UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-graduação em Bioinformática

JÚLIA SANTOS PEREIRA

GENOMIC IDENTIFICATION, GENE EXPRESSION AND ENZYME ACTIVITY OF CAZYMES FROM BIOTROPHIC FUNGUS *Hemileia vastatrix* DURING COFFEE LEAF INFECTION

Belo Horizonte
2021

Júlia Santos Pereira

GENOMIC IDENTIFICATION, GENE EXPRESSION AND ENZYME ACTIVITY OF CAZYMES FROM BIOTROPHIC FUNGUS *Hemileia vastatrix* DURING COFFEE LEAF INFECTION

Versão Final

Dissertação apresentada ao Programa de Pós-graduação em Bioinformática da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Bioinformática.

Orientador: Prof. Dr. Tiago Antônio de Oliveira Mendes

Belo Horizonte
2021

043 Pereira, Júlia Santos.
Genomic identification, gene expression and enzyme activity of CAZymes from biotrophic fungus Hemileia vastatrix during coffee leaf infection [manuscrito] / Júlia Santos Pereira. – 2021.
60 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Tiago Antônio de Oliveira Mendes.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Bioinformática.

1. Genômica. 2. Café. 3. Infecção. 4. Carboidratos. 5. Celulase. 6. Poligalacturonase. I. Mendes, Tiago Antônio de Oliveira. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 573:004

Ficha catalográfica elaborada pela bibliotecária Fabiane C M Reis – CRB 6 - 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Bioinformática

ATA DA DEFESA DE DISSERTAÇÃO

JÚLIA SANTOS PEREIRA

Às oito horas do dia **25 de fevereiro de 2021**, reuniu-se, no aplicativo Zoom, a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho de Júlia Santos Pereira intitulado: **"Genomic identification, gene expression and enzyme activity of CAZymes from biotrophic fungus Hemileia vastatrix during coffee leaf infection"**, requisito para obtenção do grau de Mestre em **Bioinformática.** Abrindo a sessão, o Presidente da Comissão, **Dr. Tiago Antonio de Oliveira Mendes**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	**Instituição**	**Indicação**
Dr. Tiago Antônio de Oliveira Mendes	UFV	**Aprovada**
Dra. Eveline Teixeira Caixeta	EMBRAPA	**Aprovada**
Dr. Siomar de Castro Soares	UFTM	**Aprovada**

Pelas indicações, a candidata foi considerada: **Aprovada**

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 25 de fevereiro de 2021.

Dr. Tiago Antonio de Oliveira Mendes - Orientador

Dra. Eveline Teixeira Caixeta

Dr. Siomar de Castro Soares

Documento assinado eletronicamente por **Siomar de Castro Soares**, **Usuário Externo**, em 25/02/2021, às 11:17, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Eveline Teixeira Caixeta Moura**, **Usuário Externo**, em 25/02/2021, às 11:19, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Tiago Antônio de Oliveira Mendes**, **Usuário Externo**, em 25/02/2021, às 11:19, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0584755** e o código CRC **95728DDC**.

Referência: Processo nº 23072.209268/2021-08 SEI nº 0584755

AGRADECIMENTOS

A Deus, por me sustentar em todos os momentos.

Aos meus pais, Cássia e Roberto, por tanto amor e por estarem sempre ao meu lado.

Ao meu irmão e amigo, Vítor, pela parceria de sempre.

Ao Gabriel, pelo companheirismo, carinho e por acreditar tanto em mim.

Aos meus amigos e toda a minha família pelo suporte, incentivo, compreensão e estímulo constante.

À Universidade Federal de Minas Gerais e ao Programa Interunidades de Pós Graduação em Bioinformática, pela qualidade do ensino.

À Universidade Federal de Viçosa e ao Programa de Pós Graduação em Bioquímica Aplicada, pela infraestrutura disponibilizada para o desenvolvimento desta dissertação.

Ao meu orientador, Tiago Mendes, pelos ensinamentos, confiança e dedicação durante esses dois anos.

À Rafaela Ventorim e ao Túlio Morgan, pela enorme contribuição na execução deste trabalho.

A todos os colegas do LBM que me ensinaram e ajudaram tanto durante o mestrado, em especial à Grazi, Ananda, Lucas Maia, Lucas Filipe, Renato, Higor, Edson, Larissa, Igor e Lis.

Aos secretários do PPG da Bioinformática, Tiago e Sheila, e ao Marco Aurélio, da Bioquímica (UFV), por serem sempre tão solícitos e por toda a boa vontade em ajudar.

Às agências financiadoras, CAPES, CNPq, FAPEMIG e Consórcio Café.

RESUMO

A ferrugem, causada pelo fungo *Hemileia vastatrix*, é uma das principais doenças do cafeeiro e gera grandes prejuízos econômicos. O *H. vastatrix*, por ser um fungo biotrófico, tem seu crescimento e reprodução totalmente dependentes das células do hospedeiro vivo. Sendo assim, infecta o tecido pelos estômatos, sem causar necrose. Apesar disso, sabe-se que alguns fungos durante a interação com a planta podem expressar genes envolvidos na formação de estruturas infecciosas, bem como sintetizar enzimas que auxiliam na penetração pela parede celular do hospedeiro. Pouco ainda se sabe sobre a importância das enzimas que degradam parede celular durante a infecção por fitopatógenos biotróficos. Aqui, realizamos uma análise genômica com o objetivo de identificar CAZymes (em inglês, Carbohydrate Active EnZymes) produzidas por *H. vastatrix*, além da análise da expressão gênica e atividade enzimática durante a infecção do cafeeiro. A primeira etapa foi a predição *ab initio* do genoma de *H. vastatrix*, utilizando o programa AUGUSTUS ajustado com dados de RNA-seq do próprio fungo para definição do modelo. A anotação funcional para identificar CAZymes foi realizada utilizando o programa dbCAN2. Dentre as 345 CAZymes encontradas, 162 pertencem às Glicosídeo Hidrolases (GH), incluindo 34 endoglucanases, xilanases e poligalacturonases das famílias GH 5, 9, 10, 12 e 28 - que podem atuar durante infecção fúngica degradando componentes da parede celular, representados por celulose, xilana e pectina, respectivamente. Parâmetros físico-químicos dessas enzimas foram preditos, incluindo identificação de resíduos catalíticos e domínios padrões de CAZymes na sequência primária e estrutura modelada. Além disso, análises de expressão das enzimas foram realizadas utilizando dados de RNA-seq de um genótipo de *C. arabica* suscetível à infecção, portanto apresentando interação compatível com o patógeno, nos tempos 0, 12, 24 e 96 horas após a inoculação (hai) com o fungo. Ensaios de atividade enzimática foram realizados para verificar se essas enzimas também atuam durante a infecção fúngica. As enzimas apresentaram características físico-químicas e estruturais típicas de enzimas ativas, incluindo a conservação dos sítios catalíticos. Genes que codificam endoglucanases e xilanases são expressos em diferentes momentos da infecção por *H. vastatrix*, sendo o primeiro mais expresso em 12 hai e o segundo em 24 hai. Em concordância com os dados de expressão, as endoglucanases começam a ser produzidas precocemente e são mais ativas a 24 hai, enquanto as xilanases têm maior atividade a 96 hai. Esses resultados sugerem que fungos biotróficos também utilizam enzimas ativas sob carboidratos durante a infecção. No caso de *H. vastatrix*, essas enzimas começam a ser produzidas na fase pré-haustorial da infecção

do fungo, mas são mais ativas na fase pós-haustorial, quando o haustório já começa a ser formado nas membranas do hospedeiro e na parede celular da planta.

Palavras-chave: Ferrugem da folha do café, CAZymes, carboidratos, xilanase, celulase, pectinase.

ABSTRACT

Coffee leaf rust, caused by the fungus *Hemileia vastatrix*, is a major disease that affects many coffee producers around the world, causing a huge economic loss. *H. vastatrix* is a biotrophic fungus, hence, its growth and reproduction are totally dependent on the cells of the living host. Because of that, they infect the tissue without causing necrosis. Nevertheless, it is known that some fungi during plant interaction can express genes involved in the formation of infectious structures, as well as synthesize enzymes responsible for the degradation of the host cell wall. However, little is known about the importance of cell wall degrading enzymes (CWDE) for biotrophic phytopathogens. In this work, we performed genomic analysis to identify CAZymes (Carbohydrate Active EnZymes) from *H. vastatrix* and the analysis of expression and enzyme activity during experimental coffee infection. The first step was the *ab initio* prediction of *H. vastatrix* genome, using AUGUSTUS program, which was adjusted with *H. vastatrix* RNA-Seq for model determination. Functional annotation to identify CAZymes was performed using dbCAN2 meta server. Among the 345 CAZymes found, 162 belong to Glycoside Hydrolases (GH) including 34 endoglucanases, xylanases and polygalacturonases from families GH 5, 9, 10, 12 and 28 - which may act during fungal infection degrading the highest prevalent components of cell wall, represented by cellulose, xylan and pectin, respectively. Physico-chemical parameters of these enzymes were checked, including identification of catalytic residues and standard CAZymes motifs. The enzyme structures were also modelled. Moreover, analysis of expression of enzymes were performed using RNA-seq data from a genotype of *C. arabica* susceptible to infection, therefore presenting compatible interaction with the pathogen, after 0, 12, 24 and 96 hours after inoculation (hai). Enzyme activity assays were carried out in order to verify if these enzymes were also active during the fungal infection. The enzymes showed typical physico-chemical characteristics and structures of active enzymes including conservation of the catalytic sites. Genes encoding endoglucanases and xylanases were expressed during different times of *H. vastatrix* infection, being the first more expressed at 12 hai and the other at 24 hai. In agreement with expression data, endoglucanases started to be produced early, and were more active at 24 hai, whereas xylanases had greater activity at 96 hai. The results suggest that biotrophic fungus also use active enzymes on carbohydrates during the infection. In relation to *H. vastatrix*, these enzymes start being produced at the pre-haustorial phase of the fungal infection, but are more active at the post-haustorial phase, when the haustorium begins to be formed in the host membranes and in the plant cell wall.

Keywords: coffee leaf rust, cazymes, carbohydrates, xylanase, cellulase, pectinase.

Sumário

1. REVISÃO DE LITERATURA

1.1. Café

A história da origem do café não é muito precisa, mas registros históricos de 575 d.C. permitem estimar que esse grão tenha sido observado pela primeira vez na província de Kaffa, na Etiópia (ABIC, 2021; ICOb, 2021). Seu processo de cultivo, no entanto, teve início nos monastérios islâmicos do Iêmen e a bebida da forma como é conhecida hoje, e sua consequente popularização, foi possibilitada somente após o desenvolvimento do seu processo de torrefação, na Pérsia, no século XIV (ABIC, 2021; CECAFE, 2021).

O café começou a ser altamente difundido para outras regiões no início do século XVI, quando viajantes europeus conheceram o grão durante as navegações ao Oriente em busca de produtos exóticos que pudessem ter valor na Europa (CECAFE, 2021; SANTIAGO, 2021). Cerca de um século depois, em 1727, as primeiras mudas de café foram trazidas para o Brasil por um oficial português vindo da Guiana Francesa (ABIC, 2021). As mudas foram plantadas inicialmente em Belém do Pará e devido às condições climáticas favoráveis, rapidamente seu cultivo se propagou para outras regiões do Brasil, principalmente no Sudeste, tornando-se em 1830 o principal produto exportado no país. No entanto, uma geada no centro-sul brasileiro marcou o primeiro abalo sofrido pela indústria do café, em 1870. O segundo, que resultou no fim do auge do ciclo do café, foi em 1929 devido à quebra da bolsa de Nova York (CECAFE, 2021).

Atualmente o café não é mais o principal produto de exportação no Brasil. Apesar disso, o país ainda é o maior exportador mundial do produto, tendo sido responsável em 2019 por cerca de um terço da produção do planeta, com aproximadamente 60 milhões de sacas produzidas (ICOa, 2021). Além disso, no mesmo período, foi o segundo maior consumidor de café, atrás apenas dos EUA (ICO, 2020). De acordo com o relatório de acompanhamento da safra brasileira de café do último trimestre de 2020 da CONAB (2021), a área cultivada com café arábica e canéfora foi de 2,16 milhões de hectares, e o principal estado produtor foi Minas Gerais, com 34,65 milhões de sacas.

Dentre as espécies de cafeeiro, duas se destacam pela sua importância econômica: *Coffea arabica* (Café Arábica) *e Coffea canephora* (Café Canéfora) (PARTELLI *et al.*, 2020). O café Arábica, cultivado em regiões de altitude acima de 800 m, representou 77% da safra brasileira no quarto trimestre de 2020. Enquanto que o Canéfora, predominante em regiões de temperaturas mais altas e de baixa altitude, foi responsável por 23% (CABRAL *et al.*, 2016;

CONAB, 2021). Apesar do *C. canephora* apresentar muitas qualidades agronômicas desejadas e ser altamente resistente a pragas (MCCOOK; VANDERMEER, 2015), seu sabor e qualidade são inferiores ao *C. arabica* e por este motivo, esta última espécie é a mais cultivada (RIBEIRO *et al.*, 2014).

Dada toda a importância do café para a economia brasileira, alguns fatores bióticos e abióticos que afetam a qualidade e a produção deste grão devem ser observados. As condições favoráveis para o cultivo do café incluem temperaturas entre 15ºC a 24ºC para o arábica e até 30ºC para o canéfora, além de um pH ótimo entre 4.5 e 5.5 (MAGRACH; GHAZOUL, 2015; NAIK *et al.*, 2019). Dessa forma, alguns estresses abióticos são devidos a ocorrência de geadas, mesmo que esporádicas, bem como exposição direta ao sol e a altas temperaturas. Há também o estresse hídrico, causado por limitação de chuva e baixa capacidade de retenção de umidade do solo (NAIK *et al.*, 2019). Já grandes quantidades de chuva não causam danos tão significativos ao plantio do café, desde que haja sistemas de drenagem adequados. Todavia, áreas com precipitação maior que 3000mm/ano podem gerar consequências secundárias como a erosão e dificuldade de secagem dos grãos ao sol (DESCROIX; SNOECK, 2009). Assim, como a produção do café é muito sensível a esses fatores abióticos, há uma preocupação em relação às mudanças climáticas, cujas projeções apontam para a redução das áreas cultiváveis, consequentemente gerando um grande impacto econômico (MAGRACH; GHAZOUL, 2015).

Já os fatores bióticos incluem algumas pragas que afetam significativamente a produção de café no mundo, como a mariposa do café ou bicho mineiro (*Leucoptera coffeella*), que cria lesões necrotróficas na folha, provocando sua queda e reduzindo a eficiência fotossintética (AGROLINK, 2021). O fungo *Colletotrichum kahawae* causa a antracnose do cafeeiro e leva à formação de manchas escuras nos frutos verdes e em consequência, à sua queda precoce (SILVA *et al.*, 2006). O fungo *Fusarium xylarioides* causa a murcha do cafeeiro, doença que bloqueia o sistema vascular da planta, limitando o fornecimento de água e culminando na morte da planta (PFENNING; MARTINS, 2000). No entanto, a doença mais grave que afeta a indústria cafeeira é a ferrugem da folha do café, já que 95% das variedades plantadas são suscetíveis à doença (CABRAL *et al.*, 2016; CHEMURA; MUTANGA; DUBE, 2016). A ferrugem é causada pelo fungo *H. vastatrix* e seus sintomas começam com o aparecimento de pequenos pontos amarelados na folha, que vão se expandindo, formando massas de esporos alaranjados na superfície inferior da folha e conduzindo à sua queda (ALFONSI *et al.*, 2019a; TALHINHAS *et al.*, 2017).

1.2. A ferrugem do café

A ferrugem do café foi identificada pela primeira vez em 1869 no Lago Vitória, no Quênia (ALFONSI *et al.*, 2019a), mas emergiu como um grande problema 8 anos mais tarde, no Ceilão (Sri Lanka). Esta região possui um clima úmido que favoreceu a propagação do fungo, levando à destruição de cerca de 90% da safra de café (CRESSEY, 2013). No Brasil, constatou-se a primeira aparição da ferrugem do café em *C. arabica* em janeiro de 1970, na Bahia. Estima-se que hoje a doença cause perdas na produção que variam entre 30 e 50%, de acordo com o nível de resistência do genótipo (ZAMBOLIM, 2016).

Figura 1: Início da infecção da folha de café por *H. vastatrix*. Os sintomas começam com o aparecimento de pequenos pontos amarelados na superfície inferior da folha, que vão se expandindo, formando massas de esporos alaranjados e conduzindo à sua queda. **Fonte:** 3R LAB, 2021.

Mais de 50 raças diferentes do fungo já foram descritas no mundo, e no Brasil, foram identificadas 15: I, II, III, VII, X, XIII, XV, XVI, XVII, XXI, XXII, XXIII, XXIV, XXV ou XXXI e XXXVII (CAPUCHO *et al.*, 2012, PORTO *et al.*, 2019). O uso de cultivares resistentes foi até hoje a forma mais sustentável e menos danosa ao meio ambiente e aos produtores encontrada de controlar a doença, porém tem sido observado que novas raças identificadas do fungo são capazes de suplantar essa resistência. Apesar de a raça II, que infecta variedades de *C. arabica* não resistentes ser a mais predominante na maioria dos países (CRISTANCHO *et al.*, 2014), outras raças são objeto de estudos mais aprofundados por infectarem variedades que foram inicialmente lançadas como resistentes, mas tiveram a resistência suplantada e passaram a ser suscetíveis. Um exemplo é a raça XXXIII, detectada recentemente em cultivares derivados dos principais doadores parentais de resistência (PORTO *et al.*, 2019).

O fungo *H. vastatrix* pertence ao filo *Basidiomycota* e à ordem *Puccinales*, que representa os fungos da ferrugem (SILVA *et al.*, 2012), e são patógenos biotróficos obrigatórios (RAMIRO *et al.*, 2009). Enquanto os fungos necrotróficos matam as células do tecido hospedeiro e se desenvolvem em tecido morto, os fungos biotróficos se desenvolvem no tecido sem causar necrose, pois dependem inteiramente do hospedeiro para seu crescimento e reprodução, o que ocorre por meio dos haustórios, estruturas que possibilitam a obtenção de nutrientes por meio do contato com as células vivas da planta (GARCÍA-GUZMÁN; HEIL, 2013, RAMIRO *et al.*, 2009).

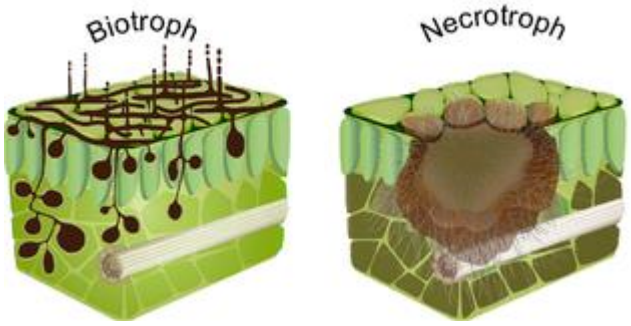


Figura 2: Dois estilos de vida de fungos que infectam plantas: biotróficos (que desenvolvem estruturas infecciosas no tecido vivo) e necrotróficos (causam a morte do tecido do hospedeiro). **Fonte:** García-Guzmán e Heil (2013).

O ciclo de vida do *H. vastatrix* consiste em dois diferentes estágios: o primeiro dicariótico, em que são formadas as uredínias (que darão origem aos urediniósporos), e o segundo monocariótico, representado pelos basidiósporos, que não têm ainda um papel bem conhecido no ciclo de vida deste patógeno (FERNANDES, 2007; RAMIRO *et al.*, 2009).

O mecanismo infeccioso pode ser visto na Figura 1. Existe uma variação em relação aos tempos de duração de cada etapa desse processo, de acordo com a suscetibilidade do genótipo em estudo e diferentes raças do patógeno. Desse modo, o mecanismo inicia-se com a adesão dos urediniósporos (UR) na folha do café, seguida de sua germinação e o alongamento de um tubo germinativo a fim de encontrar um estômato, por onde acontece a penetração. Considerando genótipos de café suscetíveis ao *H. vastatrix* de diferentes raças e que compõem interação compatível, esse processo pode ocorrer de 12 hai a 18 hai (SILVA *et al.*, 2002; SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). Nesse momento, há a diferenciação de um apressório (AP) maduro, estrutura que fixa o fungo à planta e que se diferencia em hifa primária (HiP) (TALHINHAS *et al.*, 2017). A HiP é formada por volta das 24 hai, penetrando o estômato e

alcançando a câmara subestomática (SILVA *et al.*, 2002; SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). Essa hifa gera então duas ramificações laterais em formato de âncora (AN), que formam a célula mãe do haustório (CMH) de 24 a 36 hai (SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). A CMH origina, por sua vez, o haustório primário (HP) dentro das células subsidiárias, de 36 a 48 horas após a infecção (SILVA *et al.*, 2002; SILVA *et al.* 2008). A partir daí o fungo continua se desenvolvendo, há a formação de novas âncoras, CMHs e os haustórios secundários (HS) são produzidos nas células do mesófilo a partir de 36 hai (RAMIRO *et al.* 2009).

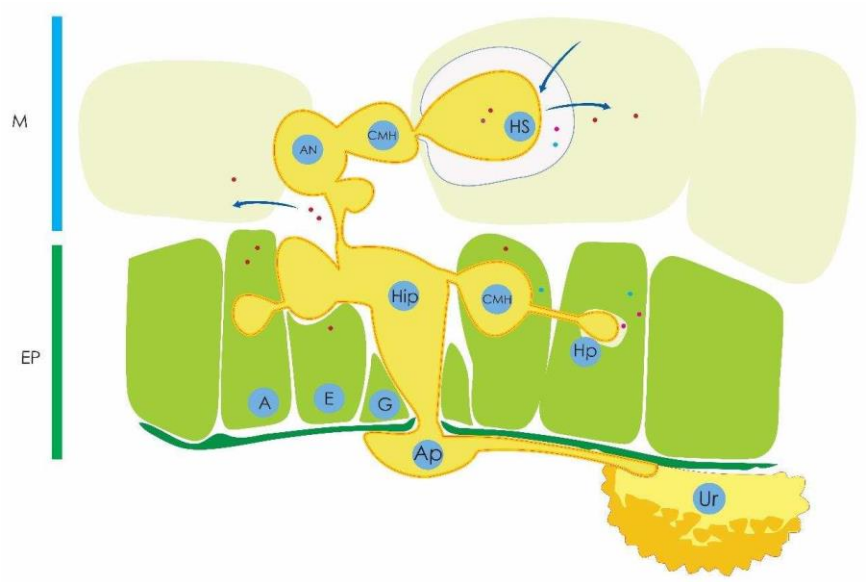


Figura 3: Representação esquemática do mecanismo de infecção do fungo *H. vastatrix* nos tecidos da folha do cafeeiro. O processo se inicia pela adesão de urediniósporos à folha, seguida da penetração do apressório pelos estômatos, até o desenvolvimento dos haustórios - estruturas de nutrição do fungo. Abreviações - Ur: urediniósporo, Ap: apressório, Hip: haustório primário, AN: âncora, CMH: célula mãe do haustório, HP: haustório primário, A: célula subsidiária, E: estômato, G: células guarda, HS: haustório secundário, M: mesófilo, EP: epiderme. **Fonte:** RAMIREZ (2015 apud RAMIRO *et al.*, 2009).

1.3. Parede celular vegetal

A parede celular vegetal é constituída em sua maior parte (mais de 90%) por uma fração polissacarídica que engloba celulose, hemicelulose e pectina, e uma outra não polissacarídica, formada por lignina, compostos fenólicos e proteínas (KUMAR; NARAIAN, 2019; MOTA *et*

al., 2018). A celulose é o principal componente da parede celular das plantas e é formada por monômeros de glicose unidos entre si por ligações β-1,4 glicosídicas, que resultam num arranjo cristalino. A celulose tem um papel de sustentação da parede. Os arranjos de celulose são interconectados com cadeias de hemicelulose, que por sua vez, são formados da união de diferentes monômeros como xilose, arabinose e galactose, e por isso, resultam em estruturas amorfas. Esses arranjos descritos estão incorporados numa matriz de pectina, um carboidrato mais complexo, rico em ácido galacturônico (MADEIRA JUNIOR *et al.*, 2017; VORWERK; SOMERVILLE; SOMERVILLE, 2004). Por fim, a lignina é um polímero formado por álcoois aromáticos, que aumenta a rigidez da planta e a resistência à entrada de água e microorganismos (NATIONAL HIGH MAGNETIC FIELD LABORATORY, 2021).

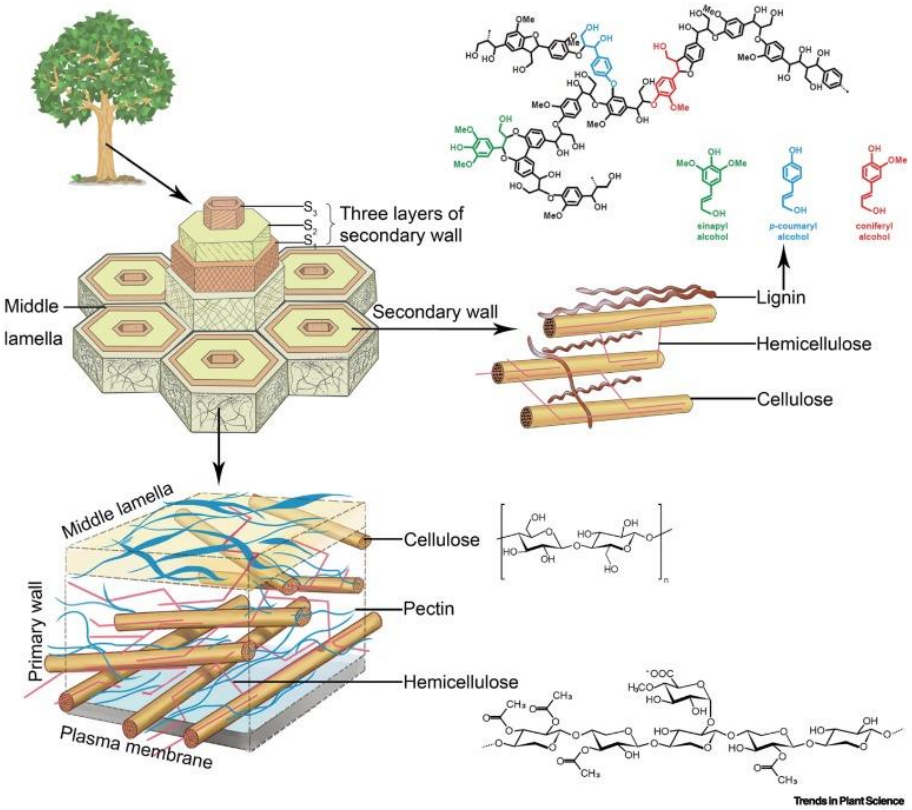


Figura 4: Representação das três principais camadas da parede celular vegetal: as paredes primária e secundária e a lamela média que une as células vegetais entre si. A parede celular primária precisa ser flexível e resistente ao mesmo tempo, para permitir o crescimento celular e é constituída por celulose (representada em tubos marrons), hemiceluloses (fibrilas vermelhas) e pectina (fibrilas em azul). A parede secundária é mais espessa e rígida e contém, além desses três componentes, também lignina (polímero formado por álcoois aromáticos). **Fonte:** Zhao *et al.* (2019).

1.4. Enzimas relacionadas à degradação da parede celular vegetal

Para que fitopatógenos consigam colonizar seu hospedeiro de maneira efetiva, precisam atravessar a parede celular vegetal, que contém os 3 polissacarídeos principais: celulose, hemicelulose e pectina. Por esse motivo, um arsenal enzimático que auxilia na hidŕolise desses carboidratos é encontrado em diversos fungos, incluindo biotróficos e necrotróficos, sendo significativamente maior para o segundo grupo (KUBICEK; STARR; GLASS, 2014; ZHAO *et al*., 2013).

Ainda que seja descrito na literatura que o processo de infecção do café por *H. vastatrix* se inicie pelos estômatos (aberturas naturais da planta), há evidências de que outros fungos biotróficos e hemibiotróficos apresentam enzimas acessórias durante a invasão dos tecidos para que se espalhem mais rapidamente sem causar muitos danos ao hospedeiro (GIBSON *et al.*, 2011).

As principais enzimas que degradam os carboidratos essenciais da parede celular vegetal, encontradas nos fungos, são celulases, hemicelulases e pectinases (ZHAO *et al*., 2013). O complexo de enzimas responsável pela hidrólise de celulose, é composto por endoglucanases, exoglucanases e beta-glucosidases (NAVYA; PUSHPA, 2012). A primeira é responsável pela quebra das ligações β-1,4 glicosídicas na cadeia de celulose, e as exoglucanases hidrolisam sua fração cristalina, liberando celobiose, que será degradada pelas β-glucosidases (FARINAS, 2011). Endo- β-1,4-xylanases são as principais e mais abundantes hemicelulases existentes, e são responsáveis pela hidrólise das ligações β-1,4 glicosídicas da xilana, liberando como subprodutos diversos oligossacarídeos, como xilose, xilobiose e xilotriose (KUMAR; NARAIAN, 2019). Já a degradação da pectina, é um processo mais complexo e requer a ação de diversas enzimas, entre elas, poligalacturonases e rhamnogalacturonases (MADEIRA JUNIOR *et al.*, 2017).

1.5. CAZymes

Há uma enorme diversidade de carboidratos encontrados na natureza, que desempenham diferentes funções, e são controlados por enzimas responsáveis por desde a sua síntese até a quebra. Essas enzimas são designadas como CAZymes: enzimas ativas em carboidratos (do inglês: Carbohydrate-Active enZymes) e são divididas em 5 principais classes: i) Glicosídeo Hidrolases (GHs) - agem na hidrólise e rearranjo de ligações glicosídicas; ii) Glicosil

Transferases (GTs) - auxiliam na formação das ligações glicosídicas; iii) Polissacarídeo Liases (PLs) - quebra sem hidrólise das ligações glicosídicas; iv) Carboidrato Esterases (CEs) - hidrólise de ésteres de carboidrato; v) Atividades Auxiliares (AAs) - enzimas que atuam na degradação de lignina, que foram adicionadas posteriormente, justamente porque a lignina é sempre encontrada junto com os polissacarídeos da parede celular vegetal (CAZy, 2020; LOMBARD *et al.*, 2013).

Dentro dessas classes, as CAZy são ainda subdivididas em famílias, que reúnem enzimas que têm em comum o mecanismo catalítico e que se dobram de forma similar, apresentando assim um bom potencial preditivo em se tratando das novas enzimas que vêm sendo constantemente descobertas (GARRON; HENRISSAT, 2019).

As principais enzimas que degradam parede celular vegetal, pertencem às Glicosídeo Hidrolases, em sua maioria. Das celulases, há 12 famílias que contemplam as β-1,4-endoglucanases: GH 5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74 e 124 (KUMAR; NARAIAN, 2019), sendo a GH5 a maior delas. Zhao *et al*. (2013) demonstraram que enzimas pertencentes à GH5 parecem ter um papel importante na degradação de materiais lignocelulósicos, uma vez que foi a família mais comum encontrada nos fungos examinados com potencial de degradação da parede celular vegetal.

Já as endo- β-1,4-xylanases pertencem em maior parte às famílias GH10 e GH11, que são bem diferentes entre si em diversos aspectos, tais como a massa molecular que é maior para as GH10 e os pontos isoelétricos que são mais ácidos para enzimas desta família e mais básicos para as GH11 (KUMAR; NARAIAN, 2019). Além disso, as xilanases GH10 se dobram numa conformação TIM-barrel e apresentam uma maior versatilidade catalítica que as GH11, de conformação β-jelly roll. Desta forma, as GH10 agem numa maior diversidade de substratos e diferentes subprodutos são obtidos desta hidrólise (BIELY *et al.*, 1997).

A pectina, por ser uma estrutura mais complexa, tem uma maior diversidade de enzimas envolvidas na sua hidrólise. Por exemplo, as poligalacturonases e as rhamnogalacturonases que pertencem à família GH28, as pectina e pectato liases, pertencentes à PL1-4 e a pectina metilesterase, que pertence à CE8 (MADEIRA JUNIOR *et al*., 2017).

A variedade e o número dessas enzimas é diferente para cada grupo de fungos. Os biotróficos, hemibiotróficos e necrotróficos apresentam, respectivamente, uma quantidade crescente do número de CAZymes (ZHAO *et al*., 2013).

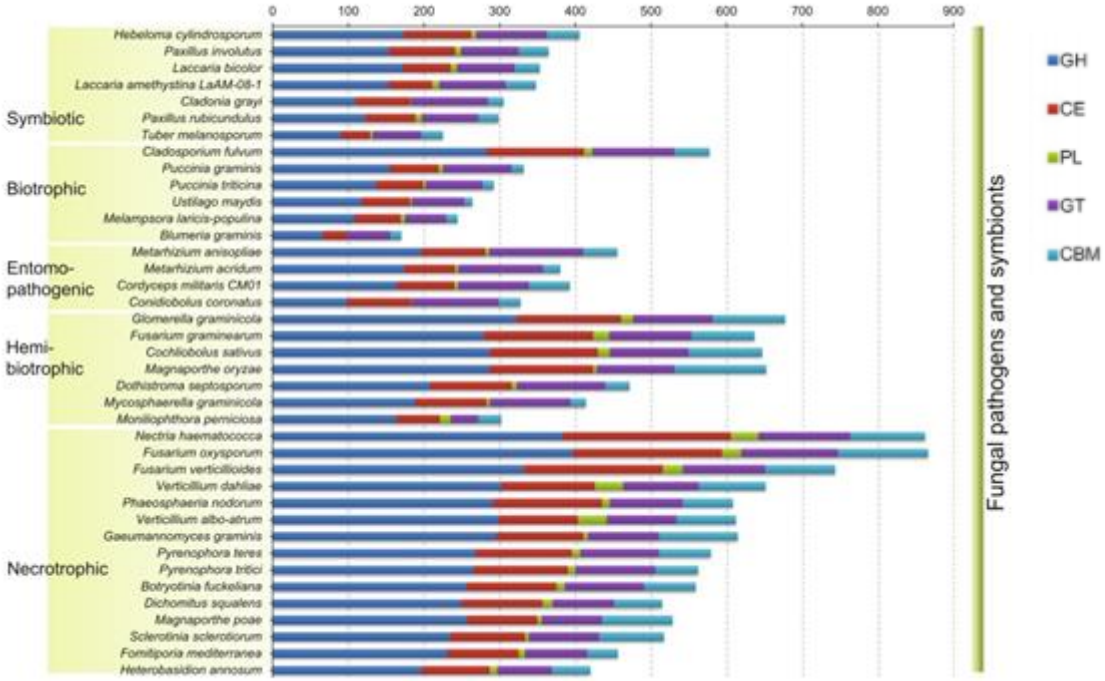


Figura 5: Gráfico mostrando a representatividade das classes de CAZymes em diferentes fungos e patossistemas. Apesar da quantidade de CAZymes ser diferente, o grupo das GHs é o maior em todos os casos e os fungos necrotróficos são os que apresentam maior quantidade dessas enzimas. **Fonte:** ZHAO *et al* (2013).

1.6. Estudos sobre CAZymes em fungos fitopatogênicos a nível de genoma, transcriptoma e proteoma

Dentre todos os tipos de CAZymes que podem ser produzidas por fungos, as que estão envolvidas na degradação da parede celular vegetal têm sido objeto de diversos estudos, dada sua importância para a efetiva infecção do hospedeiro. Com um maior número de genomas de fungos sequenciados nos últimos tempos, foram descobertos novos genes que codificam CAZymes que degradam parede celular de plantas e estão envolvidos na sua infecção (ZHAO *et al*., 2013). No entanto, ainda pouco se sabe sobre a função dessas enzimas durante a infecção e em quais momentos elas estão presentes, ou se estão sendo expressas nesses fungos (HERBERT *et al.*, 2004)

Alguns estudos de expressão de genes codificadores de CAZymes foram feitos para fungos hemibiotróficos. Foi demonstrado por Herbert *et al.* (2004) que o fungo *Colletotrichum*

lindemuthianum apresenta atividade enzimática durante o estágio biotrófico, direcionada à degradação de biopolímeros da parede celular do hospedeiro. Zhao *et al*. (2013) também conclui que os perfis de expressão do fungo *Fusarium graminearum* mostram genes codificadores de CAZymes com expressão aumentada durante a infecção à planta. Além disso, análises bioquímicas e moleculares de fungos de ferrugem demonstraram a presença de enzimas envolvidas com uma quebra localizada da parede celular vegetal, sugerindo que esses fungos possuem mecanismos que permitem a lise da parede celular do hospedeiro, facilitando a penetração (VOEGELE; HAHN; MENDGEN, 2009). Rauscher, Mendgen e Deising (1995) demonstraram, num estudo com o fungo biotrófico *Uromyces viciae-fabae* - responsável pela ferrugem da ervilha - que pressão celular (turgor) e a secreção controlada de enzimas líticas agem em conjunto para que haja uma penetração bem-sucedida na parede celular do hospedeiro.

Os genomas de fungos de ferrugem são muito grandes quando comparados a fungos não biotróficos, especialmente o de *H. vastatrix*, que é considerado um dos maiores deste grupo (PORTO *et al.*, 2019). Por este motivo, somente recentemente foi realizado o sequenciamento do genoma do fungo da ferrugem do café, possibilitando estudos mais aprofundados sobre sua interação com o hospedeiro. Cristancho *et al.* (2014) foi o primeiro grupo a depositar a montagem do genoma (333 Mb) com base num conjunto de 8 diferentes isolados de *H. vastatrix* (HvCat) utilizando as plataformas Illumina e 454. Porto *et al.* (2019) foram responsáveis pelo depósito do genoma da raça XXXIII (Hv33), de 547 Mb, combinando estratégias de long reads (PacBio) e de short reads (Illumina), para alcançar uma alta qualidade do sequenciamento. No entanto, a anotação destes genomas não está publicamente disponível, o que dificulta a investigação da existência de CAZymes que possivelmente estão sendo codificadas, e também se e como atuam durante a infecção da folha de café.

Talhinhas *et al.* (2014) avaliaram alguns mecanismos moleculares associados ao patossistema *H. vastatrix* - cafeeiro por meio da análise transcriptômica (sequenciado por 454), durante duas fases iniciais da infecção: germinação de urediniósporos e formação do apressório. Esse estudo incluiu ainda uma comparação com dados de outro trabalho previamente publicado (FERNANDEZ *et al.*, 2011) referentes a um estágio mais tardio da infecção, quando haustórios já foram formados. Por meio de uma comparação por blastp das proteínas preditas com a base de dados CAZy, eles conseguiram identificar 148 CAZymes putativas codificadas no genoma de *H. vastatrix*. Foi observado que a maioria dos genes que codificam essas enzimas estão sendo expressos nas fases iniciais de infecção, e a maior parte delas pertence à família das Glicosídeo Hidrolases.

Apesar de saber que existem genes que codificam CAZymes no genoma de *H. vastatrix*, ainda não foi comprovado que estes realmente produzem as enzimas, se elas são secretadas e se desempenham alguma função durante a infecção, e se sim, qual é essa função. Além disso, se faz necessário um estudo mais aprofundado para um melhor entendimento do papel de cada família de CAZymes durante a infecção, e se esse papel está relacionado à degradação de carboidratos da parede celular vegetal.

2. PERGUNTAS E HIPÓTESE

Propusemos as seguintes perguntas neste trabalho: sendo o *H. vastatrix* um fungo biotrófico, (i) ele expressa genes codificadores de CAZymes durante a infecção e (ii), caso sejam de fato produzidas, as enzimas estarão ativas?

Hipóteses a serem testadas: (i) as principais enzimas que degradam parede celular vegetal são produzidas e (ii) estão ativas durante a infecção do fungo *H. vastatrix* à folha de café.

3. OBJETIVOS

3.1. Objetivo geral

Predizer e identificar genes codificadores de CAZymes e avaliar se essas enzimas estão sendo produzidas e estão ativas durante a infecção do fungo *H. vastatrix* à folha de café.

3.2. Objetivos específicos

- Predizer e identificar genes codificadores de CAZymes no genoma do fungo *H. vastatrix*;
- Avaliar padrões de conservação e diversidade nas sequências e estruturas modeladas das enzimas;
- Avaliar padrões de expressão dos genes codificadores dessas enzimas nas diferentes fases de infecção do fungo;
- Avaliar a atividade das enzimas selecionadas ao longo do processo de infecção.

O texto foi preparado no formato de manuscrito a ser submetido ao periódico Fungal Genetics and Biology (Fator de impacto: 3.071)

4. ARTIGO

Genomic identification, gene expression and enzyme activity of CAZymes from biotrophic fungus *Hemileia vastatrix* during coffee leaf infection

Júlia Santos Pereira[1,2], Túlio Morgan[1,2], Isabel Samila Lima Castro[2], Lucas Moreira Maia[1], Rafaela Zandonade Ventorim[3], Danúbia Rodrigues Alves[4], Dênia Pires de Almeida[4], Eveline Teixeira Caixeta[4,5], Tiago Antônio de Oliveira Mendes[1,2,4*]

1 Programa de Pós-graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brasil

2 Departamento de Bioquímica e Biologia Molecular, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brasil

3 Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brasil

4 Instituto de Biotecnologia Aplicada à Agropecuária - BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, 36590-900. Brasil

5 Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA Café, Brasília, DF, 70770-901, Brasil

Corresponding author.

*E-mail address: tiagoaomendes@ufv.br

Keywords: coffee leaf rust, cazymes, carbohydrates, xylanase, cellulase, pectinase

Abstract

Coffee leaf rust, caused by the fungus *Hemileia vastatrix*, is a major disease that affects many coffee producers around the world, causing a huge economic loss. *H. vastatrix* is a biotrophic fungus, hence, its growth and reproduction are totally dependent on the cells of the living host. Because of that, they infect the tissue without causing necrosis. Nevertheless, it is known that some fungi during plant interaction can express genes involved in the formation of infectious structures, as well as synthesize enzymes responsible for the degradation of the host cell wall. However, little is known about the importance of cell wall degrading enzymes (CWDE) for biotrophic phytopathogens. In this work, we performed genomic analysis to identify CAZymes (Carbohydrate Active EnZymes) from *H. vastatrix* and the analysis of expression and enzyme activity during experimental coffee infection. The first step was the *ab initio* prediction of *H. vastatrix* genome, using AUGUSTUS program, which was adjusted with *H. vastatrix* RNA-Seq for model determination. Functional annotation to identify CAZymes was performed using dbCAN2 meta server. Among the 345 CAZymes found, 162 belong to Glycoside Hydrolases (GH) including 34 endoglucanases, xylanases and polygalacturonases from families GH 5, 9, 10, 12 and 28 - which may act during fungal infection degrading the highest prevalent components of cell wall, represented by cellulose, xylan and pectin, respectively. Physico-chemical parameters of these enzymes were checked, including identification of catalytic residues and standard CAZymes motifs. The enzyme structures were also modelled. Moreover, analysis of expression of enzymes were performed using RNA-seq data from a genotype of *C. arabica* susceptible to infection, therefore presenting compatible interaction with the pathogen, after 0, 12, 24 and 96 hours after inoculation (hai). Enzyme activity assays were carried out in order to verify if these enzymes were also active during the fungal infection. The enzymes showed typical physico-chemical characteristics and structures of active enzymes including conservation of the catalytic sites. Genes encoding endoglucanases and xylanases were expressed during different times of *H. vastatrix* infection, being the first more expressed at 12 hai and the other at 24 hai. In agreement with expression data, endoglucanases started to be produced early, and were more active at 24 hai, whereas xylanases had greater activity at 96 hai. The results suggest that biotrophic fungus also use active enzymes on carbohydrates during the infection. In relation to *H. vastatrix*, these enzymes start being produced at the pre-haustorial phase of the fungal infection, but are more active at the post-haustorial phase, when the haustorium begins to be formed in the host membranes and in the plant cell wall.

4.1. Introduction

Global coffee market was valued at approximately 102 billion US dollars in 2019 (RESEARCH AND MARKETS, 2020) and the world coffee consumption was around 168 million bags of 60kg in the same year (ICO, 2020). Furthermore, coffee is produced in more than 50 countries around the world (NATIONAL COFFEE ASSOCIATION, 2020) and it is the source of income of more than 100 million people (TALHINHAS *et al.*, 2017).

Coffee leaf rust (CLR), caused by the fungus *Hemileia vastatrix (H. vastatrix)*, is one of the major diseases of coffee plantations. It was first observed in 1861 in wild coffee trees near Lake Victoria in Kenya (East Africa) and, in the 1860s, caused destruction on coffee crops in Ceylon (Sri Lanka). After 1900, epidemics of CRL were detected in Africa, Philippines, Madagascar and Americas, showing that rust had spread to most coffee-growing countries worldwide (ZAMBOLIM; CAIXETA, 2021). In Brazil, it was first detected in January 1970, infecting *Coffea arabica* crops in the southern region of Bahia state (ZAMBOLIM, 2016). Today, it is estimated that rust causes losses of one to two billion US dollars annually (MCCOOK, 2006).

The severity of this disease is associated with the speed reproduction of the pathogen, which is influenced by temperature - being urediniospores germination rate optimum in 23ºC. In regions of Brazil that experience warm weather, the disease can reduce yields by 30 to 50% depending on some factors such as the level of resistance of the plant genotype and the level of epidemics in the previous year (ZAMBOLIM, 2016).

The disease symptoms start with the appearance of small patches of pale yellow color in the leaves, which expand from around 1 to 20mm in diameter. The chlorotic spots precedes the differentiation of uredinia, where the urediniospores, the main primary inoculum of the coffee leaf rust, are produced (ZAMBOLIM, 2016).

The growth and reproduction of *H. vastatrix* is totally dependent on the cells of the living host, since they are a biotrophic fungus. For this reason, the coffee rust fungus presents a colonization process that differs from the necrotrophic and endosymbiotic fungi (PANSTRUGA; DODDS, 2009). This pathogen life cycle consists of two different stages. The first is a dikaryotic stage, in which urediniospores are formed as structures of dispersion and infection. In the monokaryotic stage is produced basidiospores, in which coffee infectivity has not yet been demonstrated. Thus, the dikaryotic phase plays the most important role during infection (GOLD; MEDGEN, 1991; RAMIRO *et al.*, 2009).

The infectious mechanism can be divided in some steps, each one with its period of time varying according to genotype susceptibility and pathogen race. It starts with urediniospores adhesion in the coffee leaf followed by the elongation of the germinative tube until it meets a stomata, where the penetration happens. Considering coffee genotypes susceptible to *H. vastatrix* of different races, which will compose compatible interaction with the host, that process occurrs from 12 h to 18 hai (SILVA *et al.*, 2002; SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). At this point, there is the differentiation of a mature appressorium, an accessory that fixes the plant to the fungus, which will in turn differentiate into primary hypha (PHy) (TALHINHAS *et al.*, 2017). According to Silva *et al.* (2002), at 24 post-infection, the percentage of spore germination is around 74%, whereas an average 54% of the appressorium is formed over the stomata of a susceptible coffee genotype. The PHy is formed around 24 hai, penetrating the stomata and reaching the substomatic chamber (SILVA *et al.*, 2002; SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). This primary hypha will then generate two anchor-like lateral ramifications, forming the haustorium mother cell (HMC) 24 h to 36 hai (SILVA *et al.*, 2008; RAMIRO *et al.*, 2009). HMC gives rise to the primary haustorium (PH) inside subsidiary cells, 36 h to 48 hai (SILVA *et al.*, 2002; SILVA *et al.*, 2008). Then, the fungus keeps developing, forming new anchors, HMC's and secondary haustoriums (SH), produced at the mesophyll cells from 36 hai (RAMIRO *et al.*, 2009).

Many microorganisms such as fungus produce and secrete enzymatic cocktails with activity under vegetal biomass during colonization resulting in degradation of the primary and secondary plant cell wall. The primary cell wall needs to be flexible and at the same time mechanically stable to allow cell expansion without breaking them. The primary cell wall is constituted by cellulose, hemicelluloses and pectin. The secondary cell wall requires a higher mechanical stability and for this reason, it contains lignin in addition to the other components (TAIZ; ZEIGER, 2015). Cellulose is a linear biopolymer formed by monomers of D-glucose linked by β-1,4 glycosidic bonds, presenting a crystalline structure. In contrast, hemicelluloses are formed by pentoses and hexoses linked in an amorphous arrangement. Pectin is a family of galacturonic acid-rich polysaccharides and it is structurally and functionally the most complex carbohydrate of plant cell wall (O'NEILL; ALBERSHEIM; DARVILL; MOHNEN, 2008; TAIZ; ZEIGER, 2015). Lignin, on the other hand, is composed of phenolic compounds, presenting an aromatic structure and chemical nature quite different from carbohydrates (TAIZ; ZEIGER, 2015).

Cellulose is hydrolysed by a cellulolytic enzyme system composed of endoglucanase, exoglucanase and beta-glucosidase, reducing cellulose to glucose units (NAVYA; PUSHPA, 2012). Endoglucanase acts over β-1,4 glycosidic bonds in the cellulose chain. Exoglucanase hydrolyzes crystalline cellulose, releasing cellobiose, which will be degraded by β-glucosidases (ZHANG *et al*., 2006). β-1,4-endoglucanases are classified in 12 different GH families: GH5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74 and 124 (KUMAR; NARAIAN, 2019).

The most abundant group of hemicelluloses is xylan, which is also the second-most abundant polysaccharide after cellulose (SAHA, 2003; KUMAR; NARAIAN, 2019). Because of that, its hydrolysis is essential in biomass degradation (DODD; CANN, 2009). Endo- β-1,4-xylanases are the enzymes responsible for xylan hydrolysis, and most of them belong to the families GH10 and GH11, which are quite different from each other in physicochemical and molecular terms. When compared to GH11, GH10 xylanases act on a greater diversity of substrates, since they present catalytic flexibility. Because of that, their subproducts are many different oligosaccharides such as xylose, xylobiose, xylotriose (KUMAR; NARAIAN, 2019).

The major pectin structures are very complex and its deconstruction includes the action of many enzymes, such as polygalacturonases and rhamnogalacturonases, which are included in the GH 28 family (MADEIRA JUNIOR *et al.*, 2017).

Biotrophic fungi such as *H. vastatrix* need to obtain nutrients from the host cells and developed mechanisms to infect them without causing necrosis. However, Bellincampi *et al.* (2014) showed that small quantities of cell wall degrading enzymes (CWDE) are produced by biotrophic fungus in order to soften and loosen the cell walls and to facilitate fungal invasion.

It was noticed that there is almost no information about how CWDE can act during *H. vastatrix* infection to the coffee leaf. Because of that, we proposed in this work a search for CAZymes encoded in the genome of *H. vastatrix*, followed by a screening of CWDEs which may be expressed and enzymatically active during the CLR infection.

4.2. Material and Methods

4.2.1. Gene Prediction and Genome Annotation

Since the protein-coding genes of *Hemileia vastatrix* were not publicly available, the program AUGUSTUS v.3.2.2 (STANKE; MORGENSTERN, 2005) was used to perform the *ab initio* gene prediction. Genes were predicted in the genome of *H. vastatrix* HvCat isolate (GenBank assembly accession number: GCA_003057935.1) using transcripts assembled from RNA-seq data from *H. vastatrix* HvCat (SRA accession: SRR1124793) as gene models. The output was a general feature format (GFF3) file, then a FASTA file with the amino acid sequences was obtained using in-house Python scripts.

The functional annotation of Carbohydrate-Active Enzymes (CAZymes) was performed using dbCAN2 v.8.0 (ZHANG *et al.*, 2018) with default parameters. The enzymes from families GH 5, 9 and 12 (endoglucanases), GH 10 (xylanases) and GH 28 (polygalacturonase) were also submitted to a BLASTp similarity search (version 2.9.0) against NCBI RefSeq database using the cutoff parameters E-value smaller than 10^{-10}, identity and coverage greater than 30% and 70%, respectively, to provide another source of evidence for protein annotation.

4.2.2. Coding regions mapping and structure prediction

4.2.2.1. Gene mapping

The structures of the genes of interest were observed by plotting their CDS (CoDing Sequence) and introns positions. First, an R script was used to filter the GFF3 file obtained from AUGUSTUS to get the beginning and the end of each region, from each gene. The R packages “gggenes” and “ggplot2” were used to plot the CDS and intron regions.

4.2.2.2. Multiple sequence alignment with reference enzymes

Multiple sequence alignments of the *H. vastatrix* CAZyme sequences against a well-characterized enzyme from the same family were performed using Clustal Omega v2.1 with default settings. PDB (Protein Data Bank) codes of the best matches for GH5.5 endoglucanases, GH9 endoglucanase, GH12 endoglucanases, GH10 xylanases and GH28 polygalacturonase, respectively: 3QR3 (LEE *et al.*, 2011), 1GA2 (MANDELMAN *et al.*, 2003), 2NLR (SULZENBACHER *et al.*, 1999), 1R85 (ZOLOTNITSKY *et al.*, 2004), 4C2L (ROZEBOOM *et al.*, 2013).

4.2.2.3. Prediction of signal peptides and protein physical and chemical parameters

SignalP 5.0 server (ARMENTEROS *et al.*, 2019) was used to predict signal peptide and cleavage sites in each predicted enzyme. The following tools used default parameters. NetNGlyc 1.0 Server (GUPTA; JUNG; BRUNAK, 2004) and DictyOGlyc 1.1 Server (GUPTA *et al.*, 1999) were used to predict N and O- Glycosylation Sites, respectively. ProtParam tool (GASTEIGER *et al.*, 2005) was used to predict the physical and chemical parameters of each protein including the number of amino acids, molecular weight and theoretical pI.

4.2.2.4. Homology modelling and quality checking

The Phyre2 v2.0 web server (KELLEY *et al.*, 2015) was used in the intensive mode to obtain the 3D models of the selected enzymes. The structures and stereochemistry quality was checked by ProSA-web (WIEDERSTEIN; SIPPL, 2007), and Saves 6.0 - Verify3D (BOWIE *ET AL.*, 1991; LUETHY *ET AL.*, 1992), ERRAT (COLOVOS;YEATES, 1993), Procheck (LASKOWSK *et al.*, 2001). Energy minimization was performed by YASARA Energy Minimization Server (KRIEGER *et al., 2009*), using the YASARA force field. The VMD v1.9.3 was used to analyse the three-dimensional structure of modeled proteins, as well as overlap the 3D structures.

4.2.3. RNA-seq analysis

The four libraries of the susceptible coffee genotype, inoculated with spores of *H. vastatrix* race XXXIII (*C. arabica* cv. caturra vermelho CIFC 19/1) used in this study - 0, 12, 24 and 96 hai - were obtained from a previous work (Florez *et al*, 2017) and are available at Sequence Read Archive (SRA) database (SRR5417837, SRR5417838, SRR5417839, SRR5417840). Read quality was checked with FastQC v0.11.5 and trimmed with Trimmomatic v0.36, using a phred+33 quality score cutoff of 25. Next the Tuxedo pipeline (TRAPNELL C. *et al.*,2012) was employed using the softwares: bowtie2 v2.2.8, bowtie v1.1.2, tophat v2.1.1, and cufflinks v2.2.1.

For transcriptome assembly, we used the genome of *H. vastatrix* HvCat as the reference, and the GFF file obtained from Augustus gene prediction as the reference annotation. The

expression of the enzymes of interest was measured by FPKM (fragments per kilobase of exon model per million reads mapped) and the heatmap was plotted in Tableau Software v2020.3.1.

4.2.4. Experimental infection and plant protein extraction

The susceptible cultivar *C. arabica* cv. Caturra CIFC 19/1 was transferred to an acclimatization room and was infected with urediniospore of *Hemileia vastatrix* race XXXIII as previously described by CAPUCHO *et al.*, 2009 and performed at the Coffee Biotechnology Lab (BioCafé/UFV) . Leaves without infection were used as control groups. For each time, three leaves were infected for plant protein extraction. In addition to time zero immediately before the infection, the leaves were collected at 12, 24 and 96 hours after inoculation (hai). First, the leaves were vacuum infiltrated according to Guerra-Guimarães *et al*. (2009). A square slice of about 2 cm² of leaves was vacuum impregnated in a 100 mM Tris-HCl buffer (pH 7.6) solution containing 50 mM L-ascorbic acid, 500 mM KCl and 25 mM Thermo Scientific Halt™ Protease Inhibitor Cocktail. The blotted sections were centrifuged at 5000 g, for 15 minutes at 4°C, and the collected protein extract was frozen. This fraction was desalted, concentrated and purified (GUERRA-GUIMARÃES *et al*., 2014). Bradford assay method was used to quantify the soluble proteins. In order to verify if the material was not degraded, an SDS-PAGE was carried out.

4.2.5. Enzyme activity assays

Endoglucanase, xylanase and polygalacturonase activities were measured by the release of reducing sugars from carboxymethylcellulose (CMC), beechwood xylan and polygalacturonic acid, respectively, using the dinitrosalicylic acid reagent (DNS).

Endoglucanase assay was performed in a 96-well plate containing 90ul of CMC (1.25% w/v) diluted in 100mM sodium phosphate buffer pH 5.0 and 10 µL of *H. vastatrix* protein extract. D-glucose (2 g/L) was used as standard.

Polygalacturonase assay was performed in a 96-well plate containing 90 µL of polygalacturonic acid (0.25% w/v) diluted in 100 mM sodium phosphate buffer pH 5.0 and 10 µL of *H. vastatrix* protein extract. Galacturonic acid (2 g/L) was used as standard.

Xylanase assay was performed in eppendorfs containing 100 µL of beechwood xylan (1% w/v) diluted in 100 mM sodium phosphate buffer pH 5.0 and 25 µL of *H. vastatrix* protein extract. D-xilose (2 g/L) was used as standard.

All reactions were carried out for 30 min at 30 ºC and were stopped with the addition of 100 µL of DNS reagent for the reactions in the plate, and 125 µL for those in eppendorfs. The plates and the eppendorfs were then boiled for 5 min and cooled to room temperature. The amount of reducing sugars was determined at 540 nm. One unit of activity was defined as the amount of enzyme that released 1 µmol/min of the reducing sugar under assay conditions. The graphs were plotted using GraphPad Prism 5.0 as mean values of triplicate assays, being zero time (plat with no infection) subtracted from the others to eliminate noise signal from plant material. The ANOVA method with Bonferroni correction was applied to evaluate the significance difference between experimental groups.

4.3. Results

4.3.1. CAZymes identified by dbCAN2 and homology modelling

The figure 1 shows the result of dbCAN2 search, in which 345 CAZymes were found in the predicted proteome of *H. vastatrix* HvCat, being the Glycoside Hydrolases (GH) the biggest group, followed by the GlycosylTransferases (GT), Carbohydrate Esterases (CE), Auxiliary Activities (AA), Polysaccharide Lyases (PL) and Carbohydrate Binding Modules (CBM).

Due to the importance and high abundance of GH in fungal phytopathogens, GH families have been studied more thoroughly. From the 162 GH submitted to BLASTp similarity search, 103 presented best hits with the previously assigned family, resulting in the distribution shown in Figure 2. The endoglucanases GH5, GH9 and GH12, xylanases GH10 and polygalacturonases GH28 were selected because they may act during fungal infection degrading the highest prevalent components of cell wall represented by cellulose, xylan and pectin, respectively.

The position of the CDS in the contigs were plotted (Figure 3), enabling the detection of fragmented genes and gene position within the contigs and helping to identify possible problems in the sequences. Some sequences were possibly incomplete, due to evidence such as the absence of start codon and signal peptide. For example, the xylanase gene g3152 (Figure 3b) did not show those structural signals and was located at the very beginning of a contig in the genome. By searching the transcript derived from this gene in the RNA-seq data we were able to find the missing fragment and correct the sequence. The physicochemical properties of proteins were also compatible with those expected for different families of glycosylhydrolases (Supplementary Table 1).

We modeled 34 protein sequences from GH families 5, 9, 10, 12 and 28, using the Phyre2 web server. These enzymes were then separated into groups based on the templates used to construct the models. Since the GH 5 endoglucanases group was composed of many and highly diverse sequences, we divided it into subfamilies. The enzymes of subfamily 5.5 (Figure 4a) aligned well with an endoglucanase from *Trichoderma reesei*, whose catalytic sites are two glutamate residues in the positions 239 and 350 (proton donor and nucleophile, respectively). Data showed that the nucleophile catalytic site was conserved in all the three enzymes, but the proton donor was conserved only in g7743. The same analysis was done with the other four groups. Endoglucanase GH 9 (Figure 4b), aligned to an endoglucanase from *Ruminiclostridium cellulolyticum* and three of the four catalytic sites were conserved. GH 12 endoglucanases (Figure 4c) aligned with the first predicted structure of a fungal GH 12 enzyme (from

Trichoderma reesei) and showed conservation only in the catalytic site (a nucleophile). GH 10 xylanases (Figure 4d) aligned with the endo-1,4-beta-xylanase from *Geobacillus stearothermophilus* and showed good conservation in the glutamate residues at the catalytic site. GH 28 polygalacturonase (Figure 4e) aligned with an endo-xylogalacturonan hydrolase from *Aspergillus tubingensis* and also showed conservation at the catalytic residues.

To visualize the structural similarities between the predicted models and the templates, we overlapped those structures using VMD. Figures 5a and 5b showed that the regions around the catalytic sites in predicted models showed a folding similar to their respective templates. On the other hand, Figure 5c shows that despite the structures aligned well, the regions around the catalytic sites did not overlap.

The programs used to check the stereochemical and structural parameters of the predicted models indicated an overall low quality for many structures. For this reason, they were submitted to YASARA Energy Minimization Server, which improved the structural quality of the models. It was observed that the most problematic structures were derived from incomplete sequences, like the previously mentioned xylanase g3152. The 3D model of this protein obtained with its complete sequence showed a better quality than the model from the incomplete sequence, even after energy minimization.

4.3.2. Analysis of expression of genes that encode CAZymes

In the conditions tested, about half of the endoglucanases genes of *H. vastatrix* were little or not expressed, showing less than 30 FPKM (Figure 6). Nevertheless, many GH5 endoglucanases genes were expressed after fungal inoculation. The ones of IDs g9307, g4848, g1441, g8871 and g9077 have a peak of expression in 12 hai that decreases in the next times (24 hai and 96 hai), whereas g4364 showed a peak at 24 hai. The bigger expression value (548 FPKM) was observed at 24 hai by g284. Just g3177 had a peak at 96 hai.

On the other hand, figure 6e shows that the 3 xylanases genes were expressed at 12 hai, but they present their peaks at 24 hai, except for g3152 that was expressed at this time. It is important to notice that the bigger expression value (230 FPKM) was also observed at 24 hai (g9612). At 96 hai, the expression value was zero for g1 and very low for the others.

None of these genes were expressed at 0 hai, what was expected since the plant was not infected yet. Moreover, the only gene coding polygalacturonase was not expressed during any time point.

4.3.3. CAZymes activity

The specific endoglucanase activity of *H. vastatrix* protein extract was significantly higher at 24 hai - around 200 U/g - when compared to 12 hai, which was almost undetectable (Figure 7a). At 96 hai, activity drops to less than half.

Conversely, xylanases (Figure 7b) have at 96 hai an activity value close to that of endoglucanases at 24 hai, which is significantly higher when compared to xylanases activity at 24 hai. At 12 hai these enzymes were not active.

Polygalacturonases showed a growing activity with time, reaching its maximum at 96 hai, which was smaller than the peak of the other enzymes.

4.4. Discussion

4.4.1. The genome of biotrophic fungus *H. vastatrix* encodes several CAZymes families potentially functional

Although the number of CAZyme coding genes predicted in the *H. vastatrix* genome (345) was more than the double found by Talhinhas *et al.* (2014) (148), its proportion in each CAZyme family remained similar. The higher number found may be explained by the different approaches used in each work. Talhinhas et al (2014) only used blastp comparison, while we used dbCAN2, which integrates 3 tools/databases for CAZyme annotation: DIAMOND, that generates fast blast hits in the CAZy database, HMMER to annotate CAZyme domain boundaries and Hotpep for short conserved motifs in the Peptide Pattern Recognition. This procedure increases the sensitivity of the searches, leading to a wider range of identified enzymes.

Among the 345 CAZymes found, nearly half belong to the GH group, an interesting finding, once this enzymes group hydrolyses glycosidic bonds between carbohydrates (CAZy, 2020). Biotrophic fungi do not tend to extensively hydrolyze the vegetal cell wall causing necrosis, as they establish a constant feeding relationship with the host's living cells instead of killing it (DEACON, 2020). On the other hand, CAZymes distribution and quantity found here were congruent with those of other biotrophic fungi observed in Zhao *et al.* (2013) work.

Experimental determination of a protein function usually depends on considerable time and financial resources (LEE; REDFERN; ORENGO, 2007). For this reason, computational approaches were used in this work in order to find some patterns and conserved regions in the sequences, in addition to other evidence to infer whether the predicted CAZymes are functional. The multiple sequence alignments performed allowed us to identify residues that may be crucial for the function of the protein and gave a good indication that the aligned enzymes could have a similar function, due to the conservation shown in the regions of the catalytic sites. Likewise, the structural comparison of the predicted CAZymes with well characterized ones indicated a similar fold of these proteins around the catalytic sites, suggesting that they have the same or a similar function.

According to Punta and Ofran (2008), the sequence and structural similarity can not be used alone to state the function of a protein, because it may have come from the homologous protein used in the modelling. Furthermore, it is known that two proteins could have the same structure

with the functional residues conserved and present different functions, as well as have the same function besides their totally different structures. Nevertheless, some information such as residue conformation and cavities can provide good clues about properties of the protein associated with the function. These approaches allow computational prediction of the function making them a good resource that precedes bench experiments, saving time and money in numerous cases.

4.4.2. CAZymes from *H. vastatrix* are expressed during coffee infection

The endoglucanase- and xylanase-coding genes presented different expression profiles during *H. vastatrix* infection. The only gene coding for a polygalacturonase was not detected in RNA-seq data.

Almost all of these genes were expressed mostly in the early hours of infection (12 and 24 hai), which is congruent with what was found by Talhinhas *et al.* (2014). In their work, it was observed a higher expression of genes that encode CAZymes during urediniospore germination and appressorium formation phases, that occur in the early hours of the infection. This trend was also verified for two resistance gene analogs (RGAs) during compatible and incompatible coffee and *H. vastatrix* interaction, when the major expression peak occurred 24 hai (ALMEIDA *et al.*, 2021). When compared to the later response (infected leaves after 21 days after inoculation), in which the haustorium is already formed, the quantity of genes encoding CAZymes goes from 90 to 57.

It is known that some fungi, during plant interaction, can express genes involved in the formation of infectious structures as well as synthesize enzymes responsible for the degradation of the host cell wall (KOLATTUKUDY *et al*., 1995; KAHMANN; BASSE, 2001; EMAMI; HACK, 2002; ZHAO *et al*., 2013; KUBICEK; STARR; GLASS, 2014;). These genes can be expressed at different stages of infection: during the initial contact or as a response to plant defense, for example. Combining this fact with what is known about the infectious process of *H. vastatrix*, in the compatible interaction, it can be inferred that around 12 h and 17 hai, when hyphae enter the plant through stomata (SILVA *et al.*, 2008), there is a stimulus for the production of some accessory enzymes, which will facilitate penetration (GIBSON *et al.*, 2011). At 24 hai, a time when it was previously found that about 54% of the appressorium is already

formed over the stomata of susceptible coffee leaves (SILVA *et al.*, 2002.), the production of other enzymes is still occurring and some peaks of expression are observed (Figure 6).

In addition, endoglucanases and xylanases have their expression peak at different times, 12 and 24 hai, respectively. This should be related only to the fact that the expression of these enzymes is a response to the type and concentration of available carbohydrates (KUMAR; NARAIAN, 2019).

4.4.3. The biotrophic fungus *H. vastatrix* produces active CAZymes in the post-haustorial phase during coffee infection

CAZyme activities were congruent with the gene expression data, since most endoglucanases started to be produced at 12 hai, the enzyme activity was higher at 24 hai, and decreased at 96 hai. Although xylanases were expressed at 12 hai, its peak was at 24 hai, which explains a high enzyme activity at 96 hai.

Polygalacturonase presents a lower activity than the other groups of enzymes and has its peak at 96 hai (Figure 7c) - although its expression has not been detected, probably due to the RNA-seq coverage value used.

It is worth noting that these enzymes were more active in the post-haustorial phase of the infection, which begins around 24 hai (RAMIRO *et al.*, 2009). This contrasts with the mechanism described by Bellincampi (2014), that shows that these CWDEs are present in the early stages of infection (the pre-haustorial phase) in order to soften and loosen the cell walls and facilitate biotrophic fungal penetration. Although gene expression started in the pre-haustorial phase, *H. vastatrix* enzymes were active when haustorium started to breach the host cell wall and kept their activity even after haustorium had invaded the host membranes and plant cell wall

4.5. Conclusion

Among the 345 CAZymes predicted in the genome of the biotrophic fungus *H. vastatrix* using dbCAN2, 162 were from the Glycoside Hydrolase family. Furthermore, 103 also showed BLASTp hits to sequences that belong to the previously assigned families, giving more reliability to the functional annotation. From these, 34 endoglucanases, xylanases and polygalacturonases from families GH 5, 9, 10, 12 and 28 were selected. These enzymes showed acceptable physico-chemical parameters and conserved structural properties compared to enzymes experimentally characterized. Furthermore, the multiple sequence alignments with well characterized proteins from the same families indicated a great conservation between sequences, especially in the regions of the catalytic sites. In addition, some of the predicted models and the templates were structurally overlapped, showing that the predicted proteins can fold in a similar way, and suggesting that they are functional.

The expression of the CAZyme-coding genes suggested that they can play different roles during *H. vastatrix* infection. During infection, *H. vastatrix* initially produced more cellulases (12 hai) and, later on, xylanases (24 hai). This coordinated pattern of expression of CAZymes can be inferred as a strategy to establish the infection efficiently, so that the plant tissue does not suffer necrosis and the plant's immune response is not so severe as to inhibit the growth of the fungus.

Figures

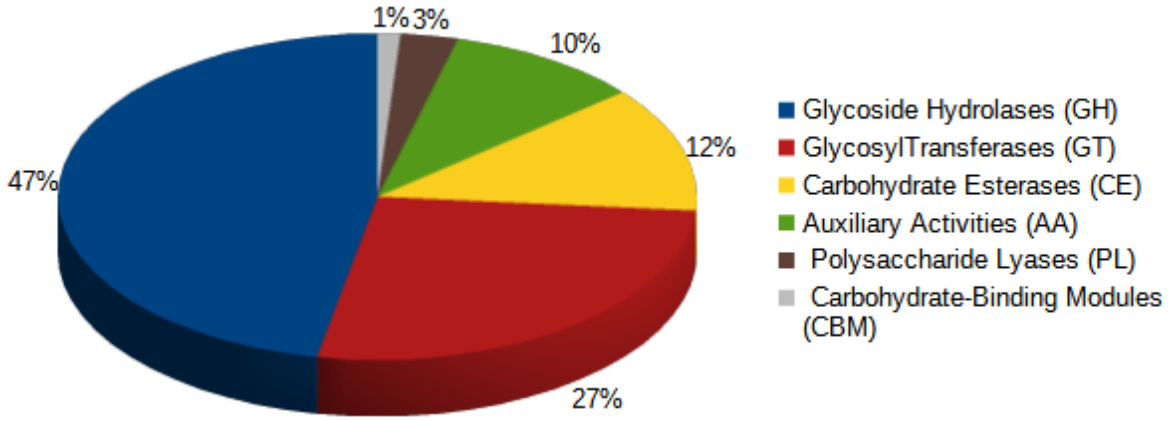


FIGURE 1: Distribution of CAZymes encoded in *H. vastatrix* genome.

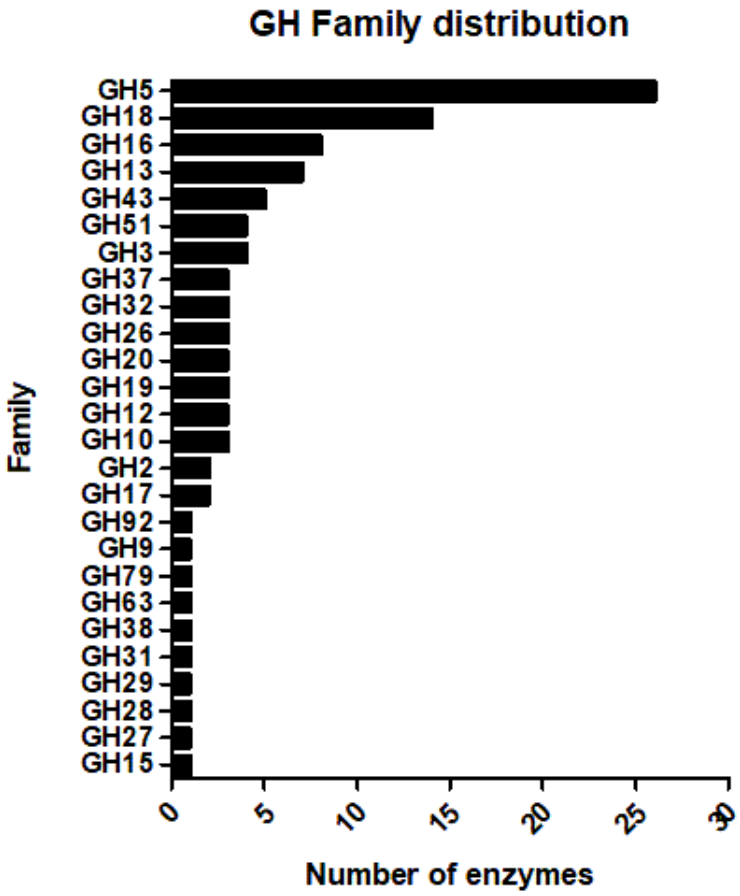
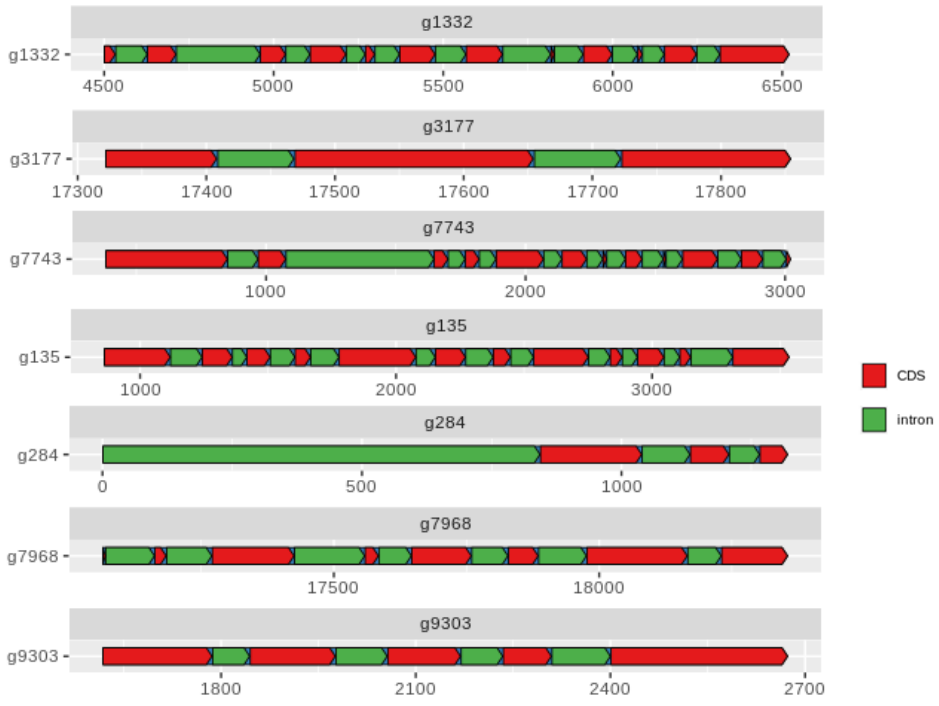
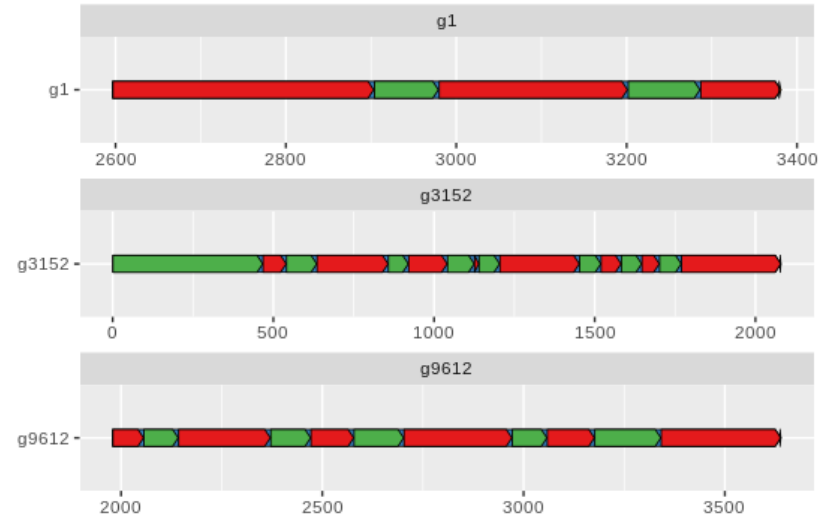


FIGURE 2: Distribution of Glycoside Hydrolases families coded in the genome of *H. vastatrix.*

3(a)



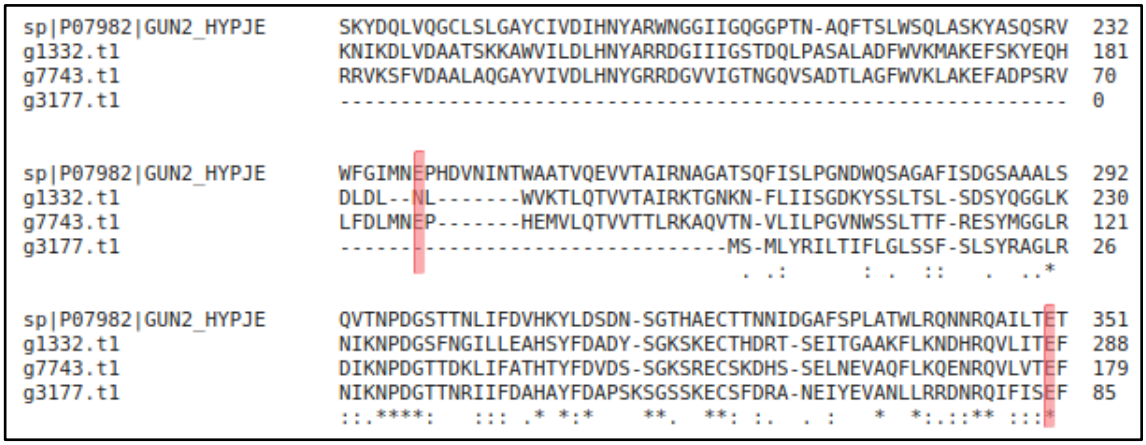
3(b)



3(c)

FIGURE 3: Gene structure and mapping of selected CAZymes. The scale shows the locus in the contigs analyzed. Regions in red represent the coding sequences (CDS) and sections in green the non coding ones (intron). Plots shows (a) some selected endoglucanases from GH 5 (g1332, g3177, g7743), GH 9 (g135) and GH 12 endoglucanases (g284, g7968, g9303), (b) GH 10 xylanases, (c) GH 28 polygalacturonase.

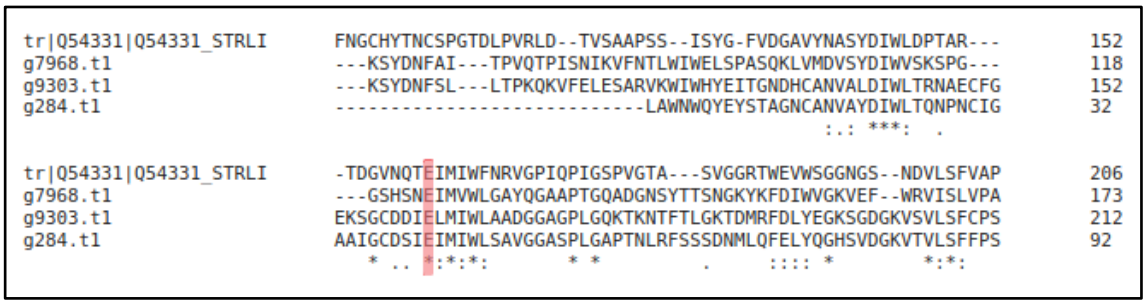
4(a)


sp|P07982|GUN2_HYPJE SKYDQLVQGCLSLGAYCIVDIHNYARWNGGIIGQGGPTN-AQFTSLWSQLASKYASQSRV 232
g1332.t1 KNIKDLVDAATSKKAWVILDLHNYARRDGIIIGSTDQLPASALADFWVKMAKEFSKYEQH 181
g7743.t1 RRVKSFVDAALAQGAYVIVDLHNYGRRDGVVIGTNGQVSADTLAGFWVKLAKEFADPSRV 70
g3177.t1 ---------- 0
sp|P07982|GUN2_HYPJE WFGIMNEPHDVNINTWAATVQEVVTAIRNAGATSQFISLPGNDWQSAGAFISDGSAAALS 292
g1332.t1 DLDL--NL-------WVKTLQTVVTAIRKTGNKN-FLIISGDKYSSLTSL-SDSYQGGLK 230
g7743.t1 LFDLMNEP-------HEMVLQTVVTTLRKAQVTN-VLILPGVNWSSLTTF-RESYMGGLR 121
g3177.t1 ----------MS-MLYRILTIFLGLSSF-SLSYRAGLR 26
sp|P07982|GUN2_HYPJE QVTNPDGSTTNLIFDVHKYLDSDN-SGTHAECTTNNIDGAFSPLATWLRQNNRQAILTET 351
g1332.t1 NIKNPDGSFNGILLEAHSYFDADY-SGKSKECTHDRT-SEITGAAKFLKNDHRQVLITEF 288
g7743.t1 DIKNPDGTTDKLIFATHTYFDVDS-SGKSRECSKDHS-SELNEVAQFLKQENRQVLVTEF 179
g3177.t1 NIKNPDGTTNRIIFDAHAYFDAPSKSGSSKECSFDRA-NEIYEVANLLRRDNRQIFISEF 85
sp|P37700|GUNG_RUMCH ----------FVVGYGVNPPQHPHHRTAHGS--WTDQMTSPTYHRHTIYGALVGGP 438
g135.t1 RPYNQNLQSNSRYHYCHKVHPNSPKNPHSASASGGTDINDIDDSPPDTAHVLYGAVVGGP 448
sp|P37700|GUNG_RUMCH DNADGYTDEINNYVNNEIACDYNAGFTGALAKMYKHSGGDPIPNFKAIEKITNDEVIIKA 498
g135.t1 NEQDEYFDQRSDWAQSEVALDYNAPFLTLVSYNLLNKAVDPP--Y---VKLMGPPVILKR 503

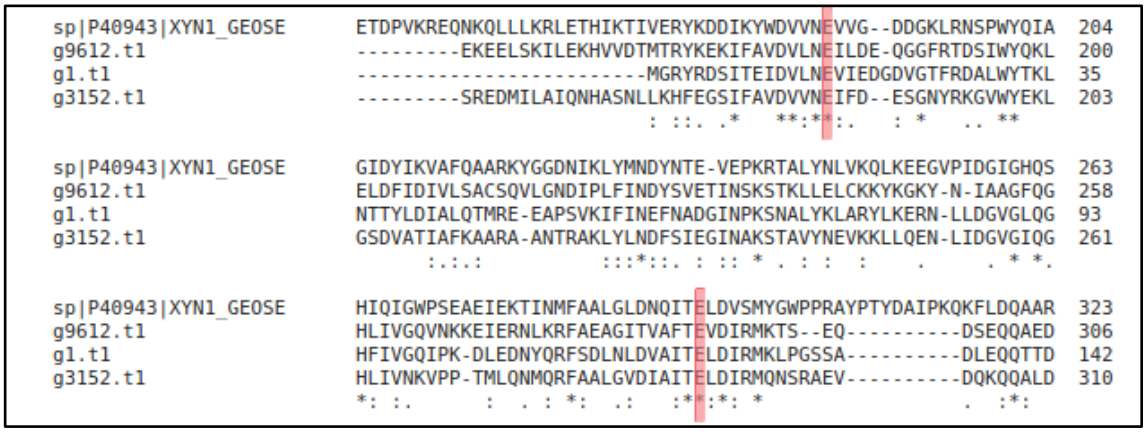
4(b)


sp|P37700|GUNG_RUMCH ----------FVVGYGVNPPQHPHHRTAHGS--WTDQMTSPTYHRHTIYGALVGGP 438
g135.t1 RPYNQNLQSNSRYHYCHKVHPNSPKNPHSASASGGTDINDIDDSPPDTAHVLYGAVVGGP 448
sp|P37700|GUNG_RUMCH DNADGYTDEINNYVNNEIACDYNAGFTGALAKMYKHSGGDPIPNFKAIEKITNDEVIIKA 498
g135.t1 NEQDEYFDQRSDWAQSEVALDYNAPFLTLVSYNLLNKAVDPP--Y---VKLMGPPVILKR 503

4(c)


tr|Q54331|Q54331_STRLI FNGCHYTNCSPGTDLPVRLD--TVSAAPSS--ISYG-FVDGAVYNASYDIWLDPTAR--- 152
g7968.t1 ---KSYDNFAI---TPVQTPISNIKVFNTLWIWELSPASQKLVMDVSYDIWVSKSPG--- 118
g9303.t1 ---KSYDNFSL---LTPKQKVFELESARVKWIWHYEITGNDHCANVALDIWLTRNAECFG 152
g284.t1 ----------LAWNWQYEYSTAGNCANVAYDIWLTQNPNCIG 32
tr|Q54331|Q54331_STRLI -TDGVNQTEIMIWFNRVGPIQPIGSPVGTA---SVGGRTWEVWSGGNGS--NDVLSFVAP 206
g7968.t1 ---GSHSNEIMVWLGAYQGAAPTGQADGNSYTTSNGKYKFDIWVGKVEF--WRVISLVPA 173
g9303.t1 EKSGCDDIELMIWLAADGGAGPLGQKTKNTFTLGKTDMRFDLYEGKSGDGKVSVLSFCPS 212
g284.t1 AAIGCDSIEIMIWLSAVGGASPLGAPTNLRFSSSDNMLQFELYQGHSVDGKVTVLSFFPS 92

4(d)

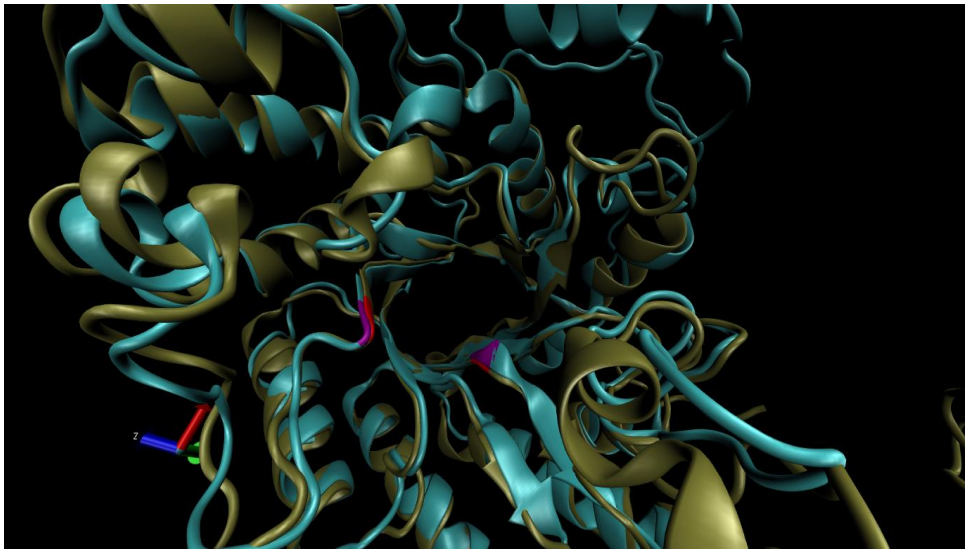

sp|P40943|XYN1_GEOSE ETDPVKREQNKQLLLKRLETHIKTIVERYKDDIKYWDVVNEVVG--DDGKLRNSPWYQIA 204
g9612.t1 ---------EKEELSKILEKHVVDTMTRYKEKIFAVDVLNEILDE-QGGFRTDSIWYQKL 200
g1.t1 ----------MGRYRDSITEIDVLNEVIEDGDVGTFRDALWYTKL 35
g3152.t1 ---------SREDMILAIQNHASNLLKHFEGSIFAVDVVNEIFD--ESGNYRKGVWYEKL 203
sp|P40943|XYN1_GEOSE GIDYIKVAFQAARKYGGDNIKLYMNDYNTE-VEPKRTALYNLVKQLKEEGVPIDGIGHQS 263
g9612.t1 ELDFIDIVLSACSQVLGNDIPLFINDYSVETINSKSTKLLELCKKYKGKY-N-IAAGFQG 258
g1.t1 NTTYLDIALQTMRE-EAPSVKIFINEFNADGINPKSNALYKLARYLKERN-LLDGVGLQG 93
g3152.t1 GSDVATIAFKAARA-ANTRAKLYLNDFSIEGINAKSTAVYNEVKKLLQEN-LIDGVGIQG 261
sp|P40943|XYN1_GEOSE HIQIGWPSEAEIEKTINMFAALGLDNQITELDVSMYGWPPRAYPTYDAIPKQKFLDQAAR 323
g9612.t1 HLIVGQVNKKEIERNLKRFAEAGITVAFTEVDIRMKTS--EQ----------DSEQQAED 306
g1.t1 HFIVGQIPK-DLEDNYQRFSDLNLDVAITELDIRMKLPGSSA----------DLEQQTTD 142
g3152.t1 HLIVNKVPP-TMLQNMQRFAALGVDIAITELDIRMQNSRAEV----------DQKQQALD 310

4(e)

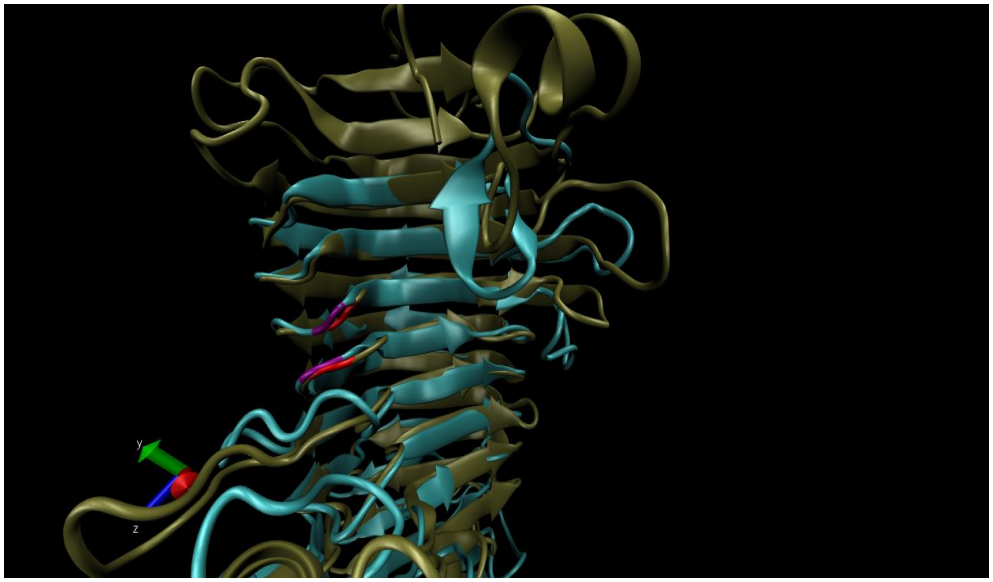
```
sp|Q9UUZ2|XGHA_ASPTU      PPNVFNSVKGGATNVVFSNLKMDANSKSDNPPKNTDGFDIGESTYVTITEVTVVNDDDCV  231
g12069.t1                 NSKLFHVNLFGCNNITFQHFTITAP----GDSPNTDGIHIGHSTGVVITDSNIGTGDDCI  196
                            ::*:    *..*:.*.::.: *     .   ****:.**.** *.**: .: ..***:

sp|Q9UUZ2|XGHA_ASPTU      AFKPSSNYVTVDTISCTGSHGISVGSLGKSSDD-SVKNIYVTGATMINSTKAAGIKTYPS  290
g12069.t1                 SIGDGAKQVNISKVTCGPGHGISVGSLGRYDNELPVEGIFVTDCTISGTLNGVRVKSWPA  256
                          ::  .:: *.:..::*  .********: .::  *:.*:**..*: .: :.. :*::*:
```

FIGURE 4: Alignment of the selected *H. vastatrix* CAZymes with other well characterized enzymes from the same family. (a) GH 5 endoglucanases (g1332, g3177, g7743) aligned with Endoglucanase EG-II from *Hypocrea jecorina* (PBD ID: 3QR3). (b) GH 9 (g135) endoglucanase aligned with Endoglucanase G from *Ruminiclostridium cellulolyticum* (PBD ID: 1GA2). (c) GH 12 endoglucanases (g284, g7968, g9303) aligned with Endoglucanase from *Streptomyces lividans* (PBD ID: 2NLR). (d) GH 10 xylanases (g1, g3152, g9612) aligned with Endo-1,4-beta-xylanase from *Geobacillus stearothermophilus* (PBD ID: 1R85). (e) GH 28 polygalacturonase (g12069) aligned with endo-xylogalacturonan hydrolase A from *Aspergillus tubingensis* (PBD ID: 4C2L).

5(a)



5(b)



5(c)

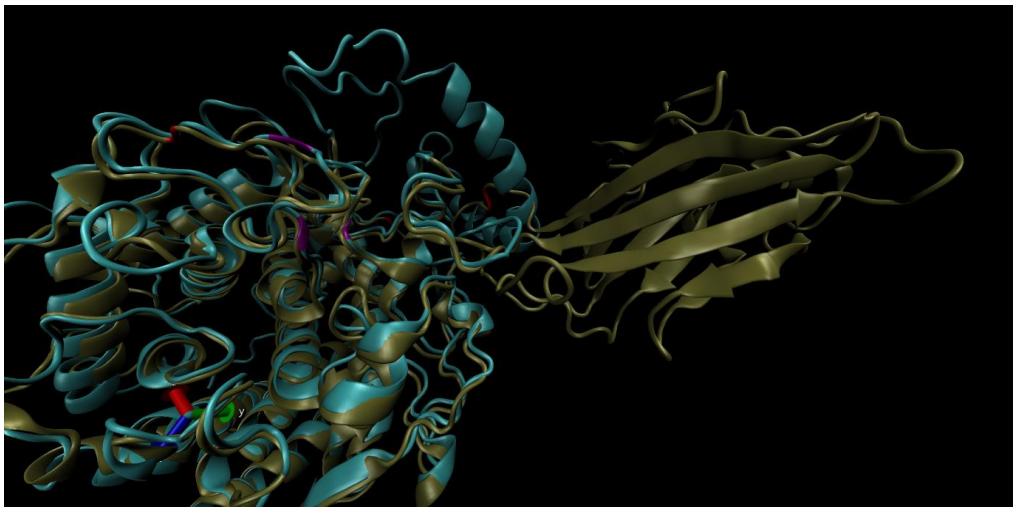


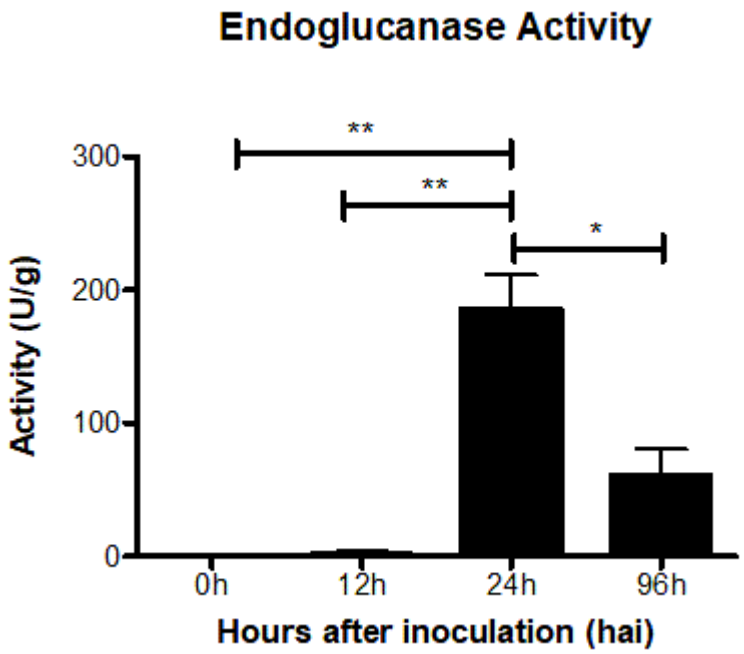
FIGURE 5: Overlapped structures of *H. vastatrix* CAZymes and previously characterized enzymes. (a) The blue structure is the predicted GH10 xylanase from *H. vastatrix* (g9612) and the bronze is the well characterized one GH 10 xylanase from *Cellvibrio japonicus* (PDB ID: 1US2). The catalytic sites, marked in red for the first one and in purple for the other one, show the superposition in this region. (b) Following the same color pattern as in a, GH 28 polygalacturonase (g12069) aligns with endo-xylogalacturonan hydrolase A from *Aspergillus tubingensis* (PBD ID: 4C2L). (c) Endoglucanase GH 9 (g135) aligned with Endoglucanase G from *Ruminiclostridium cellulolyticum* (PBD ID: 1GA2), showing some overlapping of the structures but not in the catalytic residues.

Enzyme	Family	ID	Hours after inoculation 0 hai	12 hai	24 hai	96 hai
Endoglucanase		g1332	0,0	0,0	0,0	0,0
		g1441	0,0	173,1	116,3	62,1
		g3177	0,0	0,0	54,6	131,0
		g3389	0,0	0,0	0,0	0,0
		g3729	0,0	0,0	11,7	0,0
		g4364	0,0	271,1	320,1	92,3
		g4848	0,0	157,8	17,2	10,9
		g5621	0,0	2,6	0,0	0,0
		g5660	0,0	25,9	0,0	0,0
		g6021	0,0	0,0	0,0	0,0
		g6446	0,0	0,0	0,0	0,0
		g7743	0,0	0,0	0,0	0,0
	GH5	g7803	0,0	0,0	0,0	0,0
		g7804	0,0	1,8	0,0	8,8
		g8435	0,0	0,0	0,0	0,0
		g8603	0,0	67,2	11,9	0,0
		g8843	0,0	0,0	0,0	0,0
		g8872	0,0	66,3	48,5	17,9
		g8910	0,0	0,0	0,0	0,0
		g9077	0,0	58,0	33,8	25,8
		g9307	0,0	31,8	6,4	4,5
		g9743	0,0	0,0	0,0	0,0
		g10543	0,0	2,7	0,0	0,0
		g10617	0,0	2,3	0,0	0,0
		g11342	0,0	7,5	0,0	0,0
		g11585	0,0	0,0	6,2	0,0
	GH9	g135	0,0	0,0	13,7	4,5
	GH12	g284	0,0	0,0	548,4	57,7
		g7968	0,0	0,0	0,0	0,0
		g9303	0,0	11,3	0,0	0,0
Xylanase	GH10	g1	0,0	57,5	93,3	0,0
		g3152	0,0	41,7	0,0	7,1
		g9612	0,0	18,5	230,1	12,5

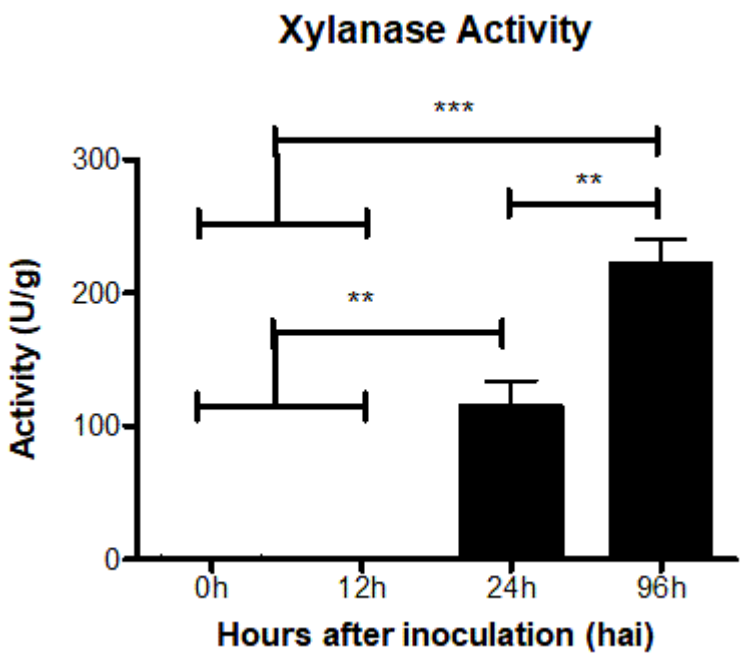
Expression (FPKM)
0,0 548,4

FIGURE 6: Analysis of expression of the enzymes during coffee leaf infection by *H. vastatrix*. 0h represents the non infected plant, and for this reason no expression values were observed for this time. The bigger expression value was observed in the GH12 endoglucanase g284 at 24 hours after inoculation.

7(a)



7(b)



7(c)

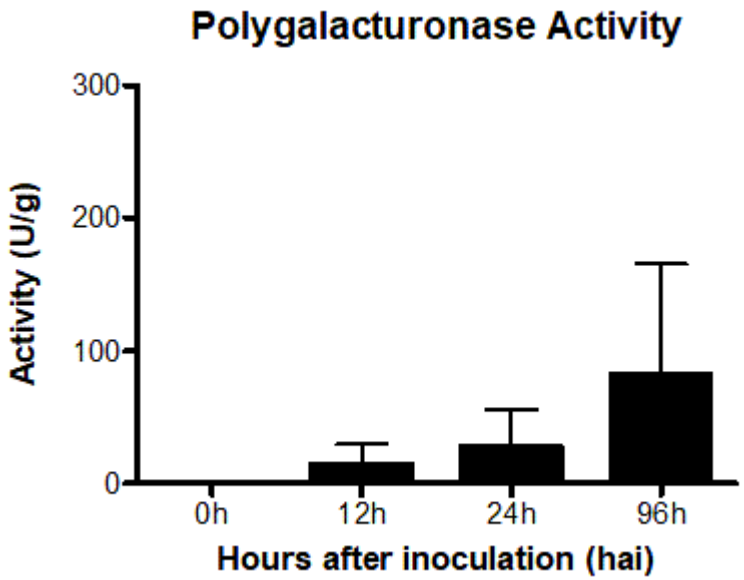


FIGURE 7: Analysis of activity of (a) Endoglucanases, (b) Xylanases, (c) Polygalacturonases in coffee plant protein extract during infection with *H. vastatrix*. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05 in ANOVA test with Bonferroni correction for multiple hypothesis test.

Supplementary data

Table 1 - Physical and chemical parameters of the CAZymes identified in the *Hemileia vastatrix* genome.

Gene ID	Family	Number of amino acids	Molecular weight	Thoretical pI	Signal Peptide	Number of N-Glycosylation Sites	Number of O-Glycosylation Sites
g3389	GH5	621	70533.64	8.44	ND	5	ND
g9743		621	70248.70	6.51	ND	4	3
g7803		626	71696.54	6.55	ND	6	ND
g7804		561	63123.80	6.03	ND	6	2
g4364		541	60546.62	6.68	ND	8	ND
g9307		562	63195.88	6.85	ND	8	1
g8910		438	50794.06	9.59	1-25	3	ND
g5660		562	63376.87	8.90	ND	5	2
g8843		491	55232.64	6.28	ND	3	2
g11342		222	25732.84	8.83	ND	3	1
g8435		330	38293.14	8.62	ND	3	ND
g1332		315	34847.52	9.12	ND	ND	ND
g3177		134	14925.01	9.64	1-21	ND	ND
g7743		427	47292.38	9.42	ND	1	1
g8603		326	36851.29	8.97	ND	1	1
g10543		427	47585.75	8.79	1-19	2	0
g3729		587	65416.38	4.87	ND	3	ND
g6446		277	31127.10	5.13	ND	ND	ND
g4848		273	30274.14	9.08	ND	1	ND
g10617		456	51558.95	6.83	1-20	1	ND
g1441		715	77093.15	5.40	ND	3	6
g6021		772	85679.99	7.55	ND	2	6
g8872		850	92996.16	5.94	ND	6	1
g9077		967	108969.06	5.78	ND	5	ND
g5621		716	82633.06	6.85	ND	5	ND
g11585		308	34812.25	5.25	ND	ND	ND
g135	GH9	546	60783.31	7.16	ND	6	ND
g7968	GH12	230	25145.53	6.28	ND	ND	ND
g9303		253	29167.89	8.07	ND	1	ND
g284		108	11889.22	4.16	ND	ND	1
g1	GH10	207	23390.38	4.60	ND	1	ND
g3152		384	43497.59	9.02	1-28	ND	ND
g9612		365	41773.55	8.93	1-21	2	ND
g12069	GH28	296	31747.98	5.22	ND	5	1

Conflicts of interest

The authors reports no conflicts of interest

Acknowledgments

FAPEMIG, Capes, CNPq and Consórcio Pesquisa Café

References

3R LAB. Ferrugem do cafeeiro (hemileia vastatrix) compreenda essa doença e saiba como combatê-la. Disponível em: https://www.3rlab.com.br/2020/07/20/ferrugem-do-cafeeiro-hemileia-vastatrix-compreenda-essa-doenca-e-saiba-como-combate-la/. Acesso em: 21 abr. 2021.

ABIC - Associação Brasileira da Indústria do Café -. **História do café**. Disponível em: https://www.abic.com.br/o-cafe/historia/. Acesso em: 04 jan. 2021.

AGROLINK. **Bicho mineiro (Leucoptera coffeella)**. Disponível em: https://www.agrolink.com.br/problemas/bicho-mineiro_471.html. Acesso em: 08 jan. 2021.

ALFONSI, Waldenilza Monteiro Vital *et al*. Período de incubação da ferrugem do cafeeiro. **Summa Phytopathologica**, v. 45, n. 2, p. 134-140, 10 jul. 2019. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-54052019000200134&tlng=pt. Acesso em: 14 dez. 2020.

ALFONSI, Waldenilza Monteiro Vital *et al*. Geographical distribution of the incubation period of coffee leaf rust in climate change scenarios. **Pesquisa Agropecuária Brasileira**, v. 54, p. 1-10, 2019b. DOI: 10.1590/s1678-3921.pab2019.v54.00273.

ALMEIDA, Dênia Pires de *et al.* Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. **Scientia Agricola**, v. 78, n. 6, 2021.

ARMENTEROS, José Juan Almagro *et al*. SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature Biotechnology**, [S.L.], v. 37, n. 4, p. 420-423, 18 fev. 2019. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/s41587-019-0036-z.

BELLINCAMPI, Daniela; CERVONE, Felice; LIONETTI, Vincenzo. Plant cell wall dynamics and wall-related susceptibility in plant-pathogen interactions. **Frontiers In Plant Science**, [S.L.], v. 5, p. 1-8, 28 maio 2014. Frontiers Media SA. http://dx.doi.org/10.3389/fpls.2014.00228. Disponível em: https://www.frontiersin.org/articles/10.3389/fpls.2014.00228/ful:l. Acesso em: 12 dez. 2020.

BIELY, P. *et al*. Endo-β-1,4-xylanase families: differences in catalytic properties. **Journal Of Biotechnology**, [S.L.], v. 57, n. 1-3, p. 151-166, set. 1997. Elsevier BV. http://dx.doi.org/10.1016/s0168-1656(97)00096-5.

BORUAH, Paranjoli *et al*. Xylanase from Penicillium meleagrinum var. viridiflavum – a potential source for bamboo pulp bleaching. **Journal Of Cleaner Production**, [S.L.], v. 116, p. 259-267, mar. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.jclepro.2015.12.024.

BOWIE, J.; LUTHY, R; EISENBERG, D. A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, [S.L.], v. 253, n. 5016, p. 164-170, 12 jul. 1991. American Association for the Advancement of Science (AAAS). http://dx.doi.org/10.1126/science.1853201.

CABRAL, P. G. C. *et al*. Genetic diversity and structure of *Hemileia vastatrix* populations onCoffeaspp. **Plant Pathology**, [S.L.], v. 65, n. 2, p. 196-204, 2016. Wiley. http://dx.doi.org/10.1111/ppa.12411

CAPUCHO, A. S. *et al*. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. **Australasian Plant Disease Notes**, [S.L.], v. 7, n. 1, p. 189-191, 27 out. 2012. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s13314-012-0081-7.

CAPUCHO, A. S. *et al*. Herança da resistência do Híbrido de Timor UFV 443-03 à ferrugem-do-cafeeiro. **Pesquisa Agropecuária Brasileira**, [S.L.], v. 44, n. 3, p. 276-282, mar. 2009. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s0100-204x2009000300009.

CAZy: CARBOHYDRATE-ACTIVE enZYmes Database. Disponível em: http://www.cazy.org/. Acesso em: 12 dez. 2020.

CECAFE - Conselho dos Exportadores de Café do Brasil -. **História do café**. Disponível em: https://www.cecafe.com.br/sobre-o-cafe/historia-do-cafe/. Acesso em: 04 jan. 2021.

CHEMURA, Abel; MUTANGA, Onisimo; DUBE, Timothy. Separability of coffee leaf rust infection levels with machine learning methods at Sentinel-2 MSI spectral resolutions. **Precision Agriculture**, [S.L.], v. 18, n. 5, p. 859-881, 27 dez. 2016. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s11119-016-9495-0.

COLOVOS, C.; YEATES T.O.. Verification of protein structures: patterns of nonbonded atomic interactions. **Protein Sci**. 1993 Sep;2(9):1511-9. doi: 10.1002/pro.5560020916. PMID: 8401235; PMCID: PMC2142462.

CONAB - Companhia Nacional de Abastecimento -. **Boletim da Safra Brasileira de Café - Dezembro 2020**. Disponível em: https://www.conab.gov.br/info-agro/safras/cafe. Acesso em: 04 jan. 2021.

CRESSEY, Daniel. Coffee rust regains foothold. **Nature**, X, v. 493, p. 587-587, 30 jan. 2013. Disponível em: https://www.nature.com/news/coffee-rust-regains-foothold-1.12320. Acesso em: 06 jan. 2021.

CRISTANCHO, Marco 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**, [S.L.], v. 5, n. , p. 1-11, 31 out. 2014. Frontiers Media SA. http://dx.doi.org/10.3389/fpls.2014.00594.

DEACON, Jim. **The Microbial World: Biotrophic plant pathogens**. Disponível em: http://archive.bio.ed.ac.uk/jdeacon/microbes/biotroph.htm. Acesso em: 07 dez. 2020.

DESCROIX, F; SNOECK, J. Environmental factors suitable for coffee cultivation. In: WINTGENS, J. N.. **Coffee: growing, processing, sustainable production**: a guidebook for growers, processors, traders and researchers. Weinheim: Wiley-Vch, 2009. p. 168-181.

DODD, Dylan.; CANN, Isaac. K. O. Enzymatic deconstruction of xylan for biofuel production. **GCB Bioenergy**, v. 1, p. 2–17, 2009.

EMAMI, Kaveh; HACK, Ethan. Conservation of XYN11A and XYN11B xylanase genes in *Bipolaris sorghicola*, *Cochliobolus sativus*, *Cochliobolus heterostrophus*, and *Cochliobolus spicifer*. **Current Microbiology**, v. 45, p. 303–306, 2002. DOI: 10.1007/s00284-002-3618-8.

FARINAS, Cristiane Sanchez. A parede celular vegetal e as enzimas envolvidas na sua degradação. **Embrapa Instrumentação**, São Carlos, SP, 2011. Disponível em: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/919344/a-parede-celular-vegetal-e-as-enzimas-envolvidas-na-sua-degradacao. Acesso em: 27 fev. 2019.

FERNANDES, Ronaldo de Castro. **REAVALIAÇÃO DO CICLO DE VIDA DE Hemileia vastatrix NO BRASIL**. 2007. 64 f. Tese (Doutorado) - Curso de Fitopatologia, Universidade Federal de Viçosa, Viçosa, 2007. Disponível em: https://www.locus.ufv.br/handle/123456789/4420. Acesso em: 07 out. 2020.

FERNANDEZ, Diana *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**, [S.L.], v. 13, n. 1, p. 17-37, 1 jun. 2011. Wiley. http://dx.doi.org/10.1111/j.1364-3703.2011.00723.x.

FLOREZ, Juan Carlos *et al*. High throughput transcriptome analysis of coffee reveals prehaustorial resistance in response to Hemileia vastatrix infection. **Plant Molecular Biology**, [S.L.], v. 95, n. 6, p. 607-623, 1 nov. 2017. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s11103-017-0676-7. Disponível em: https://pubmed.ncbi.nlm.nih.gov/29094279/. Acesso em: 21 out. 2020.

FREITAS, R. L. *et al*. Cytological Evaluation of the Infection Process of Hemileia vastatrix (race XXXIII) in Resistant and Susceptible Coffee. In: 25TH INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 25., 2014, Armenia (Colombia). **25th International Conference on Coffee Science, ASIC 2014, Armenia, Colombia, September 2014.** Paris (France): Association Scientifique International Du Café, 2014. p. 8-13.

GARCÍA-GUZMÁN, Graciela; HEIL, Martin. Life histories of hosts and pathogens predict patterns in tropical fungal plant diseases. **New Phytologist**, [S.L.], v. 201, n. 4, p. 1106-1120, 31 out. 2013. Wiley. http://dx.doi.org/10.1111/nph.12562.

GASTEIGER, Elisabeth *et al*. Protein Identification and Analysis Toolson the ExPASy Server. In: WALKER, John M.. **The Proteomics Protocols Handbook**. Humana Press, 2005. Cap. 52. p. 571-607.

GIBSON, Donna M *et al*. Plant pathogens as a source of diverse enzymes for lignocellulose digestion. **Current Opinion In Microbiology**, [S.L.], v. 14, n. 3, p. 264-270, jun. 2011. Elsevier BV. http://dx.doi.org/10.1016/j.mib.2011.04.002.

GOLD, Randall E.; MENDGEN, Kurt. Rust basidiospore germlings and disease initiation. Em: COLE, Garry T.; HOCH, Harvey C. **The Fungal Spore and Disease Initiation in Plants and Animals**. Nova York: 1991. DOI: 10.1007/978-1-4899-2635-7.

GUERRA-GUIMARÃES, L, et al. (2009). Chitinases of *Coffea arabica* genotypes resistant to orange rust *Hemileia vastatrix*. ***Biol. Plant***. 53, 702–706. doi: 10.1007/s10535-009-0126-8.

GUERRA-GUIMARÃES, L, et al. (2014). Effect of greenhouse conditions on the leaf apoplastic proteome of *Coffea arabica* plants. ***J. Proteomics*** 104, 128–139. doi: 10.1016/j.jprot.2014.03.024.

GUPTA, R. *et al*. Scanning the available Dictyostelium discoideum proteome for O-linked GlcNAc glycosylation sites using neural networks. **Glycobiology**, [S.L.], v. 9, n. 10, p. 1009-1022, 1 out. 1999. Oxford University Press (OUP). http://dx.doi.org/10.1093/glycob/9.10.1009.

GUPTA, R.; JUNG, E.; BRUNAK, S.. Prediction of N-glycosylation sites in human proteins. 2004.

HERBERT, Corentin *et al*. Production of a cell wall-associated endopolygalacturonase by Colletotrichum lindemuthianum and pectin degradation during bean infection. **Fungal Genetics And Biology**, [S.L.], v. 41, n. 2, p. 140-147, fev. 2004. Elsevier BV. http://dx.doi.org/10.1016/j.fgb.2003.09.008.

ICO - INTERNATIONAL COFFEE ORGANIZATION. **World coffee consumption**. Disponível em: http://www.ico.org/prices/new-consumption-table.pdf. Acesso em: 15 dez. 2020.

ICO a - International Coffee Organization -. **Coffee year production by country**. Disponível em: http://www.ico.org/prices/po-production.pdf. Acesso em: 04 jan. 2021.

ICO b - International Coffee Organization. **A história do café**. Disponível em: http://www.ico.org/pt/coffee_storyp.asp. Acesso em: 04 jan. 2021.

KAHMANN, Regine; BASSE, Christoph. Fungal gene expression during pathogenesis-related development and host plant colonization. Current Opinion in Microbiology, v. 4, p. 374-80, 2001. DOI: 10.1016/s1369-5274(00)00220-4. PMID: 11495797.

KELLEY, Lawrence A. *et al*. The Phyre2 web portal for protein modeling, prediction and analysis. **Nature Protocols**, [S.L.], v. 10, n. 6, p. 845-858, 7 maio 2015. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nprot.2015.053.

KOLATTUKUDY, Pappachan E.; ROGERS, Linda M.; LI, Daoxin; HWANG, Cheng-Shine; FLAISHMAN, Moshe A. Surface signaling in pathogenesis. Proceedings of the National Academy of Sciences of the United States of America, v. 92, p. 4080-4087, 1995. DOI:10.1073/pnas.92.10.4080.

KRIEGER E. *et al.* Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. **Proteins**. 2009;77 Suppl 9(Suppl 9):114-22. doi: 10.1002/prot.22570. PMID: 19768677; PMCID: PMC2922016.

KUBICEK, Christian P.; STARR, Trevor L.; GLASS, N. Louise. Plant Cell Wall–Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. **Annual Review Of**

Phytopathology, [S.L.], v. 52, n. 1, p. 427-451, 4 ago. 2014. Annual Reviews. http://dx.doi.org/10.1146/annurev-phyto-102313-045831.

KUMAR, Arvind; NARAIAN, Ram. Differential Expression of the Microbial β-1,4-Xylanase, and β-1,4-Endoglucanase Genes. **New And Future Developments In Microbial Biotechnology And Bioengineering: Microbial Genes Biochemistry And Applications**, [S.L.], p. 95-111, 2019. Elsevier. http://dx.doi.org/10.1016/b978-0-444-63503-7.00006-1. Disponível em: https://www.scopus.com/results/citedbyresults.uri?sort=plf-f&cite=2-s2.0-85082267189&src=s&imp=t&sid=05ca2c14ccff06c05e8f375c5af0ef56&sot=cite&sdt=a&sl=0&origin=inward&editSaveSearch=&txGid=c5e3d5eb9694abb1f0cb0942474a5be3. Acesso em: 12 dez. 2020.

LASKOWSKI R A, MACARTHUR M W, THORNTON J M. Procheck: validation of protein structure coordinates, in ***International Tables of Crystallography***, *Volume F. Crystallography of Biological Macromolecules*, eds. Rossmann M G & Arnold E, Dordrecht, Kluwer Academic Publishers, The Netherlands, 2001, pp. 722-725.

LEE, David; REDFERN, Oliver; ORENGO, Christine. Predicting protein function from sequence and structure. **Nature Reviews Molecular Cell Biology**, [S.L.], v. 8, n. 12, p. 995-1005, dez. 2007. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nrm2281.

LEE, Toni M. *et al*. A structural study of Hypocrea jecorina Cel5A. **Protein Science**, [S.L.], v. 20, n. 11, p. 1935-1940, 27 set. 2011. Wiley. http://dx.doi.org/10.1002/pro.730.

LOMBARD, Vincent *et al*. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, [S.L.], v. 42, n. 1, p. 490-495, 21 nov. 2013. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gkt1178. Disponível em: https://academic.oup.com/nar/article/42/D1/D490/1057423. Acesso em: 05 jan. 2021.

LÜTHY, Roland; BOWIE, James U.; EISENBERG, David. Assessment of protein models with three-dimensional profiles. **Nature**, [S.L.], v. 356, n. 6364, p. 83-85, mar. 1992. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/356083a0.

MADEIRA JUNIOR, Jose Valdo *et al*. Agro-Industrial Residues and Microbial Enzymes: An Overview on the Eco-Friendly Bioconversion into High Value-Added Products. In: BRAHMACHARI, Goutam *et al*. **Biotechnology of Microbial Enzyme**. Academic Press, 2017. Cap. 18. p. 475-500. Disponível em: https://www.elsevier.com/books/biotechnology-of-microbial-enzymes/brahmachari/978-0-12-803725-6. Acesso em: 07 dez. 2020.

MCCOOK, Stuart. Global rust belt: Hemileia vastatrix and the ecological integration of world coffee production since 1850. Journal of Global History, v. 1, p. 177–195, 2006. DOI: 10.1017/S174002280600012X.

MAGRACH, Ainhoa; GHAZOUL, Jaboury. Climate and Pest-Driven Geographic Shifts in Global Coffee Production: implications for forest cover, biodiversity and carbon storage. **Plos One**, [S.L.], v. 10, n. 7, p. 1-15, 15 jul. 2015. Public Library of Science (PLoS). http://dx.doi.org/10.1371/journal.pone.0133071.

MANDELMAN, David *et al*. X-Ray Crystal Structure of the Multidomain Endoglucanase Cel9G from Clostridium cellulolyticum Complexed with Natural and Synthetic Cello-Oligosaccharides. **Journal Of Bacteriology**, [S.L.], v. 185, n. 14, p. 4127-4135, 15 jul. 2003. American Society for Microbiology. http://dx.doi.org/10.1128/jb.185.14.4127-4135.2003.

MCCOOK, Stuart; VANDERMEER, John. The Big Rust and the Red Queen: long-term perspectives on coffee rust research. **Phytopathology®**, [S.L.], v. 105, n. 9, p. 1164-1173, set. 2015. Scientific Societies. http://dx.doi.org/10.1094/phyto-04-15-0085-rvw.

MOHNEN, Debra. Pectin structure and biosynthesis. **Current Opinion In Plant Biology**, [S.L.], v. 11, n. 3, p. 266-277, jun. 2008. Elsevier BV. http://dx.doi.org/10.1016/j.pbi.2008.03.006. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S1369526608000630?casa_token=gS_2V9VDd0oAAAAA:OWgrKaPsUfC-Gte-uHWWwmI98K4TPSZYfqRAJ_2yJ8fXlSIU7xQSR2inFf42TxqRlTAo2skWvIiZ. Acesso em: 14 dez. 2020.

NAIK, Banavath Jayanna *et al*. Responses to Biotic and Abiotic Stresses and Transgenic Approaches in the Coffee Plant. **Journal Of The Korean Society Of International Agricultue**, [S.L.], v. 31, n. 4, p. 359-377, 31 dez. 2019. Korean Society of International Agriculture. http://dx.doi.org/10.12719/ksia.2019.31.4.359.

NATIONAL COFFEE ASSOCIATION EUA. **Coffee Around the World**. Disponível em: https://www.ncausa.org/about-coffee/coffee-around-the-world. Acesso em: 15 dez. 2020.

National High Magnetic Field Laboratory of Florida State University. **Plant Cell Wall**. Disponível em: https://micro.magnet.fsu.edu/cells/plants/cellwall.html. Acesso em: 08 jan. 2021.

NAVYA, P. N.; PUSHPA, S. Murthy. Production, statistical optimization and application of endoglucanase from Rhizopus stolonifer utilizing coffee husk. **Bioprocess And Biosystems Engineering**, [S.L.], v. 36, n. 8, p. 1115-1123, 7 dez. 2012. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s00449-012-0865-3. Disponível em: https://www.researchgate.net/publication/233886486_Production_statistical_optimization_and_application_of_endoglucanase_from_Rhizopus_stolonifer_utilizing_coffee_husk. Acesso em: 25 jul. 2019.

O'CONNELL, R.J.; PANSTRUGA, R. Tête-à-tête inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. **New Phytologist, v.** 171, p. 699–718, 2006. doi : 10.1111/j.1469-8137.2006.01829.x

O'NEILL, Malcolm; ALBERSHEIM, Peter; DARVILL, Alan. The pectic polysaccharides of primary cell walls. In: DEY, P. M. Methods in Plant Biochemistry. 2. Ed. London: Academic Press; 1990. p. 415-441.

PANSTRUGA, Ralph; DODDS, Peter N. Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens. **Science**. 2009 May v. 324 (5928), p. 748–750. doi:10.1126/science.1171652.

PARTELLI, Fabio Luiz *et al*. Spatial distribution of the root system of Conilon and Arabica coffee plants. **Pesquisa Agropecuária Brasileira**, [S.L.], v. 55, p. 1-9, 29 jun. 2020. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1678-3921.pab2020.v55.01333. Disponível em: https://www.scielo.br/scielo.php?pid=S0100-204X2020000102601&script=sci_arttext. Acesso em: 05 jan. 2021.

PFENNING, Ludwig Heinrich; MARTINS, Melissa Faria. Espécies de *Fusarium* associadas ao cafeeiro na região sul de Minas Gerais. Em: **SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL,** 01., 2000, Poços de Caldas, 2000. p. 283-286.

PORTO, Brenda Neves *et al*. Genome sequencing and transcript analysis of Hemileia vastatrix reveal expression dynamics of candidate effectors dependent on host compatibility. **Plos One**, p. 1-23, 18 abr. 2019. Disponível em: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0215598. Acesso em: 06 jan. 2021.

PUNTA, Marco; OFRAN, Yanay. The Rough Guide to In Silico Function Prediction, or How To Use Sequence and Structure Information To Predict Protein Function. **Plos Computational Biology**, [S.L.], v. 4, n. 10, p. 1-7, 31 out. 2008. Public Library of Science (PLoS). http://dx.doi.org/10.1371/journal.pcbi.1000160.

RAMIREZ, Gustavo Adolfo Marin. **IDENTIFICAÇÃO DE GENES DE Hemileia vastatrix CANDIDATOS A EFETORES CAPAZES DE ATIVAR OU SUPRIMIR RESPOSTAS DE DEFESA NO CAFEEIRO**. 2015. 66 f. Tese (Doutorado) - Curso de Fitopatologia, Universidade Federal de Viçosa, Viçosa, 2015. Disponível em: https://www.locus.ufv.br/handle/123456789/6476. Acesso em: 02 out. 2020.

RAMIRO, D. A. *et al*. Biphasic haustorial differentiation of coffee rust (Hemileia vastatrixrace II) associated with defence responses in resistant and susceptible coffee cultivars. **Plant Pathology**, [S.L.], v. 58, n. 5, p. 944-955, out. 2009. Wiley. http://dx.doi.org/10.1111/j.1365-3059.2009.02122.x. Disponível em: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/880368/biphasic-haustorial-differentiation-of-coffee-rust-hemileia-vastatrix-race-ii-associated-with-defense-responses-in-resistant-and-susceptible-coffee-cultivars. Acesso em: 07 out. 2020.

RESEARCH AND MARKETS. **Global Coffee Market - By Product: Global Industry Perspective, Comprehensive Analysis and Forecast 2020 - 2026**. Disponível em: https://www.researchandmarkets.com/reports/5143799/global-coffee-market-by-product-global?utm_source=BW&utm_medium=PressRelease&utm_code=qb5cxz&utm_campaign=1448119+-+Global+Coffee+Market+(2020+to+2026)+-+Industry+Perspective%2c+Comprehensive+Analysis+and+Forecast&utm_exec=jamu273prd. Acesso em: 15 dez. 2020.

RIBEIRO, Bruno Batista *et al*. AVALIAÇÃO QUÍMICA E SENSORIAL DE BLENDS DE Coffea canephora Pierre E Coffea arabica L. **Coffee Science**, Lavras, v. 9, n. 2, p. 178-186, 12 jun. 2014. Disponível em: http://www.sbicafe.ufv.br/bitstream/handle/123456789/8027/Coffee%20Science_v9_n2_p178-186_2014.pdf?sequence=1&isAllowed=y. Acesso em: 05 jan. 2021.

ROZEBOOM, Henriëtte J. *et al*. Crystal structure of endo-xylogalacturonan hydrolase fromAspergillus tubingensis. **Febs Journal**, [S.L.], v. 280, n. 23, p. 6061-6069, 11 out. 2013. Wiley. http://dx.doi.org/10.1111/febs.12524.

SAHA, Badal C. Hemicellulose bioconversion. **Journal of Industrial Microbiology and Biotechnology**, v. 30, p. 279–291, 2003. DOI 10.1007/s10295-003-0049-x.

SANTIAGO, Emerson. **Rota das Especiarias**. Disponível em: https://www.infoescola.com/historia/rota-das-especiarias/. Acesso em: 04 jan. 2021.

SILVA, Diogo N. *et al*. Phylogenetic analysis of Hemileia vastatrix and related taxa using a genome-scale approach. In: INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 24., 2012, San Jose. **ASIC 2012.** X: Al- 01267895, 2020. p. 1-5.

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). **The University Of Chicago Press Journal.** na, p. 79-91. jan. 1999. Disponível em: https://www.journals.uchicago.edu/doi/abs/10.1086/314113. Acesso em: 09 out. 2020.

SILVA, M.C. *et al*. Involvement of peroxidases in the coffee resistance to orange rust (Hemileia vastatrix). **Physiological And Molecular Plant Pathology**, [S.L.], v. 72, n. 1-3, p. 29-38, jan. 2008. Elsevier BV. http://dx.doi.org/10.1016/j.pmpp.2008.04.004.

SILVA, M.C.; NICOLE, M.; GUERRA-GUIMARÃES, L.; RODRIGUES Jr., C.J. 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. doi: 10.1006/pmpp.2002.0389.

SILVA, Maria do Céu *et al*. Coffee resistance to the main diseases: leaf rust and coffee berry disease. **Brazilian Journal Of Plant Physiology**, [S.L.], v. 18, n. 1, p. 119-147, mar. 2006. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1677-04202006000100010.

STANKE, M.; MORGENSTERN, B.. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. **Nucleic Acids Research**, [S.L.], v. 33, n. , p. 465-467, 1 jul. 2005. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gki458.

SULZENBACHER, Gerlind *et al*. The Crystal Structure of a 2-Fluorocellotriosyl Complex of theStreptomyce lividans Endoglucanase CelB2 at 1.2 Å Resolution†,‡. **Biochemistry**, [S.L.], v. 38, n. 15, p. 4826-4833, abr. 1999. American Chemical Society (ACS). http://dx.doi.org/10.1021/bi982648i.

TAIZ, L.; ZEIGER, E.; MOLLER, I. M.; MURPHY, A. “Cell walls: structure, formation, and expansion” in **Plant Physiology and Development**, 6th Edition, p. 379-405, 2015. Sinauer Associates.

TALHINHAS, Pedro *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**, [S.L.], v. 5, n. 0, p. 1-18, 14 mar. 2014. Frontiers Media SA. http://dx.doi.org/10.3389/fpls.2014.00088.

TALHINHAS, Pedro *et al*. The coffee leaf rust pathogen Hemileia vastatrix: one and a half centuries around the tropics. **Molecular Plant Pathology**, [S.L.], v. 18, n. 8, p. 1039-1051, 25 jan. 2017. Wiley. http://dx.doi.org/10.1111/mpp.12512. Disponível em: https://pubmed.ncbi.nlm.nih.gov/27885775/. Acesso em: 14 dez. 2020.

TRAPNELL C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nat Protoc**. 2012 Mar 1;7(3):562-78. doi: 10.1038/nprot.2012.016. Erratum in: Nat Protoc. 2014 Oct;9(10):2513. PMID: 22383036; PMCID: PMC3334321.

VOEGELE, Ralf T.; MENDGEN, Kurt. Rust haustoria: nutrient uptake and beyond. **New Phytologist**, v. 159, p. 93–100, 2003. doi: 10.1046/j.1469-8137.2003.00761.x

VOEGELE, Ralf T.; HAHN, Matthias; MENDGEN, Kurt. The Uredinales: Cytology, Biochemistry, and Molecular Biology. In: ESSER, Karl. **The Mycota**: a comprehensive treatise on fungi as experimental systems for basic and applied research. 2. ed. X: Springer, 2009. p. 69-98.

VORWERK, Sonja; SOMERVILLE, Shauna; SOMERVILLE, Chris. The role of plant cell wall polysaccharide composition in disease resistance. **Trends In Plant Science**, [S.L.], v. 9, n. 4, p. 203-209, abr. 2004. Elsevier BV. http://dx.doi.org/10.1016/j.tplants.2004.02.005.

WIEDERSTEIN, M.; SIPPL, M. J.. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, [S.L.], v. 35, n. , p. 407-410, 8 maio 2007. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gkm290.

ZAMBOLIM, Laércio. Current status and management of coffee leaf rust in Brazil. **Tropical Plant Pathology**, [S.L.], v. 41, n. 1, p. 1-8, 25 jan. 2016. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s40858-016-0065-9. Disponível em: https://link.springer.com/article/10.1007/s40858-016-0065-9. Acesso em: 14 dez. 2020.

ZHANG, Han *et al*. DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. **Nucleic Acids Research**, [S.L.], v. 46, n. 1, p. 95-101, 16 maio 2018. Oxford University Press (OUP). http://dx.doi.org/10.1093/nar/gky418.

ZHAO, Yuanyuan *et al*. Advances in Imaging Plant Cell Walls. **Trends In Plant Science**, [S.L.], v. 24, n. 9, p. 867-878, set. 2019. Elsevier BV. http://dx.doi.org/10.1016/j.tplants.2019.05.009.

ZHAO, Zhongtao *et al*. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. **Bmc Genomics Volume**, v. 274, n. 14, p. 1-15, 23 abr. 2013. Disponível em: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-274. Acesso em: 07 dez. 2020.

ZOLOTNITSKY, G. *et al*. Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. **Proceedings Of The National Academy Of Sciences**, [S.L.], v. 101, n. 31, p. 11275-11280, 26 jul. 2004. Proceedings of the National Academy of Sciences. http://dx.doi.org/10.1073/pnas.0404311101.

5. PRODUÇÃO ACADÊMICA DURANTE O MESTRADO

Resumos expandidos publicados em anais de congressos

1. PAULINO, G. S. ; PEREIRA, J. S. ; RIBON, A. O. B. ; CAIXETA, E. T. ; MENDES, T. A. O. . PRODUÇÃO DE BIOPLÁSTICO A PARTIR DE RESÍDUOS DE CAFÉ E SEU USO COMO MATRIZ DE LIBERAÇÃO DE ANTIMICROBIANOS. In: X Simpósio de Pesquisa dos Cafés do Brasil, 2019, Vitória - ES. **X Simpósio de Pesquisa dos Cafés do Brasil, 2019.**

Apresentações de Trabalho

1. PEREIRA, JULIA S. Identificação de CAZymes ativas durante a infecção do fungo biotrófico H. vastatrix à folha de café por genome mining. **Seminário do PPG em Bioinformática da UFMG. 2020.**

2. PEREIRA, JULIA S. Otimização do processo de extração de celulose do café utilizando dados do genoma do fungo *Hemileia vastatrix*. **Workshop de Avaliação Discente do PPG em Bioinformática da UFMG. 2020.**

3. PEREIRA, JULIA S.; MORGAN, T. ; VENTORIM, R. Z. ; SENRA, R. L. ; CASTRO, I. S. L. ; CAIXETA, E. T. ; MENDES, T. A. O. Analysis of expression of xylanases encoded in the genome of rust coffee fungus during different stages of infection. 2020. (Apresentação de Poster) **X-Meeting Experience 2020.**

4. PAULINO, G. S. ; PEREIRA, J. S. ; RIBON, A. O. B. ; CAIXETA, E. T. ; MENDES, T. A. O. . PRODUÇÃO DE BIOPLÁSTICO A PARTIR DE CELULOSE EXTRAÍDA DE CAFÉ E AVALIAÇÃO DA LIBERAÇÃO DE COMPOSTOS ANTIMICROBIANOS. 2019. (Apresentação Oral). **X Simpósio de Pesquisa dos Cafés do Brasil, 2019.**

Participação como ouvinte em eventos, congressos, exposições e feiras

1. Minicurso: Analysis of Molecular Dynamics Simulations Data Using Python. Women in Bioinformatics and Data Science LA. 2020. (Workshop).

2. Workshop de imunologia: atualizações imunológicas em tempos de COVID-19. 2020. (Workshop).

3. X-meeting eXperience 2020. 2020. (Encontro).

4. Workshop de Avaliação Discente do PPG em Bioinformática da UFMG. 2020.

5. X Simpósio de Pesquisa dos Cafés do Brasil. 2019. (Simpósio).