

REJANE DO LIVRAMENTO FREITAS LOPES

**ANÁLISE CITOLÓGICA E PERFIL DE EXPRESSÃO GÊNICA
DE *Hemileia vastatrix* (RAÇA XXXIII) NA INTERAÇÃO COM O
CAFEIRO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

VIÇOSA
MINAS GERAIS - BRASIL
2015

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

L864a
2015
Lopes, Rejane do Livramento Freitas, 1983-
Análise citológica e perfil de expressão gênica de *Hemileia vastatrix* (raça XXXIII) na interação com o cafeeiro / Rejane do Livramento Freitas Lopes. – Viçosa, MG, 2015.
x, 64f. : il. (algumas color.) ; 29 cm.

Orientador: Ney Sussumu Sakiyama.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. *Coffea*. 2. Café - Citogética. 3. Café - Doenças e pragas.
4. *Hemileia vastatrix*. 5. Ferrugem. 6. Expressão gênica.
7. Transcriptoma. I. Universidade Federal de Viçosa.
Departamento de Fitotecnia. Programa de Pós-graduação em
Genética e Melhoramento. II. Título.

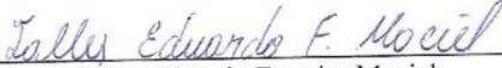
CDD 22. ed. 633.73

REJANE DO LIVRAMENTO FREITAS LOPES

**ANÁLISE CITOLÓGICA E PERFIL DE EXPRESSÃO GÊNICA
DE *Hemileia vastatrix* (RAÇA XXXIII) NA INTERAÇÃO COM O
CAFEIEIRO**

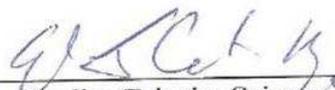
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

APROVADA: 12 de agosto de 2015.


Talles Eduardo Ferreira Maciel


Antônio Carlos Baião de Oliveira


Eunize Maciel Zambolim
(Coorientadora)


Eveline Teixeira Caixeta
(Coorientadora)


Ney Sussumu Sakiyama
(Orientador)

À memória do meu pai que, desde a
graduação, já me chamava
carinhosamente de “minha Doutora”

DEDICO

AGRADECIMENTOS

Em primeiro lugar, a Deus, pelo dom da vida e pela força nos momentos mais difíceis da caminhada.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realização do Curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudo.

Ao professor Ney Sakiyama, pela orientação e confiança em mim depositada. Mas, acima de tudo, por compreender meus anseios e apoiar as minhas decisões.

À Dra. Eunize Maciel Zambolim, com quem pude conviver durante os últimos oito anos e que tanto me ensinou. Agradeço o carinho, a amizade e toda a ajuda durante a realização deste trabalho.

À Dra. Eveline Caixeta, pela valiosa colaboração dada ao trabalho, especialmente na fase final.

Ao Dr. Talles Maciel, cuja ajuda foi fundamental para a realização deste trabalho.

Ao Dr. Antônio Baião, pela disponibilidade em participar da banca.

Ao professor Dr. Laércio Zambolim, pela oportunidade de me integrar à equipe do Biocafé, me concedendo a bolsa de pesquisa do Consórcio e, posteriormente, me contemplando com a bolsa do INCT-Café. Agradeço a consideração e confiança.

Ao INCT-Café, na pessoa do prof. Dr. Mário Lúcio Vilela Resende, que possibilitou a realização deste trabalho, por custear o sequenciamento das amostras.

Aos colegas do Laboratório Biocafé, pelo agradável convívio e amizade. De modo especial, à amiga Daniela Lelis, que muito colaborou na execução dos experimentos; ao Juan e à Brenda, pela ajuda nas análises finais; e ao casal Patrícia e Alexandre, pela amizade.

Aos bioinformatas Cleydson e Otávio e ao prof. Moysés, pela valiosa ajuda.

À toda a equipe do CIFIC/Oeiras/Portugal, pela hospitalidade com que me receberam, e por terem me incentivado a iniciar o Doutorado. Em especial, à Dra.

Maria do Céu, que me ensinou as técnicas de citologia e à Paula Leandro (in memoriam) que me auxiliou diretamente no laboratório durante o treinamento. À Dra. Leonor, Andreia e Vânia, pela amizade e carinho, não só durante este período, mas até os dias de hoje.

À minha família, que sempre me apoiou e torceu por mim. Ao meu pai, que, mesmo não estando mais entre nós, continua vivo em nossos pensamentos. À minha mãe, meu porto-seguro, com quem sempre pude contar, principalmente através das orações. Aos meus irmãos, pelo amor incondicional, amizade e apoio. Às cunhadas e sobrinhas, por estarem sempre presentes na minha vida.

Aos meus sogros, Sr. Pedrinho e Dona Rita, aos cunhados e sobrinhas, pelo apoio, pela preocupação e pelas orações.

Ao meu esposo Ueder, por estar sempre ao meu lado, me dando a força necessária para não desanimar diante das dificuldades. Por enxugar minhas lágrimas, ouvir minhas lamentações, compreender minhas preocupações e vibrar com minhas conquistas. E, ainda, pelas valiosas sugestões e correções da tese.

Aos casais de amigos César e Iris, Glessner e Régia, Luciano e Daniela, pela amizade e pela força na fase final.

A todos os amigos e familiares, que de alguma forma, contribuíram para meu crescimento profissional e pessoal. E a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Muito Obrigada!

BIOGRAFIA

REJANE DO LIVRAMENTO FREITAS LOPES, filha de Honório José Vicente Neto (in memorian) e Simone do Livramento Moreira Neto, nasceu no dia quinze de junho de 1983, em Barroso, MG.

Em março de 2001, ingressou no curso de Bioquímica, na Universidade Federal de Viçosa, MG, graduando-se como Bacharel em Bioquímica em janeiro de 2005.

Em março de 2005, ingressou no Programa de Pós-Graduação em Fisiologia Vegetal, em nível de mestrado, na mesma instituição, concentrando seus estudos na área de Biologia Molecular de Plantas. Obteve o título de Magister Scientiae em abril de 2007.

Em agosto de 2011, iniciou o doutorado no Programa de Pós-Graduação em Genética e Melhoramento da UFV, submetendo-se à defesa de tese em agosto de 2015.

SUMÁRIO

RESUMO	vii
ABSTRACT	ix
1. INTRODUÇÃO GERAL	1
2. REFERÊNCIAS BIBLIOGRÁFICAS	4
CAPÍTULO I	6
1. INTRODUÇÃO	7
2. MATERIAL E MÉTODOS	9
2.1. Multiplicação e preservação de esporos de <i>H. vastatrix</i>	9
2.2. Avaliação citológica da interação cafeeiro-ferrugem	10
2.2.1. Avaliação da viabilidade dos esporos	10
2.2.2. Avaliação do processo de colonização	11
2.2.3. Avaliação das respostas de defesa da planta	11
3. RESULTADOS E DISCUSSÃO	13
4. REFERÊNCIAS BIBLIOGRÁFICAS	17
CAPÍTULO II	18
1. INTRODUÇÃO	19
2. MATERIAL E MÉTODOS	21
2.1. Sequenciamento do transcriptoma de <i>Hemileia vastatrix</i>	21
2.1.1. Preparo das amostras	21
2.1.2. Extração de RNA total	21
2.1.3. Envio das amostras de RNA para sequenciamento	22
2.2. Análise de qualidade e tratamento dos reads	22
2.3. Montagem dos contigs	23
2.4. Predição de genes de <i>H. vastatrix</i>	24
2.5. Anotação dos genes de <i>H. vastatrix</i>	24
2.6. Caracterização de sequências “exclusivas” de <i>H. vastatrix</i>	25
3. RESULTADOS E DISCUSSÃO	26
4. REFERÊNCIAS BIBLIOGRÁFICAS	36
CAPÍTULO III	38
1. INTRODUÇÃO	39
2. MATERIAL E MÉTODOS	42
2.1. Inoculação das plantas de cafeeiro com a raça XXXIII de <i>H. vastatrix</i>	42
2.2. Sequenciamento das bibliotecas de cDNA	43
2.3. Mapeamento das bibliotecas contra o transcriptoma de <i>H. vastatrix</i>	45
2.4. Análise da expressão de transcritos de <i>H. vastatrix</i> presentes na interação com o cafeeiro	46
2.5. Análise de genes exclusivos das bibliotecas da interação	47
3. RESULTADOS E DISCUSSÃO	49
4. REFERÊNCIAS BIBLIOGRÁFICAS	59
CONCLUSÕES GERAIS	61
GLOSSÁRIO	63

RESUMO

LOPES, Rejane do Livramento Freitas, D.Sc., Universidade Federal de Viçosa, agosto de 2015. **Análise citológica e perfil de expressão gênica de *Hemileia vastatrix* (raça XXXIII) na interação com o cafeeiro.** Orientador: Ney Sussumu Sakiyama. Coorientadoras: Eunize Maciel Zambolim e Eveline Teixeira Caixeta.

A raça XXXIII de *Hemileia vastatrix* foi recentemente identificada no Brasil, infectando algumas cultivares resistentes à ferrugem do cafeeiro. Embora o processo infeccioso de *H. vastatrix* seja bem estudado, não existem informações acerca da interação de cafeeiros com esta raça do fungo. Diante disso, um dos objetivos deste trabalho foi avaliar o processo de colonização do fungo e as respostas de defesa da planta em genótipos de cafeeiro resistente (Híbrido de Timor – HDT CIFC 832/1) e suscetível (*C. arabica* cv. Caturra CIFC 19/1) infectados pela raça XXXIII. As primeiras respostas citológicas induzidas pelo fungo foram observadas particularmente nas células estomáticas de ambos os genótipos, às 17 horas após inoculação (hai), e corresponderam à morte celular e ao acúmulo de compostos fenólicos. No genótipo suscetível, o fungo foi capaz de colonizar os tecidos do hospedeiro e produziu um grande número de haustórios, culminando na esporulação cerca de 20 dias após a inoculação (dai). Ao contrário, no genótipo resistente, o crescimento do fungo foi impedido, com alta frequência no estágio pré-haustorial. Estas observações citológicas indicam que a resistência de HDT CIFC 832/1 à raça XXXIII de *H. vastatrix* é pré-haustorial, ao contrário da resistência pós-haustorial que é geralmente descrita para interações cafeeiro - *H. vastatrix*. Com base nesta análise, foram definidos os tempos mais adequados para o estudo de genes expressos ao longo do processo infeccioso. Embora o sequenciamento do transcriptoma tenha sido bastante utilizado em estudos de interação planta-patógeno, uma das maiores dificuldades consiste na separação dos transcritos provenientes da planta e do fungo, principalmente quando não existe genoma de referência. Por esta razão, foi realizado o sequenciamento e a montagem do transcriptoma de esporos hidratados e germinados da raça XXXIII de *H. vastatrix*. A montagem resultou em 32.882 contigs, a partir dos quais foi gerado um conjunto de 11.989 genes preditos. Plantas de *C. arabica* cv. Caturra Vermelho (CIFC 19/1) e Híbrido de Timor (CIFC 832/1)

foram inoculadas com esporos frescos de *H. vastatrix* (raça XXXIII) para estabelecer uma interação compatível e incompatível, respectivamente. Como controle, foram utilizadas folhas não inoculadas. Nos tempos de coleta pré-definidos (12, 24, 96 horas e 17 dias após a inoculação), as folhas foram removidas e utilizadas para extração de RNA. O sequenciamento das dez bibliotecas de cDNA foi realizado na plataforma Illumina MiSeq, gerando um total de 206 milhões de reads. Após tratamento dos reads, foi realizado o mapeamento das bibliotecas da interação, utilizando como referência o transcriptoma de *H. vastatrix* previamente montado. Os dados obtidos mostraram que muitos genes de *H. vastatrix* são similares aos genes do cafeeiro, o que dificulta o estudo da expressão de genes na interação. Por outro lado, genes não conservados entre os dois organismos puderam ser avaliados quanto à sua expressão. Foi verificada a presença de muitos genes comuns às duas interações (compatível e incompatível), sendo que a grande maioria não apresentou expressão diferencial, principalmente nas fases iniciais do processo infeccioso. As maiores diferenças entre as interações foram encontradas aos 17 dias, consistindo principalmente de genes “no hit”, ou seja, genes que não apresentam similaridade com sequências dos bancos de dados de proteínas. Estas sequências podem corresponder a genes específicos de *H. vastatrix*. As informações geradas nesta pesquisa poderão auxiliar no melhor entendimento da interação entre *H. vastatrix* e o cafeeiro. O conhecimento dos genes expressos durante a infecção poderá auxiliar no entendimento dos mecanismos moleculares que levam à suplantação da resistência por novas raças do fungo.

ABSTRACT

LOPES, Rejane do Livramento Freitas, D.Sc., Universidade Federal de Viçosa, August, 2015. **Cytological analysis and gene expression profiling of *Hemileia vastatrix* (race XXXIII) in the interaction with coffee.** Adviser: Ney Sussumu Sakiyama. Co-advisers: Eunize Maciel Zambolim and Eveline Teixeira Caixeta.

The race XXXIII of *Hemileia vastatrix* was recently identified in Brazil, infecting some cultivars that had been released as resistant to coffee leaf rust. To gain new insights into the interaction between coffee plants and this race of the fungus, we evaluated the fungal growth and the plant defence responses in coffee genotypes resistant (Híbrido de Timor – HDT CIFC 832/1) and susceptible (*C. arabica* cv. Caturra CIFC 19/1) to race XXXIII. The first cytological responses induced by the fungus were observed particularly in stomatal cells of both coffee genotypes by 17 hours post-inoculation (hpi) and corresponded to hypersensitive-like response (HR) and accumulation of phenolic-like compounds. In the susceptible genotype, the fungus was able to colonize the host tissues, produced a large number of haustoria and sporulated around 20 days post-inoculation (dpi). Reversely, in the leaves of resistant genotype, the fungus stopped its growth with higher frequency at pre-haustorial stages. These cytological observations indicate that the resistance of HDT CIFC 832/1 to race XXXIII of *H. vastatrix* is pre-haustorial, contrary to the post-haustorial resistance that is generally described for coffee - *H. vastatrix* interactions. Based on this analysis, the most suitable key time points of the infection process were defined, aiming at the study of differentially expressed genes in the interaction. Although the transcriptome sequencing has been widely used in plant-pathogen interaction studies, a major challenge is to separate transcripts from the plant and the fungus, especially in the absence of reference genomic sequences. For this reason, we performed the transcriptome sequencing and assembly of *H. vastatrix* (race XXXIII) hydrated and germinated spores. A total of 32.882 contigs were assembled, from which 11.989 genes were predicted. Plants of *C. arabica* cv. Caturra (CIFC 19/1) and Híbrido de Timor (CIFC 832/1) were inoculated with fresh spores of *H. vastatrix* race XXXIII to establish a compatible and an incompatible interaction, respectively. No inoculated leaves were used as control. In each time point (12, 24, 96 hours and 17 days post-

inoculation), the leaves were removed and used for RNA extraction. Illumina MiSeq sequencing of ten cDNA libraries generated 206,000,000 reads. High quality reads of the *H. vastatrix* and coffee interacting transcriptomes were mapped against the *H. vastatrix* (race XXXIII) hydrated and germinated spores transcriptome previously assembled. The data showed that many genes of *H. vastatrix* are similar to coffee genes, which makes difficult the study of gene expression in the interaction. On the other hand, genes not conserved between the organisms could be evaluated for their expression. The presence of many common genes to both interactions (compatible and incompatible) was verified, and the majority did not show differential expression, especially in the early stages of the infection process. The major differences between interactions were found at 17 dpi, consisting mainly of genes that showed no similarity to protein databases. Those sequences can correspond to *H. vastatrix* specific genes. This investigation provides new insights into the coffee-rust interaction. The knowledge of gene expression profiling during the infection process may contribute to understanding the molecular mechanisms that lead to the breakdown of resistance by new physiological races of the fungus.

1. INTRODUÇÃO GERAL

A cafeicultura está entre as atividades agrícolas de maior importância econômica e social, proporcionando meios de subsistência para milhões de pessoas em todo o mundo. O Brasil ocupa posição de destaque como principal produtor e exportador deste produto, tendo sido responsável por 32% da produção mundial na safra 2014/15, com mais de 45 milhões de sacas (ICO, 2015). Atualmente, o café é o quinto item agrícola mais exportado do país e tem como principais países importadores os Estados Unidos e a Alemanha (CONAB, 2015). As duas espécies que apresentam importância econômica são *Coffea arabica* L. (café arábica) e *C. canephora* Pierre ex Froehner (café robusta). A produção de café arábica corresponde a mais de 70% da produção brasileira, tendo como maior produtor o estado de Minas Gerais, responsável por mais de 50% da produção nacional (CONAB, 2015).

Embora seja a espécie mais cultivada e comercializada no mundo devido à superioridade da qualidade de bebida, *C. arabica* é altamente suscetível a doenças. Dentre as doenças, destaca-se a ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix* Berkeley & Broome. Folhas infectadas tendem a cair precocemente, levando à desfolha e posterior seca dos ramos. Com isso, a produção do ano seguinte é drasticamente afetada, com perdas que podem variar de 35 a 50%, dependendo das condições climáticas e de outros fatores, como cultivar plantada, estado nutricional da planta, tipo de irrigação, dentre outros (Zambolim et al., 1999).

O controle químico com uso de fungicidas protetores e/ou sistêmicos é uma das principais estratégias utilizadas no controle da doença. No entanto, apresenta uma série de desvantagens, principalmente com relação ao impacto ambiental e

contaminação do alimento e de aplicadores. Em função disso, o plantio de cultivares resistentes constitui a medida de controle mais indicada, por ser mais eficiente e causar menor impacto ao ambiente (Zambolim et al., 1999).

Diversas instituições de pesquisa têm desenvolvido programas de melhoramento genético para obtenção de cultivares de cafeeiro resistentes à ferrugem. No entanto, um grande desafio dos melhoristas tem sido alcançar a resistência durável, uma vez que a alta variabilidade genética de *H. vastatrix* favorece o surgimento de novas raças do patógeno que, ao longo do tempo, pode levar à suplantação da resistência no campo (Várzea & Marques, 2005). As raças de *H. vastatrix* mais recentemente descobertas têm sido capazes de infectar derivados de Híbrido de Timor (HDT), a principal fonte de resistência utilizada nos programas de melhoramento genético (Gichuru et al., 2012). Por exemplo, a cultivar Oeiras MG 6851, originada do cruzamento entre *C. arabica* cv. Caturra (CIFC 19/1) e HDT (CIFC 832/1), foi lançada como resistente à ferrugem (Pereira et al., 2000). No entanto, 12 anos após sua liberação, teve sua resistência suplantada pela raça XXXIII de *H. vastatrix* (Capucho et al., 2012). Portanto, conhecer os mecanismos moleculares envolvidos na interação entre o fungo e a planta durante o processo infeccioso é fundamental para auxiliar na compreensão da suplantação da resistência por novas raças fisiológicas de *H. vastatrix*.

Estudos nas áreas de genômica e transcriptômica têm contribuído para o entendimento dos mecanismos da resistência do cafeeiro à ferrugem. A análise in silico de 214.964 ESTs (Expressed Sequence Tags) geradas pelo Projeto Brasileiro do Genoma Café levou à identificação de 14.060 ESTs potencialmente associadas à resistência do cafeeiro a doenças (Alvarenga et al., 2010). Tais sequências apresentaram similaridade com diversas proteínas relacionadas ao processo de defesa

da planta, incluindo cinases, quitinases, citocromo P450, proteínas R, proteínas contendo domínios NBS (nucleotide binding site) e LRR (leucine-rich repeat), dentre outras. Análises da expressão gênica em genótipos de cafeeiro inoculados com *H. vastatrix* permitiram a identificação de genes da planta relacionados com os processos de reconhecimento, sinalização, comunicação celular, defesa e metabolismo (Fernandez et al., 2004; Guzzo et al., 2009; Diola et al., 2013).

Para o fungo, tem sido demonstrada a expressão de genes envolvidos principalmente nos processos de sinalização, estabelecimento e manutenção da biotrofia (Fernandez et al., 2012; Vieira et al., 2012; Talhinhos et al., 2014). A análise parcial do transcriptoma de *H. vastatrix* (raça XIV), expresso na fase de esporulação (21 dias após inoculação), permitiu a identificação de alguns efetores candidatos, dentre os quais homólogos de HESPs (haustorially expressed secreted proteins) e da proteína RTP1 (rust transferred protein 1), quitina desacetilases e endoglucanases (Fernandez et al., 2012). Além disso, o transcriptoma de *H. vastatrix* (raça XIV) foi analisado durante as fases de germinação e formação do apressório, obtidas *in vitro* (Talhinhos et al., 2014). Nos esporos germinados, foi verificado um intenso transporte, atividade secretória e multiplicação celular, enquanto que no apressório foi observado um metabolismo ativo, atividade traducional, produção de novas estruturas e sinalização. Estas observações sugerem que a comunicação entre a planta e o fungo inicia precocemente, antes mesmo de ocorrer a penetração. Portanto, visando entender melhor a interação entre cafeeiro e *H. vastatrix*, é importante avaliar as diferentes etapas, desde a fase de germinação até a esporulação. Diante disso, o objetivo deste trabalho foi estudar a interação entre cafeeiros e a raça XXXIII de *H. vastatrix*, tanto em nível citológico quanto molecular, ao longo do processo infeccioso.

2. REFERÊNCIAS BIBLIOGRÁFICAS

ALVARENGA, S. M. et al. In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases. **Genetics and Molecular Biology**, v. 33, n. 4, p. 795-806, 2010.

CAPUCHO, A. S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Notes**, v. 7, n. 1, p. 189-191, 2012.

CONAB, 2015. <http://www.conab.gov.br/> (Acesso em 03-07-2015).

DIOLA, V. et al. A new set of differentially expressed signaling genes is early expressed in coffee leaf rust race II incompatible interaction. **Functional & Integrative Genomics**, v. 13, n. 3, p. 379-389, 2013.

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**, v. 13, n. 1, p. 17-37, 2012.

FERNANDEZ, D. et al. Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). **Molecular Plant Pathology**, v. 5, n. 6, p. 527-536, 2004.

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

GUZZO, S. D.; HARAKAVA, R.; TSAI, S. M. Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. **Journal of Phytopathology**, v. 157, n. 10, p. 625-638, 2009.

ICO 2015. <http://www.ico.org/> (Acesso em 03-07-2015).

PEREIRA, A. A. et al. Cultivar de café resistente à Ferrugem: Oeiras-MG 6851. **Revista Ceres**, v. 46, p. 121-124, 2000.

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**, v. 5, n. 88, p. 1-17, 2014.

VÁRZEA, V. M. P.; MARQUES, D. Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: Zambolim, L.; Zambolim, E. M.; Várzea, V. M. P. (Eds). **Durable Resistance to Coffee Leaf Rust**. Viçosa: Editora UFV, 2005, p. 53-74.

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**, v. 133, n. 1, p. 261-277, 2012.

ZAMBOLIM, L. et al. Manejo integrado das doenças do cafeeiro. In: Zambolim, L. (Ed). **Produção de café com qualidade**. Viçosa: Editora UFV, 1999, p. 134-215.

CAPÍTULO I

**Avaliação citológica do processo infeccioso de *Hemileia vastatrix*
(raça XXXIII) em cafeeiros resistente e suscetível**

1. INTRODUÇÃO

A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* Berkeley & Broome, é considerada a principal doença da cultura. O fungo está amplamente distribuído nas regiões produtoras de café, causando severos danos econômicos. Embora o uso de cultivares resistentes seja uma das principais estratégias para o controle da doença, o surgimento de novas raças do fungo tem levado à suplantação da resistência no campo. Recentemente, a raça XXXIII foi identificada no Brasil, infectando algumas cultivares que haviam sido liberadas como resistentes à ferrugem do cafeeiro, incluindo a cultivar Oeiras MG 6851 (Pereira et al., 2000; Capucho et al., 2012). Esta cultivar é derivada de um cruzamento entre *C. arabica* cv. Caturra Vermelho CIFC 19/1 e o Híbrido de Timor (HDT) CIFC 832/1, uma das principais fontes de resistência nos programas de melhoramento do café.

Durante o processo infeccioso, *H. vastatrix* diferencia várias estruturas de infecção especializadas, como tubo germinativo, apressório, hifa de penetração e haustórios. Inicialmente, os esporos formam o tubo germinativo, em cuja extremidade ocorre a diferenciação do apressório sobre o estômato. Esta etapa garante a adesão no ponto de penetração. Por meio de uma hifa delgada, conhecida como hifa de penetração, o fungo alcança a câmara subestomática, onde ocorre a ramificação em duas hifas laterais, formando uma estrutura em forma de âncora. Na ponta de cada ramo lateral, é formada a célula mãe do haustório (CMH), que invade as células subsidiárias do estômato, formando os haustórios primários. A diferenciação prossegue, com a formação de haustórios secundários nas células do mesófilo. A partir daí ocorre o crescimento de hifas intercelulares e a formação de haustórios, formando um micélio denso que toma todo o mesófilo foliar (Ramiro et

al., 2009). Em genótipos suscetíveis, este processo culmina com a esporulação na superfície externa da folha via estômatos, cerca de 20 dias após a infecção (Silva et al., 1999). Por outro lado, em genótipos resistentes, o crescimento do fungo é retardado e cessa nas primeiras fases do processo infeccioso, geralmente após a formação do haustório primário. Observações citológicas sugerem que esta resposta seja expressa por uma reação de hipersensibilidade (HR), que se caracteriza por uma morte rápida e localizada das células do hospedeiro no local da infecção. Além disso, várias respostas de defesa pós-haustoriais estão associadas à HR, incluindo revestimento do haustório com calose e β -1,4-glucanos, deposição de compostos fenólicos, lignificação da parede celular da planta, acúmulo de material intercelular contendo polissacarídeos e fenóis, dentre outros. Estas modificações atuam como uma barreira, impedindo a penetração do patógeno ou limitando o seu crescimento nos tecidos da planta (Silva et al., 2002; 2006; 2008).

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. Até o momento, não existe nenhuma informação sobre a interação entre o cafeeiro e a raça XXXIII de *H. vastatrix*. Portanto, o objetivo deste estudo foi avaliar as diferentes etapas do processo infeccioso em cafeeiros resistente e suscetível, infectados pela raça XXXIII. Além de melhorar o entendimento sobre a interação, as informações geradas poderão ser utilizadas para definir os tempos mais adequados do processo infeccioso para a construção de bibliotecas de cDNA, visando ao estudo de genes diferencialmente expressos nestas interações.

2. MATERIAL E MÉTODOS

2.1. Multiplicação e preservação de esporos de *H. vastatrix*

Neste estudo, foi utilizado o isolado de *H. vastatrix* caracterizado como raça XXXIII, mantido na micoteca do laboratório Biotecnologia do Cafeeiro/BIOAGRO/UFV. A escolha desta raça fisiológica se deve ao fato de ter sido identificada no Brasil, suplantando a resistência da cultivar Oeiras MG 6851, a qual foi originada do cruzamento entre *C. arabica* cv. Caturra (CIFC 19/1) e o HDT CIFC 832/1 (Pereira et al., 2000; Capucho et al., 2012).

A multiplicação do isolado foi feita em mudas de cafeeiro *C. arabica* cv. Caturra (CIFC 19/1), pela deposição dos esporos na face abaxial de folhas jovens e completamente desenvolvidas, com auxílio de um pincel. As folhas foram aspergidas com água e cobertas com saco plástico, a fim de manter o ambiente úmido necessário para a germinação do fungo. As plantas foram mantidas a 22°C, no escuro, por 48 horas, e após este período, foram transferidas para câmara de crescimento a 22°C e fotoperíodo de 12 horas. Aos 30 dias após a inoculação (dai), os esporos foram coletados em cápsulas de gelatina e armazenados em dessecador, a 4°C e 50% de umidade. Inoculações sucessivas foram realizadas, de forma a garantir inóculo ‘fresco’ (in planta), com alta porcentagem de germinação, para os trabalhos posteriores. Eventualmente, o isolado (raça XXXIII) foi inoculado no clone diferenciador (CIFC 7963/117, portador dos genes $S_{H5,7}$ ou $S_{H5,7,9}$), com intuito de preservar a sua integridade genética.

2.2. Avaliação citológica da interação cafeeiro-ferrugem

Plantas de *C. arabica* cv. Caturra Vermelho (CIFC 19/1, contendo o gene de resistência S_H5) e Híbrido de Timor (CIFC 832/1, contendo os genes $S_H5,6,7,8,9,?$) foram inoculadas com esporos frescos de *H. vastatrix* (raça XXXIII, genótipo v_5v_7 ou $v_5v_7v_9$) para estabelecer uma interação compatível e incompatível, respectivamente. A inoculação foi feita conforme descrito no item 2.1, sendo inoculadas oito folhas jovens e completamente desenvolvidas, para cada interação. O processo de colonização do fungo e as respostas de defesa da planta foram avaliados em diferentes tempos após a inoculação: 10, 17, 24, 48, 72 e 96 horas; 17 e 21 dias. Em cada tempo de coleta, uma folha foi removida e, com auxílio de um criomicrotomo, foram obtidos cortes com espessura de 22-25 μm . Simultaneamente, foi realizada a avaliação da viabilidade dos esporos (% de germinação e formação de apressório). O experimento foi realizado duas vezes, a fim de confirmar os resultados obtidos.

2.2.1. Avaliação da viabilidade dos esporos

Para que a infecção possa ocorrer com sucesso, é crucial que os esporos estejam viáveis, ou seja, com alta porcentagem de germinação. Uma maneira de fazer esta avaliação é por meio da obtenção de um molde da superfície foliar (Silva et al., 1999). Com esse objetivo, foi feita a avaliação de folhas coletadas às 24 horas após a inoculação (hai). Uma porção da folha inoculada foi retirada e a superfície abaxial foi coberta com esmalte incolor. Após a secagem, a película foi removida, corada com solução de azul de algodão e montada em lâmina para observação ao microscópio óptico (Figura 1). Foram realizadas seis avaliações, cada uma consistindo da contagem de 100 esporos. A porcentagem de germinação foi dada pelo número de esporos germinados em relação ao total, sendo obtida a média das

seis avaliações. Além disso, foi feita a avaliação da formação de apressórios, sendo a porcentagem dada pelo número de esporos germinados contendo apressório em relação ao número total de esporos germinados.

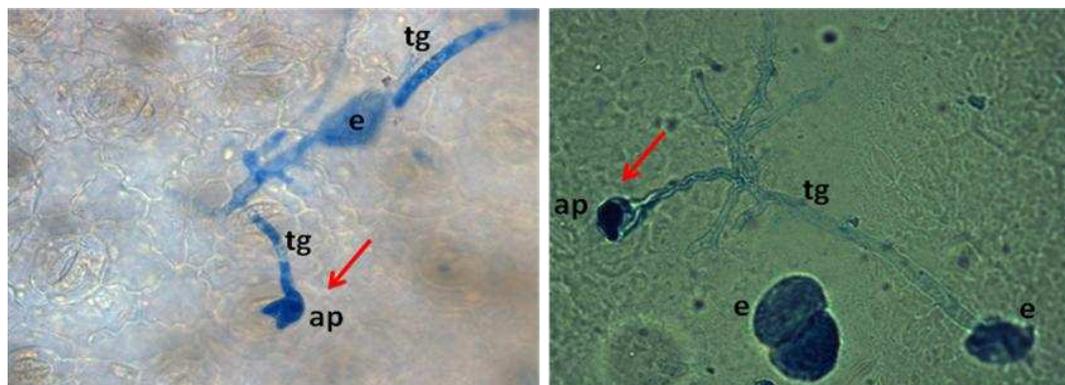


Figura 1. Micrografia de moldes da superfície foliar de cafeeiro inoculado com *H. vastatrix*. Após coloração com azul de algodão, os moldes foram observados ao microscópio óptico. As setas apontam os apressórios diferenciados sobre os estômatos. ap: apressório; e: esporo; tg: tubo germinativo.

2.2.2. Avaliação do processo de colonização

Os cortes histológicos foram clarificados em lactofenol e corados com uma mistura de lactofenol e azul de algodão (Silva et al., 1999). A avaliação foi realizada com auxílio de um microscópio óptico, onde foram observadas as estruturas do patógeno formadas ao longo do processo infeccioso (apressório, hifa de penetração, célula âncora, célula mãe de haustório, haustório, hifa intercelular e hifa esporogênica).

2.2.3. Avaliação das respostas de defesa da planta

Os cortes foliares foram armazenados em tampão fosfato de potássio (K_2HPO_4) 0,07M pH 8,9, montados no mesmo tampão e observados em microscópio de fluorescência sob luz UV (excitação 330-385nm) (Silva et al., 2002). Para cada tempo de coleta, foram analisados 60 pontos de infecção, onde foi avaliado o número de sítios de infecção apresentando autofluorescência e/ou acastanhamento das

células. A autofluorescência pode indicar a presença de compostos fenólicos, estando frequentemente associada com a morte celular no local da infecção (Bennett et al., 1996; Heath, 1998).

3. RESULTADOS E DISCUSSÃO

Às 24 horas após inoculação (hai), tanto nas folhas da planta suscetível quanto na resistente, os esporos apresentaram uma alta porcentagem de germinação (em torno de 75%) e 50% dos esporos diferenciaram apressório sobre o estômato. Este resultado era esperado, uma vez que os esporos utilizados foram coletados diretamente de plantas inoculadas. Caso contrário, se fossem utilizados esporos armazenados em dessecador, provavelmente a viabilidade seria menor, o que poderia comprometer o sucesso da infecção. Não houve diferença entre os valores encontrados para a planta suscetível e a resistente, como também observado por Silva et al. (2002; 2006). Isso porque estes eventos iniciais dependem do reconhecimento da superfície do hospedeiro. Características topográficas da cutícula da planta, como a dimensão das células-guarda do estômato, agem como um sinal para a diferenciação das estruturas de infecção. A ponta do tubo germinativo é capaz de reconhecer as características topográficas da superfície da planta, aumentando a probabilidade de encontrar uma abertura estomática (Mendgen & Voegelé, 2005).

Após a formação do apressório sobre o estômato, *H. vastatrix* desenvolve uma hifa de penetração que cresce na câmara subestomática, produzindo na sua extremidade duas ramificações laterais, em forma de âncora. De cada ramo lateral é formada a célula-mãe do haustório (CMH), que invade as células subsidiárias do estômato, produzindo o haustório primário. Em folhas suscetíveis, na maioria dos sítios de infecção, o fungo prossegue o seu crescimento, com a formação de muitas hifas intercelulares e um grande número de haustórios nas células do mesófilo, chegando até à epiderme superior. Por outro lado, em genótipos resistentes, o crescimento do fungo dentro dos tecidos foliares cessa em diferentes etapas da

infecção, com maior frequência no estágio de CMH com formação de poucos haustórios (Silva et al., 2002; 2006).

Neste estudo, o estágio de hifa de penetração foi o que ocorreu com a mais alta frequência (cerca de 70%), tanto nas folhas resistentes quanto nas suscetíveis, às 24 hai. Entretanto, às 96 hai, o crescimento do fungo foi significativamente maior na planta suscetível, sendo encontradas várias CMHs com haustórios e hifas intercelulares. Ao contrário, na planta resistente, a maioria dos sítios de infecção (65%) permaneceram como hifa de penetração (Figura 2A). Aos 17 dias após inoculação (dai), muitas hifas intercelulares e haustórios estavam presentes na planta suscetível e aos 21 dai foram também encontradas urédias, que representam a etapa final do processo infeccioso (Figura 2B).

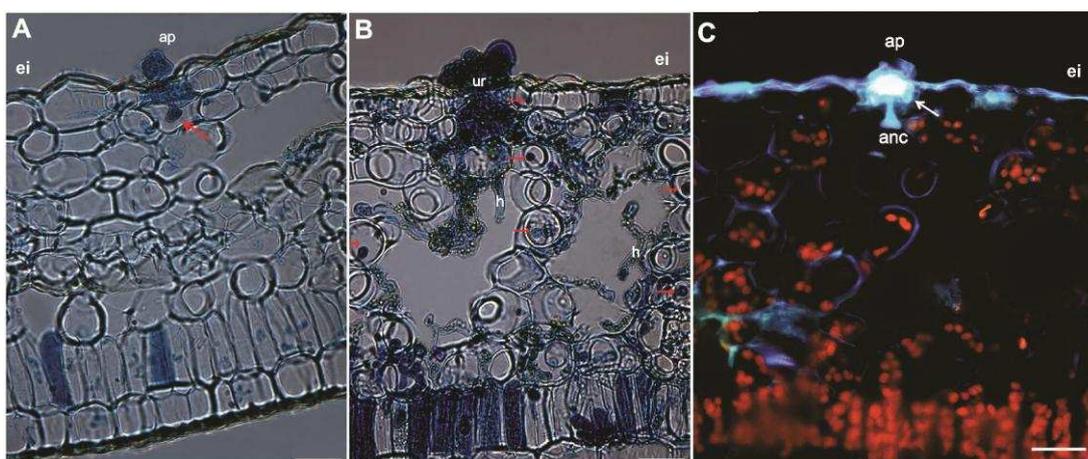


Figura 2. Observações em microscopia de luz. **A e B.** Colonização de tecidos foliares pela raça XXXIII de *H. vastatrix* em plantas resistente (A) e suscetível (B), às 96hai e 21dai, respectivamente. Coloração com azul de algodão e lactofenol: Fig. A. Apressório sobre o estômato e uma hifa de penetração dentro da câmara subestomática (seta). Fig. B. Hifa (h) e haustórios (setas) dentro das células da epiderme inferior (ei) e mesófilo. **C.** Respostas citológicas induzidas pelo fungo no genótipo resistente, às 24hai. Teste de epifluorescência (luz UV): uma âncora (anc) associada com a autofluorescência de células-guarda e subsidiárias (seta), indicando morte celular (HR) e acúmulo de compostos fenólicos. Barra=20 μ m.

Em genótipos resistentes, muitas respostas de defesa são ativadas para prevenir o crescimento do fungo no interior dos tecidos do hospedeiro. Neste trabalho, as primeiras respostas citológicas induzidas pelo fungo (Tabela 1) foram

observadas particularmente nas células estomáticas (Fig. 1C) de ambos os genótipos, às 17 hai, e corresponderam à morte celular e ao acúmulo de compostos fenólicos. No genótipo resistente, estas respostas foram observadas em 18% dos sítios de infecção às 17 hai, alcançando 65% e 93% às 24 e 96 hai, respectivamente (Tabela 1). No genótipo suscetível, tais respostas foram observadas em cerca de 30% dos sítios de infecção, desde as 17 até às 96 hai (Tabela 1). Embora estes valores sejam maiores do que têm sido relatado em outras interações compatíveis entre cafeeiro e *H. vastatrix* (Silva et al., 2006; 2008), o fungo foi capaz de crescer e colonizar os tecidos do hospedeiro (Figura 1B).

Tabela 1. Porcentagem de sítios de infecção com morte celular (HR) e acúmulo de compostos fenólicos (monitorado pela presença de células autofluorescentes) em folhas de genótipos de cafeeiro resistente (HDT832/1) e suscetível (cv. Caturra) à *H. vastatrix* (raça XXXIII)

Horas após inoculação	% de sítios de infecção com células autofluorescentes	
	HDT832/1 x ± DP	cv. Caturra x ± DP
10	0	0
17	18 ± 3	30 ± 5
24	65 ± 13	30 ± 5
48	73 ± 8	30 ± 5
72	83 ± 8	28 ± 6
96	93 ± 3	28 ± 3

x ± DP = média ± desvio padrão

Dados obtidos de 60 sítios de infecção em cada tempo foram apresentados como os valores combinados de dois experimentos, uma vez que não houve diferença significativa entre eles.

Tipicamente, a HR está associada à resistência raça-específica, envolvendo interações gene-a-gene (Heath, 1998; Prats et al., 2007). No entanto, vários trabalhos concordam que o acúmulo de fenóis pode estar associado com a morte celular na resistência do hospedeiro sendo, portanto, uma das primeiras linhas de defesa contra a infecção (Silva et al., 2002; Diniz et al., 2012). Neste estudo, as observações

citológicas indicaram que a resistência de HDT CIFC 832/1 à raça XXXIII de *H. vastatrix* é pré-haustorial, ao contrário da resistência pós-haustorial que é geralmente descrita para interações cafeeiro - *H. vastatrix* (Silva et al., 2002). Isto é consistente com os resultados de Diniz et al. (2012), os quais sugerem que a rápida resposta de resistência, impedindo a formação do haustório, pode ser a base para a durabilidade da resistência. As informações geradas neste trabalho permitirão um melhor entendimento do processo infeccioso envolvendo esta raça de *H. vastatrix*, que foi capaz de suplantar a resistência de cultivares de café derivados do HDT. Além disso, estas informações serão utilizadas para definir pontos-chave ao longo do processo infeccioso, para a construção de bibliotecas de cDNA, visando ao estudo de genes diferencialmente expressos na interação.

4. REFERÊNCIAS BIBLIOGRÁFICAS

BENNETT, M. et al. The hypersensitive reaction, membrane damage and accumulation of autofluorescent phenolics in lettuce cells challenged by *Bremia lactucae*. **The Plant Journal**, v. 9, n. 6, p. 851-865, 1996.

CAPUCHO, A. S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Notes**, v. 7, n. 1, p. 189-191, 2012.

DINIZ, I. et al. Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and nonhost resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. **European Journal of Plant Pathology**, v. 133, n. 1, p. 141-157, 2012.

HEATH, M. C. Involvement of reactive oxygen species in the response of resistant (hypersensitive) or susceptible cowpeas to the cowpea rust fungus. **New Phytologist**, v. 138, p. 251-263, 1998.

MENDGEN, K.; VOEGELE, R. T. Biology of rusts and mechanisms of infection. In: Zambolim, L.; Zambolim, E. M.; Várzea, V. M. P. (Eds). **Durable Resistance to Coffee Leaf Rust**. Viçosa: Editora UFV, 2005, p. 233-248.

PEREIRA, A. A. et al. Cultivar de café resistente à Ferrugem: Oeiras-MG 6851. **Revista Ceres**, v. 46, p. 121-124, 2000.

PRATS, E. et al. Differential effects of phenylalanine ammonia lyase, cinnamyl alcohol dehydrogenase, and energetic metabolism inhibition on resistance of appropriate host and nonhost cereal-rust interactions. **Phytopathology**, v. 97, n. 12, p. 1578-1583, 2007.

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

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

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

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

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**, v. 160, n. 1, p. 79-91, 1999.

CAPÍTULO II

**Transcriptoma da raça XXXIII de *Hemileia vastatrix*:
sequenciamento, montagem e anotação de genes**

1. INTRODUÇÃO

O transcriptoma consiste no conjunto de genes expressos na célula em uma determinada condição biológica. Seu sequenciamento fornece um alto conteúdo de informação funcional, podendo revelar o mecanismo molecular de genes funcionais (Fu et al., 2013).

As tecnologias de sequenciamento de nova geração (NGS – Next-generation Sequencing) têm contribuído para os estudos de transcriptoma, pois permitem gerar milhares de sequências por ensaio e utilizam o sistema de clonagem *in vitro*, que elimina o intensivo trabalho para a produção de clones bacterianos (Morozova et al., 2009). A plataforma Illumina MiSeq, lançada em 2011, é baseada na química sequencing-by-synthesis, gerando reads paired-end, ou seja, fragmentos sequenciados em ambas as direções, de 150-250 bases de comprimento. Após o sequenciamento dos cDNAs, os milhões de reads gerados devem ser montados de novo ou alinhados a um banco de dados de referência, sendo esta uma etapa crucial e que deve ser realizada de forma criteriosa, a fim de se obter resultados confiáveis (Auer et al., 2011). Esta técnica de sequenciamento apresenta uma baixa taxa de erros, o que resulta em maior qualidade dos reads. Portanto, tem sido considerada eficiente, confiável e adequada para o sequenciamento de organismos cujo genoma de referência não está disponível (Loman et al., 2012; Fu et al., 2013).

A existência de um genoma de referência é fundamental para os estudos de interação planta-patógeno. Para *H. vastatrix*, até recentemente, poucas sequências genômicas estavam disponíveis na literatura e nos bancos de dados públicos. Porém, a partir de estudos de transcriptômica realizados com a raça XIV do fungo (Fernandez et al., 2012; Talhinhos et al., 2014) e do sequenciamento parcial do

genoma de *H. vastatrix* (Cristancho et al., 2014), novas informações tornaram-se disponíveis. Apesar disso, sabe-se que diferentes raças fisiológicas podem conter diferentes genes, os quais vão conferir a capacidade de infectar ou não determinado genótipo de cafeeiro. Portanto, embora as sequências genômicas existentes possam auxiliar nos estudos da interação cafeeiro-Hemileia, o mais adequado seria utilizar sequências específicas da raça de interesse.

Até o momento, não existe nenhuma informação genômica referente à raça XXXIII de *H. vastatrix*. Diante do exposto, o objetivo deste trabalho foi realizar o sequenciamento e a montagem do transcriptoma da raça XXXIII de *H. vastatrix*, por meio da plataforma Illumina MiSeq. Com as sequências resultantes, espera-se gerar um banco de dados, para ser utilizado como referência nos estudos da interação entre cafeeiros e esta raça do fungo.

2. MATERIAL E MÉTODOS

2.1. Sequenciamento do transcriptoma de *Hemileia vastatrix*

2.1.1. Preparo das amostras

Esporos de *H. vastatrix* (raça XXXIII) recém-coletados de plantas de cafeeiro foram suspensos em água destilada esterilizada contendo 0,02% de Tween (Sigma) e distribuídos de forma homogênea em placas de poliestireno, contendo uma fina camada de água. As placas foram incubadas a 22°C, no escuro, por 2 horas (para obtenção de esporos hidratados) e 12 horas (para esporos germinados). Após esse período, o material foi coletado e centrifugado a 18.000 x g, por 5 min. O sobrenadante foi descartado e o pellet contendo os esporos foi imediatamente congelado em nitrogênio líquido e armazenado a -80°C, para posterior extração de RNA.

2.1.2. Extração de RNA total

O RNA total das amostras foi extraído utilizando o RNeasy Plant Mini Kit (Qiagen), segundo as instruções do fabricante. A avaliação da quantidade e qualidade do RNA extraído foi realizada espectrofotometricamente em Nanodrop e em gel de agarose a 1,2%. Além disso, o RNA foi quantificado com o Quant-iTTM RiboGreen[®] RNA Reagent (Invitrogen) e avaliado quanto à integridade (RIN) com o Bioanalyzer, utilizando o RNA 6000 Nano Kit (Agilent Technologies).

2.1.3. Envio das amostras de RNA para sequenciamento

As amostras de RNA foram precipitadas em etanol e enviadas para o laboratório Genomics (North Carolina State University – NCSU), onde foram realizadas as etapas de síntese de cDNA, preparo das bibliotecas e sequenciamento.

A síntese de cDNA foi feita utilizando o kit Mint-2 cDNA Synthesis (Evrogen). Após essa etapa, as amostras foram normalizadas, com o kit TRIMMER – cDNA Normalization (Evrogen), visando reduzir os transcritos abundantes. Finalmente, foi feito o preparo das bibliotecas, com o kit TruSeq DNA Sample Preparation (Illumina). As bibliotecas, denominadas “EH” (para “esporo hidratado”) e “EG” (para “esporo germinado”), foram sequenciadas na plataforma Illumina MiSeq, gerando reads paired-end (R1 e R2).

2.2. Análise de qualidade e tratamento dos reads

A fim de avaliar a qualidade dos dados de sequenciamento, foi feita a análise estatística dos dados brutos, utilizando o programa FastQC. Os principais parâmetros avaliados foram “qualidade das bases” e “conteúdo das bases”. A partir desta análise, foram definidos os tratamentos a serem aplicados, visando melhorar a qualidade dos reads. Os tratamentos foram realizados utilizando a ferramenta Prinseq (script Prinseq-lite) e consistiram em:

- eliminação de reads de baixa qualidade (valor de Phred médio inferior a 20);
- eliminação de reads contendo mais de uma base ambígua (“N”);
- eliminação de reads menores que o tamanho mínimo estabelecido (n=50);
- clivagem de adaptadores e vetores;
- clivagem das extremidades 5’ e 3’ dos reads.

Após o tratamento, foi realizada uma nova análise no FastQC, de modo a garantir que os dados (reads) estavam adequados para a etapa posterior de montagem dos contigs.

2.3. Montagem dos contigs

O processo de montagem consiste no alinhamento dos reads, na tentativa de encontrar regiões que sejam suficientemente idênticas e que possam ser unidas em uma única sequência denominada contig. Para isto, foi utilizado o programa CLC Genomics Workbench 6.5.2, com os parâmetros default de 50% de cobertura e 80% de identidade, ajustado para análises de paired-reads. Os parâmetros de distância (mínima e máxima) foram definidos com base no tamanho do inserto sequenciado (436 ± 100 pb). Esta informação é importante para evitar erros na montagem, uma vez que determina a distância entre as sequências nas extremidades de cada clone sequenciado e, portanto, favorece a sobreposição dos paired-reads (R1 e R2). Após a montagem, os contigs gerados foram filtrados por tamanho (> 199 pb).

A fim de enriquecer a montagem, foram incluídas neste processo outras sequências de *H. vastatrix* disponíveis na literatura, sendo 3.161 sequências do NCBI (National Center for Biotechnology and Information), 7.218 sequências do estudo de Fernandez et al. (2012) e 7.894 sequências provenientes do estudo de Talhinas et al. (2014). As 18.273 sequências foram filtradas por tamanho (> 199 pb) e, então, agrupadas por meio do programa CAP3 (Huang et al., 2003), utilizando os parâmetros default (40% de cobertura e 90% de identidade). O arquivo gerado, contendo 10.214 sequências, foi utilizado para o enriquecimento dos dados, visando aumentar o tamanho dos contigs. Para isto, foi utilizado o programa CAP3, com parâmetros mais estridentes (60% de cobertura e 90% de identidade). As sequências

da literatura que permaneceram isoladas (singlets) foram excluídas das análises posteriores.

2.4. Predição de genes de *H. vastatrix*

Para identificar possíveis sequências codificadoras (Open Reading Frames - ORFs) foi utilizada a ferramenta Find Open Reading Frames do programa CLC Genomics Workbench 6.5.2. Foram consideradas apenas as ORFs completas, contendo códon de iniciação e de parada, com número mínimo de 100 códons. As sequências duplicadas foram eliminadas utilizando o programa CD-Hit, cujo parâmetro “c=1” elimina sequências com 100% de identidade.

As ORFs encontradas foram traduzidas para proteínas no programa CLC Genomics Workbench 6.5.2 e passaram a ser denominadas “genes preditos”.

2.5. Anotação dos genes de *H. vastatrix*

Inicialmente, foi realizada a busca de similaridade de todas as sequências proteicas contra o banco de dados de proteínas (nr) do NCBI. Foi utilizado o algoritmo BLAST (Basic Local Alignment Search Tool), disponível no programa CLC Genomics Workbench 6.5.2. Foram considerados como possíveis homólogos os hits com e-value menor que 1×10^{-5} . Sequências com alto valor de e-value (superior a 1×10^{-5}) e aquelas que não apresentaram hit com o banco de dados foram consideradas como “exclusivas” (possíveis genes de *H. vastatrix* ainda não descritos).

As sequências que apresentaram similaridade significativa com sequências do banco de dados foram utilizadas na análise de categorização funcional, por meio da plataforma BLAST2GO (Conesa et al., 2005). Foram analisadas as categorias de

processo biológico, função molecular e componente celular, nos diferentes níveis hierárquicos do GO (Gene Ontology). Além disso, foi realizada a busca de mapas metabólicos no banco de dados KEGG (Kyoto Encyclopedia of Genes and Genomes) e a busca de similaridade contra o banco de dados PHI (Pathogen-Host Interaction, disponível em <http://www.phi-base.org>), contendo genes relacionados à patogenicidade/virulência.

2.6. Caracterização de sequências “exclusivas” de *H. vastatrix*

As sequências que não apresentaram hit contra o banco de dados foram consideradas como “exclusivas” (possíveis genes de *H. vastatrix* ainda não descritos). Assim, a fim de caracterizar as proteínas deduzidas a partir destas ORFs, foi realizada a predição de peptídeo sinal e de localização subcelular, utilizando os algoritmos SignalP (Petersen et al., 2011) e TMHMM (Krogh et al., 2001), visando identificar possíveis proteínas secretadas pelo fungo.

3. RESULTADOS E DISCUSSÃO

O sequenciamento da biblioteca EH gerou 19.095.474 sequências (paired-reads - R1 e R2), com tamanho médio de 250 bases. Após a clivagem das extremidades (65 bases à esquerda e 15 bases à direita) e remoção de reads de baixa qualidade, foram obtidas 7.153.431 sequências cujos reads paired-end foram sobrepostos (R1 + R2), apresentando tamanho médio de 370 pb. Do restante das sequências, 2.168.085 permaneceram como singlets (ou seja, sequência isolada cujo par foi descartado), com tamanho médio de 185 bases, e 2.620.527 foram descartadas (Tabela 1). Para a biblioteca EG foram geradas 11.854.206 sequências (paired-reads - R1 e R2), com tamanho médio de 250 bases. Após o tratamento de qualidade, foram descartados 1.385.573 reads, restando 4.666.293 sequências sobrepostas (R1 + R2) com tamanho médio de 370 pb e 1.136.047 singlets com tamanho médio de 185 bases.

Tabela 1. Dados referentes aos reads, antes e após o tratamento visando melhorar sua qualidade

	EH	EG
Nº de sequências antes do tratamento	19.095.474	11.854.206
Tamanho médio (bases) das sequências	250	250
Nº de sequências sobrepostas (R1 + R2)	7.153.431	4.666.293
Tamanho médio (pb) das sequências sobrepostas	370	370
Nº de singlets	2.168.085	1.136.047
Tamanho médio (bases) dos singlets	185	185
Sequências filtradas pelos parâmetros especificados:		
Baixa qualidade (Phred < 20) na extremidade 5'	518.047	283.409
Tamanho inferior a 50 nucleotídeos	987.021	526.531
Baixa qualidade (valor médio de Phred < 20)	1.096.919	564.317
Mais de uma base ambígua ("N")	18.540	11.316
Nº de sequências após o tratamento	16.474.947	10.468.633

A fim de obter uma maior quantidade de informações a respeito do transcriptoma da raça XXXIII de *H. vastatrix*, os dois conjuntos de dados (EH e EG) foram unidos, gerando um único arquivo com 26.943.580 reads. Após a montagem, foi realizada a filtragem por tamanho (> 199 pb), uma vez que sequências curtas apresentam maior probabilidade de alinhar com sequências não-homólogas, o que poderia comprometer os resultados. Foram gerados 35.430 contigs, com tamanho variando de 200 a 9.028 pb, e tamanho médio de 692 pb. O valor de N50 foi 922 pb, o que significa que 50% de toda a montagem está contida em contigs ≥ 922 pb. Sequências variando de 200 a 2.000 pb totalizaram 93,6%, enquanto que 6,2% dos contigs apresentaram tamanho entre 2.000 e 5.000 pb e apenas 0,2% foram maiores que 5.000 pb (Tabela 2).

Visando aumentar o tamanho dos contigs, foi realizada uma nova montagem, incluindo 10.214 sequências de *H. vastatrix* obtidas da literatura. Como resultado, foram gerados 4.530 contigs e 34.344 singlets. Dentre estes singlets, 28.352 sequências são contigs gerados na etapa anterior e que fazem parte do conjunto final de montagem. As 5.992 sequências restantes são provenientes da literatura e, portanto, foram descartadas por não fazerem parte do conjunto de dados (Figura 1).

Tabela 2. Número e tamanho dos contigs obtidos após as montagens

Montagem	Nº de contigs	Tamanho médio (pb)	N50	< 2.000pb	2.000 – 5.000pb	> 5.000pb
1 ^a	35.430	692	922 pb	33.160 (93,6%)	2.194 (6,2%)	76 (0,2%)
2 ^a	32.882	757	1.112 pb	30.385 (92,4%)	2.363 (7,2%)	132 (0,4%)

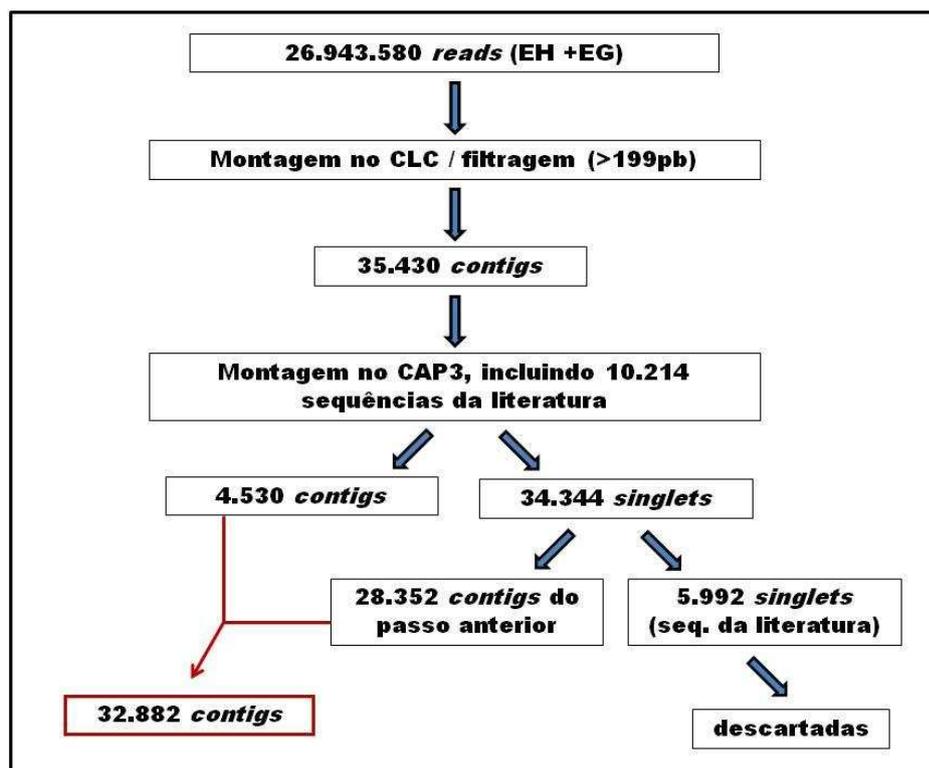


Figura 1. Fluxograma ilustrando os passos das etapas de montagem.

Desta forma, foi obtido o conjunto final de montagem contendo 32.882 contigs, com tamanho variando de 200 a 13.254 pb. A distribuição do comprimento dos contigs (de 250 a 4.799 pb) está ilustrada na Figura 2. É possível observar que a maior parte das sequências estão concentradas na faixa de 200 a 500 pb, referindo-se aos paired-reads sobrepostos, o que mostra que a montagem foi eficiente. Sequências variando de 200 a 2.000 pb totalizaram 92,4%, enquanto 7,2% apresentaram tamanho entre 2.000 e 5.000 pb e apenas 0,4% dos contigs foram maiores que 5.000 pb (Tabela 2). O tamanho médio foi de 757 pb e o valor de N50 foi de 1.112 pb.

Os resultados mostraram que a inclusão de outras sequências de *H. vastatrix* no processo de montagem foi uma estratégia eficiente, uma vez que levou ao aumento do tamanho das sequências (maior N50 e maior média), devido à maior cobertura dos transcritos e consequente redução do número de gaps no alinhamento.

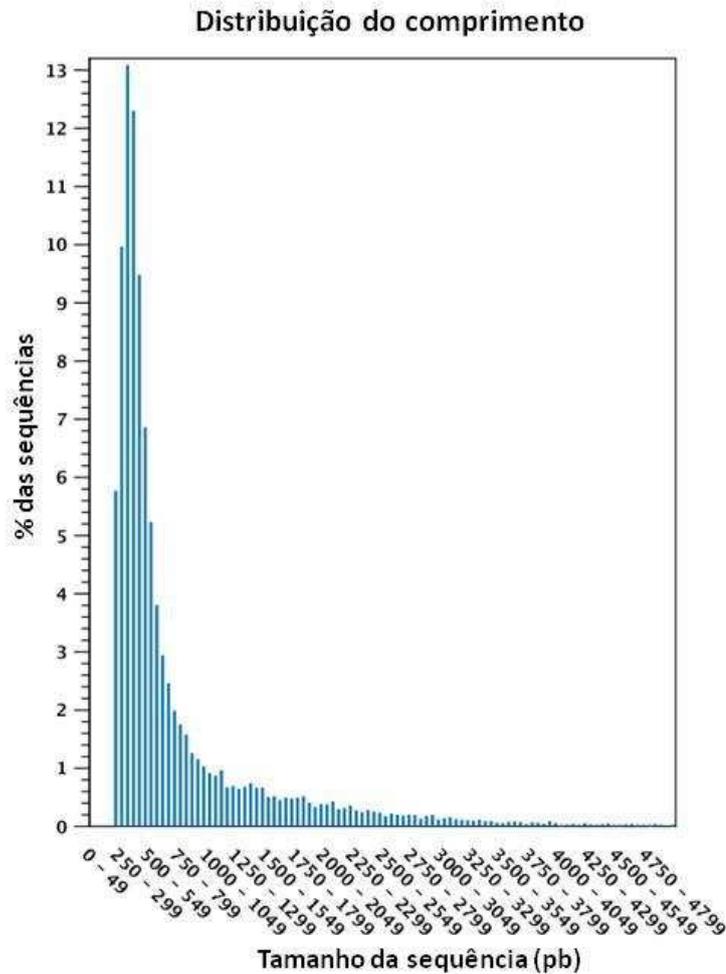


Figura 2. Distribuição do comprimento dos contigs (de 250 a 4.799 pb) após a montagem final.

Partindo das 32.882 seqüências montadas, foram encontradas 12.721 possíveis ORFs completas (contendo códon de iniciação e de parada) e com número mínimo de 100 códons. Após a eliminação de seqüências duplicadas, foram obtidas 11.989 ORFs, que foram traduzidas para proteínas e analisadas pela similaridade contra o banco de dados de proteínas não redundantes do NCBI (NCBI-nr), por meio do algoritmo BLASTp. Do total, 4.084 seqüências não apresentaram similaridade com o banco de dados, sendo que 2.137 seqüências (17,8%) não apresentaram similaridade significativa ($e\text{-value} > 10^{-5}$) e 1.947 seqüências (16,2%) não apresentaram hit contra o banco de dados. Estas seqüências podem representar genes de *H. vastatrix* ainda não descritos ou genes muito divergentes, como também

observado por Talhinhos et al. (2014), ao avaliar o transcriptoma de esporos germinados e apressórios da raça XIV de *H. vastatrix*.

Por outro lado, 7.905 sequências (66,0%) apresentaram similaridade significativa ($e\text{-value} < 10^{-5}$), principalmente com genes de *Melampsora larici-populina* (30,96%) e *Puccinia graminis* (28,05%) (Figura 3). Esta alta similaridade pode ser explicada por dois motivos: i) assim como *H. vastatrix*, os dois patógenos são teliomicetos, com estilos de vida semelhantes e, conseqüentemente, com um conjunto de proteínas similares; ii) existe uma grande quantidade de sequências depositadas em bancos de dados, após a disponibilização dos genomas de ambas as espécies (Duplessis et al., 2011). Somente 23 genes preditos (0,33%) apresentaram similaridade com sequências de *H. vastatrix*, provavelmente devido à baixa quantidade de sequências depositadas no banco de dados.

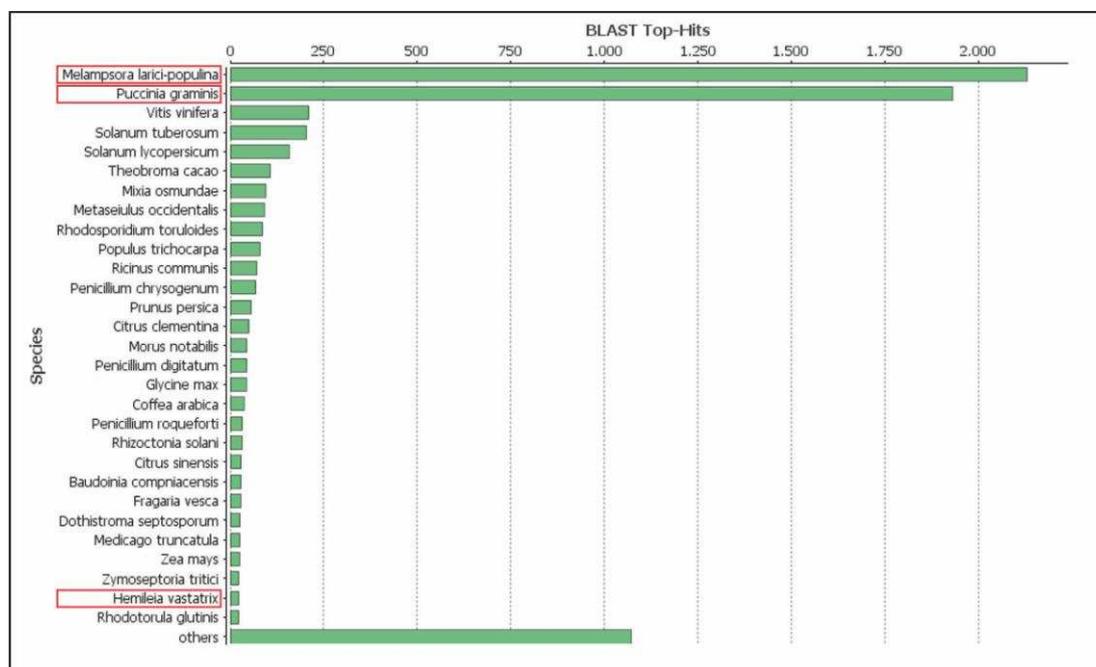


Figura 3 – Distribuição das espécies encontradas em maior frequência na busca de similaridade pelo BLASTp. Em destaque, as duas espécies com a mais alta similaridade (*Melampsora larici-populina* e *Puccinia graminis*) e *Hemileia vastatrix*.

Dentre os 7.905 genes preditos que apresentaram similaridade com as sequências do NCBI, foram eliminados 1.022, por apresentarem o mesmo hit no banco de dados. Desta forma, após eliminar a redundância, restaram 6.883 genes, que foram então utilizados na análise de categorização funcional, por meio da plataforma BLAST2GO. Foram obtidos 12.042 termos GO, distribuídos nas três categorias, o que significa que alguns genes foram associados a dois ou mais termos (em média, 1,75 termos/gene). A categoria mais representada foi “Processo Biológico”, com 5.219 termos (43,3%), seguida por “Função Molecular” e “Componente Celular”, com 3.992 (33,2%) e 2.831 termos (23,5%), respectivamente.

Na categoria “Processo Biológico”, os termos mais representados, considerando o nível hierárquico ‘3’, foram “processo metabólico celular” (14,4%), “processo metabólico de substâncias orgânicas” (13,5%) e “processo metabólico primário” (13,5%) (Figura 4). Este resultado foi confirmado pela análise das vias metabólicas geradas pelo KEGG, que revelou a presença de inúmeros genes codificadores de enzimas do metabolismo primário, como glicólise, ciclo do ácido cítrico, via das pentoses-fosfato, biossíntese e degradação de ácidos graxos, e metabolismo de aminoácidos e nucleotídeos. Além disso, a presença de enzimas envolvidas em outros processos, como sinalização, metabolismo de espécies reativas de oxigênio (ROS), replicação, transcrição, tradução, biogênese de ribossomos e processamento de proteínas, demonstram um metabolismo bastante ativo. Isso foi também observado na análise de BLAST contra o banco de dados PHI (Pathogen-Host Interaction), onde 78 genes preditos de *H. vastatrix* apresentaram alta similaridade a genes relacionados à patogenicidade e virulência.

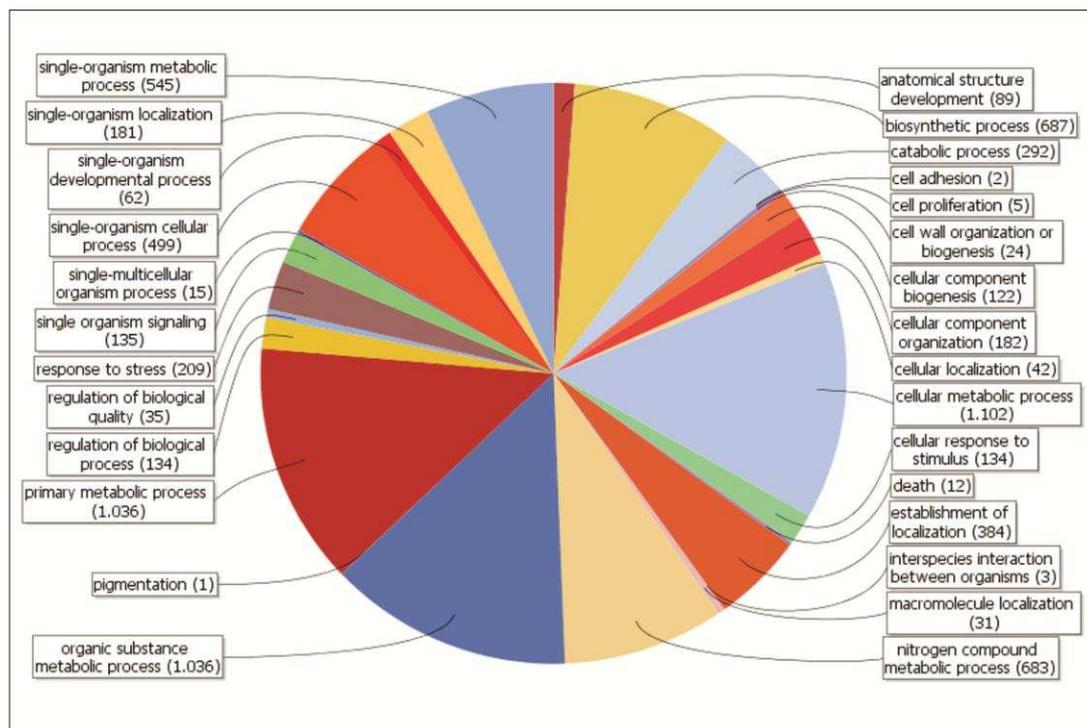


Figura 4 – Distribuição dos termos GO na categoria Processo Biológico, nível ‘3’.

Na categoria “Função Molecular”, considerando o nível hierárquico ‘3’, os termos mais representados foram “ligação a íons” (26,4%), “atividade hidrolase” (13,3%), “atividade transferase” (12,1%) e “atividade oxidorreductase” (10,0%) (Figura 5). Estes resultados foram consistentes com a anotação de ECs (Enzyme Codes), onde as classes enzimáticas mais representadas foram as transferases (30,8%), hidrolases (27,1%) e oxidorreductases (24,0%) (Figura 6). Na classe das hidrolases, destacam-se as enzimas modificadoras de carboidratos, como as glicosil-hidrolases, que podem atuar tanto na remodelagem da parede celular fúngica quanto na degradação da parede celular do hospedeiro, a fim de favorecer a infecção. As enzimas quitinase e quitina sintase atuam em conjunto, em um processo de degradação e síntese, para garantir o crescimento da hifa no tecido do hospedeiro. Outra enzima de grande importância para o fungo é a quitina desacetilase, capaz de converter quitina (principal componente da parede celular fúngica) a quitosana,

impedindo o reconhecimento por receptores de quitina da planta. 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), *Uromyces viciae-fabae* (Deising & Siegrist, 1995), *Puccinia graminis* f. sp. *tritici* e *M. larici-populina* (Duplessis et al., 2011). Outras enzimas como glucanases, manosidases e galactosidases parecem desempenhar importantes papéis na fase inicial da interação fungo-planta, como tem sido observado em vários patossistemas (Garnica et al., 2013).

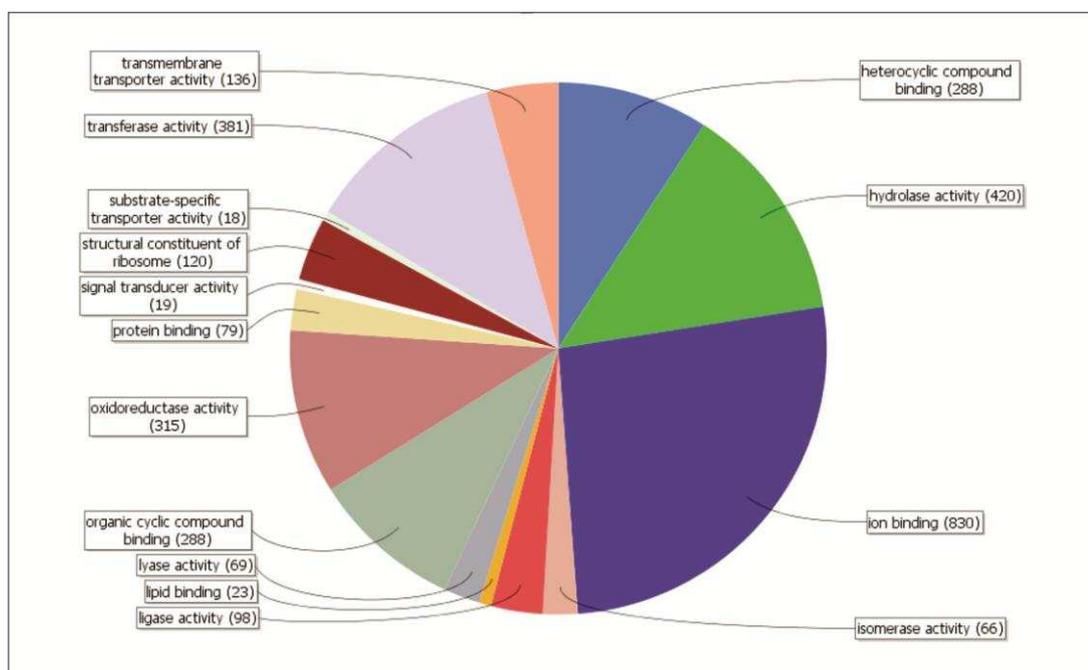


Figura 5 – Distribuição dos termos GO na categoria Função Molecular, nível '3'.

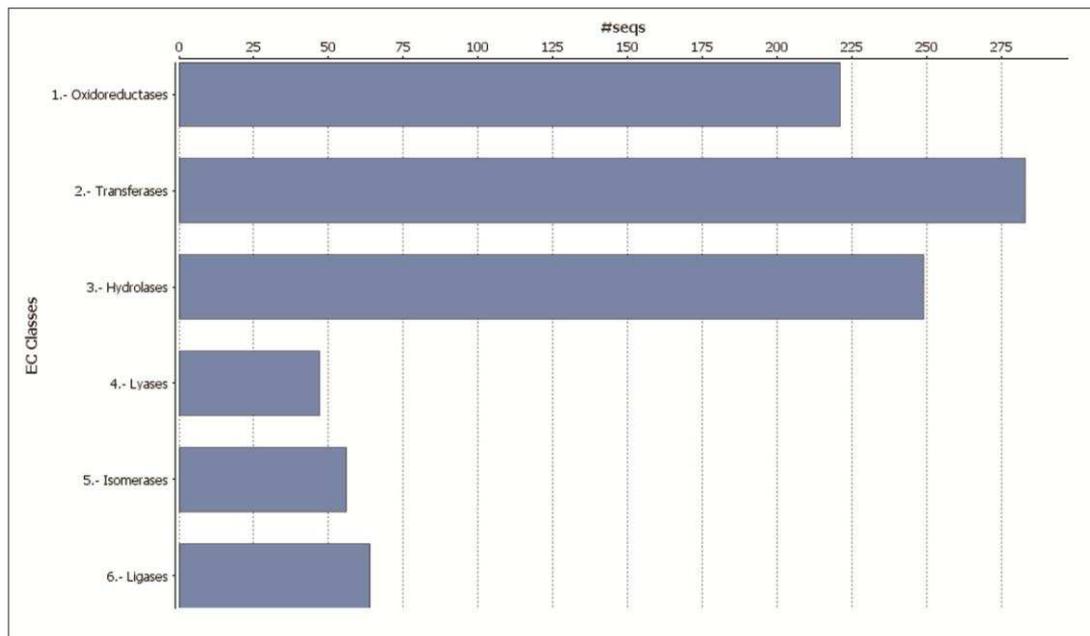


Figura 6 – Distribuição da anotação das classes de enzimas (Enzyme Codes).

Por fim, na categoria “Componente Celular”, considerando o nível hierárquico ‘3’, os termos mais representados foram “parte celular” (47,5%), seguido por “organela membranososa” (27,0%), “complexo de proteínas” (12,1%) e “organela não-membranososa” (9,8%) (Figura 7). 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, que revelou que os genes estão associados principalmente ao núcleo, citosol, mitocôndria, ribossomo, complexo de Golgi e retículo endoplasmático. Nestes componentes ocorrem os mais variados processos biológicos, como os mencionados anteriormente, o que justifica sua maior representação.

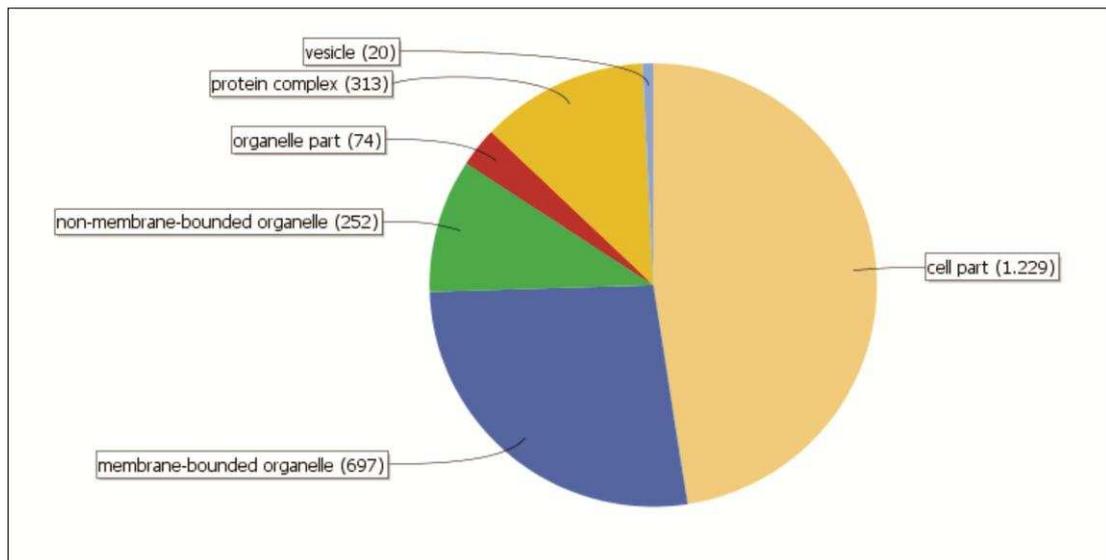


Figura 7 – Distribuição dos termos GO na categoria Componente Celular, nível ‘3’.

A fim de gerar informações acerca dos transcritos “no hit” (possíveis genes “exclusivos” de *H. vastatrix*), foi realizada a análise de localização subcelular. Do total de 1.947 genes preditos, foram encontrados 150 contendo sequências de peptídeo sinal. O tamanho destas proteínas variou de 99 a 1.316 aminoácidos, sendo a maioria (77,6%) com menos de 200 aminoácidos. Pela análise de domínios, foram encontradas 32 proteínas de membrana (sendo 15 proteínas transmembrana e 17 proteínas associadas à membrana) e 118 proteínas extracelulares. As proteínas de membrana podem atuar como receptores, transportadores e na sinalização celular, enquanto que as proteínas secretadas podem estar envolvidas em diferentes funções, como ligação a substratos, comunicação celular, interação entre patógeno e hospedeiro, dentre outras. Tais observações sugerem que estes 150 genes preditos podem representar efetores de *H. vastatrix* ainda não descritos, uma vez que não apresentam similaridade com sequências de espécies relacionadas (Ellis et al., 2009).

4. REFERÊNCIAS BIBLIOGRÁFICAS

AUER, P. L.; SRIVASTAVA, S.; DOERGE, R. W. Differential expression - the next generation and beyond. **Briefings in Functional Genomics**, p. elr041, 2011.

CONESA, A. et al. BLAST2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674-3676, 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**, v. 5, p. 1-11, 2014.

DEISING, H.; SIEGRIST, J. Chitin deacetylase activity of the rust *Uromyces viciae-fabae* is controlled by fungal morphogenesis. **FEMS Microbiology Letters**, v. 127, n. 3, p. 207-211, 1995.

DUPLESSIS, S. et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proceedings of the National Academy of Sciences**, v. 108, n. 22, p. 9166-9171, 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**, v. 12, n. 4, p. 399-405, 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**, v. 13, n. 1, p. 17-37, 2012.

FU, N.; WANG, Q.; SHEN, H-L. De novo assembly, gene annotation and marker development using Illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). **PLoS One**, v. 8, n. 2, p. e57686, 2013.

GARNICA, D. P. et al. Strategies for wheat stripe rust pathogenicity identified by transcriptome sequencing. **PLoS One**, v. 8, n. e67150, 2013.

HUANG, X. et al. PCAP: a whole-genome assembly program. **Genome Research**, v. 13, n. 9, p. 2164-2170, 2003.

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

LOMAN, N. J. et al. Performance comparison of benchtop high-throughput sequencing platforms. **Nature Biotechnology**, v. 30, n. 5, p. 434-439, 2012.

MOROZOVA, O.; HIRST, M.; MARRA, M. A. Applications of new sequencing technologies for transcriptome analysis. **Annual Review of Genomics and Human Genetics**, v. 10, p. 135-151, 2009.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785-786, 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**, v. 5, n. 88, p. 1-17, 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**, v. 133, n. 1, p. 261-277, 2012.

CAPÍTULO III

**Perfil de expressão de genes de *Hemileia vastatrix* (raça XXXIII)
presentes na interação com o cafeeiro**

1. INTRODUÇÃO

A interação cafeeiro-Hemileia é governada pela teoria gene-a-gene, sendo a resistência condicionada por pelo menos nove genes dominantes (S_{H1} a S_{H9}) sozinhos ou associados, embora outros genes de menor efeito possam estar envolvidos (Rodrigues Jr et al., 1975; Bettencourt & Rodrigues Jr, 1988). Proteínas codificadas por estes genes são capazes de reconhecer, direta ou indiretamente, efetores específicos codificados pelos genes de avirulência (Avr) do patógeno, ativando os mecanismos de defesa. Neste caso, o fungo não consegue causar uma infecção estável nos tecidos do hospedeiro, caracterizando uma interação incompatível. Por outro lado, o fungo pode secretar proteínas efetoras, que são capazes de impedir o reconhecimento pela planta e suprimir suas respostas de defesa, além de promover o parasitismo (Ellis et al., 2009). Nesta situação, a doença é desencadeada, resultando em uma interação compatível. Portanto, o conhecimento dos genes expressos pelo fungo e pela planta pode facilitar o entendimento dos eventos moleculares que levam ao desenvolvimento da doença ou à resistência da planta.

Com base nos genes de resistência do cafeeiro, foram caracterizadas 49 raças de *H. vastatrix* em todo o mundo, sendo que 16 delas estão presentes no Brasil (Capucho et al., 2012; Cabral et al., 2009; Zambolim et al., 2005). As raças mais recentemente descobertas têm sido capazes de infectar derivados de Híbrido de Timor (HDT), a principal fonte de resistência utilizada nos programas de melhoramento genético (Gichuru et al., 2012). É o caso da raça XXXIII de *H. vastatrix*, identificada no Brasil infectando plantas da cultivar Oeiras MG 6851,

originada do cruzamento entre *C. arabica* cv. Caturra (CIFC 19/1) e HDT (CIFC 832/1) (Pereira et al., 2000; Capucho et al., 2012).

Embora a fase sexuada de *H. vastatrix* seja desconhecida, algumas evidências têm demonstrado a ocorrência de eventos de recombinação no uredíniosporo, sugerindo que uma reprodução sexuada oculta (criptossexualidade), disfarçada na fase assexuada, pode ser uma das responsáveis pela alta diversidade genética do patógeno (Carvalho et al., 2011). Além disso, a grande proporção de elementos repetitivos encontrados no genoma de *H. vastatrix* (Cristancho et al. 2014) e de outros fungos causadores de ferrugens (Duplessis et al., 2014), pode explicar a alta variabilidade encontrada nestes organismos. Esta alta variabilidade genética favorece o surgimento de novas raças do patógeno que, ao longo do tempo, pode levar à suplantação da resistência no campo (Várzea & Marques, 2005). Portanto, conhecer os mecanismos moleculares utilizados pelo fungo para infectar o cafeeiro é fundamental para auxiliar no entendimento da suplantação da resistência por novas raças fisiológicas de *H. vastatrix*.

Alguns estudos da interação cafeeiro-Hemileia têm permitido a identificação de genes diferencialmente expressos ao longo do processo infeccioso. Em plantas de cafeeiro infectadas pelo fungo tem sido verificada a expressão de genes relacionados com os processos de reconhecimento, sinalização, comunicação celular, defesa e metabolismo (Fernandez et al., 2004; Guzzo et al., 2009; Diola et al., 2013). Para o fungo, tem sido demonstrada a expressão de genes envolvidos principalmente nos processos de sinalização, estabelecimento e manutenção da biotrofia (Fernandez et al., 2012; Vieira et al., 2012; Talhinhos et al., 2014). A análise do transcriptoma de *H. vastatrix*, usando o sequenciamento em larga escala (454-GS-FLX Titanium) de cDNAs obtidos de folhas de cafeeiro inoculadas com o fungo, aos 21 dias após

inoculação (dai), permitiu a identificação de efetores candidatos, dentre os quais homólogos de HESPs (haustorially expressed secreted proteins) e da proteína RTP1 (rust transferred protein 1), quitinases e endoglucanases. Além disso, foram detectados novos transcritos tanto da planta quanto do fungo, os quais não apresentaram similaridade com sequências depositadas em bancos de dados genômicos (Fernandez et al., 2012). Embora tais informações sejam bastante interessantes, os estudos foram concentrados em uma única etapa do processo infeccioso (21 dai). Portanto, é fundamental uma nova abordagem, que abranja diferentes etapas, permitindo assim um melhor entendimento da interação cafeeiro-*H. vastatrix*. Neste sentido, o objetivo deste trabalho foi sequenciar e avaliar o transcriptoma de plantas de cafeeiro infectadas, a fim de gerar informações acerca dos genes diferencialmente expressos por *H. vastatrix* (raça XXXIII) em diferentes etapas do processo infeccioso, em interações compatível e incompatível.

2. MATERIAL E MÉTODOS

2.1. Inoculação das plantas de cafeeiro com a raça XXXIII de *H. vastatrix*

Com base nas avaliações citológicas do processo de colonização do fungo e das respostas de defesa da planta (apresentadas no Capítulo 1), foram definidos os tempos de amostragem para o estudo de genes expressos durante a interação cafeeiro-*H. vastatrix*. Três plantas de *C. arabica* cv. Caturra Vermelho (CIFC 19/1, contendo o gene S_{H5}) e três de Híbrido de Timor (CIFC 832/1, portador dos genes $S_{H5,6,7,8,9,?}$) foram inoculadas com esporos frescos de *H. vastatrix* (raça XXXIII, contendo os genes v_{5v7} ou v_{5v7v9}) para estabelecer uma interação compatível e incompatível, respectivamente. Em cada planta foram inoculadas quatro folhas, e uma folha não inoculada foi utilizada como controle. No total, foram obtidas dez amostras, conforme descrito a seguir:

C-controle - Caturra controle (não inoculada)

C-12h - Caturra x raça XXXIII - 12 horas após inoculação (hai)

C-24h - Caturra x raça XXXIII - 24 hai

C-96h - Caturra x raça XXXIII - 96 hai

C-17d - Caturra x raça XXXIII - 17 dias após inoculação (dai)

HT-controle – Híbrido de Timor controle (não inoculada)

HT-12h - Híbrido de Timor x raça XXXIII - 12 hai

HT-24h - Híbrido de Timor x raça XXXIII - 24 hai

HT-96h - Híbrido de Timor x raça XXXIII - 96 hai

HT-17d - Híbrido de Timor x raça XXXIII - 17 dai

No tempo de coleta pré-definido, as folhas foram removidas, imediatamente congeladas em nitrogênio líquido e acondicionadas a -80°C para posterior extração de RNA. Cada amostra consistiu de um pool composto por três folhas, provenientes de três diferentes plantas.

2.2. Sequenciamento das bibliotecas de cDNA

As etapas de extração de RNA, preparo das bibliotecas, obtenção e tratamento das sequências foram realizadas conforme descrito nas páginas 21 e 22. Para o sequenciamento, foi adicionado um barcode a cada uma das amostras, que foram então reunidas e sequenciadas em dez faixas (lanes).

O conjunto de reads de cada biblioteca foi utilizado para avaliar o perfil de expressão dos genes de *H. vastatrix* durante a interação com cafeeiro. Para facilitar o entendimento, foi montado um fluxograma (Figura 1) que apresenta todas as análises realizadas neste trabalho.

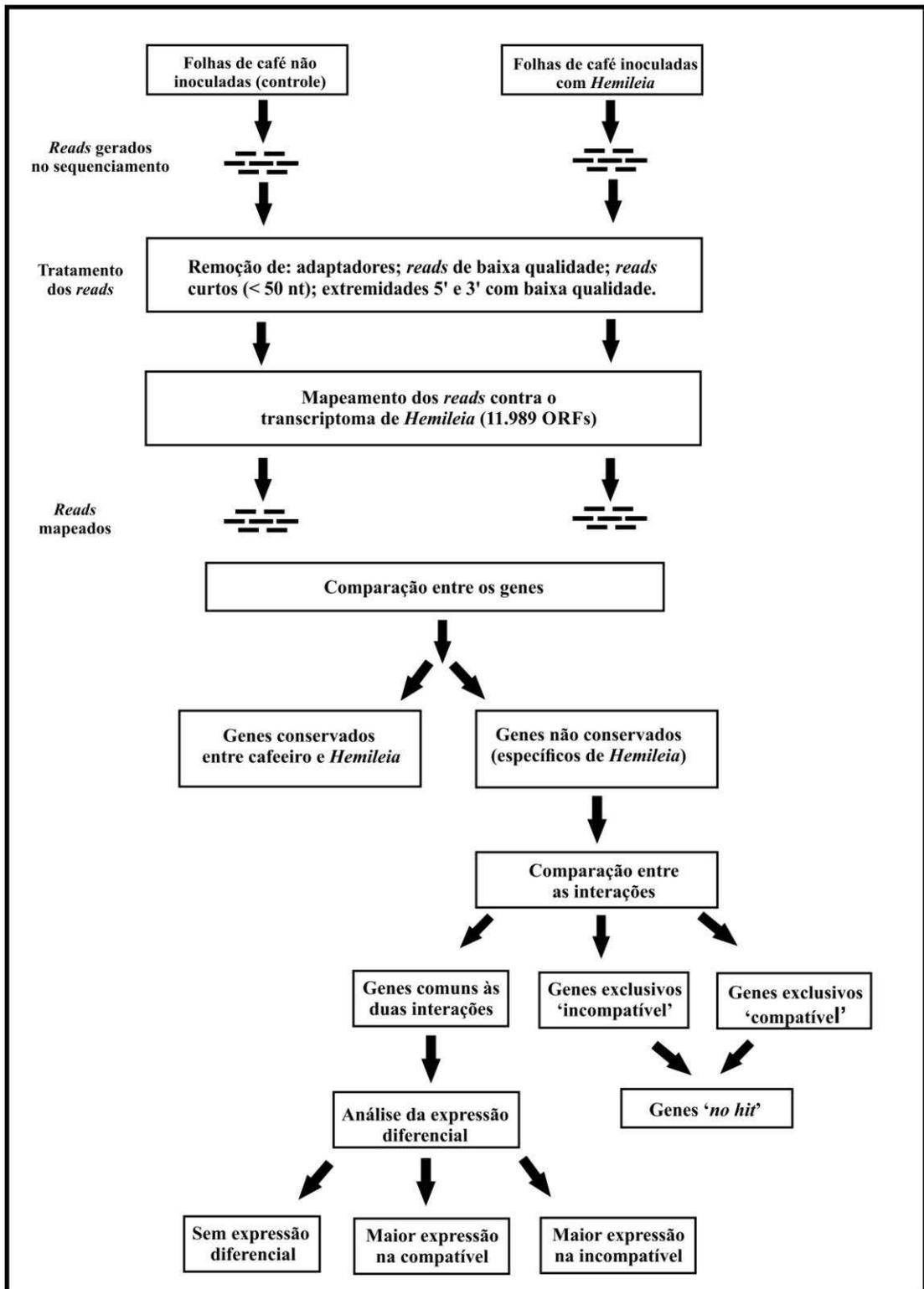


Figura 1. Fluxograma ilustrando as etapas realizadas neste estudo, desde o sequenciamento das amostras até o estudo do perfil de expressão dos genes.

2.3. Mapeamento das bibliotecas contra o transcriptoma de *H. vastatrix*

Devido à ausência do genoma referência de *H. vastatrix* e considerando o interesse em estudar especificamente a raça XXXIII, foram utilizados como referência os dados previamente obtidos pelo sequenciamento e montagem do transcriptoma desta raça (Capítulo 2). Assim, os conjuntos de reads das dez bibliotecas (oito da interação e duas do controle) foram mapeados nas 11.989 ORFs do transcriptoma de *H. vastatrix* previamente montado. Foi utilizado o programa CLC Workbench 6.5.2, com os parâmetros default (50% de cobertura e 80% de identidade). Após o mapeamento, foi feita a comparação entre as ORFs presentes nas bibliotecas da interação e na biblioteca controle. Desta forma, foram gerados 16 conjuntos de dados, sendo dois para cada biblioteca da interação: um contendo os genes conservados entre *H. vastatrix* e cafeeiro (ou seja, genes presentes em ambos os organismos) e outro contendo os genes não conservados (isto é, genes específicos de *H. vastatrix*). Com estes dados, foi realizada uma comparação entre as interações, em cada tempo após a inoculação (Figura 2).

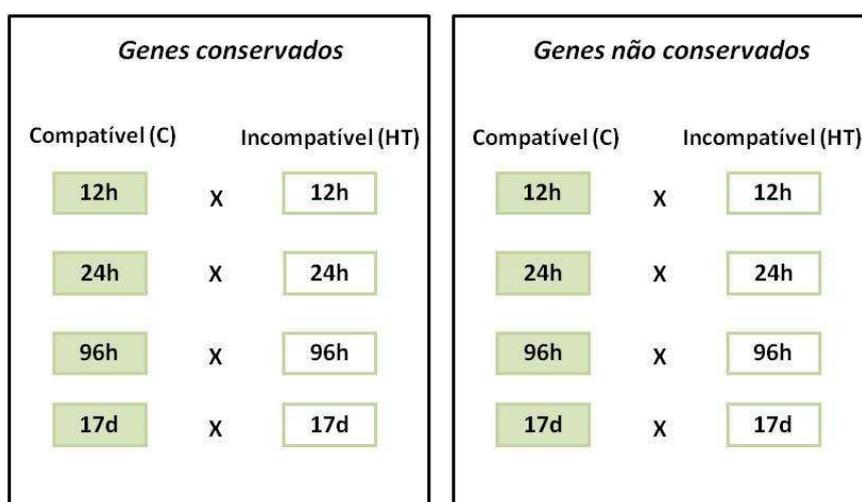


Figura 2. Esquema ilustrando as comparações realizadas entre as bibliotecas, visando separar os genes de *H. vastatrix* (raça XXXIII) encontrados apenas na interação compatível (C) e incompatível (HT) e os genes comuns a ambas.

O objetivo desta análise foi separar os genes comuns a ambas as interações, genes encontrados exclusivamente na interação compatível (Caturra - C) e genes exclusivos da interação incompatível (Híbrido de Timor - HT).

2.4. Análise da expressão de transcritos de *H. vastatrix* presentes na interação com o cafeeiro

O objetivo desta análise foi identificar genes de *H. vastatrix* diferencialmente expressos entre as duas interações (compatível e incompatível) ao longo do tempo. Foram considerados nesta análise apenas os genes não conservados, uma vez que a expressão dos genes conservados poderia ser superestimada devido à presença de reads da planta. Partindo dos contrastes estabelecidos na Figura 2, foi gerado o conjunto de genes comuns às duas interações. A partir daí, foram geradas as tabelas de contagem, contendo o número de reads mapeados em cada uma das 11.989 ORFs do transcriptoma (item 2.3). Com estes dados, foi calculada a expressão relativa, com base no valor de RPKM (Reads Per Kilobase of exon model per Million mapped reads) (Mortazavi et al., 2008). Esta análise faz uma normalização do número de reads mapeados, levando em consideração o número total de reads em cada amostra e o comprimento do gene em questão.

Com estas informações, foi realizada a análise de expressão diferencial, por meio do algoritmo GFOLD (Generalized Fold Change) (Feng et al., 2012). Este método foi escolhido por permitir a estimativa da diferença relativa da expressão gênica em dados de RNA-seq sem repetição biológica. O programa estabelece um equilíbrio entre o grau da variação e a significância, de forma a penalizar genes com baixo nível de expressão e favorecer genes com alta expressão (mesmo que possuam um menor valor de fold change). Como resultado, é gerado um valor de GFOLD,

que corresponde ao “log2 fold change”. Os genes com GFOLD > 0 são diferencialmente up-regulados, enquanto os genes com GFOLD < 0 são down-regulados. Genes com valor de GFOLD = 0 não apresentam expressão diferencial entre as amostras.

O programa foi executado usando a linha de comando:

```
./gfold diff -s1 Sample_1 -s2 Sample_2 -suf .csv -o Sample_1xSample_2.diff
```

Neste caso, valores de GFOLD positivos (GFOLD > 0) representam genes diferencialmente up-regulados na amostra “**Sample_2**”, ao passo que valores negativos (GFOLD < 0) referem-se a genes down-regulados nesta amostra. O raciocínio inverso é válido para a amostra “**Sample_1**”.

É importante considerar que, como as bibliotecas de cDNA foram normalizadas antes do sequenciamento, os níveis de expressão dos transcritos foram alterados, principalmente para aqueles altamente expressos. Neste caso, o valor de ‘fold change’ não deve ser considerado como a expressão real do gene, mas sim como um indicativo da sua expressão em determinada condição. A partir dos valores de GFOLD obtidos em cada contraste, foi feita uma classificação em ordem decrescente, gerando um ranking dos genes diferencialmente expressos. Como ponto de corte foi estabelecido o valor absoluto de GFOLD $\geq 1,0$ (correspondente ao fold change ≥ 2). Assim, genes cujo valor absoluto de GFOLD foi menor que 1,0 não foram considerados diferencialmente expressos.

2.5. Análise de genes exclusivos das bibliotecas da interação

Os conjuntos de dados contendo os genes exclusivos de cada interação foram avaliados quanto à presença de genes “no hit”, ou seja, que não apresentam similaridade com sequências dos bancos de dados. O objetivo desta análise foi

identificar possíveis novos genes de *H. vastatrix*, que estão presentes apenas em uma das interações (compatível ou incompatível), em determinada etapa do processo infeccioso.

3. RESULTADOS E DISCUSSÃO

O sequenciamento das dez bibliotecas de cDNA (C-controle, C-12h, C-24h, C-96h, C-17d, HT-controle, HT-12h, HT-24h, HT-96h e HT-17d) gerou um total de 206 milhões de reads. Os dados referentes ao número de reads de cada biblioteca, antes e após o tratamento para remoção de sequências de baixa qualidade, estão apresentados na Tabela 1. Em todas as bibliotecas, a maior parte dos reads paired-end (76% em média) foram sobrepostos (R1 + R2), gerando sequências com tamanho médio de 370 pb. Do restante das sequências, em média 11% permaneceram na forma de singlets (com tamanho médio de 184 bases) e 13% foram descartadas.

Tabela 1. Número de reads das bibliotecas, antes e após a remoção de sequências de baixa qualidade

Biblioteca	N° total de reads	Reads limpos		% de reads de boa qualidade	N° de reads descartados*
		Reads sobrepostos (R1 + R2)	Singlets		
C-controle	12.258.442	8.637.512	1.111.852	79,5	2.509.078
C-12h	24.937.788	20.535.754	2.060.964	90,6	2.341.070
C-24h	19.149.370	13.859.864	2.478.476	85,3	2.811.030
C-96h	13.416.692	9.792.002	1.664.885	85,4	1.959.805
C-17d	24.571.784	19.002.254	2.577.739	87,8	2.991.791
HT-controle	27.295.036	20.712.516	4.360.856	91,9	2.221.664
HT-12h	22.328.618	17.892.790	2.111.404	89,6	2.324.424
HT-24h	31.061.884	22.796.908	3.857.373	85,8	4.407.603
HT-96h	12.359.240	8.404.730	1.821.760	82,7	2.132.750
HT-17d	18.684.474	15.329.014	1.595.350	90,6	1.760.110

*Reads descartados por apresentarem baixa qualidade. C: Caturra; HT: Híbrido de Timor

Os reads sobrepostos (R1 + R2) foram mapeados contra as 11.989 ORFs do transcriptoma da raça XXXIII de *H. vastatrix* previamente montado (Capítulo 2). Para todas as bibliotecas, a porcentagem de reads mapeados foi em torno de 11%, exceto para a biblioteca C-17d, que teve 23,8% dos reads mapeados (Tabela 2). Isso se deve ao fato de que, na planta suscetível, já aos 17 dias após a inoculação, o intenso crescimento de hifas intercelulares e haustórios culminam na formação de um micélio denso que toma todo o mesófilo foliar (Ramiro et al., 2009). Portanto, devido a essa maior biomassa do fungo no tecido infectado, uma maior quantidade de mRNA foi extraído e, conseqüentemente, sequenciado.

Tabela 2. Número de reads mapeados ao transcriptoma de *H. vastatrix* para as dez bibliotecas da interação

Biblioteca	Reads sobrepostos (R1 + R2)		% de reads mapeados	Nº de ORFs correspondentes*
	Total	Mapeados		
C-controle	8.637.512	926.805	10,7	6.012
C-12h	20.535.754	2.486.701	12,1	10.161
C-24h	13.859.864	1.430.864	10,3	10.163
C-96h	9.792.002	1.049.591	10,7	9.928
C-17d	19.002.254	4.519.619	23,8	10.924
HT-controle	20.712.516	2.174.814	10,5	6.359
HT-12h	17.892.790	1.915.617	10,7	9.679
HT-24h	22.796.908	2.655.100	11,6	9.896
HT-96h	8.404.730	828.738	9,9	8.714
HT-17d	15.329.014	1.460.055	9,5	7.451

*Nº de ORFs representadas pelos reads mapeados, dentre as 11.989 ORFs usadas como referência; C: Caturra; HT: Híbrido de Timor

Curiosamente, para as bibliotecas controle, a porcentagem de reads mapeados ao transcriptoma de *H. vastatrix* também ficou em torno de 11%. Apesar disso, o número de ORFs correspondentes aos reads mapeados foi bem menor (6.012 ORFs na biblioteca C-controle e 6.359 ORFs na HT-controle) do que nas bibliotecas da

interação (em torno de 10.000 ORFs). Este resultado sugere que muitos genes são conservados entre o fungo e a planta, enquanto outros genes são exclusivos do fungo. A fim de investigar esta possibilidade foram gerados oito conjuntos de dados para a interação compatível (Caturra x raça XXXIII) e oito para a incompatível (Híbrido de Timor x raça XXXIII). Para cada tempo, foram obtidos dois conjuntos de dados: um contendo as ORFs conservadas entre *H. vastatrix* e cafeeiro (ou seja, genes similares entre os dois organismos) e outro contendo as ORFs não conservadas (isto é, genes encontrados apenas em *H. vastatrix*).

Como mostrado na Tabela 3, considerando as duas interações, a porcentagem de genes conservados foi maior do que a de genes não conservados (em média, 62% e 38%, respectivamente). Uma possível explicação para este resultado é que os genes conservados devem ser mais abundantes na célula, o que favoreceu o seu sequenciamento. Além disso, como o mapeamento foi realizado com reads provenientes das bibliotecas da interação (planta + fungo), é possível que reads pertencentes à planta também tenham sido mapeados, contribuindo para a maior porcentagem dos genes conservados.

Esta observação dificulta a análise de expressão diferencial dos genes conservados, uma vez que os níveis de expressão poderiam ser superestimados devido à presença de reads da planta, comprometendo os resultados. Além disso, traz à tona uma das grandes dificuldades de se estudar genes expressos na interação entre dois organismos, que é exatamente separar a sua origem. No caso do patossistema cafeeiro-Hemileia, esta tarefa é ainda mais complicada, devido à inexistência de genoma de referência. Recentemente, foi liberado o genoma de referência de *C. canephora*, com 710 Mb (Denoëud et al., 2014). Embora este seja um grande avanço para os estudos envolvendo o cafeeiro, ainda não é a solução, já que diferentes

espécies podem apresentar genes divergentes. Da mesma forma, para *H. vastatrix* foi liberado um genoma parcial, contendo 333 Mb (Cristancho et al., 2014), e algumas sequências do transcriptoma estão disponíveis na literatura (Fernandez et al., 2012; Talhinhos et al., 2014). No entanto, como o genoma da espécie tem sido estimado em 733 Mb (Carvalho et al., 2014), ainda existe muita informação a ser descoberta.

Tabela 3. Número de ORFs conservadas e não conservadas entre as bibliotecas da interação e as bibliotecas controle

Biblioteca	Nº de ORFs mapeadas*	Nº de ORFs conservadas**	Nº de ORFs não conservadas***
C-controle	6.012	-	-
C-12h	10.161	5.890 (58%)	4.271 (42%)
C-24h	10.163	5.901 (58%)	4.262 (42%)
C-96h	9.928	5.867 (59%)	4.061 (41%)
C-17d	10.924	5.971 (55%)	4.953 (45%)
HT-controle	6.359	-	-
HT-12h	9.679	6.135 (63%)	3.544 (37%)
HT-24h	9.896	6.169 (62%)	3.727 (38%)
HT-96h	8.714	5.944 (68%)	2.770 (32%)
HT-17d	7.451	5.450 (73%)	2.001 (27%)

*Nº de ORFs representadas pelos reads mapeados, dentre as 11.989 ORFs usadas como referência; **Nº de ORFs presentes tanto nas bibliotecas da interação quanto no controle, ou seja, genes conservados entre *H. vastatrix* e cafeeiro; ***Nº de ORFs presentes apenas nas bibliotecas da interação, ou seja, genes não conservados, encontrados apenas em *H. vastatrix*; C: Caturra (interação compatível); HT: Híbrido de Timor (interação incompatível).

A comparação entre as bibliotecas revelou que boa parte dos genes são comuns às duas interações, tanto para os genes conservados quanto para os não conservados (Figura 3). Como discutido anteriormente, para os genes conservados não é possível afirmar a origem dos reads mapeados, o que impede de fazer inferências a seu respeito. Portanto, a discussão será voltada apenas para os genes não conservados, que foram considerados como exclusivos do fungo.

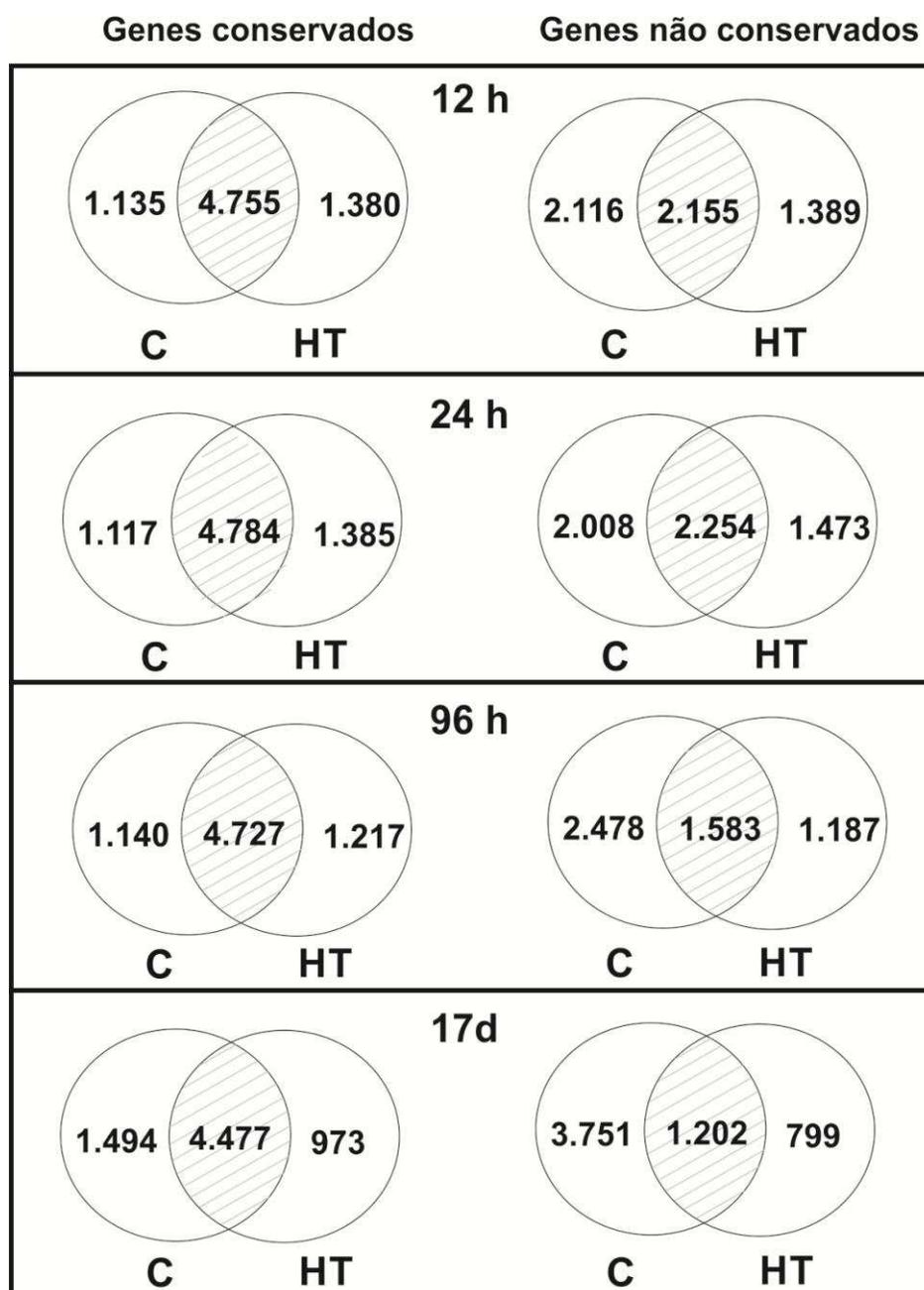


Figura 3. Diagrama de Venn mostrando o número de genes encontrados exclusivamente na interação compatível (Caturra - C) e incompatível (Híbrido de Timor – HT) e o número de genes comuns a ambas (hachurado), em cada tempo após a inoculação (12h, 24h, 96h e 17d). Os diagramas à esquerda referem-se aos genes conservados e os da direita correspondem aos genes não conservados.

A partir dos dados apresentados no diagrama (Figura 3) foram obtidas as porcentagens referentes aos genes não conservados, presentes em cada interação e em cada tempo avaliado (Tabela 4).

Tabela 4. Porcentagem de genes de *H. vastatrix* (não conservados) presentes em cada interação ao longo do tempo

Contraste	Compatível (C)		Incompatível (HT)	
	% genes comuns*	% genes exclusivos**	% genes comuns*	% genes exclusivos***
C-12h x HT-12h	50,4	49,6	60,8	39,2
C-24h x HT-24h	52,9	47,1	60,5	39,5
C-96h x HT-96h	39,0	61,0	57,1	42,9
C-17d x HT-17d	24,3	75,7	60,1	39,9
Média	41,7	58,4	59,6	40,4

*Genes comuns às duas interações (compatível – “C” e incompatível – “HT”); **Genes presentes apenas na interação compatível; ***Genes presentes apenas na interação incompatível. C: Caturra; HT: Híbrido de Timor.

Para a interação compatível, até às 24 hai, cerca de 50% dos genes são comuns com a interação incompatível. No entanto, com o avanço do processo infeccioso, essa porcentagem diminui, chegando a 24,3% aos 17 dai. Por outro lado, para a interação incompatível, a porcentagem de genes comuns se mantém em torno de 60% em todos os tempos avaliados. Tais observações sugerem que estes genes comuns às duas interações correspondam a genes envolvidos nas etapas iniciais do processo infeccioso, como pré-penetração (germinação e formação de apressório) e penetração. Esta hipótese é suportada pela análise citológica dos dois patossistemas (Capítulo 1) que demonstrou que, às 24 hai, o estágio de hifa de penetração foi o que ocorreu com a mais alta frequência (cerca de 70%), em ambas as interações. Além disso, como também verificado nesta análise, na planta resistente, a maioria dos sítios de infecção (65%) permaneceram como hifa de penetração, às 96 hai. Ao contrário, na planta suscetível, o fungo prosseguiu seu crescimento, com a formação de célula-mãe de haustório (CMH) com haustórios e hifas intercelulares. Isso explicaria o fato de, na interação incompatível, a porcentagem desses genes comuns

permanecer praticamente inalterada, enquanto que na interação compatível seu valor diminui com o avanço da infecção.

Com relação aos genes exclusivos, a porcentagem foi sempre maior na interação compatível do que na incompatível (em média, 58,4% e 40,4%, respectivamente), além de aumentar ao longo do tempo. A maior diferença foi observada aos 17 dai, com 75,7% dos genes expressos na interação compatível exclusivos, contra apenas 39,9% da incompatível. Este resultado está de acordo com o esperado, uma vez que, nesta fase avançada da infecção, o desenvolvimento do fungo nos tecidos do hospedeiro suscetível é muito diferente do hospedeiro resistente, o que certamente se reflete na expressão dos genes. Nesta fase, muitos haustórios são encontrados nas células do mesófilo. Estas estruturas são hifas altamente especializadas, responsáveis pela absorção de nutrientes e pela secreção de proteínas efetoras. Estas proteínas são translocadas para o interior da célula do hospedeiro, induzindo uma série de alterações estruturais, que facilitam a infecção (Ellis et al., 2009). Portanto, os genes expressos exclusivamente na interação compatível devem estar envolvidos no estabelecimento e na manutenção da biotrofia, podendo corresponder a possíveis efetores de *H. vastatrix*, inclusive secretados pelo haustório.

A análise da expressão dos genes comuns às duas interações revelou que a maior parte deles não apresentou expressão diferencial (Tabela 5). Como pode ser observado, até às 96 hai, a porcentagem de expressão diferencial foi muito baixa (em média, 4%). Isto reforça a hipótese de que estes genes comuns estejam envolvidos nos eventos iniciais do processo infeccioso (pré-penetração e penetração), sendo igualmente importantes para ambas as interações. A análise de categorização funcional revelou que os genes estão envolvidos em diversos processos biológicos,

como processo celular, processo metabólico, organização de componente celular, resposta a estímulos, localização, regulação biológica e sinalização. Isto foi observado para os quatro contrastes avaliados, sugerindo que os genes expressos em cada condição estão envolvidos em processos biológicos semelhantes.

Tabela 5. Porcentagem de expressão diferencial dos genes comuns a ambas as interações, nos diferentes tempos

Contraste	Sem expressão diferencial*	Maior expressão na compatível	Maior expressão na incompatível
C x HT - 12h	94,9	2,0	3,1
C x HT - 24h	97,1	1,3	1,6
C x HT - 96h	96,5	1,3	2,2
C x HT - 17d	70,9	4,2	24,9
Média	89,9	2,2	7,9

*Genes cujo valor absoluto de GFOLD foi menor que 1,0, ou seja, fold change < 2. C: Caturra; HT: Híbrido de Timor.

Aos 17 dai, a porcentagem de expressão diferencial foi de 29,1%, sendo 4,2% dos genes mais expressos na interação compatível e 24,9% na interação incompatível. Em um primeiro momento, este resultado parece incoerente, uma vez que, na interação incompatível, o crescimento do fungo é impedido, o que compromete a expressão dos genes. No entanto, essa maior porcentagem observada na interação incompatível deve refletir a expressão dos genes nas etapas iniciais da infecção, que se mantêm ao longo do tempo. Já na interação compatível, a expressão destes genes tende a diminuir com o avanço da infecção, em detrimento a novos genes expressos na fase tardia.

Dentre os genes diferencialmente expressos (GDEs), em média 20% não apresentaram similaridade com o banco de dados de proteínas, permanecendo sem anotação (“no hit”) (Tabela 6). Para a interação compatível, a porcentagem de GDEs “no hit” foi de 25%, enquanto na incompatível foi de 17%, em média. Esta ausência

de similaridade com proteínas de espécies relacionadas é uma das características de proteínas efetoras (Ellis et al., 2009). Esta observação, aliada ao fato de terem apresentado expressão diferencial entre as duas interações, sugere que os genes “no hit” podem representar possíveis efetores de *H. vastatrix* ainda não descritos.

Tabela 6. Número de genes diferencialmente expressos (GDEs) em cada biblioteca

Biblioteca	Total de GDE	Genes “no hit”
C-12h	43	9 (20,9%)
C-24h	29	9 (31%)
C-96h	20	5 (25%)
C-17d	50	11 (22%)
HT-12h	67	8 (11,9%)
HT-24h	36	9 (25%)
HT-96h	35	4 (11,4%)
HT-17d	300	61 (20,3%)

C: Caturra; HT: Híbrido de Timor

Além destes genes comuns às duas interações, foram também avaliados os genes encontrados exclusivamente em cada interação, para os quatro contrastes (Tabela 7). Para a interação compatível, o número de genes exclusivos foi sempre maior do que na incompatível, sendo a maior diferença observada aos 17 dai. No total, foram encontrados 10.353 genes expressos exclusivamente na interação compatível, sendo 2.110 genes (20%) “no hit”. Estes genes podem ser considerados candidatos a efetores, pois além de não apresentarem similaridade com proteínas de outras espécies relacionadas, foram encontrados somente na interação compatível, podendo corresponder a proteínas expressas pelo haustório. Na interação incompatível, foram encontrados 4.848 genes exclusivos, dos quais 872 (18%) são “no hit”. Estes genes podem corresponder a efetores liberados pelos haustórios primários, no início do processo infeccioso. De fato, como verificado na avaliação

citológica da interação incompatível (Capítulo 1), o contato do fungo com as células-guarda e subsidiárias do estômato é suficiente para induzir a reação de hipersensibilidade (HR). Portanto, efetores secretados pelos haustórios nestas células podem ser os responsáveis por desencadear as respostas logo no início da infecção.

Tabela 7. Número de genes encontrados exclusivamente em cada interação, considerando os diferentes tempos

Contraste	Compatível		Incompatível	
	Nº de genes exclusivos	% genes “no hit”	Nº de genes exclusivos	% genes “no hit”
C x HT - 12h	2.116	20,1	1.389	19,6
C x HT - 24h	2.008	18,5	1.473	18,5
C x HT - 96h	2.478	20,3	1.187	18,1
C x HT - 17d	3.751	21,6	799	14,1
Média	2.588	20	1.212	18

C: Caturra; HT: Híbrido de Timor.

A análise de categorização funcional dos genes anotados não revelou diferenças entre as interações, sendo os genes atribuídos aos mesmos processos biológicos. Tomados juntos, os resultados apresentados sugerem que as maiores diferenças entre as duas interações residem nos genes “no hit”. A caracterização destes novos genes, candidatos a possíveis efetores de *H. vastatrix*, permitirá inferir sobre sua função biológica durante o processo infeccioso.

4. REFERÊNCIAS BIBLIOGRÁFICAS

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

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

CAPUCHO, A.S. et al. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Notes**, v. 7, n. 1, p. 189-191, 2012.

CARVALHO, G. M. A. et al. Coffee rust genome measured using flow cytometry: does size matter?. **Plant Pathology**, v. 63, n. 5, p. 1022-1026, 2014.

CARVALHO, C. R. et al. Cryptosexuality and the genetic diversity paradox in coffee rust, *Hemileia vastatrix*. **PLoS One**, v. 6, n. 11, p. e26387, 2011.

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**, v. 5, p. 1-11, 2014.

DENOEUDE, F. et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, v. 345, n. 6201, p. 1181-1184, 2014.

DIOLA, V. et al. A new set of differentially expressed signaling genes is early expressed in coffee leaf rust race II incompatible interaction. **Functional & Integrative Genomics**, v. 13, n. 3, p. 379-389, 2013.

DUPLESSIS, S.; BAKKEREN, G.; HAMELIN, R. Advancing knowledge on biology of rust fungi through genomics. **Advances in Botanical Research**, v. 70, p. 173-209, 2014.

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**, v. 12, n. 4, p. 399-405, 2009.

FENG, J. et al. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. **Bioinformatics**, v. 28, n. 21, p. 2782-2788, 2012.

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**, v. 13, n. 1, p. 17-37, 2012.

- FERNANDEZ, D. et al. Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). **Molecular Plant Pathology**, v. 5, n. 6, p. 527-536, 2004.
- GICHURU, E. K. et al. Additional physiological races of coffee leaf rust (*Hemileia vastatrix*) identified in Kenya. **Tropical Plant Pathology**, v. 37, n. 6, p. 424-427, 2012.
- GUZZO, S. D.; HARAKAVA, R.; TSAI, S. M. Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. **Journal of Phytopathology**, v. 157, n. 10, p. 625-638, 2009.
- MORTAZAVI, A. et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, v. 5, n. 7, p. 621-628, 2008.
- PEREIRA, A. A. et al. Cultivar de café resistente à Ferrugem: Oeiras-MG 6851. **Revista Ceres**, v. 46, p. 121-124, 2000.
- RAMIRO, D. A. et al. Biphasic haustorial differentiation of coffee rust (*Hemileia vastatrix* race II) associated with defence responses in resistant and susceptible coffee cultivars. **Plant Pathology**, v. 58, n. 5, p. 944-955, 2009.
- RODRIGUES JR, C. J.; BETTENCOURT, A. J.; RIJO, L. Races of the pathogen and resistance to coffee rust. **Annual Review of Phytopathology**, v. 13, n. 1, p. 49-70, 1975.
- 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**, v. 5, n. 88, p. 1-17, 2014.
- VÁRZEA, V. M. P.; MARQUES, D. Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: Zambolim, L.; Zambolim, E. M.; Várzea, V. M. P. (Eds). **Durable Resistance to Coffee Leaf Rust**. Viçosa: Editora UFV, 2005, p. 53-74.
- 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**, v. 133, n. 1, p. 261-277, 2012.
- ZAMBOLIM, L. et al. Physiological races of *Hemileia vastatrix* in Brazil: physiological variability, current situation and future prospects. In: Zambolim, L.; Zambolim, E. M.; Várzea, V. M. P. (Eds). **Durable Resistance to Coffee Leaf Rust**. Viçosa: Editora UFV, 2005, pp. 75-98.

CONCLUSÕES GERAIS

A análise citológica do processo infeccioso de cafeeiros infectados com a raça XXXIII de *H. vastatrix* revelou que a resistência do HDT CIFC 832/1 a esta raça é pré-haustorial, ao contrário da resistência pós-haustorial geralmente descrita para interações cafeeiro - *H. vastatrix*. Esta observação pode ter uma implicação direta na durabilidade da resistência, devido à rápida resposta de defesa, que impede a formação do haustório e, conseqüentemente, o avanço da infecção.

A partir do sequenciamento do transcriptoma de esporos hidratados e germinados, foi possível gerar um banco de dados contendo 11.989 genes preditos. Estas sequências foram utilizadas como referência para o estudo de genes diferencialmente expressos pelo fungo na interação com o cafeeiro. A análise do transcriptoma permitiu a identificação de 150 possíveis candidatos a efetores de *H. vastatrix* ainda não descritos. Além disso, os resultados de categorização funcional sugeriram que a comunicação entre o fungo e a planta ocorre logo no início do processo infeccioso, ainda durante a fase de germinação dos esporos.

Para avaliar os genes do fungo expressos durante o processo infeccioso, foi realizado o sequenciamento do transcriptoma de plantas de cafeeiro (resistente e suscetível) infectadas com a raça XXXIII de *H. vastatrix*. Esta abordagem foi possível devido à normalização das bibliotecas de cDNA, que reduziu os transcritos abundantes, favorecendo o sequenciamento de transcritos raros. Esta estratégia permitiu o sequenciamento de transcritos do fungo, apesar da maior abundância de mRNA da planta no tecido infectado. Os resultados obtidos mostraram que muitos genes presentes em *H. vastatrix* são similares aos genes do cafeeiro, o que dificulta o estudo da expressão gênica na interação. Por outro lado, genes não conservados entre

os dois organismos apresentaram expressão diferencial entre as interações. As maiores diferenças foram encontradas na fase avançada da infecção (17 dai), consistindo principalmente de genes “no hit”, ou seja, genes que não apresentam similaridade com sequências dos bancos de dados de proteínas de outras espécies.

Nesta pesquisa, foram realizados os primeiros estudos acerca da interação entre cafeeiros e a raça XXXIII de *H. vastatrix*. As informações geradas são bastante relevantes, pois servirão de base para estudos que visam ao melhor entendimento dos mecanismos moleculares envolvidos na suplantação da resistência por novas raças do fungo.

GLOSSÁRIO

BLAST: Basic Local Alignment Search Tool; algoritmo de alinhamento entre regiões de sequências.

BLAST2GO: plataforma utilizada para classificação funcional dos genes nas categorias “Processo Biológico”, “Função Molecular” e “Componente Celular”.

cDNA: DNA complementar, sintetizado a partir de mRNA, por meio da enzima transcriptase reversa.

Contig: sequência de DNA resultante da união de vários fragmentos menores (reads), a partir do alinhamento de regiões sobrepostas.

Efector: proteína secretada pelo patógeno, cujo reconhecimento pelas proteínas de resistência do hospedeiro leva à ativação dos mecanismos de defesa. Por outro lado, na ausência de genes de resistência, estas proteínas são capazes de suprimir as respostas de defesa do hospedeiro e promover o parasitismo.

e-value: refere-se à probabilidade do alinhamento ter ocorrido ao acaso.

Gene Ontology: plataforma para a anotação funcional baseada na ontologia gênica, que permite a categorização dos genes em classes funcionais.

Genes comuns: neste estudo, refere-se aos genes encontrados tanto na interação compatível quanto na incompatível.

Genes conservados: neste estudo, refere-se a genes do fungo que apresentam similaridade com genes da planta.

Genes exclusivos: neste estudo, refere-se aos genes encontrados exclusivamente em uma das interações, compatível ou incompatível.

Genes não conservados: neste caso, trata-se de genes encontrados apenas no fungo e que não apresentaram similaridade com nenhum gene da planta.

Genes no hit: genes que não apresentam similaridade com sequências de proteínas dos bancos de dados.

GFOLD: Generalized Fold Change; programa usado para estimar a diferença relativa da expressão gênica em dados de RNA-seq sem repetição biológica.

Haustório: hifa altamente especializada, responsável pela absorção de nutrientes a partir da célula hospedeira e pela secreção de proteínas efetoras do patógeno.

Hit: termo usado para se referir à similaridade significativa entre determinada sequência e sequências dos bancos de dados.

Interação compatível: ocorre quando o patógeno consegue suprimir as respostas de defesa da planta, desencadeando a doença. A planta é dita suscetível.

Interação incompatível: ocorre quando o patógeno não consegue estabelecer uma infecção estável nos tecidos do hospedeiro, devido ao reconhecimento dos efetores por proteínas de resistência. A planta é dita resistente.

N50: medida estatística que indica que 50% da montagem dos reads (todas as bases sequenciadas) está contida em contigs iguais ou maiores a este valor (em pares de base).

Normalização do cDNA: procedimento realizado antes do sequenciamento, a fim de reduzir os transcritos abundantes.

ORF: Open Reading Frame; região compreendida entre um códon de iniciação e um de terminação, com potencial para codificar uma proteína.

Phred: valor de qualidade associado a cada base sequenciada; quanto maior esse valor, menor é a probabilidade erro de sequenciamento.

Raça fisiológica: variante do patógeno capaz de infectar determinado genótipo da espécie hospedeira, devido à presença de genes de avirulência específicos.

Reads paired-end: fragmentos de DNA sequenciados em ambas as direções, partir das extremidades 5' e 3'.

Similaridade: grau de semelhança entre duas sequências de proteínas.

Singlet: neste caso e considerando o sequenciamento de paired-reads, trata-se da sequência que permaneceu isolada, não sendo encontrada a outra sequência sobreposta.

Transcriptoma: conjunto de genes expressos na célula em uma determinada condição biológica.