

Identificação e caracterização de genes de patogenicidade do fungo *Colletotrichum kahawae*, agente causal da antracnose dos frutos verdes do cafeeiro

Jessica Pereira Carvalho

Dissertação para obtenção do Grau de Mestre em
Engenharia Agronómica, especialização em Proteção de Plantas

Orientadoras: Doutora Ana Cristina Garcia Pereira Cabral;

Doutora Helena Manuela Gil Azinheira.

Júri:

Presidente: Doutora Elisabete Tavares Lacerda de Figueiredo Oliveira, Professora auxiliar do Instituto Superior de Agronomia da Universidade de Lisboa.

Vogais: Doutora Luísa Cristina dos Mártires Ferreira de Carvalho, Professora auxiliar do Instituto Superior de Agronomia da Universidade de Lisboa;

Doutora Ana Cristina Garcia Pereira Cabral, Investigadora Júnior do Instituto Superior de Agronomia da Universidade de Lisboa, orientadora.

Abstract

Colletotrichum kahawae is a specialized hemibiotrophic plant pathogen, responsible for the Coffee Berry Disease (*Antracnose dos frutos verdes do cafeeiro*) and it presents itself as one of the biggest threats to the production of Arabica coffee in Africa. This disease can account for a loss of up to 80% of production if no measures are taken, in years of strong epidemics. The main goal of this work is to study the interaction between *Coffea arabica* – *C. kahawae*, with particular interest in the infection process and the genes involved in the occurrence of a successful infection, mainly during the first biotrophic steps and the switch to necrotrophy. The results obtained through bioinformatics analysis of the proteome, generated from the genome sequencing, of *C. kahawae* and through the analysis of the expression profiles of seven transcripts potentially involved in pathogenicity and in the establishment of infection, showed that cutinases are essentially involved at the beginning of the infection, probably resulting in the penetration of the host tissues. Pectate lyases and lysine domain proteins (LysM) probably play an important role during the switch to necrotrophy and the colonization of host tissues during the necrotrophic phase. The ceramidases and at least one of the peroxidases studied revealed activity during the colonization of the host tissues throughout the infection process. This study offers a deeper understanding of the mechanisms involved in the interaction between *Coffea arabica* - *C. kahawae*, which contributes to developing coffee culture protection strategies based on this knowledge of host resistance and the pathogenicity of the pathogen.

Key words

Pathogenicity genes, *Colletotrichum kahawae*, effectors, bioinformatics, RT-qPCR.

Resumo

Colletotrichum kahawae é um fungo hemibiotrófico especializado, responsável pela antracnose dos frutos verdes do cafeeiro, doença mais conhecida como CBD (do inglês *Coffee Berry Disease*). Esta doença apresenta-se como a maior ameaça para a produção de *Coffea arabica* no continente africano, podendo levar a perdas de 80% de produção quando nenhuma medida é tomada e em anos de fortes epidemias. Este trabalho foca-se essencialmente na interação *Coffea arabica* – *C. kahawae*, com particular interesse nas etapas mais importantes do processo de infeção e os genes possivelmente envolvidos no estabelecimento da doença (fase biotrófica) e no *switch* para a necrotrofia. Os resultados obtidos através da análise bioinformática do proteoma, previsto a partir da sequenciação do genoma, de *C. kahawae* e através da análise dos perfis de expressão de sete transcritos potencialmente envolvidos na patogenicidade e no estabelecimento da infeção, mostraram que as cutinases estão essencialmente envolvidas no início da infeção, provavelmente relacionadas com a penetração dos tecidos do hospedeiro, as pectato liases e as proteínas com domínios de lisina (LysM) provavelmente têm alguma função durante o *switch* para a necrotrofia e a colonização dos tecidos do hospedeiro durante a fase necrotrófica. As ceramidases e pelo menos uma peroxidase estudada revelam também ação na colonização dos tecidos do hospedeiro ao longo de todo o processo de infeção. Em conjunto, os resultados permitem aumentar o conhecimento sobre os mecanismos da interação *Coffea arabica* - *C. kahawae* contribuindo para que as estratégias de proteção da cultura do café se baseiem no conhecimento da resistência do hospedeiro e da patogenicidade do agente patogénico.

Palavras-chave

Genes de patogenicidade, *Colletotrichum kahawae*, Efetores, Bioinformática, RT-qPCR.

Agradecimentos

Ao dar por concluído este trabalho e esta etapa do meu percurso acadêmico e pessoal, não posso deixar de agradecer a todos aqueles que me apoiaram e acompanharam durante estes anos.

A ambas as minhas orientadoras, à doutora Ana Cristina Garcia Pereira Cabral e à doutora Helena Manuela Gil Azinheira, dirijo o meu sincero agradecimento por toda a paciência para as minhas crises existenciais e por toda a ajuda que me prestaram durante este longo percurso. Foram incansáveis e extremamente compreensíveis no sentido de se adaptar às mudanças exigidas por este ano atípico e também às mudanças que ocorreram na minha vida.

Um agradecimento muito especial a toda a minha família que ouviu todas as lamentações e inseguranças e que me deram força para enfrentar as adversidades encontradas. Um obrigado enorme à minha mãe que me ajudou a ser quem sou e que meteu o mundo ao meu dispor. Um obrigada ao meu padrasto por ser o meu pai de coração e por ter a paciência de me ir buscar à estação sempre que eu chegava cansada da faculdade. Um obrigada à minha irmã por ter sido sempre a minha melhor amiga e por estar lá sempre que eu precisava. E não podia faltar um obrigada aos meus queridos avós.

Aos meus amigos agradeço todos os incentivos, todos os abraços apertados (ainda que às vezes de longe) e toda a força que me deram. Obrigada por serem os meus pilares e por todos os momentos que passamos juntos. Sei que estas amizades levo para a vida.

Por fim não podia faltar um obrigada gigante ao meu namorado que nunca hesita em mandar uma piada para me fazer sorrir quando tudo parece estar mal.

A todas estas pessoas um obrigada enorme e de coração, porque sem todos vocês estes anos universitários não teriam sido o que foram e a conclusão desta dissertação não teria sido possível.

Índice	
Abstract	I
Resumo	II
Agradecimentos	III
Índice de Figuras	V
Índice de Anexos	VIII
I. Introdução	1
II. Revisão Bibliográfica	2
a. <i>Coffea arabica</i>	2
b. Antracnose dos frutos verdes do cafeeiro	3
c. Interação hospedeiro-agente patogénico	6
III. Materiais e métodos	10
a. Material biológico	10
b. Inoculação hipocótilos e recolha de amostras para estudos de expressão genética	10
c. Bioinformática	11
d. Extração de RNA e síntese de cDNA	14
e. PCR quantitativo	14
IV. Resultados e discussão	17
a. Secretoma e efetoroma de <i>Colletotrichum kahawae</i>	17
i. <i>Enzimas ativas de hidratos de carbono</i>	19
ii. <i>Proteases</i>	22
iii. <i>Metabólitos secundários</i>	24
iv. <i>Homologia com diferentes géneros</i>	27
v. <i>Homologia com proteínas da base de dados PHI</i>	27
vi. <i>Seleção das proteínas a estudar</i>	28
b. PCR quantitativo	30
i. <i>Cutinases</i>	30

ii.	Pectato liases	32
iii.	Ceramidasas	34
iv.	Peroxidasas	36
v.	Proteínas contendo motivos de lisina (LysM)	38
V.	Conclusão	41
VI.	Referências bibliográficas	A
VII.	Anexos	R

Índice de Figuras

Figura 1 - Sintomas da antracnose dos frutos verdes do cafeeiro. Fonte: Bedimo <i>et al.</i> , 2010.	4
Figura 2 – Estados fenológicos do desenvolvimento dos frutos do cafeeiro, associadas às fases de suscetibilidade à doença. Adaptado de Macedo, 1994.....	5
Figura 3 - Representação esquemática do ciclo da doença antracnose dos frutos verdes do cafeeiro. Adaptada de Diniz <i>et al.</i> , 2017; Loureiro <i>et al.</i> , 2012; Silva <i>et al.</i> , 2006 e Vieira, 2018.	6
Figura 4 – Esquema explicativo das ferramentas bioinformáticas utilizadas na análise <i>in silico</i>	13
Figura 5 – Diagrama de Euler das proteínas previstas como apresentando sinais peptídicos e que não contêm domínios transmembranares de acordo com as predições efetuadas pelos programas Phobius, SignalP 5.0, TargetP 0.2 e TMHMM Server v2.0 num total de 19181 proteínas analisadas.	17
Figura 6 – Diagrama de Euler das proteínas previstas como excretadas para o espaço extracelular de acordo com as predições efetuadas pelos programas Wolf PSORT, DeepLoc-1.0 e ProtComp 9.0.	18
Figura 7 – Percentagem das diferentes classes de CAZymes previstas no proteoma de <i>Colletotrichum kahawae</i> (círculo exterior), no secretoma (círculo do meio) e para os possíveis efetores presentes no secretoma (círculo interior). Classes de CAZymes: glicosiltransferases (GT); esterases auxiliares (AA); esterases de hidratos de carbono (CE); Glicosil hidrolases (GH); polissacarídeos liases (PL); módulos de ligação a glícidos (CBM).....	20

Figura 8 – Representatividade das diversas famílias de <i>CAZymes</i> , e respectivas classes, no proteoma de <i>Colletotrichum kahawae</i> . Glicosiltransferases (GT); esterases auxiliares (AA); esterases de hidratos de carbono (CE); Glicosil hidrolases (GH); polissacarídeos liases (PL); módulos de ligação a glícidos (CBM).	22
Figura 9 - Percentagem das diferentes famílias de proteases do total de proteínas presentes no proteoma de <i>Colletotrichum kahawae</i> (círculo exterior), no secretoma (círculo do meio) e para os possíveis efetores presentes no secretoma (círculo interior). Família das proteases aspárticas (A), cisteínicas (C), glutâmicas (G), metaloproteases (M), asparágicas (N), mistas (P); serínicas (S); treónicas (T), inibidoras (I) e desconhecidas (U).....	24
Figura 10 - Percentagem das diferentes famílias de <i>backbone enzymes</i> identificadas no proteoma de <i>Colletotrichum kahawae</i> , de acordo com a previsão efetuada pelo programa SMURF. Família das prenilttransferases (DMAT), das sintases peptídicas não-ribossomais (NRPS), das policetido sintases (PKS) e híbridos entre NRPS e PKS (Híbridos).	25
Figura 11 - Percentagem das diferentes famílias de <i>backbone enzymes</i> presentes no proteoma de <i>Colletotrichum kahawae</i> de acordo com a previsão do programa antiSMASH 5.0. Família das Indole alcaloides (Indole), das sintases peptídicas não-ribossomais (NRPS), das policetido sintases (PKS); híbridos entre NRPS e PKS (Híbrido), das terpeno ciclases (TC) e Fungal-RiPP e betalactona (Outro).....	26
Figura 12 – Representatividade da homologia das proteínas em estudo com alguns géneros importantes.	27
Figura 13 – Fluxograma representativo da sequência de seleção das proteínas potencialmente interessantes para serem estudadas.....	29
Figura 14 – Perfil de expressão génica obtido do gene CKA01_04076 para o isolado Que2 de <i>Colletotrichum kahawae</i> obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com <i>Coffea arabica</i> , e em esporos, em apressórios <i>in vivo</i> (APV) e <i>in vitro</i> (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste <i>t-student</i>) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).	31
Figura 15 – Perfil de expressão génica obtido do gene CKA01_09212 para o isolado Que2 de <i>Colletotrichum kahawae</i> obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com <i>Coffea arabica</i> , e em esporos, em apressórios <i>in vivo</i> (APV) e <i>in vitro</i> (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620,	

utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).33

Figura 16 – Perfil de expressão génica obtido para o gene CKA01_12047 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dai) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).34

Figura 17 – Esquema representativo da reação de hidrólise da ceramida em esfingosina, por ação da enzima ceramidase.....35

Figura 18 – Perfil de expressão génica obtido do gene CKA01_09762 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).36

Figura 19 – Perfil de expressão génica obtido do gene CKA01_11318 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).37

Figura 20 – Perfil de expressão génica obtido do gene CKA01_15250 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio

padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controle (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).39

Figura 21 – Perfil de expressão gênica obtido do gene CKA01_15357 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infecção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controle. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controle (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).40

Índice de Anexos

Anexo 1 - Localização das proteínas segundo Wolf PSORT.....	R
Anexo 2 - Localização das proteínas segundo DeepLoc-1.0.....	R
Anexo 3 - Localização das proteínas segundo o ProtComp 9.0, com base na previsão “Neural Nets”.....	S
Anexo 4 – Comparação da representatividade das diferentes classes de CAZymes em diversos genomas do género <i>Colletotrichum</i> (Baroncelli <i>et al.</i> , 2016; Rao & Nandineni, 2017).	S
Anexo 5 - Comparação da representatividade das diferentes famílias de proteases em diversos genomas do género <i>Colletotrichum</i> (Baroncelli <i>et al.</i> , 2016; Rao & Nandineni, 2017).	T
Anexo 6 - Primers dos genes escolhidos, desenhados com recurso aos PerlPrimer v1.1.21, e dos genes de referência (*) (Vieira <i>et al.</i> , 2016) com as devidas temperaturas de emparelhamento, tamanho dos amplicões e eficiência da reação de qPCR.....	U
Anexo 7 - Primers dos restantes genes escolhidos, desenhados com recurso aos PerlPrimer v1.1.21, com as devidas temperaturas de emparelhamento e a razão pela qual não foi possível obter resultados conclusivos.	V
Anexo 8 - Exemplo de eletroforese em gel de agarose a 1,2% em 1xtampão Tris-Acetato-EDTA para verificação da qualidade do RNA extraído; o RNA foi visualizado após coloração com GreenSafe Premium (NZYTech) sob luz ultravioleta (A); Curva de dissociação (esquerda) e eletroforese em gel de agarose a 2% em 0,5xtampão Tris-Borato-EDTA (direita) de alguns	

dos produtos de RT-PCR amplificados para os diferentes genes estudados; o DNA foi visualizado por coloração com GreenSafe Premium (NZYTech) sob luz ultravioleta; marcador molecular 100 bp Plus (ThermoScientific): CKA01_04076 (B); CKA01_09212 (C); CKA01_09762 (D); CKA01_11318 (E); CKA01_12047 (F); CKA01_15250 (G); CKA01_15357 (H)..... W

Anexo 9 - Resultados da expressão génica, relativa ao controlo (Teste *t-student*), obtidos através do estudo da expressão realizados no programa Bio-Rad CFX Maestro 2.0 v.5.0.021.0616, na função “Gene Study” X

I. Introdução

A produção de café é uma atividade de elevado interesse económico a nível mundial, sendo uma das principais receitas dos países em desenvolvimento (António, 2008). Como em todas as culturas agrícolas, o aparecimento de doenças prejudica gravemente os níveis de produção e rendimentos económicos da atividade agrícola.

O fungo *Colletotrichum kahawae* é o agente causal de antracnose dos frutos verdes do cafeeiro (*Coffee Berry Disease* na versão em inglês, CBD). Este agente patogénico emprega uma estratégia de infeção hemibiotrófica, ou seja, num primeiro estágio apresenta-se como um fungo biotrófico e em média, entre as 48 e as 72h, passa a necrotrófico. Esta doença causa grandes perdas, uma vez que, ao infetar os frutos verdes, impede o desenvolvimento da semente (Batista *et al.*, 2017; van der Vossen & Walyaro, 2009). A CBD é considerada a maior ameaça à produção de café arábica em África, levando a perdas de produção entre os 50-80% em anos de epidemias severas e se nenhuma medida de controlo for aplicada. Até agora, a CBD está restrita a África, mas o risco de sua introdução na América e na Ásia representa uma séria ameaça à cafeicultura nestas regiões e faz com que o *C. kahawae* seja um organismo de quarentena (Silva *et al.*, 2006; van der Vossen & Walyaro, 2009).

Neste trabalho pretende-se estudar a interação *Coffea arabica* - *Colletotrichum kahawae*, com particular ênfase nas principais etapas do processo de infeção e nos genes que determinam a transição para a necrotrofia e para o estabelecimento da doença. Para tal foram utilizados hipocótilos de cafeeiro para a inoculação com suspensões conidiais do fungo, permitindo a recolha de amostras para estudos de expressão genética ao longo do processo de infeção.

Apesar do pouco conhecimento existente sobre os mecanismos moleculares que regulam o estabelecimento desta interação, existem alguns trabalhos com outras interações com fungos do género *Colletotrichum* que referem alguns genes potencialmente relacionados com a patogenicidade (Gan *et al.*, 2013; Vieira *et al.*, 2016). Esta informação permite, em paralelo, a análise de dados de RNAseq e de sequenciação do genoma por NGS (Illumina), anteriormente obtidos, permite, com recurso a diversas ferramentas bioinformáticas, selecionar um conjunto de possíveis efetores para este estudo.

II. Revisão Bibliográfica

a. *Coffea arabica*

O cafeeiro é uma planta arbustiva, da família botânica Rubiaceae, cujo produto economicamente importante são as sementes, os grãos de café. No gênero *Coffea* existem mais 100 espécies diferentes, no entanto apenas duas têm importância econômica, *C. arabica* Linnaeus e *C. canephora* Pierre ex Frohener (Robusta) (Brisdon & Vercourt, 1988; Cros *et al.*, 1998; Sondahl & Baumann, 2001) que representam respectivamente 61% e 39% da produção mundial de café (ICO, 2019).

A espécie *C. canephora* foi descrita pela primeira vez em 1895, sendo natural da África Ocidental. Esta espécie é diploide ($2n=2x=22$ cromossomas) e alogâmica, a sua temperatura ótima de crescimento varia entre 24 e 30 °C e a baixa altitude, 0 a 700 metros, com uma precipitação anual de 2000-3000 mm. Esta espécie é cultivada essencialmente na África central e ocidental, no sul e ocidente Asiático e no Brasil e produz um café que habitualmente é considerado de menor qualidade do que *C. arabica* mas com um teor de cafeína mais elevado (entre 1,7 e 4%) (Clifford & Willson, 1985).

A espécie *C. arabica* foi descrita por Linnaeus em 1735 e é oriunda das zonas montanhosas do Sul da Etiópia, próximas da linha equatorial (Righi, 2005). Está adaptada a uma temperatura de 15 a 24 °C, em altitudes que variam entre 1000-2000 metros com uma precipitação anual de 1500-2000 mm. Esta espécie é atualmente cultivada na América latina, África oriental, Índia e Indonésia, produz um café de melhor qualidade com um teor de cafeína entre 0,8 e 1,4%. Apresenta-se como sendo a única espécie, do seu gênero, alotetraploide ($2n=4x=44$ cromossomas) e autogâmica (Batista de Sá, 2010; Davis *et al.*, 2006). Segundo Lashermes *et al.* (1999), esta espécie poderá ter origem da hibridação de outras duas espécies, *C. eugenoides* e *C. canephora*.

A planta do cafeeiro pode atingir entre 8 e 10 metros de altura, desenvolvendo-se a partir de um caule principal, de onde em seguida irão crescer pares de ramos laterais, horizontais. A partir destes ramos laterais poderão ainda desenvolver-se ramos secundários e terciários (António, 2008).

Em 2018 os maiores produtores de café foram o Brasil, o Vietname e a Colômbia que produziram respetivamente 35,15; 29,48 e 12,8 milhões de sacas de café (saca igual a 60 kg). Os maiores consumidores foram a Finlândia, a Noruega e a Islândia com o consumo *per capita* de, respetivamente, 12 kg, 9,9 kg e 9 kg (ICO, 2019).

b. Antracnose dos frutos verdes do cafeeiro

A antracnose dos frutos verdes do cafeeiro foi detetada pela primeira vez em 1922 no oeste do Quênia (McDonald, 1926). Inicialmente o agente causal desta doença foi identificado como *C. coffeanum*, uma espécie descrita em 1901 por Noack e que foi isolado de manchas nas folhas do cafeeiro arábica no Brasil. Posteriormente, Rayner (1952) criou o táxon *C. coffeanum* var. *virulans* para acomodar os isolados de *C. coffeanum* responsáveis pela sintomatologia nos frutos verdes e assim diferenciá-los dos que causam sintomatologia nas folhas. Este agente patogénico também era referenciado como *C. coffeanum* Noack *sensu* Hindorf (Hindorf, 1974, 1970). Waller *et al.* em 1993, tendo por base estudos morfológicos e bioquímicos, definiram este fungo como uma espécie independente, de nome científico *Colletotrichum kahawae* J.M. Waller & P.D. Bridge, utilizando como epíteto específico a palavra suaíli para café (*kahawae*).

A CBD espalhou-se do Quênia, para Angola em 1930, para o Zaire em 1937, para os Camarões entre 1955 e 1957, para o Uganda em 1959, para a Tanzânia em 1964, para a Etiópia em 1971 (van der Graaff, 1981), para o Malawi em 1985 (Lutzeyer, 1993) e para o Zimbabué e Zâmbia em 1985 (Masaba & Waller, 1992). O principal fator para a dispersão desta doença pelas principais regiões produtoras de café arábica do continente africano foi a falta de controlo na movimentação de material vegetal proveniente de zonas infetadas com CBD (Giddisa, 2016).

O fungo *C. kahawae* em termos taxonómicos pertence ao filo Ascomycota, à classe Sordariomycetes, à ordem Glomerellales e à família Glomerellaceae (*National Center for Biotechnology Information Taxonomy Browser*, <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>). É um fungo hemibiotrófico, ou seja, o seu desenvolvimento inicia-se de forma biotrófica, dependendo dos tecidos vivos do hospedeiro, e, depois de instalado, passa a necrotrófico, subsistindo a partir de tecidos vegetais mortos (Silva *et al.*, 2006).

A variabilidade genética de *C. kahawae* foi estudada com recurso a diferentes marcadores moleculares tendo sido observados baixos níveis de polimorfismo genético (Derso & Waller, 2003; Bridge *et al.*, 2008; Loureiro *et al.*, 2011; Silva *et al.*, 2012; Vieira *et al.*, 2018). Silva *et al.* (2012), através de estudos filogenéticos e de genética populacional, distinguiram três haplótipos que nomearam de acordo com a sua origem geográfica: Angola, Camarões e África Oriental. Verificaram que a população angolana seria a mais ancestral e as populações do este africano seriam as mais recentes.

Mais recentemente Vieira *et al.* (2018), analisando um painel de milhares de polimorfismos nucleotídicos únicos (SNPs), obtidos pela técnica de “RAD-seq” também identificou três

populações geneticamente distintas (Angola, Camarões e África Oriental). No entanto, os resultados destes autores sugerem que a população de Angola e da África Oriental tenham surgido praticamente ao mesmo tempo, e, só mais tarde, a população dos Camarões terá emergido a partir da população de Angola. Além disso, encontraram também duas linhagens clonais na população angolana. Este estudo, confirmou a baixa variabilidade genética do fungo *C. kahawae*, onde a ocorrência de eventos de recombinação é rara, parecendo ser um verdadeiro fungo clonal, ainda que possam ocorrer eventuais trocas de material genético, como deverá ter ocorrido na população angolana. As populações clonais têm uma grande vantagem a nível demográfico uma vez que um único indivíduo pode colonizar um habitat livre e gerar descendência em elevado número (Bazin *et al.*, 2014; Hansen *et al.*, 2016). No entanto, o seu potencial de adaptação a novos habitats é inferior devido à raridade dos eventos de recombinação genética (Dutech *et al.*, 2017). Paralelamente, alguns autores referem que fungos assexuados têm tendência a acumular mutações de deleção que os podem levar à extinção (Bazin *et al.*, 2014; Weir *et al.*, 2016).

Em termos patológicos, *C. kahawae* afeta todos os estados fenológicos e órgãos da planta, incluindo folhas, flores e frutos maduros, mas os maiores impactos na redução da produtividade devem-se a infeções nos frutos verdes, afetando a produção dos grãos de café (Firman, 1964; Masaba & Waller, 1992; Silva & Várzea, 2006).

Os sintomas nos frutos verdes podem ser de dois tipos, lesões ativas ou lesões *scab* (ou sarna) (Figura 1). As primeiras correspondem aos sintomas que se observam em interações suscetíveis e caracterizam-se por pequenas manchas negras, em depressão, que podem desenvolver-se rapidamente e cobrir toda a superfície dos frutos, que podem cair ou permanecer mumificados na planta. Em condições de humidade elevada, sob a superfície

lesada, o fungo produz uma massa de esporos de tom rosado/alaranjado e esta capacidade de esporulação pode manter-se durante vários meses após a infeção (Masaba & Waller, 1992). As lesões do tipo *scab* apresentam-se sob a forma de lesões de pequena dimensão, cor de camurça e em depressão, de formato variável e geralmente superficiais e correspondem aos sintomas macroscópicos de



Figura 1 - Sintomas da antracnose dos frutos verdes do cafeeiro. Fonte: Bedimo *et al.*, 2010.

resistência. Sobre estas lesões podem ou não aparecer acérvulos dispersos em círculos concêntricos (Várzea *et al.*, 2002). Este tipo de lesões poderá evoluir para lesões ativas, na fase de amadurecimento dos frutos, sem afetar o seu crescimento ou desenvolvimento e sem ter efeitos aparentes na produção (Masaba & Waller, 1992), sendo então o único tipo de lesões em variedades de cafeeiro resistentes (Firman, 1964; Masaba & Waller, 1992; Várzea *et al.*, 2002).

Nas variedades suscetíveis, a suscetibilidade dos frutos do cafeeiro a esta doença depende do seu estado de desenvolvimento. Os frutos são suscetíveis na fase de expansão (8 a 12 semanas após a floração), muitos suscetíveis na fase do endosperma (15 a 24 semanas após a floração) e na fase de maturação. Na fase de pré-maturação (25 a 31 semanas após floração) os frutos são menos suscetíveis, apresentando até alguma capacidade de resistência e podendo haver paragem na evolução de lesões já existentes (Muller, 1964; Muller, 1980; Várzea *et al.*, 2002) (Figura 2).

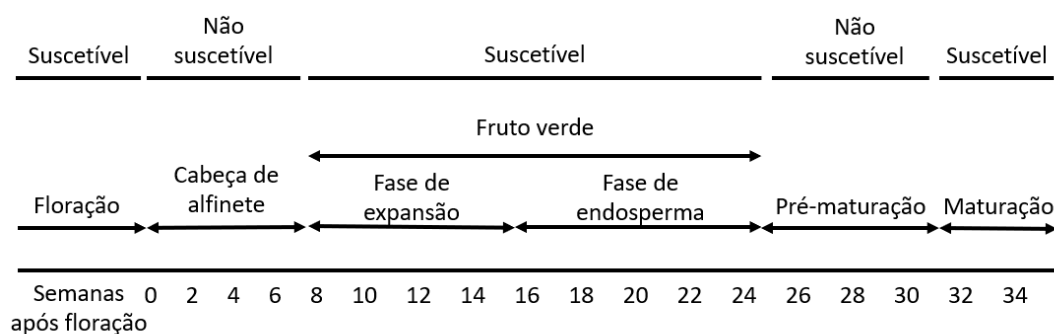


Figura 2 – Estados fenológicos do desenvolvimento dos frutos do cafeeiro, associadas às fases de suscetibilidade à doença. Adaptado de Macedo, 1994.

A infeção inicia-se pelo desenvolvimento de conídios a partir dos acérvulos presentes nos ramos ou frutos mumificados, do ciclo produtivo anterior (Firman, 1964). Segundo Gibbs (1969), os frutos infetados são a principal fonte de inóculo, uma vez que a progressão da doença está mais dependente do inóculo presente nestes órgãos. A germinação dos esporos dá-se em condições de humidade relativa de cerca de 100% e é essencial a existência de uma fina camada de água resultante de precipitação, neblina ou orvalho (Nutman & Roberts, 1960a; Nutman & Roberts, 1960b). A temperatura de germinação ótima é de 22°C (Nutman & Roberts, 1960b; Várzea *et al.*, 2002).

O primeiro passo para a infeção é a germinação dos conídios de *C. kahawae* originando tubos germinativos, na extremidade dos quais se diferenciam apressórios, que irão ser melanizados, permitindo a penetração e colonização dos tecidos vegetais do hospedeiro, diretamente através das células da epiderme (Garcia, 1999; Chen, 2002; Silva *et al.*, 2006; Loureiro, 2008). A hifa de infeção penetra o lúmen da célula do hospedeiro e forma uma

vesícula de infecção. O período de biotrofia, em que o fungo depende de células vivas para o seu desenvolvimento vegetativo, pode durar entre 48 e 72 horas após inoculação, dependendo do nível de agressividade do isolado. Neste período não são observáveis sintomas (Várzea, 1990; Várzea *et al.*, 2002; Silva *et al.*, 2006). De seguida observa-se a ramificação intra e intercelular das hifas e o fungo passa para a fase necrotrófica em que ocorre morte das células do hospedeiro (Garcia, 1999; Silva *et al.*, 2006; Loureiro, 2008). O ciclo da doença está representado na Figura 3.

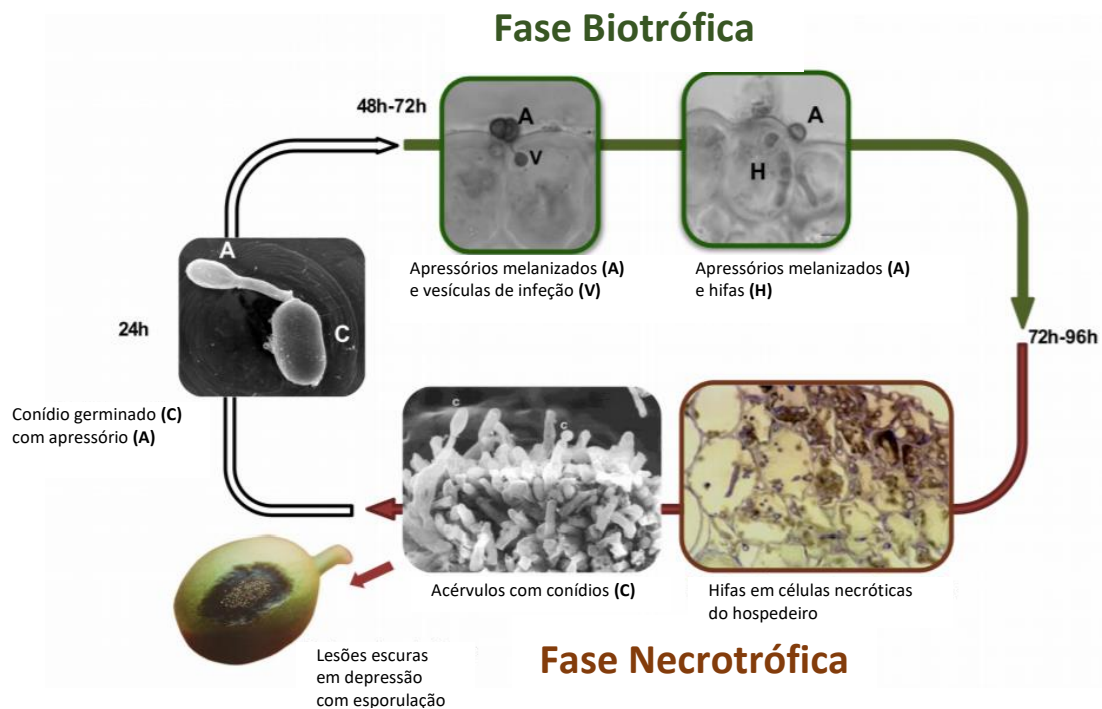


Figura 3 - Representação esquemática do ciclo da doença antracnose dos frutos verdes do cafeeiro. Adaptada de Diniz *et al.*, 2017; Loureiro *et al.*, 2012; Silva *et al.*, 2006 e Vieira, 2018.

c. Interação hospedeiro-agente patogénico

Na luta pela sobrevivência, as plantas desenvolveram diversas estratégias de defesa que lhes conferem a capacidade de resistir aos ataques por parte de microrganismos e insetos, entre outros. Estas estratégias de defesa dividem-se essencialmente em dois tipos: as defesas constitutivas e as defesas induzidas (Greenberg & Yao, 2004).

As defesas constitutivas são aquelas que se encontram sempre presentes, independentemente do tipo de agressão. São características estruturais das plantas como ceras, pelos, a produção de um composto antifúngico, entre outras, que impedem ou dificultam a penetração de microrganismos.

As defesas induzidas, tal como o nome indica, são ativadas em resposta a um ataque por parte de um agente patogénico. Este tipo de defesa envolve diversos mecanismos de transmissão de sinal e requer a rápida modulação da expressão génica, de modo a produzir proteínas essenciais no combate ao microrganismo. Estas defesas consistem essencialmente em dois tipos de respostas (Vleeshouwers *et al.*, 2008). As primeiras ocorrem na face externa da célula hospedeira e são desencadeadas por padrões moleculares característicos dos agentes patogénicos, *Pathogen-associated molecular pattern* (PAMP), reconhecidos por recetores de reconhecimento PRR (do inglês *Pattern recognition receptors*), resultando na imunidade desencadeada por PAMP (PTI, do inglês *PAMP-triggered immunity*). Os segundos tipos de resposta ocorrem já no espaço intracelular da célula hospedeira e envolvem a perceção de moléculas associadas à virulência, excretadas pelo agente patogénico. Estas moléculas são comumente denominadas efetores e o seu reconhecimento induz à imunidade (ETI, do Inglês *Effector-triggered immunity*). Se não forem reconhecidos, os efetores funcionam como determinantes de virulência e são essenciais para que a infeção se estabeleça com sucesso (Gan *et al.*, 2016, 2013; Presti *et al.*, 2015; Rao & Nandineni, 2017). Se reconhecidos, funcionam como fatores de avirulência (Jones & Dangl, 2006).

Ao longo da evolução, os fungos e outros agentes patogénicos desenvolveram também eles diversas estratégias para infetar e colonizar os seus hospedeiros, intervindo em processos bioquímicos, fisiológicos e morfológicos. O sucesso ou insucesso da infeção está, portanto, relacionado com as proteínas efetoras de patogenicidade, em particular com o seu reconhecimento ou não por parte do hospedeiro (Hogenhout *et al.*, 2009; Stergiopoulos & de Wit, 2009). Estas proteínas podem ser agrupadas em duas classes, consoante o seu alvo – efetores do apoplasto ou do citoplasma (Kamoun, 2006). Os fungos biotróficos e hemibiotróficos instalam-se dependendo de tecidos vegetais vivos e excretam ambos os tipos de efetores. Muitos dos efetores são proteínas ricas em cisteína que permitem a sua estabilidade no apoplasto do hospedeiro, ambiente hostil rico em proteases e outras hidrolases (Kamoun, 2006; Selin *et al.*, 2016; Stergiopoulos & de Wit, 2009). A expressão destas proteínas é específica durante a penetração celular e colonização da fase biotrófica do fungo (de Jonge *et al.*, 2011; Djamei & Kahmann, 2012; Rafiqi *et al.*, 2012). A maioria das proteínas efetoras são muito diversas em termos de sequências e não contêm zonas conservadas, pelo que a sua previsão tem vindo a ser realizada com base em diferentes critérios, baseados em características observadas em efetores conhecidos (Kamoun, 2006).

Os efeitos da infeção de *C. arabica* por *C. kahawae* podem ser devastadores levando a perdas de produção na ordem dos 80% quando nenhum tipo de medida de controlo é utilizado (Silva *et al.*, 2006; Bedimo *et al.*, 2010; Hindorf & Omondi, 2011). Por esta razão, além de um enfoque nos programas de melhoramento das plantas do cafeeiro, que constitui uma

importante ferramenta para o controlo desta doença (Vieira *et al.*, 2018), é também importante realizar estudos que permitam conhecer os mecanismos moleculares e funcionais do agente patogénico que lhe permitem estabelecer com sucesso a infeção (Vieira *et al.*, 2016).

Nos últimos anos foram publicados os genomas de cerca de 32 espécies de *Colletotrichum* (Baroncelli *et al.*, 2017), tendo sido realizados diversos estudos de genómica comparativa, uns focados nos genes relacionados com a patogenicidade (Gan *et al.*, 2013), outros na especialização em relação ao hospedeiro, (Baroncelli *et al.*, 2016; Gan *et al.*, 2016), outros nos mecanismos de transição da biotrofia para a necrotrofia (O'Connell *et al.*, 2012) e outros entre estilos de vida endofíticos e parasitários (Hacquard *et al.*, 2016; Hiruma *et al.*, 2016). Paralelamente, diversas análises comparativas dos secretomas revelaram que a maioria das proteínas previstas como sendo segregadas possuem atividade enzimática, e muitas são candidatas a serem efetores (Lelwala *et al.*, 2019, Rao & Nandineni, 2017, Baroncelli *et al.*, 2016; Gan *et al.*, 2013; O'Connell *et al.*, 2012; Sanz-Martín *et al.*, 2016). A variedade de hospedeiros existente em *Colletotrichum* sp. parece estar associada à perda ou ganho de famílias de genes, tais como as famílias de enzimas ativas de hidratos de carbono (CAZymes, do inglês *Carbohydrate Active Enzymes*) e as proteases (Baroncelli *et al.*, 2016, 2017) que no caso do género *Colletotrichum* estão altamente expandidas em comparação com outros Sordariomycetes (Baroncelli *et al.*, 2016; Gan *et al.*, 2013), sugerindo uma importância crescente dessas famílias de genes nos processos de infeção em diferentes hospedeiros e na adaptação aos diferentes estilos de vida (Gan *et al.*, 2016). A análise comparativa dos secretomas de diferentes espécies, sugere a existência de algumas estratégias de infeção conservativas, uma vez que alguns dos efetores identificados são homólogos conhecidos noutros agentes patogénicos (Baroncelli *et al.*, 2017). Entre eles encontram-se proteínas como NPP1, enzima indutora de necroses em *Phytophthora* sp., a proteína segregada BAS2, associada com a biotrofia em *Magnaporthe oryzae* (Mosquera *et al.*, 2009), a PEP1, proteína relacionada com a patogenicidade de *Fusarium* sp. em ervilhas (Coleman *et al.*, 2011). No genoma de *C. orbiculare* são ainda codificadas pequenas proteínas semelhantes a SIX1 e SIX6 (*Secreted In Xylem*) de *Fusarium oxysporum* e um homólogo do efetor de avirulência Ave1 de *Verticillium dahliae*, *F. oxysporum* e *C. higginsianum* (de Jonge *et al.*, 2012).

Estudos dos perfis de expressão génica revelaram que as CAZymes, juntamente com outras classes de proteínas segregadas (por exemplo proteínas contendo domínios de lisina - domínio LysM), são altamente moduladas durante o processo de infeção, sugerindo um papel importante na patogenicidade (Gan *et al.*, 2013; Kleemann *et al.*, 2012; O'Connell *et al.*, 2012). Dados de RNAseq para a interação *C. fructicola* - *Fragaria x ananassa* revelaram também a expressão diferencial ao longo do processo de infeção de genes envolvidos na degradação da parede celular da planta, metabolismo secundário e desintoxicação bem como

a presença e a ativação de numerosos efetores, confirmando que a transcrição é regulada de acordo com a fase da infecção sendo acompanhada por uma resposta de defesa do hospedeiro, a qual é seguida pela regulação do sistema de evasão do agente patogénico (Zhang *et al.*, 2018). Em trabalhos anteriores foram também referidos genes envolvidos na modulação do pH da célula hospedeira com um papel importante na patogenicidade (Fernandes *et al.*, 2017; Miyara *et al.*, 2010a) e que podem atuar durante a transição para a necrotrofia (Miyara *et al.*, 2010b), assim como diversos fatores de transcrição (Dufresne *et al.*, 2000).

Um dos métodos mais utilizado para analisar a expressão génica é o PCR quantitativo a partir de mRNA, uma vez que permite a deteção de pequenas variações de expressão (mesmo em transcritos pouco abundantes) com elevado rigor e precisão. No entanto, para se obterem resultados corretos é necessário proceder-se a uma normalização dos dados obtidos (Bustin *et al.*, 2009). Essa normalização deve ser feita em relação a genes de referência cuidadosamente estudados e selecionados (Vandesompele *et al.*, 2002; Andersen *et al.*, 2004). Genes de referência são genes constitutivamente expressos que são essenciais para manter as funções celulares básicas, ou seja, são genes cuja expressão se mantém constante, pelo menos durante as condições em análise (Lin *et al.*, 2014). O trabalho de Vieira *et al.* (2016) permitiu selecionar os melhores genes de referência (PP1, Actinae ck34620) para o estudo dos perfis de expressão de genes de *C. kahawae* ao longo do processo de infecção em *Coffea* spp.

III. Materiais e métodos

a. Material biológico

No presente estudo foram utilizados hipocótilos de *C. arabica* variedade Caturra (CIFC 19/1) e conídios de *C. kahawae* do isolado Que2 (proveniente do Quênia) ambos provenientes das coleções do CIFC (Centro de Investigação de Ferrugens do Cafeeiro, Instituto Superior de Agronomia, Universidade de Lisboa).

b. Inoculação hipocótilos e recolha de amostras para estudos de expressão genética

Os hipocótilos da var. Caturra foram inoculados com uma suspensão de conídios de Que2 com a concentração 2×10^6 de acordo com a técnica descrita por van der Vossen *et al.* (1976), com ligeiras modificações. Resumidamente os hipocótilos de *C. arabica* var. Caturra (CIFC 19/1; suscetível) foram dispostos em placas de esferovite perfuradas, que, por sua vez, foram colocadas em tabuleiros, de modo que a raiz ficasse em contacto com água. Posteriormente os hipocótilos foram inoculados com a suspensão de conídios de *C. kahawae*. Após a inoculação os tabuleiros foram colocados dentro de sacos de plástico para assegurar a manutenção de câmara húmida. Os sintomas observados foram registados ao longo do tempo. Foram recolhidas amostras de 15 hipocótilos (três réplicas biológicas) em diferentes tempos: 1, 2, 5, 7 e 10 dias após a inoculação (dai), as quais foram imediatamente congeladas em azoto líquido e guardadas a $-80\text{ }^{\circ}\text{C}$ para posterior extração de RNA.

Amostras de *C. kahawae*, representando diferentes estágios de desenvolvimento fúngico [conídios não germinados, conídios germinados com apressórios melanizados formados após 18 a 22 h em poliestireno ou em folhas de *C. arabica* (doravante referidos como apressórios)] foram obtidas da seguinte forma: conídios não germinados – colhidos a partir de culturas de 7 a 10 dias de idade em meio gelose de batata dextrosada (PDA, BD-Difco EUA) com um fotoperíodo de 12 h, em água estéril e separados do micélio usando um cadinho com placa filtrante de vidro sintetizado integrada com porosidade 1 (100 - 160 μm); apressórios *in vitro* - obtidos seguindo o método de Kleemann *et al.* (2008) com algumas modificações. Os apressórios formados em placas de poliestireno após 18 a 22 h foram raspados com 10 mL de água estéril contendo Tween 20 a 0,02% (Fisher Scientific, EUA). A suspensão foi centrifugada a 5000 *g* por 10 min, e o precipitado contendo os apressórios foi recolhido, liofilizado e armazenado a $-80\text{ }^{\circ}\text{C}$; apressórios *in planta* – obtidos na superfície abaxial de 12 folhas jovens de *C. arabica* (var. Caturra – CIFC 19/1) após pulverização com uma suspensão de conídios (2×10^6 conídios.mL⁻¹), usando um atomizador. As folhas inoculadas foram mantidas em câmara húmida por 24 h a $22\text{ }^{\circ}\text{C}$ para permitir a germinação dos conídios e a formação de apressórios. Após esse período, as folhas foram secas ao ar e uma fina camada

de verniz foi aplicada (Loureiro *et al.*, 2015). O verniz, contendo as estruturas fúngicas, foi deixado secar por 24 h e então cuidadosamente removido das folhas e armazenado a -80 °C até à extração de RNA.; micélio saprófita – obtido após a inoculação de 5 mL de suspensão conidial (2×10^6 conidia.mL⁻¹) do fungo em 25 mL de caldo de batata dextrosada (PDB, BD-Difco EUA). As culturas líquidas foram mantidas a uma temperatura de 22±1 °C, sob agitação orbital (100 rpm, agitador Agitorb 160 E, Aralab) durante três dias. Após esse tempo o micélio foi filtrado, liofilizado e congelado a -80 °C, para posterior extração de RNA. Esta amostra foi utilizada como controlo.

c. Bioinformática

Com o objetivo de identificar alguns genes envolvidos no processo de infecção de *C. arabica* pelo fungo *C. kahawae*, foi realizado um estudo *in silico*, essencialmente focado no estabelecimento da infecção e no *switch* para a necrotrofia. Foram utilizados os dados de sequenciação do genoma do isolado Que2 de *C. kahawae* cuja assemblagem foi guiada a partir dos dados de RNAseq anteriores [interação cafeeiro - *Colletotrichum kahawae* em, dois e três dias após a inoculação em interações compatíveis e incompatíveis (Fino *et al.*, 2014)]. Foram utilizadas diversas ferramentas bioinformáticas cujos *outputs* permitiram a seleção dos genes de interesse e contribuíram para a anotação do genoma do *C. kahawae*.

A análise *in silico* utilizou a sequência de aminoácidos de 19181 proteínas previstas pela anotação do genoma de *C. kahawae* isolado Que2 (*BioProject*: PRJNA564749; *BioSample*: SAMN12721764), o qual tem tamanho nuclear de 59142311 pb (~59Mpb) com 886 contigs, enquanto que o tamanho do genoma mitocondrial é 65534 bp (~65kbp) com um contig.

A identificação do secretoma foi baseada na previsão simultânea dos programas: SignalP 5.0 (Almagro Armenteros *et al.*, 2019; <https://services.healthtech.dtu.dk/service.php?SignalP-5.0>), TargetP 2.0 (Armenteros *et al.*, 2019; <https://services.healthtech.dtu.dk/service.php?TargetP-2.0>), Phobius (Käll *et al.*, 2004; <http://phobius.sbc.su.se/>) e TMHMM Server v2.0 (Krogh *et al.*, 2001; Sonnhammer & Krogh, 1998; <https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>), baseado no modelo HMM (Hidden Markov models). Estes programas permitem identificar a presença de sinais peptídicos “*N-terminal targeting peptides*”, responsáveis por encaminhar as proteínas para as respetivas vias secretoras (Armenteros *et al.*, 2019), excluindo as proteínas com regiões transmembranares cujas regiões *N-terminal* podem resultar em previsões erradas (Reynolds *et al.*, 2008).

A previsão da localização sub-celular foi feita com base na predição simultânea em pelo menos dois dos três programas utilizados: o WoLF PSORT (Horton *et al.*, 2007;

<https://wolfsort.hgc.jp/>), cujo algoritmo classifica as proteínas como pertencentes a 10 localizações distintas, e ainda identifica o grupo das proteínas que podem transitar entre o citosol e o núcleo; o DeepLoc-1.0 (Armenteros *et al.*, 2017; <https://services.healthtech.dtu.dk/service.php?DeepLoc-1.0>), que não se baseia em *machine learning* nem na comparação das sequências com outras homologas, mas que faz a predição apenas com base na informação da sequência (Armenteros *et al.*, 2017); o ProtComp 9.0 (<http://www.softberry.com>), que faz a predição usando o mesmo princípio que o DeepLoc, mas inclui a comparação com sequências homologas.

Para o estabelecimento do secretoma foram ainda retiradas as proteínas com sinal nuclear (PS50079) e com sinal para o retículo endoplasmático (PS00014), recorrendo para isso ao PROSITE (Hulo, 2006; <https://prosite.expasy.org/>), base de dados com uma vasta coleção de informação biológica de proteínas. Foram também retiradas todas as proteínas com potencial para formar âncoras GPI (glicosilfosfatidilinositol) de ligação às membranas. Para tal utilizaram-se os softwares PredGPI (Pierleoni *et al.*, 2008; <http://gpcr.biocomp.unibo.it/predgpi/>) e NetGPI 1.0 (Gíslason *et al.*, 2019; <https://services.healthtech.dtu.dk/service.php?NetGPI-1.0>).

O proteoma de *C. kahawae* previsto foi analisado pelo *software* EffectorP, que inclui agora duas versões, EffectorP 1.0 e EffectorP 2.0 (Sperschneider *et al.*, 2018, 2016; <http://effectorp.csiro.au/>). Este *software* está assente em *machine learning*, com o objetivo de identificar potenciais proteínas efetoras, cuja identificação tem sido baseada em critérios diversos como sequências com menos de 200 aminoácidos, elevado número de cisteínas, evidência de seleção positiva, localização do gene em regiões do genoma em constante evolução e/ou expressão genética diferencial durante a infeção, devido ao facto das sequências conterem poucas regiões conservadas (Kamoun, 2006). A previsão de efetores feita com base nas características dos efetores atualmente conhecidos limitam bastante a pesquisa, impedindo até que se descubram novas e diferentes proteínas efetoras (Petre *et al.*, 2014). Deste modo, a versão mais recente do EffectorP tem então o potencial de melhorar a previsão de proteínas efetoras, nomeadamente, ao fazer uma melhor distinção entre proteínas excretadas que são ou não efetoras, apresentado uma baixa percentagem de falsos positivos (Sperschneider *et al.*, 2018). Para uma seleção mais restritiva, neste trabalho foram cruzados os resultados de ambas as versões deste *software* (Figura 4).

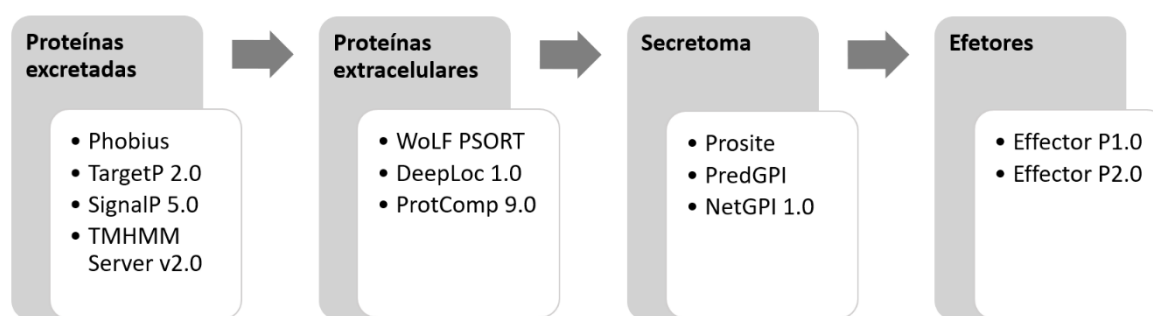


Figura 4 – Esquema explicativo das ferramentas bioinformáticas utilizadas na análise *in silico*.

A previsão da função das proteínas e das suas famílias, foi realizada com recurso a diversas ferramentas bioinformáticas, nomeadamente, pHMMER (<https://www.ebi.ac.uk/Tools/hmmer/search/phmmer>) e HMMScan (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>), fornecidas pelo HMMER 3.3 (Finn *et al.*, 2011; Potter *et al.*, 2018; <http://hmmer.org/>), eggNOG v5.0 (Huerta-Cepas *et al.*, 2019; <http://eggno5.embl.de/#/app/home>), KEGG Mapper v4.1 (Kanehisa & Sato, 2020; <https://www.genome.jp/kegg/>), BlastKOALA v2.2 (Kanehisa *et al.*, 2016; <https://www.kegg.jp/blastkoala/>) e Pfam (Finn *et al.*, 2014; <https://pfam.xfam.org/>) que permitem identificar: domínio conservados, famílias, PFAMs, termos GO, KOG, KEEGs, entre outras características, que, em conjunto, permitem inferir características como a função ou a via metabólica, por exemplo, de cada proteína.

A identificação das proteases foi feita recorrendo ao pHMMER, tirando partido da base de dados MEROPS (Rawlings *et al.*, 2018; <https://www.ebi.ac.uk/merops/>), cuja classificação é baseada em *clusters* que formam uma família de sequências homólogas. Adicionalmente, a identificação por parte deste *software* consiste na análise da parte da sequência de aminoácidos diretamente responsável pela atividade peptídica ou inibidora (Rawlings *et al.*, 2018), ou seja, o centro ativo.

As proteínas pertencentes às famílias das enzimas ativas de hidratos de carbono (CAZymes), foram anotadas usando o servidor dbCAN2 (Zhang *et al.*, 2018; <http://bcb.unl.edu/dbCAN2/index.php>), no sentido de fornecer uma previsão mais precisa, engloba três ferramentas diferentes. A ferramenta HMMER faz a homologia em relação à base de dados de domínios de CAZYme HMM sendo anotadas as proteínas que apresentem um valor de E Value < $1e^{-15}$ e uma cobertura > 0,35. A ferramenta DIAMOND faz um blast rápido com a base de dados de CAZymes anotadas (<http://www.cazy.org/>), anotando as proteínas que apresentem um E-Value < $1e^{-102}$ e um *hit per query* (-k) = 1 e o HotPep faz a homologia em relação à base de dados de pequenos péptidos CAZymes conservados, anotando as proteínas que apresentem uma frequência > 1,6 e *hits* > 6. Com o resultado destas três

ferramentas, consideram-se *CAZymes* aquelas proteínas que foram anotadas como tal em pelo menos duas destas três ferramentas.

Os genes envolvidos no metabolismo dos metabolitos secundários foram identificados pelas plataformas SMURF (do inglês *Secondary Metabolite Unknown Regions Finder*) (Khaldi *et al.*, 2010; http://smurf.icvi.org/run_smurf.php) e antiSMASH 5.0 (do inglês *antibiotics and Secondary Metabolite Analysis Shell*) (Blin *et al.*, 2019; <https://fungismash.secondarymetabolites.org/#!/start>) baseados na presença das principais enzimas envolvidas no metabolismo secundário (*backbone enzymes*) e no facto delas constituírem *clusters* com as enzimas auxiliares dessas vias e, por vezes, com proteínas relacionadas com o seu transporte (Blin *et al.*, 2019, Khaldi *et al.*, 2010).

Finalmente, realizou-se um *blast* com a base de dados PHI-base (*Pathogen-Host Interactions data base*; Urban *et al.*, 2020; <http://www.phi-base.org/>), onde estão anotadas proteínas estudadas noutros patossistemas como estando associadas à interação patógeno-hospedeiro e à patogenicidade.

d. Extração de RNA e síntese de cDNA

A extração de RNA foi efetuada recorrendo ao kit Spectrum Plant Total RNA kit (Sigma-Aldrich, USA), de acordo com as instruções do fabricante. Os resíduos de DNA genómico foram digeridos com DNase I (On-column DNase I Digestion Set, Sigma-Aldrich, USA) para assegurar que não existe DNA nas amostras. A concentração e qualidade do RNA foi avaliada por espectrofotometria a 260/280 nm e 260/230 nm no leitor de microplacas Biotek Synergy HT, enquanto que a integridade do RNA foi avaliada em gel de agarose a 1,2% em 1× tampão Tris-Acetato-EDTA (TAE, 40 mM Tris-acetato e 1 mM EDTA, pH 8,3) contendo GreenSafe Premium (NZYTech). O DNA complementar (cDNA) foi sintetizado de acordo como o protocolo “First Strand cDNA Synthesis” (Thermo Scientific) utilizando Oligo(dT)₁₈ e a transcriptase reversa (RevertAid Reverse Transcriptase, Thermo Scientific), a partir de 1 µg de RNA total para um volume final de 20 µL. O cDNA foi diluído 1/50 com água ultrapura esterilizada e armazenado a -20 °C para posterior utilização.

e. PCR quantitativo

O PCR quantitativo (RT-qPCR) é uma ferramenta essencial neste tipo de análise pois permite acompanhar a amplificação dos genes em tempo real e assim quantificar a expressão desse gene. Para uma reação de PCR é essencial conhecer bem os genes a amplificar e desenhar primers específicos desses mesmos genes.

Os primers foram desenhados com recurso ao programa PerlPrimer v1.1.21 (Marshall, 2004; <http://perlprimer.sourceforge.net/>) de modo a obter amplicões entre 80 e 200 pares de bases. O ideal é que os primers selecionados tenham entre 20 e 24 pb e estejam entre dois exões, para se garantir que a amplificação só ocorre no cDNA. Se não for possível satisfazer esta última condição, os primers podem apenas estar em exões distintos. A percentagem de GC deve ser entre 40 a 60% e a temperatura de emparelhamento deve estar entre os 60 e os 63 °C. Para evitar erros de emparelhamento, os primers devem conter, na extremidade 3', 2 ou 3 nucleótidos C ou G. Uma vez que as amostras contêm material genético tanto do fungo como da planta, é essencial proceder-se a uma pesquisa no Primer-BLAST (Ye *et al.*, 2012; <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>), para garantir a especificidade dos primers para *C. kahawae*. Isto pode ser especialmente importante para as fases iniciais da infeção em que a quantidade de DNA do fungo é reduzida, em relação ao da planta.

Um passo essencial após a obtenção dos primers é fazer a sua otimização, ou seja, testar as concentrações de cDNA que resultam melhor e as temperaturas adequadas para um bom emparelhamento. Desta otimização resultou a mistura de 5 ng de cDNA, 7,5 µL de EvaGreen Supermix (Bio-Rad) e 500 µM de cada primer num volume final de 15 µL.

As amplificações por RT-qPCR foram realizadas num termociclador MJ Mini thermocycler (Bio-Rad) equipado com MiniOpticon Real Time PCR 610 System (Bio-Rad). A reação iniciou-se com 3 min a 95 °C para ativar a polimerase e os ciclos reacionais consistiram em 45 ciclos de desnaturação a 95 °C durante 10 s, seguidos de 20 s à temperatura de emparelhamento específica de cada gene e, por fim, 10 s de alongamento a 72 °C.

Foi realizada a análise das curvas de dissociação obtidas no final das reações de PCR, numa amplitude de temperaturas entre 65 e 95 °C, com um aumento de 0,5 °C a cada 10 s. Cada conjunto de reações incluiu um controlo negativo sem cDNA e foram realizadas duas réplicas técnicas de cada amostra. As amostras em estudo incluem fungo *in planta*, micélio, apressórios *in vitro* e *in vivo* e esporos. Além disso fizeram-se também as reações de amplificação dos genes de referência escolhidos através do estudo realizado por Vieira *et al.* (2016) (actina, PP1 e ck 34620).

A eficiência de cada par de primers (E – eficiência de amplificação) foi testada experimentalmente com recurso ao programa LinRegPCR v2020.1 (Ruijter *et al.*, 2009), que realiza uma análise de regressão linear a partir dos dados da fluorescência da fase exponencial da amplificação por PCR.

A análise dos resultados de expressão e análise estatística foram realizados através do programa Bio-Rad CFX Maestro 2.0 v.5.0.021.0616, na função *Gene Study*. Para avaliação

da expressão génica, foram calculadas as quantidades relativas (RQ) para ambos os genes de referência e de interesse, utilizando-se a fórmula

$$RQ = E^{\Delta\Delta Cq}$$

em que E representa a eficiência de amplificação para cada gene e ΔCq a diferença entre os Cq da amostra em estudo e da amostra controlo

$$\Delta Cq = Cq_{micélio} - Cq_{amostra}$$

para cada um dos genes de interesse. Procedeu-se à normalização das quantidades relativas obtidas através de um fator de normalização

$$NF = (RQ_{ref1} \times RQ_{ref2} \times \dots \times RQ_{refn})^{\frac{1}{n}}$$

calculado pela média geométrica das quantidades relativas dos genes de referência.

Os valores de expressão normalizados obtidos para as amostras e o controlo (micélio 3 dias) foram analisados recorrendo ao teste *t-student* (com nível de significância de 0,05). Os valores de expressão normalizados obtidos para os nove grupos biológicos em estudo foram submetidos a análise de variância (ANOVA) utilizando um delineamento a um fator e um nível de significância de 0,05 em que se testou a hipótese de a média de expressão entre dois grupos biológicos ser igual. A comparação das médias entre dois grupos biológicos foi efetuada recorrendo ao teste de Tukey HSD (com nível de significância de 0,05). Foi também efetuado o teste de normalidade de Shapiro-Wilk para determinar se os valores de expressão das amostras dentro de cada grupo biológico foram obtidos a partir de populações com distribuição normal. Calculou-se ainda o *Fold-change* entre duas amostras (x e y), definido como a razão entre a expressão génica normalizada da amostra x e a expressão génica normalizada da amostra y, utilizando-se um *Fold-change threshold* de 4 para a regulação do gene e um nível de significância de 0,05. Estes valores foram observados num *Volcano Plot*.

IV. Resultados e discussão

a. Secretoma e efetoroma de *Colletotrichum kahawae*

Com base na sequenciação e montagem do genoma de *C. kahawae*, no qual foram previstas 19181 proteínas, procedeu-se à análise *in silico* e à comparação dos resultados com os obtidos em estudos de outros proteomas de *Colletotrichum* sp. De modo a descobrir algumas proteínas potencialmente envolvidas na patogenicidade foi necessário fazer uma seleção detalhada a partir desta análise.

Primeiramente, selecionaram-se as proteínas excretadas, sendo selecionadas aquelas que simultaneamente apresentavam sinais peptídicos e não incluíam domínios transmembranares nas predições efetuadas pelos programas Phobius, SignalP 5.0, TargetP 2.0 e TMHMM Server v2.0. Segundo o programa Phobius, 2599 proteínas são excretadas, apresentando sinais peptídicos e não contendo domínios transmembranares, enquanto que os programas SignalP 5.0 e o TargetP 2.0 previram que 2268 e 2596 proteínas secretadas, respetivamente. Por fim, o TMHMM Server v2.0 previu que 15006 proteínas não contêm domínios transmembranares. Cruzando a informação destes programas obteve-se uma previsão mais restritiva, mas menos propícia a erro, sendo então previstas 1774 proteínas excretadas, o que representa cerca de 9,3% do proteoma de *C. kahawae* (Figura 5).

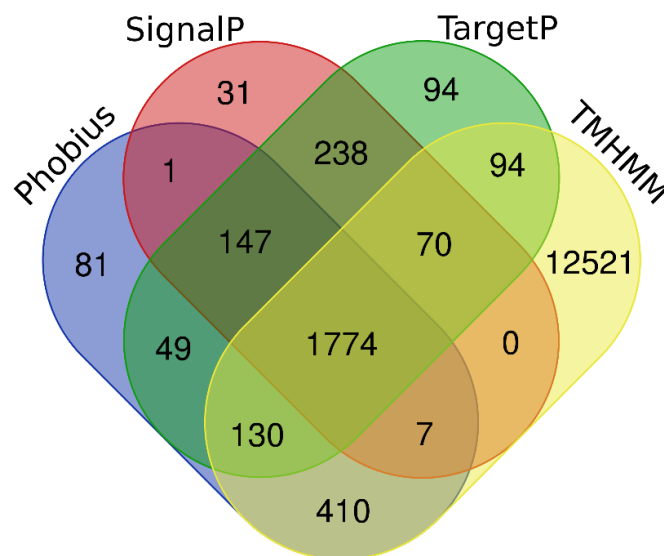


Figura 5 – Diagrama de Euler das proteínas previstas como apresentando sinais peptídicos e que não contêm domínios transmembranares de acordo com as predições efetuadas pelos programas Phobius, SignalP 5.0, TargetP 0.2 e TMHMM Server v2.0 num total de 19181 proteínas analisadas.

Diagrama obtido através de <http://bioinformatics.psb.ugent.be/webtools/Venn/>

Uma vez que nem todas as proteínas excretadas se destinam ao espaço extracelular, o passo seguinte foi prever a localização destas proteínas dentro da célula, sendo que, numa

primeira fase, as proteínas excretadas para o espaço extracelular seriam a fração mais interessante a estudar. Os programas utilizados para prever a localização apresentam resultados bastante idênticos e estes podem ser observados nos anexos 2, 3 e 4. Das 1774 proteínas potencialmente excretadas 1661 foram previstas com localização extracelular em pelo menos dois dos programas utilizados e na figura 6 apresentam-se as frações de proteínas extracelulares prevista por cada um dos programas.

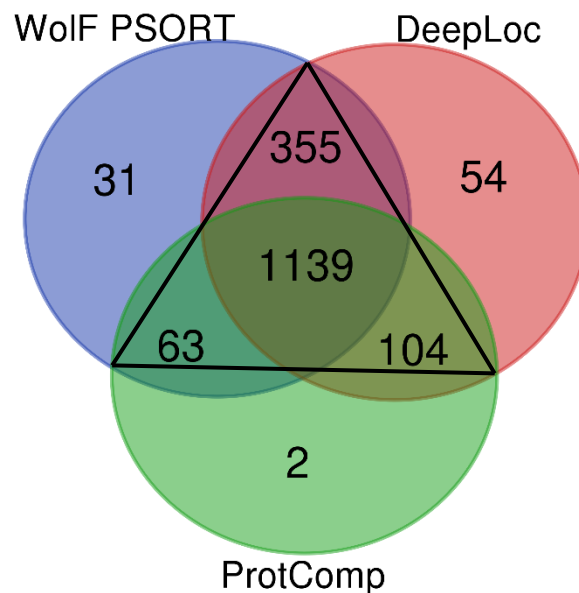


Figura 6 – Diagrama de Euler das proteínas previstas como excretadas para o espaço extracelular de acordo com as predições efetuadas pelos programas Wolf PSORT, DeepLoc-1.0 e ProtComp 9.0.

Diagrama obtido através de <http://bioinformatics.psb.ugent.be/webtools/Venn/>

Para a previsão do secretoma foram ainda excluídas as proteínas previstas pelo PROSITE com sinal nuclear (PS50079) e com sinal para o retículo endoplasmático (PS00017), e também as previstas como tendo âncoras de GPI pelos programas PredGPI e NetGPI. Daqui resultou que 1549 proteínas constituem o secretoma de *C. kahawae*, o que correspondem a 8,1% do total do proteoma. Este valor é muito semelhante ao valor previsto para o *C. truncatum*, onde o secretoma representa 9,2% do proteoma de 13724 proteínas (Rao & Nandineni, 2017), sendo que nos dois casos foram utilizadas praticamente as mesmas ferramentas bioinformáticas. No entanto, aplicando um método diferente, Baroncelli *et al.* (2016) obtiveram secretomas cuja percentagem, relativa ao total do proteoma, se situava em cerca de 15% para várias espécies de *Colletotrichum* sp. De estudos mais recentes resultou que o secretoma de *C. fruticola* representa cerca de 8,1% do proteoma, com 1440 genes de um total de 17827, respetivamente (Liang *et al.*, 2020) e que no secretoma de *C. gloeosporioides* existem 2047 genes no secretoma de 15744 genes previstos, representando 13,0% do proteoma (Fu *et al.*, 2020).

As proteínas efetoras foram previstas recorrendo ao programa EffectorP incluindo as duas versões disponíveis. Das 19181 proteínas que constituem o proteoma de *C. kahawae*, 4206 (21,9%) foram previstas como efetoras pelo programa EffectorP 1.0, enquanto que apenas 2397 (12,5%) foram previstas como tal pelo programa EffectorP 2.0. Cruzando a informação de ambas as versões, o total de proteínas simultaneamente previstas como efetoras é de 2094 (10,9%) proteínas. No âmbito deste trabalho, interessam aquelas proteínas que sendo excretadas são também previstas como efetoras. Deste modo, das 1549 proteínas que formam o secretoma de *C. kahawae*, 358 proteínas são previstas como efetoras (EffectorP 1.0 e 2.0), representando 1,9% do proteoma e 23,1% do secretoma deste fungo. Em comparação com outras espécies, *C. truncatum* tem 310 efetores secretados (Rao & Nandineni, 2017) e *C. fruticola* tem 372 (Liang *et al.*, 2020).

i. Enzimas ativas de hidratos de carbono

As enzimas envolvidas na biossíntese, modificação, ligação e catabolismo de glúcidos (CAZymes) desempenham um papel importante na degradação das paredes celulares de fungos e plantas. As CAZymes encontram-se divididas por seis famílias estruturalmente relacionadas nas quais se incluem todas as proteínas que possuem atividade catalítica relacionada com glúcidos ou módulos de ligação a glúcidos. O envolvimento destas proteínas na degradação, criação ou modificação de ligações glicosídicas torna-as muito importantes na patogenicidade, nomeadamente no caso dos fungos fitopatogénicos por possibilitarem a utilização de tecidos vegetais durante a colonização do hospedeiro (Lyu *et al.*, 2015).

De acordo com o programa dbCAN2, cerca de 718 (3,74%) genes previstos no genoma de *C. kahawae* codificam CAZymes. Alguns destes genes codificam mais do que uma família de CAZyme, resultando num total de 967 CAZymes identificadas no proteoma de *C. kahawae*. Dessas proteínas, 353 estavam presentes no secretoma previsto neste trabalho, contendo 524 domínios de CAZymes. Deste conjunto 13 proteínas foram previstas como efetoras, as quais contêm 18 domínios de CAZymes.

A classe da glicosil transferases (GT) representa 10,5% das CAZymes do proteoma, mas não tem representatividade no secretoma ou nas proteínas efetoras. Tanto no proteoma como no secretoma, a classe das glicosil hidrolases (GH) é que tem maior representatividade, no entanto, de entre as efetoras, é das menos representadas. A classe das polissacarídeo liases (PL) tem baixa representatividade no proteoma e no secretoma, mas uma representatividade significativa de entre as proteínas efetoras. As esterases auxiliares (AA) têm representatividades semelhantes no proteoma e no secretoma e ligeiramente mais baixa nas proteínas efetoras. As esterases de hidratos de carbono (CE) têm também

representatividades semelhantes no proteoma e no secretoma, mas apresentam maior importância nas proteínas efetoras. Por fim, as proteínas com módulos de ligação a glicídios (CBM) têm representatividades constantes ao longo do proteoma, do secretoma e dos efetores (Figura 7).

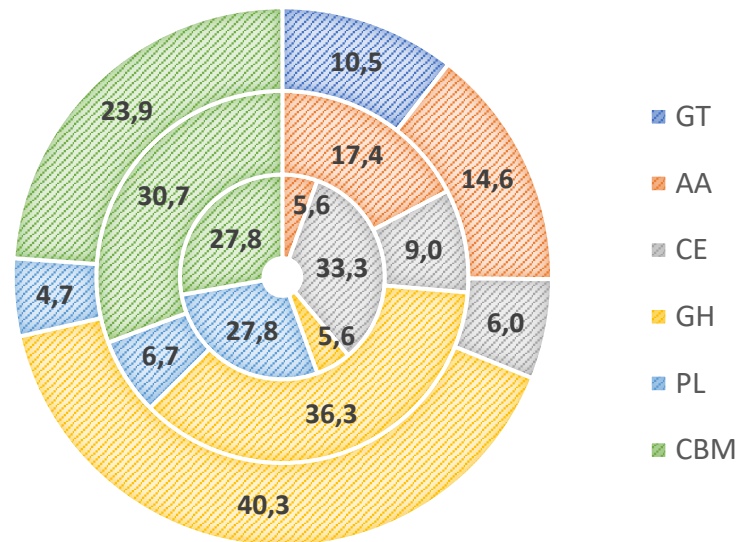


Figura 7 – Percentagem das diferentes classes de CAZymes previstas no proteoma de *Colletotrichum kahawae* (círculo exterior), no secretoma (círculo do meio) e para os possíveis efetores presentes no secretoma (círculo interior). Classes de CAZymes: glicosiltransferases (GT); esterases auxiliares (AA); esterases de hidratos de carbono (CE); Glicosil hidrolases (GH); polissacarídeos liases (PL); módulos de ligação a glicídios (CBM).

Os resultados obtidos para *C. kahawae* quando comparados com *C. truncatum* mostraram que *C. kahawae* codifica menos CAZymes no proteoma total (1120 para 967, respetivamente), mas mais no secretoma (477 para 524, respetivamente), e menos nos efetores (29 para 18, respetivamente) (Rao & Nandineni, 2017). Em *C. fruticola* foram identificadas 1211 CAZymes, sendo que esta identificação foi feita com recurso apenas a um *blast* com a base de dados das CAZymes (Liang *et al.*, 2020). Em *C. gloeosporioides* foram identificadas 2344, com recurso a um ferramenta distinta, que poderá ser menos restritiva (Fu *et al.*, 2020).

Nos diversos secretomas do género *Colletotrichum* sp. disponíveis para comparação, a percentagem das diferentes classes de CAZymes é ligeiramente diferente. Em média, a classe das AA representa 21,3% do secretoma, com algumas diferenças entre espécies, enquanto que no secretoma em estudo representa ligeiramente menos, cerca de 17,5% (Baroncelli *et al.*, 2016; Rao & Nandineni, 2017). A classe das CBM apresenta a maior diferença uma vez que no secretoma de *C. kahawae* representam cerca de 30,7% das CAZymes e, em geral, nos restantes secretomas, representa em média apenas 9,8%. Esta diferença pode ser potencialmente interessantes uma vez que proteínas que contêm estes domínios estão geralmente associadas à degradação da parede celular dos hospedeiros e à obtenção de glicídios como fonte de alimento para o agente patogénico (Gilbert *et al.*, 2013).

As CE possuem uma menor representatividade no genoma em estudo (9,0%) em comparação com os genomas anteriormente estudados (em média representa 15,9% do secretoma). A classe das GH é aquela que têm maior representatividade tanto no secretoma de *C. kahawae* (36,3%), como nos restantes secretomas (em média 42,9%). As GT não têm representatividade no secretoma aqui estudado, mas, em média, representa 4,1% dos restantes secretomas do mesmo género. Por fim, as PL têm representatividades semelhantes entre o genoma em estudo (6,7%) e todos os outros já estudados (6,1%) (Baroncelli *et al.*, 2016; Rao & Nandineni, 2017). Em anexo apresenta-se a comparação do número de proteínas das diferentes classes de *CAZymes* em diversas espécies de *Colletotrichum* sp. (anexo 4).

Dentro de cada uma das classes, as *CAZymes* estão divididas em diferentes famílias, sendo possível analisar a distribuição dos domínios destas famílias pelos genes em estudo. Rao & Nandineni (2017) fizeram uma análise do mesmo tipo para *C. truncatum*, o que permite comparar os resultados.

Na figura 8, apresentam-se as várias famílias de *CAZymes* presentes em cada uma das seis principais classes identificadas. Como se pode observar na classe das AA a família mais representada é AA3, seguida de AA9 e AA1. Em *C. truncatum* a mais abundante é AA7, seguida de AA3 e AA9, com 56, 43 e 31 genes contendo domínios destas famílias, respetivamente. Dentro das CE, as famílias têm quase todas a mesma representatividade, sendo as famílias CE3, CE4 e CE5 as mais abundantes. O mesmo não se verifica para *C. truncatum*, que apresenta uma família dominante, CE10, presente em 87 genes, a qual não está presente em *C. kahawae*. As famílias CE3, CE4 e CE5 existem em 11, 16 e 17 genes contendo domínios destas famílias, respetivamente. Na classe das GH a família mais abundante é GH43 em ambos os proteomas, com 42 genes em *C. truncatum* e 39 genes em *C. kahawae*.

A família de GT mais abundante é GT2, seguida de GT1 com apenas menos um gene. Em *C. truncatum* acontece o contrário, sendo GT1 a família mais abundante, em 26 genes, seguida de GT2 em 15 genes. A família de CBM mais abundante é CBM18, seguida de CBM2. Em *C. truncatum* as famílias com maior representatividade são CBM50 (que neste estudo está incluída nas “outras”) e CBM1 (sem representatividade neste estudo), ambas em 28 genes, seguidas de CBM18, em 24 genes. Por fim, as famílias de PL mais abundantes são PL1 e PL3, seguidas de PL4, semelhante ao que acontece em *C. truncatum*.

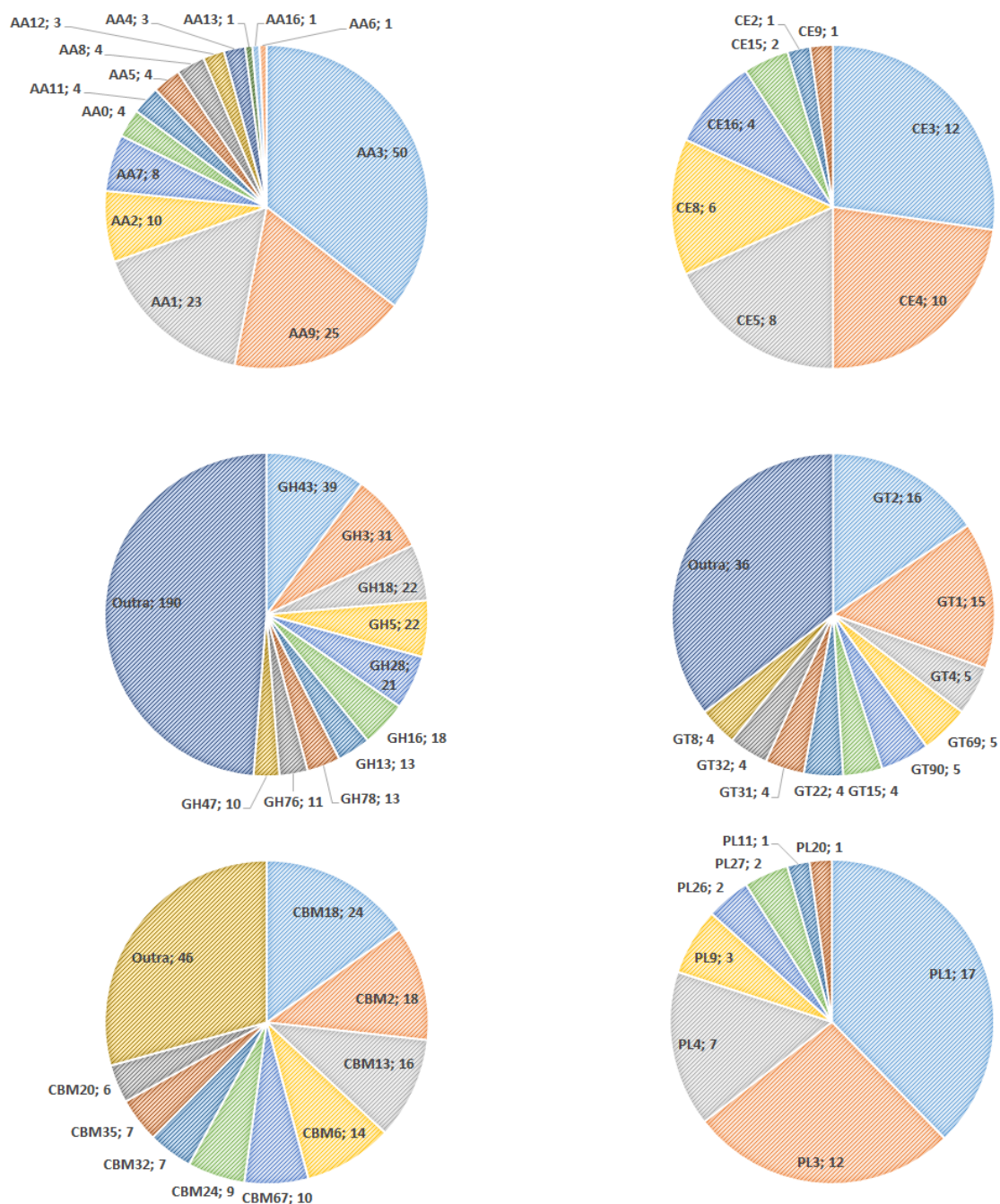


Figura 8 – Representatividade das diversas famílias de CAZymes, e respetivas classes, no proteoma de *Colletotrichum kahawae*. Glicosiltransferases (GT); esterases auxiliares (AA); esterases de hidratos de carbono (CE); Glicosil hidrolases (GH); polissacarídeos liases (PL); módulos de ligação a glicídios (CBM).

ii. Proteases

Várias proteases extracelulares foram até agora identificadas por contribuírem para o crescimento fúngico, formação das estruturas de infeção, degradação da parede celular, processamento proteolítico de proteínas relacionadas com a patogenicidade (PRs) ou atuando como elicitadores das respostas de defesa. Foi também observada uma correlação positiva entre a secreção de proteases e a agressividade da doença (Chalfoun *et al.*, 2013;

Chandrasekaran *et al.*, 2016; Chu *et al.*, 2015; Olivieri *et al.*, 2004; Soberanes-Gutiérrez *et al.*, 2015).

A identificação das proteases, feita recorrendo à base de dados MEROPS (E value < 1⁻¹⁰) do programa pHMMER, resultou num total de 869 proteases, 4,5% do proteoma de *C. kahawae*. De entre as 1549 proteínas que constituem o secretoma, apenas 194 proteínas são proteases, representando 12,5% do secretoma. No grupo das proteínas potencialmente efetoras, apenas 4 são putativamente proteases.

Na figura 9 podemos observar a distribuição das diferentes famílias de proteases no proteoma, no secretoma e nos efetores. A família com maior representatividade no proteoma e no secretoma é a das proteases serínicas, sendo que representa também 25% nos efetores, o que corresponde a uma de quatro proteínas. Os inibidores apresentam-se em muito baixo número no proteoma e no secretoma, mas são 50% das proteases efetoras (duas proteínas). Outra família de proteases importante é a das metaloproteases que tem uma representatividade bastante semelhante nos três grupos aqui comparados. As restantes famílias não têm representatividade nos efetores e apresentam-se em baixo número também no proteoma e no secretoma. As proteases cisteínicas têm ainda representatividade significativa no proteoma, que decresce no secretoma e é nula nos efetores.

Os resultados obtidos foram comparados com a bibliografia e esta comparação apresenta-se na tabela em anexo 5. A família de proteases com maior representatividade em todas as espécies do género *Colletotrichum* é a das proteases serínicas, representando em média 43,0% das proteases presentes no secretoma, seguida das metaloproteases, (36,9%). Estes resultados são semelhantes aos observados para *C. kahawae*, apesar da percentagem de proteases serínicas serem bastante mais elevadas (61,9 %), e a das metaloproteases ligeiramente mais baixa (26,8 %). A percentagem de proteases aspárticas é ainda relevante na generalidade das espécies, em comparação, representando 12,2%, das proteases presentes nesses secretomas, no entanto é menos representativa no secretoma de *C. kahawae* (5,7%). As restantes famílias apresentam percentagens semelhantes (Baroncelli *et al.*, 2016; Rao & Nandineni, 2017).

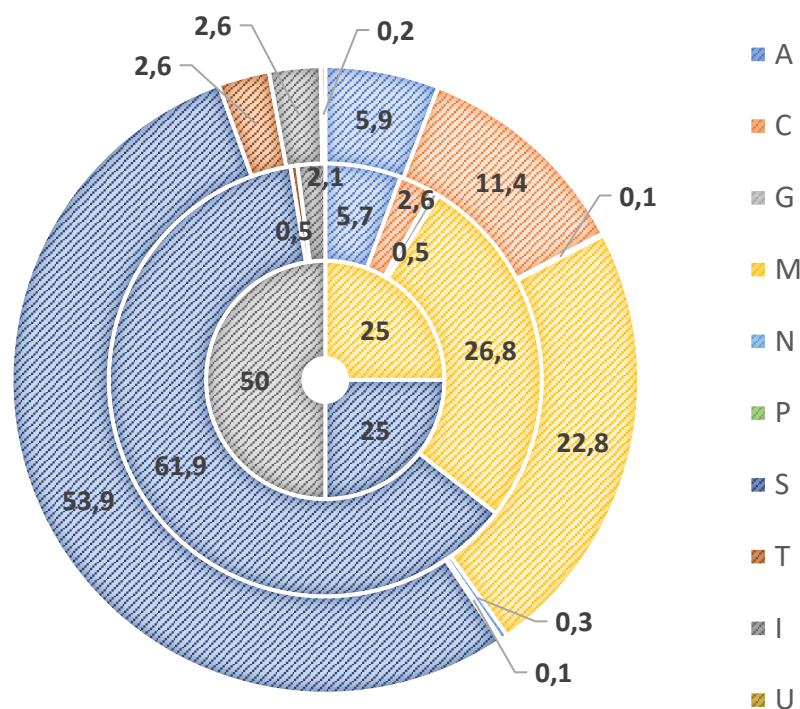


Figura 9 - Percentagem das diferentes famílias de proteases do total de proteínas presentes no proteoma de *Colletotrichum kahawae* (círculo exterior), no secretoma (círculo do meio) e para os possíveis efetores presentes no secretoma (círculo interior). Família das proteases aspárticas (A), cisteínicas (C), glutâmicas (G), metaloproteases (M), asparágicas (N), mistas (P); serínicas (S); treónicas (T), inibidoras (I) e desconhecidas (U).

iii. Metabolitos secundários

Tal como o nome sugere, os metabolitos secundários não são essenciais para o crescimento, desenvolvimento ou reprodução do fungo, mas são importantes na proteção contra stresses ambientais ou durante a interação com o hospedeiro, e a produção destes compostos é característica dos fungos filamentosos (Inglis *et al.*, 2013). A biossíntese dos metabolitos secundários tem várias vias metabólicas, as quais geralmente estão organizadas em *clusters* que incluem as enzimas principais, frequentemente referidas como *backbone enzymes* (Khaldi *et al.*, 2010) e as enzimas auxiliares que se encontram num *cluster* na proximidade das respetivas *backbone enzymes* (Hoffmeister & Keller, 2007). O programa de previsão SMURF inclui os grupos das sintases peptídicas não-ribossomais (NPRS), policetido sintases (PKS), proteínas híbridas das NPRS e PKS (NPRS-PKS), as prenilttransferases conhecidas como dimetilalil triptofano sintase (DMAT) e as terpeno ciclases (TC). No entanto, este programa não faz a previsão das terpeno ciclases por serem sequencias muito variáveis e difíceis de detetar bioinformaticamente (Khaldi *et al.*, 2010). O programa antiSMASH apesar

de se basear na procura das *backbone enzymes* e nos *clusters* por estas formados, considera mais e diferentes tipos de proteínas responsáveis pela síntese de metabolitos secundários. Desse modo permite a identificação de 52 tipos diferentes de *secondary metabolite 'biosynthetic gene clusters'* (BGCs). Por esta razão, a comparação dos resultados obtidos pelos dois programas é bastante dificultada.

O programa SMURF previu, no proteoma de *C. kahawae*, a existência de 58 *clusters* de genes associados aos metabolitos secundários e a existência de 67 *backbone enzymes*. Destas enzimas as que tem maior representatividade são as PKS (34 genes PKS e um gene PKS-like, 51,5%), seguidas das NRPS (11 genes NRPS e 13 genes NRPS-like, 35,3%). As proteínas menos presentes são as DMAT (sete genes DMAT, 10,3%) e os híbridos entre NRPS e PKS (2 genes híbridos, 2,9%) (Figura 10).

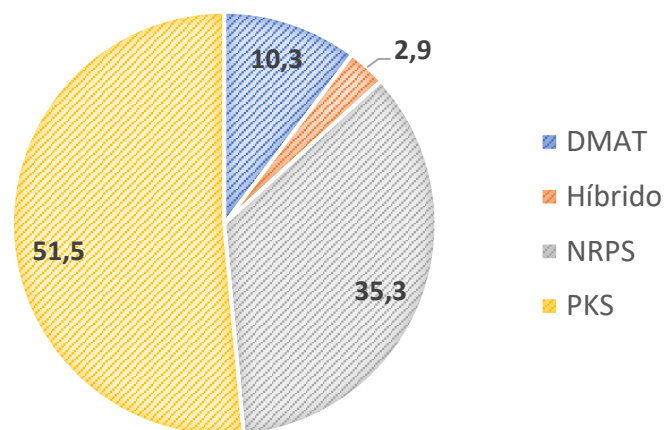


Figura 10 - Percentagem das diferentes famílias de *backbone enzymes* identificadas no proteoma de *Colletotrichum kahawae*, de acordo com a previsão efetuada pelo programa SMURF. Família das prenilttransferases (DMAT), das sintases peptídicas não-ribossomais (NRPS), das policetido sintases (PKS) e híbridos entre NRPS e PKS (Híbridos).

No trabalho de Rao & Nandineni (2017) e com recurso ao SMURF, foram previstas, no proteoma de *C. truncatum*, 90 *backbone enzymes* agrupadas em 73 *clusters* e divididas em quatro tipos. As PKS dominam com uma percentagem de 55,6%, seguidas das NRPS, com 30%, tal como observado em *C. kahawae*. As DMAT e os híbridos entre NRPS e PKS estão presentes numa percentagem de 10% e 4,4%, respetivamente. Relativamente a estas enzimas, o mesmo estudo fez a comparação entre diferentes espécies do mesmo género e o número de *clusters* pode variar entre 41 e 73 em *C. salicis* e *C. truncatum*, respetivamente. Em *C. fruticola* foram detetados 88 *backbone enzymes* (Liang *et al.*, 2020).

O programa antiSMASH 5.0 previu a existência de 82 *clusters* de genes contendo 97 enzimas envolvidas na síntese de metabolitos secundários. As enzimas com maior

representatividade são as PKS, com 34 genes PKS do tipo I (T1PKS) e dois genes PKS do tipo III (T3PKS), representando 37,1%, seguidas das NRPS com 15 genes NRPS e 12 genes NRPS-like (27,8%) e das TC com 19 genes TC (19,6%) As proteínas menos presentes são as indole com apenas 8 genes (8,2%), os híbridos entre NRPS e PKS com dois genes híbridos (2,1%) e outros tipos de proteínas como β lactona (dois genes) e Fungal-RiPP (dois genes) (Figura 11).

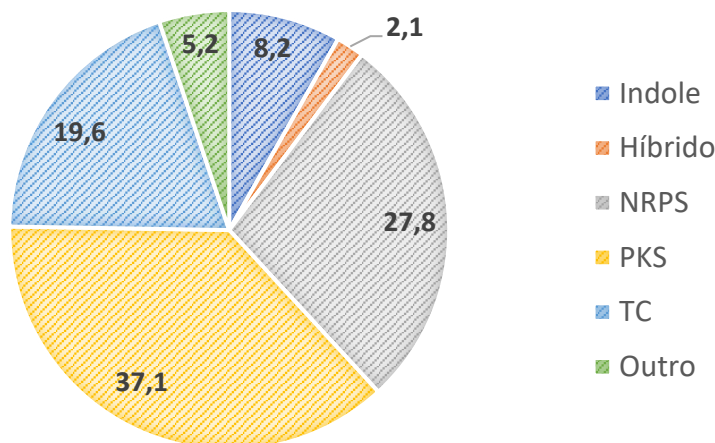


Figura 11 - Percentagem das diferentes famílias de *backbone enzymes* presentes no proteoma de *Colletotrichum kahawae* de acordo com a previsão do programa antiSMASH 5.0. Família das Indole alcaloides (Indole), das sintases peptídicas não-ribossômicas (NRPS), das policetido sintases (PKS); híbridos entre NRPS e PKS (Híbrido), das terpeno ciclastes (TC) e Fungal-RiPP e betalactona (Outro).

No trabalho de Rao & Nandineni (2017) recorreram também ao programa antiSMASH 4.0 para efetuarem a predição de *backbone enzymes* no proteoma de *C. truncatum*. Este programa previu a existência de 73 *clusters de metabolitos secundários*, os quais contêm 29 PKS do tipo I (39,7%), 12 NRPS (16,4%), 7 TC (9,6%), sete indole (9,6%) e diversos tipos de híbridos, nomeadamente sete híbridos NRPS-T1PKS, um híbrido indole-T1PKS, um híbrido T1PKS-T3PKS e um híbrido indole-T1PKS-NRPS (13,7% de híbridos) e oito *clusters* identificados como outros. Com esta análise podemos observar que as percentagens de PKS e indole não diferem muito entre os genomas de *C. truncatum* e *C. kahawae*, mas que as percentagens de NRPS e TC são bastante mais baixas em *C. truncatum* do que em *C. kahawae*. Por outro lado, a percentagem de híbridos é bastante mais elevada em *C. truncatum* do que em *C. kahawae*.

iv. Homologia com diferentes géneros

Paralelamente, a totalidade das 19181 proteínas previstas no genoma de *C. kahawae* foi comparada com a base de dados de sequências proteicas depositadas no NCBI por *blastP*, com vista à sua anotação. Como esperado a maioria das proteínas (84,59%) apresenta uma maior semelhança com proteínas pertencentes ao género *Colletotrichum* (F12). No entanto, o mais interessante é analisar a informação das proteínas não descritas para outros géneros de fungos, ou seja, aquelas para as quais não se obteve resultado no *blastP*, uma vez que podem ser únicas e específicas da espécie em estudo. Das 2272 proteínas que não obtiveram resultado no *blastP* (11,85% do genoma), 437 são previstas como efetoras, o que corresponde a cerca de 27% dos efetores identificados, 46 como excretadas para o espaço extracelular (de entre as quais 16 são efetoras) e 19 como proteases.

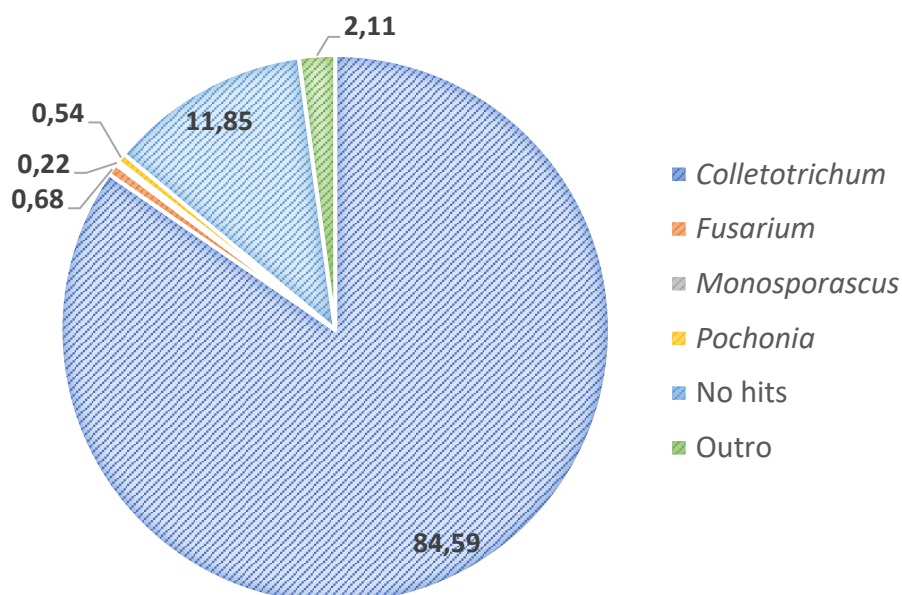


Figura 12 – Representatividade da homologia das proteínas em estudo com alguns géneros importantes.

v. Homologia com proteínas da base de dados PHI

Foi também realizado o *blastP* com a base de dados PHI que inclui as proteínas anteriormente estudadas noutros patossistemas e associadas à interação patogénio-hospedeiro. Este *blastP* mostrou que 5965 proteínas apresentaram homologia a outras proteínas já descritas como envolvidas no processo de infeção, o que representa cerca de 31,1% das proteínas identificadas. Das 358 proteínas efetoras pertencentes ao secretoma, 29

apresentam homologia com proteínas anotadas nesta base de dados, representando 1,9% do secretoma.

Segundo Rao & Nandineni (2017), em *C. truncatum*, 4165 proteínas apresentam homologia com proteínas já anotadas nesta base de dados, representando 30,6% do proteoma, semelhante ao que acontece com *C. kahawae*. Em *C. gloeosporioides*, 4273 proteínas apresentam homologia com a base de dados PHI, representando apenas 27,1% do proteoma (Fu *et al.*, 2020).

vi. Seleção das proteínas a estudar

O processo de seleção das proteínas a estudar está representado no fluxograma da figura 13. Inicialmente foram selecionadas as proteínas excretadas, cruzando-se esta informação com a sua localização sub-celular, removendo as proteínas com sinal nuclear (PS50079) ou do retículo endoplasmático (PS00014) e capazes de formar âncoras de GPI, resultaram 1549 proteínas pertencentes ao secretoma. Dessas proteínas, 358 são efetoras. Com recurso a outras ferramentas bioinformáticas que nos permitiram selecionar proteínas com homologia a proteínas de outros fungos, e homologia a famílias filogeneticamente mais próximas de *C. kahawae*, foi possível reduzir a seleção a 40 proteínas das quais se selecionaram proteínas potencialmente interessantes devido à sua família e função. Simultaneamente foram cruzadas as proteínas efetoras com o base de dados PHI, de onde resultaram 29 proteínas homólogas a outras já anotadas como estando envolvidas na interação hospedeiro – agente patogénico. Daqui foi selecionada uma única proteína que se julgou interessante, CKA_15357, uma vez que continha um domínio LysM e estas proteínas já foram descritas como sendo importantes neste tipo de interação (Gan *et al.*, 2013; Kleemann *et al.*, 2012; O’Connell *et al.*, 2012).

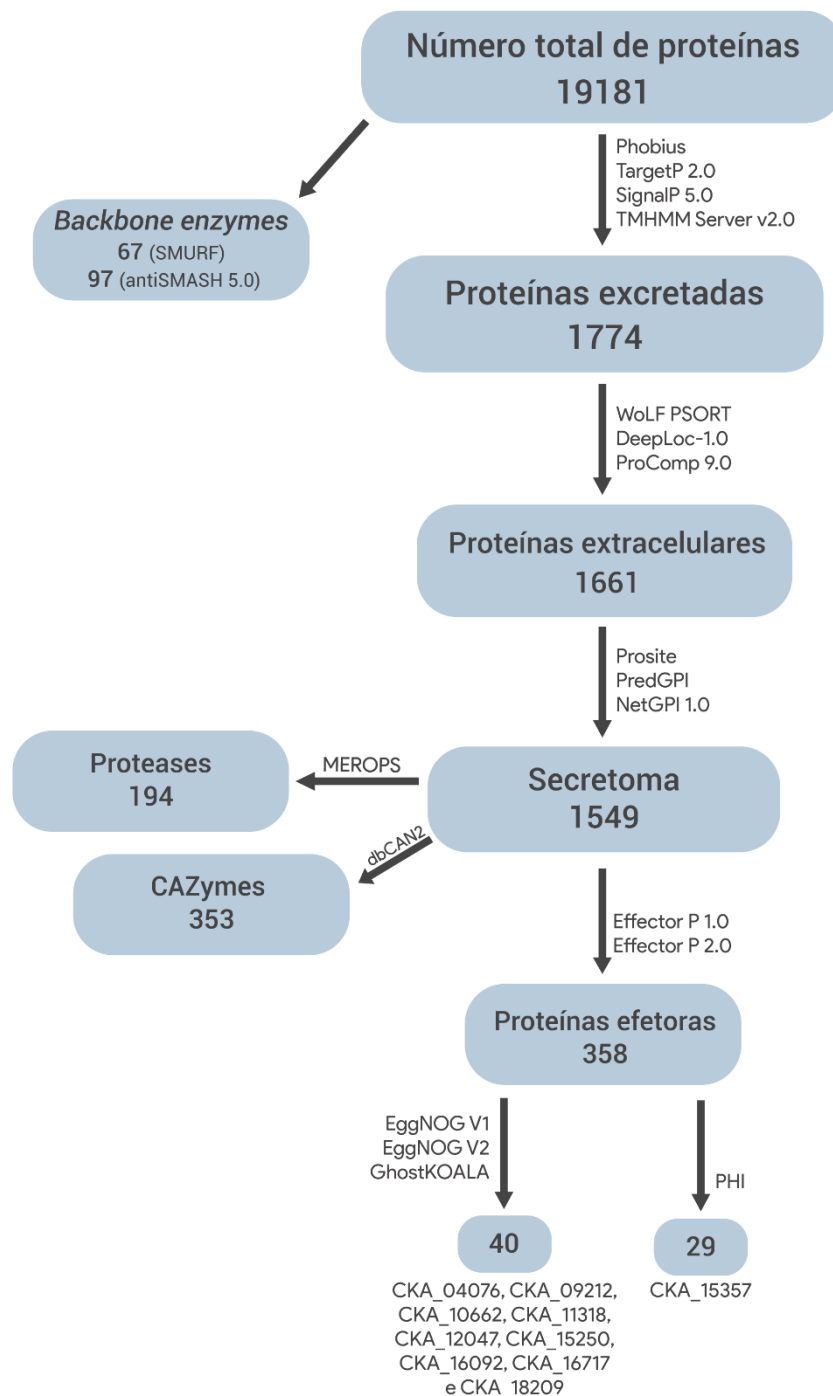


Figura 13 – Fluxograma representativo da sequência de seleção das proteínas potencialmente interessantes para serem estudadas.

Além das proteínas selecionadas com base na análise *in silico*, foram ainda selecionadas as proteínas, CKAH01_09762 e CKAH01_11483, que foram identificadas em dois mutantes de *C. kahawae* obtidos com recurso à transformação mediada por *Agrobacterium tumefaciens* e que apresentaram agressividade reduzida ou nula para os frutos verdes do cafeeiro devido a estes genes terem ficados truncados após a inserção do T-DNA (Cabral *et al.*, 2016); e as proteínas CKAH01_11118 e CKAH01_14472 que foram selecionadas por serem homólogas

das proteínas TDZ26243 (ENH82067; Cob_09713) e TDZ18289 (ENH88842; Cob_02999) as quais foram referidas por Gan *et al.* (2013) como sendo umas das 100 proteínas mais expressas um dia após a inoculação de *C. orbiculare* em *Nicotiana benthamiana*.

b. PCR quantitativo

Na interação *C. kahawae* – *Coffea* sp., os conídios são as entidades que dão início à infeção. Após o contacto com os frutos verdes dá-se a germinação com a emissão de um tubo germinativo, no fim do qual se diferenciam os apressórios e posteriormente ocorre a sua melanização (24 h após inoculação), iniciando-se a penetração dos tecidos (por volta das 48 horas após inoculação) com a formação das vesículas de infeção e o desenvolvimento de hifas no interior do hospedeiro. Durante este período de tempo o fungo tem um crescimento biotrófico, até cerca de 72 a 96 h após inoculação, altura em que se dá o *switch* para a fase necrotrófica (Silva *et al.*, 2006). Neste trabalho, em que se pretende estudar os determinantes moleculares que permitem o estabelecimento, com sucesso, da infeção, avaliando a expressão de genes de potenciais efetores ao longo dos diferentes estádios da infeção, recorreu-se à técnica do PCR quantitativo (RT – qPCR) com amostras recolhidas em diferentes fases fulcrais da infeção: 1, 2, 5, 7 e 10 dias após a inoculação (dai) tempos médios que correspondem respetivamente à formação e melanização dos apressórios, fase de penetração e crescimento biotrófico, *switch* para a necrotrofia e crescimento necrotrófico.

Inicialmente foram selecionados 14 genes para serem estudados, mas foi apenas possível obter resultados de expressão para sete desses genes, resultados esses que são apresentados nas figuras 14 a 16 e 18 a 21. Na tabela apresentada no anexo 6 são apresentados os primers com as respetivas temperaturas de emparelhamento, tamanho dos amplicões e a eficiência para as reações de RT-qPCR dos genes escolhidos e dos genes de referência obtidos de Vieira *et al.* (2016). Os restantes genes estudados são apresentados no anexo 7, listando os problemas encontrados que levaram ao seu abandono. No anexo 8 pode-se observar a eletroforese efetuada para verificar a integridade do RNA, as curvas de dissociação e as eletroforeses de algumas amostras dos genes estudados. No anexo 9 apresentam-se os dados relativos aos resultados da expressão génica.

i. Cutinases

A primeira barreira física imposta aos agente patogénicos pelos seus hospedeiros é a cutícula presente à superfície dos tecidos vegetais (Kolattukudy *et al.*, 1995). Esta cutícula é constituída por cutinas e por isso é necessário que os fungos excretem uma enzima que lhes

permita ultrapassar esta barreira. Estas enzimas são denominadas cutinases e são essencialmente responsáveis por catalisar a hidrólise das ligações de éster dos polímeros de cutina (Degani *et al.*, 2006; Dickman *et al.*, 1989).

A cutinase (EC 3.1.1.74) em estudo é expressa pelo gene CKAH01_04076 e a sua expressão relativa normalizada está representada na figura 14. Nesta figura podemos observar que apenas existem diferenças significativas (Teste de *t-student*, com um nível de significância de 0,05) entre 1, 7, 10 dai e o controlo. Contudo verifica-se que a expressão deste gene é ativada 1 dai, altura em que se observa uma maior expressão relativa, com um *Fold change* > 4 com um p-value < 0,05 em relação a todas as fases estudadas. Aos 2 dai a expressão diminui, mas mantém-se ativada comparativamente aos 10 dai (*Fold change* > 4 e p-value < 0,05). A expressão, nos apressórios *in vivo* é, aparentemente, ativada, no entanto não é possível determinar as diferenças significativas por só termos uma réplica biológica. No entanto, este resultado é expectável, uma vez que a força exercida pelo apressório poderá ser auxiliada pelas cutinases que deverão ser a chave para o sucesso na penetração do hospedeiro, como já foi descrito em Podila *et al.* (1995) e Chen *et al.* (2007). Nos apressórios *in vitro* a expressão é bastante mais baixa, possivelmente porque na ausência do hospedeiro, não há necessidade do fungo produzir estas enzimas. Por fim, é interessante observar que a expressão desta enzima é reprimida nos esporos, sugerindo que é apenas necessária após a infeção do hospedeiro, para permitir a penetração dos tecidos.

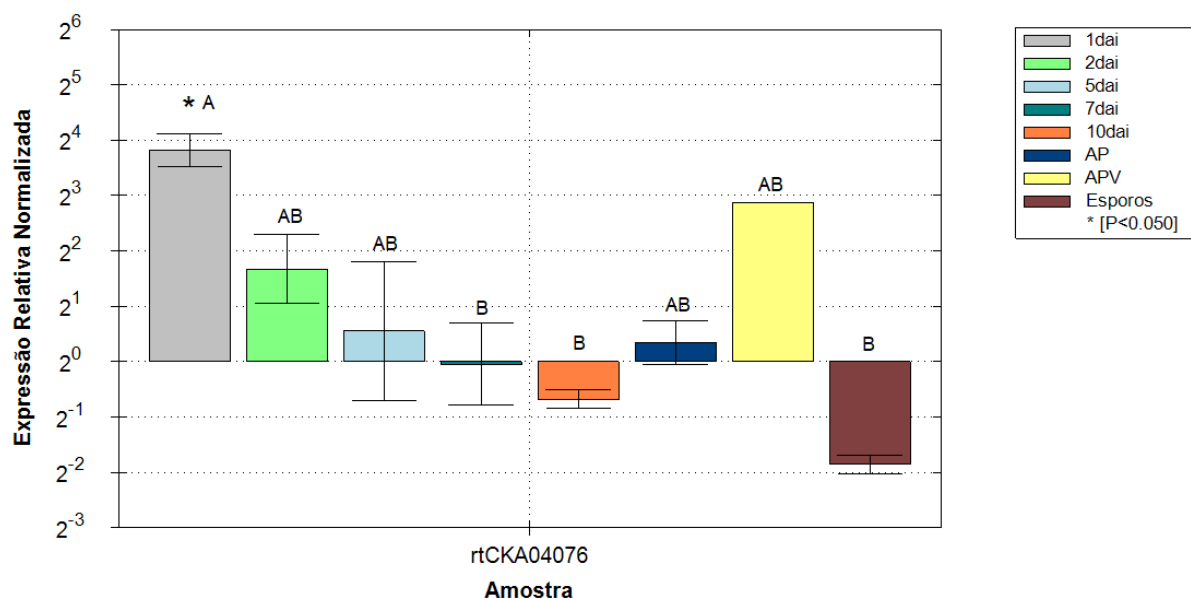


Figura 14 – Perfil de expressão génica obtido do gene CKA01_04076 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

A importância das cutinases nas primeiras etapas da infecção já tinha sido descrita por Dickman & Patil (1986) cujos trabalhos revelaram efetivamente o seu papel na penetração dos hospedeiros ao verificarem a incapacidade de infectar o hospedeiro por parte de mutantes de *C. gloeosporioides* para a cutinase, a menos que sejam feitas lesões no hospedeiro, *à priori* da inoculação. Em concordância com estes resultados, está o estudo realizado em *Mycosphaerella* sp., um fungo capaz de infectar o hospedeiro através de lesões, no qual a inserção de um gene codificante de uma cutinase resultou num mutante capaz de infectar os tecidos intactos do hospedeiro (Dickman *et al.*, 1989).

ii. Pectato liases

As pectato liases (EC 4.2.2.2) são enzimas que catalisam a quebra das ligações α -(1-4) entre os resíduos de ácido galacturônico nas pectinas constituintes das paredes celulares das plantas (Carpita & Gibeaut, 1993), tendo ação na degradação da parede celular dos hospedeiros. Assim sendo, as proteínas responsáveis pela degradação das paredes celulares apresentam um papel importante na patogenicidade (Annis & Goodwin, 1997). Em *C. magna* a inibição de um gene codificante de uma pectato liase resultou numa diminuição no desenvolvimento dos sintomas provocados por este mutante (Wattad *et al.*, 1995a), e em *C. gloeosporioides* a inibição de outro gene também codificante deste tipo de enzimas não demonstrou qualquer efeito na germinação de esporos, formação do tubo germinativo ou formação de apressórios, mas observou-se uma redução nos sintomas provocados pelo mutante (Wattad *et al.*, 1997). Segundo Yakoby *et al.* (2000a) num estudo envolvendo estes dois fungos, as pectato liases têm não só a ação de maceração dos tecidos vegetais, mas possivelmente também têm influência sobre outros fatores de virulência, uma vez que a expressão destas enzimas resultou numa maior patogenicidade e virulência. Noutro estudo do mesmo autor, sobre estas proteínas, verificou-se que a sua secreção é inibida a um pH baixo (Yakoby *et al.*, 2000b). Esta característica já tinha sido referida anteriormente tendo-se observado que determinadas enzimas extracelulares, como as pectato liases, requerem um ambiente mais alcalino para serem ativadas (Wattad *et al.*, 1995, 1997). Segundo Prusky *et al.* (2001), o próprio agente patogénico tem a capacidade de alcalinizar o hospedeiro ao acumular amónio, contribuindo para a expressão de proteínas como as pectato liases e aumentando a sua virulência.

Neste trabalho foram estudadas duas pectato liases, codificadas pelos genes CKAH01_09212 e CKAH01_12047 as quais apresentam um padrão de expressão bastante distinto (Figuras 15 e 16). No gene CKAH01_09212 não foi detetada expressão 1 dai e o pico de expressão ocorre 2 dai, havendo uma ativação do gene comparativamente às restantes

amostras tendo-se observado um valor de Fold change > 4 e p-value < 0,05. Aos 5, 7 e 10 dai não se observaram diferenças significativas entre si, mas a expressão relativa é significativamente maior do que nos esporos ou nos apressórios *in vitro*. O facto de não ter sido detetada expressão deste gene 1 dai e nos apressórios *in vivo* sugere que a função deste gene esteja relacionada com a penetração para o interior dos tecidos do hospedeiro e o estabelecimento da fase necrotrófica. O facto dos apressórios *in vitro* e *in vivo* serem diferentes não é de estranhar uma vez que já em 2012 O'Connell *et al.* referiram diferenças de expressão em estruturas desenvolvidas *in planta* e *in vitro*.

A finalizar, é de notar que não existem diferenças significativas na expressão relativa normalizada ($p < 0,05$) ao longo do processo de infeção na planta, no entanto observaram-se diferenças significativas na comparação da expressão relativa pelo Teste de Tukey HSD ($p < 0,05$) e uma ativação deste gene (Fold change > 4 e p-value < 0,05) relativamente ao controlo (exceto aos 5 dai).

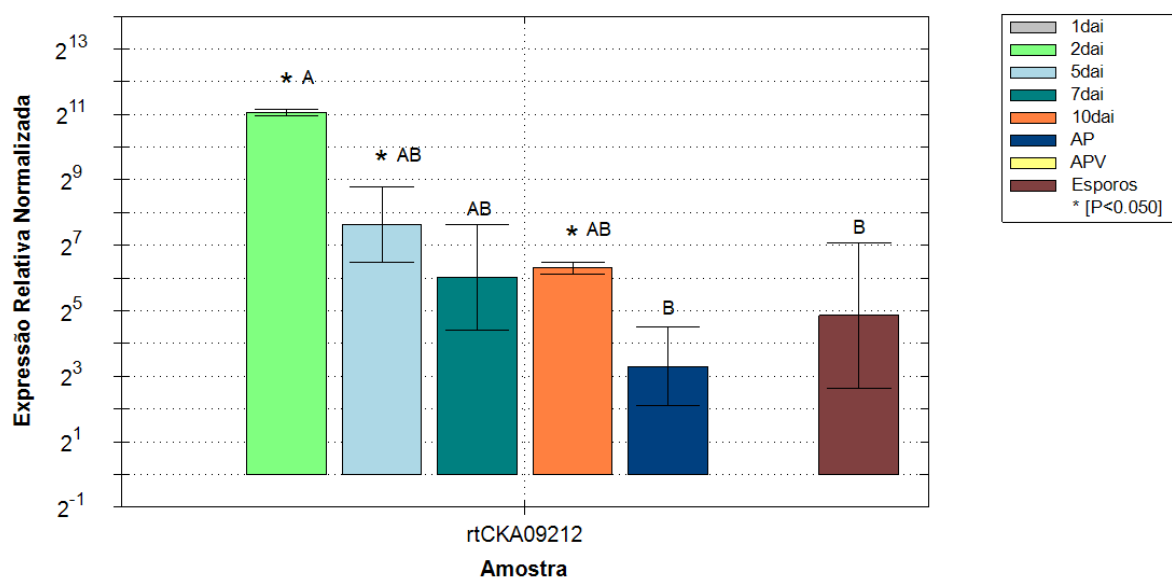


Figura 15 – Perfil de expressão génica obtido do gene CKA01_09212 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

O gene CKAH01_12047 não é expresso 1 e 2 dai. A maior expressão ocorre aos 5 dai, tendo-se observado diferenças significativas comparativamente às restantes amostras, apresentando-se ativado (Fold change > 4 e p-value < 0,05) neste tempo. A expressão diminuí significativamente a partir dos 7 dai. Nos apressórios *in vitro* este gene é altamente reprimido. A grande ativação aos 5dai a sugere o seu possível envolvimento no *switch* para a necrotrófica

e no seu estabelecimento. Tal como não se observa expressão nos apressórios *in vivo* o que está de acordo com os resultados obtidos 1 dai.

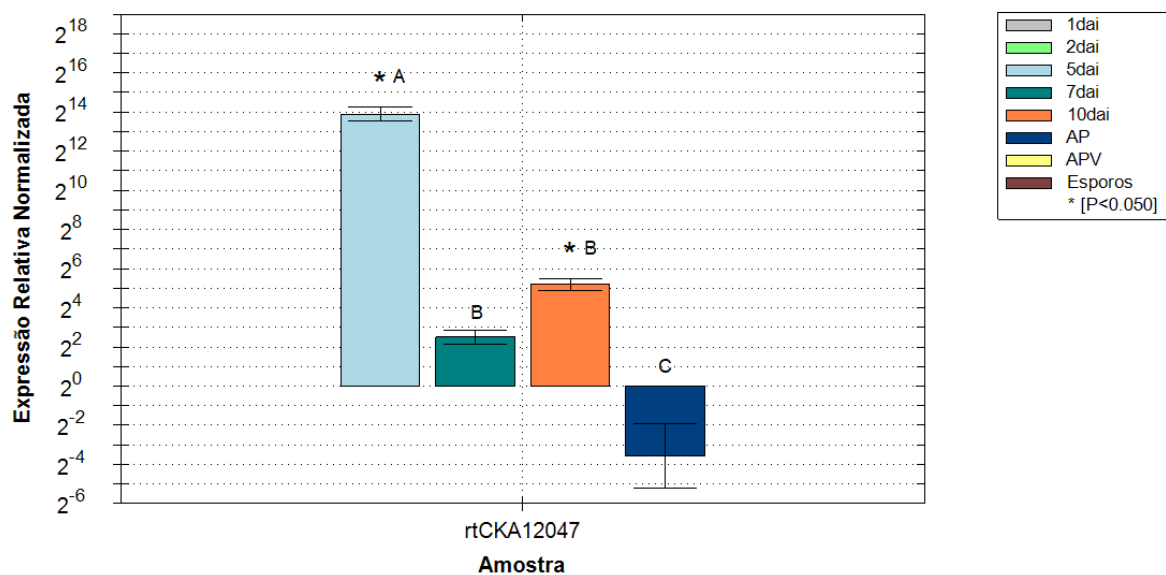


Figura 16 – Perfil de expressão génica obtido para o gene CKA01_12047 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dai) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

iii. Ceramidases

As ceramidases (EC 3.5.1.23) são enzimas responsáveis por metabolizar ceramidas em esfingosinas, ao hidrolisarem as ligações N-C (não incluindo as ligações peptídicas) entre a base e o ácido gordo (Figura 17). No entanto, foi descoberta, no arroz, uma ceramidase, cuja função é contrária àquela que seria a sua função normal, ou seja, catalisa a reação inversa, dando origem a fitoceramidas a partir de ácidos gordos de cadeia longa (Pata *et al.*, 2008).

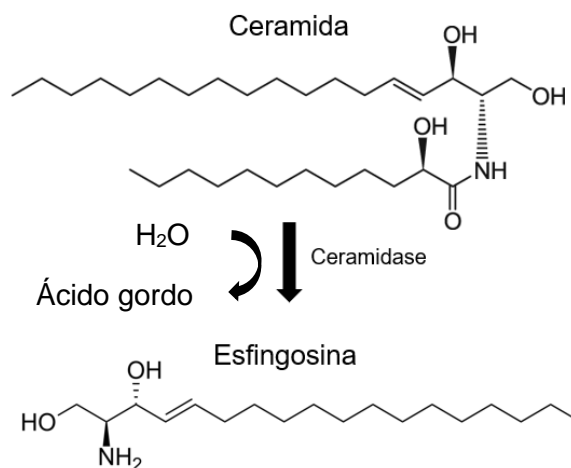


Figura 17 – Esquema representativo da reação de hidrólise da ceramida em esfingosina, por ação da enzima ceramidase.

Adaptado de (Canals *et al.*, 2011).

Existem três tipos de ceramidases, as ácidas, as alcalinas e as neutras, dependendo do pH ótimo para serem funcionais (Mao & Obeid, 2002).

Neste estudo analisámos os perfis de expressão de uma ceramidase, codificada pelo gene CKAH01_09762. Esta proteína apresenta uma maior expressão 1 dai mostrando-se ativada relativamente às restantes amostras com exceção dos 2 dai (Fold change > 4 e p-value < 0,05), contudo, os valores de expressão quer dos apressórios *in vivo* quer dos apressórios *in vitro* não confirmam esta ativação durante as primeiras 24 horas da infeção. A partir de 1 dai a expressão vai decrescendo ao longo da infeção e aos 2 dai e 5 dai a diferença de expressão só é significativa comparativamente aos esporos, aos apressórios e ao controlo, este gene mostra-se ainda ativado nestes dois tempos em comparação com as amostras referidas (Figura 18). Resumindo, este gene é ativado no início da infeção, sugerindo o seu envolvimento nas fases iniciais do processo de infeção, mas os elevados valores ainda observados em tempos mais tardios preconiza que se trate de um gene com um papel relevante ao longo de todo o processo de infeção. Em *C. gloeosporioides* foi estudada uma glicosilceramidase que se apresentou como sendo importante na regulação da germinação, no desenvolvimento do fungo, na esporulação e na patogenicidade, ao afetar a expressão de diversos outros genes (Huang *et al.*, 2019).

Adicionalmente, em *Aspergillus* sp. as ceramidases foram detetadas como sendo necessárias para o desenvolvimento de hifas e germinação dos esporos (Leverly *et al.*, 2002) enquanto que em *Penicillium digitatum* eram essenciais para a virulência (Zhu *et al.*, 2014). Em *Fusarium graminearum* estas enzimas eram importantes para a morfologia dos conídios e o desenvolvimento das hifas, sugerindo que apresentam diferentes papéis em diferentes fungos (Ramamoorthy *et al.*, 2007).

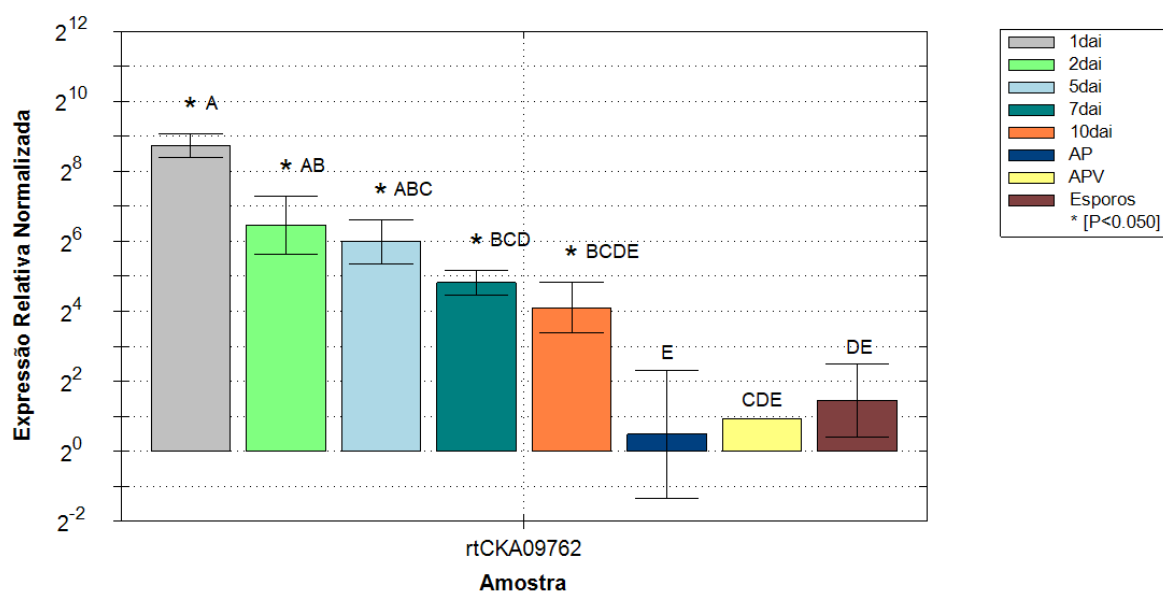


Figura 18 – Perfil de expressão génica obtido do gene CKA01_09762 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

Em trabalhos futuros deverá ser reavaliada a expressão nos apressórios de modo a confirmar a hipótese de que este gene é importante na adesão ao hospedeiro, na germinação e na formação do apressório.

iv. Peroxidases

As peroxidases (EC 1.11.1.7) são enzimas responsáveis pelo transporte de eletrões para diversos aceitadores, catalisando reações de oxidação/redução (Passardi *et al.*, 2007) e de acordo com diversos estudos, são capazes de interferir na ação dos ROS nos mecanismos de defesa das plantas.

O gene CKAH01_11318 codifica uma peroxidase e a sua expressão está representada na figura 19. De acordo com os resultados, podemos verificar que esta proteína apresenta uma expressão muito elevada e relativamente constante ao longo de todo o processo de infeção, incluindo nos apressórios *in vivo*, o que sugere a sua importância para o sucesso da infeção. Este gene está ativado nestas amostras quando comparadas com os apressórios *in vitro*, os esporos e o controlo e as diferenças de expressão são significativas.

As diferenças de expressão nos apressórios *in vivo* e *in vitro* é bastante significativa, sugerindo que este gene seja regulado em resultado do contacto com o hospedeiro. O papel do hospedeiro na regulação génica foi já referido por O'Connell *et al.* (2012) em *C.*

higginsianum, que observou inúmeras diferenças de expressão consoante as amostras eram obtidas *in vivo* ou *in vitro*.

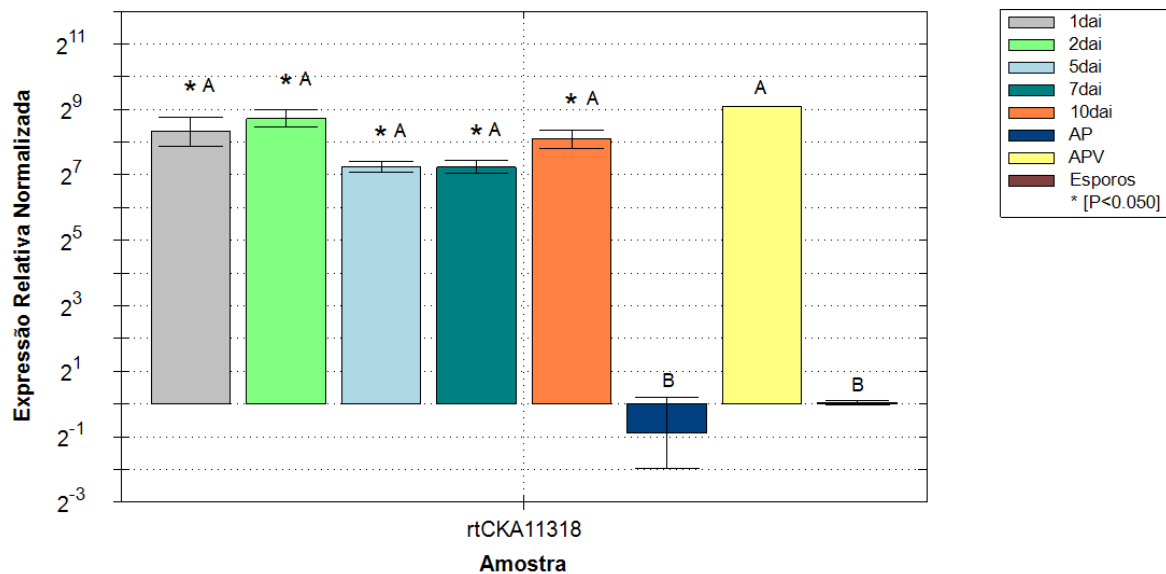


Figura 19 – Perfil de expressão génica obtido do gene CKA01_11318 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

Segundo o estudo realizado por Fernandez & Wilson (2014), em *Magnaporthe oryzae*, a deleção de três genes codificantes de diferentes peroxidases, permitiu a penetração do fungo nos tecidos do hospedeiro, mas não houve desenvolvimento de sintomas ou movimento do fungo entre células. Num outro estudo envolvendo o mesmo fungo, foram suprimidos 7 genes codificantes de peroxidases, sendo que um dos mutantes, $\Delta Mopr1$, apresentou um atraso no seu desenvolvimento, tendo dificuldade em penetrar as células do hospedeiro, indicando que as peroxidases podem ter uma ação na fase pós-penetração por parte do apressório. Além disso, foi testada ainda a sensibilidade deste mutante a H_2O_2 , que revelou a ação das peroxidases contra as ROS, durante as fases iniciais de infeção (Mir *et al.*, 2015).

As peroxidases têm funções muito diversas, sendo assim natural a existência de pelo menos 31 proteínas, no genoma de *C. kahawae*, previstas como peroxidases (resultados obtidos através do programa HMMScan). Várias destas peroxidases podem ter funções redundantes desta aqui estudada ou funções bastante diferentes, estando envolvidas noutros processos, nomeadamente em fases de desenvolvimento do fungo distintas.

v. **Proteínas contendo motivos de lisina (LysM)**

As fases iniciais de infecção são marcadas por um grande aumento na secreção de proteínas efectoras (O'Connell *et al.*, 2012), possivelmente envolvidas na supressão dos mecanismos de defesa das plantas (Stergiopoulos & de Wit, 2009). Nesta fase, em *C. higginsianum*, há uma elevada secreção de proteínas contendo motivos de lisina (*Lysin motifs* – LysM) (O'Connell *et al.*, 2012).

A quitina é um componente essencial da constituição da parede dos fungos e um dos mecanismos com que está relacionada é a resistência associada aos PAMP (PTI), por poder gerar reações de imunidade por parte do hospedeiro ao ser reconhecida pelos PRR. A maioria dos PRR conhecidos como estando envolvidos no reconhecimento da quitina são proteínas que contêm o domínio LysM, localizadas na membrana plasmática (Dodds & Rathjen, 2010; Takahara *et al.*, 2016).

Neste trabalho estudaram-se duas proteínas com domínios LysM, codificadas pelos genes CKAH01_15250 e CKAH01_15357, cuja expressão pode ser observada nas figuras 20 e 21, respetivamente. De acordo com essas figuras, os perfis de expressão, de ambas as proteínas, são bastante semelhantes, sugerindo que possam ter funções muito semelhantes.

Em *Cladosporium fulvum* foi identificada uma proteína, Ecp6, contendo o domínio LysM. Esta proteína é altamente expressa *in planta* e tem a capacidade de se ligar à quitina e impedir o reconhecimento da mesma por parte do hospedeiro, permitindo então a infecção por parte do fungo (Bolton *et al.*, 2008; de Jonge & Thomma, 2009). Outras proteínas do mesmo tipo já foram identificadas noutros patossistemas envolvendo os fungos *Magnaporthe oryzae* e *M. graminicola*, sendo que interferem também no reconhecimento da quitina por parte dos recetores de quitina do hospedeiro mas protegem ainda as hifas do fungos das quitinases excretadas pelas plantas (Marshall *et al.*, 2011; Mentlak *et al.*, 2012; Lee *et al.*, 2014). Para compreender a função das proteínas que contêm os domínios LysM no género *Colletotrichum* sp., Takahara *et al.* (2016) estudaram duas proteínas deste tipo em *C. higginsianum*, concluindo que estas são excretadas durante a fase biotrófica e apresentam um duplo papel, na supressão de respostas de imunidade desencadeadas pela quitina e na contribuição para a manutenção da função dos apressórios.

Ao contrário do que seria de esperar, neste trabalho, o pico de expressão foi aos 5 dai diminuindo ligeiramente aos 7 e 10 dai. Aos 1 e 2 dai os genes estão ativados, mas com uma expressão menor que aos 5, 7 e 10 dai. Tendo-se observado diferenças significativas na expressão ao longo do processo de infecção para o gene CKAH01_15357 ao contrário do que se observa para o gene CKAH01_15250. Nos esporos e apressórios *in vitro* a expressão está bastante reprimida em ambos os genes. Nos apressórios *in vivo* a expressão é bastante

semelhante à observada 1 dai. O perfil obtido para estes dois genes sugere o seu envolvimento no *switch* para a necrotrofia e ao seu estabelecimento.

De acordo com a bibliografia seria na fase biotrófica que estas proteínas seriam mais importantes para suprimir respostas de imunidade por parte da planta e auxiliar na função dos apressórios. Os resultados obtidos sugerem que para além dessa função as LysM possam estar relacionadas com o *switch* para a necrotrofia podendo, ou não, a sua função estar relacionada com a proteção do fungo contra os mecanismos de defesa da planta, nomeadamente impedindo o reconhecimento da quitina pelos PPR.

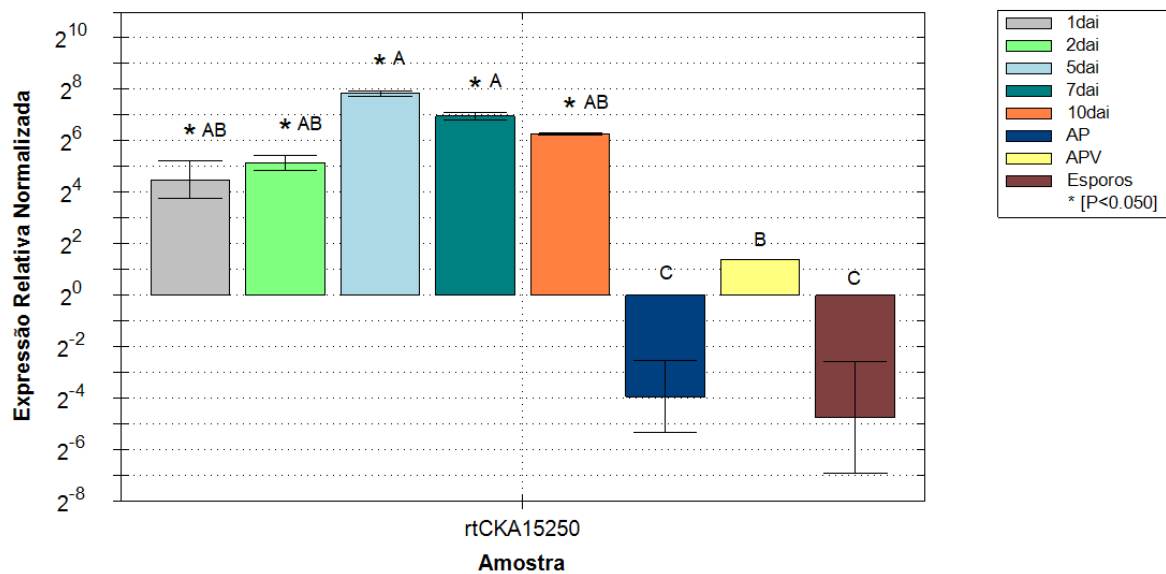


Figura 20 – Perfil de expressão génica obtido do gene CKA01_15250 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

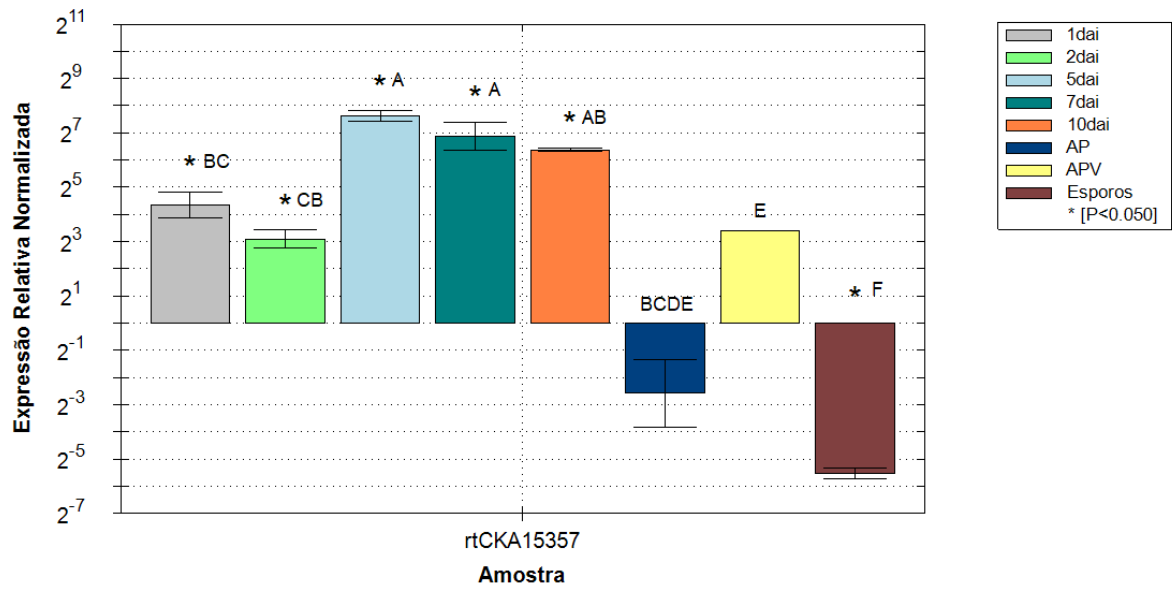


Figura 21 – Perfil de expressão génica obtido do gene CKA01_15357 para o isolado Que2 de *Colletotrichum kahawae* obtidos em diferentes dias do processo de infeção (1, 2, 5, 7 e 10 dias após inoculação) na interação com *Coffea arabica*, e em esporos, em apressórios *in vivo* (APV) e *in vitro* (AP). A normalização foi efetuada utilizando os genes PP1, Actina e ck34620, utilizando o micélio com três dias como controlo. As barras de erro representam o desvio padrão da média, os asteriscos representam diferenças de expressão significativas em relação ao controlo (Teste *t-student*) e as barras assinaladas com letras diferentes são significativamente diferentes para $p < 0,05$ (Teste de Tukey HSD).

V. Conclusão

Este trabalho teve como principal objetivo estudar a interação *Coffea arabica* - *Colletotrichum kahawae* e descobrir genes possivelmente envolvidos no estabelecimento da infecção e na transição para a necrotrofia. Depois de uma análise *in silico* detalhada foi possível selecionar um número reduzido de genes potencialmente interessantes. Esses genes foram escolhidos tendo em conta diferentes características anteriormente estudadas e associadas à interação hospedeiro – agente patogénico, desde serem proteínas efetoras a serem proteínas pertencentes a determinadas famílias importantes.

Foi então possível selecionar diversas proteínas interessantes que revelaram estar envolvidas no processo de infecção. Estas proteínas pertencem a famílias diferentes e, provavelmente, desempenham funções distintas e em diferentes fases do processo de infecção. Foi possível compreender que as cutinases são proteínas essencialmente importantes nas fases iniciais do estabelecimento da infecção, para penetrar a cutícula das plantas, enquanto que as pectato liases são ativadas numa fase mais avançada do processo de infecção com o objetivo de degradar a parede celular do hospedeiro. A ceramidase estudada apresenta-se como sendo especialmente importante no estabelecimento da infecção, mas o facto de ser expressa durante todo o processo de infecção indica que estas proteínas podem ter uma ação importante ao longo desse processo. Da mesma maneira, a peroxidase aqui em estudo apresenta-se expressa ao longo de todo o processo de infecção, sugerindo que estas podem ser proteínas com diferentes funções e importantes em diferentes fases da infecção. Por fim, as proteínas com domínios de lisina são expressas durante todo o processo de infecção, tendo o seu pico de expressão no início da fase necrotrofica, sugerindo que podem estar envolvidas no *switch* para a necrotrofia, fator que carece de estudos futuros, com amostras às 72 horas após inoculação, para se verificar a expressão durante a fase do *switch*.

Para concluir este trabalho é importante referir que o genoma em estudo ainda carece de trabalho futuro e seria interessante estudar os genes cujos primers desenhados não se mostraram adequados, mas cujas proteínas apresentam características interessantes para o estudo da patogenicidade de *C. kahawae*. É importante ainda recorrer à análise *in silico* aqui realizada no sentido de se selecionarem diferentes genes potencialmente interessantes, de famílias diferentes daquelas aqui estudadas ou das mesmas famílias, para estudar mais aprofundadamente as proteínas com domínios de lisina ou os metabolitos secundários, comparando, neste último caso, os resultados obtidos por ambos os *softwares* utilizados e combinar essa informação com outros dados como por exemplo os termos GO.

O estudo de novos e diferentes efetores apresenta-se como uma estratégia interessante para a compreensão desta interação uma vez que estas proteínas (efetoras) têm sido

descritas como importantes em diferentes agentes patogénicos, mas seria também interessante estudar outras proteínas com homologia em bases de dados como o PHI, uma vez que estas foram já referidas como importantes na interação hospedeiro – agente patogénico. Numa outra perspetiva, poderá ser interessante estudar proteínas que não apresentam qualquer homologia com outras bases de dados, no sentido de se identificarem efetores específicos da espécie em estudo e que tenham contribuído para a especiação ocorrida.

VI. Referências bibliográficas

- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. <https://doi.org/10.1038/s41587-019-0036-z>
- Andersen, C., Jensen, J., Orntoft, T., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64, 5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
- Annis, S.L., Goodwin, P.H., 1997. Recent advances in the molecular genetics of plant cell wall-degrading enzymes produced by plant pathogenic fungi. *Eur. J. Plant Pathol.* 103, 1–14. <https://doi.org/10.1023/A:1008656013255>
- António, J., 2008. O micro-crédito como ferramenta para o relançamento da cultura do café na região agrícola do Libolo e Amboim (Angola). Dissertação para a obtenção do Grau de Mestre em Economia Agrária e Sociologia Rural. Instituto Superior de Agronomia, Universidade de Técnica de Lisboa. 111 pp.
- Armenteros, J.J.A., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., Nielsen, H., 2019. Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* 2, 1–14. <https://doi.org/10.26508/lsa.201900429>
- Armenteros, J.J.A., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. <https://doi.org/10.1093/bioinformatics/btx431>
- Baroncelli, R., Amby, D.B., Zapparata, A., Sarrocco, S., Vannacci, G., le Floch, G., Harrison, R.J., Holub, E., Sukno, S.A., Sreenivasaprasad, S., Thon, M.R., 2016. Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. *BMC Genomics* 17, 1–17. <https://doi.org/10.1186/s12864-016-2917-6>
- Baroncelli, R., Talhinhos, P., Pensec, F., Sukno, S.A., le Floch, G., Thon, M.R., 2017. The *Colletotrichum acutatum* species complex as a model system to study evolution and host specialization in plant pathogens. *Front. Microbiol.* 8:2001. <https://doi.org/10.3389/fmicb.2017.02001>
- Batista, D., Silva, D.N., Vieira, A., Cabral, A., Pires, A.S., Loureiro, A., Guerra-Guimarães, L., Pereira, A.P., Azinheira, H., Talhinhos, P., Silva, M.D.C., Várzea, V., 2017. Legitimacy and implications of reducing *Colletotrichum kahawae* to subspecies in plant pathology.

Front. Plant Sci. 7, 2051. <https://doi.org/10.3389/fpls.2016.02051>

- Batista De Sá, M.F., 2010. Avaliação do papel dos ácidos salicílico, abscísico e jasmónico na interacção *Coffea arabica* – *Hemileia vastatrix*: implementação de um método de análise. Dissertação para a obtenção do Grau de Mestre em Engenharia Alimentar. Instituto Superior de Agronomia, Universidade de Técnica de Lisboa, Lisboa. 88 pp.
- Bazin, É., Mathé-Hubert, H., Facon, B., Carlier, J., Ravigné, V., 2014. The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. *Biol. Invasions* 16, 875–886. <https://doi.org/10.1007/s10530-013-0544-6>
- Bedimo, J., Bieysse, D., Nyasse, S., Nottéghem, J., Cilas, C., 2010. Role of rainfall in the development of Coffee Berry Disease in *Coffea arabica* caused by *Colletotrichum kahawae*, in Cameroon. *Plant Pathol.* 59, 324–329. <https://doi.org/10.1111/j.1365-3059.2009.02214.x>
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H., Weber, T., 2019. AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. <https://doi.org/10.1093/nar/gkz310>
- Bolton, M.D., van Esse, H.P., Vossen, J.H., de Jonge, R., Stergiopoulos, I., Stulemeijer, I.J.E., van den Berg, G.C.M., Borrás-Hidalgo, O., Dekker, H.L., de Koster, C.G., de Wit, P.J.G.M., Joosten, M.H.A.J., Thomma, B.P.H.J., 2008. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Mol. Microbiol.* 69, 119–136. <https://doi.org/10.1111/j.1365-2958.2008.06270.x>
- Bridge, P.D., Waller, J.M., Davies, D., Buddie, A.G., 2008. Variability of *Colletotrichum kahawae* in relation to other *Colletotrichum* species from tropical perennial crops and the development of diagnostic techniques. *J. Phytopathol.* 156, 274–280. <https://doi.org/10.1111/j.1439-0434.2007.01354.x>
- Brisdon, D., Vercourt, B., 1988. Flora of tropical East Africa: Rubiaceae, Part 2, in: Polhill, R.M. (Ed.), . CRC Press, pp. 703–723.
- Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. <https://doi.org/10.1373/clinchem.2008.112797>
- Cabral, A., Talhinas, P., Nascimento, T., Várzea, V., Silva, M.D.C., Azinheira, H.G., Oliveira, H., 2016. *Agrobacterium tumefaciens*-mediated transformation: an efficient tool for pathogenicity studies on *Colletotrichum kahawae*, in: 26th International Conference on Coffee Science. Association for Science and Information on Coffee, 13-19 November

2016, Kunming, China, p. 102.

- Canals, D., Perry, D.M., Jenkins, R.W., Hannun, Y.A., 2011. Drug targeting of sphingolipid metabolism: sphingomyelinases and ceramidases. *Br. J. Pharmacol.* 163, 694–712. <https://doi.org/10.1111/j.1476-5381.2011.01279.x>
- Carpita, N.C., Gibeaut, D.M., 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. <https://doi.org/10.1111/j.1365-313X.1993.tb00007.x>
- Chalfoun, N.R., Grellet-Bournonville, C.F., Martínez-Zamora, M.G., Díaz-Perales, A., Castagnaro, A.P., Díaz-Ricci, J.C., 2013. Purification and characterization of AsES protein: a subtilisin secreted by *Acremonium strictum* is a novel plant defense elicitor. *J. Biol. Chem.* 288, 14098–14113. <https://doi.org/10.1074/jbc.M112.429423>
- Chandrasekaran, M., Thangavelu, B., Chun, S.C., Sathiyabama, M., 2016. Proteases from phytopathogenic fungi and their importance in phytopathogenicity. *J. Gen. Plant Pathol.* 82, 233–239. <https://doi.org/10.1007/s10327-016-0672-9>
- Chen, Z.J., 2002. Morphocultural and pathogenic comparisons between *Colletotrichum kahawae* and *Colletotrichum gloeosporioides* isolated from coffee berries. Dissertação para Obtenção do Grau de Doutor em Engenharia Agronómica. Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa. 163 pp.
- Chen, Z., Franco, C.F., Baptista, R.P., Cabral, J.M.S., Coelho, A. V., Rodrigues, C.J., Melo, E.P., 2007. Purification and identification of cutinases from *Colletotrichum kahawae* and *Colletotrichum gloeosporioides*. *Appl. Microbiol. Biotechnol.* 73, 1306–1313. <https://doi.org/10.1007/s00253-006-0605-1>
- Chu, J., Li, W.F., Cheng, W., Lu, M., Zhou, K.H., Zhu, H.Q., Li, F.G., Zhou, C.Z., 2015. Comparative analyses of secreted proteins from the phytopathogenic fungus *Verticillium dahliae* in response to nitrogen starvation. *Biochim. Biophys. Acta - Proteins Proteomics* 1854, 437–448. <https://doi.org/10.1016/j.bbapap.2015.02.004>
- Clifford, M., Willson, K. (Eds.), 1985. Coffee: botany, biochemistry and production of beans and beverage. London, Croom Helm.
- Coleman, J.J., Wasmann, C.C., Usami, T., White, G.J., Temporini, E.D., McCluskey, K., van Etten, H.D., 2011. Characterization of the gene encoding pisatin demethylase (FoPDA1) in *Fusarium oxysporum*. *Mol. Plant-Microbe Interact.* 24, 1482–1491. <https://doi.org/10.1094/MPMI-05-11-0119>
- Cros, J., Combes, M.C., Trouslot, P., Anthony, F., Hamon, S., Charrier, A., Lashermes, P.,

1998. Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. *Mol. Phylogenet. Evol.* 9, 109–117. <https://doi.org/10.1006/mpev.1997.0453>
- Davis, A.P., Govaerts, R., Bridson, D.M., Stoffelen, P., 2006. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot. J. Linn. Soc.* 152, 465–512. <https://doi.org/10.1111/j.1095-8339.2006.00584.x>
- de Jonge, R., Bolton, M.D., Thomma, B.P.H.J., 2011. How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. *Curr. Opin. Plant Biol.* 14, 400–406. <https://doi.org/10.1016/j.pbi.2011.03.005>
- de Jonge, R., Thomma, B.P.H.J., 2009. Fungal LysM effectors: extinguishers of host immunity? *Trends Microbiol.* 17, 151–157. <https://doi.org/10.1016/j.tim.2009.01.002>
- de Jonge, R., van Esse, H.P., Maruthachalam, K., Bolton, M.D., Santhanam, P., Saber, M.K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., Thomma, B.P.H.J., 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5110–5115. <https://doi.org/10.1073/pnas.1119623109>
- Degani, O., Salman, H., Gepstein, S., Dosoretz, C.G., 2006. Synthesis and characterization of a new cutinase substrate, 4-nitrophenyl (16-methyl sulfone ester) hexadecanoate. *J. Biotechnol.* 121, 346–350. <https://doi.org/10.1016/j.jbiotec.2005.08.011>
- Derso, E., Waller, J.M., 2003. Variation among *Colletotrichum* isolates from diseased coffee berries in Ethiopia. *Crop Prot.* 22, 561–565. [https://doi.org/10.1016/S0261-2194\(02\)00191-6](https://doi.org/10.1016/S0261-2194(02)00191-6)
- Dickman, M.B., Patil, S.S., 1986. Cutinase deficient mutants of *Colletotrichum gloeosporioides* are nonpathogenic to papaya fruit. *Physiol. Mol. Plant Pathol.* 28, 235–242. [https://doi.org/10.1016/S0048-4059\(86\)80067-4](https://doi.org/10.1016/S0048-4059(86)80067-4)
- Dickman, M.B., Podila, G., Kolattukudy, P., 1989. Insertion of cutinase gene into a wound pathogen enables it to infect intact host. *Nature* 446–448. <https://doi.org/doi.org/10.1038/342446a0>
- Diniz, I., Figueiredo, A., Loureiro, A., Batista, D., Azinheira, H., Varzea, V., Pereira, A.P., Gichuru, E., Moncada, P., Guerra-Guimarães, L., Oliveira, H., Silva, M.D.C., 2017. A first insight into the involvement of phytohormones pathways in coffee resistance and susceptibility to *Colletotrichum kahawae*. *PLoS One* 12, 1–20. <https://doi.org/10.1371/journal.pone.0178159>
- Djamei, A., Kahmann, R., 2012. *Ustilago maydis*: dissecting the molecular interface between

- pathogen and plant. PLoS Pathog. 8. <https://doi.org/10.1371/journal.ppat.1002955>
- Dodds, P.N., Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. Nat. Rev. Genet. 11, 539–348. <https://doi.org/doi.org/10.1038/nrg2812>
- Dufresne, M., Perfect, S., Pellier, A.L., Bailey, J.A., Langin, T., 2000. A GAL4-like protein is involved in the switch between biotrophic and necrotrophic phases of the infection process of *Colletotrichum lindemuthianum* on common bean. Plant Cell 12, 1579–1589. <https://doi.org/10.1105/tpc.12.9.1579>
- Dutech, C., Labbé, F., Capdevielle, X., Lung-Escarmant, B., 2017. Genetic analysis reveals efficient sexual spore dispersal at a fine spatial scale in *Armillaria ostoyae*, the causal agent of root-rot disease in conifers. Fungal Biol. 121, 550–560. <https://doi.org/10.1016/j.funbio.2017.03.001>
- Fernandes, T.R., Segorbe, D., Prusky, D., Di Pietro, A., 2017. How alkalization drives fungal pathogenicity. PLoS Pathog. 13, 1–8. <https://doi.org/10.1371/journal.ppat.1006621>
- Fernandez, J., Wilson, R.A., 2014. Characterizing roles for the glutathione reductase, thioredoxin reductase and thioredoxin peroxidase-encoding genes of *Magnaporthe oryzae* during rice blast disease. PLoS One 9. <https://doi.org/10.1371/journal.pone.0087300>
- Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: The protein families database. Nucleic Acids Res. 42, 222–230. <https://doi.org/10.1093/nar/gkt1223>
- Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39, 29–37. <https://doi.org/10.1093/nar/gkr367>
- Fino, J., Figueiredo, A., Loureiro, A., Gichuru, E.K., Várzea, V., 2014. Transcriptional profiling of compatible and incompatible coffee - *Colletotrichum kahawae* interactions through RNA-Seq analysis. 25th Int. Conf. Coffee Sci. 42–46.
- Firman, I.D., 1964. Screening of coffee for resistance to Coffee Berry Disease. East African Agric. For. J. 29, 192–194. <https://doi.org/10.1080/00128325.1964.11661923>
- Fu, F.F., Hao, Z., Wang, P., Lu, Y., Xue, L.J., Wei, G., Tian, Y., Hu, B., Xu, H., Shi, J., Cheng, T., Wang, G., Yi, Y., Chen, J., 2020. Genome sequence and comparative analysis of *Colletotrichum gloeosporioides* isolated from liriodendron leaves. Phytopathology 110, 1260–1269. <https://doi.org/10.1094/PHYTO-12-19-0452-R>

- Gan, P., Ikeda, K., Irieda, H., Narusaka, M., O'Connell, R.J., Narusaka, Y., Takano, Y., Kubo, Y., Shirasu, K., 2013. Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi. *New Phytol.* 197, 1236–1249. <https://doi.org/10.1111/nph.12085>
- Gan, P., Narusaka, M., Kumakura, N., Tsushima, A., Takano, Y., Narusaka, Y., Shirasu, K., 2016. Genus-wide comparative genome analyses of *Colletotrichum* species reveal specific gene family losses and gains during adaptation to specific infection lifestyles. *Genome Biol. Evol.* 8, 1467–1481. <https://doi.org/10.1093/gbe/evw089>
- Garcia, I.V.A.N., 1999. Histologia e ultra-estrutura do processo de infecção de *Colletotrichum kahawae* e *Colletotrichum gloesporioides* em *Coffea arabica*. Dissertação para a obtenção do Grau de Mestre em Engenharia Agronómica. Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa. 96 pp.
- Gibbs, J.N., 1969. Inoculum sources for Coffee Berry Disease. *Ann. Appl. Biol.* 64, 515–522. <https://doi.org/10.1111/j.1744-7348.1969.tb02900.x>
- Giddisa, G., 2016. A review on the status of Coffee Berry Disease (*Colletotrichum kahawae*) in Ethiopia. *J. Biol. Agric. Healthc.* 6, 140–151.
- Gilbert, H.J., Knox, J.P., Boraston, A.B., 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* 23, 669–677. <https://doi.org/10.1016/j.sbi.2013.05.005>
- Gíslason, M.H., Nielsen, H., Almagro Armenteros, J.J., Johansen, A.R., 2019. Prediction of GPI-Anchored proteins with pointer neural networks. <https://doi.org/10.1101/838680>
- Greenberg, J.T., Yao, N., 2004. The role of regulation of programmed cell death in plant-pathogen interactions. *Cell. Microbiol.* 6, 201–211. <https://doi.org/10.1111/j.1462-5822.2004.00361.x>
- Hacquard, S., Kracher, B., Hiruma, K., Münch, P.C., Garrido-Oter, R., Thon, M.R., Weimann, A., Damm, U., Dallery, J.F., Hainaut, M., Henrissat, B., Lespinet, O., Sacristán, S., van Themaat, E. V. L., Kemen, E., McHardy, A.C., Schulze-Lefert, P., O'Connell, R.J., 2016. Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. *Nat. Commun.* 7. <https://doi.org/10.1038/ncomms11362>
- Hansen, Z.R., Everts, K.L., Fry, W.E., Gevens, A.J., Grünwald, N.J., Gugino, B.K., Johnson, D.A., Johnson, S.B., Judelson, H.S., Knaus, B.J., McGrath, M.T., Myers, K.L., Ristaino, J.B., Roberts, P.D., Secor, G.A., Smart, C.D., 2016. Genetic variation within clonal lineages of *Phytophthora infestans* revealed through genotyping-by-sequencing, and implications for late blight epidemiology. *PLoS One* 11, 1–22.

<https://doi.org/10.1371/journal.pone.0165690>

- Hindorf, H., 1974. *Colletotrichum*-Arten aus den Kaffeeanbaugebieten von Kiambu in Kenia. Z. Pflanzenkr. Pflanzenschutz 81, 108–113.
- Hindorf, H., 1970. *Colletotrichum* spp. isolated from *Coffea arabica* L. in Kenya. Z. Pflanzenkr. Pflanzenschutz 77, 328–331. <https://doi.org/stable/43215857>
- Hindorf, H., Omondi, C., 2011. A review of three major fungal diseases of *Coffea arabica* L. in the rainforests of Ethiopia and progress in breeding for resistance in Kenya. J. Adv. Res. 2, 109–120. <https://doi.org/10.1016/j.jare.2010.08.006>
- Hiruma, K., Gerlach, N., Sacristán, S., Nakano, R.T., Hacquard, S., Kracher, B., Neumann, U., Ramírez, D., Bucher, M., O'Connell, R.J., Schulze-Lefert, P., 2016. Root endophyte *Colletotrichum tofieldiae* confers plant fitness benefits that are phosphate status dependent. Cell 165, 464–474. <https://doi.org/10.1016/j.cell.2016.02.028>
- Hoffmeister, D., Keller, N.P., 2007. Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat. Prod. Rep. 24, 393–416. <https://doi.org/10.1039/b603084j>
- Hogenhout, S.A., van der Hoorn, R.A.L., Terauchi, R., Kamoun, S., 2009. Emerging concepts in effector biology of plant-associated organisms. Mol. Plant-Microbe Interact. 22, 115–122. <https://doi.org/10.1094/MPMI-22-2-0115>
- Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: Protein localization predictor. Nucleic Acids Res. 35, 585–587. <https://doi.org/10.1093/nar/gkm259>
- Huang, Y., Li, B., Yin, J., Yang, Q., Sheng, O., Prusky, D.B., 2019. CgGCS , encoding a glucosylceramide synthase, is required for growth, conidiation and pathogenicity in *Colletotrichum gloeosporioides*. Front. Microbiol. 10, 1–12. <https://doi.org/10.3389/fmicb.2019.01016>
- Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 47, 309–314. <https://doi.org/10.1093/nar/gky1085>
- Hulo, N., 2006. The PROSITE database. Nucleic Acids Res. 34, D227–D230. <https://doi.org/10.1093/nar/gkj063>
- ICO, 2019. Coffee Market Report. <http://www.ico.org/>
- Inglis, D.O., Binkley, J., Skrzypek, M.S., Arnaud, M.B., Cerqueira, G.C., Shah, P., Wymore, F.,

- Wortman, J.R., Sherlock, G., 2013. Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. BMC Microbiol. 13, 1471–2180. <https://doi.org/10.1186/1471-2180-13-91>
- Jones, J.D.G., Dangl, J.L., 2006. The plant immune system. Nature 444, 323–329. <https://doi.org/doi.org/10.1038/nature05286>
- Käll, L., Krogh, A., Sonnhammer, E.L.L., 2004. A combined transmembrane topology and signal peptide prediction method. J. Mol. Biol. 388, 1027–1036. <https://doi.org/doi.org/10.1016/j.jmb.2004.03.016>
- Kamoun, S., 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. Annu. Rev. Phytopathol. 44, 41–60. <https://doi.org/10.1146/annurev.phyto.44.070505.143436>
- Kanehisa, M., Sato, Y., 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci. 29, 28–35. <https://doi.org/10.1002/pro.3711>
- Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J. Mol. Biol. 428, 726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>
- Khalidi, N., Seifuddin, F.T., Turner, G., Haft, D., Nierman, W.C., Wolfe, K.H., Fedorova, N.D., 2010. SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet. Biol. 47, 736–741. <https://doi.org/10.1016/j.fgb.2010.06.003>
- Kleemann, J., Rincon-Rivera, L.J., Takahara, H., Neumann, U., van Themaat, E.V.L., van der Does, H.C., Hacquard, S., Stüber, K., Will, I., Schmalenbach, W., Schmelzer, E., O’Connell, R.J., 2012. Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. PLoS Pathog. 8. <https://doi.org/10.1371/journal.ppat.1002643>
- Kleemann, J., Takahara, H., Stüber, K., O’Connell, R., 2008. Identification of soluble secreted proteins from appressoria of *Colletotrichum higginsianum* by analysis of expressed sequence tags. Microbiology 154, 1204–1217. <https://doi.org/10.1099/mic.0.2007/014944-0>
- Kolattukudy, P.E., Rogers, L.M., Li, D., Hwang, C.S., Flaishman, M.A., 1995. Surface signaling in pathogenesis. Proc. Natl. Acad. Sci. U. S. A. 92, 4080–4087. <https://doi.org/10.1073/pnas.92.10.4080>
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J. Mol.

- Biol. 305, 567–580. <https://doi.org/10.1006/jmbi.2000.4315>
- Lashermes, P., Combes, M.C., Robert, J., Trouslot, P., D'Hont, A., Anthony, F., Charrier, A., 1999. Molecular characterisation and origin of the *Coffea arabica* L. Genome. Mol. Gen. Genet. 261, 259–266. <https://doi.org/10.1007/s004380050965>
- Lee, W.S., Rudd, J.J., Hammond-Kosack, K.E., Kanyuka, K., 2014. *Mycosphaerella graminicola* lysM effector-mediated stealth pathogenesis subverts recognition through both cerk1 and cebip homologues in wheat. Mol. Plant-Microbe Interact. 27, 236–243. <https://doi.org/10.1094/MPMI-07-13-0201-R>
- Leverly, S.B., Momany, M., Lindsey, R., Toledo, M.S., Shayman, J.A., Fuller, M., Brooks, K., Doong, R.L., Straus, A.H., Takahashi, H.K., 2002. Disruption of the glucosylceramide biosynthetic pathway in *Aspergillus nidulans* and *Aspergillus fumigatus* by inhibitors of UDP-Glc:ceramide glucosyltransferase strongly affects spore germination, cell cycle, and hyphal growth. FEBS Lett. 525, 59–64. [https://doi.org/https://doi.org/10.1016/s0014-5793\(02\)03067-3](https://doi.org/https://doi.org/10.1016/s0014-5793(02)03067-3)
- Liang, X., Cao, M., Li, S., Kong, Y., Rollins, J.A., Zhang, R., Sun, G., 2020. Highly contiguous genome resource of *Colletotrichum fructicola* generated using long-read sequencing. Mol. Plant-Microbe Interact. 33, 790–793. <https://doi.org/10.1094/MPMI-11-19-0316-A>
- Lin, F., Jiang, L., Liu, Y., Lv, Y., Dai, H., Zhao, H., 2014. Genome-wide identification of housekeeping genes in maize. Plant Mol. Biol. 86, 543–554. <https://doi.org/10.1007/s11103-014-0246-1>
- Loureiro, A., Azinheira, H.G., Silva, M.D.C., Talhinhos, P., 2015. A method for obtaining RNA from *Hemileia vastatrix* appressoria produced in planta, suitable for transcriptomic analyses. Fungal Biol. 119, 1093–1099. <https://doi.org/10.1016/j.funbio.2015.08.008>
- Loureiro, A., Guerra-Guimarães, L., Lidon, F.C., Bertrand, B., Silva, M.D.C., Várzea, V., 2011. Isoenzymatic characterization of *Colletotrichum kahawae* isolates with different levels of aggressiveness. Trop. Plant Pathol. 36, 287–298. <https://doi.org/10.1590/S1982-56762011000500003>
- Loureiro, A., Nicole, M.R., Várzea, V., Moncada, P., Bertrand, B., Silva, M.D.C., 2012. Coffee resistance to *Colletotrichum kahawae* is associated with lignification, accumulation of phenols and cell death at infection sites. Physiol. Mol. Plant Pathol. 77, 23–32. <https://doi.org/10.1016/j.pmpp.2011.11.002>
- Loureiro, A.S.B.C., 2008. Caracterização da resistência à antracnose do frutos verdes do cafeeiro. Dissertação apresentada para obtenção do Grau de Doutor em Biologia, especialidade Biologia Vegetal, Faculdade de Ciências e Tecnologia, Universidade de

Lisboa, Lisboa. 173 pp.

- Lutzeyer, H., 1993. Neue Erkenntnisse über pflanzenschutz in plantagenkulturen dargestellt am beispiel kaffee. Weltforumverlag, Koln.
- Lyu, X., Shen, C., Fu, Y., Xie, J., Jiang, D., Li, G., Cheng, J., 2015. Comparative genomic and transcriptional analyses of the carbohydrate-active enzymes and secretomes of phytopathogenic fungi reveal their significant roles during infection and development. *Sci. Rep.* 5, 1–16. <https://doi.org/10.1038/srep15565>
- Macedo, F.W.V.P., 1994. A antracnose dos frutos do cafeeiro induzida pelo fungo *Colletotrichum coffeanum* Noack. Relatório Final de Estágio. Universidade de Trás-os-Montes e Alto Douro, Vila Real. 41 pp.
- Mao, C., Obeid, L.M., 2002. Ceramidases: regulators of turnover of ceramide and ceramide-mediated responses, in: Futerman, A.H. (Ed.), *Ceramide Signaling*. Landes Bioscience, Georgetown, Texas, pp. 29–40.
- Marshall, O.J., 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 20, 2471–2472. <https://doi.org/10.1093/bioinformatics/bth254>
- Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K.E., Thomma, B.P.H.J., Rudd, J.J., 2011. Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiol.* 156, 756–769. <https://doi.org/10.1104/pp.111.176347>
- Masaba, D.M., Waller, J.M., 1992. Coffee Berry Disease: the current status, in: Bailey, J.A., Jeger, M.J. (Eds.), *Colletotrichum: Biology, Pathology and Control*. CAB international, Wallingford, UK, pp. 237–249.
- McDonald, J., 1926. A preliminary account of a disease of green coffee berries in Kenya Colony. *Trans. Br. Mycol. Soc.* 11, 145–154. [https://doi.org/10.1016/s0007-1536\(26\)80033-6](https://doi.org/10.1016/s0007-1536(26)80033-6)
- Mentlak, T.A., Kombrink, A., Shinya, T., Ryder, L.S., Otomo, I., Saitoh, H., Terauchi, R., Nishizawa, Y., Shibuya, N., Thomma, B.P.H.J., Talbot, N.J., 2012. Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. *Plant Cell* 24, 322–335. <https://doi.org/10.1105/tpc.111.092957>
- Mir, A.A., Park, S.Y., Sadat, M.A., Kim, S., Choi, J., Jeon, J., Lee, Y.H., 2015. Systematic characterization of the peroxidase gene family provides new insights into fungal

- pathogenicity in *Magnaporthe oryzae*. *Sci. Rep.* 5, 1–14.
<https://doi.org/10.1038/srep11831>
- Miyara, I., Shafran, H., Davidzon, M., Sherman, A., Prusky, D., 2010a. pH regulation of ammonia secretion by *Colletotrichum gloeosporioides* and its effect on appressorium formation and pathogenicity. *Mol. Plant-Microbe Interact.* 23, 304–316.
- Miyara, I., Shafran, H., Davidzon, M., Sherman, A., Prusky, D., 2010b. PH regulation of ammonia secretion by *Colletotrichum gloeosporioides* and its effect on appressorium formation and pathogenicity. *Mol. Plant-Microbe Interact.* 23, 304–316.
<https://doi.org/10.1094/MPMI-23-3-0304>
- Mosquera, G., Giraldo, M.C., Khang, C.H., Coughlan, S., Valent, B., 2009. Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy-associated secreted proteins in rice blast disease. *Plant Cell* 21, 1273–1290.
<https://doi.org/10.1105/tpc.107.055228>
- Muller, R.A., 1964. L'antracnose des baies du caféier d'arabie (*Coffea arabica*) due à *Colletotrichum coffeanum* Noack au Cameroun. *Inst. Français du Café du Cacao. Bull.* n°6.
- Muller, R.A., 1980. Contribution à la connaissance de la phytomycocénose constituée par *Coffea arabica* L., *Colletotrichum coffeanum* Noack (*sensu* Hindorf), *Hemileia vastatrix* Ber. et Br., *Hemileia coffeicola* Maublanc et Roger. *Inst. Français du Café du Cacao. Bull.* n°15.
- Nutman, F. J., Roberts, F.M., 1960. Investigations on a disease of *Coffea arabica* caused by a form of *Colletotrichum coffeanum* Noack: I. Some factors affecting infection by the pathogen. *Trans. Br. Mycol. Soc.* 43, 489–505. [https://doi.org/10.1016/S0007-1536\(60\)80031-9](https://doi.org/10.1016/S0007-1536(60)80031-9)
- Nutman, F.J., Roberts, F.M., 1960. Investigations on a disease of *Coffea arabica* caused by a form of *Colletotrichum coffeanum* Noack. II. Some factors affecting germination and infection and their relation to disease distribution. *Trans. Br. Mycol. Soc.* 43, 643–659.
[https://doi.org/10.1016/S0007-1536\(60\)80055-1](https://doi.org/10.1016/S0007-1536(60)80055-1)
- O'Connell, R.J., Thon, M.R., Hacquard, S., Amyotte, S.G., Kleemann, J., Torres, M.F., Damm, U., Buiate, E.A., Epstein, L., Alkan, N., Altmüller, J., Alvarado-Balderrama, L., Bauser, C.A., Becker, C., Birren, B.W., Chen, Z., Choi, J., Crouch, J.A., Duvick, J.P., Farman, M.A., Gan, P., Heiman, D., Henrissat, B., Howard, R.J., Kabbage, M., Koch, C., Kracher, B., Kubo, Y., Law, A.D., Lebrun, M.H., Lee, Y.H., Miyara, I., Moore, N., Neumann, U., Nordström, K., Panaccione, D.G., Panstruga, R., Place, M., Proctor, R.H., Prusky, D.,

- Rech, G., Reinhardt, R., Rollins, J.A., Rounsley, S., Schardl, C.L., Schwartz, D.C., Shenoy, N., Shirasu, K., Sikhakolli, U.R., Stüber, K., Sukno, S.A., Sweigard, J.A., Takano, Y., Takahara, H., Trail, F., van der Does, H.C., Voll, L.M., Will, I., Young, S., Zeng, Q., Zhang, J., Zhou, S., Dickman, M.B., Schulze-Lefert, P., van Themaat, E. V. L., Ma, L.J., Vaillancourt, L.J., 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nat. Genet.* 44, 1060–1065. <https://doi.org/10.1038/ng.2372>
- Olivieri, F.P., Maldonado, S., Tonón, C. V., Casalongué, C.A., 2004. Hydrolytic activities of *Fusarium solani* and *Fusarium solani* f. sp. *eumartii* associated with the infection process of potato tubers. *J. Phytopathol.* 152, 337–344. <https://doi.org/10.1111/j.1439-0434.2004.00851.x>
- Passardi, F., Zamocky, M., Cosio, C., Rouhier, N., Teixeira, F., Margis-pinheiro, M., Ioannidis, V., Penel, C., Falquet, L., Dunand, C., 2007. PeroxiBase: the peroxidase database. *Phytochemistry* 68, 1605–1611. <https://doi.org/10.1016/j.phytochem.2007.04.005>
- Pata, M.O., Wu, B.X., Bielawski, J., Xiong, T.C., Hannun, Y.A., Science, E., 2008. Molecular cloning and characterization of OsCDase, a ceramidase enzyme from rice. *Plant J.* 1000–1009. <https://doi.org/10.1111/j.1365-313X.2008.03569.x>
- Petre, B., Joly, D.L., Duplessis, S., 2014. Effector proteins of rust fungi. *Front. Plant Sci.* 5, 1–7. <https://doi.org/10.3389/fpls.2014.00416>
- Pierleoni, A., Martelli, P., Casadio, R., 2008. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* 9, 1–11. <https://doi.org/10.1186/1471-2105-9-392>
- Podila, G.K., Rosen, E., San Francisco, M.J., Kolattukudy, P.E., 1995. Targeted secretion of cutinase in *Fusarium solani* f.sp. *pisi* and *Colletotrichum gloesporioides*. *Phytopathology*. <https://doi.org/10.1094/Phyto-85-238>
- Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D., 2018. HMMER web server: 2018 update. *Nucleic Acids Res.* 46, 200–204. <https://doi.org/10.1093/nar/gky448>
- Presti, L.L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., Kahmann, R., 2015. Fungal effectors and plant susceptibility. *Annu. Rev. Phytopathol.* 66, 513–545. <https://doi.org/doi.org/10.1146/annurev-arplant-043014-114623>
- Prusky, D., McEvoy, J.L., Leverentz, B., Conway, W.S., 2001. Local modulation of host pH by *Colletotrichum* species as a mechanism to increase virulence. *Mol. Plant-Microbe Interact.* 14, 1105–1113. <https://doi.org/10.1094/MPMI.2001.14.9.1105>

- Rafiqi, M., Ellis, J.G., Ludowici, V.A., Hardham, A.R., Dodds, P.N., 2012. Challenges and progress towards understanding the role of effectors in plant-fungal interactions. *Curr. Opin. Plant Biol.* 15, 477–482. <https://doi.org/10.1016/j.pbi.2012.05.003>
- Ramamoorthy, V., Cahoon, E.B., Li, J., Thokala, M., Minto, R.E., Shah, D.M., 2007. Glucosylceramide synthase is essential for alfalfa defensin-mediated growth inhibition but not for pathogenicity of *Fusarium graminearum*. *Mol. Microbiol.* 66, 771–786. <https://doi.org/10.1111/j.1365-2958.2007.05955.x>
- Rao, S., Nandineni, M.R., 2017. Genome sequencing and comparative genomics reveal a repertoire of putative pathogenicity genes in chilli anthracnose fungus *Colletotrichum truncatum*. *PLoS One* 12, 1–31. <https://doi.org/10.1371/journal.pone.0183567>
- Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., Finn, R.D., 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* 46, D624–D632. <https://doi.org/10.1093/nar/gkx1134>
- Rayner, R.W., 1952. Coffee Berry Disease - a survey of investigations carried out up to 1950. *East African Agric. J.* 17, 130–158. <https://doi.org/10.1080/03670074.1952.11664802>
- Reynolds, S.M., Käll, L., Riffle, M.E., Bilmes, J.A., Noble, W.S., 2008. Transmembrane topology and signal peptide prediction using dynamic Bayesian networks. *PLoS Comput. Biol.* 4. <https://doi.org/10.1371/journal.pcbi.1000213>
- Righi, C.A., 2005. Avaliação ecofisiológica do cafeeiro (*Coffea arabica* L.) em sistema agroflorestal e em monocultivo. Tese apresentada para obtenção do título de Doutor em Agronomia, Área de Concentração: Fitotecnia. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, São Paulo, Brasil. 101 pp.
- Ruijter, J.M., Ramakers, C., Hoogaars, W.M.H., Karlen, Y., Bakker, O., van den Hoff, M.J.B., Moorman, A.F.M., 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37. <https://doi.org/10.1093/nar/gkp045>
- Sanz-Martín, J.M., Pacheco-Arjona, J.R., Bello-Rico, V., Vargas, W.A., Monod, M., Díaz-Mínguez, J.M., Thon, M.R., Sukno, S.A., 2016. A highly conserved metalloprotease effector enhances virulence in the maize anthracnose fungus *Colletotrichum graminicola*. *Mol. Plant Pathol.* 17, 1048–1062. <https://doi.org/10.1111/mpp.12347>
- Selin, C., de Kievit, T.R., Belmonte, M.F., Fernando, W.G.D., 2016. Elucidating the role of effectors in plant-fungal interactions: progress and challenges. *Front. Microbiol.* 7, 1–21. <https://doi.org/10.3389/fmicb.2016.00600>

- Silva, D.N., Talhinhos, P., Cai, L., Manuel, L., Gichuru, E.K., Loureiro, A., Várzea, V., Paulo, O.S., Batista, D., 2012. Host-jump drives rapid and recent ecological speciation of the emergent fungal pathogen *Colletotrichum kahawae*. *Mol. Ecol.* 21, 2655–2670. <https://doi.org/10.1111/j.1365-294X.2012.05557.x>
- Silva, M.D.C., Várzea, V., 2006. Coffee Berry Disease (*Colletotrichum kahawae*). PaDIL. Updated on 10/01/2006. <https://www.padil.gov.au/pests-and-diseases/pest/main/136610>
- Silva, M.D.C., Várzea, V., Guerra-Guimarães, L., Azinheira, H.G., Fernandez, D., Petitot, A.S., Bertrand, B., Lashermes, P., Nicole, M., 2006. Coffee resistance to the main diseases: Leaf rust and Coffee Berry Disease. *Brazilian J. Plant Physiol.* 18, 119–147. <https://doi.org/10.1016/j.ijepes.2005.11.005>
- Soberanes-Gutiérrez, C. V., Juárez-Montiel, M., Olgún-Rodríguez, O., Hernández-Rodríguez, C., Ruiz-Herrera, J., Villa-Tanaca, L., 2015. The pep4 gene encoding proteinase A is involved in dimorphism and pathogenesis of *Ustilago maydis*. *Mol. Plant Pathol.* 16, 837–846. <https://doi.org/10.1111/mpp.12240>
- Sondah, M.R., Baumann, T.W., 2001. Agronomy II: Developmental and Cell Biology, in: Clarke, R., Vitzthum, O. (Eds.), *Coffee Recent Developments*. Malden, MA: Blackwell Science, pp. 203–223. <https://doi.org/10.1002/9780470690499.ch10>
- Sonnhammer, E.L.L., Krogh, A., 1998. A hidden markov model for predicting transmembrane helices in protein sequence, in: Glasgow, J., Littlejohn, T., Major, F., Lathrop, R. (Eds.), *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*. AAAI Press, Menlo Park, California, pp. 175–182.
- Sperschneider, J., Dodds, P.N., Gardiner, D.M., Singh, K.B., Taylor, J.M., 2018. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.* 19, 2094–2110. <https://doi.org/10.1111/mpp.12682>
- Sperschneider, J., Gardiner, D.M., Dodds, P.N., Tini, F., Covarelli, L., Singh, K.B., Manners, J.M., Taylor, J.M., 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytol.* 210, 743–761. <https://doi.org/10.1111/nph.13794>
- Stergiopoulos, I., de Wit, P., 2009. Fungal effector proteins. *Annu. Rev. Phytopathol.* 47, 233–263. <https://doi.org/10.1146/annurev.phyto.112408.132637>
- Takahara, H., Kombrink, A., Hughes, H.B., Halder, V., Robin, G.P., Hiruma, K., Neumann, U., Shinya, T., Kombrink, E., Shibuya, N., Thomma, B.P.H.J., Connell, R.J.O., 2016. *Colletotrichum higginsianum* extracellular LysM proteins play dual roles in appressorial function and suppression of chitin-triggered plant immunity. *New Phytol.* 211, 1323–1337.

<https://doi.org/10.1111/nph.13994>

- Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S.Y., de Silva, N., Martinez, M.C., Pedro, H., Yates, A.D., Hassani-Pak, K., Hammond-Kosack, K.E., 2020. PHI-base: the pathogen-host interactions database. *Nucleic Acids Res.* 48, 613–620. <https://doi.org/10.1093/nar/gkz904>
- van der Graaff, N.A., 1981. Selection of Arabica coffee types resistance to Coffee Berry Disease in Ethiopia. Meded. Landbouwhogeschool wageningen.
- van der Vossen, H.A.M., Cook, R.T.A., Murakaru, G.N.W., 1976. Breeding for resistance to Coffee Berry Disease caused by *Colletotrichum coffeanum* Noack (*sensu* Hindorf) in *Coffea arabica* L. I. Methods of preselection for resistance. *Euphytica* 25, 733–745. <https://doi.org/10.1007/BF00041613>
- van der Vossen, H.A.M., Walyaro, D.J., 2009. Additional evidence for oligogenic inheritance of durable host resistance to Coffee Berry Disease (*Colletotrichum kahawae*) in arabica coffee (*Coffea arabica* L.). *Euphytica* 165, 105–111. <https://doi.org/10.1007/s10681-008-9769-3>
- Vandesompele, J., de Preter, K., Pattyn, F., Poppe, B., van Roy, N., de Paepe, A., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3. <https://doi.org/10.1186/gb-2002-3-7-research0034>
- Várzea, V.M.P., 1990. A antracnose dos frutos do cafeeiro causada pelo fungo *Colletotrichum coffeanum*. Trabalho de síntese apresentado ao Instituto de Investigação Científica Tropical para acesso à categoria de assistente de investigação. Instituto de Investigação Científica Tropical. Centro de Investigação das Ferrugens do Cafeeiro, 32pp.
- Várzea, V.M.P., Silva, M.C.M.L., Rodrigues Jr., C.J., 2002. Resistência do cafeeiro à antracnose dos frutos verdes, in: Zambolim, L. (Ed.), *O Estado de Arte de Tecnologias Na Produção de Café*. Universidade Federal Viçosa, Brasil, pp. 321–366.
- Vieira, A.C.M., 2018. Unraveling the genomics of adaptation of *Colletotrichum kahawae* to *Coffea arabica*. Dissertação para obtenção do grau de Doutor em Biologia, Especialidade de Biologia Evolutiva, Faculdade de Ciências da Universidade de Lisboa, Lisboa. 187 p.
- Vieira, A.C.M., Cabral, A., Fino, J., Azinheira, H.G., Loureiro, A., Talhinhos, P., Pires, A.S., Varzea, V., Moncada, P., Oliveira, H., Silva, M.D.C., Paulo, O.S., Batista, D., 2016. Comparative validation of conventional and RNA-seq data-derived reference genes for qPCR expression studies of *Colletotrichum kahawae*. *PLoS One* 11, 1–18. <https://doi.org/10.1371/journal.pone.0150651>

- Vieira, A.C.M., Silva, D.N., Várzea, V., Paulo, O.S., Batista, D., 2018. Novel insights on colonization routes and evolutionary potential of *Colletotrichum kahawae*, a severe pathogen of *Coffea arabica*. *Mol. Plant Pathol.* 19, 2488–2501. <https://doi.org/10.1111/mpp.12726>
- Vleeshouwers, V.G.A.A., Rietman, H., Krenek, P., Champouret, N., Young, C., Oh, S.K., Wang, M., Bouwmeester, K., Vosman, B., Visser, R.G.F., Jacobsen, E., Govers, F., Kamoun, S., van der Vossen, E.A.G., 2008. Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora Infestans* avirulence genes. *PLoS One* 3. <https://doi.org/10.1371/journal.pone.0002875>
- Waller, J.M., Bridge, P.D., Black, R., Hakiza, G., 1993. Characterization of the Coffee Berry Disease, *Colletotrichum kahawae* sp. nov. *Mycol. Res.* 97, 989–994. [https://doi.org/10.1016/S0953-7562\(09\)80867-8](https://doi.org/10.1016/S0953-7562(09)80867-8)
- Wattad, C., Freeman, S., Dinooor, A., Prusky, D., 1995a. A nonpathogenic mutant of *Colletotrichum magna* is deficient in extracellular secretion of pectate lyase. *Mol. Plant-Microbe Interact.* 8, 621–626. <https://doi.org/10.1094/MPMI-8-0621>
- Wattad, C., Freeman, S., Dinooor, A., Prusky, D., 1995b. A nonpathogenic mutant of *Colletotrichum magna* is deficient in extracellular secretion of pectate lyase. *Mol. Plant-Microbe Interact.* 8, 621–626. <https://doi.org/10.1094/MPMI-8-0621>
- Wattad, C., Kobilier, D., Dinooor, A., Prusky, D., 1997. Pectate lyase of *Colletotrichum gloeosporioides* attacking avocado fruits: cDNA cloning and involvement in pathogenicity. *Physiol. Mol. Plant Pathol.* 197–212. <https://doi.org/doi.org/10.1006/pmpp.1997.0080>
- Weir, W., Capewell, P., Foth, B., Clucas, C., Pountain, A., Steketee, P., Veitch, N., Koffi, M., de Meeûs, T., Kaboré, J., Camara, M., Cooper, A., Tait, A., Jamonneau, V., Bucheton, B., Berriman, M., Macleod, A., 2016. Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *Elife* 5, 1–14. <https://doi.org/10.7554/eLife.11473>
- Yakoby, N., Freeman, S., Dinooor, A., Keen, N.T., Prusky, D., 2000a. Expression of pectate lyase from *Colletotrichum gloeosporioides* in *C. magna* promotes pathogenicity. *Mol. Plant-Microbe Interact.* 13, 887–891. <https://doi.org/10.1094/MPMI.2000.13.8.887>
- Yakoby, N., Kobilier, I., Dinooor, A., Prusky, D., 2000b. pH regulation of pectate lyase secretion modulates the attack of *Colletotrichum gloeosporioides* on avocado fruits. *Appl. Environ. Microbiol.* 66, 1026–1030. <https://doi.org/10.1128/AEM.66.3.1026-1030.2000>
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC*

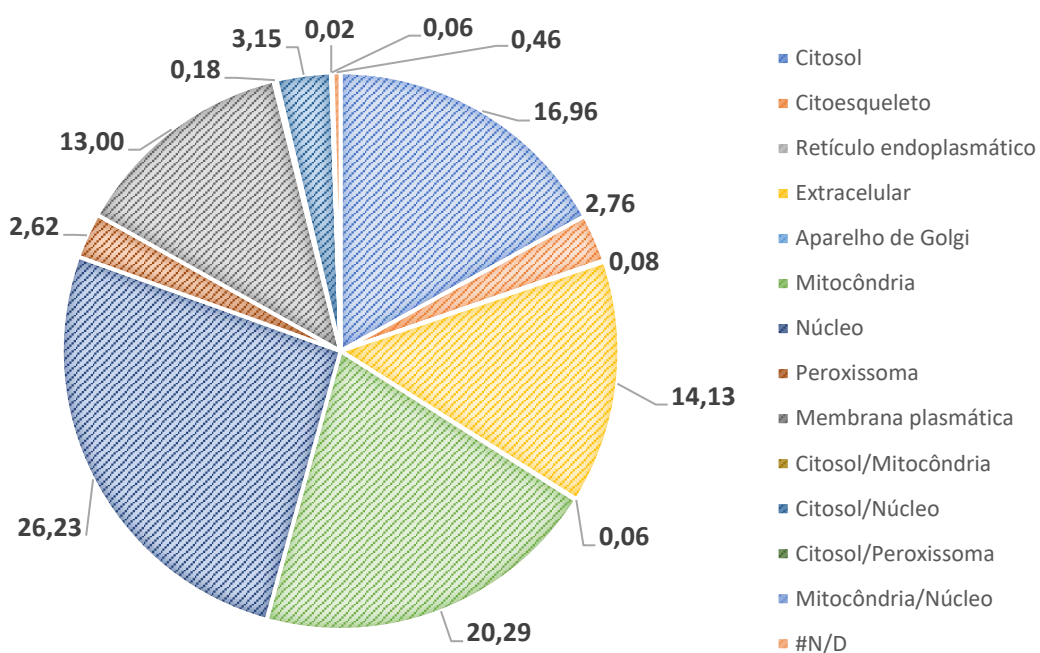
Bioinformatics 13, 134. <https://doi.org/doi.org/10.1186/1471-2105-13-134>

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P.K., Xu, Y., Yin, Y., 2018. DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, 95–101. <https://doi.org/10.1093/nar/gky418>

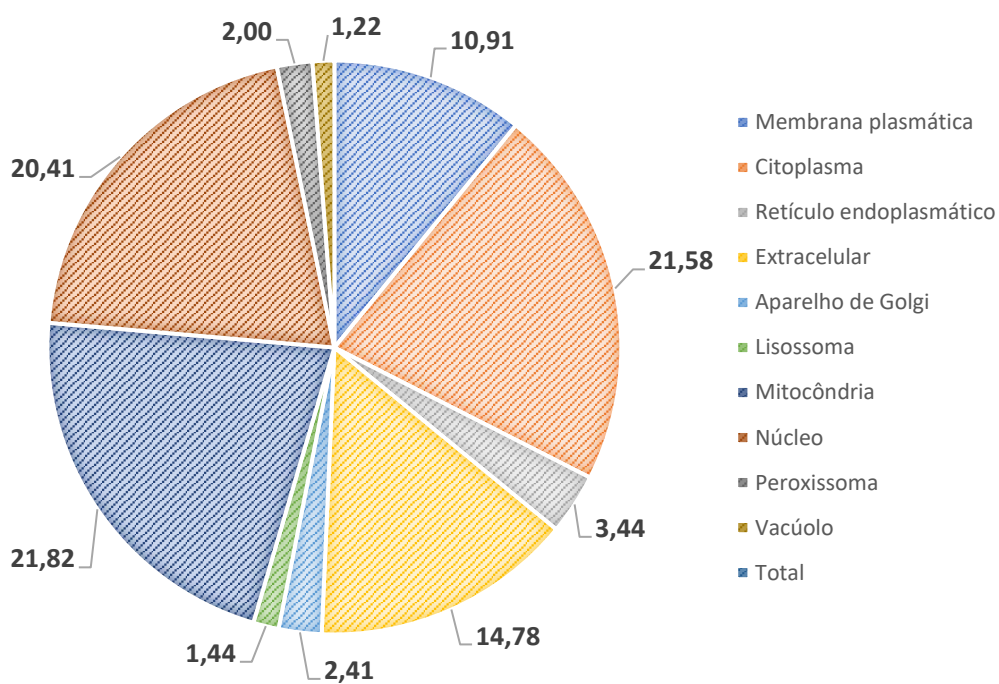
Zhang, L., Huang, X., He, C., Zhang, Q.Y., Zou, X., Duan, K., Gao, Q., 2018. Novel fungal pathogenicity and leaf defense strategies are revealed by simultaneous transcriptome analysis of *Colletotrichum fructicola* and strawberry infected by this fungus. *Front. Plant Sci.* 9, 1–19. <https://doi.org/10.3389/fpls.2018.00434>

Zhu, C., Wang, M., Wang, W., Ruan, R., Ma, H., Mao, C., Li, H., 2014. Glucosylceramides are required for mycelial growth and full virulence in *Penicillium digitatum*. *Biochem. and Biophys. Res. Commun.* 455, 165–171. <https://doi.org/10.1016/j.bbrc.2014.10.142>

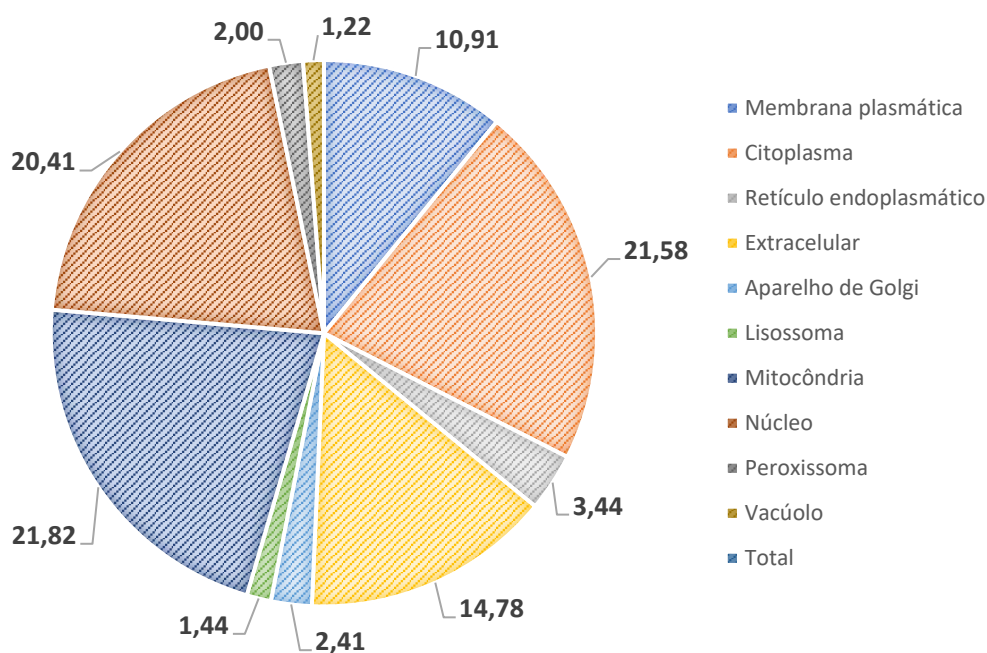
VII. Anexos



Anexo 1 - Localização das proteínas segundo WoLF PSORT.



Anexo 2 - Localização das proteínas segundo DeepLoc-1.0.



Anexo 3 - Localização das proteínas segundo o ProtComp 9.0, com base na previsão "Neural Nets".

Anexo 4 – Comparação da representatividade das diferentes classes de CAZymes em diversos genomas do género *Colletotrichum* (Baroncelli *et al.*, 2016; Rao & Nandineni, 2017).

Espécie	AA	CBM	CE	GH	GT	PL
<i>Colletotrichum nymphaeae</i>	22,94	8,89	15,67	42,33	4,04	6,14
<i>Colletotrichum simmondii</i>	23,21	8,88	15,89	42,52	3,58	5,92
<i>Colletotrichum fioriniae</i>	22,83	9,29	15,43	42,52	3,94	5,98
<i>Colletotrichum salicis</i>	21,89	10,27	16,67	41,58	3,3	6,23
<i>Colletotrichum graminicola</i>	23,34	11,78	14,56	44,11	3	3,21
<i>Colletotrichum sublineola</i>	20,08	11,7	14,81	45,42	4,68	3,31
<i>Colletotrichum higginsianum</i>	19,04	8,97	17,72	42,67	4,38	7,22
<i>Colletotrichum fruticola</i>	22,75	9,13	16,47	42,96	2,99	5,69
<i>Colletotrichum gloeosporioides</i>	22,95	8,84	16,12	42,95	3,1	6,05
<i>Colletotrichum orbiculare</i>	20,93	9,3	14,67	42,93	4,29	7,87
<i>Colletotrichum truncatum</i>	13,84	10,48	16,77	41,93	7,55	9,43
Média	21,3	9,8	15,9	42,9	4,1	6,1

Anexo 5 - Comparação da representatividade das diferentes famílias de proteases em diversos genomas do género *Colletotrichum* (Baroncelli et al., 2016; Rao & Nandineni, 2017).

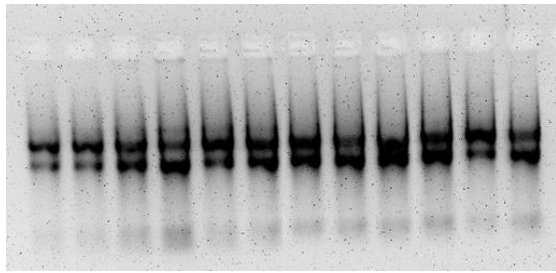
Espécie	A	C	G	I	M	P	S	T
<i>Colletotrichum nymphaeae</i>	14,15	1,89	0,94	3,77	37,74	0,00	40,57	0,94
<i>Colletotrichum simmondii</i>	12,96	2,78	0,93	3,70	39,81	0,00	37,96	1,85
<i>Colletotrichum fioriniae</i>	13,89	1,85	0,93	3,70	37,04	0,00	41,67	0,93
<i>Colletotrichum salicis</i>	12,38	0,95	0,95	3,81	38,10	0,00	42,86	0,95
<i>Colletotrichum graminicola</i>	14,46	1,20	1,20	2,41	40,96	0,00	36,14	3,61
<i>Colletotrichum sublineola</i>	13,64	1,14	1,14	3,41	35,23	0,00	43,18	2,27
<i>Colletotrichum higginsianum</i>	13,83	2,13	2,13	6,38	30,85	0,00	42,55	2,13
<i>Colletotrichum fruticola</i>	10,71	0,89	0,00	2,68	38,39	0,00	46,43	0,89
<i>Colletotrichum gloeosporioides</i>	10,38	1,89	3,77	0,00	34,91	0,94	47,17	0,94
<i>Colletotrichum orbiculare</i>	11,54	0,00	0,96	4,81	33,65	0,00	47,12	1,92
<i>Colletotrichum truncatum</i>	6,76	1,35	1,35	4,05	39,19	0,00	47,30	0,00
Média	12,2	1,5	1,3	3,5	36,9	0,1	43,0	1,5

Anexo 6 - Primers dos genes escolhidos, desenhados com recurso aos PerlPrimer v1.1.21, e dos genes de referência (*) (Vieira et al., 2016) com as devidas temperaturas de emparelhamento, tamanho dos amplicões e eficiência da reação de qPCR.

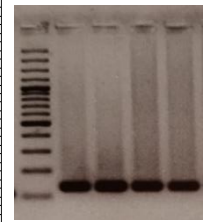
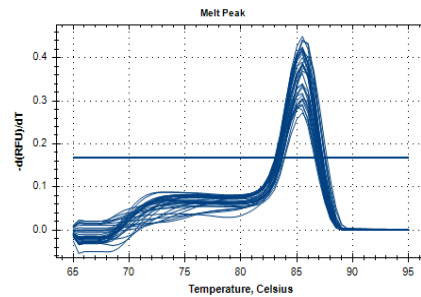
Gene	Descrição	Primer		Temperatura de anelamento (°C)	Comprimento do amplicão	Eficiência
Actina	Actina	F	CAACATTGTCATGTCTGGTGG	63	202	102,1
		R	GTACTCCTGCTTGGAGATCC			
PP1	Proteína de serina/treonina fosfatase	F	CACTGGTTGGAGCGAAAACG	60	250	100,2
		R	CAGGATCTGGAACGAGCAAAG			
ck34620	Proteína não caracterizada homóloga a CGGC5_3535	F	CCCGACTTCCACTTCCATTACC	63	208	100,9
		R	CGCCGACCAGGATGAACTTG			
CKAH01_04076	Cutinase	F	GCCAAACGAAGGGATTTGCTC	62,5	125	101,7
		R	GGAGATGTGCTACTGGATCGT			
CKAH01_09212	Pectato liase	F	GTGAACATTATCAACTTCTACGCC	61	106	99,7
		R	GCAGTAGTACCTTCGACATACAC			
CKAH01_09762	Fitoceramidase alcalina	F	TCCAAACCTCAAAGCAAGTCTG	62	209	101,4
		R	AACAAATGCCACCATCCATGAC			
CKAH01_11318	Peroxidase	F	CCGATGACTTTCGTGGTCCT	62	198	100,3
		R	GTAAAGAACGTGGCATTAGGCTC			
CKAH01_12047	Pectato liase	F	CCAGGTCTATGACGGAAAGTGG	63	87	103
		R	GACGGAGTCCTTGTAATCACCC			
CKAH01_15250	Peptidase da família C40 (com domínio LysM)	F	GAGGGCTGGTACTACACGAC	62	92	100,1
		R	GTTGTTCCACTGCTTGATGACC			
CKAH01_15357	Peptidase da sub-família M23b (com domínio LysM)	F	GGGCTGGTACTACATCGTCAC	62,5	99	100,1
		R	GTTGGGAATGCCGTTGATCTC			

Anexo 7 - Primers dos restantes genes escolhidos, desenhados com recurso aos PerlPrimer v1.1.21, com as devidas temperaturas de emparelhamento e a razão pela qual não foi possível obter resultados conclusivos.

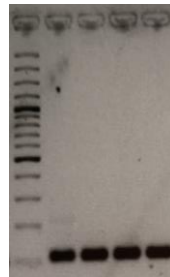
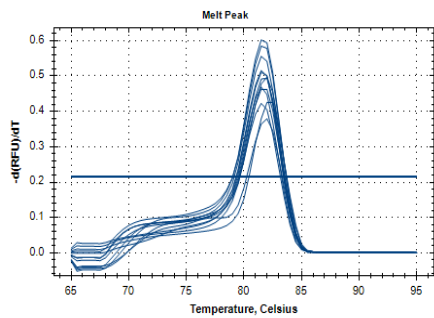
Gene	Descrição	Primer	Temperatura de anelamento (°C)	Problemas encontrados	
CKAH01_10662	Proteína indutora de necrose induzida pelo etileno (NPP1)	F	GAGATTCAACCCTACTCTGCAC	60,92	Início da amplificação muito tardia-após 35 ciclos; <i>melting curves</i> muito variadas
		R	TTGCCATTGAGCGTTGTTCC	62,29	
CKAH01_11118	Proteína não caracterizada semelhante a CGGC5_704	F	GCAGAAGAACACCACCCTCC	62,77	Diversas bandas no gel de eletroforese; <i>melting curves</i> muito variadas
		R	TGTTGCCCTTGAACCTGTCGT	62,76	
CKAH01_11483	Proteína transportadora de membrana (<i>Major facilitator superfamily</i>)	F	GCGGATCGACTTCTTTACCT	60,02	Início da amplificação muito tardia-após 35 ciclos; <i>melting curves</i> muito variadas
		R	GGACAAGTTGCTGTAGACCC	60,80	
CKAH01_14472	Proteína não caracterizada semelhante a CGGC5_2236	F	CCTGTCCTGAGAATGCTCTACC	62,38	Diversas bandas no gel de eletroforese; <i>melting curves</i> muito variadas
		R	CTGTCCGTATCGCTTCGTCC	63,14	
CKAH01_16092	Proteína com módulo de ligação a quitina	F	GCCCACGGAAACATCATCTC	61,58	Início da amplificação muito tardia-após 40 ciclos; <i>melting curves</i> muito variadas
		R	TACCTGGGCCTTATTGTCCTC	61,49	
CKAH01_16717	Proteína não caracterizada com possível módulo LysM	F	GCTTGATGTGCGGTGATGTG	62,74	Sem amplificação; possível erro na anotação do gene
		R	GCACCTACCTCGTTGGACTG	62,99	
CKAH01_18209	Proteína do tipo "Germin-like"	F	GAACACCTTCCAGATGTCCGT	62,56	Grandes diferenças na amplificação entre réplicas biológicas; diversas bandas no gel de eletroforese
		R	CACCAAGAGTAGCAGAGACGAC	62,84	



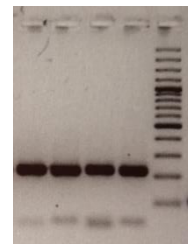
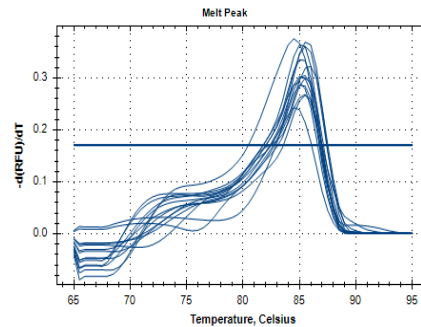
(A)



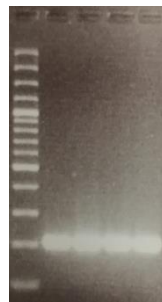
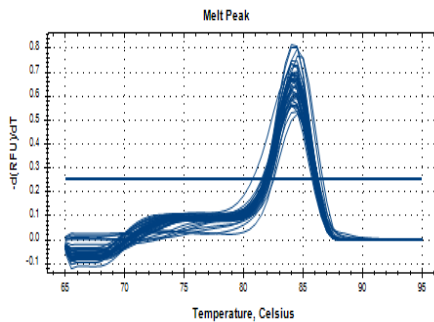
(B)



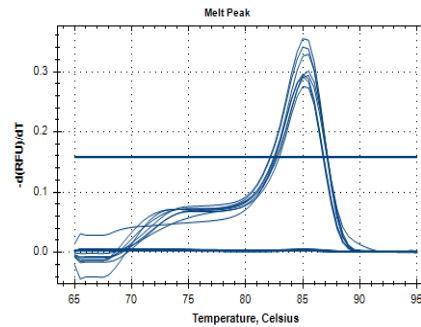
(C)



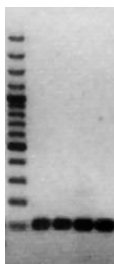
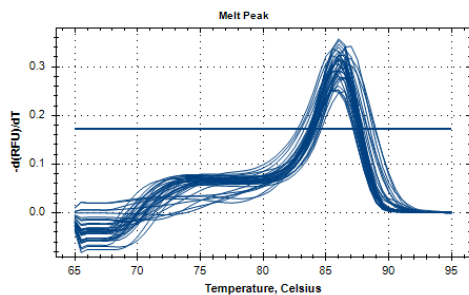
(D)



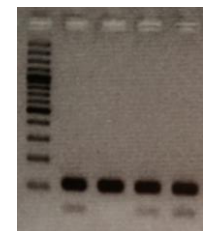
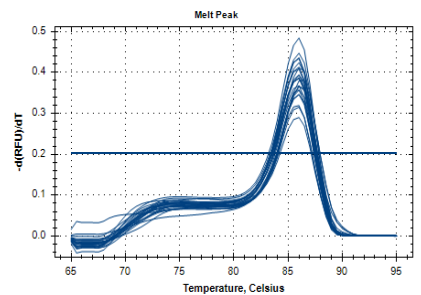
(E)



(F)



(G)



(H)

Anexo 8 - Exemplo de eletroforese em gel de agarose a 1,2% em 1xtampão Tris-Acetato-EDTA para verificação da qualidade do RNA extraído; o RNA foi visualizado após coloração com GreenSafe Premium (NZYTech) sob luz ultravioleta (A); Curva de dissociação (esquerda) e eletroforese em gel de agarose a 2% em 0,5xtampão Tris-Borato-EDTA (direita) de alguns dos produtos de RT-PCR amplificados para os diferentes genes estudados; o DNA foi visualizado por coloração com GreenSafe Premium (NZYTech) sob luz ultravioleta; marcador molecular 100 bp Plus (ThermoScientific): CKA01_04076 (B); CKA01_09212 (C); CKA01_09762 (D); CKA01_11318 (E); CKA01_12047 (F); CKA01_15250 (G); CKA01_15357 (H).

Anexo 9 - Resultados da expressão génica, relativa ao controlo (Teste *t-student*), obtidos através do estudo da expressão realizados no programa Bio-Rad CFX Maestro 2.0 v.5.0.021.0616, na função "Gene Study".

Gene	Amostra	Expressão	P-Value
CKAH01_04076	1dai	14,15169	0,004822
	2dai	3,18541	0,152446
	5dai	1,46735	0,758963
	7dai	0,96978	0,967651
	10dai	0,6256	0,191321
	AP	1,26694	0,627264
	APV	7,31352	N/A
	Esporos	0,27726	0,061348
	Micélio	1	N/A
CKAH01_09212	2dai	2118,87808	0,00074
	5dai	197,60187	0,019853
	7dai	65,14196	0,07299
	10dai	79,67374	0,004165
	AP	9,86369	0,169806
	Esporos	28,9345	0,185185
	Micélio	1	N/A
CKAH01_09762	1dai	425,7533	0,000283
	2dai	87,92304	0,009363
	5dai	63,58488	0,005372
	7dai	28,16603	0,002067
	10dai	17,17417	0,022492
	AP	1,40199	0,816347
	APV	1,88673	N/A
	Esporos	2,74151	0,299656
	Micélio	1	N/A
CKAH01_11318	1dai	320,08905	0,001127
	2dai	420,34538	0,000384
	5dai	151,07605	0,0004
	7dai	150,43762	0,000487
	10dai	272,82015	0,00052
	AP	0,54103	0,530562
	APV	548,85222	N/A
	Esporos	1,03131	0,934999
	Micélio	1	N/A
CKAH01_12047	5dai	15255,63682	0,000578
	7dai	5,67067	0,068639
	10dai	36,3498	0,009754
	AP	0,08396	0,207654
	Micélio	1	N/A
CKAH01_15250	1dai	22,52925	0,032827
	2dai	35,54725	0,00795

	5dai	230,15701	0,001751
	7dai	124,50536	0,002548
	10dai	76,24103	0,00325
	AP	0,066	0,148435
	APV	2,61654	N/A
	Esporos	0,03779	0,185312
	Micélio	1	N/A
CKAH01_15357	1dai	20,27326	0,009563
	2dai	8,57252	0,012293
	5dai	195,65814	0,000498
	7dai	117,36063	0,003048
	10dai	83,29481	0,000538
	AP	0,16713	0,198509
	APV	10,4331	N/A
	Esporos	0,0215	0,010038
	Micélio	1	N/A