



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

BEATRIZ LORENA COMIN DA COSTA

**CARACTERIZAÇÃO DE EFETORES DA FAMÍLIA EGH16-  
LIKE DE *Phakopsora pachyrhizi***

---

Londrina  
2021

BEATRIZ LORENA COMIN DA COSTA

**CARACTERIZAÇÃO DE EFETORES DA FAMÍLIA EGH16-  
LIKE DE *Phakopsora pachyrhizi***

Trabalho de Dissertação apresentado ao Departamento de Agronomia da Universidade Estadual de Londrina, como requisito à obtenção do título de Mestre em Agronomia.

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Maria Isabel Balbi-Peña.  
Co-orientadora: Dr<sup>ª</sup> Francismar Corrêa Marcelino-Guimarães.

Londrina  
2021

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

B369c Costa, Beatriz Lorena Comin da.  
Caracterização de efetores da família Egh16-like de *Phakopsora pachyrhizi* / Beatriz Lorena Comin da Costa. - Londrina, 2021.  
84 f. : il.

Orientador: Maria Isabel Balbi-Peña.  
Coorientador: Francismar Corrêa Marcelino-Guimarães.  
Dissertação (Mestrado em Agronomia) - Universidade Estadual de Londrina, Centro de Ciências Agrárias, Programa de Pós-Graduação em Agronomia, 2021.  
Inclui bibliografia.

1. Fitopatologia - Tese. 2. Ferrugem asiática da soja - Tese. 3. Interação planta-patógeno - Tese. 4. Transcriptoma - Tese. I. Balbi-Peña, Maria Isabel . II. Marcelino-Guimarães, Francismar Corrêa . III. Universidade Estadual de Londrina. Centro de Ciências Agrárias. Programa de Pós-Graduação em Agronomia. IV. Título.

CDU 63

BEATRIZ LORENA COMIN DA COSTA

**CARACTERIZAÇÃO DE EFETORES DA FAMÍLIA EGH16-  
LIKE DE *Phakopsora pachyrhizi***

Trabalho de Dissertação apresentado ao Departamento de Agronomia da Universidade Estadual de Londrina, como requisito à obtenção do título de Mestre em Agronomia.

**BANCA EXAMINADORA**

---

Orientadora: Prof<sup>ª</sup>. Dra. Maria Isabel Balbi-Peña  
Universidade Estadual de Londrina - UEL

---

Prof<sup>ª</sup>. Dra. Mayra C. C. Gallo de Carvalho  
Universidade Estadual do Norte do Paraná - UENP

---

Dra. Viviane Yumi Baba  
Universidade Estadual de Londrina - UEL

Londrina, 31 de Agosto de 2021.

*Aos meus pais, Helena e Laudenor,  
e a minha avó Placidina,  
que são a minha maior fonte de inspiração e de amor.*

*DEDICO.*

## AGRADECIMENTOS

A Deus, por tudo que me proporciona e por ter me sustentado em todos os momentos, não me deixando desistir.

Aos meus pais, Helena e Laudenor, por todo amor, apoio e por serem meu exemplo. A distância nos faz mais fortes, assim como o amor que nos une. Vocês estão em tudo que faço e em cada conquista.

Às outras duas mulheres importantes da minha vida, minha “vovó” Placidina, que é meu exemplo de mulher batalhadora e resiliente, e que mesmo longe me fortalece dia após dia, e à minha irmã Bruna, que durante este período enfrentou uma dura batalha em sua vida, me ensinando a não desanimar e a não abaixar a cabeça para nenhum problema.

Ao meu companheiro de vida, Gustavo Okano, que me fortaleceu em cada momento, que enxugou minhas lágrimas, dividiu comigo os bons momentos deste período e me tira da zona de conforto, me incentivando sempre a buscar o melhor e ir além.

À minha orientadora, Prof<sup>a</sup>. Dra. Maria Isabel Balbi Peña, por todo o conhecimento e amor pela Fitopatologia transmitidos. Agradeço pela oportunidade, por todas as considerações e incentivo, pelos desafios propostos e pela paciência que teve comigo. ¡Gracias, profe!

À minha co-orientadora, Dra. Francismar Corrêa Marcelino Guimarães, por ter me recebido tão gentilmente em seu laboratório e me permitido vivenciar a Biologia Molecular. Obrigada pelas oportunidades, por cada palavra de incentivo, por toda experiência compartilhada, por buscar sempre o melhor para seus alunos e por todos os feedbacks que fizeram toda diferença!

À Dra. Valéria Yukari Abe, pelo auxílio no desenvolvimento deste trabalho, pela paciência e dedicação em transmitir todo seu conhecimento para que eu pudesse me desenvolver profissionalmente e por todas as conversas e experiências trocadas.

À Embrapa Soja, em especial ao laboratório de Biotecnologia Vegetal, local onde tive o privilégio de passar grande parte dessa jornada; pela oportunidade de desenvolver minha pesquisa em um ambiente acolhedor e que respira ciência.

Aos funcionários da Embrapa Soja, em especial às analistas de laboratório, Adriana Brombini e Silvana Rockenbach, por todo o suporte em cada etapa da minha pesquisa, por compartilharem comigo seus conhecimentos e por todas as conversas. Agradeço também à Vera Pierotti por me alegrar e zelar por mim, e a todos os funcionários que contribuíram de alguma forma no meu desenvolvimento.

Aos colegas de laboratório, Fernanda Castanho, Roberta Costa e Everton Ferreira, que não só contribuíram para a execução deste trabalho, como também contribuíram para que essa jornada fosse mais leve e divertida. Obrigada pelas conversas, ensinamentos, “cafés” e almoços compartilhados.

A todos que passaram por mim e contribuíram de alguma forma para minha formação, seja em sala de aula, em laboratório, em congressos ou em comemorações.

À Universidade Estadual de Londrina e ao Programa de Pós-Graduação em Agronomia, pela oportunidade e por todo conhecimento adquirido.

À CAPES pelo suporte financeiro.

*“Querer liberta: pois querer é criar. E somente para  
criar deves aprender.”*

*(Friedrich Nietzsche)*

DA COSTA, Beatriz Lorena Comin. **Caracterização de efetores da família Egh16-like de *Phakopsora pachyrhizi***. 2021. 84 f. Dissertação (Mestrado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2021.

## RESUMO

A ferrugem-asiática da soja (FAS), causada pelo fungo *Phakopsora pachyrhizi* constitui uma das principais ameaças à cultura da soja em regiões tropicais, tais como nos principais países produtores dessa cultura na América do Sul. Estratégias de controle a este patógeno tem se mostrado ineficientes devido à sua alta variabilidade e capacidade adaptativa, sendo recorrente a perda da sensibilidade aos fungicidas e aos genes de resistência *Rpp*. Para ser capaz de infectar e causar doença, este fungo biotrófico necessita estabelecer um íntimo contato com seu hospedeiro, a fim de secretar moléculas capazes de favorecer o processo infeccioso. Tais moléculas são denominadas efetores e atuam suprimindo e/ou manipulando as respostas de defesa da planta em favor da infecção. Até o momento, diferentes estudos de transcriptomas do fungo têm fornecido listas de potenciais efetores, sendo que poucos trabalhos tem efetivamente validado tais candidatos funcionalmente, principalmente devido à dificuldade de manipulação do fungo *in vitro*. Recentemente, com a disponibilidade de seu genoma de referência, o conjunto completo de efetores pode ser predito favorecendo os estudos de diversidade e evolução de famílias de efetores. Neste trabalho, um conjunto de proteínas do fungo candidatas a efetores apresentando alta similaridade entre si e compartilhando o mesmo domínio funcional Egh16-like foi avaliado *in silico* com base nas proteínas preditas no genoma de referência do isolado MT2006. Tal domínio já foi descrito e caracterizado em outras espécies de fitopatógenos, sendo demonstrada sua importância para a penetração na cutícula do hospedeiro. Um total de 22 modelos gênicos compartilhando elevada similaridade e presença do domínio Egh16-like foram identificados e anotados manualmente no genoma de referência MT2006. A análise do perfil de expressão de pelo menos um membro de cada clado da árvore filogenética, demonstrou sete clados contendo efetores potencialmente ativos nas fases iniciais da infecção, coincidindo com a formação do haustório. A análise comparativa desses efetores com as proteínas dos genomas de outras espécies de fungos revelou ainda efetores conservados entre as diferentes espécies, bem como aqueles exclusivos de Pucciniales e ainda aqueles apenas identificados em *P. pachyrhizi*. Finalmente, dois efetores desta família, apresentando diferentes cópias no genoma do fungo e organizadas lado a lado num mesmo scaffold, apresentaram-se sintênicos entre dois isolados (MT2006 e UFV02). Com base nestes resultados, acredita-se que esta família de efetores tenha um papel importante na virulência do fungo durante a interação com a soja.

**Palavras-chave:** Biotecnologia. Ferrugem-asiática da soja. *Glycine max*. Transcriptoma.

DA COSTA, B.L.C. **Characterization of effectors of the Egh16-like family of *Phakopsora pachyrhizi***. 2021. 84 p. Dissertation (Master in Agronomy) – Universidade Estadual de Londrina, Londrina, 2021.

## ABSTRACT

The asian soybean rust (ASR), caused by the fungus *Phakopsora pachyrhizi* is one of the main threats to the soybean crop in tropical regions, such as in the main producing countries of this crop in South America. Control strategies for this pathogen have been shown inefficient due to their high variability and adaptive capacity, with recurrent loss of sensitivity to fungicides and *Rpp* resistance genes. To be able to infect and cause disease, this biotrophic fungus needs to establish close contact with its host, in order to secrete molecules capable of favoring the infectious process. Such molecules are called effectors and act by suppressing and/or manipulating the plant's defense responses in favor of infection. Until now, different studies based on the analysis of the reference genomes and the fungus transcriptome have provided lists of potential effectors, and few studies have effectively validated such candidates functionally, mainly due to the difficulty of manipulating the fungus *in vitro*. Recently, with the availability of its reference genome, the complete set of effectors can be predicted favoring studies of diversity and evolution of effector families. In this work, a set of effector candidate proteins from the fungus showing high similarity and sharing the same functional Egh16-like domain was evaluated *in silico* based on the predicted proteins in the reference genome of the MT2006 isolate. This domain has already been described and characterized in other species of plant and insect pathogens, and its importance for the penetration of the host cuticle by pathogens has been demonstrated. A total of 22 gene models sharing high similarity and the Egh16-like domain were identified and manually annotated in the reference genome MT2006. The analysis of the expression profile of at least one member of each clade of the phylogenetic tree showed seven clades containing potentially active effectors in the initial stages of the infection, coinciding with the formation of the haustorium. The comparative analysis of these effectors with the proteins from the genomes of other fungal species also revealed effectors conserved among the different species, as well as those unique to Pucciniales and those only identified in *P. pachyrhizi*. Finally, two effectors from this family, presenting different copies in the fungus genome organized side by side in the same scaffold, were syntenic between two isolates (MT2006 and UFV02). Based on these results, it is believed that this family of effectors plays an important role in the virulence of the fungus during interaction with soybean.

**Keywords:** Asian soybean rust. Biotechnology. *Glycine max*. Transcriptome.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b>	Estágios de desenvolvimento do fungo <i>Phakopsora pachyrhizi</i> durante interação com a superfície foliar do hospedeiro.....	20
<b>Figura 2</b>	Imunidade inata e defesa das plantas durante o ataque de fitopatógenos.....	24
<b>Figura 3</b>	Árvore filogenética dos membros preditos no genoma de <i>Phakopsora pachyrhizi</i> que possuem o domínio Egh16-like.....	44
<b>Figura 4</b>	Representação da estrutura molecular dos candidatos a proteínas efetoras de acordo com o modelo gênico no genoma de referência MT2006.....	48
<b>Figura 5</b>	Análise filogenética dos candidatos a efetores de <i>Phakopsora pachyrhizi</i> comparados com outras espécies de fungos fitopatogênicos.....	52
<b>Figura 6</b>	Análise da sintenia dos genomas de referência de <i>Phakopsora pachyrhizi</i> contendo genes candidatos a efetores.....	60
<b>Figura 7</b>	Análise da sintenia dos genomas de referência de <i>Phakopsora pachyrhizi</i> contendo proteínas candidatas a efetores.....	61

## LISTA DE TABELAS

<b>Tabela 1</b>	Códigos das sequências transcritas preditas, membros da família 1, utilizados para buscas contra o genoma “ <i>Phakopsora pachyrhizi</i> MT2006 v1.0” e os modelos gênicos correspondentes.....42
<b>Tabela 2</b>	Candidatos a efetores de <i>Phakopsora pachyrhizi</i> escolhidos para o estudo. Os códigos referem-se a identidade das proteínas em cada um dos três genomas de referência, disponíveis no banco de dados do Joint Genome Institute.....46
<b>Tabela 3</b>	Dados oriundos da busca pelo domínio conservado nas sequências preditas dos candidatos de <i>Phakopsora pachyrhizi</i> selecionados para o estudo.....49
<b>Tabela 4</b>	Agrupamento das proteínas preditas de <i>Phakopsora pachyrhizi</i> nos três genomas de referência disponíveis e proteínas preditas de 14 genomas de fungos fitopatogênicos.....50
<b>Tabela 5</b>	Análise de expressão dos oito candidatos a efetores de <i>Phakopsora pachyrhizi</i> durante a interação com a soja.....54
<b>Tabela 6</b>	Expressão gênica obtida por RNA-Seq dos modelos gênicos correspondentes aos candidatos a efetores de <i>Phakopsora pachyrhizi</i> durante a interação com a soja e na ausência dela.....56
<b>Tabela 7</b>	Identificação dos candidatos a efetores de <i>Phakopsora pachyrhizi</i> , sua proteína correspondente e sua localização nos genomas de referência.....59

## ABREVIATURAS E SIGLAS

cDNA	DNA complementar
CMES	Coleção de Microrganismos da Embrapa Soja
Ct	<i>Cycle Threshold</i>
DMI	<i>Demethylation Inhibitor</i>
DPM	Duração do Período de Molhamento Foliar
DUF	<i>Domain of Unknown Function</i>
EDV	<i>Effector Detector Vector</i>
ETI	<i>Effector-Triggered Immunity</i>
ETS	<i>Effector Triggered Susceptibility</i>
FAS	Ferrugem Asiática da Soja
Gbp	<i>Giga base pairs</i>
Hpi	Horas pós inoculação
HR	<i>Hypersensitive Response</i>
<i>hrp</i>	<i>hypersensitive response and pathogenicity</i>
JGI	<i>Joint Genome Institute</i>
LGE	Laboratório de Genômica e Expressão
MAMPs	<i>Microbe-Associated Molecular Patterns</i>
NLR	<i>Nucleotide-binding Leucine-Rich Repeat</i>
ORF	<i>Open Reading Frame</i>
PAMPs	<i>Pathogen-Associated Molecular Patterns</i>
<i>PpEC</i>	<i>P. pachyrhizi Effector Candidate</i>
PRRs	<i>Pattern Recognition Receptors</i>
PTI	<i>PAMP-triggered immunity</i>
QoIs	<i>Quinone Outside Inhibitors</i>
RB	<i>Reddish Brown</i>
RNA-Seq	<i>RNA Sequencing</i> (Sequenciamento de alto rendimento de mRNAs)
ROS	Reactive Oxygen Species
<i>Rpp</i>	<i>Resistance to Phakopsora pachyrhizi</i>
RT-qPCR	<i>Real Time Quantitative Polimerase Chain Reaction</i>
SDHIs	<i>Succinate Dehydrogenase Inhibitors</i>
TAN	<i>Tannish</i>
TEs	<i>Transposable Elements</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO GERAL</b> .....	14
<b>2</b>	<b>REVISÃO DE LITERATURA</b> .....	17
2.1	IMPORTÂNCIA DA CULTURA DA SOJA.....	17
2.2	Ferrugem-Asiática da Soja.....	18
	2.2.1 O Fungo <i>Phakopsora pachyrhizi</i> .....	19
	2.2.2 Sintomatologia.....	21
	2.2.3 Controle.....	21
2.3	Interação planta-patógeno.....	22
	2.3.1 Proteínas efetoras: caracterização e métodos de análise.....	25
	2.3.2 Proteínas efetoras de fungos causadores de ferrugem.....	27
2.4	Domínios funcionais.....	29
	2.4.1 Domínio Egh16-like.....	29
<b>3</b>	<b>ARTIGO A: CARACTERIZAÇÃO DE EFETORES DA FAMÍLIA EGH16 DE <i>Phakopsora pachyrhizi</i></b> .....	
3.1	RESUMO.....	31
3.2	ABSTRACT.....	33
3.3	INTRODUÇÃO.....	35
3.4	MATERIAIS E MÉTODOS.....	36
	3.4.1 Análise e seleção dos candidatos a efetores.....	36
	3.4.2 Comparação filogenética entre espécies de fungos fitopatogênicos e <i>P. pachyrhizi</i> .....	37
	3.4.3 Material biológico.....	38
	3.4.4 Perfil de expressão de genes candidatos a efetores de <i>P. pachyrhizi</i> durante o ciclo infeccioso.....	39
	3.4.4.1 Extração de RNA total e síntese de cDNA.....	39
	3.4.4.2 Análise RT-qPCR.....	40
	3.4.4.3 Validação dos níveis de expressão gênica de RNA-Seq por RT- qPCR.....	41
	3.4.5 Análise de sintenia.....	41
3.5	RESULTADOS E DISCUSSÃO.....	41

3.5.1 Predição de candidatos a efetores de <i>P. pachyrhizi</i> contendo o domínio Egh16-like com base no genoma do isolado MT2006.....	41
3.5.2 Análise de diversidade genética entre os membros da família Egh16-like.....	43
3.5.3 Caracterização molecular dos genes candidatos e efetores Egh16-like.....	46
3.5.4 Análises comparativas de proteínas candidatas a efetores de <i>P. pachyrhizi</i> em diferentes filos e espécies.....	49
3.5.5 Perfil de expressão de genes candidatos a efetores de <i>P. pachyrhizi</i> durante interação com a soja.....	53
3.5.6 Validação dos efetores via expressão gênica a partir de dados de RNA-Seq.....	54
3.5.7 Análise de sintonia.....	59
3.6 CONCLUSÃO.....	62

<b>REFERÊNCIAS</b> .....	64
--------------------------	----

<b>APÊNDICES</b> .....	75
------------------------	----

APÊNDICE A – Sequências dos primers dos candidatos a efetores utilizados para análise de expressão gênica por RTqPCR, tamanho aproximado do fragmento amplificado (pb) e sua respectiva temperatura de anelamento (TA).....	76
---	----

APÊNDICE B – Modelos gênicos do genoma de <i>Phakopsora pachyrhizi</i> MT2006 com similaridades aos candidatos secretados descritos por De Carvalho et al. (2006), contendo o domínio PFAM11327/DUF3129.....	77
--	----

APÊNDICE C – Alinhamento da sequência do de_novo_555 e o seu modelo gênico correspondente MT2006_4582009.....	78
---	----

APÊNDICE D – Alinhamento da sequência do de_novo_635 e o scaffold expandido do seu modelo gênico correspondente no genoma MT2006.....	80
---	----

APÊNDICE E – Folhas de soja inoculadas com esporos de <i>Phakopsora pachyrhizi</i> e lesões do tipo “TAN”, características de genótipos suscetíveis a este fungo, utilizadas para o estudo da expressão dos candidatos a efetores.....	82
--	----

APÊNDICE F – Análise de expressão dos oito candidatos a efetores de <i>Phakopsora pachyrhizi</i> durante a interação com a soja.....	83
--	----

APÊNDICE G – Expressão gênica obtida por RNA-Seq dos modelos gênicos correspondentes aos candidatos a efetores de <i>Phakopsora pachyrhizi</i> .....	84
--	----

## 1 INTRODUÇÃO GERAL

O fungo biotrófico *Phakopsora pachyrhizi* (Syd. & P. Syd) é o agente causal de uma das doenças mais importantes que acometem a cultura da soja, a ferrugem-asiática da soja (FAS). Trata-se de um patógeno biotrófico obrigatório que necessita das células vivas do hospedeiro para sobreviver e se reproduzir de forma assexuada, formando os uredósporos (HARTMAN et al., 2015). Em condições climáticas favoráveis, esses uredósporos originam novas infecções e servem como fonte de inóculo para outras lavouras de soja, podendo ainda sobreviver em plantas voluntárias ou em hospedeiros alternativos (ZAMBOLIM, 2006).

Uma vez presente na lavoura, os sintomas da doença são visíveis em 7-9 dias após a infecção e são caracterizados por lesões na superfície foliar, que podem variar de coloração castanho clara (lesões TAN) ou castanho avermelhada (lesões RB) (MILES et al., 2011). Na parte abaxial da folha, cada lesão produz urédias que envolvem os uredósporos produzidos no interior da folha e que são dispersados quando há o rompimento da cutícula foliar. Conforme o progresso da doença, ocorre o aumento do número de lesões, promovendo o amarelecimento das folhas e, posteriormente, a desfolha precoce. A perda da área fotossintética ocasiona a redução do número de vagens por planta e também a capacidade de enchimento dos grãos, resultando em consequências na produtividade (REIS; BRESOLIN; CARMONA, 2006).

Desde seu surgimento no Brasil, os custos com a doença ultrapassam US\$2 bilhões por safra, relativos ao custo com fungicidas e às perdas na produção (CAF, 2018). Sem que haja a adoção de medidas de controle, as perdas de rendimento podem chegar a 90% (GODOY et al., 2016a; HARTMAN et al., 2015). As principais medidas de controle são práticas culturais e o uso de cultivares com ciclo mais reduzido para escapar do período mais propenso à doença; além do uso de fungicidas e cultivares contendo genes de resistência. A prática cultural de maior eficiência tem sido o vazio sanitário, que consiste na erradicação de plantas de soja durante 60-90 dias, na entressafra, visando a eliminação total do hospedeiro, associado com o uso de cultivares resistentes e precoces (GODOY et al., 2016b). O uso de fungicidas é o método de controle mais utilizado para a doença e estes atuam em sítios específicos que desempenham uma função vital na célula do patógeno, ocasionando sua morte. Entretanto, a rápida perda de eficiência destes produtos devido a adaptação do patógeno impulsiona a pesquisa por novas formas de controle do fungo e evidencia que nenhuma alternativa de controle é eficaz se utilizada de maneira isolada (GODOY et al., 2016b). Finalmente, cultivares resistentes à ferrugem-asiática estão disponíveis contendo um ou uma

combinação de genes *Rpp* (Resistência a *Phakopsora pachyrhizi*) em alguns países da América do Sul, como no Brasil.

Para causar a infecção, o fungo secreta proteínas no interior da célula hospedeira, denominadas de efetores. Estas proteínas são secretadas durante o contato mais íntimo entre o fungo e seu hospedeiro, por uma estrutura denominada haustório, sendo lançadas no apoplasto ou no interior das células da planta. Ao serem reconhecidas, resultam na imunidade desencadeada por efetores (ETI), que induz o sistema imune a ativar de forma rápida e robusta os seus recursos de defesa, como o acúmulo de calose, produção de ROS e reação de hipersensibilidade, que limitam a progressão do patógeno (PASCHOLATI; DALIO, 2018). Conhecer como atuam os efetores na ativação das respostas de defesa é de suma importância, porém, o isolamento e caracterização de efetores de fungos de ferrugem ainda é escasso, devido à dificuldade de estudos *in vitro* com esses microrganismos e à falta de genomas de referência (BAKKEREN; SZABO, 2020; DUPLESSIS et al., 2021; PETRE et al., 2014).

Previamente, De Carvalho et al. (2016) identificaram um grupo de proteínas preditas como secretadas obtidas através do transcriptoma de *P. pachyrhizi* expresso durante a interação com a soja. Na época, o estudo comparativo entre esta família de proteínas e as preditas nos genomas de espécies de *Puccinia graminis* f. sp. *tritici* e *Melampsora larici-populina*, ambas espécies de ferrugens, revelou uma família conservada presente nos fungos causadores de ferrugem, contendo 39 membros, sendo 22 destes específicos de *P. pachyrhizi*. A maioria dos membros descritos apresentaram sequências de tamanho reduzido ( $\leq 250$  aminoácidos), resíduos ricos em cisteínas ( $\geq 3\%$ ), além da presença do domínio Egh16-like (DUF3129/PF11327) em todas as sequências, características estas de proteínas amplamente codificadas nos genomas de alguns fungos fitopatogênicos e que tem ação na virulência desses microrganismos (HUANG et al., 2019).

O domínio Egh16-like é um fator de virulência que desempenha papel importante durante a fase inicial de infecção em patógenos fúngicos (XUE et al., 2002). É característico de *Blumeria graminis* f. sp. *hordei* mas também já foi observado em proteínas oriundas de *P. graminis* f. sp. *tritici* e *M. larici-populina* (DUPLESSIS et al., 2011; SAUNDERS et al., 2012), entretanto sua função ainda não é totalmente compreendida. No fungo patogênico de insetos, *Metarhizium robertsii*, a função deste domínio está associada a infecção dos hospedeiros, sendo que a deleção desses genes interfere na capacidade desses fungos em penetrar a cutícula do seu hospedeiro, afetando sua virulência (HUANG et al., 2019; XUE et al., 2002). Por abranger o genoma de diversas espécies de fungos fitopatogênicos, compreender

a coevolução dos microrganismos que possuem este domínio e seus hospedeiros deve ser ponderada.

Recentemente, três genomas de referência de isolados Sul Americanos de *P. pachyrhizi* estão disponíveis para a comunidade científica, fornecendo informação sobre todo o arsenal gênico que caracteriza este parasita. A partir destes dados, todo o conjunto de efetores pode ser caracterizado, inclusive famílias ou membros relacionados, com diferentes cópias e como é sua ordem ao longo dos genomas. Tais dados disponíveis associados, por exemplo, dados de expressão em larga escala em estudos transcriptômicos e filogenéticos, podem revelar candidatos com papéis importantes na (a)virulência e no parasitismo deste fitopatógeno.

Assim, o objetivo dessa pesquisa foi caracterizar estruturalmente membros geneticamente diversos da família de proteínas candidatas a efetores de *P. pachyrhizi* contendo o domínio Egh16-like (DUF3129/PF11327), validar sua função efetora via caracterização do seu perfil de expressão durante a infecção, além de caracterizar filogeneticamente suas relações com outros fungos.

## 2. REVISÃO DE LITERATURA

### 2.1 IMPORTÂNCIA DA CULTURA DA SOJA

A soja (*Glycine max* (L.) Merrill) é uma planta herbácea, da classe *Magnoliopsida*, pertencente à ordem *Fabales* e à família *Fabaceae* (*Leguminosae*) (EMBRAPA-CNPSO, 2008). É uma das culturas mais produzidas no Brasil e no mundo, sendo que a produção brasileira é uma das mais expressivas entre os países produtores (CONAB, 2019b). Grande parte da soja produzida no Brasil é destinada a países orientais, sendo a China o maior país importador da soja brasileira, além de ser negociada por produtores de diversos países, apresentando forte liquidez na sua comercialização e alta demanda (CONAB, 2020; LAZZAROTTO; HIRAKURI, 2010). Além de a produção ser destinada para exportação, os preços da soja brasileira são definidos por cotações de outros países por ser uma *commodity*, ou seja, a produção no país gira em torno da economia mundial, principalmente Estados Unidos e China.

Desde os anos 1970, a soja figura como uma das grandes culturas produtoras de grãos, sendo que a oleaginosa foi a que teve o maior percentual de crescimento mundialmente. De 1970 a 2017, a produção global de soja cresceu 7,73 vezes – de 44 para 340 milhões de toneladas (GAZZONI; DALL’AGNOL, 2018). O grande aumento na produção mundial de soja pode ser atribuído a diversos fatores, como o elevado teor de proteínas (aproximadamente 40%) e óleos (em torno de 20%) encontrados no grão, além do complexo de vitaminas (A, E, B1, B2 e niacina); e, principalmente, ao desenvolvimento expressivo das tecnologias de produção, que permitiram a exploração de novas fronteiras agrícolas, ampliando significativamente a área cultivada e a produtividade da oleaginosa (LAZZAROTTO; HIRAKURI, 2010).

Devido ao seu valor e sua importância ao redor do mundo, os produtores brasileiros investem cada vez mais na produção do grão de soja e em novos locais de cultivo, abrangendo 19 dos 26 estados brasileiros, além do Distrito Federal (USDA, 2020). Na safra brasileira de 2018/19 houve crescimento de 1,9% na área de plantio e redução de 3,7% na produção, atingindo 114,8 milhões de toneladas. As regiões Centro-Oeste e Sul representam mais de 78% dessa produção, sendo o estado do Mato Grosso o principal produtor de soja, com produção de 35.434,5 mil toneladas na safra 2019/20, incremento de 9,2% em relação à safra anterior (CONAB, 2020). O estado do Paraná retomou a posição de segundo maior produtor nacional nessa safra, visto que na anterior, o estado do Rio Grande do Sul tornou-se o segundo

maior produtor, produzindo quase 3.000 toneladas a mais que o Paraná (CONAB, 2020; CONAB, 2019b). A quebra de produção no estado do Paraná na safra 2018/19 e no Rio Grande do Sul na safra 2019/20 deveu-se às condições climáticas adversas durante o ciclo da cultura, principalmente à falta do regime de chuvas, diminuindo a eficiência produtiva da planta (USDA, 2019).

Além das condições climáticas desfavoráveis, fatores bióticos, como as doenças e pragas, também contribuem com a diminuição da eficiência produtiva das plantas. Entre eles, as doenças causam grandes prejuízos aos produtores, sendo a ferrugem-asiática da soja a mais agressiva.

## 2.2 Ferrugem Asiática da Soja

A cultura da soja pode ser acometida por duas espécies fúngicas pertencentes ao gênero *Phakopsora*, que são agentes causais da ferrugem: *Phakopsora pachyrhizi* (Sydow & P. Sydow) e *Phakopsora meibomiae* (Arthur) Arthur. Até 1992, acreditava-se que apenas uma espécie, *P. pachyrhizi*, era a causadora da doença. Entretanto, através de estudos sobre a morfologia de populações destas espécies, evidenciou-se que o fungo *P. meibomiae* correspondia as populações que aconteciam naturalmente em diversas espécies de leguminosas na América, além de ser menos agressivo; enquanto que, *P. pachyrhizi* era o agente causal que ocorria em populações da Ásia e Austrália (GODOY et al., 2016; USDA, 2015).

Os primeiros relatos de ferrugem-asiática foram no Japão, em 1902, enquanto que, no continente americano, os primeiros relatos surgiram mais de 70 anos depois, sendo relatada em Porto Rico em 1976. A espécie *P. pachyrhizi*, que até então, estava limitada aos países que cultivavam soja na Ásia e Austrália, foi detectada pela primeira vez fora desses países no Hawaii, em 5 de maio de 1994 (BONDE; PETERSON, 1996; KILLGORE, 1996; USDA, 2015; YORINORI et al., 2004). Já no Brasil, os primeiros esporos de ferrugem da soja surgiram em 2001 (GODOY et al., 2016). Entretanto, o fungo já foi encontrado em diversos países, como Austrália, China, Coreia, Índia, Japão, Nepal, Taiwan, Tailândia, Filipinas, Moçambique, Nigéria, Ruanda, Uganda, Zimbábwe, África do Sul, Estados Unidos, Brasil, Argentina e Paraguai (USDA, 2015).

Soja e kudzu (*Pueraria montana* var. *lobata*) são os principais hospedeiros, porém, existem mais de 90 hospedeiros alternativos conhecidos de *P. pachyrhizi*, como algumas espécies de feijões (*Cajanus cajan*, *Vigna radiata*, *Phaseolus coccineus*, entre outros) e algumas espécies de trevos (*Alysicarpus vaginalis*, *Trifolium repens*, *Trifolium aureum*, entre

outros) (RUPE; SCONYERS, 2008).

A doença é considerada, atualmente, como a mais severa da cultura da soja, podendo acarretar mais de 80% de perda de produtividade (GODOY et al., 2016). Na safra 2001/2002, em seu primeiro ano de ocorrência no Brasil, a perda em grãos foi de 569,2 mil toneladas, aproximadamente U\$125,5 milhões de dólares, abrangendo apenas os estados de Goiás, Mato Grosso, Mato Grosso do Sul, Paraná e Rio Grande do Sul. Tais perdas foram muito expressivas devido ao despreparo dos produtores em diagnosticar a doença e à inexistência de produtos registrados para o controle. Além disso, os produtos disponíveis que eram utilizados para controlar outras doenças eram do grupo dos benzimidazóis, que são ineficientes no controle dessa doença (CAF, 2019).

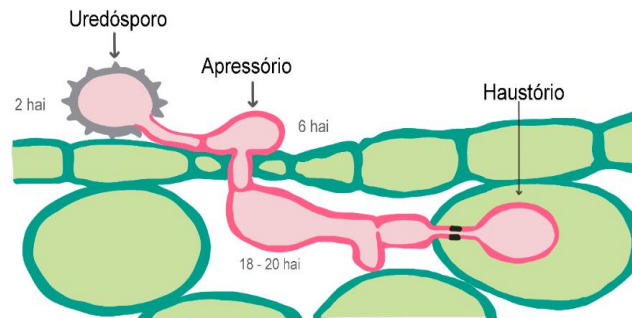
### 2.2.1 O fungo *Phakopsora pachyrhizi*

O agente causal da ferrugem-asiática da soja, *Phakopsora pachyrhizi*, pertence ao reino *Fungi*, filo *Basidiomycota*, classe *Pucciniomycetes*, ordem *Pucciniales*, família *Phakopsoraceae*. É um parasita obrigatório necessitando sempre da presença do tecido hospedeiro para seu estabelecimento. O fungo coloniza intercelularmente os tecidos vegetais e nutre-se pela absorção de nutrientes diretamente do interior do hospedeiro, por meio de estruturas especializadas, os haustórios (BEDENDO, 2018).

O processo infeccioso inicia com a deposição e germinação do uredósforo na superfície foliar do hospedeiro, sendo esta fase denominada de fase de infecção. Assim, ocorre a formação do tubo germinativo e do apressório, que é o responsável pela penetração direta (cutícula e epiderme) do fungo na superfície do hospedeiro, provocando a desorganização e necrose das células epidérmicas penetradas, iniciando a fase de colonização (KOCH et al., 1983). Após sua penetração, a hifa forma o haustório, que é uma estrutura especializada na retirada de nutrientes e que cresce intracelularmente, provocando a desorganização da parede celular e posterior invaginação da membrana celular, porém causa o menor dano possível ao hospedeiro (AMORIM; PASCHOLATI, 2018). Além de provocar mudanças estruturais na célula para a absorção de nutrientes, o haustório é também responsável pela produção e liberação de proteínas efetoras no citoplasma vegetal (AMORIM; PASCHOLATI, 2018; MENDGEN; HAHN, 2002). Desta forma, o patógeno estabelece relações parasitárias com o hospedeiro, sendo esta fase denominada de fase biotrófica. A última fase do ciclo do processo de infecção é a fase reprodutiva ou de esporulação, que é quando o patógeno se reproduz no tecido infectado, garantindo a perpetuação da espécie, e é então dispersado para tecidos não

parasitados (BROMFIELD et al., 1980; KOCH et al., 1983). O fungo *P. pachyrhizi* coloniza poucas porções do hospedeiro para rapidamente produzir suas estruturas reprodutivas e dispersá-las para partes ainda saudáveis, visto que é um parasita obrigatório e necessita da célula vegetal viva para sobreviver. Assim, a ferrugem-asiática é considerada uma doença policíclica, devido aos diversos ciclos de vida que o fungo possui em uma mesma safra da cultura (Figura 1) (AMORIM; PASCHOLATI, 2018).

**Figura 1** – Estágios de desenvolvimento do fungo *Phakopsora pachyrhizi* durante interação com a superfície foliar do hospedeiro. Os uredósporos germinam duas horas após inoculação (hpi), emitindo o tubo germinativo que desenvolve o apressório. Para penetrar na célula, uma hifa de penetração é emitida percorrendo os espaços intercelulares da célula vegetal. Dentro da célula do mesófilo forma-se o primeiro haustório, onde se estabelece a relação biotrófica entre patógeno e hospedeiro.



**Fonte:** Adaptado de Catanzariti et al. (2007)

Os fungos são microrganismos que dependem de condições climáticas favoráveis para sobrevivência, como umidade e temperaturas amenas. Além disso, a duração do período de molhamento foliar (DPM) também é um fator que contribui com a ocorrência de doenças fúngicas em plantas. Assim, as condições ideais para o agente causal da ferrugem germinar e formar lesões são temperaturas entre 18 e 26,5°C (temperatura ótima) e DPM foliar mínimo de seis horas (HARTMAN et al., 2015; HENNING et al., 2014; SEIXAS et al., 2018). Outro fator que influi sobre o fungo é o estágio fenológico da planta, sendo que a maior probabilidade de desenvolvimento da doença em condições de campo/lavoura é nos estágios mais avançados da cultura, onde ocorre o fechamento do dossel das plantas, o que cria um microclima favorável para a manutenção da umidade, além da proteção dos esporos contra efeitos da radiação ultravioleta e da radiação solar direta pelas folhas superiores do dossel (GODOY et al., 2016a). Portanto, lesões de ferrugem podem ser encontradas em qualquer estágio fenológico das plantas, entretanto, altos índices da doença não são encontrados até o período de floração da soja, caso não haja grande quantidade de inóculo na fase vegetativa

(IGARASHI et al., 2014; RUPE; SCONYERS, 2008).

### 2.2.2 Sintomatologia

Os sintomas da ferrugem-asiática podem ser observados nas folhas, cotilédones e hastes, sendo as folhas o local de maior ocorrência. Os primeiros sintomas são caracterizados por minúsculos pontos, de coloração esverdeada, porém, mais escuros que o tecido sadio da planta (YORINORI et al., 2003). A doença é confirmada com a visualização das urédias, que são as estruturas de reprodução do fungo e são pequenas lesões localizadas na face abaxial das folhas. As urédias têm de 2 a 5 mm de diâmetro, são circulares ou elípticas e de coloração amarelada, estando cobertas pela cutícula foliar. Entretanto, com o progresso da doença, as manchas adquirem uma coloração castanha, aumentam seu tamanho e a cutícula foliar é rompida, liberando os urediniósporos (BEDENDO, 2018). Assim, formam-se as pústulas, que são urédias formadas nas lesões e que produzem uma grande massa de urediniósporos. Tais urediniósporos possuem uma coloração diferente em relação aos outros patógenos causadores de ferrugem (vermelho-amarronzada), possuindo uma coloração amarelo-amarronzada e possuem uma superfície densamente equinulada e ornamentada (RUPE; SCONYERS, 2008).

As lesões são caracterizadas pela sua coloração e definem o grau de suscetibilidade da planta. Plantas suscetíveis apresentam lesões TAN (*Tanish*), que possuem coloração castanho-claro, com esporulação abundante e sem necrose extensiva. As lesões RB (*Reddish Brown*) são lesões marrom-avermelhadas, possuem menor número de urédias por lesão e pouca esporulação, e apresentam diversos pontos de necrose foliar, indicando uma reação de resistência. Plantas que não apresentam lesões macroscópicas em suas folhas são classificadas como imunes (GODOY et al., 2016a).

As folhas da soja infectadas pelo fungo perdem sua área fotossintética, diminuindo a eficiência no enchimento dos grãos e, conseqüentemente, o peso dos grãos.

### 2.2.3 Controle

As estratégias de manejo da ferrugem-asiática são fundamentalmente três: uso de práticas culturais, controle genético e controle químico. Dentre as práticas culturais, o vazio sanitário é o mais importante, pois consiste na eliminação de plantas de soja e ausência de cultivo na entressafra, durante 60 a 90 dias, eliminando qualquer fonte de inóculo do

patógeno. Ademais, o monitoramento das áreas para indicar com precisão o momento ideal para iniciar o controle com fungicidas, o uso de cultivares precoces a fim de reduzir o tempo da cultura no campo e a semeadura no início da época recomendada para diminuir a taxa de inóculo primário, auxiliam na minimização da doença (GODOY et al., 2016a).

Quanto ao controle genético, trata-se de uma medida eficiente e econômica, pois, comparado com os produtos fitossanitários, possui baixo impacto ambiental (CAMARGO, 2018a), entretanto, para sua obtenção é necessário investimento financeiro em pesquisas (BEDENDO, 2018). Além disso, sabe-se que os genes *Rpp* não são efetivos para a ampla variabilidade gênica do fungo, não proporcionando resistência ampla e durável para todos os isolados. Para a ferrugem-asiática, sete diferentes *loci* de resistência (*Rpp1* (HYTEN et al., 2007), *Rpp2* (SILVA et al., 2008), *Rpp3* (HYTEN et al., 2009), *Rpp4* (SILVA et al., 2008), *Rpp5* (GARCIA et al., 2008), *Rpp6* (LI et al., 2012) e *Rpp7* (CHILDS et al., 2018) – *Resistance to Phakopsora pachyrhizi*) já foram identificados e mapeados no genoma da soja.

O controle químico, através do uso de moléculas orgânicas ou inorgânicas para a proteção das plantas (SILVA JUNIOR; BEHLAU, 2018), é o mais utilizado e foi o mais eficiente até a safra 2007/08 (GODOY et al., 2016a; RUPE; SCONYERS, 2008). Os principais grupos de fungicidas utilizados no controle da FAS são os inibidores da quinona externa (QoI, “estrobilurinas”), inibidores da desmetilação de esteróis (DMI, “triazóis”) e os inibidores da enzima succinato desidrogenase (SDHI, “carboxamidas”) (SILVA JUNIOR; BEHLAU, 2018). Entretanto, desde o início do uso dos fungicidas no Brasil, há evidências da diminuição da eficácia dos produtos usados no controle dos fitopatógenos em cada safra. Tal fato deve-se a pressão de seleção exercida nesses microrganismos pelo uso excessivo dos produtos. Além disso, o tempo de semeadura da soja no país, propicia duas safras em algumas regiões brasileiras, aumentando, assim, o número de aplicações de fungicidas (GODOY et al., 2016b).

### 2.3 Interação Planta-Patógeno

As plantas são organismos sésseis e, portanto, não possuem células de defesa móveis e nem sistema imune adaptativo, sendo dependentes da troca de sinais entre si e os microrganismos para ativar seu sistema imune inato (ABE, 2012; CHISHOLM et al., 2006; DALIO et al., 2014; JONES; DANGL, 2006). Os mecanismos de defesa vegetal podem ser constitutivos (passivos) ou pós-formados (ativos). Cutícula, tricomas, fenóis e alcaloides são alguns dos fatores de resistência constitutivos e que estão presentes na planta independente de aspectos bióticos. Os fatores pós-formados são aqueles inexistentes ou que estão presentes em

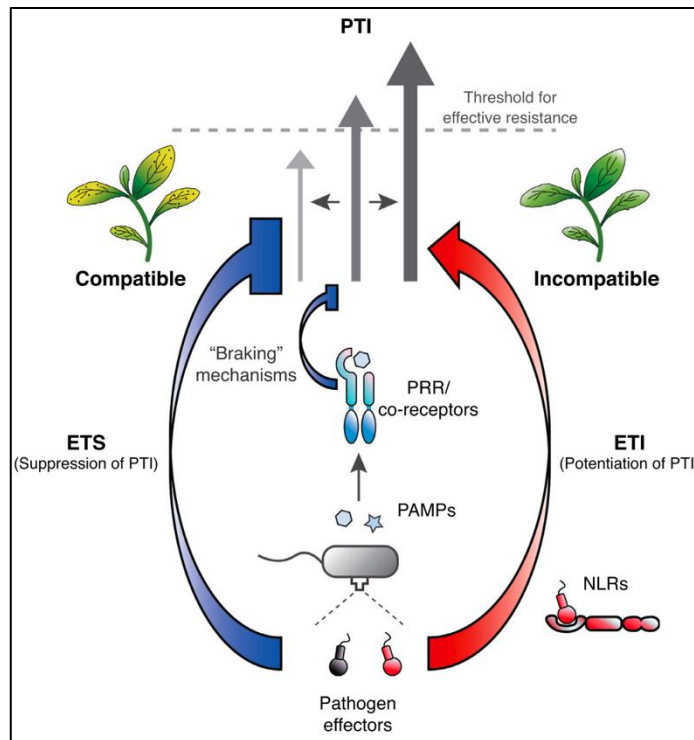
baixas quantidades antes da infecção, sendo ativados ou produzidos após a identificação do ataque (PASCHOLATI; DALIO, 2018). A fim de interferir no avanço dos fitopatógenos pelos tecidos, mudanças estruturais são acionadas, como a formação de papilas e halos. Compostos bioquímicos também são produzidos ou ativados, como a liberação de espécies reativas de oxigênio (ROS), que podem atuar na inibição do desenvolvimento fúngico nos tecidos vegetais ou então aumentando a integridade da membrana plasmática (JONES; DANGL, 2006). Além disso, as fitoalexinas também são compostos antimicrobianos e são consideradas prejudiciais a uma ampla gama de organismos, como fungos, bactérias, nematoides, animais e vegetais superiores (AHUJA et al., 2012).

Qualquer microrganismo que entre em contato com a célula vegetal é detectado inicialmente por meio de receptores de reconhecimento padrão (PRRs), localizados na membrana plasmática e que reconhecem características microbianas conservadas, denominadas de PAMPs ou MAMPs (*Pathogen- ou Microbe-Associated Molecular Patterns*) e que ativam o sistema imune das plantas (BOLLER; FELIX, 2009). Os PAMPs são moléculas conservadas e ditas essenciais na sobrevivência dos patógenos, mas não existem no hospedeiro, como a quitina para os fungos, que é componente principal da parede celular, e a flagelina para as bactérias, que é uma proteína que compõe os flagelos e é essencial para a virulência bacteriana (BOUTROT; ZIPFEL, 2017; DALIO et al., 2014; MEDZHITOV; JANEWAY, 1997; TAIZ et al., 2017). Esse reconhecimento confere a imunidade desencadeada por PAMPs (PTI – *PAMP-triggered immunity*), que ativa rapidamente o sistema de defesa basal contra uma ampla gama de patógenos e restringe o processo de colonização no hospedeiro (CHISHOLM et al., 2006; JONES; DANGL, 2006; ROUX et al., 2014).

Para suprimir as defesas primárias do hospedeiro ou PTI, os patógenos secretam proteínas efetoras diretamente no interior da célula vegetal que são reconhecidas pelos receptores intracelulares NLRs (*Nucleotide-binding Leucine-rich Repeat*). Quando há a interação dos efetores com os receptores NLRs ocorrem alterações conformacionais desses receptores que, posteriormente, são ativados e resultam na imunidade desencadeada por efetores (YUAN et al., 2021). Apesar de serem vias de sinalização distintas e desencadeadas por diferentes receptores, as respostas de PTI e ETI atuam sinergicamente no sistema imunológico das plantas e estão funcionalmente ligadas e com várias sobreposições nas rotas e cascatas de transdução de sinal. Ambas as respostas de defesa podem ocasionar a morte celular nos sítios de infecção, conhecida como resposta de hipersensibilidade (HR – *Hypersensitive response*), evitando o crescimento e desenvolvimento do patógeno nos tecidos vegetais, bem como a produção de espécies reativas de oxigênio (ROS) e o acúmulo de calose nos tecidos

(BENTHAM et al., 2020; NGOU et al., 2020; YUAN et al., 2021). Evidências sugerem que a ETI pode potencializar a PTI em resposta a múltiplos PAMPs e ainda restaurar a PTI atenuada pelos efetores. Em estudos com *Pseudomonas syringae*, foi observado que a ativação de apenas uma resposta imune isolada é insuficiente para controlar o patógeno (NGOU et al., 2020). Além disso, a ativação da resposta ETI mediada por diferentes genes NLR mostrou-se dependente dos receptores de membrana, sendo comprometida em linhas mutantes de diferentes PRRs oriundas de *Arabidopsis thaliana* em resposta a infecção por *P. syringae* pv. *tomato* (Pst) DC3000 (YUAN et al., 2021).

**Figura 2** – Imunidade inata e defesa das plantas durante o ataque de fitopatógenos. As plantas detectam padrões moleculares associados a patógenos (PAMPs) por meio de receptores de reconhecimento de padrões (PRRs) que induzem a imunidade desencadeada por padrões (PTI). Proteínas efectoras são secretadas pelos patógenos no interior da célula vegetal para suprimir a resposta PTI. Essas proteínas são reconhecidas por receptores NLRs que resultam em suscetibilidade desencadeada por efetores (ETS) e ativam a imunidade desencadeada por efetores (ETI), que potencializa e restaura a PTI suprimida pelo patógeno. PTI e ETI atuam sinergicamente para fornecer uma imunidade robusta contra fitopatógenos.



Fonte: Yuan et al., 2021.

Dois modelos exemplificam a interação entre as proteínas efectoras do patógeno e as proteínas de resistência do hospedeiro. A teoria Gene-a-Gene, formulado por Flor (1942) em estudos com o patossistema linho-*Melampsora lini*, propõe que cada proteína de

avirulência (Avr) do patógeno, possui um gene cognato de reconhecimento na planta, denominado de gene *R*, e a interação destes genes desencadeia a resposta de defesa da planta, ou seja, ocorre uma interação direta de receptor-ligante. Em *Arabidopsis*, a proteína efetora ATR1 é reconhecida diretamente pelo receptor RPP1 da planta, sendo que este reconhecimento ativa o sistema de defesa contra o oomiceto *Peronospora parasitica* (JONES; VANCE; DANGL, 2016; SCHREIBER et al., 2016). Outro modelo é o da “hipótese guarda”, que sugere o reconhecimento indireto por meio da interação entre um efector e uma proteína hospedeira, chamada *guardiee*. As proteínas *R* supervisionam a presença de efetores do patógeno sendo capazes de detectar modificações nas proteínas receptoras (STERGIOPOULOS; DE WIT, 2009). O receptor NBS-LRR é ativado quando o efector secretado no citoplasma pelo patógeno interage com a proteína *guardiee*. Essa interação modifica a conformação do receptor, devido a dissociação da proteína e do receptor, desencadeando uma cascata de sinalização que ativa o sistema de defesa do hospedeiro (TAIZ et al., 2017). Um exemplo deste modelo é o reconhecimento indireto de proteínas do nematoide *Globodera rostochiensis* no tomateiro. Na presença da proteína Cf-2 no hospedeiro, o reconhecimento do eliciador Gr-VAP1 disparado pelo nematoide é mediado pela proteína Rcr3<sup>pim</sup>, desencadeando a resposta de defesa. Entretanto, na ausência da Cf-2 ocorre um aumento da suscetibilidade da planta ao patógeno (LOZANO-TORRES et al., 2012).

### 2.3.1 Proteínas efectoras: caracterização e métodos de análise

Para a maioria dos microrganismos tornarem-se patogênicos, principalmente bactérias, fungos e oomicetos, é necessária sua penetração no interior das plantas, tanto por penetração direta, por ferimentos ou aberturas naturais, como os estômatos e hidatódios. Alguns nematoides tornam-se patogênicos pela inserção do estilete através da célula das plantas (CHISHOLM et al., 2006; JONES; DANGL, 2006). Apesar das diferentes estratégias de ataque, todos os patógenos secretam proteínas efectoras para dentro das células do respectivo hospedeiro com a função de suprimir as respostas de defesa da planta e permitir o estabelecimento das relações parasíticas (ABE, 2012; DANGL et al., 2013). Proteínas efectoras são moléculas patogênicas que podem modificar a estrutura e funções celulares do hospedeiro, facilitando a infecção e colonização, bem como a supressão das respostas de defesa (DALIO et al., 2017; GÖHRE; ROBATZEK, 2008). Apesar de serem amplamente associados com a patogenicidade, alguns efetores ao serem detectados pelo hospedeiro, ativam as respostas de defesa. Analisando candidatos a efetores de *Melampsora larici-populina* de forma transiente em folhas de

*Arabidopsis thaliana*, Germain et al. (2018) observaram que os candidatos MLP37347 e MLP123227 promoveram a diminuição no crescimento de *P. syringae* inoculada nas folhas, possivelmente pelo reconhecimento dos receptores imunológicos da planta e que resultou na ativação da resposta de defesa. Além disso, em estudo similar com *Nicotiana benthamiana* e 20 candidatos a efetores de *M. larici-populina*, Petre et al. (2015) observaram a especificidade entre o candidato MLP124017 e a proteína co-repressora TPL/TPR, oriunda da planta hospedeira, e que está intimamente ligada ao desencadeamento da resposta imune vegetal. Também foram observadas respostas satisfatórias quanto à indução da defesa em plantas de pimentão testadas com 17 candidatos a efetores de *P. pachyrhizi*. A maioria dos candidatos observados induziu HR e também não suprimiu PTI, indicando que existe uma especificidade no reconhecimento dessas proteínas em plantas de pimentão (QI et al., 2018). Sendo assim, é correto definirmos efetores como “moléculas liberadas ou associadas a um organismo que modificam a fisiologia de outro organismo” (DALIO et al., 2014).

Os efetores são divididos em efetores apoplásticos e citoplasmáticos, sendo classificados de acordo com o local em que são secretados. Efetores apoplásticos ou extracelulares são secretados no espaço extracelular do hospedeiro, enquanto que, os efetores citoplasmáticos ou intracelulares são translocados para dentro da célula do hospedeiro acumulando-se no citoplasma ou em compartimentos da célula (DALIO et al., 2014).

Cada classe de fitopatógenos possui um mecanismo específico para exportar efetores do citoplasma do patógeno para a célula do hospedeiro, sendo este transporte uma etapa fundamental para a patogênese (ZARAMELA, 2011). Bactérias possuem diversos sistemas de secreção de efetores especializados (TSENG et al., 2009), sendo o Tipo III, IV e VI os sistemas mais comuns (DALIO, et al., 2017; MA et al., 2014; MCCANN; GUTTMAN, 2008). O Tipo III (TTSS) é o mais estudado, visto que seus efetores secretados podem modificar diversas funções moleculares, mesmo sem a presença da célula do hospedeiro (DESLANDES; RIVAS, 2012). Fungos e oomicetos secretam proteínas através do sistema de secreção Tipo II, que envolve o retículo endoplasmático e a exocitose de vesículas secretoras derivadas do complexo de Golgi (GRIFFITHS; SIMONS, 1986). A maioria dos efetores fúngicos e oriundos de oomicetos apresentam características conservadas, como a presença de um peptídeo sinal na região N-terminal, tamanho reduzido (<250 aminoácidos), natureza rica em cisteína (>3%), ausência de semelhança com proteínas que possuem funções já conhecidas e descritas nos bancos de dados (DE CARVALHO et al., 2016; ELLIS et al., 2009; LINK et al., 2014).

Conhecer as características dos efetores e utilizar ferramentas para predizê-los é um passo fundamental na detecção de novos candidatos a efetores, com o objetivo de obter

maior entendimento sobre as interações patógeno-hospedeiro para, posteriormente, alcançar êxito na obtenção de novas e sustentáveis formas de controle das doenças de plantas.

A identificação de genes que codificam proteínas efetoras é conduzida inicialmente por análises *in silico*, que são as análises feitas com o uso da bioinformática em bancos de dados de genomas, transcriptomas e proteomas de microrganismos. A partir desses dados, os candidatos são escolhidos de acordo com alguns critérios específicos para a seleção, tais como a presença de peptídeo sinal indicando sua secreção, presença de domínios conservados, pontes de dissulfeto, entre outros (DE CARVALHO et al., 2016). Após a identificação dos genes, é necessário validar estes candidatos através de ensaios funcionais. Em ferrugens, onde estratégias de transformação gênica são limitantes, tais ensaios baseiam-se por exemplo em testar a capacidade de estes suprimirem ou induzirem as respostas de defesa do hospedeiro, tanto PTI como ETI em sistemas modelo, através da expressão transiente e a avaliação de atividades semelhantes a efetores (KAMOUN, 2006).

A expressão transiente em plantas pode ser analisada de diversas formas, sendo que a translocação de proteínas efetoras através do sistema EDV é amplamente utilizada (*Effector Detector Vector*) (ABE, 2012). Esse sistema baseia-se no Sistema de Secreção Tipo III (SST3) de bactérias e que são codificados por genes *hrp* (*hypersensitive response and pathogenicity*) (SOHN et al., 2007). Esse sistema secreta proteínas efetoras através da membrana bacteriana diretamente para o citoplasma da célula vegetal, desencadeando respostas de hipersensibilidade em plantas não hospedeiras, e causando patogenicidade nas plantas hospedeiras. A reação HR em plantas não hospedeiras acontece devido a reação de genes *R* do hospedeiro interagindo com o gene *hrp* correspondente (AGRIOS, 2005).

### 2.3.2 Proteínas efetoras de fungos causadores de ferrugem

Candidatos a efetores são amplamente estudados em secretomas de fungos causadores de ferrugem, como *Melampsora* spp., *Puccinia* spp. e *Uromyces* spp. (SPERSCHNEIDER et al., 2017). O efector RTP1 (*Rust Transferred Protein*) de *U. fabae*, causador da ferrugem no feijão e de *U. striatus*, agente causal da ferrugem em alfafa, foi caracterizado e descrito como fator importante na manutenção da interação biotrófica entre patógeno e o hospedeiro, através da sua liberação do haustório diretamente para o meio intracelular da célula hospedeira durante a infecção (KEMEN et al., 2005; PRETSCHKE et al., 2013). Assim como os efetores AvrL567, AvrM, AvrP123 e AvrP4 secretados por *M. lini*, que quando expressos transientemente demonstraram funções efetoras de patogenicidade e também

a ativação das respostas de resistência, como HR (ELLIS et al., 2007). Ademais, o efector PEC6 secretado por *P. striiformis* f. sp. *tritici*, demonstrou-se eficiente na supressão das respostas de defesa em plantas de trigo (hospedeira) e em plantas de *Nicotiana benthamiana* (não-hospedeira), além de ser classificado como sendo rico em cisteína (LIU et al., 2016).

Diversos estudos descrevem a predição *in silico* de candidatos de *P. pachyrhizi*, que são posteriormente validados funcionalmente, por exemplo, através da localização subcelular em plantas modelo, pelos níveis de expressão gênica na presença do hospedeiro, bem como supressão das respostas de defesa em diferentes não hospedeiros, como *N. benthamiana*, *N. tabacum*, *Arabidopsis thaliana*, *Capsicum annuum* e *Solanum lycopersicum* (KUNJETI et al., 2016; QI et al., 2016; QI et al., 2018; RINCÃO et al., 2018).

Avaliando candidatos a efetores de *P. pachyrhizi* em *N. benthamiana*, Qi et al. (2018) observaram que nenhum dos candidatos estudados induziu HR nos pontos em que foram infiltrados, sugerindo que nenhum deles pode ser reconhecido por *N. benthamiana*. Resultado semelhante também foi observado em plantas de tomate, sendo que apenas um candidato induziu HR nas folhas infiltradas. Entre os 82 candidatos, 17 suprimiram as respostas de defesa em *N. benthamiana* e *A. thaliana*.

Além de suprimir PTI e ETI, alguns candidatos também podem desempenhar papéis fundamentais na virulência do seu hospedeiro. Os candidatos a efetores de *P. pachyrhizi* CSEP-07 e CSEP-09 foram capazes de aumentar a virulência de *Pseudomonas infestans* em folhas de *N. benthamiana* e em estudos de expressão gênica demonstraram padrões de virulência, através da supressão das respostas de defesa do hospedeiro (KUNJETI et al., 2016).

Apenas um efector de *P. pachyrhizi* foi caracterizado funcionalmente de forma ampla, demonstrando sua capacidade de supressão das respostas de defesa, bem como de seus interactores na soja (QI et al., 2016). Interessantemente, foi demonstrado através de ensaios de co-imunoprecipitação (Co-IP) que o efector *PpEC23* interage com o fator de transcrição GmSPL121, que é membro da família SPL de fatores de transcrição da soja e que desempenham funções no crescimento e desenvolvimento da planta (WANG et al., 2014). No trabalho, quando o gene foi silenciado transientemente via VIGS (*Virus Induced Gene Silencing*), as plantas de soja demonstraram imunidade constitutivamente ativa, evidenciando que este fator é um regulador negativo do sistema de defesa da planta, ou seja, não permite a ativação da imunidade pela planta (QI et al., 2016).

Em estudo de De Carvalho et al. (2016) foram preditas 851 proteínas secretadas pelo fungo *P. pachyrhizi* que foram distribuídas em famílias de efetores em base no estudo do transcriptoma do fungo expresso durante a interação com a soja e análises

comparativas com outras espécies. Dentre estas, a família 1, contendo 22 membros oriundos de *P. pachyrhizi*, foi a maior família de candidatos a efetores de *P. pachyrhizi* identificada. Análises de BLASTP mostraram similaridade entre os 35 membros desta família com fungos fitopatogênicos da ordem Pucciniales, e do mesmo filo de *P. pachyrhizi*. Além disso, um membro demonstrou similaridade com *Magnaporthe grisea* do filo *Ascomycota*. Entretanto, esta família não mostra igualdade com nenhum outro grupo de ferrugem já citado por e Kunjeti et al. (2016), Link et al. (2014) e Saunders et al. (2012) (DE CARVALHO et al., 2016).

Todos os membros da família 1 descritos por De Carvalho et al. (2016) possuem o domínio Egh16-like (DUF3129/PF11327), característico de proteínas amplamente codificadas nos genomas de fungos patogênicos dos filios *Ascomycota* e *Basidiomycota* (HUANG et al., 2019).

## 2.4 Domínios funcionais

Domínios funcionais são responsáveis por uma função ou interação específica em uma proteína. Geralmente, são identificados como regiões altamente conservadas da sequência proteica e são responsáveis por uma função ou interação específica, contribuindo para o papel geral de uma proteína. Portanto, é através dos domínios funcionais que as proteínas adquirem novas funções, resultando na sua evolução. Entender como funcionam e a função das proteínas é um dos principais objetivos da bioinformática (ROJANO et al., 2022; WANG et al., 2021). Atualmente, é possível a detecção com alta precisão de domínios de proteínas através de diversos bancos de dados, como Pfam (FINN et al., 2008) e INTERPRO (MULDER et al., 2007), que usam métodos baseados nas sequências ou estruturas proteicas que pertencem a famílias de domínio específicas (BASU et al., 2009).

### 2.4.1 Domínio Egh16

O domínio Egh16-like é característico de *Erysiphe graminis* f. sp. *hordei* mas também foi observado em proteínas efetoras de *Melampsora larici-populina* e *Puccinia graminis* f. sp. *tritici*, indicando ser conservado entre algumas espécies de ferrugem e mais especificamente em sequências secretadas por esses fungos, atuando principalmente durante a fase inicial de infecção (DUPLESIS et al., 2011; SAUNDERS et al., 2012; XUE et al., 2002).

Apesar disso, a função que desempenha em alguns patógenos ainda não é bem compreendida. Em estudo com proteínas de *M. grisea*, a proteína GAS1 que contém o domínio

Egh16-like, atuou na penetração do apressório e na formação de lesões em plantas de arroz (XUE et al., 2002). Além disso, o domínio Egh16-like foi observado em transcritos altamente expressos nos fungos nematófagos *Arthrobotrys dactyloides*, *A. oligospora* e *Monacrosporium cionopagum* (ANDERSSON et al., 2014). No fungo entomopatogênico *Metarhizium robertsii*, as proteínas Egh16-like atuam como TLDD (Targeting Lipid Droplets for Degradation), gerando turgor do apressório para romper as cutículas, reforçando o papel dessas proteínas no processo infeccioso, na etapa de penetração através da cutícula do hospedeiro (HUANG et al., 2019). Sendo assim, proteínas desta classe constituem potenciais determinantes na patogênese de *P. pachyrhizi* em plantas de soja, sendo este um fungo que infecta diretamente as plantas de soja pela epiderme.

### **3. ARTIGO A: CARACTERIZAÇÃO DE EFETORES DA FAMÍLIA EGH16-LIKE DE *Phakopsora pachyrhizi***

#### 3.1 RESUMO

Atualmente, o fungo biotrófico *Phakopsora pachyrhizi* é uma das ameaças mais importantes na agricultura brasileira por se tratar do agente causal da Ferrugem-Asiática da soja, a principal doença da cultura. Para seu sucesso, o patógeno modula o metabolismo do hospedeiro através da secreção de pequenas proteínas na célula vegetal, que são chamadas de efetores. Através da liberação dos efetores e o reconhecimento pelo hospedeiro, ocorre a ativação do sistema imune da planta, chamado de ETI (Imunidade Desencadeada por Efetores). Os efetores são classificados de acordo com seus domínios conservados, que são responsáveis por uma função ou interação específica, contribuindo para o papel geral de uma proteína. Uma extensa família de efetores contendo o domínio Egh16-like foi previamente descrita em *P. pachyrhizi* com base em dados de RNA-Seq. Pouco sabe-se sobre este domínio, mas estudos indicam que está relacionado, principalmente, com as etapas iniciais da infecção dos patógenos na célula do seu hospedeiro. Neste estudo, os membros da família de efetores contendo o domínio Egh16-like (DUF3129/PF11327) foram identificados com base no genoma de referência MT2006 do fungo, com o objetivo de caracterizar estruturalmente membros geneticamente diversos desta família, além de validar sua função efetora via caracterização do seu perfil de expressão durante a infecção. A partir de um conjunto de sequências expressas (dados de transcriptoma) inicialmente preditas por De Carvalho et al. (2016), seis de 22 membros preditos foram identificados no genoma recentemente anotado do fungo. Novos membros foram identificados via Blast contra o genoma do fungo, sendo possível identificar um total de 22 modelos gênicos preditos contendo este domínio. Esses membros variam de 195 a 500 aminoácidos, além de apresentarem elevado conteúdo de cisteína e a presença de peptídeo sinal limitado aos 24 primeiros aminoácidos, sendo que três destes modelos não apresentaram peptídeo sinal, mas foram preditos como secretados por vias não convencionais. Análises filogenéticas com base na sequência de proteínas, agrupou tais membros em oito clusters e ao menos um membro de cada ramo foi selecionado para avaliar o perfil de expressão. Assim, foram analisados os perfis de expressão gênica destes candidatos por RT-qPCR em diferentes tempos de infecção em uma cultivar suscetível, através de dados oriundos dos modelos gênicos correspondentes aos candidatos preditos nos três genomas de referência do fungo. Alguns candidatos tiveram picos de expressão nas fases iniciais de infecção, que coincidem com a formação dos haustório,

enquanto outros foram expressos apenas nos momentos tardios da infecção. Análises filogenéticas entre os candidatos selecionados e proteínas oriundas de diferentes espécies fúngicas indicaram que as proteínas selecionadas possuem um ancestral comum com diferentes espécies, mas todas evoluíram singularmente de acordo com o seu hospedeiro. Quatro clados da árvore filogenética agruparam apenas membros de ferrugens, sendo que em cada um deles havia um membro oriundo de *P. pachyrhizi* e ao menos um clado agrupou membros exclusivos de *P. pachyrhizi*, indicando que, possivelmente, este é um patógeno restritivo à diversificação, evidenciando sua coevolução com seu hospedeiro em comparação a outras espécies de Pucciniales. Análises da região gênica de dois efetores com diversas cópias no genoma revelaram a ocorrência de regiões sintênicas entre as regiões contendo os genes, indicando uma certa conservação. Portanto, através dos dados observados, acredita-se que esta família de efetores tenha um papel importante na virulência do fungo durante a interação com a soja.

**PALAVRAS-CHAVE:** Biotecnologia. Ferrugem-asiática da soja. *Glycine max*. Interação planta-patógeno. Transcriptoma.

## ARTICLE A: CHARACTERIZATION OF EFFECTORS OF THE EGH16 FAMILY OF *Phakopsora pachyrhizi*

### 3.2 ABSTRACT

Currently, the biotrophic fungus *Phakopsora pachyrhizi* is one of the most important threats in Brazilian agriculture because it is the causal agent of Asian soybean rust, the main disease of the crop. For its success, the pathogen modulates the host's metabolism through the secretion of small proteins in the plant cell, which are called effectors. Through the release of the effectors and the recognition by the host, the activation of the plant's immune system occurs, called ETI (Effector Triggered Immunity). Effectors are classified according to their conserved domains, which are responsible for a specific function or interaction, contributing to the overall role of a protein. An extensive family of effectors containing the Egh16-like domain was previously described in *P. pachyrhizi* based on RNA-Seq data. Little is known about this domain, but studies indicate that it is mainly related to the initial stages of infection of pathogens in their host cell. In this study, members of the effector family containing the Egh16-like domain (DUF3129/PF11327) were identified based on the MT2006 reference genome of the fungus, with the aim of structurally characterizing genetically diverse members of this family, in addition to validating their effector function via characterization of its expression profile during infection. From a set of expressed sequences (transcriptome data) initially predicted by De Carvalho et al. (2016), six of 22 predicted members were identified in the recently annotated genome of the fungus. New members were identified via Blast against the fungus genome, making it possible to identify a total of 22 predicted gene models containing this domain. These members range from 195 to 500 amino acids, in addition to having high cysteine content and the presence of a signal peptide limited to the first 24 amino acids, and three of these models did not present a signal peptide, but were predicted to be secreted by unconventional pathways. Phylogenetic analyzes based on protein sequence grouped such members into eight clusters and at least one member from each branch was selected to assess the expression profile. Thus, the gene expression profiles of these candidates were analyzed by RT-qPCR at different times of infection in a susceptible cultivar, using data from the gene models corresponding to the predicted candidates in the three reference genomes of the fungus. Some candidates had peaks of expression in the early stages of infection, which coincide with the formation of haustoria, while others were expressed only at late stages of infection. Phylogenetic analyzes between the selected candidates and proteins from different fungal species indicated that the selected

proteins have a common ancestor with different species, but all evolved uniquely according to their host. Four clades of the phylogenetic tree grouped only members of rusts, and in each of them there was a member from *P. pachyrhizi* and at least one clade grouped exclusive members from *P. pachyrhizi*, indicating that, possibly, this is a diversification-restrictive pathogen, evidencing its coevolution with its host compared to other Pucciniales species. Analysis of the gene region of two effectors with several copies in the genome revealed the occurrence of syntenic regions between the regions containing the genes, indicating a certain conservation. Therefore, through the observed data, it is believed that this family of effectors plays an important role in the virulence of the fungus during its interaction with soybean.

KEYWORDS: Asian soybean rust. Biotechnology. *Glycine max*. Plant-pathogen interaction. Transcriptome.

### 3.3 INTRODUÇÃO

Entender como os fungos modulam o sistema de defesa das plantas hospedeiras para sobreviverem e dessa forma levarem ao desenvolvimento das doenças é parte fundamental da Fitopatologia, pois é através destes conceitos que é possível mitigar e controlar doenças fitopatogênicas que acarretam em prejuízos na produção agrícola. Dentre as doenças de grande importância na agricultura nacional, a Ferrugem Asiática da Soja (FAS) destaca-se por ser causada por um patógeno altamente adaptável, mesmo possuindo reprodução assexuada, e que desafia os métodos de controle convencionais.

O agente causal da FAS é o fungo biotrófico *Phakopsora pachyrhizi*, que depende exclusivamente do seu hospedeiro para sobreviver, possuindo mais de 90 espécies de leguminosas como hospedeiras, sendo a soja a sua principal hospedeira (RYTTER, et al., 1984; SLAMINKO et al., 2008). Trata-se de um fungo que possui um dos maiores genomas entre os fitopatógenos, com cerca de 1,27 Gbp, sendo que grande parte enriquecido com elementos transponíveis (TEs). Adicionalmente, o genoma deste fungo apresenta um secretoma potencial com cerca de duas mil proteínas preditas (dados do Consórcio para o Genoma de *P. pachyrhizi*), dentre estes candidatos a efetores. Tais proteínas desempenham funções tanto na virulência, sendo importantes por modularem o sistema imune inato do hospedeiro e, dessa forma, favorecendo o processo infeccioso. Estas proteínas são alvos de pesquisas referentes às suas funções e como atuam no favorecimento do patógeno e consequente estabelecimento das doenças. Sua identificação e caracterização fornecem informações sobre a patogênese, gerando *insights* para novas formas de controle de doenças (LINK et al., 2014; SPERSCHNEIDER et al., 2015).

A predição de proteínas candidatas a efetores de *P. pachyrhizi* através do seu transcriptoma já foram descritas (DE CARVALHO et al., 2017; ELMORE et al., 2020; KUNJETI et al., 2016; LINK et al., 2014; RINCÃO et al., 2018). Entre estes, alguns demonstraram a supressão do sistema imune em plantas por estes candidatos. Um candidato a efector, *PpEC23*, suprimiu as respostas de defesa em plantas de *Nicotiana benthamiana*, *N. tabacum* e de soja (QI et al., 2016). Em outro estudo, os candidatos CSEP-07 e CSEP-09 tornaram plantas de *N. benthamiana* mais suscetíveis a *Phytophthora infestans* (KUNJETI et al., 2016). Além disso, 17 *PpECs* foram capazes de suprimir a imunidade em plantas de *N. benthamiana* e *Arabidopsis thaliana*. Entretanto, trabalhos envolvendo a análise ampla de potenciais membros de uma família gênica, de sua variabilidade genética, integrando dados de transcriptômica e utilizando dados disponíveis a partir de um genoma de referência, são

inéditos.

Neste trabalho, membros da família de proteínas candidatas a efetores de *P. pachyrhizi* contendo o domínio Egh16-like (DUF3129/PF11327) foram identificados com base no genoma de referência do isolado MT2006, bem como seus correspondentes nos genomas dos isolados K8108 e UFV02. Tais membros foram caracterizados molecularmente além de terem seu perfil de expressão avaliado durante a infecção. Análises filogenéticas foram ainda conduzidas para demonstrar as relações filogenéticas desses membros com outros fungos causadores de ferrugem ou distantes geneticamente. Finalmente, famílias de TE foram identificados associados a membros destas famílias.

### 3.4 MATERIAIS E MÉTODOS

#### 3.4.1 Análise e seleção dos candidatos a efetores

Para seleção dos genes candidatos a efetores de *P. pachyrhizi* utilizados neste estudo, foram utilizados como referência os efetores preditos por De Carvalho et al. (2016), que por meio de microdissecção a laser e do sequenciamento de RNA, predizeram o secretoma do fungo, obtendo uma lista de 851 proteínas candidatas secretadas, que foram agrupadas em famílias por análises comparativas de similaridade a nível de proteína. Com base nas análises *in silico*, tais proteínas demonstraram características efetoras, como a presença de peptídeo sinal, tamanho reduzido ( $\leq 250$  aminoácidos) e resíduos ricos em cisteína ( $> 3\%$ ) (De Carvalho et al., 2016).

Os 22 membros da Família 1 descrita por De Carvalho et al. (2016) apresentando o domínio Egh16-like (DUF3129/PF11327) foram utilizados para busca dos genes correspondentes no genoma de referência do fungo “*Phakopsora pachyrhizi* MT2006 v1.0”, depositado no banco de dados do Joint Genome Institute (JGI) (<https://mycocosm.jgi.doe.gov/Phapa1/Phapa1.info.html>) (GRIGORIEV et al., 2011).

Para seleção de cada gene correspondente, bem como novos possíveis genes preditos correlacionados no genoma de referência, análises foram conduzidas com cada um dos 22 membros descritos. Inicialmente, cada um dos membros passou por uma verificação de predição ORF (Open Reading Frames) utilizando o software ORF-Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>). As sequências dos candidatos foram então analisadas pelo software SignalP (<http://www.cbs.dtu.dk/services/SignalP-5.0/>) para confirmar a presença de peptídeo sinal. Posteriormente, foi realizado o alinhamento contra o genoma de

referência (BLASTX, e-value  $1e^{-5}$ ) e sendo os modelos gênicos considerados como correspondentes aos transcritos quando apresentassem identidade  $\geq 70\%$  e cobertura  $\geq 90\%$  com o genoma. Para anotação manual dos candidatos a efetores, além dos critérios de identidade e cobertura, também foram considerados os tamanhos das proteínas dos modelos gênicos e dos transcritos, sendo descartados aqueles que possuíam tamanho muito reduzido, visto que alguns transcritos com sequências muito pequenas correspondiam a um mesmo gene predito no genoma, o que pode ser indício de uma montagem errada. Devido a isto, foram escolhidos os transcritos que melhor representassem o modelo gênico. Considerou-se também a identidade e cobertura dos transcritos em relação aos modelos gênicos dos outros dois genomas de referência, para inferir qual sequência (modelo gênico ou transcrito) poderia estar montada da forma correta (Tabela 1).

Posteriormente, foi verificada a estrutura de cada gene com base na predição gênica conduzida pelo JGI. O alinhamento entre as sequências dos transcritos dos candidatos a efetores e a sequência nucleotídica dos genes correspondentes foi realizada através do software ClustalW ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/)) visando observar a semelhança entre as sequências e também identificar regiões de éxons e de íntrons. Além disso, foi realizada a busca por domínios conservados utilizando o domínio PFAM11327, que corresponde ao domínio proteico Egh16-like, característico dos membros dessa família. Todas as proteínas que apresentaram este domínio, os transcritos correspondentes e os transcritos que não tiveram modelo gênico foram agrupados por um alinhamento múltiplo e assim foi construída a árvore filogenética utilizando o programa IQ-TREE (MINH et al., 2020).

Os genes identificados no isolado MT2006 como correspondentes aos transcritos da família contendo o domínio proteico Egh16-like também foram analisados nos outros dois genomas de referência de *P. pachyrhizi* (“*Phakopsora pachyrhizi* K8108 v2.0” e o “*Phakopsora pachyrhizi* UFV02 v2.1”). O critério para identificação dos genes homólogos foi similar, selecionando aqueles com elevada identidade e cobertura. Os dois genomas adicionais disponíveis no JGI são: o genoma UFV02, cujo o isolado é oriundo da região central do Brasil, e foi sequenciado pela tecnologia PromethION; e o isolado K8108, também coletado de folhas de soja no ano de 2015, mas proveniente da região de Colonia, no Uruguai, sendo sequenciado pela tecnologia Pacbio. Os genes identificados nos três genomas foram então utilizados para estudos comparativos e para inferir o nível de expressão em cada isolado.

### 3.4.2 Comparação filogenética entre espécies de fungos fitopatogênicos e *P. pachyrhizi*

O agrupamento comparativo entre as proteínas preditas no genoma de *P. pachyrhizi* e de outras espécies de fungos foi conduzida pelo Consórcio do Genoma de *P. pachyrhizi* visando determinar as relações evolutivas das famílias de genes. A clusterização foi conduzida utilizando o programa ORTHO-MCL (LI et al., 2003) e as sequências proteicas preditas nos genomas de 15 espécies de fungos fitopatogênicos, entre eles os pertencentes à ordem Pucciniales, que são os causadores de ferrugens, como “*Puccinia triticina* 1-1 BBBD Race 1”, “*Puccinia graminis* f. sp. *tritici* v2.0”, “*Melampsora allii-populina* 12AY07 v1.0”, “*Melampsora larici-populina* v2.0” e “*Cronartium quercuum* f. sp. *fusiforme* G11 v1.0”, além dos três genomas de *P. pachyrhizi*, já citados anteriormente. “*Sporobolomyces roseus* v1.0”, “*Microbotryum lychnidis-dioicae* p1A1 Lamole”, “*Mixia osmundae* IAM 14324 v1.0”, “*Ustilago maydis* 521 v2.0”, “*Sporisorium reilianum* SRZ2”, “*Laccaria bicolor* v2.0” e “*Heterobasidion annosum* v2.0”, que não são causadores de ferrugem mas também fazem parte do filo Basidiomycota. Por último, “*Magnaporthe oryzae* 70-15 v3.0” e “*Blumeria graminis* f. sp. *hordei* DH14”, ambos pertencentes ao filo Ascomycota.

Os oito genes previamente identificados como contendo o domínio Egh16-like foram utilizados para identificar novos genes que possam ter se agrupado nos clusters. Após a identificação dos clusters de interesse, as sequências preditas dos membros de cada um deles foram agrupadas, alinhadas pelo software MUSCLE (EDGAR, 2007) e então utilizadas para a construção da árvore filogenética, através do método *Neighbor-Joining* (*Bootstrap* 1000x), com o software MEGA11 (TAMURA; STECHER; KUMAR; 2021).

### 3.4.3. Material biológico

Para este trabalho foi utilizado o isolado monospórico de *P. pachyrhizi* PF2016.1M (CMES1962) pertencente à Coleção de Microrganismos da Embrapa Soja. O isolado foi coletado em Passo Fundo, RS na safra 2016 e apresenta perfil de virulência aos genes Rpp1 (PI200492), Rpp2 (PI230970), Rpp3 (PI462312) e de avirulência aos genes Rpp4 (PI 459025) e Rpp5 (PI 200487) (DARBEN et al., 2019). Para manutenção de esporos foram realizados *rounds* de multiplicação em folhas destacadas de soja do genótipo suscetível COODETEC 219 RR (CD219 RR), seguindo a metodologia de Castanho (2017). As plantas foram cultivadas em casa de vegetação de fluxo negativo para evitar a presença de qualquer tipo de esporo fúngico. Quando as plantas atingiam o estágio V3 - segunda folha trifoliolada completamente desenvolvida (RITCHIE et al., 1994), o segundo e terceiro trifólios eram coletados. No laboratório, as folhas foram lavadas em água corrente e para hidratá-las foram

mantidas com a face abaxial em contato com água destilada por uma hora. Posteriormente, as folhas foram transferidas para placas de petri contendo meio ágar-água 1%. Para inoculação, foi preparada uma suspensão de esporos do isolado, contendo 0,5 mL de água destilada + Tween 20, gentilmente espalhada pela superfície foliar com auxílio de um pincel. As placas foram acondicionadas em câmara de crescimento com 14h de luz e 10h de escuro e temperatura de 23°C ( $\pm$  2°C). Após 15 dias de incubação, foram realizadas coletas dos esporos para novas multiplicações e para armazenamento em freezer a -80°C em tubos de 2 mL. As folhas eram mantidas por, aproximadamente, 30 dias na câmara de crescimento, sendo possível até três coletas de esporos a partir de uma única inoculação.

#### 3.4.4 Perfil de expressão de genes candidatos a efetores de *P. pachyrhizi* durante o ciclo infeccioso

Com base nas análises estruturais e de diversidade entre os genes preditos no genoma do isolado MT2006, oito candidatos foram selecionados para caracterização do perfil de expressão durante o início do ciclo infeccioso, utilizando o genótipo de soja suscetível (BRS184). Para análise, foi utilizado o isolado PF16.1M (CMES1962), descrito anteriormente.

Para a análise durante o ciclo infeccioso, plantas de soja da cultivar BRS184, suscetível a *P. pachyrhizi*, foram cultivadas em vasos de plástico de 0,5 L preenchidos com solo autoclavado e peneirado, e que foram mantidos em câmara climática de crescimento (Fitotron<sup>®</sup>) até as plantas atingirem o estágio V3. Assim, com o auxílio de um borrifador foi inoculada uma suspensão de esporos diretamente no primeiro e segundo trifólio de cada planta. A suspensão foi preparada pela adição de 15 mg de urediniósporos do isolado PF16.1M em 30 mL de água destilada com adição de Tween 20 (0,05%), obtendo uma concentração final de  $1 \times 10^4$  esporos.mL<sup>-1</sup>. As folhas infectadas foram coletadas nos diferentes tempos 0, 6, 12, 24 e 48 horas pós inoculação (hpi) para representar o início da infecção até a colonização dos tecidos do hospedeiro pelo fungo. Após a coleta, as folhas foram acondicionadas em sacos de papel alumínio, imediatamente congeladas em nitrogênio líquido e, posteriormente armazenadas em ultra freezer a -80°C. O experimento foi instalado no delineamento inteiramente casualizado constituído por três repetições biológicas, sendo que cada uma das repetições era composta por três plantas.

##### 3.4.4.1 Extração de RNA total e síntese de cDNA

As amostras de folhas de soja inoculadas com a suspensão de esporos foram

maceradas em cadinhos de porcelana e a extração do RNA total foi realizada através do uso do reagente Trizol™ (1 mL / 100 mg tecido), de acordo com o protocolo do fabricante (Invitrogen™). O RNA foi ressuspensionado em 35 µL de água UltraPure™ DNase/RNase-Free (Invitrogen™) e armazenado a -80°C. As concentrações de RNA foram quantificadas em Espectrofotômetro NanoDrop™ 1000 (Thermo Scientific) e a integridade das amostras analisada em gel de agarose 1% corado com brometo de etídio. Para eliminar todo e qualquer DNA genômico presente nas amostras, uma alíquota de 5 µg de RNA foi tratada com a enzima Desoxirribonuclease (DNase I - Invitrogen™) e visualizada em gel de agarose em 1% para confirmar a eficiência do tratamento. Posteriormente, o cDNA foi sintetizado utilizando o kit SuperScript III First-Strand (Invitrogen™), seguindo as instruções do fabricante. A amplificação do cDNA foi realizada para confirmação da síntese das amostras e o cDNA foi armazenado a -20°C.

#### 3.4.4.2 Análise RT-qPCR

Os primers para as reações de RT-qPCR foram desenhados pelo software Primer3Plus com base nas sequências do transcriptoma de *P. pachyrhizi* (Apêndice A) e dos dois modelos gênicos selecionados. Para otimizar a concentração de cDNA e conhecer a eficiência de cada um dos primers desenhados, foram feitas curvas padrões com cinco pontos de diluição seriada de cDNA. As eficiências de amplificação de cada um deles foi calculada de acordo com a equação  $E = [2^{(-1/\text{slope})}] - 1$  descrita por Pfaffl (2001). As análises foram conduzidas no equipamento 7900HT Fast Real-Time PCR System (Applied Biosystems™), utilizando o kit Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen). As condições para a qPCR foram: 50°C por 2 minutos, 95°C por 10 minutos (desnaturação), seguido de 40 ciclos de 95°C por 15 segundos (anelamento) e a 60°C por 1 minuto (extensão). Posteriormente, também foi realizada a etapa de dissociação, que consistiu na elevação da temperatura para 95°C por 15 segundos, seguida da diminuição até 60°C por 15 segundos e, por fim, a elevação até 95°C por 15 segundos. Essa etapa foi realizada para a construção da curva de dissociação com a intenção de verificar a especificidade dos produtos da amplificação. As amostras foram analisadas em duas corridas independentes e cada reação de qPCR foi realizada em triplicata técnica para cada uma das três repetições biológicas, a fim de minimizar erros experimentais.

Os níveis de expressão dos genes alvos foram determinados pela quantificação relativa, através do software REST® (PFAFFL; HORGAN; DEMPFLER, 2002).

Os dados foram submetidos a transformação por log e usados para gerar o *heat map* com o software Morpheus (<https://software.broadinstitute.org/morpheus/>). O gene endógeno *CytB* foi utilizado como gene normalizador para o fungo *P. pachyrhizi* (LINK et al., 2015), enquanto o tempo 0 hpi foi empregado como calibrador. Os valores de Ct dos esporos germinados nos tempos 6, 12, 24, 48 hpi foram comparados com os valores de Ct do tempo 0 hpi em cada um dos genes alvos.

#### 3.4.4.3 Validação dos níveis de expressão gênica de RNA-Seq por RT-qPCR

Os dados de expressão gênica obtidos por RT-qPCR foram contrastados com dados de RNA-Seq gerados pela equipe do Consórcio internacional de pesquisa do genoma do fungo *P. pachyrhizi*, a fim de validarmos o perfil de expressão dos candidatos obtidos neste estudo em pelo menos outros três estudos independentes. Os dados de RNA-Seq contêm informações da expressão diferencial dos três isolados do fungo: MT200, UFV02 e o K8108 nas condições de interação com a soja e na ausência dela, sendo: PPUFV: esporo, 12, 24, 36, 72 e 168 hpi; MT2006: esporo, apressório, 192 e 288 hpi; K8108: esporo, apressório, 10, 24, 72 e 196 hpi. As análises estão disponíveis à equipe do Consórcio.

#### 3.4.5 Análise de sintenia

Devido ao elevado número de cópias, algumas próximas no mesmo scaffold, a região física contendo os modelos gênicos correspondentes aos candidatos a efetores de *P. pachyrhizi* de *de\_novo\_1784* e *de\_novo\_2238* foi analisada manualmente nos três genomas de referência visando analisar sua organização. Na região, foi avaliada a organização dos modelos gênicos utilizando a análise de sintenia do portal JGI, comparando essas regiões específicas ou próximas (100 Kbp) de cada genoma com a região correspondente nos outros dois genomas.

### 3.5 RESULTADOS E DISCUSSÃO

#### 3.5.1 Predição de candidatos a efetores de *P. pachyrhizi* contendo o domínio Egh16-like com base no genoma do isolado MT2006

Para identificação de genes preditos no genoma de *P. pachyrhizi* similares aos transcritos da Família 1 descrita por De Carvalho et al. (2016), todas as sequências expressas

foram utilizadas para buscas contra o genoma “*Phakopsora pachyrhizi* MT2006 v1.0”. Com base na análise, foram identificados um total de 22 modelos gênicos, sendo que para 11 sequências expressas foi possível identificar um modelo gênico correspondente com identidade de pelo menos 70% e cobertura de 90%. Estes transcritos estão em destaque negrito na Tabela 1.

As análises de BlastX utilizando as sequências dos transcritos retornaram 13 modelos gênicos correspondentes. Os nove modelos gênicos adicionais foram obtidos através de buscas pelo domínio Egh16-like (PFAM11327) utilizando o genoma de referência do isolado MT2006, totalizando os 22 modelos gênicos preditos no genoma do fungo. O tamanho das proteínas preditas no genoma variou de 195 a 500 aminoácidos. Os resultados de BlastX e as características de cada modelo gênico encontram-se no Apêndice B.

**Tabela 1** – Modelos gênicos correspondentes e códigos das sequências transcritas preditas, membros da família 1, utilizados para buscas contra o genoma “*Phakopsora pachyrhizi* MT2006 v1.0”. Transcritos em negrito referem-se aqueles com identidade de pelo menos 70% e cobertura de 90% em relação ao modelo gênico correspondente.

Modelo gênico com maior similaridade	Transcrito de_novo	Tamanho da proteína predita (aa)	Identidade (%)	Cobertura (%)
6413136/6427169	<b>251</b>	306	100,00	100,00
4635428/4578427	<b>1784</b>	202	100,00	100,00
4578227	<b>2595</b>	247	100,00	100,00
	<b>1303</b>	247	100,00	100,00
6898838	635	289	100,00	85,00
	8112	132	100,00	35,00
8120199/7639239/4594856/4594859	<b>2238</b>	203	100,00	100,00
	<b>1402</b>	203	100,00	100,00
	<b>5381</b>	102	100,00	91,18
	<b>5849</b>	178	88,00	100,00
6435004	789	269	52,00	95,91
	<b>2037</b>	281	74,00	92,53
	2721	218	84,00	72,94
4582009/4597124	555	430	100,00	45,35
	<b>6062</b>	160	94,00	99,38
	<b>7507</b>	121	99,00	95,87
	5333	97	98,00	89,69
	2227	291	100,00	67,01
	5672	58	100,00	75,86
	3838	66	100,00	84,85
6643	99	97,00	65,66	
1770	372	100,00	52,42	

Fonte: JGI MycoCosm, 2021.

Foi observado que dentre os 11 transcritos em que foi encontrada correspondência no genoma, quatro destes (em *itálico*) embora apresentassem elevada identidade e cobertura, correspondiam aos mesmos modelos gênicos, sendo os transcritos 6062 e 7507 correspondendo aos genes 4582009/4597124, e 5381 e 5849 aos modelos 8120199/7639239/4594856/4594859. Considerando o tamanho das proteínas preditas, este fato indica que tais transcritos possivelmente configuram erros de montagem e, nestes casos, o maior transcrito, garantido uma maior cobertura do modelo gênico com o transcrito, foi o validado. Além disso, muitos desses transcritos estão agrupados no mesmo ramo da árvore filogenética, ratificando que são semelhantes (Figura 3). Para os casos acima, os transcritos 555 e 2238, respectivamente, foram considerados os que melhor correspondiam aos modelos gênicos indicados. De modo similar, o transcrito 2037 não foi considerado para avaliação, sendo o seu modelo gênico o 6435004, que possui muita semelhança com o gene 4582009 e, como descrito anteriormente, foi selecionado o maior transcrito entre os transcritos semelhantes.

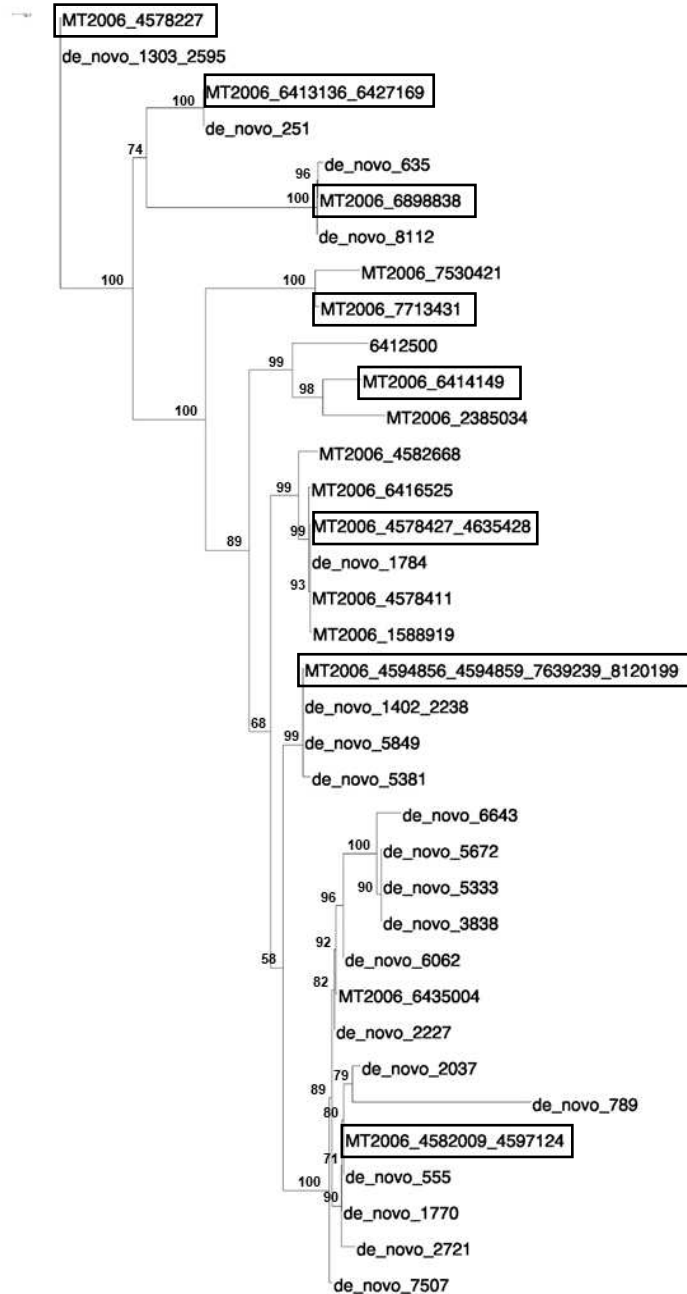
Dois transcritos, 555 e 635, embora não apresentassem valores de identidade e cobertura que atendessem os critérios estabelecidos, na análise fina do seu alinhamento contra os modelos gênicos observou-se que, na verdade, ocorreram erros na predição dos genes no genoma do fungo. Para o transcrito 555, o modelo gênico mais similar não apresentou códon de parada em nenhum dos três genomas, sendo então escolhida a sequência do transcrito a mais completa e, portanto, correta. De modo similar, o candidato 635, que possui em sua sequência um T a mais e que possivelmente altera a fase de leitura, resultando em uma menor cobertura. O alinhamento entre os transcritos e seus respectivos modelos gênicos podem ser observados nos Apêndices C e D.

Desse modo, para um total de oito transcritos de novo foram identificados corretamente um modelo gênico correspondente. Destaca-se que os transcritos 2238 e 1401 são 100% idênticos, assim como os transcritos 1303 e 2595.

### 3.5.2 Análise de diversidade genética entre os membros da família Egh16-like de *P. pachyrhizi*

Todas as sequências dos 22 modelos gênicos preditos no genoma do fungo, bem como as sequências dos transcritos (*de novo*) foram utilizadas para a construção da árvore de diversidade genética (Figura 3).

**Figura 3** – Árvore filogenética dos membros preditos que possuem o domínio Egh16-like no genoma de *Phakopsora pachyrhizi*. Árvore gerada pela análise *Neighbour-Joining* (Bootstrap 1000x).



\*Os retângulos referem-se aos candidatos selecionados nesse estudo.

Fonte: IQ-TREE, 2021.

Como esperado, a árvore revelou elevada similaridade entre as proteínas, sendo possível ainda assim observar oito clusters. Nos três primeiros clusters são observadas as proteínas secretadas preditas com base no estudo do transcriptoma e seus respectivos modelos gênicos, sendo estes, em sua maioria, idênticos entre si, como MT2006\_4578227 e as

sequências de\_novo\_2595 e 1303 que estão no primeiro cluster da árvore. Adiante, o de\_novo\_251 foi o mais similar ao MT2006\_6413136 e MT2006\_6427169, com 100% de identidade e cobertura. Já no terceiro cluster, agruparam-se um modelo gênico e duas sequências do transcriptoma, sendo a de\_novo\_635 a mais similar e com maior porcentagem de cobertura com MT2006\_6898838.

Os clusters 4 e 5 são compostos por modelos gênicos que possuem o domínio Egh16-like, mas não tiveram correspondência com as sequências transcritas, sendo eles o MT2006\_6414149 e MT2006\_7713431. Além disso, os dois apresentaram menor similaridade entre os outros membros e, em análises de alinhamento, ambas as sequências demonstraram menor similaridade com as sequências de\_novo, e por isso, foram selecionadas para o estudo.

O cluster 6 agrupou, principalmente, sequências dos modelos gênicos identificados no genoma, sendo que a maioria dos membros apresentou similaridade com a proteína secretada previamente predita de\_novo\_1784. Logo abaixo, encontram-se modelos gênicos pertencentes ao cluster 7, sendo MT2006\_4594856, MT2006\_4594859, MT2006\_7639239 e MT2006\_8120199 idênticos entre si e com o de\_novo\_2238 e 1402.

Por último, são agrupadas principalmente sequências secretadas previstas com base no estudo do transcriptoma e apenas três modelos gênicos no genoma do fungo, sendo que dois deles, MT2006\_4582009 e MT2006\_4597124 são idênticos entre si. Neste grupo, o de\_novo\_555 destaca-se por ter a maior proteína predita entre todos deste cluster, além de possuir o domínio completo em sua sequência.

Com base nos resultados, considerando alta cobertura e alta identidade, bem como a presença do domínio Egh16, oito candidatos descritos por De Carvalho et al. (2016) foram considerados validados com base neste estudo, utilizando o genoma de referência do isolado MT2006. Destes oito, 2238 e 1402 foram idênticos entre si, assim como 1303 e 2595. Desse modo, estes seis genes foram selecionados para os estudos *in vivo*, além de representarem diferentes clusters na árvore filogenética.

Assim, foram selecionados seis candidatos (de\_novo\_251, de\_novo\_2595, de\_novo\_555, de\_novo\_635, de\_novo\_2238, de\_novo\_1784 para estudos moleculares e *in vivo* do perfil de expressão. Adicionalmente, dois candidatos preditos no genoma, MT2006\_6414149 e MT2006\_7713431, que possuem menor similaridade com as sequências dos transcritos, estando em ramos distintos dos demais candidatos na árvore filogenética, também foram selecionados.

### 3.5.3 Caracterização molecular dos genes candidatos e efetores Egh16-like

As sequências dos candidatos selecionados também foram analisadas contra outros dois genomas de referência de *P. pachyrhizi* (UFV02 e K8108). Os genes no genoma de MT2006 e sua correspondência nos outros dois genomas de referência, bem como detalhes característicos de proteínas efetores, podem ser visualizados na Tabela 2.

**Tabela 2** – Candidatos a efetores de *Phakopsora pachyrhizi* selecionados para o estudo. Os códigos referem-se à identidade das proteínas em cada um dos três genomas de referência, disponíveis no banco de dados do Joint Genome Institute.

Candidatos	Código	Identidade <sup>1</sup> (%)	Nº cisteínas	Peptídeo Sinal	Tamanho proteína
<b>de_novo_251</b>	MT2006_6413136/6427169	100.0	6	1-20	306
	K8108_3093100	96.0	6	1-20	301
	UFV02_1582762	96.0	6	1-20	301
<b>de_novo_2595</b>	MT2006_4578227	100.0	6	1-24	247
	K8108_10880951	100.0	6	1-24	247
	UFV02_1579692	100.0	6	1-24	247
<b>de_novo_555</b>	MT2006_4582009 4597124	100.0	7	1-20	195
	K8108_3117622/3092552	100.0	7	1-20	195
	UFV02_1577018	100.0	7	1-20	195
<b>de_novo_635</b>	MT2006_6898838	100.0	7	Não	290
	K8108_3167062	100.0	9	1-24	313
	UFV02_3375195	93.0	9	1-24	331
<b>de_novo_1784</b>	MT2006_4635428/4578427 4578411	100.0	7	1-24	202
	K8108_3115284	100.0	7	1-24	202
	UFV02_3267052	100.0	7	1-24	202
<b>de_novo_2238</b>	MT2006_8120199/7639239/ 4594856/4594859	100.0	7	1-23	203
	K8108_3140583/11437327/10878620	100.0	7	1-23	203
	UFV02_3064936/3064159/3064092	100.0	7	1-23	203
<b>6414149</b>	MT2006_6414149	-	6	1-24	500
	K8108_11701857	-	6	Não	459
	UFV02_1573666	-	6	1-24	258
<b>7713431</b>	MT2006_7713431	-	8	Não	220
	K8108_3169489/9393185	-	8	Não	239
	UFV02_3030949/3093441	-	7	1-28	205

<sup>1</sup>Identidade obtida através do BlastN entre as sequências do transcriptoma de *P. pachyrhizi* e as sequências do genoma de referência.

Fonte: JGI MycoCosm, 2021.

Cada um dos transcritos selecionados neste estudo possui uma alta similaridade com as sequências preditas nos três genomas de referência. De modo similar, a estrutura dos genes e o tamanho da proteína predita foi coincidente.

Outra característica importante para caracterização de efetores é a presença de peptídeo sinal para secreção (PETRE et al., 2014; STERGIPOULOS; DE WIT, 2009). Nas proteínas analisadas, foi constatada a presença de peptídeo sinal ao menos em um genoma de referência e, em sua maioria, entre os aminoácidos 1 e 24. Em estudo de identificação de candidatos a efetores de *P. pachyrhizi*, Link et al. (2014) basearam-se na presença de peptídeo sinal, entre outras características, para identificar os melhores candidatos para sua pesquisa. Entretanto, a presença de peptídeo sinal pode não ser um fator limitante para a identificação de novos candidatos a efetores, visto que estudos demonstraram que algumas proteínas preditas secretadas podem não apresentar peptídeo sinal, e mesmo assim serem secretadas por vias não clássicas. Collins et al. (2013), em um estudo sobre proteínas de *Armillaria mellea*, verificou que apenas 50,2% das 293 proteínas secretadas identificadas possuíam peptídeo sinal. Além disso, 34% possuíam um sistema de secreção não convencional. Em outro estudo sobre duas proteínas efetoras secretadas, a PsIsc1 e VdIsc1 de *Phytophthora sojae* e *Verticillium dahliae*, que estão associadas à supressão da resposta imune vegetal inata, foram preditas para serem secretadas por um sistema não convencional por não possuírem peptídeo sinal (LIU et al., 2014).

Para os genes selecionados foi conduzida ainda uma análise estrutural dos genes preditos (Figura 4). Os genes apresentam os tamanhos entre 585 a 1500 nucleotídeos, sendo que o número de éxons variou de 3 a 12. Os tamanhos das proteínas preditas pelos oito candidatos variaram de 195 a 500 aminoácidos, sendo que o tamanho reduzido das proteínas efetoras é coincidente na maioria dos candidatos a efetores já estudados (SAUNDERS et al., 2012; STERGIPOULOS; DE WIT, 2009).

**Figura 4** – Representação da estrutura molecular dos candidatos a proteínas efetoras de acordo com o modelo gênico no genoma de referência MT2006. Estão representadas as regiões codificantes de proteínas (CDS), bem como os éxons (amarelo), íntrons (preto) e regiões *upstream* e *downstream* (azul). Além disso, também estão representados o peptídeo sinal (vermelho) e o domínio conservado Egh16-like (verde), juntamente com o tamanho e seu alinhamento na CDS.



**Fonte:** Adaptado de JGI MycoCosm, 2021.

Em todas as proteínas identificadas, o domínio Egh16-like (DUF3129/PF11327) foi localizado, em sua maioria, entre as posições 18 e 195, e abrangeu 167 aminoácidos em média. Nos candidatos em que o domínio estava completo, como o de\_novo\_555, de\_novo\_2238 e o de\_novo\_1784, o tamanho do domínio foi de 172 aminoácidos, enquanto que, nos demais candidatos, o tamanho variou de 151 a 166 (Tabela 3). Um estudo recente demonstrou que este domínio é requerido pelos fungos na penetração da

cutícula vegetal, através da deleção dos genes *Mras1-Mras7*, oriundos de *M. oryzae*, resultando na incapacidade do fungo penetrar a cutícula do hospedeiro (HUANG et al., 2019), que é a principal barreira de defesa constitutiva das plantas (PASCHOLATI; DALIO, 2018). Ainda nesse estudo, foram utilizadas proteínas que continham o domínio DUF3129 de diferentes tamanhos, inferindo-se que os tamanhos podem variar entre as proteínas de uma mesma espécie (HUANG et al., 2019).

**Tabela 3** – Dados oriundos da busca pelo domínio conservado Egh16-like nas sequências preditas dos candidatos a efetores de *Phakopsora pachyrhizi* selecionados para o estudo.

Candidato	Alinhamento <sup>1</sup>		Envelope <sup>2</sup>	
	Start	Stop	Start	Stop
de_novo_251	18	183	18	193
de_novo_2595	26	187	23	194
de_novo_555	19	191	18	191
de_novo_635	16	151	7	154
de_novo_2238	22	194	22	194
de_novo_1784	21	193	21	193
MT2006_6414149	22	195	22	195
MT2006_7713431	41	207	41	211

<sup>1</sup>Valores de alinhamento referem-se à região sobre a qual o alinhamento da sequência ao perfil HMM está correto.

<sup>2</sup>Valores de envelope delimitam a região na sequência onde a correspondência foi determinada probabilisticamente.

Fonte: Xfam, 2020

### 3.5.4 Análises comparativas de proteínas candidatas a efetores de *P. pachyrhizi* em diferentes filós e espécies

Sequências de proteínas preditas de 17 genomas fúngicos, presentes nos filós Basidiomycota e Ascomycota, foram agrupados de acordo com a sua similaridade e assim foi possível identificar os agrupamentos de proteínas correspondentes para cada candidato a efetor selecionado. Foram utilizadas as proteínas dos modelos gênicos correspondentes a cada um dos candidatos e que pertencem ao genoma de referência MT2006.

A partir desta análise foram identificados oito clusters distintos, sendo que cada um deles corresponde a um candidato selecionado; nenhum dos candidatos se agrupou no mesmo cluster (Tabela 4). Dentre eles, quatro clusters possuem membros oriundos apenas de fungos causadores de ferrugens (5375, 5918, 6583 e 5950) onde estão localizados os efetores 251, 2595, 635 e 6414149, enquanto outros dois clusters contém proteínas exclusivas de *P. pachyrhizi* (12374 e 12811), onde estão localizados os efetores 1784 e 7713431. Em contrapartida, dois clusters (3677 e 5230) possuem membros de *P. pachyrhizi* como único

patógeno causador de ferrugem e espécies patogênicas causadoras de outras doenças, como podridão de raiz, e membros que oriundos do filo Ascomycota, consistindo em efetores mais conservados entre as espécies.

**Tabela 4** – Agrupamento das proteínas preditas de *Phakopsora pachyrhizi* nos três genomas de referência disponíveis e proteínas preditas de 14 genomas de fungos fitopatogênicos pertencentes ao filo Ascomycota e Basidiomycota (Pucciniales). Cada cluster corresponde a um candidato a efetor selecionado.

Genomas de Referência <sup>1</sup>	Cluster ID							
	3677	5230	5375	5918	6583	5950	12374	12811
<i>Phakopsora pachyrhizi</i> MT2006 v1.0	2	4	2	1	1	1	3	1
<i>Phakopsora pachyrhizi</i> UFV02 v2.1	1	3	1	1	1	1	1	1
<i>Phakopsora pachyrhizi</i> K8108 v2.0	2	3	1	1	1	1	0	2
Puctr1 - <i>Puccinia triticina</i> 1-1 BBBB Race 1	0	0	2	1	1	1	0	0
Pucgr2 - <i>P. graminis</i> f. sp. <i>tritici</i> v2.0	0	0	1	2	1	2	0	0
Mellp2_3 - <i>Melampsora larici-populina</i> v2.0	0	0	1	1	1	1	0	0
Melap1finSC_191 - <i>M. allii-populina</i> 12AY07	0	0	1	1	1	1	0	0
Croqu1 - <i>Cronartium quercuum</i> f. sp. <i>fusiforme</i> G11	0	0	1	1	1	1	0	0
Sporo1 - <i>Sporobolomyces roseus</i> v1.0	0	0	0	0	0	0	0	0
Micld1 - <i>Microbotryum lychnidis-dioicae</i> p1A1	0	0	0	0	0	0	0	0
Mixos1 - <i>Mixia osmundae</i> IAM 14324 v1.0	1	0	0	0	0	0	0	0
Ustma2_2 - <i>Ustilago maydis</i> 521 v2.0	0	0	0	0	0	0	0	0
Spore1 - <i>Sporisorium reilianum</i> SRZ2	0	0	0	0	0	0	0	0
Lacbi2 - <i>Laccaria bicolor</i> v2.0	0	0	0	0	0	0	0	0
Hetan2 - <i>Heterobasidion annosum</i> v2.0	1	0	0	0	0	0	0	0
Magor1 - <i>Magnaporthe oryzae</i> 70-15 v3.0	1	1	0	0	0	0	0	0
Blugr2 - <i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14	7	0	0	0	0	0	0	0
<b>Total</b>	15	11	10	9	8	9	4	4
<b>Candidatos</b>	555	2238	251	2595	6414149	635	1784	7713431

<sup>1</sup>Genomas disponíveis no portal MycoCosm.

Todas as sequências foram obtidas e utilizadas para construção de uma árvore filogenética (Figura 5). O candidato de\_novo\_555 pertencente ao cluster 3677 é o que tem maior número de membros e de número de espécies, como *P. pachyrhizi*, *M. osmundae*, *H. annosum*, *M. oryzae* e *B. graminis* f. sp. *hordei*. Pela análise filogenética, observou-se que *M. oryzae* e *B. graminis* f. sp. *hordei*, ambos ascomicetos, estão intimamente relacionados entre si e igualmente relacionados com *M. osmundae*, e todos estes por sua vez têm um ancestral comum com *H. annosum* e *P. pachyrhizi*.

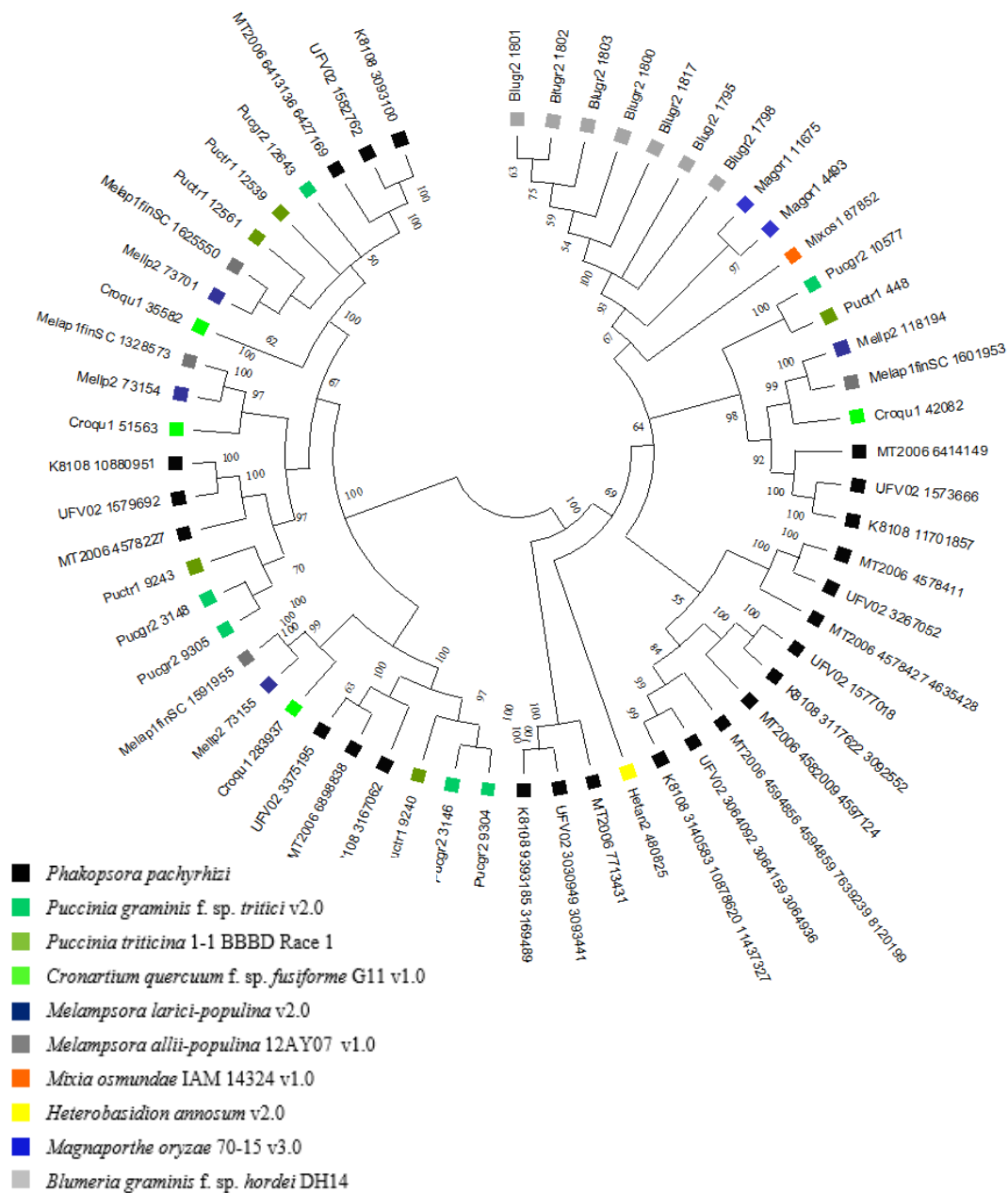
Outro cluster com perfil semelhante ao 3677 é o cluster 5230, do candidato 2238, pois possui apenas proteínas oriundas de *P. pachyrhizi* e de *M. oryzae*. Apesar de apresentarem fungos relacionados filogeneticamente, os dois clusters apresentados possuem membros muito diversos em relação ao filo e também ao seu modo de sobrevivência, como

fungos biotróficos e necrotrófico (*H. annosum*). Essa diversidade genética deve-se a regulação e ao conteúdo gênico diferenciado entre espécies, visto que, todas elas possuem um ancestral comum, mas adaptaram-se a plantas hospedeiras e estilos de vida completamente diferentes. Dados semelhantes foram observados por de Wit et al. (2012), que através de análises filogenéticas dos genomas de *Cladosporium fulvum*, fungo biotrófico que infecta o tomateiro, e *Dothistroma septosporum*, patógeno hemibiotrófico de pinheiros, constataram que dois os patógenos estão intimamente relacionados, entretanto, possuem características evolutivas muito divergentes devido ao conteúdo gênico de cada espécie.

Os clusters 5375, 5918, 6583 e 5950 são aqueles que apresentaram proteínas preditas oriundas apenas de Pucciniales. Cada um destes grupos abrangeu um cluster da árvore filogenética, e observou-se que todos possuem uma mesma conformação, contendo proteínas preditas oriundas dos três genomas de *P. pachyrhizi*. Em um sub-cluster são encontradas proteínas de *P. triticina* e *P. graminis* f. sp. *tritici*, agentes causais das ferrugens da folha e do colmo de trigo, respectivamente. Enquanto que, as proteínas de *M. allii-populina*, *M. larici-populina* – causadores de ferrugem em álamos, e *C. quercuum* f. sp. *fusiforme* – agente causal da ferrugem fusiforme em pinheiros, também se encontram agrupadas. Assim, observou-se que ocorre uma associação entre as espécies patogênicas e as plantas hospedeiras, indicando que estes candidatos não são generalistas e possuem especificidade com o seu hospedeiro. A similaridade entre candidatos a efetores de *P. pachyrhizi* e proteínas secretadas de fungos fitopatogênicos dos gêneros *Puccinia* e *Melampsora* também foi observada por Rincão et al. (2018). De modo similar, Utiyama (2020) constatou a similaridade entre proteínas secretadas preditas nos genomas de *C. quercuum* f. sp. *fusiforme*, *M. americana*, *M. lini*, *M. larici-populina*, *M. medusae* f. sp. *deltoidae*, *P. graminis* f. sp. *tritici*, *P. striiformis* f. sp. *tritici*, *P. triticina* e *P. pachyrhizi*, evidenciando que estas espécies possuem um ancestral comum.

Os dois clusters específicos de *P. pachyrhizi*, 12374 e 12811, são observados em clusters distintos e específicos para estas proteínas. O candidato de\_novo\_1784 pertence ao cluster 12374, enquanto que, o candidato 7713431 ao 12811. Por estarem inseridos em clusters que contém apenas proteínas oriundas de *P. pachyrhizi*, infere-se que estes dois candidatos a efetores são pouco conservados entre espécies, o que pode ser devido à necessidade de driblar o sistema de defesa da soja, sofreram modificações durante a evolução do patógeno. Uma família gênica tão diversa como Egh16-like, com membros conservados, mas também específicos entre as ferrugens e mesmo na espécie *P. pachyrhizi*, pode indicar caminhos evolutivos diferentes, sendo que embora tenham um domínio em comum, seu papel no parasitismo e virulência também pode ser diverso.

**Figura 5** – Análise filogenética dos candidatos a efetores de *Phakopsora pachyrhizi* comparados com outras espécies de fungos fitopatogênicos. Os clusters foram obtidos de acordo com a análise MCL Cluster de proteínas preditas nos genomas de referência de *P. pachyrhizi* e oriundas de 14 genomas de fungos. O método utilizado para montagem da árvore foi o *Neighbor-Joining* (1000x), sendo que os números apresentados próximos aos nós indicam a porcentagem em que as sequências aparecem agrupadas. As sequências são designadas pelo nome da espécie e por seus respectivos números de acesso na base de dados do JGI. As espécies são identificadas por cores diferentes, que estão apresentadas na legenda.



### 3.5.5 Perfil de expressão de genes candidatos a efetores de *P. pachyrhizi* durante interação com a soja

O perfil transcricional dos oito genes alvos selecionados de *P. pachyrhizi* (251, 2595, 555, 635, 1784, 2238, 6414149 e 7713431) foi avaliado em diferentes tempos do ciclo de infecção do patógeno no genótipo suscetível (BRS184). Após a coleta do material biológico para as análises de expressão, as folhas remanescentes das plantas inoculadas foram mantidas e apresentaram lesões TAN, típicas de genótipos suscetíveis ao fungo (Figura 1 – Apêndice C).

Os níveis de expressão variaram entre os candidatos nos diferentes tempos e, de modo geral, foi observada uma elevada concordância entre os níveis de expressão entre as repetições biológicas, bem como nas duas repetições das corridas de RT-qPCR realizadas. De uma forma geral, os maiores níveis de expressão foram observados nos tempos 12 e 24 hpi, enquanto que, os menores, no tempo de 48 hpi (Tabela 5).

O 1784 destacou-se entre os genes analisados por apresentar os maiores níveis de expressão em todos os tempos, demonstrando expressão significativa desde os tempos de 12 hpi até 48 hpi em relação ao tempo inicial de infecção. O pico de expressão foi atingido em 12 hpi, enquanto que, nos dois tempos posteriores, o nível de expressão apresentou um perfil decrescente, mas ainda assim, induzida. Estes resultados indicam a importância desse gene desde as etapas iniciais do processo infeccioso, sendo possivelmente importante nas etapas de germinação e penetração.

Um perfil semelhante foi observado no candidato 635, que apresentou expressão crescente após 12 hpi, sendo detectados os níveis mais elevados e significativos de regulação nos tempos de 24 e 48 hpi, e também no 7713431, que teve a expressão induzida em 12 e 24 hpi. Tal como, o candidato 2238 também demonstrou ter uma alta indução em 12 hpi, mas sofre uma drástica queda nos dois últimos tempos analisados, apresentando uma regulação negativa 48 hpi. Estes tempos coincidem com as fases de crescimento das hifas do fungo no espaço intercelular e formação do haustório e tais resultados ratificam a provável função efetora, onde sua indução coincide com eventos importantes no parasitismo e biotrofia.

O candidato 555 apresentou os menores níveis de expressão entre todos os selecionados, sendo que a sua maior expressão foi observada no tempo de 12 hpi. Desse modo, este gene candidato foi considerado pouco induzido na presença do hospedeiro, pelo menos nos tempos e isolado avaliados.

Níveis de expressão e perfis semelhantes foram observados nos candidatos 251, 2595 e 6414149, que não foram induzidos em nenhum dos tempos avaliados, indicando

ou que são induzidos tardiamente na interação com a soja, em tempos não avaliados neste ensaio, ou podem não ocorrer no isolado utilizado neste estudo.

**Tabela 5** – Análise de expressão dos oito candidatos a efetores de *Phakopsora pachyrhizi* durante a interação com a soja. São apresentados os níveis de expressão gênica dos candidatos em folhas de soja infectadas e coletadas em 6, 12, 24 e 48 horas pós inoculação (hpi). Os dados obtidos por RT-qPCR são apresentados como expressão relativa ao gene endógeno *CytB* e o tempo 0h utilizado como calibrador.

Genes	Diferentes tempos de infecção			
	6 hpi	12 hpi	24 hpi	48 hpi
251	1,023	0,591	1,054	0,09
2595	1,147	1,01	1,064	0,152
555	0,709	1,001	0,772	0,037
635	0,991	1,113	<b>1,235*</b>	<b>1,373*</b>
1784	2,136	<b>59,2*</b>	<b>34,846*</b>	<b>4,539*</b>
2238	1,252	<b>2,369*</b>	0,696	0,41
6414149	1,108	1,426	1,018	0,173
7713431	<b>1,602*</b>	<b>1,563*</b>	<b>2,885*</b>	1,25

\*Representa expressão diferencial estatisticamente significativa.

### 3.5.6 Validação dos efetores via expressão gênica a partir de dados de RNA-Seq

Dados de RNA-Seq disponíveis pelo Consórcio do Genoma de *P. pachyrhizi* também foram utilizados para validação dos efetores analisados neste trabalho, visto se tratar de três isolados distintos, e cobrindo coletas ao longo de todo o ciclo infeccioso.

Os dados oriundos do transcriptoma do isolado MT2006 foram gerados nos tempos tardios de infecção, 192 e 288 hpi, além da expressão nas estruturas do fungo: esporo e apressório. A maioria dos modelos gênicos correspondentes aos candidatos selecionados apresentou uma indução negativa ou baixa expressão em 192 e 288 hpi. Os modelos gênicos 4578427, 6898838 e 7713431 foram os únicos que foram induzidos nos dois tempos, sendo que o último modelo gênico citado apresentou a maior indução, demonstrando ser requerido nos momentos tardios da infecção. Os modelos 4578427 e 7713431 também foram induzidos no esporo apressório, respectivamente.

Quanto aos dados referentes ao transcriptoma do isolado UFV02, observou-se uma concordância com os dados obtidos por RT-qPCR gerados neste trabalho,

principalmente para os modelos gênicos 3093441 e 3030949, que correspondem ao candidato 7713431. Os modelos apresentaram uma expressão crescente durante a interação, inclusive nos tempos 12 e 24 hpi, que também foram analisados por RT-q PCR. Neste transcriptoma também foi possível validar a expressão tardia do gene 1573666, correspondente ao gene 6414149 no MT2006, o que pode explicar o fato deste candidato não ter sido detectado nos ensaios de RT-qPCR devido a análise apenas nos tempos iniciais de infecção avaliados.

Por último, o transcriptoma do K8108 revelou picos de expressão nos tempos 10 e 24 hpi para 4 modelos gênicos, dois deles correspondentes aos genes 7713431, e aos transcritos 1784 e 251. Neste transcriptoma então foi possível novamente verificar a indução do 1784 e validar a expressão do 251. De modo similar ao observado para o gene 6414149, o modelo 10880951, que corresponde ao candidato 2595, também apresentou os maiores níveis e indução nos tempos tardios, após às 72 h, indicando que estes candidatos podem ser acionados em momentos mais tardios da infecção. Finalmente, os modelos gênicos 3169489 e 9393185, correspondentes ao candidato 7713431, apresentaram um perfil semelhante ao observado nos outros dois genomas, com pico de expressão nos tempos 10 e 24 hpi. Este gene teve o perfil de expressão mais concordante entre todos os transcriptomas, e demonstra ser uma proteína requerida ao longo do processo infeccioso.

As divergências entre os perfis de expressão observadas são esperadas e podem ser relacionadas a diferentes fatores, dentre estes o fato de ser experimentos diferentes, utilizando isolados e diferentes acessos de soja em cada experimento, além de se tratar de diferentes metodologias de análise. No entanto, dos oito efetores selecionados, com exceção do alvo 555, todos os demais foram induzidos em pelo menos uma fase do ciclo infeccioso, com destaque para 1784, 635 e 7713431.

**Tabela 6** – Expressão gênica obtida por RNA-Seq dos modelos gênicos correspondentes aos candidatos a efetores de *Phakopsora pachyrhizi* durante a interação com a soja e na ausência dela. Perfil de expressão dos genes do genoma MT2006 durante 192 e 288 hpi e na ausência do hospedeiro (esporo e apressório). Perfil de expressão das proteínas oriundas do isolado UFV02 durante 12, 24, 36, 72 e 168 hpi, além do esporo. Perfil de expressão dos genes oriundos do isolado K8108 durante 10, 24, 72 e 196 hpi, além do esporo e apressório.

Transcrito	MT2006					UFV02							K8108						
	Modelo Gênico	Esporo	Apressório	192hpi	288hpi	Modelo Gênico	Esporo	12 hpi	24 hpi	36 hpi	72 hpi	168 hpi	Modelo Gênico	Esporo	Apressório	10 hpi	24 hpi	72 hpi	196 hpi
<b>2595</b>	4578227	-0,29	-0,10	-4,12	-1,38	1579692	-0,34	-0,65	-0,64	-0,87	-1,12	<b>3,70</b>	10880951	-0,03	-0,33	-1,20	-4,36	-7,01	-4,49
<b>1784</b>	4578427	0,78	0,25	<b>1,47</b>	<b>1,83</b>	3267052	-0,48	0,56	-0,89	-3,13	-2,19	-0,02	3115284	0,95	<b>11,34</b>	<b>7,70</b>	<b>7,29</b>	<b>5,13</b>	<b>1,28</b>
<b>555</b>	4582009	-0,28	-0,10	-4,12	-1,38	1577018	0,04	0,26	0,05	0,16	-0,1	-0,06	3092552	0,06	-0,26	-2,38	-4,46	-9,73	-14,42
<b>251</b>	6413136	-0,47	-0,42	-8,61	-2,39	1582762	-0,26	-0,59	-0,51	-0,50	-1,54	-0,80	3093100	1,20	-1,10	<b>7,70</b>	<b>7,29</b>	<b>5,13</b>	<b>1,28</b>
<b>635</b>	6898838	0,28	-0,34	0,55	0,65	3375195	0,13	-1,80	-0,50	0,77	<b>2,68</b>	<b>7,51</b>	3167062	-0,08	-1,08	-1,55	-1,96	-1,47	-0,50
<b>2238</b>	4594856	-0,68	0,41	-2,71	-0,92	3064092	0,30	-1,19	-2,28	-0,49	-1,82	-0,5	3140583	-0,29	-0,77	-2,23	-2,64	-4,80	-8,65
<b>6414149</b>	6414149	-1,43	0,18	-2,04	-1,70	1573666	-1,14	0,37	<b>3,73</b>	<b>2,73</b>	<b>3,52</b>	<b>4,01</b>	11701857	0,17	-0,07	<b>3,29</b>	<b>1,91</b>	-0,24	-4,09
<b>7713431</b>	7713431	0,03	<b>1,27</b>	<b>5,37</b>	<b>4,97</b>	3030949	-2,45	<b>3,81</b>	<b>5,33</b>	<b>5,05</b>	<b>6,81</b>	<b>11,25</b>	3169489	-0,74	0,79	<b>4,57</b>	<b>4,16</b>	<b>2,01</b>	<b>2,21</b>

Os modelos gênicos dos três genomas agrupados na mesma linha correspondem a um mesmo candidato a efetor.

Conforme previamente descrito, genes contendo o domínio Egh16-like atuam especialmente nas etapas iniciais de infecção pelos patógenos, período em que ocorre a formação do tubo germinativo e apressório (GRELL et al., 2003; JUSTESEN et al., 1996; XUE et al., 2002). O apressório possui grande importância para os fungos patogênicos, pois através desta estrutura é que ocorre a penetração direta no hospedeiro. Em análises de expressão dos genes de *B. graminis* f. sp. *hordei* na infecção de plantas de cevada, Grell et al. (2003) observaram uma regulação positiva e um nível alto de expressão nos estágios iniciais de infecção de genes contendo o Egh16-like, enquanto no terceiro dia após inoculação, estes níveis decresceram.

Através da deleção dos genes GAS2 e GAS1, ortólogos ao Egh16-like em *Magnaporthe grisea*, Xue et al. (2002) constataram a redução na penetração do apressório bem como no desenvolvimento de lesões, evidenciando a alta expressão deste domínio na fase de penetração do fungo e seus estágios iniciais de infecção. Outro ortólogo ao Egh16-like, Magas1, também foi requerido especificamente durante o período de penetração, no fungo entomopatogênico *Metarhizium acridum* (CAO et al., 2012). Além disso, genes com domínio Egh16-like, também foram altamente expressos durante a formação do apressório no fungo patogênico *M. robertsii* em estudos de expressão gênica (HUANG et al., 2019). Sendo assim, os resultados de expressão gênica observadas neste estudo demonstram a relação dos candidatos selecionados com o domínio conservado, que é altamente expresso nos momentos iniciais de infecção, tanto em fungos patogênicos biotróficos, hemibiotróficos e necrotróficos, mas que necessitam penetrar direta ou indiretamente seu hospedeiro.

Vários autores verificaram o perfil de expressão gênica ao validar candidatos a efetores de *P. pachyrhizi* (DE CARVALHO et al., 2016; KUNJETI et al., 2016; QI et al., 2016; RINCÃO et al., 2018; STONE et al., 2015). O candidato de\_novo\_2238 foi um dos transcritos mais expressos em análises de candidatos a efetores de *P. pachyrhizi* na interação com diferentes genótipos de soja (RINCÃO et al., 2018), evidenciando sua alta indução no período de infecção. Nas análises de expressão dos transcriptomas da Família 1 em diferentes tempos de inoculação, De Carvalho et al. (2016) observaram uma baixa regulação durante os tempos associados a formação do haustório de *P. pachyrhizi*. Entretanto, os genes podem demonstrar sua atividade efetora ao longo de todo o ciclo de infecção, como o candidato 5381 descrito no estudo, que apresentou um nível de expressão em 192 hpi semelhante ao observado em 48 hpi, enquanto que, seu menor nível foi em 96 hpi. Assim, esta seria uma hipótese para a baixa indução do candidato 555, que possui todos os atributos de uma proteína efetora, porém teve uma indução negativa nos tempos avaliados por RT-qPCR e foi pouco expresso nas

análises de RNA-Seq.

Além disso, alguns efetores podem ter sua expressão influenciada pela interação ou presença de outros efetores na célula ou então não serem necessários no início da infecção, sendo requeridos em outros momentos da infecção (MA et al., 2017). O papel da proteína no processo infeccioso também pode refletir os picos de expressão; alguns efetores podem ter funções distintas, como induzir a liberação de nutrientes para o patógeno ou ainda causar alterações nas células ou tecidos infectados, indicando que a expressão gênica de efetores pode ser altamente diversa em todo o período infeccioso (QI et al., 2019). Análises do perfil de expressão gênica do efector de *P. pachyrhizi* PpEC23, demonstraram altos níveis de transcrição durante a formação do apressório, mas esses níveis também continuaram a aumentar ao longo do curso de infecção, indicando que este efector não é exclusivamente expresso em haustórios e nos períodos iniciais de infecção (QI et al., 2016). Em outro estudo com candidatos a efetores de *P. pachyrhizi* na interação com a soja, os candidatos CSEP-06, CSEP-08 e CSEP-09 tiveram altos níveis de indução apenas nas primeiras 24 hpi, ao contrário dos candidatos CSEP-03 e CSEP-07, que tiveram níveis de transcrição indetectáveis nos períodos iniciais da infecção, mas foram induzidos após 72 hpi (KUNJETI et al., 2016). Perfil semelhante ao observado para o candidato, que foi induzido apenas em 168 hpi, ao contrário do candidato 1784 e 2238 que foram requeridos nos períodos iniciais da infecção e diminuem sua atividade nos mais tardios.

Desse modo, podemos especular que os candidatos 251, 2595 e 6414149, embora não tenham sido induzidos nos tempos iniciais da interação, podem ainda atuar como efetores, sendo suas funções associadas às fases mais tardias do desenvolvimento do fungo. Além disso, membros de uma mesma família de proteínas podem ter funções em diferentes estágios durante o ciclo infeccioso. O efector de *Uf*-HSP44, de *U. fabales*, é altamente regulado durante os estágios iniciais de infecção, enquanto *Uf*-HSP39 é fracamente expresso na emissão do haustório (DUPLESSIS et al., 2011). Além disso, Qi et al. (2019), analisando candidatos a efetores de *U. appendiculatus*, causador da ferrugem no feijão, constataram que apesar de não possuir uma forte indução durante a patogênese, o candidato *Uaca\_22* suprimiu HR em plantas inoculadas com o patógeno, indicando que além de estudos sobre a expressão gênica dos candidatos, outros ensaios funcionais são importantes para inferir a função efetora.

Sendo assim, a correlação da indução da expressão de um efector com as etapas de infecção do patógeno é uma importante estratégia para avaliar o envolvimento do gene no parasitismo. Porém, além das análises de expressão, estudos referentes a supressão das respostas de defesa PTI e ETI são necessários e capazes de validar a função efetora de genes preditos como secretados por fungos fitopatogênicos.

### 3.5.7 Análise de sintenia

Ao analisar os candidatos de *de\_novo\_2238* e de *de\_novo\_1784* observou-se que ambos possuem identidade com diferentes modelos gênicos, sendo que algumas destas cópias se encontram localizadas fisicamente próximas no genoma. Dos oito modelos gênicos similares ao efector 1784, três deles estão localizados na mesma região genômica, enquanto o mesmo foi observado para o candidato 2238, onde dos cinco modelos gênicos similares, duas cópias idênticas estão próximas, como pode ser observado na descrição da localização física de cada gene (Tabela 7).

**Tabela 7** – Identificação das regiões contendo os modelos gênicos correspondentes aos candidatos a efetores de *Phakopsora pachyrhizi* de *de\_novo\_1784* e de *de\_novo\_2238* no genoma MT2006.

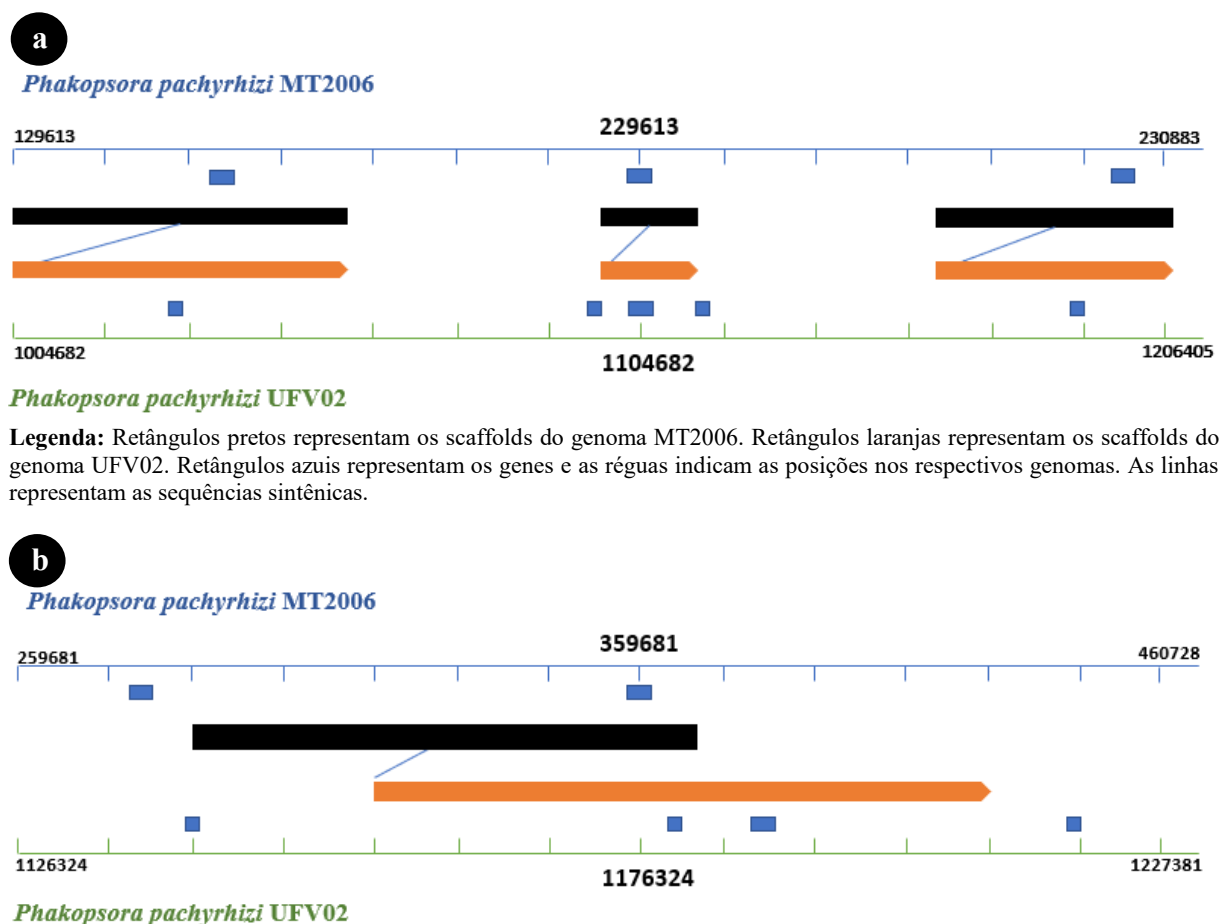
Modelo Gênico	Transcrito de <i>de_novo</i> correspondente	Localização	Genoma de referência
4635428	1784	scaffold_150: 229613-230883	<b>MT2006</b>
4578411	1784	scaffold_150: 359681-360738	
4578427	1784	scaffold_150: 632138-633201	
4594856	2238	scaffold_2101: 56247-57189	
4594859	2238	scaffold_2101: 92746-93688	
3266447	1784	Contig_155: 286068-287634	<b>UFV02</b>
3267052	1784	Contig_155: 1104682-1106405	
3001461	1784	Contig_155: 1226324-1227381	
3064092	2238	Contig_17: 1234816-1236041	
3064159	2238	Contig_17: 1274165-1275395	
3064936	2238	Contig_17: 2353417-2354644	
9544853	1784	tig00009039_3192: 147768-149317	<b>K8108</b>
3108701	1784	tig00009039_3192: 392842-393901	
10878620	2238	tig00031356_5795: 281160-282203	
11437327	2238	tig00031922_6194: 90232-91278	

Fonte: JGI MycoCosm, 2021.

A análise comparativa destas regiões permitiu revelar uma sintenia entre os diferentes genomas (Figura 6). Assim, constatou-se que o scaffold\_150: 229.613 – 230.883, correspondente ao candidato 1784 no genoma do isolado MT2006, é sintênico ao contig\_155: 1.104.682 – 1.106.405 no genoma UFV02. Além disso, quando analisadas regiões na redondeza deste *scaffold* (100 Kbp antes e após o gene), foram identificadas outras regiões sintênicas entre

os genomas. De modo similar, o scaffold\_150: 632.138 – 633.201 revelou sintenia com o contig\_155: 1.226.324 – 1.227.381. O último *scaffold* correspondente ao genoma MT2006 não apresentou sintenia com o genoma UFV02 e K8108, o qual não possui sintenia com nenhum dos citados anteriormente.

**Figura 6** – Análise da sintenia dos genomas de referência de *Phakopsora pachyrhizi* contendo genes candidatos a efetores. (a) Regiões sintênicas entre o scaffold\_150: 229.613 - 230.883 do isolado MT2006 e o contig\_155: 1.104.682 - 1.106.405 do UFV02. Além da região em que o candidato a efector está inserido, observa-se que regiões próximas ao gene também possuem sintenia entre os genomas. (b) Regiões sintênicas entre o scaffold\_150: 359.681 – 360.738 x contig\_155: 1.226.324 – 1.227.381 entre os isolados MT2006 e UFV02.



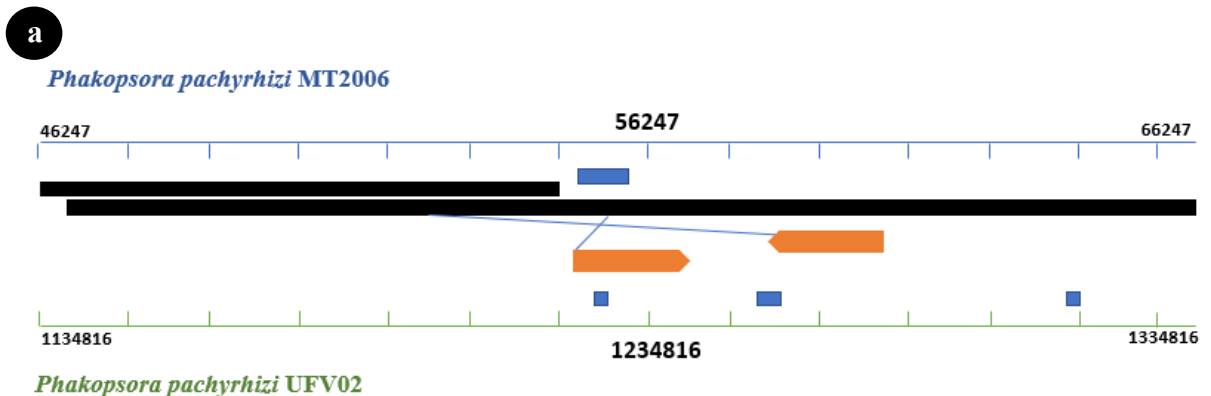
**Legenda:** Retângulos pretos representam os scaffolds do genoma MT2006. Retângulos laranjas representam os scaffolds do genoma UFV02. Retângulos azuis representam os genes e as réguas indicam as posições nos respectivos genomas. As linhas representam as sequências sintênicas.

**Fonte:** Adaptado de JGI MycoCosm, 2021.

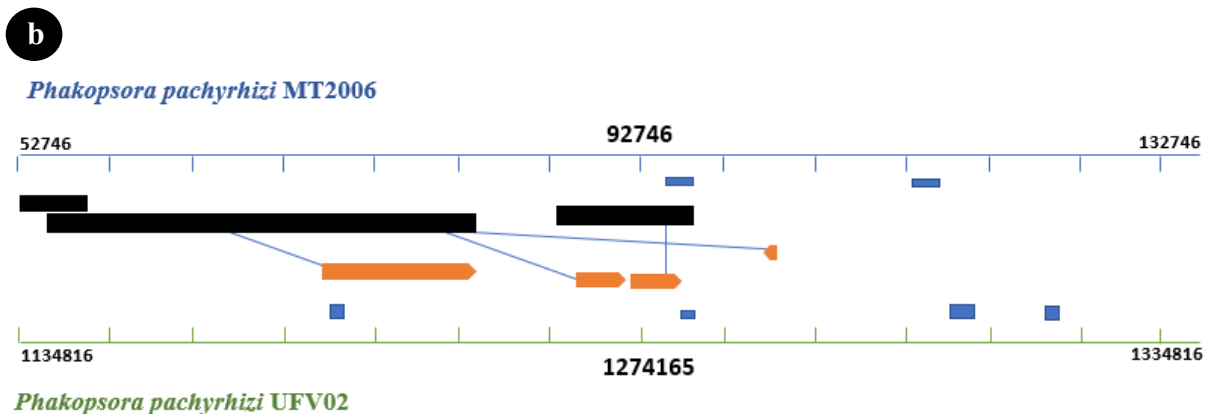
Quanto aos *scaffolds* referentes ao candidato 2238, o scaffold\_2101: 56.247 - 57.189 demonstrou ser sintênico apenas ao contig\_17: 1.234.816 - 1.236.041. Entretanto, foi encontrada uma região sintênica próxima a este *scaffold* e ao contig\_17: 1.274.165 - 1.275.395 (Figura 7 – a). O outro *scaffold* referente a este candidato também é sintênico a uma única

posição no contig\_17, localizado no intervalo 1.274.165 - 1.275.395. Também observou-se a presença de outras quatro regiões do genoma UFV02 sintênicas a uma região do MT2006 (Figura 7 – b). As cópias referentes ao genoma K8108 só foram sintênicas quando comparadas ao genoma UFV02 e somente aquelas que correspondem ao candidato a efetor 2238. Apenas o tig00031356\_5795: 281.160 - 282.203 possui sintenia com o gene inserido no contig\_17, localizado na posição 2.353.417 a 2.354.644.

**Figura 7** – Análise da sintenia dos genomas de referência de *Phakopsora pachyrhizi* contendo proteínas candidatas a efetores. (a) Regiões sintênicas entre o scaffold\_2101:56247-57189 do genoma MT2006 e o contig\_17:1234816-1236041 referente ao UFV02. Nota-se que uma região próxima ao gene também possui sintenia entre os genomas. (b) Sintenia entre o scaffold\_2101:92746-93688 e o contig\_17:1274165-1275395.



**Legenda:** Retângulos pretos representam os scaffolds do genoma MT2006. Retângulos laranjas representam os scaffolds do genoma UFV02. Retângulos azuis representam os genes e as réguas indicam as posições nos respectivos genomas. As linhas representam as sequências sintênicas.



**Legenda:** Retângulos pretos representam os scaffolds do genoma MT2006. Retângulos laranjas representam os scaffolds do genoma UFV02. Retângulos azuis representam os genes e as réguas indicam as posições nos respectivos genomas. As linhas representam as sequências sintênicas.

**Fonte:** JGI MycoCosm, 2021.

Apesar do sequenciamento com diferentes tecnologias, foram observadas regiões sintênicas entre os diferentes genomas, indicando que esses dois candidatos desempenham papéis na patogenicidade, pois são genes ricos em repetição e estão associados a compartimentos genômicos esparsos genes (RAFFAELE et al., 2010).

Com o avanço das técnicas de sequenciamento e a montagem de genomas cada vez mais complexos, avança-se também na busca e no entendimento dos eventos evolutivos que abrangem as mais diversas espécies de patógenos de plantas, que possivelmente determinarão formas de controle mais sustentáveis e menos nocivas ao homem e ao ambiente.

### 3.6 CONCLUSÃO

A família de proteínas contendo o domínio Egh16 foi caracterizada no genoma de *P. pachyrhizi*, sendo identificados pelo menos 22 modelos gênicos preditos no genoma do MT2006, todos com pelo menos um modelo correspondente nos genomas de K8108 e UFV02. Estes membros variaram em termos do tamanho da proteína predita e com relação ao tamanho do domínio Egh16, mas todos os membros apresentaram o domínio na porção N-terminal da proteína. Estes membros também revelaram um conteúdo médio de cisteína de 3% e a presença de peptídeo sinal, características comuns entre proteínas secretadas de fitopatógenos.

Com base nestes resultados, oito dos 22 transcritos previamente descritos por De Carvalho et al, (2016) puderam ser validados com base nos genomas de referência disponíveis, demonstrando que o estudo *in vivo* do transcriptoma do fungo foi capaz de revelar transcritos que são também representativos no genoma, mas que o genoma de referência é importante para dimensionar o número de membros e suas relações de similaridade e conservação.

Análises de similaridade entre os membros identificados no genoma de MT2006 revelaram pelo menos oito ramos distintos, demonstrando variabilidade nas proteínas identificadas. A análise do perfil de expressão revelou um perfil variado, com efetores sendo induzidos nos momentos iniciais da infecção (12, 24 e 48 hpi), enquanto pelo menos dois efetores, 2595 e 6414149, foram induzidos nos momentos tardios, em pelo menos um dos estudos de expressão analisados.

A análise comparativa entre oito membros da família Egh16-like com as proteínas de outras 15 espécies fúngicas revelou membros possivelmente conservados entre as

espécies, bem como comuns apenas nas espécies de ferrugens e ainda membros exclusivos de *P. pachyrhizi*, que é uma ferrugem autoécia, revelando a coevolução do patógeno com seu hospedeiro e a complexidade do grupo Pucciniales.

## REFERÊNCIAS

- ABE, V. Y. **Expressão transiente de genes de *Phakopsora pachyrhizi* em genótipos resistentes de soja visando à identificação de genes de avirulência.** 2012. 62 p. Dissertação (Mestrado em Fitopatologia) - Universidade Federal de Viçosa, Viçosa, 2012.
- AHUJA, I.; KISSEN R.; BONES, A.M. Phytoalexins in defense against pathogens. **Trends in Plant Science.** v. 17, p. 73-90, 2012.
- AMORIM, L.; PASCHOLATI, S.F. Ciclo de Relações Patógeno-Hospedeiro. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A. **Manual de Fitopatologia: Princípios e Conceitos.** 5. ed. Ouro Fino: Agronômica Ceres, 2018. p. 355-360.
- ANDERSSON, K.M.; KUMAR, D.; BENTZER, J.; FRIMAN, E.; AHRÉN, D.; TUNLID, A. Interspecific and host-related gene expression patterns in nematode-trapping fungi. **BMC Genomics.** v. 15, p. 968, 2014. <https://doi.org/10.1186/1471-2164-15-968>
- BAKKEREN, G.; SZABO, L.J. Progress on molecular genetics and manipulation of rust fungi. **Phytopathology.** v. 110, p. 532–43. 2020.
- BASU, M.K.; POLIAKO, E.; ROGOZIN, I.B. Domain mobility in proteins: functional and evolutionary implications. **Briefings in Bioinformatics.** v. 10, p. 205-216, 2009.
- BEDENDO, I.P.; Ferrugens. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A. **Manual de Fitopatologia: Princípios e Conceitos.** 5. ed. Ouro Fino: Agronômica Ceres, 2018. p. 355-360.
- BOLLER, T; FELIX, G. A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. **Ann. Rev. Plant Biol.** v. 60, p. 379-406, 2009. doi: 10.1146/annurev.arplant.57.032905.105346
- BONDE, M.R.; NESTER, S.E.; PETERSEN, G.L. Temperature effects on urediniospore germination and germ tube growth of *Phakopsora pachyrhizi* and *P. meibomia*. **Phytopathology.** v. 87, 1997.
- BOUROT, F.; ZIPFEL, C. Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. **Annual Review of Phytopathology,** v. 55, p. 257-286, 2017.
- BROMFIELD, K.R. Soybean Rust. Monograph N°. 11. St. Paul, **American Phytopathological Society,** p. 65, 1984.
- CAF. Consórcio Antiferrugem. **Custo ferrugem-asiática da soja.** Disponível em: <<http://www.consorcioantiferrugem.net/#/conteudos/view/5>>. Acesso em: 18 jun. 2019.
- CAMARGO, L.E.A. Controle Genético. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A. **Manual de Fitopatologia: Princípios e Conceitos.** 5. ed. Ouro Fino: Agronômica Ceres, 2018a. p. 229-238.

CAMARGO, L.E.A. Genética da Interação Patógeno-Hospedeiro. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A. **Manual de Fitopatologia: Princípios e Conceitos**. 5. ed. Ouro Fino: Agronômica Ceres, 2018b. p. 85-92.

CAO, Y.; ZHU, X.; JIAO, R.; XIA, Y. The Magas1 gene is involved in pathogenesis by affecting penetration in *Metarhizium acridum*. **Journal Microbioly Biotechnology**. V. 22, p. 889-893, 2012.

CASTANHO, F.M. **Caracterização estrutural e polimorfismo de genes candidatos à efetores de *Phakopsora pachyrhizi***. 2017. 67 p. Dissertação (Mestrado em Agronomia) - Universidade Estadual do Norte do Paraná, Bandeirantes, 2017.

CATANZARITI, A.M.; DODDS, P.N.; LAWRENCE, G.J.; AYLIFFE, M.A.; ELLIS, J.G. Avirulence proteins from haustoria-forming pathogens. **FEMS Microbiology Letters**, Minireview, v. 269, p. 181-188, 2007.

CHILDS, S.P.; BUCK, J.W.; LI, ZENGLU. Breeding soybeans with resistance to soybean rust (*Phakopsora pachyrhizi*). **Plant Breeding**. n. 137, p. 250–261, 2018.

CHILDS, S.P., KING, Z.R., WALKER, D.R.; HARRIS, D.K.; PEDLEY, K.F.; BUCK, J.W.; BOERMA, H.R.; LI, Z. Discovery of a seventh Rpp soybean rust resistance locus in soybean accession PI 605823. **Theoretical and Applied Genetics**. v. 131, p. 27-41, 2018. Disponível em: <<https://doi.org/10.1007/s00122-017-2983-4>>. Acesso em: 20 jun. 2019.

CHISHOLM, S.T.; COAKER, G.; DAY, B.; STASKAWICZ, B.J. Host-microbe interactions: shaping the evolution of the plant immune response. **Cell**. v. 124, p. 803-814, 2006.

COLLINS, C.; KEANE, T.M.; TURNER, D.J.; O'KEEFFE, G.; FITZPATRICK, D.A.; DOYLE, S. Genomic and proteomic dissection of the ubiquitous plant pathogen, *Armillaria mellea*: Toward a new infection model system. **Journal of Proteome Research**. v. 12, p. 2552–2570, 2013. Disponível em: <[dx.doi.org/10.1021/pr301131t](https://doi.org/10.1021/pr301131t)>.

CONAB. **Monitoramento Agrícola**. Disponível em: <<https://www.conab.gov.br/info-agro/safras/graos/monitoramento-agricola>>. Acesso em: 14 abr. 2019a.

CONAB. **Acompanhamento da Safra Brasileira: Grãos**, v. 7 - Safra 2019/20, n. 10, décimo levantamento, Brasília, p. 1-74. julho 2020b.

DALIO, R.J.D.; MAGALHÃES, D.M.; RODRIGUES, C.M.; ARENA, G.D.; OLIVEIRA, T.S.; SOUZA-NETO, R.R.; PICCHI, S.C.; MARTINS, P.M.M.; SANTOS, P.J.C.; MAXIMO, H.J.; PACHECO, I.S.; DE SOUZA, A.A.; MACHADO, M.A. PAMPs, PRRs, effectors and R-genes associated with citrus–pathogen interactions. **Annals of Botany**. v. 119, p. 749–774, 2017.

DALIO, R.J.D.; MAGALHÃES, D.M.; ATÍLIO, L.B.; RODRIGUES, C.M.; BRETON, M.C.; PICHI, S.; PASCHOLATI, S.F.; MACHADO, M.A. Efetores nas interações planta-patógeno. **Revisão Anual de Patologia de Plantas**. v. 22, p. 25-68, 2014.

DANGL, J.L.; HORVATH, D.M.; STASKAWICS, B.J. Pivoting the plant immune system from dissection to deployment. **Science**. v. 341, p. 746-751, 2013.

- DARBEN, L.M.; YOKOYAMA, A.; CASTANHO, F.M.; LOPES-CAITAR, V.S.; DE CARVALHO, M.C.C.G.; GODOY, C.V.; DE CARVALHO, S.; GONELA, A.; MARCELINO-GUIMARÃES, F.C. Characterization of genetic diversity and pathogenicity of *Phakopsora pachyrhizi* mono-uredinial isolates collected in Brazil. **European Journal of Plant Pathology**. v. 156, p. 355–372, 2020.
- DE CARVALHO, M.C.; NASCIMENTO, L.C.; DARBEN, L.M.; PODANOSQUI, A.M.P.; CAITAR, V.L.; QI, M.; ROCHA, C.S.; CARAZZOLLE, M.F.; KUWAHARA, M.K.; PEREIRA, G.A.G.; ABDELNOOR, R.V.; WHITHAM, S.A.; GUIMARÃES, F.C.M. Prediction of the *in planta* *P. pachyrhizi* secretome and potencial effector families. **Molecular Plant Pathology**. v. 18, p. 363-377, 2016.
- DE WIT, P.J.G.M.; VAN DER BURGT, A.; ÖKMEN, B.; STERGIOPOULOS, I.; ABDEL-SALAM, K.A.; AERTS, A.L.; BAHKALI, A.H.; BEENEN, H.G.; COX, M.P.; DATEMA, E.; DE VRIES, R.P.; DHILLON, B.; GANLEY, A.R.; GRIFFITHS, S.A.; GUO, Y.; HAMELIN, R.C.; HENRISSAT, B.; KABIR, M.S.; VAN DEN BURG, H.A.; ZHANG, S.; GOODWIN, S.B.; GRIGORIEV, I.V.; COLLEMARE, J.; BRADSHAW, R.E. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. **PLoS Genetics**. v. 8. ed. 11, 2012.
- DELLAPORTA, S.L.; WOOD, J. HICKS, J.B. A plant DNA miniprep: Version II. **Plant Molecular Biology Reporter**. v. 1, ed. 4, p. 19-21, 1983.
- DESLANDES, L.; RIVAS, S. Catch me if you can: Bacterial effectors and plant targets. **Trends Plant Science**. v. 17, p. 644-655, 2012.
- DUPLESSIS, S.; CUOMO, C.A.; LIN, Y.C.; AERTS, A.; TISSERANT, E.; VENEULT-FOURREY, C.; JOLY, D.L.; HACQUARD, S. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proc. Natl. Acad. Sci.** v. 108, p. 9166–9171, 2011.
- DUPLESSIS, S.; LORRAIN, C.; PETRE, B.; FIGUEROA, M.; DODDS, P.N.; AIME, M.C. Host adaptation and virulence in heteroecious rust fungi. **Annual Review of Phytopathology**. v. 59, p.17.1-17.20, 2021.
- EDGAR, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32(5), p. 1792–1797, 2004. Disponível em: <<https://doi.org/10.1093/nar/gkh340>>
- ELLIS, J. G; DODDS, P. N.; LAWRENCE, G. J. Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. **Annual Review of Phytopathology**, v. 45, p. 289-306, 2007.
- ELLIS, J.G.; RAFIQI, M.; GAN, P.; CHAKRABARTI, A.; DODDS, P.N. Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. **Current Opinion in Plant Biology**, v. 12, p. 399–405, 2009.

EMBRAPA-CNPQ. **Árvore do Conhecimento**: Soja. Disponível em: <[https://www.agencia.cnptia.embrapa.br/gestor/soja/arvore/CONTAG01\\_24\\_271020069131.html](https://www.agencia.cnptia.embrapa.br/gestor/soja/arvore/CONTAG01_24_271020069131.html)>. Acesso em: 14 abr. 2019.

FEHR, W.R., CAVINESS, C.E., BURMOOD, D.T., PENNINGTON, S.J. Stage of development descriptions for soybeans, *Glycine Max* (L.) Merrill. **Crop Science**. v. 11, p. 929–931, 1971.

FIGUEIREDO, A.C.S.; BARROSO, J.M.G.; PEDRO, L.M.G.; ASCENSÃO, L. **Histoquímica e Citoquímica em Plantas**: Princípios e Protocolos. 1. ed. Lisboa, 2007.

FINN, R.D.; TATE, J.; MISTRY, J.; COGGILL, P.C.; SAMMUT, S.J.; HOTZ, H.R.; CERIC, G.; FORSLUND, K.; EDDY, S.R.; SONNHAMMER, E.L.; BATEMAN, A. The Pfam protein families database. **Nucleic Acids Research**. v. 36, 2008.

FUNGICIDE RESISTANCE ACTION COMMITTEE, 2018. FRAC Code List. Disponível em: <<http://www.frac.info>>.

GARCIA, A. et al. Molecular mapping of soybean (*Phakopsora pachyrhizi*) resistance genes: discovery of novel locus and alleles. **Theoretical and Applied Genetics**, v.117, p.545-553, 2008.

GAZZONI, D.L.; DALL'AGNOL, A. Cenário: Paralelo entre a soja no mundo e no Brasil. In: **Soja: Quebrando Recordes**. Sorocaba: CESB, 2018. p. 37-40.

GERMAIN, H.; JOLY, D. L.; MIREAULT, C.; PLOURDE, M. B.; LETANNEUR, C.; STEWART, D.; MORENCY, M.J.; PETRE, B.; DUPLESSIS, S.; SÉGUIN, A. Infection assays in *Arabidopsis* reveal candidate effectors from the poplar rust fungus that promote susceptibility to bacteria and oomycete pathogens. **Molecular Plant Pathology**, v. 19, n. 1, p. 191-200, 2018.

GODOY, C. V.; ALMEIDA, A. M. R.; COSTAMILAN, L. M.; MEYER, M.C.; DIAS, W. P.; SEIXAS, C. D. S.; SOARES, R. M.; HENNING, A. A.; YORINORI, J. T.; FERREIRA, L. P.; SILVA, J. F.V.; Doenças da soja. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A.; CAMARGO, L. E. A. (Ed.). **Manual de fitopatologia**: Doenças das Plantas Cultivadas. 5. ed. Ouro Fino: Agronômica Ceres, 2016a. v. 2., p. 657-675.

GODOY, C.V.; MEYER, M.C.; MELLO, F.E.; UTIAMADA, C.M. Resistencia de los hongos a fungicidas en soya. In: NAKASATO, R.; SANDÓVAL, R.P. (Ed.). **Manual De Difusión Técnica de Soya 2017/2018**. Bolivia: Fundacruz, 2017. p. 18-24.

GODOY, C. V., SEIXAS, C. D. S., SOARES, R. M., MARCELINO-GUIMARÃES, F. C., MEYER, M. C., COSTAMILAN, L. M. 2016b. Asian soybean rust in Brazil: Past, present, and future. **Pesquisa Agropecuária Brasileira**. Brasília, v.51, n.5, p. 407-421, 2016b. Disponível em: <<https://doi.org/10.1590/S0100-204X2016000500002>>. Acesso em: 18 jun. 2019.

GOHRE, V.; ROBATZEK, S. Breaking the barriers: microbial effector molecules subvert plant immunity. **Annual Review of Phytopathology**. v. 46, p. 189–215, 2008.

GRAHAM, M.A. Identification and analyses of candidate genes for Rpp4-mediated resistance to Asian soybean rust in soybean. **Plant Physiology**. v. 150, p. 295–307, 2009.

GRELL, M.N.; MOURITZEN, P.; GIESE, H. A *Blumeria graminis* gene family encoding proteins with a C-terminal variable region with homologues in pathogenic fungi. **Gene**. v. 311, p. 181–192, 2003.

GRIFFITHS, G.; SIMONS, K. The trans Golgi network: sorting at the exit site of the Golgi complex. **Science**, v. 234, p. 438-444, 1986.

GRIGORIEV, I.V.; CULLEN, D.; GOODWIN, S.B.; HIBBETT, D.; JEFFRIES, T.W.; KUBICEK, C.P.; KUSKE, C.; MAGNUSON, J.K.; MARTIN, F.; SPATAFORA, J.W.; TSANG, A.; BAKER, S.E. Fueling the future with fungal genomics. **Mycology**. V. 2:3, p. 192-209, 2011. DOI:10.1080/21501203.2011.584577

GRIGORIEV, I.V.; NIKITIN, R.; HARIDAS, S.; KUO, A.; OHM, R.; OTILLAR, R.; RILEY, R.; SALAMOV, A.; ZHAO, X.; KORZENIEWSKI, F.; SMIRNOVA, T.; NORDBERG, H.; DUBCHAK, I.; SHABALOV, I. MycoCosm portal: gearing up for 1000 fungal genomes. **Nucleic Acids Research**. v. 42, p. 699–704, 2014. DOI:10.1093/nar/gkt1183

HARTMAN, G. L.; SIKORA, E. J.; RUPE, J. C. Rust. In: HARTMAN, G. L.; RUPE, J. C.; G. L.; SIKORA, E. J.; DOMIER, L. L.; DAVIS, J. A.; STEFFEY, K. L. (Ed.) **Compendium of soybean diseases**. 5. ed. Saint Paul: APS Press, 2015. p. 56-59.

HENNING, A. A.; ALMEIDA, A. M. R.; GODOY, C. V.; SEIXAS, C. D. S.; YORINORI, J. T.; COSTAMILAN, L. M.; FERREIRA, L. P.; MEYER, M. C.; SOARES, R. M.; DIAS, W. P. **Manual de identificação de doenças de soja**. 5. ed. Londrina: Embrapa Soja, 2014. 75 p. (Embrapa Soja. Documentos, 256).

HUANG, W.; HONG, S.; TANG, G.; LU, Y.; WANG, C. Unveiling the function and regulation control of the DUF3129 family proteins in fungal infection of hosts. **Royal Society**. v. 374, 2019.

HYTEN D.L.; HARTMAN, G.L.; NELSON, R.L.; FREDERICK, R.D.; CONCIBIDO, V.C. Map location of the Rpp1 locus that confers resistance to soybean rust in soybean. **Crop Science**. v. 47, p. 835-838, 2007.

HYTEN D.L.; SMITH, J.R.; FREDERICK, R.D.; TUCKER, M.L. Bulk segregant analysis using the GoldenGate assay to locate the Rpp3 locus that confers resistance to soybean rust in soybean. **Crop Science**. v. 49, p. 265–271, 2009.

IGARASHI, W.T.; FRANÇA, J.A.; SILVA, M.A.A.; IGARASHI, S.; SAAB, O.J.G.A. Application of prediction models of asian soybean rust in two crop seasons, in Londrina, Pr. **Semina: Ciências Agrárias, Londrina**, v. 37, n. 5, p. 2881-2890, 2016.

JONES, J.D.G.; DANGL, J.L. The plant immune system. **Nature**. v. 444, p. 323-329, 2006. Disponível em: < <https://www.nature.com/articles/nature05286>>. Acesso em: 20 jun. 2019.

JONES, J.D.G.; VANCE, R.E.; DANGL, J.L. Intracellular innate immune surveillance devices in plants and animals. **Science**, v. 354, 2016.

JUSTESEN, A.; SOMERVILLE, S.; CHRISTIANSEN, S.; GIESE, H. Isolation and characterization of two novel genes expressed in germinating conidia of the obligate biotroph *Erysiphe graminis* f.sp. *hordei*. **Gene**. v. 170, p. 131-135, 1996.

KAMOUN S. A catalogue of the effector secretome of plant pathogenic oomycetes. **Annual Review of Phytopathology**, v. 44, p. 41-60, 2006.

KEMEN, E.; KEMEN, A. C.; RAFIQI, M.; HEMPEL, U.; MENDGEN, K.; HAHN, M.; VOEGELE, R. T. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. **Molecular Plant-Microbe Interactions**, v. 18, n. 11, p. 1130-1139, 2005.

KILLGORE, E.M. Field notes on the detection of soybean rust. initial surveys and the current status of the disease in Hawaii. In: **Rust Workshop**, 1995, Urbana. Proceedings. Urbana: College of Agricultural, Consumer, and Environmental Sciences: National Soybean Research Laboratory, 1996. p.38-45 (Publication Number 1) Editado por J.B. Sinclair, G. L. Hartman.

KIM K.S.; UNFRIED, J.R.; FREDERICK, R.D.; HARTMAN, G.L.; NELSON, R.L.; SONG, Q.; DIERS, B.W. Molecular mapping of soybean rust resistance in soybean accession PI561356 and SNP haplotype analysis of the Rpp1 region in diverse germplasm. **Theoretical and Applied Genetics**. v. 125, p. 1339-1352, 2012.

KOCH, E.; EBRAHIMNESBAT, F.; HOPPE, H. H. Light and electron microscopic studies on the development of soybean rust (*Phakopsora pachyrhizi* Syd) in susceptible soybean leaves. **Phytopathology**, v. 106, p. 302–320, 1983.

KUNJETI, S.G.; IYER, G.; JOHNSON, E.; LI, E.; BROGLIE, K.E.; RAUSCHER, G.; RAIDAN, G.J. Identification of *Phakopsora pachyrhizi* candidate effectors with virulence activity in a distantly related pathosystem. **Frontiers in Plant Science**, v. 7, p. 269, 2016.

LAZZAROTTO, J. J.; HIRAKURI, M. H. **Evolução e perspectivas de desempenho econômico associadas com a produção de soja nos contextos mundial brasileiro**. Londrina: Embrapa Soja, p. 46, 2010. (Embrapa Soja. Documentos, 319).

LI, Y.; XIA, C.; WANG, M.; YIN, C.; CHEN, X. Whole-genome sequencing of *Puccinia striiformis* f. sp. *tritici* mutant isolates identifies avirulence gene candidates. **BMC Genomics**. v. 21, n. 247, 2020.

LINK, T.I.; LANG, P.; SCHEFFLER, B.E.; DUKE, M.V.; GRAHAM, M.A.; COOPER, B.; TUCKER, M.L.; VAN DE MORTEL, M.; VOEGELE, R.T.; MENDGEN, K.; BAUM, T.J.; WHITHAM, S.A. The haustorial transcriptomes of *Uromyces appendiculatus* and *Phakopsora pachyrhizi* and their candidate effector families. **Molecular Plant Pathology**, v. 15, n. 4, 2014.

LINK, T.; HIRSCHBURGER, D.; MÜLLER, M.; VOEGELE, R.T. Reference genes in the pathosystem *Phakopsora pachyrhizi*/soybean suitable for normalization in transcript profiling. **International Journal of Molecular Sciences**. v. 16, p. 23057-075, 2015.

LIPKA, A.E.; TIAN F.; WANG, Q.; PEIFFER, J.; LI, M.; BRADBURY, P.J.; GORE, M.A.; BUCKLER, E.S.; ZHANG, Z. GAPIT: genome association and prediction integrated tool. **Bioinformatics**. v. 28, n. 18, p. 2397-2399, 2012.

LIU, C.; PEDERSEN, C.; SCHULTZ-LARSEN, T.; AGUILAR, G.B.; MADRIZ-ORDEÑANA, K.; HOVMØLLER, M.S.; THORDAL-CHRISTENSEN, H. The stripe rust fungal effector PEC 6 suppresses pattern-triggered immunity in a host species-independent manner and interacts with adenosine kinases. **New Phytologist**, 2016.

LIU, T.; SONG, T.; ZHANG, X.; YUAN, H.; SU, L.; LI, W.; XU, J.; LIU, S. CHEN, L.; CHEN, T.; ZHANG, M.; GU, L.; ZHANG, B. Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. **Nature Communications**. V. 5, 4686, 2014. Disponível em: <<https://doi.org/10.1038/ncomms5686>>.

LOZANO-TORRES, J.L.; WILBERS, R.H.P.; GAWRONSKI, P.; BOSHOVEN, J.C.; FINKERS-TOMCZAK, A.; CORDEWENER, J.H.G.; AMERICA, A.H.P.; OVERMARS, H.A.; KLOOSTER, J.V.; BARANOWSKI, L.; SOBCZAK, M.; ILYAS, M.; VAN DER HOORN, R.A.L.; SCHOTS, A.; DE WIT, P.J.G.M.; BAKKER, J.; GOVERSE, A.; SMANT, G. Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. **Proceedings of the National Academy of Sciences**. v. 109, p. 10119–24, 2012.

MA, L.S.; HACHANI, A.; LINS, J.S.; FILLOUX, A.; LAI, E.M. *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. **Cell Host Microbe**, v. 16, p. 94-104, 2014.

MA, Z.; ZHU, L.; SONG, T.; WANG, Y.; ZHANG, Q.; XIA, Y.; QIU, M.; LIN, Y.; LI, H.; KONG, L.; FANG, Y.; YE, W.; DONG, S.; ZHENG, X.; TYLER, B.M. A paralogous decoy protects *Phytophthora sojae* apoplastic effector PsXEG1 from a host inhibitor. **Science**. v. 355, p. 710-714, 2017.

MACHEREY-NAGEL. **Genomic DNA from plant**. Disponível em: <[https://www.mnnet.com/Portals/8/attachments/Redakteure\\_Bio/Protocols/Genomic%20DNA/UM\\_gDNAPlant\\_NSPlantII.pdf](https://www.mnnet.com/Portals/8/attachments/Redakteure_Bio/Protocols/Genomic%20DNA/UM_gDNAPlant_NSPlantII.pdf)>. Acesso em: 27 jun. 2019.

MCCANN, H.C.; GUTTMAN, D.S. Evolution of the type III secretion system and its effectors in plant-microbe interactions. **New Phytology**, v. 177, p. 33-47, 2008.

MEDZHITOV, R.; JANEWAY, C.A. Innate immunity: the virtues of a nonclonal system of recognition. **Cell**. v. 91, p. 295-298, 1997.

MENDGEN, K.; HAHN, M. Plant infection and the establishment of fungal biotrophy. **Trends in Plant Science**, v. 7, p. 352-356, 2002.

MINH, B.Q.; SCHMIDT, H.A.; SCHREMPF, O.C.D.; WOODHAMS, M.D.; VON HAESLER, A; LANFEAR, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. **Molecular Biology and Evolution**, v. 37, ed. 5, p. 1530–1534, 2020. Disponível em: <<https://doi.org/10.1093/molbev/msaa015>>. Acesso em: 19 jul. 2021.

- MÖLLER, P.A. **Análise funcional de genes de *Phakopsora pachyrhizi* candidatos a efetores utilizando o sistema EDV**. 2014. 61 p. Dissertação (Mestrado em Fitopatologia) – Universidade Federal de Viçosa, Viçosa, MG.
- MONTEROS, M.J.; MISSAOUI, A.M.; PHILLIPS, D.V.; WALKER, D.R.; BOERMA, R. Mapping and confirmation of the ‘Hyyuga’ redbrown lesion resistance gene for asian soybean rust. **Crop Science**, Madison, v. 47, n. 2, p. 829-834, 2007.
- MULDER, N.J.; APWEILER, R.; ATTWOOD, T.K.; BAIROCH, A.; BATEMAN, A.; BINNS, D.; BORK, P.; BUILLARD, V.; CERUTTI, L.; COPLEY, R.; COURCELLE, E.; DAS, U.; DAUGHERTY, L.; DIBLEY, M.; FINN, R.; FLEISCHMANN, W.; GOUGH, J.; HAFT, D.; HULO, N.; HUNTER, S.; KAHN, D. New developments in the InterPro database. **Nucleic Acids Research**. v. 35, p. 224–228, 2007.
- NGOU, B.P.M.; AHN, H.; DING, P.; JONES, J.D.G. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. **bioRxiv**, 2020.
- PASCHOLATI, S.F.; DALIO, R.J.D. Fisiologia do Parasitismo: Como as plantas se defendem dos patógenos. In: AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A. **Manual de Fitopatologia: Princípios e Conceitos**. 5. ed. Ouro Fino: Agronômica Ceres, 2018. p. 423-452.
- PERSOONS, A.; MAUPETIT, A.; LOUET, C.; ANDRIEUX, A.; LIPZEN, A.; BARRY, K.W.; NA, H.; ADAM, C.; GRIGORIEV, I.V.; SEGURA, V.; DUPLESSIS, S.; FREY, P.; HALKETT, F.; DE MITA, S. Genomic signatures of a major adaptive event in the pathogenic fungus *Melampsora larici-populina*. **bioRxiv**. 2021.
- PETRE, B.; JOLY, D.L.; DUPLESSIS, S. Effector proteins of rust fungi. **Frontiers in Plant Science**. v. 5, p. 416, 2014.
- PETRE, B., SAUNDERS, D. G., SKLENAR, J., LORRAIN, C., WIN, J., DUPLESSIS, S.; KAMOUN, S. Candidate effector proteins of the rust pathogen *Melampsora larici-populina* target diverse plant cell compartments. **Molecular Plant-Microbe Interactions**. v. 28, n. 6, p. 689-700, 2015.
- PFAFFL, M.W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Research**, v. 29, n. 45, 2001.
- PFAFFL, M.W.; HORGAN, G.W.; DEMPFLER, L. Relative expression software tool (REST<sup>®</sup>) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. **Nucleic Acids Research**, v. 30, n. 9, 2002.
- QI, M.; GRAYCZYK, J.P.; DARBEN, L.M.; RIEKER, M.E.G.; SEITZ, J.M.; VOEGELE, R.T.; WHITAM, S.A.; LINK, T.I. Candidate effectors from *Uromyces appendiculatus*, the causal agent of rust on common bean, can be discriminated based on suppression of immune responses. **Frontiers in Plant Science**. v. 10, n. 1182, 2019.
- QI, M.; GRAYCZYK, J. P.; SEITZ, J. M.; LEE, Y.; LINK, T. I.; CHOI, D.; PEDLEY, K.F.; VOEGELE, R.T.; BAUM, T.J.; WHITHAM, S. A. Suppression or activation of immune

responses by predicted secreted proteins of the soybean rust pathogen *Phakopsora pachyrhizi*. **Molecular Plant-Microbe Interactions**. v. 31, n. 1, p. 163-174, 2018.

QI, M.; LINK, T.I.; MÜLLER, M.; HIRSCHBURGER, D.; PUDAKE, R.N.; PEDLEY, K.F.; BRAUN, E.; VOEGELE, R.T.; BAUM, T.J.; WHITHAM, S.A. A Small Cysteine-Rich Protein from the Asian Soybean Rust Fungus, *Phakopsora pachyrhizi*, Suppresses Plant Immunity. **PLoS Pathogens**. v. 12, 2016.

RINCÃO, M.P.; CARVALHO, M.C.D.C.G.; NASCIMENTO, L.C.; LOPES-CAITAR, V.S.; CARVALHO, K.; DARBEN, L.M.; YOKOYAMA, A.; CARAZZOLLE, M.F.; ABDELNOOR, R.V.; MARCELINO-GUIMARÃES, F.C. New insights into *Phakopsora pachyrhizi* infection based on transcriptome analysis in planta. **Genetics and Molecular Biology**. v. 41, p. 671-691, 2018.

RITCHIE, S.; HANWAY, J.J.; THOMPSON, H.E. **How a Soybean Plant Develops**. Special Report. Ames, Iowa State University of Science and Technology. n. 53, p. 20, 1982.

ROJANO, E.; JABATO, F.M.; PERKINS, J.R.; CÓRDOBA-CABALLERO, J.; GARCÍA-CRIADO, F.; SILLITOE, I.; ORENGO, C.; RANEA, J.A.G.; SEOANE-ZONJIC, P. Assigning protein function from domain-function associations using DomFun. **BMC Bioinformatics**. v. 23, 2022.

ROUX, F., VOISIN, D., BADET, T., BALAGUÉ, C., BARLET, X., HUARD-CHAUVEAU, C., ROBY, D.; RAFFAELE, S. Quantitative disease resistance in plants. **Molecular Plant Pathology**, v. 15, p. 427-432. 2014.

RUPE, J.; SCONYERS, L. Soybean Rust. **The Plant Health Instructor**. St. Paul: American Phytopathological Society, 2008. Disponível em: <<https://www.apsnet.org/edcenter/disandpath/fungalbasidio/pdlessons/Pages/SoybeanRust.aspx>>. Acesso em: 18 jun. 2019.

RYTTER, J.L.; DOWLER, W.M.; BROMFIELD, K.R. Additional alternative hosts of *Phakopsora pachyrhizi*, causal agent of soybean rust. **Plant Disease**, v.68, p.818-819, 1984. DOI: 10.1094/ PD-68-818.

SAUNDERS, D.G.O.; WIN, J.; CANO, L.M.; SZABO, L.J.; KAMOUN, S.; RAFFAELE, S. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. **PLoS One**, v. 7, e29847, 2012.

SCHREIBER, K.J.; BENTHAM, A.; WILLIAMS, S. J.; KOBE, B.; STASKAWICZ; B. J. Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. **PLOS Pathogens**, n. 12, 2016.

SEIXAS, C.D.S.; HARGER, N.; OLIVEIRA, F.T.; SERATTO, C.D.; GHELLER, J.A.; OLIVEIRA, A.B. **Monitoramento de *Phakopsora pachyrhizi* na safra 2017/2018 para tomada de decisão do controle da ferrugem-asiática da soja**. Londrina: Embrapa Soja, 2018. 19 p. (Embrapa Soja. Circular Técnica, 141).

SILVA, D.; BROGIN, R.L.; YAMANAKA, N.; ARIAS, C.A. Molecular mapping of two loci that confer resistance to Asian rust in soybean. **Theoretical and Applied Genetics**, v.117, p.57-63, 2008.

SLAMINKO, T.L.; MILES, M.R.; FREDERICK, R.D.; BONDE, M.R.; HARTMAN, G.L. New legume hosts of *Phakopsora pachyrhizi* based on greenhouse evaluations. **Plant Disease**, v.92, p.767-771, 2008. DOI: 10.1094/PDIS-92-5-0767.

SOHN, K.H.; LEI, R.; NEMRI, A.; JONES, J.D.G. The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*. **Plant Cell**, v. 19, n. 12, p. 4077-4090, 2007.

SPERSCHNEIDER, J.; DODDS, P. N.; TAYLOR, J. M.; DUPLESSIS, S. Computational Methods for Predicting Effectors in Rust Pathogens. **Wheat Rust Diseases**, v. 1659, p. 73–83, 2017.

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

TAIZ, L.; ZEIGER, E.; MOLLER, I.M.; MURPHY, A. **Fisiologia e Desenvolvimento Vegetal** [recurso eletrônico] – 6. ed. – Porto Alegre: Artmed, 2017.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular evolutionary genetics analysis version 11. **Molecular Biology and Evolution**, v. 38, ed. 7, p. 3022–3027, 2021. Disponível em: <<https://doi.org/10.1093/molbev/msab120>>.

THERMO FISHER SCIENTIFIC. **Gateway cloning technology**. Disponível em: <[https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FFTFS-Assets%2FFLSG%2Fbrochures%2F710\\_01784\\_Gateway\\_bro.pdf&title=R2F0ZXdheSBQcm9kdWN0IEJyb2NodXJl](https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FFTFS-Assets%2FFLSG%2Fbrochures%2F710_01784_Gateway_bro.pdf&title=R2F0ZXdheSBQcm9kdWN0IEJyb2NodXJl)>. Acesso em: 23 set. 2020.

TSENG, T.T.; TYLER, B.M.; SETUBAL, J.C. Protein secretion systems in bacterial-host associations, and their description in the gene ontology. **BMC Microbiology**, v. 9, n. 2, 2009.

USDA Animal and Plant Health Inspection Service. **Soybean Rust**. Disponível em: <[https://www.aphis.usda.gov/aphis/ourfocus/planthealth/plant-pest-and-disease-programs/pests-and-diseases/plant-disease/sa\\_soy\\_bean\\_rust/ct\\_soybean\\_rust](https://www.aphis.usda.gov/aphis/ourfocus/planthealth/plant-pest-and-disease-programs/pests-and-diseases/plant-disease/sa_soy_bean_rust/ct_soybean_rust)>. Acesso em: 18 jun. 2019.

USDA Foreign Agricultural Service. **Current Soybean Harvest Yields Disappoint, But Swift Recovery Forecast for 2019/20**. Disponível em: <<https://www.fas.usda.gov/regions/brazil>>. Acesso em: 14 abr. 2019.

USDA Foreign Agricultural Service. World Agricultural Production. **Circular Series**. Disponível em: <<https://apps.fas.usda.gov/psdonline/circulars/production.pdf>>. Acesso em: 29 jul. 2020.

- UTIYAMA, A.S. Caracterização funcional de proteínas candidatas a efetores de *Phakopsora pachyrhizi*. 2020. 86 p. Dissertação (Mestrado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, UEL. 2017
- VITTAL, R., YANG, H. C., & HARTMAN, G. L. Anastomosis of germ tubes and migration of nuclei in germ tube networks of the soybean rust pathogen, *Phakopsora pachyrhizi*. **European Journal of Plant Pathology**, n. 132, p. 163–167, 2011.
- XUE, C.; PARK, G.; CHOI, W.; ZHENG, L.; DEAN, R.; XU, JR. Two novel fungal virulence genes specifically expressed in appressoria of the rice blast fungus. **Plant Cell**. v. 14, p. 2107-2119, 2002.
- WANG Y.; ZHANG, H.; ZHONG, H.; XUE, Z. Protein domain identification methods and online resources. **Computational and Structural Biotechnology Journal**. v. 19, p. 1145-1153, 2021.
- YAMANAKA, N.; YAMAOKA, Y.; KATO, M.; LEMOS, N.G.; PASSIANOTTO, A.L.D.L.; SANTOS, J.V.M.; BENITEZ, E.R.; ABDELNOOR, R.V.; SOARES, R.M.; SUENAGA, K. Development of classification criteria for resistance to soybean rust and differences in virulence among Japanese and Brazilian rust populations. **Tropical Plant Pathology**. v. 35, p. 153-162, 2010.
- YOKOYAMA, A. **Caracterização estrutural, diversidade genética e perfil transcricional de candidatos a efetores de *Phakopsora pachyrhizi***. Londrina, PR, 73 p. Dissertação de Mestrado – Mestrado em Biotecnologia. Universidade Estadual de Londrina, 2016
- YORINORI, J.T.; NUNES JUNIOR, J.; LAZAROTTO, J.J. **Ferrugem "asiática" da soja no Brasil**: evolução, importância econômica e controle. Londrina: Embrapa Soja, 2004. 36 p. (Embrapa Soja. Documentos, 247).
- YORINORI, J.T.; COSTAMILAN, L.M.; PAIVA, W.M.; BERTAGNOLLI, P.F. Ferrugem da soja (*Phakopsora pachyrhizi*): identificação e controle. **Informações Agronômicas n° 104**. Disponível em: < [http://www.ipni.net/PUBLICATION/IA-BRASIL.NSF/0/EB6206BA4030ADDA83257AA2005B9BE7/\\$FILE/Page5-8-104.pdf](http://www.ipni.net/PUBLICATION/IA-BRASIL.NSF/0/EB6206BA4030ADDA83257AA2005B9BE7/$FILE/Page5-8-104.pdf)>. Acesso em: 16 jan. 2020.
- YUAN, M., JIANG, Z., BI, G.; NOMURA, K.; LIU, M.; WANG, Y.; CAI, B.; ZHOU, J.; HE, S.Y.; XIN, X. Pattern-recognition receptors are required for NLR-mediated plant immunity. **Nature**, v. 592, p. 105–109, 2021.

## **APÊNDICES**

## APÊNDICE A

Sequências dos primers (*forwad* e *reverse*) dos candidatos a efetores utilizados para análise de expressão gênica por RTqPCR, tamanho aproximado do fragmento amplificado (pb) e sua respectiva temperatura de anelamento (TA).

<b>Efedor</b>	<b>Objetivo</b>	<b>Conjunto de <i>primers</i> (F e R)</b>	<b>Tamanho (pb)</b>	<b>TA (°C)</b>
<b>de_novo_251</b>	RTqPCR	F 5'- TCGGATTTATGGCTCCAACA - 3' F 5'- ACCTCCAGCATCACCCTTCTAC - 3'	100	60
<b>de_novo_2595</b>	RTqPCR	F 5' – CATTGAACATGCCCGGAAAC - 3' F 5' – TAGCTTGTGAGCCGGCCTTA - 3'	90	60
<b>de_novo_635</b>	RTqPCR	F 5'- GCCAGCCCACATAAACTTCA - 3' R 5'- GAGCGATGCTGGAAACATCA - 3'	100	60
<b>de_novo_555</b>	RTqPCR	F 5'- ACAAACGAAACCCATTTCAA - 3' R 5'- TTCCACCAGCCAAGGTTCT - 3'	100	60
<b>de_novo_2238</b>	RTqPCR	F 5' – GCATGCAGTTTGTCTTAGCTC – 3' R 5' – GGTGATAAAGAGATGTTGCGC – 3'	100	60
<b>de_novo_1784</b>	RTqPCR	F 5' – F' CGACGGACAGACTTGTATCGTC – 3' R 5' – TCACCCAAGCCAAGTAATCTTGA – 3'	100	60
<b>Phapa1_6414149</b>	RTqPCR	R 5' – TTGCCAATCGTAGCCAGGTT – 3' R 5' – GTCCTGCTCGTGCTCCATTT – 3'	100	60
<b>Phapa1_7713431</b>	RTqPCR	F 5'- CGACCTCCTTTGGTGTGGAA - 3' R 5'- CATTACCGACTCGTCCAGATG - 3'	100	60

## APÊNDICE B

Modelos gênicos do genoma de *Phakopsora pachyrhizi* MT2006 com similaridades aos candidatos secretados descritos por De Carvalho et al. (2016), contendo o domínio PFAM11327/DUF3129.

MT2006 Protein ID	Tamanho Proteína (aa)	de_novo similar	Identidade (%)	Cobertura (%)	Localização do Gene	Domínio	Nº Cisteína	Peptídeo Sinal	Proteína idêntica
1588919	203	1784	99,5	100	scaffold_2459:59,146-60,611	DUF3129	7	1 a 21	
2385034	248	2227	34,6	84	scaffold_355:230,094-231,733	DUF3129	6	1 a 22	
<b>4578227</b>	247	2595/1303	100	100	scaffold_141:80620-81747	DUF3129	6	1 a 24	
<b>4578411</b>	204	1784	100	99	scaffold_150:359,681-360,738	DUF3129	7	1 a 24	
<b>4578427</b>	202	1784	100	100	scaffold_150:632,138-633,201	DUF3129	7	1 a 24	4635428
<b>4582009</b>	195	2227	100	100	scaffold_338:448,496-449,426	DUF3129	7	1 a 20	4597124
4582668	203	1784	78,8	98	scaffold_380:330,853-331,971	DUF3129	6	1 a 21	
<b>4594856</b>	203	1402/2238	100	100	scaffold_2101:56,247-57,189	DUF3129	7	1 a 23	4594859/7639239/8120199
<b>4594859</b>	203	1402/2238	100	100	scaffold_2101:92,746-93,688	DUF3129	7	1 a 23	4594859/7639239/8120199
4617494	278	2227	23,8	70	scaffold_618:59443-60574	DUF3129	6	1 a 29	
<b>4635428</b>	202	1784	100	100	scaffold_150:229,613-230,883	DUF3129	7	1 a 24	4578427
6412500	349	1402/2238	25	58	scaffold_212:302,681-304,242	DUF3129	7	1 a 19	
<b>6413136</b>	306	251	100	100	scaffold_260:161337-162713	DUF3129	6	1 a 20	6427169
<b>6414149</b>	500	555	52	38,0	scaffold_355:275,535-278,264	DUF3129	6	1 a 24	-
6416525	203	1784	98	100	scaffold_652:121,577-122,638	DUF3129	7	1 a 21	
6435004	258	2227	82,2	99	scaffold_3579:6,950-8,108	DUF3129	8	1 a 19	
<b>6898838</b>	246	635	100	100	scaffold_141:6701-8754	DUF3129	7	-	
7530421	278	2227	23,8	70	scaffold_1149:39389-41086	DUF3129	6	1 a 29	
<b>7639239</b>	203	1402/2238	100	100	scaffold_1491:13,188-14,715	DUF3129	7	1 a 23	4594856/4594859/8120199
7680277	204	1402/2238	100	100	scaffold_1635:147,201-148,725	DUF3129	7	1 a 22	8120199/7639239/4594856/4594859/4592831
<b>7713431</b>	220	1784	38	88,2	scaffold_1763:100,271-101,667	DUF3129	8	-	
<b>8120199</b>	203	1402/2238	100	100	scaffold_4195:30,509-32,034	DUF3129	7	1 a 23	4594856/4594859/7639239

\*Dados em negrito correspondem aos modelos gênicos selecionados para o estudo.

## APÊNDICE C

Alinhamento da sequência do de\_novo\_555 e o seu modelo gênico correspondente  
MT2006\_4582009.

555	1	ATGCAGACCAAAATTCTTCTAGCTCTTTGCCTAGTAGCTATCTCTCAGGT	50
4582009	1	ATGCAGACCAAAATTCTTCTAGCTCTTTGCCTAGTAGCTATCTCTCAGGT	50
555	51	TAATGCGCATGGTGAATTACAGCTGTCCAAGGCTCTAACGGCATGACTG	100
4582009	51	TAATGCGCATGGTGAATTACAGCTGTCCAAGGCTCTAACGGCATGACTG	100
555	101	GAGAGGCTTTGGAGTAGATCAGAGCACTCCTCGTGATGGTACAAAACGA	150
4582009	101	GAGAGGCTTTGGAGTAGATCAGAGCACTCCTCGTGATGGTACAAAACGA	150
555	151	AACCCATTTCAAACCGACTCTTCAATTATTCGTGATCGTGAGATTGCTAG	200
4582009	151	AACCCATTTCAAACCGACTCTTCAATTATTCGTGATCGTGAGATTGCTAG	200
555	201	CGCAAGTCATCAGCCTGCGGCAGAACCTTGGTGGTGGAAACAACGAGA	250
4582009	201	CGCAAGTCATCAGCCTGCGGCAGAACCTTGGTGGTGGAAACAACGAGA	250
555	251	TTGGAGCAGCAATGTCTAAGGCCGAGTCTGCTGGTATCCCAAGTGTTAGC	300
4582009	251	TTGGAGCAGCAATGTCTAAGGCCGAGTCTGCTGGTATCCCAAGTGTTAGC	300
555	301	TCCGATGAAAGGTTTCAGATGACTTTACATCAAGTTAACGGTGACGGTGG	350
4582009	301	TCCGATGAAAGGTTTCAGATGACTTTACATCAAGTTAACGGTGACGGTGG	350
555	351	GGGCCCTTACACCTGTGATGTTAATGCTTCCGGTGATGGAAAGACTTTCA	400
		.	
4582009	351	TGGCCCTTACACCTGTGATGTTAATGCTTCCGGTGATGGAAAGACTTTCA	400
555	401	CACCTATGACAATTTCTACTAACATTCGGGAAAGAACAGCAGATCTAAT	450
4582009	401	CACCTATGACAATTTCTACTAACATTCGGGAAAGAACAGCAGATCTAAT	450
555	451	GCCAAGGCTGAGGATATTCCTTTAATGCGGAGATGCCCAAAGGAATGAC	500
4582009	451	GCCAAGGCTGAGGATATTCCTTTAATGCGGAGATGCCCAAAGGAATGAC	500
555	501	TTGCACCGGTGGATCTGACGGACAGACCTGTATCGTAAGATGCAGAAACG	550
4582009	501	TTGCACCGGTGGATCTGACGGACAGACCTGTATCGTAAGATGCAGAAACG	550
555	551	CTGCCAACGCTGGTCCCTTTGGAGGATGTGTAGCTGTTACTCAGGCTACG	600
4582009	551	CTGCCAACGCTGGTCCCTTTGGAGGATGTGTAGCT-----	585
555	601	GATGCTACAAAGCAGAAGCGGGAACATCAGCAGAACAGAGAATTGATGA	650
4582009	586	-----	585
555	651	GATGATCAAACAGGCTGAGGCTGCCGAGGCATTTGGCGATGAGACTCTTG	700
4582009	586	-----	585
555	701	CAAAGCGTGAAATGTCTGCTGACGAGAGATTGGACGAAATGATCAGGCAG	750
4582009	586	-----	585
555	751	GCCGAAGCCGAGAGGCGTTCGGCGATGAGACCTTGCAAAGCGCGAGCT	800
4582009	586	-----	585
555	801	TTCTGCTGACCAGCGATTAGACGAAATGATTAACAGGATGAAGATGCTG	850
4582009	586	-----	585
555	851	AGCTGGCTGAGGAAAACCTTTCCAAGCGGAGTTGACTGCTGACCAGCGA	900
4582009	586	-----	585
555	901	TTGGATGAAATGATCAAACAGGCCGAAAGCGCTGAAATGACAGAGGAGAA	950
4582009	586	-----	585
555	951	CTTGGTTAAACGGGAATTGACTGCCGACCAAAGACTAGACGAGATGATCA	1000

4582009	586 -----	585
555	1001 AGCAAGCCGAAGCCGCTGAGCTTGCAGAGGAGAAGCTTGGCCAAGCGTGAG	1050
4582009	586 -----	585
555	1051 CTGACTGCTGACCAAAGGCTGGACCAGATGATCAAACAAGCTGAAGATGC	1100
4582009	586 -----	585
555	1101 TGAGCTGGCTGAGGAAAGCCTTTCCAAGCGGAGTTAACTGCCGATCAGC	1150
4582009	586 -----	585
555	1151 GATTGGATGAAATGATCAAACAAGCCGAAGCCGCTGAAATTGCAGAGGAG	1200
4582009	586 -----	585
555	1201 AACTTGGCTAAGCGGAATTGACTGCCGGCCAAAGACTGGACGAGATGAT	1250
4582009	586 -----	585
555	1251 TAAGCAAGCTGAAGCTGCTGAGCTTGCCGAAGAGTTGATCTAA	1293
4582009	586 -----	585

## APÊNDICE D

Alinhamento da sequência do de\_novo\_635 e o scaffold expandido do seu modelo gênico correspondente no genoma MT2006. Estão destacadas as bases em que ocorre o possível erro de montagem.

635	1	-----	0
SCAFFOLD_141	1	AATCAAGAACGTGAACATGCCTCACCCCTGTGGATATCATTATCCTCT	50
635	1	-----	0
SCAFFOLD_141	51	CTGTACTGCTTCAAGTTATACCTTTGGTTTGGGGGCACTGTAAGATCGTT	100
635	1	-----	0
SCAFFOLD_141	101	ACTGCTGCCGAAATTTAAATACTTCAAGTCAAGTCAGCTATGGATTGG	150
635	1	-----	0
SCAFFOLD_141	151	TGTGGTCAGTTTTTTGAATAGGCTACCCATAACCTCAACAACATCGTTAG	200
635	1	-----ATGGATTGGTG-	12
SCAFFOLD_141	201	TCTGATATTCGGTCTGATCTTACAAAAATTACACTTTGGCTCTGTTC	250
635	13	TGGATTACATAGCAAGTACCCGTGGTTTAAAGTCTCAAGCCGGTGAT--	60
SCAFFOLD_141	251	AGGATTACATAGCAAGTACCCGTGG-TTAAAGTCTCAAGCCGGTGATGT	299
635	61	-----	60
SCAFFOLD_141	300	CAGCTCCAGTCTATCTAAATTTTATTGCTTATTTAGTAATTTTGACCAAG	349
635	61	-----GCCGGAGCGGATTCTGAAG	79
SCAFFOLD_141	350	AACCTGTTACGCTTATGTGACTAATTAAGCCGGAGCGGATTCTGAAG	399
635	80	TTTTTACTGAATCGAAAGAGTTTGTGCGCAATCCTAATCCCCCTTGTGGC	129
SCAFFOLD_141	400	TTTTTACTGAATCGAAAGAGTTTGTGCGCAATCCTAATCCCCCTTGTGGC	449
635	130	ATGCGTGCGAAAATGGGTGCCCTAGATTTTCGACTCATCTTTTAGTCA---	176
SCAFFOLD_141	450	ATGCGTGCGAAAATGGGTGCCCTAGATTTTCGACTCATCTTTTAGTCAAGG	499
635	177	-----	176
SCAFFOLD_141	500	TTTGTAGCCATGTTCTCTTAATTTTATTTTATTCTTACAAAGTATTT	549
635	177	-----A	177
SCAFFOLD_141	550	ATAGAACTTTAAATCACTTAAATGGGATTTTTAAAAAACTGATAATTA	599
635	178	GCCGAAGCAATGGGCGTAGGAAATACTTTAGAAGATGGGTCTTTTGAGGC	227
SCAFFOLD_141	600	GCCGAAGCAATGGGCGTAGGAAATACTTTAGAAGATGGGTCTTTTGAGGC	649
635	228	ATTGATTTTTCAGGTCAACCGTATGGTGGAGGACAATGCACCTGTGAAT	277
SCAFFOLD_141	650	ATTGATTTTTCAGGTCAACCGTATGGTGGAGGACAATGCACCTGTGAAT	699
635	278	ACAACACT-----	285
SCAFFOLD_141	700	ACAACACTGTCAGTAGTTATAGCCCTTACGCTATGTTTTTAGACACATCT	749
635	286	-----GTAGGACAACCTGACAAATTTAAA	309
SCAFFOLD_141	750	AATAGCTCTTTTATCTCTGTCTAGGTAGGACAACCTGACAAATTTAAA	799
635	310	TTTTGCAAGACTCATTAATCCACCGGACAAAATGGCATTGGCCTCA	359
SCAFFOLD_141	800	TTTTGCAAGACTCATTAATCCACCGGACAAAATGGCATTGGCCTCA	849
635	360	AGATCGTGTAATCATACGGCCAAATTTCAACTACCC-----	396
SCAFFOLD_141	850	AGATCGTGTAATCATACGGCCAAATTTCAACTACCCAGGTGCGTTTTTC	899
635	397	-----	396
SCAFFOLD_141	900	GAAATGAAATTTTATCATCTTTTGTAGAATCAAACACATAGTATAATT	949
635	397	-----AGAGATACTACTTGCCTGGTGGAAATGTTCAAAG	430
SCAFFOLD_141	950	TATATCATTTACTTAGAGATACTACTTGCCTGGTGGAAATGTTCAAAG	999

```

635          431 ACAAGTGCCTTGATTAGAATCAGATGT----- 456
          |||
SCAFFOLD_141 1000 ACAAGTGCCTTGATTAGAATCAGATGTGGTAGGCACTCCATCTAGCACCTC 1049
635          457 ----- 456
SCAFFOLD_141 1050 TTGCAAAATATAATGGCTGATAACAAACCTTTTGAATATTCGGTATCTAAC 1099
635          457 -GGAGAGTTTCTAAGATTTGGAGGATGTTTGGCAATAAAGACGCCTGCCA 505
          |||
SCAFFOLD_141 1100 AGGAGAGTTTCTAAGATTTGGAGGATGTTTGGCAATAAAGACGCCTGCCA 1149
635          506 GCCCACATAAACTTCAATTGAAGGTTGTAATTGGAGGTAAAACATCAATG 555
          |||
SCAFFOLD_141 1150 GCCCACATAAACTTCAATTGAAGGTTGTAATTGGAGGTAAAACATCAATG 1199
635          556 AAAACCAAACCG----- 567
          |||
SCAFFOLD_141 1200 AAAACCAAACCGGTTAGGTTTTCATTATATTC AATTAG AATATAAAACAA 1249
635          568 ----- 567
SCAFFOLD_141 1250 CCGTCATACTAATCTACTATTTTTTATTTGAATGATGATACTGTGCGAAC 1299
635          568 -----GACTTACCAAGAAATGATGTTTCCAGCATCGCTCAAAGGTTT 610
          |||
SCAFFOLD_141 1300 CGCTAAAGACTTACCAAGAAATGATGTTTCCAGCATCGCTCAAAGGTTT 1349
635          611 TCAACACACTGAAAGCCAAAGGGCCCTTCGTGCCGATTTCTTCCGGCAGA 660
          |||
SCAFFOLD_141 1350 TCAACACACTGAAAGCCAAAGGGCCCTTCGTGCCGATTTCTTCCGGCAGA 1399
635          661 TATAAAGAAGTCTTCAATGCATATTCATGCCTTG CAGCATGTAAGAAC 710
          |||
SCAFFOLD_141 1400 TATAAAGAAGTCTTCAATGCATATTCATGCCTTG CAGCATGTAAGAAC 1449
635          711 CCTTGGCGTTACCCCTTTAAAATCAAACCTTGACTCTATGGCAGATGAAA 760
          |||
SCAFFOLD_141 1450 CCTTGGCGTTACCCCTTTAAAATCAAACCTTGACTCTATGGCAGATGAAA 1499
635          761 TTTCTTTCAAGGTGATTGAATTGATGCGAGCCAAC AAGGTTTTCCTGGTA 810
          |||
SCAFFOLD_141 1500 TTTCTTTCAAGGTGATTGAATTGATGCGAGCCAAC AAGGTTTTCCTGGTA 1549
635          811 GCGAAGAAGGCCAAATCAATATTGAAGGATTATAGGGACTCTACATCTGC 860
          |||
SCAFFOLD_141 1550 GCGAAGAAGGCCAAATCAATATTGAAGGATTATAGGGACTCTACATCTGC 1599
635          861 AGAAAGATA----- 870
          |||
SCAFFOLD_141 1600 AGAAAGATAAAAAAAAAATCTTCTTGCTCGCTTTT TTTATTTTATTTT 1649
635          871 ----- 870
SCAFFOLD_141 1650 GATGATTGAAAAAATTTTCGATGCGTTTTTCTAGAAAAATTTTGTGTA 1699
635          871 ----- 870
SCAFFOLD_141 1700 CTCTTTCAGCAGAATTATGTATGTGCTGTTTGAAAGTCTAATTTGTTTT 1749
635          871 ----- 870
SCAFFOLD_141 1750 TACTATTTTCAAATCAAACAGGGCACTGGTTGATCTGCTTAAGAGTTAT 1799
635          871 ----- 870
SCAFFOLD_141 1800 CTCTATTTAAAGATACAAATAGATTATTA ACTCTTCCAATAAAATTTCTT 1849
635          871 ----- 870
SCAFFOLD_141 1850 TCTCCATGCTTGCTTTACATCGCAATACAGTCATTTATTTTAATGTTTT 1899
635          871 ----- 870
SCAFFOLD_141 1900 GAAATGGATGCAGATAATCCAGAAACCTTGT CATCAATATCTTTGTAT 1949
635          871 ----- 870
SCAFFOLD_141 1950 TTTTGAATGGATTTTCCAAAAATCAACTGAGCCGTAGTCGCGTATGTT 1999
635          871 ----- 870
SCAFFOLD_141 2000 TTTTAAAGGGGCTATCATGTAAGAACTCAGAGATATCTTTTATTGCA 2049
635          871 ----- 870
SCAFFOLD_141 2050 ACAGG 2054

```

## APÊNDICE E

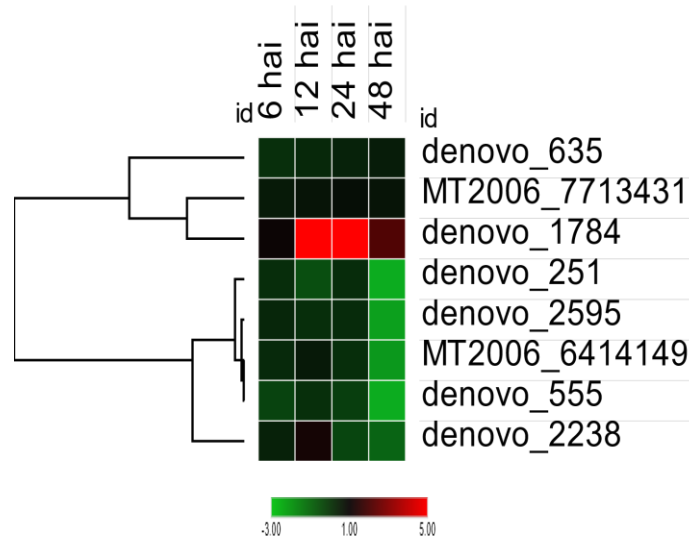
Folhas de soja inoculadas com esporos de *Phakopsora pachyrhizi* e lesões do tipo “TAN”, características de genótipos suscetíveis a este fungo, utilizadas para o estudo da expressão dos candidatos a efetores.



**Fonte:** Castanho, 2020. Embrapa Soja, Londrina/PR.

## APÊNDICE F

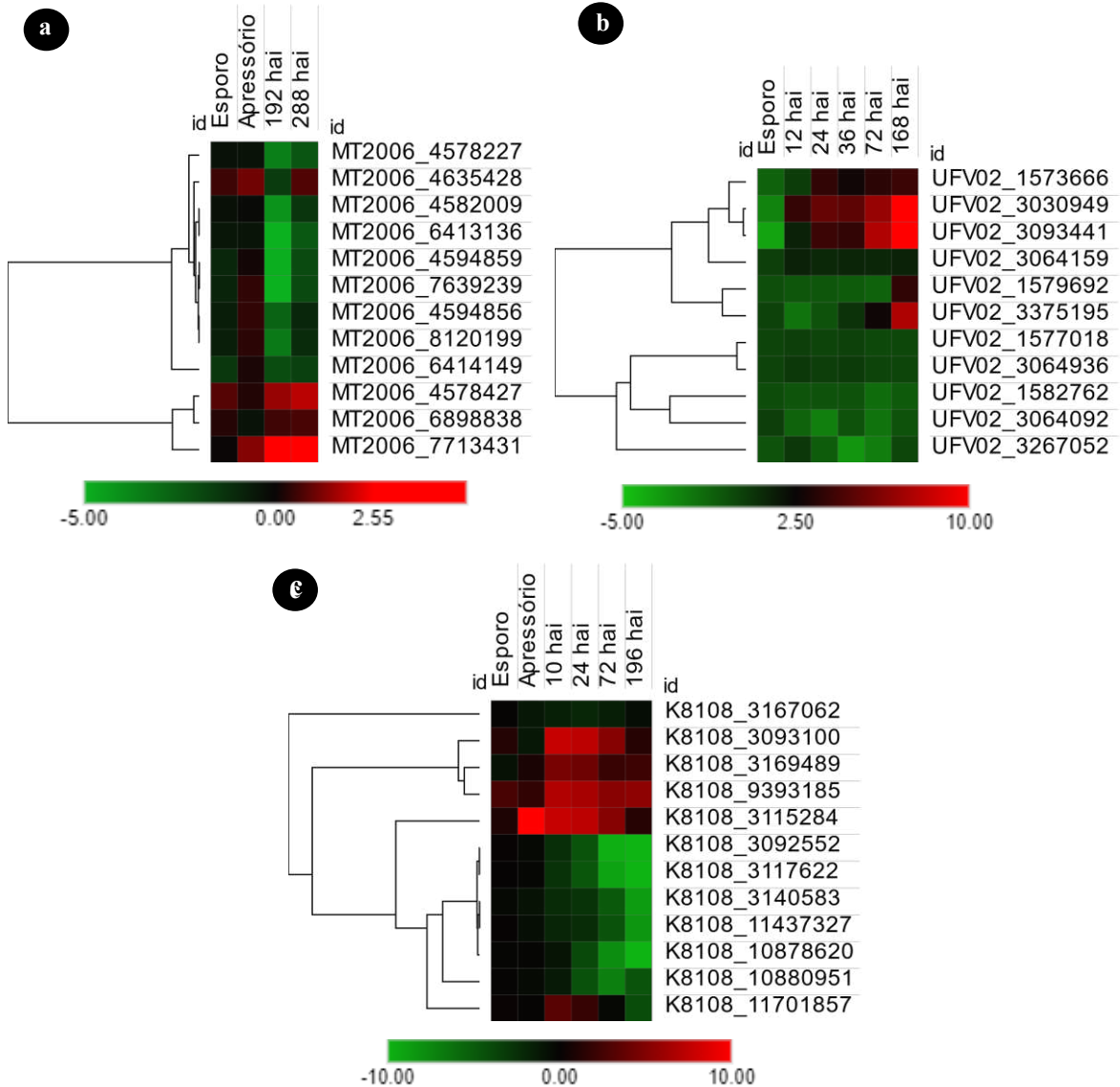
Análise de expressão dos oito candidatos a efetores de *Phakopsora pachyrhizi* durante a interação com a soja. O *heat map* mostra os níveis de expressão gênica dos candidatos em folhas de soja infectadas e coletadas em 6, 12, 24 e 48 horas após inoculação (hpi). O mapa foi gerado através da transformação log dos dados obtidos por RT-qPCR e que são apresentados como expressão relativa ao gene endógeno *CytB*.



Fonte: Morpheus, 2021.

## APÊNDICE G

Expressão gênica obtida por RNA-Seq dos modelos gênicos correspondentes aos candidatos a efetores de *Phakopsora pachyrhizi* durante a interação com a soja e na ausência dela. (a) Perfil de expressão dos genes do genoma MT2006 durante 192 e 288 hpi e na ausência do hospedeiro (esporo e apressório). (b) Perfil de expressão dos genes oriundos do isolado UFV02 no esporo e durante 12, 24, 36, 72 e 168 hpi. (c) Perfil de expressão dos genes oriundos do isolado K8108 no esporo e apressório durante 10, 24, 72 e 196 hpi. Vermelho indica altos níveis de expressão, enquanto verde indica uma baixa indução gênica.



Fonte: Morpheus, 2021.