. lini*) foi predito usando o mesmo *pipeline* descrito no item 2.4. Os secretomas foram, então, agrupados, baseando-se na similaridade de suas sequências. As sequências que apresentaram identidade

maior que 70% foram reunidas em um mesmo grupo usando o CD-HIT (LI; GODZIK, 2006).

Em cada proteína identificada como secretada, procurou-se por resíduos de cisteína e motivos de efetores, como exemplo, [YWF]xC, CxxC e CxxxC, usando *scripts* em *Perl*, executados por linha de comando em Linux.

2.7 Desenho de *primers* para RT-qPCR e teste de eficiência

Primers específicos foram desenhados, para sequência de genes que codificam proteínas candidatas a efetores de *H. vastatrix*, utilizando o programa Primer Express 2.0 (Applied Biosystem, USA) (Tabela 1). Para o desenho, foram utilizados os seguintes parâmetros: tamanho do *primer* entre 18 a 23 pb, tamanho do *amplicon* entre 70 a 150 pb, conteúdo de GC de 40-60% com ótimo em 50% e temperatura de anelamento variando entre 58°C e 62°C. Os genes beta-tubulina (*β-tubulina*), citocromo c oxidase subunidade III (*CytIII*), proteína ribossomal 40S e gliceraldeído-3-fosfato desidrogenase (*GADPH*), validados para análises de expressão de genes de *H. vastatrix* em cafeeiro (VIEIRA et al., 2011) foram testados como genes endógenos de referência. Primeiramente, foi realizada uma RT-PCR utilizando cDNA de fungo, cDNA de café e água como controle negativo da reação. Cada reação de RT-PCR foi constituída como descrita no item 2.3, contendo uma concentração de 10X de tampão, 50mM de MgCl₂, 10mM de dNTPS, 2μM de cada *primer* e 25 ng/μl de amostra em um volume final de 20μl. A integridade dos amplicons foi visualizada em gel de agarose desnaturante 1,2%.

A eficiência tanto dos genes candidatos quanto a dos genes endógenos foi analisada, por meio do desenvolvimento de uma curva padrão com cinco pontos de diluição do cDNA, seguindo um fator de diluição seriada 1:5. Com

base em cada *Ct*, obtido para cada ponto de diluição, juntamente com o valor do *slope*, foi possível calcular o valor da eficiência para cada um dos genes.

Tabela 1 Sequência dos *primers* desenhados no programa Primer Express 2.0 para análise de qRT-PCR

Nome do <i>primer</i>	Sequência do <i>primer</i>	<i>Amplicon</i> (pb)
<i>EHv33_1F</i>	5'CCCTTTGGCTCAAATGAAAA 3'	116
<i>EHv33_1R</i>	5' GTGGATTTAGCCAAGCAAGC 3'	116
<i>EHv33_2F</i>	5'ATCCCTTGTCGCATCATCTC 3'	124
<i>EHv33_2R</i>	5' ATGTGTAAGGCGGGCTAGAA 3'	124
<i>EHv33_3F</i>	5' CCTCGGCTCATTTTCATGACT 3'	100
<i>EHv33_3R</i>	5' TGAGTCTGCAAGGAGTACTGA 3'	100
<i>EHv33_4F</i>	5' TGGTGCATCATTGATTTTGG 3'	108
<i>EHv33_4R</i>	5' GACAGGATTCCAAACCAGGA 3'	108
<i>EHv33_5F</i>	5'TCAAGGCCAAAACACTTTCTG 3'	85
<i>EHv33_5R</i>	5' CAGCAAGGGAAAACCTTCAGC 3'	85
<i>EHv33_6F</i>	5' GCTGGTTTATGGGGCATATT 3'	119
<i>EHv33_6R</i>	5' GAGAATTGGCCATTTCGAAGA 3'	119
<i>EHv33_7F</i>	5' TGTTGCACTATGTTGCTTTGG 3'	93
<i>EHv33_7R</i>	5' TCCTCATTCTTTGGAAGATGG 3'	93
<i>EHv33_8F</i>	5' GAAAATGCTGCCAAGGATGT 3'	121
<i>EHv33_8R</i>	5' GCAACTTCTGGTCCCCATTA 3'	121
<i>EHv33_9F</i>	5' TCAACGCACCCTTGGTATCT 3'	102
<i>EHv33_9R</i>	5' GGCACACCCAATTTTCAAG 3'	102
<i>EHv33_10F</i>	5 TGCGGTTTCAGTCACTTTGG'3'	70
<i>EHv33_10R</i>	5' AGGCTTCATCTTATCCGTCGT 3'	70
<i>EHv33_11F</i>	5' TCCTGAGAGTTCAGTCGCAA 3'	98
<i>EHv33_11R</i>	5' TGACGGGTTTGCTTGAACAG 3'	98
<i>EHv33_12F</i>	5' ATTCTGACTTCGACAGCGA 3'	146
<i>EHv33_12R</i>	5' CCACCTGGAGGCATAATTGCTA 3'	146
<i>EHv33_13F</i>	5'AGCCCAAATACGGCTGATGA 3	96
<i>EHv33_13R</i>	5' ATACACGGGATTTCCACGCA 3'	96

“Tabela 1, conclusão”

Nome do primer	Sequência do primer	Amplicon (pb)
<i>EHv33_14F</i>	5'CGGGCAGATATACGTGAAAG 3'	165
<i>EHv33_14R</i>	5'GTCCCAGATGCTTCTTGATT 3'	165
<i>EHv33_15F</i>	5' GATTATCTGGTTCCCCAGCA 3'	186
<i>EHv33_15R</i>	5' CTACCTGGATTGGGCTTTCA 3'	186
<i>EHv33_16F</i>	5' ACCTCAGGCTGTTGCAGAAT 3'	183
<i>EHv33_16R</i>	5' AAATTCATGGTGCCACACAA 3'	183
<i>EHv33_17F</i>	5' GCCTTTGGCGTAGAATGT 3'	154
<i>EHv33_17R</i>	5'CTTCCGTCACCATCAGTATC 3'	154

(*EHv33*) genes candidatos a efetores de *Hemileia vastatrix* raça XXXIII; (F) sequência do primer forward; (R) sequência do primer reverse.

2.8 PCR em tempo real (RT-qPCR)

As análises da expressão gênica quantitativa por Reação em Cadeia da Polimerase em Tempo Real (RT-qPCR) foram conduzidas utilizando o equipamento ABI PRISM 7500 Real Time PCR Systems (Applied Biosystems), usando o sistema SYBR® Green para a detecção dos produtos gerados. A reação foi realizada em volume final de 10 µL, contendo 50ng de cDNA, 10 µM de cada primer, 5 µL de solução do kit Power SYBR® Green PCR Master Mix (Applied Biosystems) e 2,2 µL de água. As reações foram submetidas a 95°C por 10 minutos para a desnaturação inicial, seguido de 40 ciclos de 94°C por 15 segundos e 60°C por 1 minuto. A especificidade das reações foi avaliada, por meio da curva de dissociação, obtida com gradiente variando de 60 a 95°C, com a taxa de incremento de 1°C a cada 30 segundos.

Cada reação foi realizada em três réplicas técnicas, para cada uma das três repetições biológicas existentes para cada amostra (plantas resistentes e plantas suscetíveis), nos quatro diferentes tempos (12h.a.i, 24h.a.i, 48h.a.i e 72h.a.i). Os dados gerados por meio da expressão relativa foram analisados pela

metodologia proposta por Pfaffl et al. (2001). Esse método determina se existe uma diferença significativa entre as amostras e o controle, considerando a eficiência real da reação para cada *primer*, e a normalização dos genes de referência é calculada com base em uma média geométrica pelas concentrações de cada *primer* endógeno.

3 RESULTADOS

Considerando 13.034 proteínas, preditas no genoma de *H. vastatrix* raça XXXIII, um total de 615 proteínas secretadas foram encontradas utilizando um *pipeline* formado pelos programas SignalP, TargetP e TMHMM. Essas proteínas secretadas apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$), principalmente, com proteínas de *P. graminis* f. sp. *tritici* (385) e *M. larici-populina* (109). Somente 39 proteínas apresentaram similaridade com sequências de *H. vastatrix* e 72 não apresentaram *hit* contra o banco de dados (*no hits*) (Figura 1).

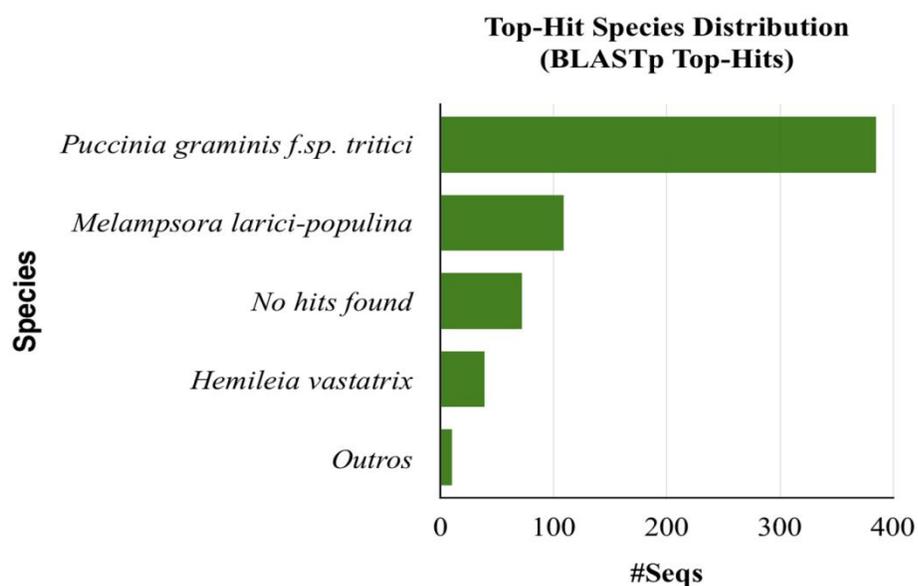


Figura 1 Resultado da busca por similaridade entre as sequências de proteínas secretadas presentes no genoma de *H. vastatrix* raça XXXIII contra os bancos de dados públicos de sequências de proteínas não redundantes (NCBI e UNIPROT), por meio do algoritmo BLASTp.

Legenda: No eixo y estão as espécies encontradas com maior frequência na busca por similaridade. No eixo x estão os números de proteínas encontradas para cada espécie.

As proteínas que apresentaram similaridade com as sequências de proteínas, depositadas nos bancos de dados do NCBI/UNIPROT, no total 543, foram analisadas quanto à categorização funcional, por meio da plataforma *BLAST2GO*. Destas, 320 proteínas foram distribuídas em três categorias (“Componente Celular”, “Função Molecular” e “Processos Biológicos”) em diferentes níveis hierárquicos do *GO*. Algumas proteínas foram anotadas em duas ou mais categorias, sendo a “Função Molecular” a categoria mais representada com 161 anotações, seguida da categoria “Processo Biológico” com 110 anotações e, por último, a categoria menos representada, “Componente Celular”, com 49 anotações.

Considerando a categoria “Função Molecular” (nível hierárquico 03), os termos “atividade transferase”, “ligação a íons” e “atividade hidrolase” foram os mais representados com 17, 15 e 11 proteínas, respectivamente (Figura 2).

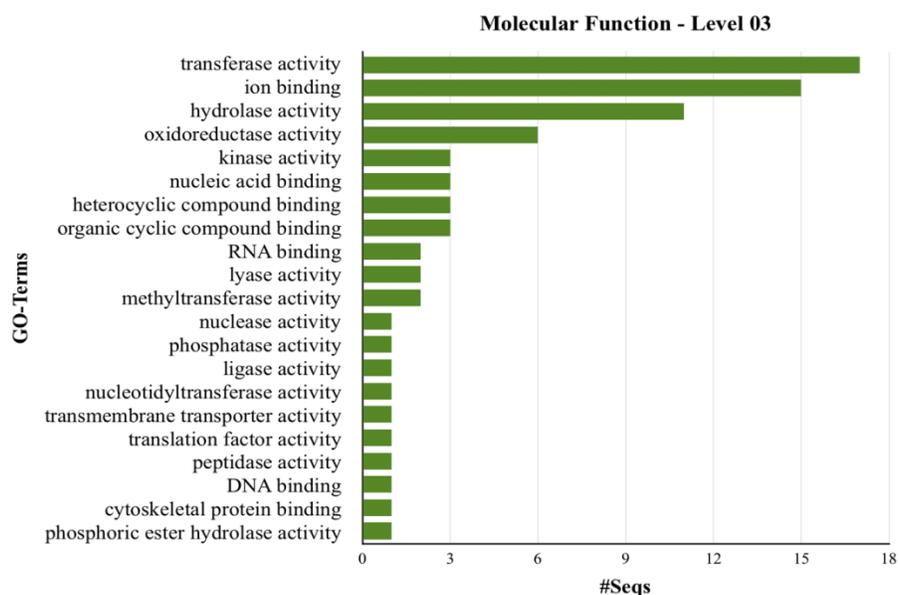


Figura 2 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO

Legenda: No eixo y estão os termos do GO descritos para a categoria Função Molecular, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria.

Para a categoria “Processo Biológico”, também, levando em consideração o nível hierárquico 03, os termos mais representados foram: “processo metabólico celular”, “processo metabólico de substâncias orgânicas”, “processo metabólico primário” e “estabelecimento de localização” com 13, 12, 12 e 10 sequências de proteínas, respectivamente (Figura 3).

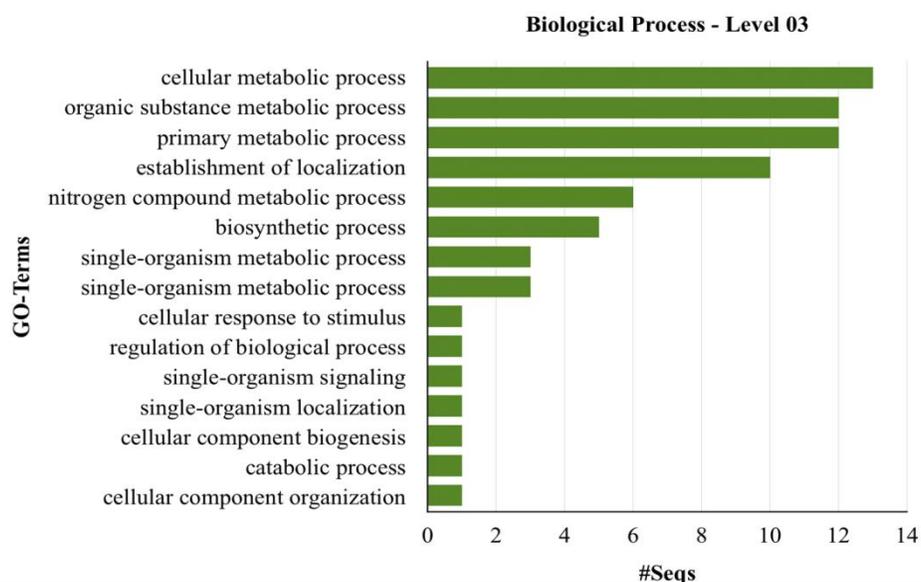


Figura 3 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software* BLAST2GO

Legenda: No eixo y estão os termos do GO descritos para a categoria Processo Biológico, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria.

Na categoria “Componente Celular”, os termos mais representados, no nível hierárquico 03, foram “parte celular”, “parte intracelular” e “citoplasma” com 24, 22 e 13 proteínas, respectivamente (Figura 4).

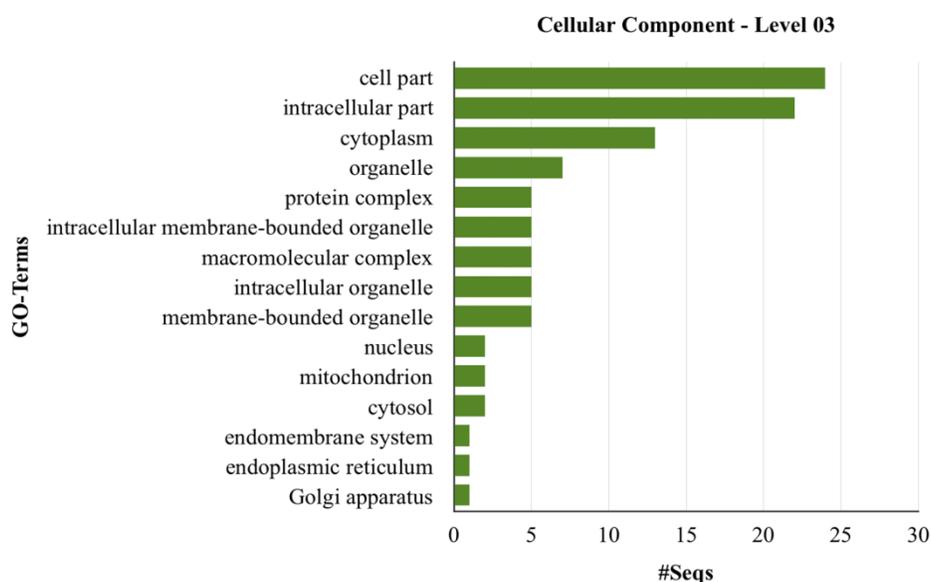


Figura 4 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$) com as sequências de proteínas depositadas nos bancos de dados (NCBI/UNIPROT), utilizando o *software BLAST2GO*

Legenda: No eixo y estão os termos do GO descritos para a categoria Componente Celular, no nível hierárquico 03. No eixo x estão os números de proteínas encontradas para cada termo do GO nessa categoria.

Cerca de 27% das 615 proteínas secretadas apresentaram assinaturas específicas nas categorias do KOG, sendo “replicação, recombinação e reparo”, “transcrição” e “transporte de amino ácido e metabolismo” as mais representadas (Figura 5).

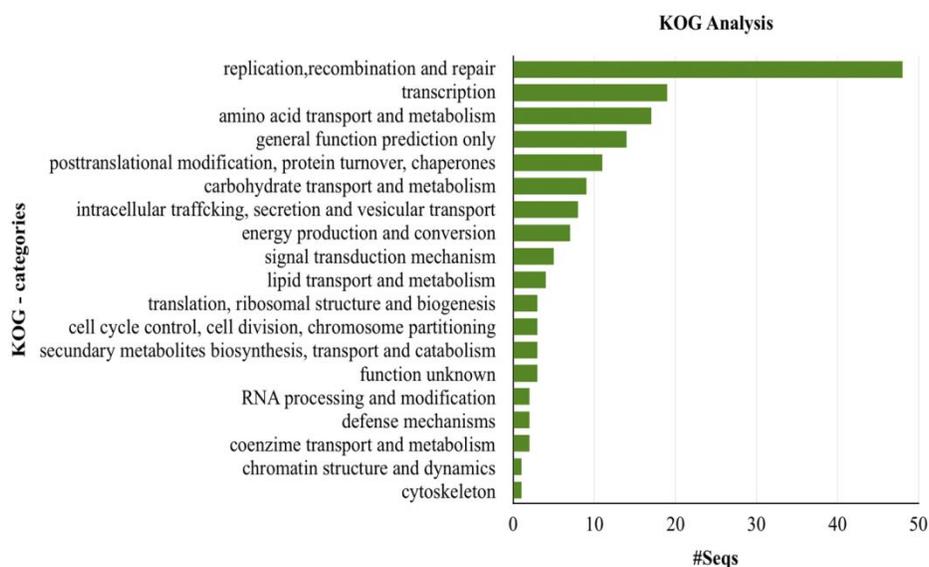


Figura 5 Resultado da análise de categorização funcional das proteínas secretadas encontradas no genoma de *H. vastatrix* raça XXXIII, que apresentaram assinaturas específicas nas categorias do KOG

Legenda: No eixo y estão as categorias do KGO. No eixo x estão os números de proteínas encontradas para cada categoria do KGO.

Observou-se, também, que dessas 615 proteínas secretadas, apenas 69 apresentaram *hit* com as proteínas de fungos relacionadas à patogenicidade, depositadas no banco de dados *PHI* e apenas 11 mostraram alta similaridade (> 50%).

As análises combinadas do secretoma das cinco ferrugens (*H. vastatrix*, *P. graminis* f. sp. *tritici*, *P. striiformis* f. sp. *tritici*, *M. larici-populina* e *M. lini*) resultou na formação de 15 grupos. *H. vastatrix* apresentou apenas 1 proteína compartilhada com *M. larici-populina* e com *P. graminis* f.sp.*tritici* e 270 exclusivas (Figura 6).

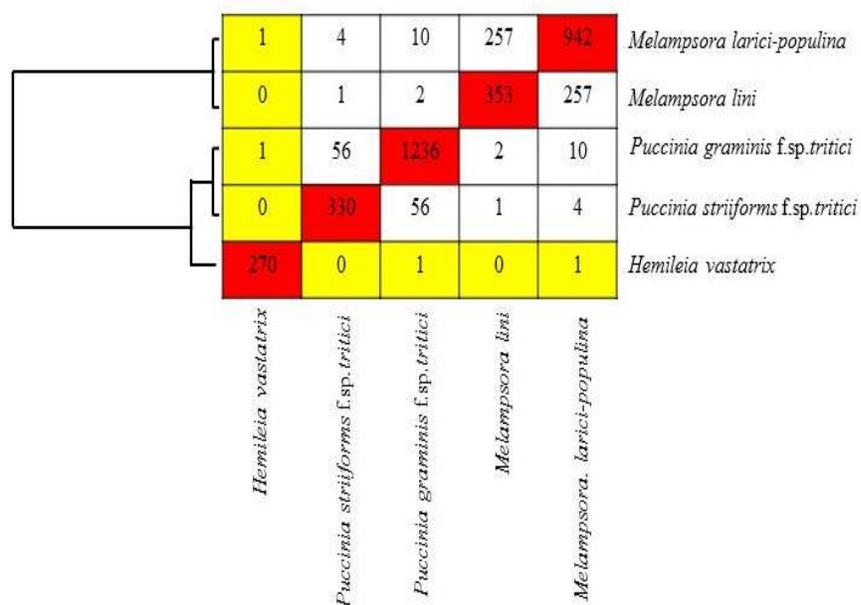


Figura 6 *Heatmap* resultante da análise do CD-HIT. Vermelho = proteínas secretadas exclusivas de cada fungo

Legenda: Amarelo = proteínas secretadas de *H. vastatrix* raça XXXIII compartilhadas com as outras ferrugens.

Cerca de 93% das proteínas secretadas de *H. vastatrix* raça XXXIII apresentaram resíduos de cisteína. A maioria com dois, três ou quatro resíduos presentes nas sequências. E 365 apresentaram os domínios [YWF]xC, CxxC e CxxxC em suas sequências, podendo ser encontrados sozinhos ou agrupados.

A caracterização do perfil de expressão de alguns genes de *H. vastatrix* raça XXXIII (*EHv33*), que codificam possíveis proteínas secretadas candidatas a efetores, foi realizada por meio de PCR em tempo real, durante a interação com o cafeeiro. Para selecioná-los, utilizaram -se duas estratégias. Primeiro, buscou-se identificar proteínas secretadas presentes nos *scaffolds* montados no genoma de referência de *H. vastatrix*. Dos 118.162 *scaffolds* obtidos nessa montagem,

193 codificam para 22 diferentes proteínas de *H. vastatrix* e 72 são proteínas secretadas (Tabela 2).

Tabela 2 Resultado da busca por proteínas nos *scaffolds* de *Hemileia vastatrix* raça XXXIII

Proteínas	Números de <i>scaffolds</i>
Putative secreted protein	72
Endoglucanase	4
Copper amine oxidase	1
Expansin	9
Macrophage activating glycoprotein	4
Aspartyl protease	3
Riboflavin aldehyde-forming	3
Fasciclin	1
DisulfideIsomerase	5
Cytochrome C peroxidase	11
Rust transferred	1
Mannitol dehydrogenase	2
Invertase	3
Hexose transporter	2
MapK	21
G Protein	6
Chitin deacetylase	4
Asparate aminotransferase	6
Amino acid transporter	10
Elongation Factor	6
Beta-tubulin	12
Cytochrome b	7

O próximo passo foi realizar o BLASTx contra o banco de dados não-redundante do NCBI e do UNIPROT, com os 72 *scaffolds* que codificam potenciais proteínas secretadas. Encontraram-se 18 potenciais proteínas secretadas de *H. vastatrix* e 11 apresentaram domínios conservados já descritos (Tabela 3).

Tabela 3 Domínios conservados presentes nas 11 proteínas de *Hemileia vastatrix* raça XXXIII, codificadas pelos 72 *scaffolds* presentes no genoma montado

Domínios conservados	Números de domínios conservados
GPI-anchored super family	3
Glyco_hydro_53 Super family	1
CFEM super family	2
Mucin super family	1
PLN03209	1
TDT super family/Noil super family	1
ManB super family	1
Oxidored_q5_N	1
Herpes_capsid super family	1
PRKI4950	1
BOPINT super family	1

A fim de confirmar se essas 18 proteínas encontradas são proteínas secretadas, analisou-se a presença de peptídeo sinal, a localização do peptídeo sinal e a presença de domínios transmembrânicos por meio dos algoritmos SignalP, TargetP e TMHMM, respectivamente. Com base na presença do peptídeo sinal, localizado na via de secreção, na ausência de domínios transmembrânicos e na presença dos domínios conservados, foram selecionados e desenhados 4 *primers* para caracterizar o perfil de expressão de possíveis genes que codificam proteínas candidatas a efetores.

A segunda estratégia partiu das 615 proteínas secretadas encontradas no proteoma predito, com base no genoma montado de *H. vastatrix* raça XXXIII. Foram, então, selecionadas as 72 proteínas que não apresentaram *hits* com os bancos de dados (*no hits*) e as 39 que apresentaram *hit* com proteínas de *H. vastatrix*. Essas 111 proteínas foram comparadas com os bancos de dados de proteínas secretadas de *H. vastatrix* descritos na literatura (FERNANDEZ et al., 2012; TALINHAS et al., 2014) e com as bibliotecas de RNA-Seq da interação cafeeiro – *H. vastatrix* construídas por Lopes (2015). Para as 72 *no hits*, observou-se que não foi encontrada similaridade com as sequências descritas por Fernandez et al. (2012) e por Talhinhos et al. (2014). Entretanto com as sequências das bibliotecas de RNA-seq, encontraram-se 23 *hits* com a biblioteca compatível 12 horas, após inoculação e 7 *hits* eram exclusivos dessa biblioteca. Foram identificados, também, 21 *hits* com a biblioteca compatível 24 horas, após inoculação, com 4 *hits* exclusivos dessa biblioteca.

Na análise das 39 proteínas que apresentaram *hits* com *H. vastatrix*, 35 foram anotadas como candidatas a efetores, no entanto, nenhuma apresentou *hit* com as sequências descritas por Fernandez et al. (2012). Nas demais comparações, 34 proteínas apresentaram similaridade com as sequências descritas por Talhinhos et al. (2014) e 20 *hits* com ambas as bibliotecas de RNA-seq.

Com base nos resultados dessas análises, foram selecionados e desenhados 13 *primers* de possíveis genes que codificam proteínas candidatas a efetores para serem caracterizados por PCR em tempo real (Tabela 4).

Tabela 4 *Primers* escolhidos com a sequência de 13 genes que codificam possíveis proteínas candidatas a efetores de *H. vastatrix*, com base nos resultados das comparações entre os diferentes bancos de dados utilizados

Nome primer	Comparação com banco de dados				
	NCB	Fernandez <i>et al.</i> (2012)	Talhinhas <i>et al.</i> (2014)	RNA-Seq 12hai	RNA-Seq 24hai
EHv33_1	No hit ¹	No hit ¹	No hit ¹	No hit ¹	No hit ¹
EHv33_2	No hit ¹	No hit ¹	No hit ¹	No hit ¹	No hit ¹
EHv33_3	No hit ¹	No hit ¹	No hit ¹	No hit ¹	No hit ¹
EHv33_4	No hit ¹	No hit ¹	No hit ¹	Hit ¹	Hit ¹
EHv33_5	No hit ¹	No hit ¹	No hit ¹	Hit ¹	Hit ¹
EHv33_6	No hit ¹	No hit ¹	No hit ¹	Hit ¹	No hit ¹
EHv33_7	No hit ¹	No hit ¹	No hit ¹	Hit ¹	No hit ¹
EHv33_8	No hit ¹	No hit ¹	No hit ¹	Hit ¹	No hit ¹
EHv33_9	No hit ¹	No hit ¹	No hit ¹	No hit ¹	Hit ¹
EHv33_10	No hit ²	No hit ²	No hit ²	Hit ²	No hit ²
EHv33_11	No hit ²	No hit ²	No hit ²	Hit ²	No hit ²
EHv33_12	No hit ²	No hit ²	No hit ²	No hit ²	Hit ²
EHv33_13	No hit ²	No hit ²	No hit ²	No hit ²	Hit ²

Legenda: **No hit¹** = 72 proteínas *no hits* não apresentaram *hits* com nenhuma proteína depositada nos diferentes bancos de dados.

No hit¹ / Hit¹ = 72 proteínas *no hits* que apresentaram *hits* com proteínas depositadas nas bibliotecas de RNA-seq de 12 h.a.i e 24 h.a.i.

No hit¹ / Hit¹ = 72 proteínas *no hits* que apresentaram *hits* com proteínas depositadas somente na biblioteca de RNA-seq de 12 h.a.i

No hit¹ / Hit¹ = 72 proteínas *no hits* que apresentaram *hits* com proteínas depositadas somente na biblioteca de RNA-seq de 24 h.a.i

No hit² / Hit² = 39 proteínas que deram *hit* com *H. vastatrix* e apresentaram, também, *hits* com proteínas depositadas somente na biblioteca de RNA-seq de 12 h.a.i

No hit² / Hit² = 39 proteínas que deram *hit* com *H. vastatrix* e apresentaram, também, *hits* com proteínas depositadas somente na biblioteca de RNA-seq de 24 h.a.i.

Com as estratégias utilizadas, foi selecionado um total de 17 genes que codificam potenciais efetores para serem caracterizados por RT-qPCR. Esses

genes foram comparados com o genoma parcial de *H. vastatrix* recentemente publicado (CRISTANCHO et al., 2014). Encontram -se 100% de similaridade para as sequências de dois gene, *EHv33_10* e *EHv33_11*.

Antes de iniciar a validação dos potenciais candidatos, quatro genes endógenos foram avaliados: beta-tubulina (*β -tubulina*), citocromo c oxidase subunidade III (*CytIII*), proteína ribossomal 40S e gliceraldeído-3-fosfato desidrogenase (*GADPH*). Desses, apenas os genes *β -tubulina* e *CytIII* foram específicos para as amostras do fungo, não amplificando amostras de café. Logo, esses dois genes foram escolhidos, para realizar os ensaios de RT-qPCR, nas interações incompatíveis e compatíveis, nos tempos 12, 24, 48 e 72 h.a.i. O tempo 12 h.a.i foi utilizado como amostra de referência.

A análise de expressão dos 17 genes *EHv33*, que codificam proteínas secretadas candidatas a efetores de *H. vastatrix*, pode ser observada nas Figuras 7, 8, 9 e 10. Pôde-se observar que 6 genes foram mais expressos, significativamente, durante a interação incompatível (Figura 7A-F), 9 durante a interação compatível (Figura 8A-D e Figura 9 A-E) e dois genes não apresentaram diferença significativa de expressão durante as diferentes interações (Figura 10).

Na interação incompatível, observou-se um pico de expressão as 24 h.a.i para os genes *EHv33_15*, *EHv33_1*, *EHv33_17*, *EHv33_13* e *EHv33_8* (Figura 7A-E). O gene *EHv33_12* mostrou um aumento da expressão em 48 h.a.i., tendo maior nível de expressão as 72 h.a.i, apresentando uma expressão tardia (Figura 7F).

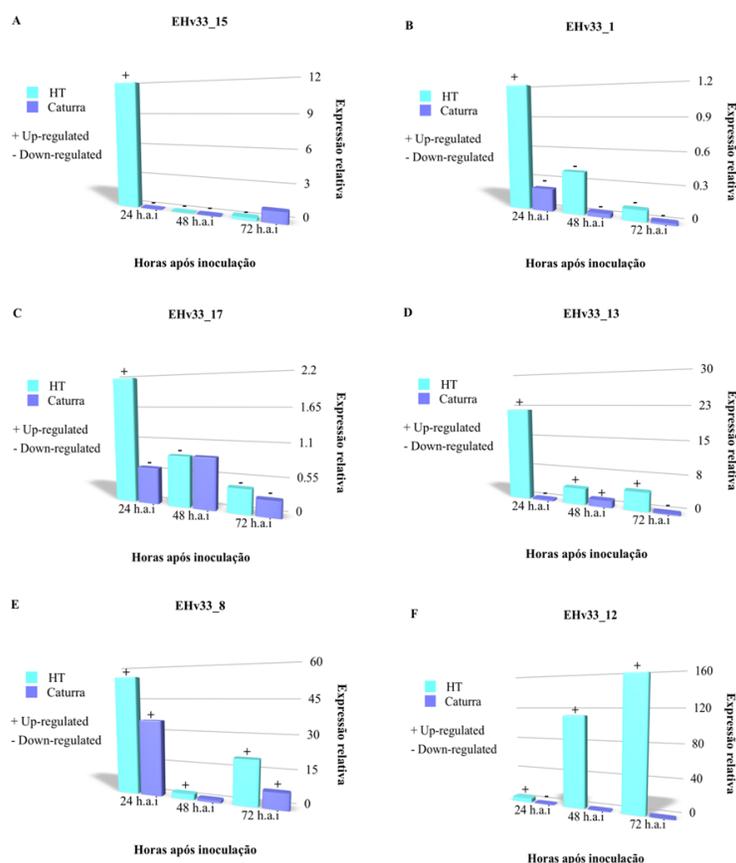


Figura 7 Análise de expressão por PCR em tempo real (qRT-PCR) de seis genes *EHv33*, que codificam possíveis proteínas secretadas candidatas a efetores de *H. vastatrix* raça XXXIII

Legenda: O padrão de expressão relativa dos genes alvo foi acessado em plantas de Híbrido de Timor (HT) e plantas de Caturra nos tempos 12 horas (12 h.a.i), 24 horas (24 h.a.i), 48 horas (48 h.a.i) e 72 horas (72 h.a.i) após a inoculação com urediniosporos da raça XXXIII de *H. vastatrix*, sendo a amostra 12 h.a.i utilizada como amostra de referência. O nível de expressão dos genes alvo foi normalizado com dois genes endógenos de *H. vastatrix*: *β-tubulina* e *CytIII*. (A) *EHv33_15* com maior nível de expressão em 24 h.a.i; (B) *EHv33_1* com maior nível de expressão em 24 h.a.i, decaindo ao longo do tempo; (C) *EHv33_17* com maior nível de expressão em 24 h.a.i, decaindo ao longo do tempo; (D) *EHv33_13* com maior nível de expressão em 24 h.a.i, decaindo as 48 h.a.i e mantendo a expressão as 72 h.a.i; (E) *EHv33_8* com maior nível de expressão em 24 h.a.i, decaindo as 48 h.a.i e tendo um aumento as 72 h.a.i; (F) *EHv33_12* aumento da expressão em 48 h.a.i., tendo maior nível de expressão as 72 h.a.i.

Os demais genes *EHv33* apresentaram maiores expressões na interação compatível. Os genes *EHv33_6*, *EHv33_5*, *EHv33_9* e *EHv33_11* apresentaram maior nível de expressão as 24 h.a.i (Figura 8A-D).

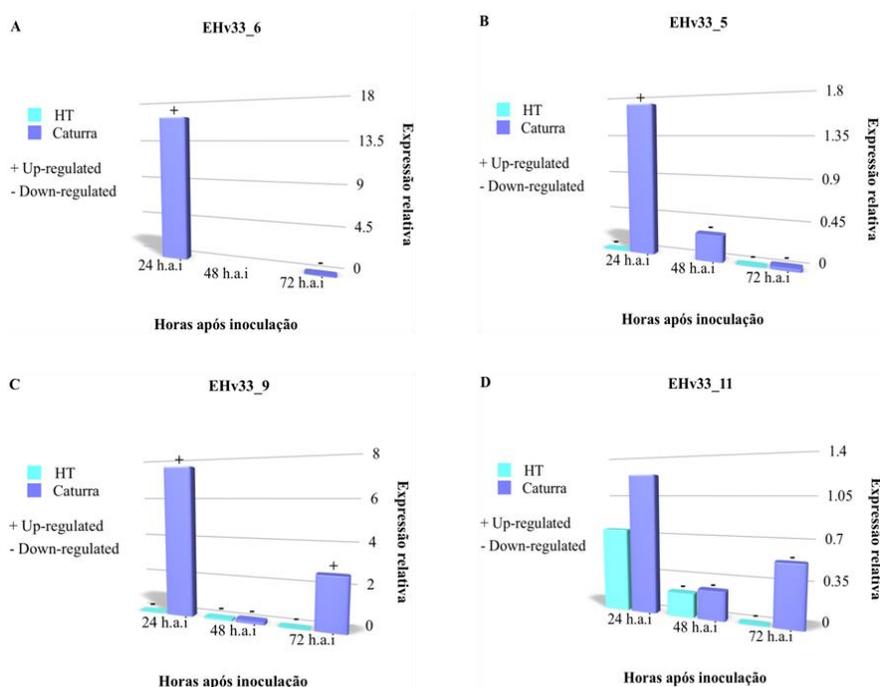


Figura 8 Análise de expressão por PCR em tempo real (qRT-PCR) de quatro genes *EHv33*, que codificam possíveis proteínas secretadas candidatas a efetores de *H. vastatrix* raça XXXIII

Legenda: O padrão de expressão relativa dos genes alvo foi acessado em plantas de Híbrido de Timor (HT) e plantas de Caturra nos tempos 12 horas (12 h.a.i), 24 horas (24 h.a.i), 48 horas (48 h.a.i) e 72 horas (72 h.a.i), após a inoculação com urediniósporos da raça XXXIII de *H. vastatrix*, sendo a amostra 12 h.a.i utilizada como amostra de referência. O nível de expressão dos genes alvo foi normalizado com dois genes endógenos de *H. vastatrix*: β -tubulina e *CytIII*. (A) *EHv33_6* com maior nível de expressão em 24 h.a.i; (B) *EHv33_5* com maior nível de expressão em 24 h.a.i, decaindo ao longo do tempo; (C) *EHv33_9* com maior nível de expressão em 24 h.a.i, decaindo às 48 h.a.i e aumentando a expressão às 72 h.a.i. (D) *EHv33_11* com maior nível de expressão em 24 h.a.i, decaindo às 48 h.a.i e aumentando a expressão as 72 h.a.i.

Já, os genes *EHv33_7*, *EHv33_2*, *EHv33_10* e *EHv33_4* apresentaram maior nível de expressão as 24 h.a.i, decaindo ao longo do tempo (Figura 9A-D). E o gene *EHv33_16* começou a apresentar níveis de expressão significativos com 24h.a.i e foi aumentando, até atingir um máximo as 72h.a.i. (Figura 9E).

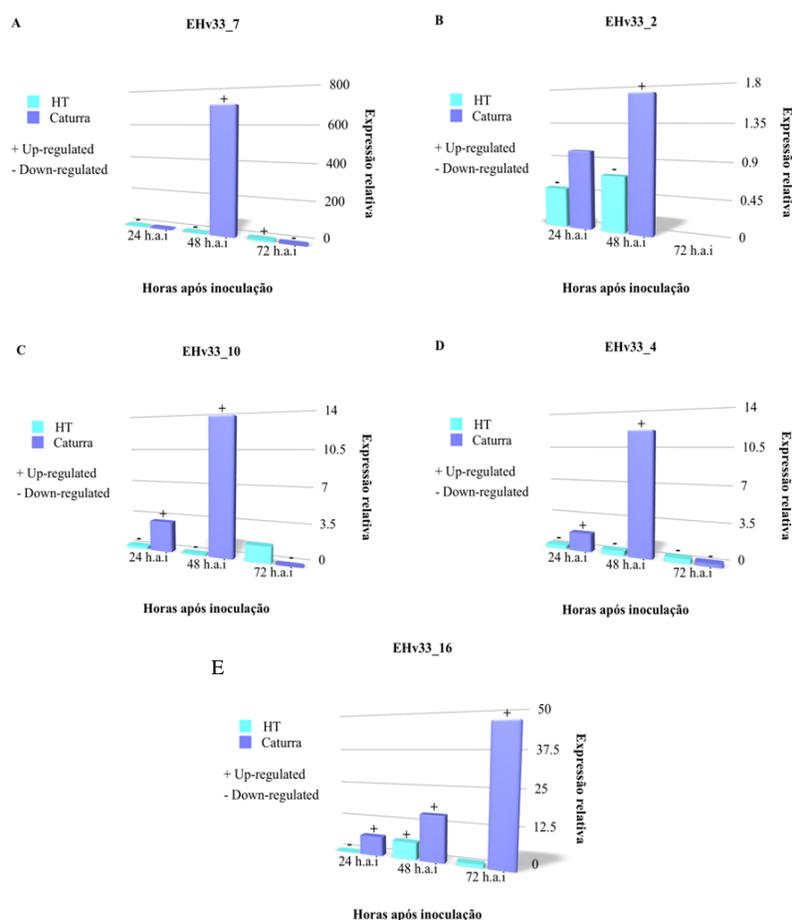


Figura 9 Análise de expressão por PCR em tempo real (qRT-PCR) de cinco genes *EHv33*, que codificam possíveis proteínas secretadas candidatas a efetores de *H. vastatrix* raça XXXIII

Legenda: O padrão de expressão relativa dos genes alvo foi acessado em plantas de Híbrido de Timor (HT) e plantas de Caturra nos tempos 12 horas (12 h.a.i), 24 horas (24 h.a.i), 48 horas (48 h.a.i) e 72 horas (72 h.a.i) após a inoculação com urediniósporos da raça XXXIII de *H. vastatrix* sendo a amostra 12 h.a.i utilizada como amostra de referência. O nível de expressão dos genes alvo foi normalizado com dois genes endógenos de *H. vastatrix*: β -tubulina e *CytIII*. (A) *EHv33_7* com maior nível de expressão em 48 h.a.i; (B) *EHv33_2* com maior nível de expressão em 48 h.a.i, decaindo ao longo do tempo; (C) *EHv33_10* com maior nível de expressão em 48 h.a.i, decaindo ao longo do tempo. (D) *EHv33_4* com maior nível de expressão em 48 h.a.i, decaindo ao longo do tempo (E) *EHv33_16* com maior nível de expressão em 72 h.a.i.

E os genes *EHv33_3* e *EHv33_14* não apresentaram diferença significativa de expressão durante as diferentes interações (Figura 10).

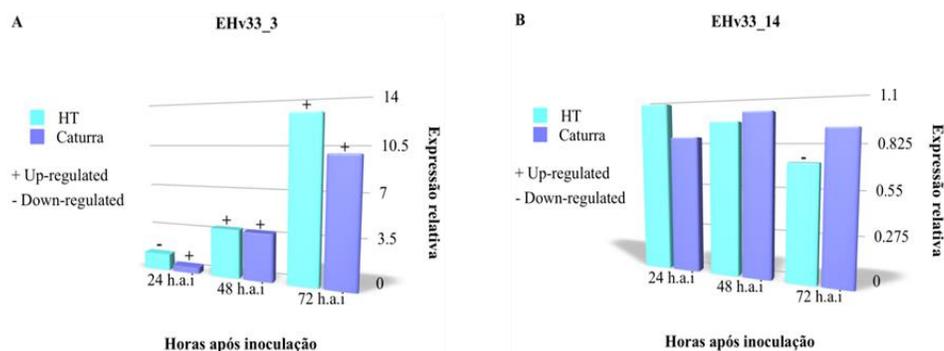


Figura 10 Análise de expressão por PCR em tempo real (qRT-PCR) de dois genes *EHv33*, que codificam possíveis proteínas secretadas candidatas a efetores de *H. vastatrix* raça XXXIII

Legenda: O padrão de expressão relativa dos genes alvo foi acessado em plantas de Híbrido de Timor (HT) e plantas de Caturra nos tempos 12 horas (12 h.a.i), 24 horas (24 h.a.i), 48 horas (48 h.a.i) e 72 horas (72 h.a.i), após a inoculação com urediniosporos da raça XXXIII de *H. vastatrix* sendo a amostra 12 h.a.i utilizada como amostra de referência. O nível de expressão dos genes alvo foi normalizado com dois genes endógenos de *H. vastatrix*: *β -tubulina* e *CytIII*. **(A)** *EHv33_3* aumenta o nível de expressão, ao longo do tempo, com maior nível em 72 h.a.i tanto na interação incompatível quanto na interação compatível. **(B)** *EHv33_14* não apresenta diferença significativa de expressão, ao longo do tempo, tanto na interação incompatível quanto na interação compatível.

4 DISCUSSÃO

Por meio das análises do proteoma de *H. vastatrix*, composto por 13.034 proteínas, foi possível identificar 615 proteínas secretadas. O número de proteínas secretadas identificado neste estudo é menor do que o encontrado em estudo recente realizado por Cristancho et al. (2014). Eles encontraram 775 proteínas secretadas, utilizando SignalP e 659 utilizando PProver, no sequenciamento parcial do genoma de *H. vastatrix*. O número maior de proteínas encontradas neste trabalho pode ser justificado pelo fato de não terem incluído análises com o programas TargetP e TMHMM. O proteoma estimado no presente trabalho foi maior que o número de proteínas efetoras de *H. vastatrix* encontradas nos estudos de Fernandez et al. (2012) e Talhinhos et al. (2014). Eles encontraram 382 e 516 proteínas secretadas, respectivamente. Nesses dois trabalhos, foram realizadas análises de sequências *ESTs* (*Expressed Sequence Tags*) de folhas de cafeeiro inoculadas com o patógeno 21 dias, após a inoculação (FERNANDEZ et al., 2012) e de *ESTs* (*Expressed Sequence Tags*), vindas de esporo germinado e apressório (TALHINHAS et al., 2014). O número inferior de proteínas encontradas pode ser justificado por terem analisado apenas as proteínas expressas nessas condições específicas e não no genoma completo como no presente trabalho.

Quando se compara com outras espécies de ferrugem, o número de proteínas secretadas por *H. vastatrix* preditas nesse estudo foi menor. No genoma de *M. larici-populina* foram encontradas 1.184 proteínas, em *P. graminis* f. sp. *tritici*, 1.106 (DUPLESSIS et al., 2011), em *P. striiformis* f.sp.*tritici* foram 2.999 (CANTU et al., 2013) e *M. lini*, 725 proteínas (NEMRI et al., 2014). Isso pode ocorrer pelo fato de que o genoma de *H. vastatrix*, ainda, não está totalmente sequenciado como o das outras ferrugens citadas. Os diferentes parâmetros utilizados com os algoritmo de previsão de secretoma (SignalP,

TargetP, e TMHMM), também, influenciam no resultado. Além disso, cada patossistema apresenta genes que são exclusivos nas suas interações. Portanto, as proteínas secretadas tendem a serem mais conservadas dentro das espécies e apresentarem funções ligadas à patogênese (DUPLESSIS et al., 2011).

Das 615 proteínas secretadas preditas no genoma de *H. vastatrix* raça XXXIII, a maior parte (385) apresentaram similaridade com proteínas de *P. graminis* f. sp. *tritici* e 109 com as proteínas de *M. larici-populina*. Todas essas ferrugens são consideradas fungos biotróficos, parasitas obrigatórios pertencentes ao filo *Basidiomycota* e à ordem *Pucciniales*. No entanto, *M. larici-populina* pertence à família *Melampsoraceae* e *P. graminis* f. sp. *tritici* e *H. vastatrix* são da família *Pucciniaceae*, explicando a maior similaridade encontrada entre essas duas espécies. Isso, também, pode ser explicado pelo número de sequências depositado nos bancos de dados públicos. Para *P. graminis* f. sp. *tritici* estão depositadas 65.082 sequências, enquanto para *M. larici-populina* estão depositadas um pouco mais da metade, 32.984 sequências. Essa alta similaridade, também, foi encontrada nos trabalhos realizados por Fernandez et al. (2012) e Talhinhos et al. (2014), com os transcritos dessa espécie.

Até o momento, *H. vastatrix* apresenta apenas 55 proteínas depositadas no *Genbank/NCBI*, o que explica a baixa similaridade encontrada para essa espécie, apenas 39 sequências. E 72 proteínas secretadas não apresentaram *hit* com nenhuma outra sequência de proteína depositada nos bancos de dados. Estas foram, portanto, consideradas proteínas exclusivas de *H. vastatrix* que, ainda, não foram caracterizadas e foram alvo de estudo neste trabalho.

Entre as proteínas anotadas com função conhecida, a maioria apresentou especificidade com “parte celular”, mostrando que os genes que codificam proteínas secretadas podem estar presentes em diferentes locais da célula (citoplasma, núcleo, mitocôndria). O mesmo foi observado no estudo funcional

do proteoma. Também, destacaram-se as proteínas com atividade de transporte de aminoácidos e envolvidas nos processos metabólicos e processos de replicação, recombinação e transcrição. Esses resultados corroboram com os estudos de Fernandez et al. (2012) e Talhinhos et al. (2014), os quais mostraram que apenas 3% das proteínas secretadas de *H. vastatrix* apresentaram anotação conhecida e, também, estão envolvidas em processos metabólicos, replicação, transcrição e tradução e no transporte de aminoácidos.

A maioria das proteínas efetoras de fungos patogênicos, caracterizada até o momento, é secretada pela rota clássica de secreção (DE WIT et al., 2009). Uma característica importante desses efetores é que suas sequências não apresentam similaridade com outras espécies relacionadas, ou seja, são espécie-específicas (ELLIS et al., 2009). Por esta razão, focou-se, neste estudo, na análise das 39 proteínas secretadas que apresentaram hit com proteínas de *H. vastatrix* e nas 72 proteínas (*no hits*) sem similaridade com as sequências proteicas depositadas nos bancos de dados públicos. Essas proteínas foram, então, classificadas como proteínas secretadas exclusivas de *H. vastatrix* e consideradas proteínas candidatas a efetores (*EHv33*).

Essas 111 proteínas secretadas candidatas a efetores de *H. vastatrix* (*EHv33*) foram anotadas no banco de dados de domínio Pfam e nenhuma função foi atribuída. Dessa forma, a ausência de domínios Pfam, para a maioria das *EHv33*, relatadas neste estudo, está de acordo com as características de efetores.

Segundo Stergiopoulos e De Wit (2009), proteínas efetoras de fungos, normalmente, são ricas em resíduos de cisteína. Esses resíduos dão estabilidade à proteína no meio extracelular, por meio da formação intramolecular de pontes dissulfeto, que são importantes, também, para sua estrutura e função. Mudanças nas sequências de aminoácidos podem ocorrer sem alterar a topologia da proteína (exceto se ocorrer nos resíduos de cisteínas), fazendo com que ela seja ideal para reconhecimento e especificidade (POVOLOTSKAYA;

KONDRASHOV, 2010). Assim, quanto maior é o número de resíduo de cisteína em uma proteína efetora, mais rápida é sua capacidade de mudança, favorecendo o surgimento de novas raças fisiológicas de fungos. Das 615 proteínas secretadas de *H. vastatrix* raça XXXIII, cerca de 93% apresentaram resíduos de cisteína. A maioria com dois, três ou quatro resíduos presentes nas sequências. Encontraram-se proteínas com até 16 resíduos, no entanto essas proteínas são consideradas muito grandes (> 150 aa) e não podem ser candidatas a efetores, uma vez que um efector é uma proteína pequena com até 150 aa presentes em sua sequência.

A genômica pode ser usada, para identificar os genes expressos pelo patógeno, durante o seu tempo de infecção. Além da presença do peptídeo sinal, as análises computacionais são capazes de identificar motivos responsáveis pelo tráfico de efetores para dentro da célula dos hospedeiros, podendo, assim, manipular o seu sistema de defesa (JONES; DANGL, 2006). Os efetores de fungos não apresentam um motivo consenso. Os motivos RXLR são característicos de efetores de oomicetos (DOU et al. 2008), já os motivos [YWF]xC, CxxC e CxxxC têm sido encontrados em famílias conservadas de efetores de patógenos pertencentes à ordem Puccinales (DUPLESSIS et al., 2011). No presente trabalho, encontraram-se esses domínios em 365 proteínas secretadas, confirmando que tais proteínas possuem funções de efetores.

Esses resultados demonstram que a estratégia adotada foi eficiente para identificar novos genes candidatos a efetores do fungo.

Segundo Spanu et al. (2010), a predição do secretoma, associada à análise de expressão temporal, permite a identificação de genes candidatos a efetores, em especial aqueles expressos durante o processo de colonização dos tecidos da planta. As análises de expressão relativa dos 17 *EHv33*, que codificam possíveis proteínas secretas, exclusivas de *H. vastatrix*, foram menores durante a interação incompatível. Observou-se maior número de *EHv33*, sendo expressos durante a interação compatível. Os dados sugerem que o

fungo, dependendo do genótipo envolvido no processo de infecção, escolhe quais efetores serão utilizados para infectar o hospedeiro e tentar suplantar o sistema de defesa da planta.

Na interação incompatível do presente estudo, onde se encontra o genótipo resistente de cafeeiro Híbrido de Timor e a raça XXXIII do fungo *H. vastatrix*, a maioria dos genes tiveram um pico de expressão no tempo 24 h.a.i., seguido de uma queda. Esse é o comportamento esperado de efetores nessa interação. Em genótipos resistentes, a planta reconhece rapidamente o patógeno e as respostas de defesa são ativadas para prevenir o crescimento e desenvolvimento do fungo no interior dos tecidos da planta. Nesse caso, a interação entre a planta resistente e o fungo inicia precocemente, antes mesmo da penetração. Muitos efetores secretados pelos esporos germinados, envolvidos nos processos de sinalização, estabelecimento e manutenção da biotrofia, tentam suplantar o sistema de defesa das plantas com a finalidade de suprimir a sua maquinaria de defesa celular (FERNANDEZ et al., 2012; TALHINHAS et al., 2014; VIEIRA et al., 2012). Esse tipo de resistência é conhecido como resistência pré-haustorial (GANESH et al., 2006 ; SILVA et al., 2008), a qual pôde ser observada na expressão do gene *EHv33_15*, *EHv33_1*, *EHv33_17*, *EHv33_13* e *EHv33_8*, os quais foram considerados genes candidatos a efetores pré- haustoriais.

Caso o fungo consiga vencer essa resistência pré-haustorial, a planta resistente busca retardar o crescimento do fungo nas primeiras fases do processo infeccioso, geralmente, após a formação do haustório primário que ocorre em torno de 24 h.a.i. Resistência conhecida como resistência pós-haustorial (SILVA et al., 2002, 2006), a qual pôde ser observada na expressão do gene *EHv33_12*.

Estudos citológicos sugerem que ambas as resistências pré e pós-haustorial estão associadas à reação de hipersensibilidade (HR), a qual se caracteriza por uma rápida e localizada morte celular das células do hospedeiro,

no local da infecção, impedindo, assim, o crescimento do patógeno nos tecidos da planta. As respostas de defesa pós-haustoriais, como, por exemplo, deposição de compostos fenólicos, lignificação da parede celular da planta, acúmulo de material intercelular com polissacarídeos e fenóis, entre outros, também, estão associados à HR (SILVA et al., 2002, 2006, 2008).

Observações citológicas, realizadas recentemente por Lopes (2015) indicaram que, ao contrário da resistência pós-haustorial que é geralmente descrita para as interações *H. vastatrix* – cafeeiro, a resistência do genótipo HDT832/1 (Híbrido de Timor - CIFC 832/1), o mesmo utilizado neste estudo, a raça XXXIII de *H. vastatrix* é pré-haustorial. O mesmo, também, foi observado por Diniz et al. (2012), que sugeriram que essa resposta rápida de defesa da planta, na qual impede a formação do haustório, pode ser a base para a durabilidade da resistência às raças desse fungo.

Durante o processo de infecção, *H. vastatrix* diferencia várias estruturas de infecção especializadas. Após a formação do apressório, desenvolve-se uma hifa de penetração que produz na sua extremidade duas ramificações laterais em forma de âncora, em que de cada ramo lateral forma-se a célula-mãe do haustório, a qual invade as células subsidiárias do estômato, formando o haustório primário. Nos genótipos suscetíveis, o fungo prossegue seu crescimento, com a formação de muitas hifas intercelulares e haustórios nas células do mesófilo, chegando até a epiderme da planta. Todas essas modificações, no desenvolvimento do fungo, ocorrem em torno de 48 a 72 h.a.i. (GANESH et al., 2006; VIEIRA et al., 2012). Isso explica o fato de ter sido observado neste estudo que os efetores que apresentaram maiores níveis de expressão, em genótipos suscetíveis, estão em 48 e 72 h.a.i. Dessa forma, pode-se dizer que os genes *EHv33_7*, *EHv33_2*, *EHv33_10*, *EHv33_4* e *EHv33_16* são candidatos a efetores translocados para o interior da célula hospedeira via haustório.

Embora alguns desses eventos estejam descritos no patossistema *H. vastatrix* – cafeeiro, dependendo do genótipo da planta e da raça do fungo envolvidos na interação, ocorrerão diferenças nas respostas de defesa da planta e no ataque do fungo.

Logo, os resultados obtidos neste trabalho, somados com a recente publicação do genoma parcial de *H. vastatrix* e outros dados que vêm sendo obtidos para esse patossistema, será possível ampliar os conhecimentos sobre os mecanismos moleculares envolvidos, durante o processo de interação, visando auxiliar os programas de melhoramento, além de ajudar na identificação e classificação dos isolados de *H. vastatrix* em diferentes raças.

5 CONCLUSÃO

Neste estudo, foram preditas 615 proteínas secretadas de *H. vastatrix*, com o proteoma anotado pelas análises realizadas, com base no sequenciamento estrutural do genoma desse isolado, fornecendo, portanto, uma informação mais abrangente das proteínas secretadas desse fungo.

A anotação funcional dessas proteínas secretadas revelou que o secretoma predito de *H. vastatrix* é constituído por um grande número de proteínas secretadas de função desconhecida. Foram selecionados 17 genes (*EHv33*) que codificam possíveis proteínas secretadas, para serem caracterizados por meio de PCR em tempo real. Suas análises de expressão revelaram uma expressão dinâmica, ao longo do tempo, associada com as diferentes interações.

Os genes *EHv33_6*, *EHv33_7*, *EHv33_10* e *EHv33_16*, expressos durante a interação compatível, serão priorizados para ensaios funcionais, para serem validados e confirmarem seus papéis durante o processo infeccioso. Resultado do fato deles se expressarem, com maior intensidade, após o desenvolvimento dos haustórios primários (24 h.a.i) e diferenciação dos haustórios secundários (48 h.a.i e 72 h.a.i) no mesófilo do cafeeiro.

A aplicação dessa abordagem pode representar um avanço no contexto da caracterização funcional de efetores de *H.vastatrix*, visando à identificação de genes de avirulência do patógeno e entendimentos dos mecanismos geradores de novas raças fisiológicas, contribuindo para o desenvolvimento de novas estratégias para obtenção de resistência durável à ferrugem do cafeeiro.

REFERÊNCIAS

- ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, London, v. 215, n. 3, p. 403-410, Oct. 1990.
- CANTU, D. et al. Genome analyses of the wheat yellow (stripe) rust pathogen *Puccinia striiformis* f. sp. *tritici* reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. **BMC Genomics**, London, v. 14, p. 270-271, Apr. 2013.
- CAPUCHO, A. S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Note**, Melbourne, v. 7, p. 189-191, Oct. 2012.
- CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, Oxford, v. 21, n. 18, p. 3674-3676, Sept. 2005.
- CRISTANCHO, M. A. et al. Annotation of a hybrid partial genome of the coffee rust (*Hemileia vastatrix*) contributes to the gene repertoire catalog of the Pucciniales. **Frontiers in Plant Science**, Lausanne, v. 5, p. 594-595, Oct. 2014.
- DALIO, R. J. D. et al. Efeitores na interação planta-patógeno. **Revisão Anual de Patologia de Plantas**, Passo Fundo, v. 22, p. 25-68, 2014.
- DANGL, J. L.; HORVATH, D. M.; STASKAWICZ, B. J. Pivoting the plant immune system from dissection to deployment. **Science**, New York, v. 341, n. 6147, p. 746-751, Aug. 2013.
- DE WIT, P. et al. Fungal effector proteins: past, present and future. **Molecular Plant Pathology**, Oxford, v. 10, n. 6, p. 735-747, Nov. 2009.
- DINIZ, I. et al. Cellular and molecular analyses of coffee resistance to and nonhost resistance to in the resistance-donor genotype HDT832/2. **European Journal of Plant Pathology**, Dordrecht, v. 133, n. 1, p. 141-157, May 2012.
- DOU, D. et al. RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. **The Plant Cell**, Rockville, v. 20, n. 7, p. 1930-1947, July 2008.

DUPLESSIS, S. et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proceeding of the National Academy of Science**, Washington, v. 108, n. 22, p. 9166-9171, May 2011.

ELLIS, J. G. et al. Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. **Current Opinion in Plant Biology**, London, v. 12, n. 4, p. 399-405, Aug. 2009.

EMANUELSSON, O. et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. **Journal of Molecular Biology**, Amsterdam, v. 300, n. 4, p. 1005-1016, July 2000.

FERNANDEZ, D. et al. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. **Molecular Plant Pathology**, Oxford, v. 13, n. 1, p. 17-37, Jan. 2012.

GAJ, T.; GERSBACH, C. A.; BARBAS, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. **Trends Biotechnology**, Barking, v. 31, n. 7, p. 397-405, July 2014.

GANESH, D. et al. Monitoring of the early molecular resistance responses of coffee (*Coffea arabica* L.) to the rust fungus (*Hemileia vastatrix*) using real-time quantitative RT-PCR. **Plant Science**, Limerick, v. 170, n. 6, p. 1045-1051, 2006.

JINEK, M. et al. RNA-programmed genome editing in human cells. **Elife**, Cambridge, v. 2, p. 00471, Jan. 2013.

JONES, J. D. G.; DANGL, J. L. The plant immune system. **Nature**, London, v. 444, n. 7117, p. 323-329, Nov. 2006.

KROGH, A. et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. **Journal of Molecular Biology**, Amsterdam, v. 305, n. 3, p. 567-580, Jan. 2001.

LI, W.; GODZIK, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. **Bioinformatics**, Oxford, v. 22, n. 13, p. 1658-1659, July 2006.

- LOPES, R. L. F. **Análise citológica e perfil de expresso gênica de *Hemileia vastarix* (raça XXXIII) na interação com o cafeeiro.** 2015. 64 p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa, MG, 2015.
- NEMRI, A. et al. The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. **Frontiers in Plant Science**, Lausanne, v. 5, p. 1-14, Mar. 2014.
- PERIYANNAN, S. et al. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. **Science**, New York, v. 341, n. 6147, p. 786-788, Aug. 2013.
- PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, New York, v. 8, n. 10, p. 785-786, Sept. 2011.
- PETRE, B; JOLY, D. L.; DUPLESSIS, S. Effector proteins of rust fungi. **Frontiers in Plant Science**, Lausanne, v. 5, p. 415-416, Aug. 2014.
- PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Research**, London, v. 29, n. 9, p. 45-46, May 2001.
- POVOLOTSKAYA, I. S.; KONDRASHOV, F. A. Sequence space and the ongoing expansion of the protein universe. **Nature**, London, v. 465, n.7300, p. 922-926, June 2010.
- SAINTENAC, C. et al. Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. **Science**, New York, v. 341, n. 6147, p. 783-786, Aug. 2013.
- SAUNDERS, D. G. et al. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. **PLoS One**, San Francisco, v. 7, n. 1, p. 29847, 2012.
- SCHORNACK, S. et al. Engineering plant disease resistance based on TAL effectors. **Annual Review Phytopathology**, Palo Alto, v. 51, p. 383-406, 2013.
- SILVA, M. C. et al. Coffee resistance to the main diseases: leaf rust and coffee berry disease. **Brazilian Journal Plant Physiology**, Londrina, v. 18, n. 1, p. 119-147, Jan./Mar. 2006.

SILVA, M. C. et al. Hypersensitive cell death and post-haustorial defense responses arrest the Orange rust (*Hemileia vastatrix*) growth in resistant coffee leaves. **Physiological and Molecular Plant Pathology**, London, v. 60, n. 4, p. 169-183, Apr. 2002.

SILVA, M. C. et al. Involvement of peroxidases in the coffee resistance to orange rust (*Hemileia vastatrix*). **Physiological and Molecular Plant Pathology**, London, v. 72, n. 1/3, p. 29-38, 2008.

SPANU, P. D. et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. **Science**, Washington, v. 330, n. 6010, p. 1543-1546, Dec. 2010.

STERGIOPOULOS, I.; DE WIT, P. J. Fungal effector proteins. **Annual Review of Phytopathology**, Palo Alto, v. 47, p. 233-263, 2009.

TALHINHAS, P. et al. Overview of the functional virulent genome of the coffee leaf rust pathogen *Hemileia vastatrix* with an emphasis on early stages of infection. **Frontiers in Plant Science**, Lausanne, v. 5, p. 87-88 Mar. 2014.

VIEIRA, A. et al. Validation of RT-qPCR reference genes for in planta expression studies in *Hemileia vastatrix*, the causal agent of coffee leaf rust. **Fungal biology Reviews**, Amsterdam, v. 115, n. 9, p. 891-901, Sept. 2011.

VIEIRA, A. et al. Expression profiling of genes involved in the biotrophic colonisation of *Coffea arabica* leaves by *Hemileia vastatrix*. **European Journal of Plant Pathology**, Dordrecht, v. 133, n. 1, p. 261-277, 2012.

WIN, J. et al. Effector biology of plant-associated organisms: concepts and perspectives. **Cold Spring Harbor Symposia on Quantitative Biology**, Cold Spring Harbor, v. 77, p. 235-247, 2012.

ZAMBOLIM, L.; CHAVES, G. M. Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredosporos de *Hemileia vastatrix* Berk. et Br. e *Uromyces phaseolityca* Arth. **Experientiae**, Viçosa, MG, v. 17, p. 151-184, 1974.