

Sulamita de Freitas Franco

“Caracterização da família gênica Taumatinas - like (*MpTLPs*) e proteínas candidatas a efetores de patogenicidade *MpCSEPs* no patossistema *T. cacao* / *M. perniciosa*”

“Characterization of the Thaumatin-like gene family (*MpTLPs*) and the putative pathogenicity effectors proteins *MpCSEPs* in the *T. cacao*/*M. perniciosa* pathosystem”

**CAMPINAS
2014**



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

Sulamita de Freitas Franco

“Caracterização da família gênica Taumatinas - like (*MpTLPs*) e proteínas candidatas a efetores de patogenicidade *MpCSEPs* no patossistema *T. cacao* / *M. pernicioso*”

“Characterization of the Thaumatin-like gene family (*MpTLPs*) and the putative pathogenicity effectors proteins *MpCSEPs* in the *T. cacao*/*M. pernicioso* pathosystem”

Tese apresentada ao Instituto de Biologia da Unicamp como parte dos requisitos exigidos para obtenção do título de Doutora em Biologia Funcional e Molecular, na área de Bioquímica.

Thesis presented to the Biology Institute of the University of Campinas in partial fulfillment of requirements of the degree of Doctor, in Functional and molecular biology, in Biochemistry.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira
Co- orientador: Dr. Jorge Mauricio Costa Mondego

Este exemplar corresponde à versão final
Tese defendida pela aluna Sulamita de
Freitas Franco e orientada pelo Prof. Dr.
Gonçalo Amarante Guimarães Pereira

CAMPINAS, 2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

F848c Franco, Sulamita de Freitas, 1981-
Caracterização da família gênica Taumatinas-like (*MpTLPs*) e proteínas candidatas a efetores de patogenicidade *MpCSEPs* no patossistema *T.cacao/M. perniciosa* / Sulamita de Freitas Franco. – Campinas, SP : [s.n.], 2014.

Orientador: Gonçalo Amarante Guimarães Pereira.
Coorientador: Jorge Maurício Costa Mondego.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. *Moniliophthora perniciosa*. 2. Taumatina. 3. Vassoura-de-bruxa (Fitopatologia). I. Pereira, Gonçalo Amarante Guimarães, 1964-. II. Mondego, Jorge Maurício Costa. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Characterization of the Thaumatin-like gene family (*MpTLPs*) and the putative pathogenicity effector proteins *MpCSEPs* in the *T. cacao/M. perniciosa* pathosystem

Palavras-chave em inglês:

Moniliophthora perniciosa

Thaumatococcus

Witches' broom disease

Área de concentração: Bioquímica

Titulação: Doutora em Biologia Funcional e Molecular

Banca examinadora:

Gonçalo Amarante Guimarães Pereira [Orientador]

Isabel Rodrigues Gerhardt

Jörg Kobarg

Carlos Augusto Colombo

Anderson Ferreira da Cunha

Data de defesa: 28-02-2014

Programa de Pós-Graduação: Biologia Funcional e Molecular

Campinas, 28 de fevereiro de 2014

BANCA EXAMINADORA

Prof. Dr. Gonçalo Amarante Guimarães Pereira
(orientador)



Assinatura

Dra. Isabel Rodrigues Gerhardt



Assinatura

Dr. Jörg Kobarg



Assinatura

Dr. Carlos Augusto Colombo



Assinatura

Prof. Dr. Anderson Ferreira da Cunha



Assinatura

Dra. Juliana Ferreira de Oliveira



Assinatura

Prof. Dr. Claudio Chrysostomo Werneck



Assinatura

Dr. Marco Aurelio Takita



Assinatura

RESUMO

A cultura do cacau é de grande importância econômica em países produtores da semente e produtores de chocolate, e move um mercado que alcança hoje ao redor de U\$ 50 bilhões. As doenças fúngicas são um fator importante na redução de produção de cacau. No Brasil, o principal responsável pela perda na produção é o basidiomiceto *Moniliophthora perniciosa*, agente etiológico da doença conhecida como vassoura de bruxa, que arrasou a produção brasileira na década de 90.

A doença tem como característica a não elicitação de resposta hipersensitiva (HR), o que permite a colonização de tecidos vegetais pelo fungo. HR é uma resposta que desencadeia a indução da expressão de proteínas denominadas PRs (*pathogenesis related*), que agem contra patógenos. Entre as PRs encontram-se as PR5, conhecidas como taumatinas.

Análise do genoma de *M. perniciosa* resultou na anotação de 13 genes que codificam proteínas Taumatina-like (*MpTLPs*). Apesar de serem bem descritas como proteínas antifúngicas em plantas, estudos em fungos basidiomicetos apontam seu envolvimento no remodelamento celular durante a formação cogumelos. *M. perniciosa* é a espécie de fungo com o maior número de *TLPs* descritos até o momento, tendo 6 deles organizados em *clusters*, indicando eventos de duplicação. Análise filogenética revelou uma maior quantidade de *TLPs* em basidiomicetos, quando comparado a ascomicetos, indicando sua participação na formação de cogumelos. No geral, *MpTLPs* são transcritas durante a fase de vassoura seca da doença, onde o galho seco está sujeito a infecção de fungos oportunistas, indicando a contribuição dessas proteínas no combate contra esses fungos, além de poder estar participando no desenvolvimento de sua parede para formação de basidiocarpo e na degradação da parede celular da planta.

Além das *MpTLPs*, análise de transcriptoma revelou um total de 35 *MpCSEPs* (Proteínas Candidatas a Efetores de Patogenicidade Secretados pelo fungo) com valor de

RPKM >50 foram identificadas em plantas de cacau infectadas. Esses genes são expressos preferencialmente na fase inicial da doença (Vassoura verde), onde encontramos o fungo ainda em fase biotrófica. Em geral, esses genes codificam proteínas pequenas e ricas em cisteínas que são características típicas de efetores de virulência (AVRs), podendo ser uma evidência de que esses genes codificam potenciais efetores.

Após triagem, 16 desses genes foram escolhidos para a caracterização estrutural. Destes, sete foram clonados para expressão de proteína heteróloga, o que resultou ao final do estudo a cristalização de 2 proteínas (MpCSEP 5 e MpCSEP 14). Os cristais obtidos até o momento foram testados em linha de luz MX1 e MX2, mas não apresentou difração e os processos de refinamento dos cristais deverão ser continuados. As proteínas obtidas nesse estudo poderão ser utilizadas em testes enzimáticos em trabalhos futuros, para sua completa caracterização estrutural e funcional.

ABSTRACT

The cocoa cultivation have great economic importance in the seed and chocolate producing countries , moving a market that presently reaches around \$50 billion. Fungal diseases are the major factor in reducing cocoa production. In Brazil, the main responsible for the production loss is the basidiomycete *Moniliophthora perniciosa*, known as the etiological agent of witches' broom disease , which devastated the Brazilian production in the 90's disease.

The disease is characterized by not eliciting a hypersensitive response (HR), which allows the plant tissue colonization by the fungus. HR is a response that triggers the induction of the expression proteins called PRs (pahtogenesis related), which act against pathogens. Among the PRs, PR5 are known as thaumatins.

M. perniciosa genome analysis resulted in the annotation of 13 genes encoding proteins Thaumatin - like (*MpTLPs*). Despite being well described as antifungal proteins in plants, fungi basidiomycetes studies suggest its involvement in cellular remodeling during mushroom formation. *M. perniciosa* fungal is the species with the highest number of *TLPs* described so far, with 6 of them organized in clusters, indicating duplication events. Phylogenetic analysis revealed a greater number of *TLPs* in basidiomycetes when compared to ascomycetes, indicating their involvement in the formation of mushrooms. Overall, *MpTLPs* are transcribed during the dry broom disease, where the dead branch is subject to opportunistic fungal infection, indicating the contribution of these proteins in the fight against these fungi, and can be participating in the development of your wallpaper for training basidiocarp and in the cell wall degradation of the plant.

Besides *MpTLPs*, transcriptome analysis revealed a total of 35 *MpCSEPs* (candidate secreted effectors of pathogenicity protein by the fungus) with RPKM value > 50 have been identified in infected cocoa plants. These genes are preferentially expressed

in the early phase of the disease (Green Broom), where we found the fungus still in biotrophic phase. In general, these genes encode small and rich in cysteine that are typical characteristics of virulence effector (AVRs) proteins, which can be evidence that these genes encode potential effectors.

After screening, 16 of these genes were chosen for structural characterization. Of these, seven were cloned for expression of heterologous protein, resulting at the end of the study the crystallization of 2 proteins (MpCSEP 5 and 14). The crystals obtained so far been tested in MX1 and MX2 light line, but showed no diffraction and processes of crystals refinement should continue. The proteins obtained in this study may be used in enzymatic assays in future work to its full structural and functional characterization.

Sumário

RESUMO	VII
ABSTRACT	IX
1- INTRODUÇÃO	1
1.1- A CULTURA DO CACAU.....	1
1.2- <i>MONILIOPHTHORA PERNICIOSA</i> E A VASSOURA DE BRUXA.....	2
1.3- MECANISMOS DE INTERAÇÃO PLANTA-PATÓGENO.....	6
1.4- PROTEÍNAS EFETORAS DE PATOGENICIDADE	10
1.5- PROTEÍNAS DO TIPO TAUMATINA	11
2- OBJETIVOS GERAIS	15
3- CAPÍTULO I: CARACTERIZAÇÃO DA FAMÍLIA GÊNICA TAUMATINA- LIKE DE <i>M. PERNICIOSA</i>	17
3.1- APRESENTAÇÃO.....	17
3.2- JUSTIFICATIVAS.....	17
3.3- OBJETIVO ESPECÍFICO.....	18
3.4- ARTIGO SUBMETIDO À REVISTA <i>MOLECULAR GENETICS AND GENOMICS</i> :.....	19
3.5- MATERIAIS E MÉTODOS.....	46
3.5.1- CLONAGEM, EXPRESSÃO E PURIFICAÇÃO DE <i>MPTLPs</i>	46
3.5.2- <i>Clonagem dos gene MPTLPs</i> :.....	47
3.5.3- <i>Teste de indução da proteína recombinante em pequena escala</i>	49
3.5.4- <i>Clonagem de MPTLP1 em vetores de expressão para células de inseto</i>	50
3.5.5- <i>Ensaio de produção de proteína em larga escala</i>	55
3.5.6- <i>Purificação de MPTLP1</i>	56
3.6- RESULTADOS E DISCUSSÃO	56
3.6.1- <i>Amplificação e clonagem de MPTLPs</i>	56
3.6.2- <i>Ensaio de expressão de proteínas MPTLPs</i>	59
3.6.3- <i>Ensaio de indução em pequena escala de MPTLP1 e MPTLP4 em Shuffle</i>	60
3.6.4- <i>Clonagem de MPTLP1 em pFastBacEGFP</i>	62
3.6.5- <i>Confirmação da transposição</i>	63
3.6.6- <i>Experimento de expressão de MPTLP1-histag em larga escala</i>	66
3.6.7- <i>Clonagem de MPTLP1+ peptídeo sinal + histag em pFastBacEGFP C-HT</i>	67
4- CAPÍTULO II: ESTUDO DE GENES CANDIDATOS A EFETORES SECRETADOS DE PATOGENICIDADE DE <i>M. PERNICIOSA</i> (<i>MPCSEPS</i>).....	71
4.1- JUSTIFICATIVA.....	72
4.2- OBJETIVOS	72
4.3- MATERIAIS E MÉTODOS.....	72
4.3.1- <i>Análise de expressão gênica de MPCSEPs</i>	72
4.3.2- <i>Clonagem dos genes MPCSEPs</i>	73
4.3.3- <i>Teste de indução e solubilidade</i>	74
4.3.4- <i>Expressão e purificação em grande escala</i>	75
4.3.5- <i>Cristalização</i>	76
4.3.6- <i>Dicroísmo Circular (CD)</i>	79
4.3.7- <i>Espalhamento dinâmico de luz (DLS)</i>	80
4.4- RESULTADOS	80
4.4.1- <i>Mineração das sequências MPCSEPs</i>	80
4.4.2- <i>Amplificação e clonagem de MPCSEPs</i>	82
4.4.3- <i>Teste de expressão, solubilidade e purificação das MPCSEPs</i>	84
4.4.4- <i>Cristalização</i>	89
4.4.5- <i>Dicroísmo Circular e Espalhamento de Luz Dinâmico</i>	91
4.5- DISCUSSÃO	93
4.6- PERSPECTIVAS	95

5- REFERÊNCIAS BIBLIOGRÁFICAS:	97
ANEXO 1: SEQUÊNCIAS CDS E DE AMINOÁCIDOS DE <i>MPTLPS</i>	103
ANEXO 2: SEQUÊNCIAS CDS E DE AMINOÁCIDOS DAS <i>MPCSEPS</i> ESTUDADAS.	107
ANEXO 3: ARTIGO EM CO-AUTORIA.....	111

DEDICATÓRIA

Dedico a vocês, meus pais, que sempre dedicaram tempo, atenção e amor em minha criação. Minhas conquistas são todas suas.

AGRADECIMENTOS

A FAPESP, agradeço pelo apoio financeiro ao projeto e a todos os projetos que o trabalho foi inserido.

Ao professor e Orientador Gonalo Pereira, que pela terceira vez me deu a oportunidade de trabalhar e superar mais esse desafio. Obrigada pela confiana depositada, pela amizade e respeito.

Ao co- orientador Jorge Mondego, por toda orientao e dedicao ao trabalho, Obrigada pelas idias, por dividir seu conhecimento e pela disposio em aprender junto comigo.

Ao Daniel Trindade, o qual tive a felicidade de conhecer em um momento crtico do trabalho. Te ver trabalhar com tanta seriedade e amor ao trabalho me ajudaram e muito a superar os problemas que enfrentava. Obrigada pela competncia, pela generosidade, dedicao e boa vontade em transmitir seus conhecimentos.

A Sandra Dias, Andr Ambrsio e Juliana Oliveira, por compartilhar de seus conhecimentos, ceder espao no laboratrio e por toda a orientao no decorrer do projeto SMOLBNET 2.

Aos membros das bancas avaliadoras da qualificao e de defesa, pela leitura crtica e toda colaborao ao texto final.

A Murphy, e sua lei que me pega o tempo todo, e se no me torna mais esperta, me deixa pelo menos, mais criativa.

A Eliane. Sem voc, esse laboratrio andaria em outro ritmo. Obrigada pela agilidade e eficincia em tudo que faz. Obrigada por toda ajuda e por muitas vezes nos acalmar, em especial quando chegamos na fase final do trabalho. E muito, muito obrigada por todo carinho e amizade.

Ao meu primo querido Jonathan que mesmo distante sempre se preocupou e me apoiou.

A Trufa e Pantufa, que acredito nunca terem entendido toda a oscilao do humor que essa fase da vida nos proporciona, mas que sempre estiveram juntas para me divertir e desviar meu foco com suas necessidades de cuidado, afeto e carinho. E obrigada pela capacidade de retribuir tudo isso em forma de companheirismo, sem nem saber a importncia que isso tem pra mim e a felicidade que isso me proporciona.

As minhas queridas amigas Anglica e Vanessa, que fazem parte de minha vida h 25 anos e sempre estiveram ao meu lado me apoiando e rindo (de mim e comigo!).

Aos colegas de laboratrio do LGE e LNBio pela ajuda e pelos momentos de descontrao.

A amiga e colega de trabalho Renata Barone. Obrigada por compartilhar seu aprendizado durante seu doutorado. Obrigada por toda conversa, idias e apoio.

Ao irmão que pude escolher Eduardo Camargo, agradeço por ter entrado em minha vida. Se pudesse, teria escolhido que isso tivesse acontecido antes. É incrível como alguém consegue traduzir melhor o que penso do que eu mesma, e sempre com aquela dose de humor seco! Obrigada por me fazer acreditar que as coisas podem dar certo, e se não der, sempre há uma alternativa. Obrigada por toda amizade, risadas, lágrimas, carinho e companheirismo que fez diferença em meu crescimento pessoal e profissional.

Aos meus pais agradeço pelo apoio, amor e preocupação. Acredito que sou aquilo que vocês fizeram. Obrigada pela educação dada a mim e minha irmã. Mesmo com tanta dificuldade e até mesmo falta de conhecimento conseguimos nos manter num caminho tranquilo e sempre aceitando e enfrentando os desafios que a vida nos proporciona. Obrigada pela simplicidade em encarar a vida e seriedade em resolver problemas. Obrigada por todo esforço e cuidado dedicado a nossas vidas. As nossas conquistas são muito mais suas do que nossas.

A minha irmã, por ter dividido sua vida comigo. Por ter me protegido tantas vezes... Obrigada pela preocupação, e por todo apoio.

A meu sogro e sogra, pelo carinho e compreensão. Fico muito feliz em poder contar com um casal tão especial que possui um equilíbrio perfeito entre doçura e firmeza. Obrigada por toda a segurança, carinho e amizade.

Ao meu marido, amigo e companheiro Celso Fiori. Agradeço primeiro por ter entrado em minha vida e desde então estar fazendo dela uma completa transformação. Obrigada por me ensinar que as coisas podem ser menos dolorosas se tratadas com mais leveza. Obrigada por me mostrar que a paciência pode nos levar a encontrar um caminho mais simples. Obrigada por toda dedicação. Obrigada por me acalmar em diversas situações e me “sacudir” em tantas outras. Obrigada por ter colocado muito mais sorriso em minha vida. Obrigada pelo colo, pela atenção, pelo carinho, por todo amor e cuidado, pela orientação e por me tornar uma pessoa mais feliz e menos ranzinza.

LISTA DE FIGURAS

Figura 1: O cacau e seu principal produto	1
Figura 2: Representação do ciclo de vida do fungo <i>M. perniciosa</i>	4
Figura 3: Desenho esquemático exemplificando a primeira linha de defesa de plantas	7
Figura 4: Desenho esquemático representando o reconhecimento de proteínas AVR do patógeno por proteínas R do hospedeiro.	9
Figura 5: Cronograma do trabalho abordado na presente tese.	16
Figura 6: Mapa da criação do vetor pFastBacEGFP	50
Figura 7: Mapa da criação do vetor pFastBacEGFP C-HT.	52
Figura 8: Digestão de pSW202- <i>MpTLP1</i> e pGEM- <i>MpTLP</i> 4, 6 e 9.	58
Figura 9: Digestão de pGEM- <i>MpTLPs</i> (2, 5, 7 e 8) com enzimas de restrição <i>EcoRI</i> e <i>HindIII</i> .	58
Figura 10: Experimento de indução em pequena escala de <i>MpTLPs</i>	60
Figura 11: Expressão em pequena escala de <i>MpTLP1</i>	61
Figura 12: Teste de solubilidade de <i>MpTLP1</i>	61
Figura 13: Confirmação de transposição da construção em pFastBacEGFP para bacmídeo	64
Figura 14: Frações insolúveis: testes de indução em quantidade crescente de MOI viral.	65
Figura 15: Frações solúveis: testes de indução em quantidade crescente de MOI viral.	65
Figura 16: Indução em larga escala <i>MpTLP1</i>	66
Figura 17: Digestão de pFastBacEGFP C-HT (1) e <i>MpTLP1</i> + peptídeo sinal	68
Figura 18: Desenho esquemático de poços de placas de gota sentada e gota pendente	79
Figura 19: Expressão gênica de proteínas efetoras candidatas	81
Figura 20: Amplificação de genes <i>MpCSEPs</i> a partir de cDNA de cacau	84
Figura 21: Teste de expressão e solubilidade das proteínas em 18°C e 37°C	85
Figura 22: Expressão e purificação de <i>MpCSEP1</i>	86
Figura 23: Expressão e purificação de <i>MpCSEP9</i> (esquerda) e <i>MpCSEP16</i> (direita)	87
Figura 24: Expressão e purificação de <i>MpCSEP5</i> (esquerda) e <i>MpCSEP14</i> (direita)	87
Figura 25: Purificação de <i>MpCSEP5</i>	88
Figura 26: Purificação de <i>MpCSEP14</i> .	88
Figura 27: Cristais de <i>MpCSEP5</i>	90
Figura 28: Cristais de <i>MpCSEP14</i>	91
Figura 29: Dicroísmo circular. <i>MpCSEP5</i> (esquerda) e <i>MpCSEP14</i> (direita).	92
Figura 30: Espalhamento de luz dinâmico. <i>MpCSEP5</i> (esquerda) e <i>MpCSEP14</i> (direita).	92

Lista de Figuras do Artigo: Capítulo I

Figure 1: Genomic organization of the <i>MpTLPs</i> gene clusters found in <i>M. pernicioso</i> .	27
Figure 2: Sequence alignment of MpTLPs with TLPs of representative organisms	28
Figure 3: Phylogenetic analysis focusing in fungal <i>TLPs</i> .	32
Figure 4: Differential expression of MpTLPs throughout the stages of fungal development and infection	34
Figure 5: 3D molecular model of MpTLP1 and MpTLP12	35

LISTA DE TABELAS

Tabela 1: Oligonucleotídeos utilizados na amplificação e clonagem de <i>MpTLPs</i>	57
Tabela 2: Kits iniciais de cristalização	77
Tabela 3: Características dos genes <i>MpCSEPs</i>	82
Tabela 4: Primer para amplificação e clonagem dos genes <i>MpCSEPs</i>	83
Tabela 5: Resumo das etapas que foram seguidas nesse capítulo	94

Tabelas do Artigo: Capítulo 1

Table 1: Characteristics of the thirteen <i>MpTLP</i> genes identified in the <i>M. perniciosus</i> genome.	29
Table 2: Number of TLPs genes in fungal species with different lifestyles.	31

1- INTRODUÇÃO

1.1- A cultura do cacau

O cacaueiro (*Theobroma cacao* L) é uma planta da família Malvacea originária da região da bacia Amazônica, que aos poucos foi dispersando-se para a região dos Andes e países da América do Sul e Central como Venezuela, Colômbia, Equador, e México. No Brasil, o cacaueiro foi levado até o sul da Bahia e essa região se tornou a maior produtora nacional, respondendo por 75% da produção brasileira (Purdy e Schmidt, 1996; Evans e Prior, 1987).

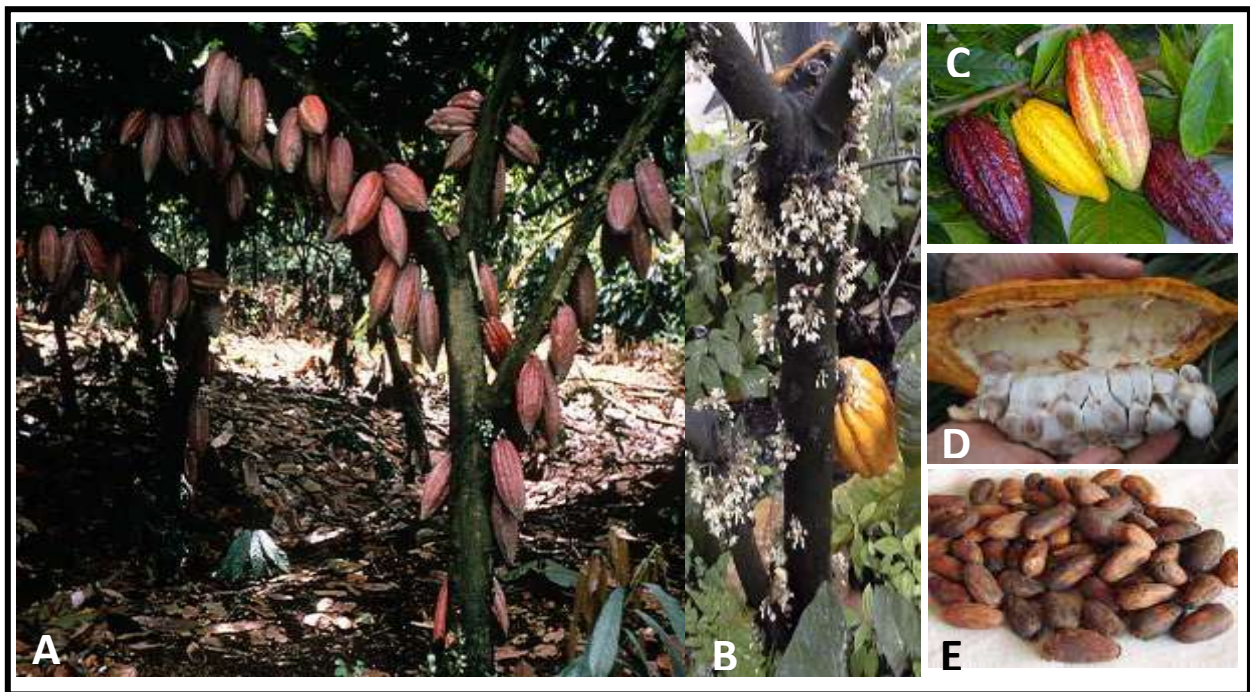


Figura 1: O cacaueiro e seu principal produto. A- árvores de cacau com frutos, B- destaque para as flores no cacaueiro, C- frutos de cacau, D- fruto de cacau cortado longitudinalmente, E- Sementes (amêndoas) do cacau.

As amêndoas produzidas pelo fruto do cacaueiro constituem a principal matéria prima para a produção do chocolate, um alimento energético muito apreciado. Além desse produto, o cacaueiro também é responsável pela produção de manteiga de cacau, licor de

cacau, sucos, e também a produção de pectina e geleias através da casca de seus frutos.

Por ser uma cultura arbórea de alta dependência de mão de obra, a produção dos frutos de cacau e seus derivados é responsável pela criação de empregos e desenvolvimento da região sul da Bahia. Entretanto, a produção dos frutos foi drasticamente afetada com o aparecimento do fungo *Moniliophthora perniciosa* na Bahia (Pereira *et al.*, 1989).

1.2- *Moniliophthora perniciosa* e a Vassoura de Bruxa

Moniliophthora perniciosa (Stahel) Sing. é um fungo endêmico originário da região amazônica responsável pela principal doença do cacau conhecida como vassoura de bruxa (*Witch's broom disease*, WBD). Este fungo foi descrito por Stahel em 1915 como *Marasmius pernicius*. Em 1942, o fungo foi reclassificado sendo transferido para o gênero *Crinipellis*, e renomeado para *Crinipellis perniciosa* (Stahel) Sing. O fungo pertence à família Tricholomataceae, ordem Agaricales, divisão Basidiomycota, e apresenta diferentes biótipos: o biótipo C (infecta *Theobroma* sp e *Herrania* SP); o biótipo S, (infecta plantas da família Solanaceae); o biótipo B (infecta Bixaceae); e o biótipo L ou liana saprotrófico, (que coloniza uma variedade de substratos). Em 2005, o biótipo C sofreu uma reclassificação devido a estudos morfológicos e moleculares, passando a ser incluído no gênero *Moniliophthora*, sendo renomeado como *Moniliophthora perniciosa* (Aime e Phillips-Mora, 2005).

M. perniciosa possui um ciclo de vida hemibiotrófico. No primeiro estágio da doença, o fungo é biotrófico (ou parasítico), com hifas monocarióticas, sem grampos de conexão, que se desenvolvem lentamente no apoplasto vegetal. No segundo estágio, o fungo é saprotrófico ou (necrotrófico), apresentando hifas dicarióticas com grampos de conexão, tendo desenvolvimento acelerado em tecidos necrosados ou mortos (Hedger *et al.*, 1987). Seu ciclo infectivo se inicia quando esporos do fungo dispersos pela água ou ar

encontram tecidos meristemáticos, almofadas florais ou frutos em desenvolvimento (Evans e Bastos, 1979). Ao atingirem a superfície da cutícula ou base de tricomas, os basidiósporos germinam, emitindo tubos germinativos que invadem os tecidos através dos estômatos ou de lesões na cutícula (Silva e Matsuoka, 1999). Os micélios biotróficos crescem intercelularmente no fluido apoplástico, um meio considerado pobre em nutrientes (Pirovani *et al.*, 2005). Esta colonização desencadeia alterações fenotípicas no hospedeiro que são os sintomas característicos da doença: clorose, crescimento hiperplásico e hipertrófico dos tecidos meristemáticos e perda da dominância apical que resulta em superbrotações e desenvolvimento anormal da planta. Esses ramos anômalos derivados dessa infecção são conhecidos como “vassouras verdes”. Tal aspecto está provavelmente associado a um desequilíbrio hormonal causado pelo patógeno na planta, mais especificamente a uma alteração da relação citocinina/auxina após a infecção. Quando a infecção se dá através dos pontos de inserção das flores ou inflorescências, ocorre a formação de frutos partenocárpicos que apresentam aspecto deformado (coloração negra e petrificação).

Dois meses após a infecção ocorre o início da necrose dos ramos infectados. Neste momento o fungo assume o estilo de vida necrotrófico, e seus micélios se tornam dicarióticos. Por ser um fungo homotático, essa transição ocorre sem a necessidade do cruzamento entre indivíduos. Depois da completa colonização nos tecidos mortos, o micélio necrotrófico dá origem aos basidiocarpos, que após períodos de seca e chuva, produzem basidiósporos, que ao serem levados pela água ou ventos, reiniciam o ciclo de vida do fungo. A liberação dos basidiósporos acontece durante a noite, provavelmente pela sensibilidade dos esporos a luz ultravioleta. Somente na fase necrotrófica do fungo, e em condições de alta umidade, os basidiósporos podem ser produzidos. Já que a proliferação do fungo depende de períodos alternados de seca e umidade, esse fungo se

estabeleceu muito bem no estado da Bahia, que possui condições climáticas favoráveis para o desenvolvimento de basidiósporos durante o ano todo.

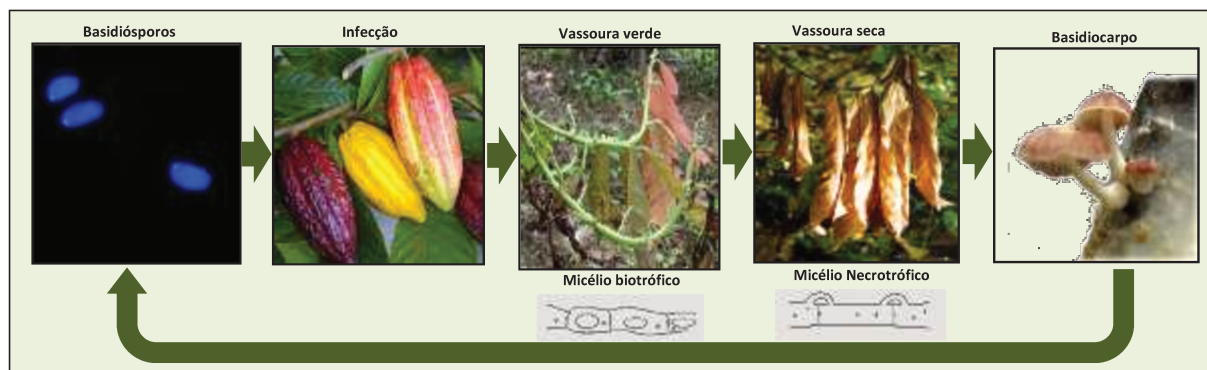


Figura 2. Representação do ciclo de vida do fungo *M. pernicioso*. Basidiósporos germinam sobre tecido meristemático do cacau resultando na formação da vassoura verde, a qual é colonizada pelo micélio biotrófico. A morte do tecido infectado é acompanhada pela mudança para a fase necrotrófica do fungo. Este tecido morto é denominado vassoura seca. Neste estágio da doença são produzidos os basidiomas, que liberam os basidiósporos, reiniciando o ciclo de vida do patógeno.

O fungo tornou-se a principal limitação à produção de cacau na América do Sul, parte da América Central e do Caribe. No final da década de 1980 detectou-se a presença do fungo no Sul da Bahia, aonde encontrou condições favoráveis para seu desenvolvimento através da infecção do cacau. *M. pernicioso* foi responsável pela perda de 70% da produtividade na cacauicultura brasileira, fazendo com que o Brasil passasse de exportador para importador de amêndoas de cacau. Atualmente, o Brasil é o quinto produtor mundial, já tendo ocupado a segunda posição. Essa queda na produção foi responsável pelo aumento do desemprego de trabalhadores do campo em áreas produtoras contribuindo para o aumento da pobreza. Alguns produtores de cacau transformaram suas fazendas em pasto, ocasionando também problemas ecológicos. Esses problemas, juntamente com o problema socioeconômico provocado pela diminuição de cacau levaram à busca de estratégias de combate e controle da doença, destacando-se a utilização de técnicas de manejo, como a poda fitossanitária (Lima *et al.*, 1991), o uso de fungicidas (Bastos, 1989), o controle biológico. (Bastos, 1992) e o

desenvolvimento de variedades de cacau resistentes (Pires, 1995). Porém, nenhuma técnica conhecida até então se mostrou eficaz ao combate da doença.

No ano de 2000, foi lançado o Projeto Genoma Vassoura de Bruxa (www.lge.ibi.unicamp.br/vassoura), sob coordenação do Prof. Dr. Gonçalo Amarante Guimarães Pereira. O projeto vem fornecendo dados importantes sobre o tamanho do genoma (Rincones *et al.*, 2003), sobre a diversidade genética do fungo (Rincones *et al.*, 2006), sobre os mecanismos de ação do fungo durante a doença (Garcia *et al.*, 2007; Ceita *et al.*, 2007) e da modificação bioquímica do cacaueiro durante a infecção (Scarpari *et al.* 2005; Rio *et al.* 2008; Zaparoli *et al.* 2011; de Oliveira *et al.* 2012; Teixeira *et al.* 2012; Barsottini *et al.* 2013). Meinhardt *et al.* (2006) desenvolveram um método de cultivo do fungo em seu estágio biotrófico, o que permitiu ao laboratório analisar a expressão de diversos genes durante essa etapa de desenvolvimento do fungo. Além disso, estudos demonstraram a produção de oxalato de cálcio (Rio *et al.*, 2008) e de proteínas indutoras de necrose (García *et al.*, 2007; Zaparoli *et al.*, 2009) pelo fungo. Em adição, dois artigos forneceram dados importantes sobre as diferenças de expressão gênica entre os micélios biotróficos e necrotróficos de *M. perniciosus* (Rincones *et al.*, 2008) e sobre a constituição do genoma deste patógeno (Mondego *et al.*, 2008).

No ano de 2009, foi aprovado pela FAPESP o projeto temático intitulado “Estudo integrado e comparativo de três doenças fúngicas do cacau – vassoura-de-bruxa, monilíase e mal do facão - visando à compreensão de mecanismos de patogenicidade para o desenvolvimento de estratégias de controle” (processo FAPESP: 2009/50119-9), no qual o presente projeto de doutorado está inserido. Tais doenças são causadas pelos fitopatógenos *Moniliophthora perniciosus*, *Moniliophthora roreri* e *Ceratocystis cacaofunesta*, respectivamente. Este conjunto de estudos visa entender os mecanismos moleculares envolvidos na interação planta-patógeno que levam ao desenvolvimento de doenças do cacaueiro, e analisar quais são as ferramentas moleculares que estes fungos

usam para driblar as defesas da planta no estabelecimento e progressão da infecção. A partir dessa compreensão, espera-se o desenvolvimento de estratégias capazes de interferir nessas interações e assim controlar as doenças que são consequências desses processos.

Análise do genoma de *M. perniciosa* mostra que o fungo possui genes que codificam proteínas similares a proteínas vegetais relacionadas à patogenicidade (proteínas PR). Em plantas, o contato com patógenos desencadeia uma cascata de sinalização que promove a expressão desse tipo de genes.

1.3- Mecanismos de interação planta-patógeno

Hoje, é estabelecido que as plantas possuem duas linhas de defesa contra patógenos. A primeira linha de defesa contra todos os potenciais patógenos desencadeia uma defesa basal e é baseada no reconhecimento de moléculas conservadas conhecidas como PAMPS (*pathogen-associated molecular patterns*), pelos PRRs (*PAMP-recognition receptors*) que ativa PTI (*PAMP-triggered immunity*) para prevenir a colonização da planta pelo hospedeiro (Zipfel 2008; Jones e Dangl 2006). Os representantes melhores descritos de Pamps são flagelinas (Felix *et al.*, 1999) e EF-Tu e harpinas produzidas por bactérias gram- negativas fitopatogênicas, e alguns elicitores de oomicetos e fungos. Um dos mais conhecidos PAMPs de microrganismos é a quitina, principal componente estrutural da parede celular de fungos.

Rincones *et al.* (2008) descrevem a expressão diferencial de genes relacionados à patogenicidade durante o desenvolvimento de *M. perniciosa*, verificando uma maior expressão de proteinases e lipases no estágio biotrófico do fungo. Mondego *et al.* (2008) verificaram a presença de genes que codificam proteínas secretadas ricas em cisteínas no genoma de *M. perniciosa*. Esses resultados corroboram com a análise do transcriptoma de WBD (Teixeira *et al.*, 2013) onde verificou-se que esses genes estão

entre os mais expressos durante a interação planta- fungo e foram nomeados *MpCSEPs* (*Moniliophthora perniciosa* candidate secreted effector proteins). Proteínas com essas características vêm sendo descritas como proteínas efetoras do tipo avr (avirulência) (Van den Ackerveken *et al.*, 1993).

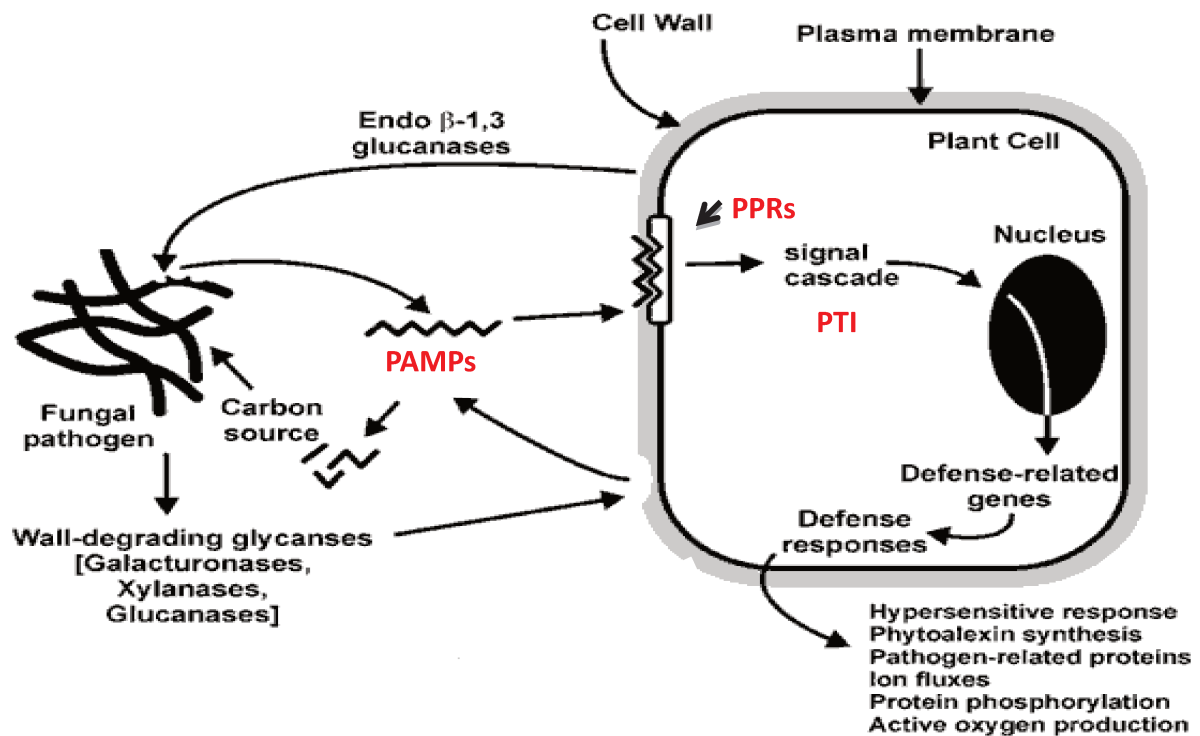


Figura 3: Desenho esquemático exemplificando a primeira linha de defesa de plantas. O ataque enzimático da planta para células fúngicas, bem como o ataque fúngico a célula vegetal liberam moléculas provenientes da digestão de sua parede celular (PAMPs), que agem como sinalizadores da defesa de plantas e são reconhecidos pelos PPR, desencadeando PTI. Figura editada a partir de <http://www.ccrc.uga.edu/~mao/plapath/PPtext.htm>. The University of Georgia

A segunda linha de defesa de plantas envolve o reconhecimento de proteínas efetoras específicas (Avrs) produzidas pelo patógeno por proteínas produzidas por genes R do hospedeiro. Quando há essa interação, ocorre a sinalização para a defesa da planta, (Pinheiro *et al.*, 1999) que desencadeia a segunda linha de defesa de planta, superando a PTI. Quando há o reconhecimento, a interação é caracterizada como incompatível, indicando que a planta é capaz de resistir ao ataque do fungo. Por outro lado, se o patógeno não apresentar o gene *avr* ou ainda, algum efector de patogenicidade

não reconhecido pelo hospedeiro, ou se a planta não possuir o gene *R*, ocorre uma interação compatível, que significa que a planta está susceptível ao ataque do patógeno (Knogge, 1996; Nimchuck *et al.*, 2003). Esse mecanismo conhecido como interação gene-a-gene vem sendo suplantado pela “teoria do guarda”. Nesse caso, as proteínas sintetizadas por genes *avr* interagem com “proteínas-guarda”, que são reconhecidas pelas proteínas *R*, e induzem a resposta de defesa. Assim sendo, a interação entre proteínas *R* e *avr* seria indireta (Dangl e Jones, 2001; Marathe e Dinesh-Kumar, 2003). O reconhecimento de proteínas *Avrs* pelas proteínas *R* ativam ETI (effector-triggered immunity), que provoca morte celular programada nas células do hospedeiro, conhecida como HR (*Hypersensitive response*). Essa resposta gera lesões necróticas próximas ao sítio de infecção que impedem que o patógeno se prolifere. A ativação da resposta da planta para o combate ao patógeno depende da eficiência do hospedeiro em reconhecer a presença do organismo invasor, através de mecanismos de percepção de transdução de sinais, que envolvem alterações transitórias no fluxo de íons através da membrana plasmática, acúmulo de espécies reativas de oxigênio (ROS) e de óxido nítrico (NO). A partir dessa cadeia de transdução de sinal ocorrem mudanças no estado de fosforilação de várias proteínas e que tem como consequência a ativação de fatores de transcrição no núcleo da célula vegetal, com subsequente expressão de genes de defesa da planta.

Após a infecção pelos patógenos, ocorre a sinalização de defesa sistêmica conhecida como SAR (*Systemic acquired resistance*), que confere uma resistência de longa duração, em alguns casos de amplo espectro, sendo efetiva contra doenças causadas por vírus, bactérias e fungos (Ryals *et al.*, 1996). SAR é uma das vias de defesa da planta que pode ter sua origem sinalizada pela HR, que cria várias rotas de defesa da planta, pois induz a uma rede de sinalizações celulares e à ativação de uma cascata de sinalização dependente de SA.

As cascatas de sinalização via HR e SAR ativam a expressão de várias proteínas PRs (*pathogenicity related proteins*) que agem no combate contra os patógenos. As PRs são classificadas em 17 famílias, às quais podemos destacar a PR-1 (função desconhecida), PR-2 (β -glucanases), PR-3 (quitinases classe I) e PR5 (Taumatina) (Stinzi *et al.*, 1993). Surpreendentemente, análises das sequências de ESTs de *M. pernicioso* demonstraram que o fungo expressa um gene que codifica uma proteína similar a proteínas da família PR-5, as proteínas do tipo taumatina. Montagens recentes no genoma de *M. pernicioso* possibilitaram encontrar mais oito sequências similares a taumatinas, número este que subiu para 13 no total depois que a montagem do genoma foi enriquecida por dados originados de RNAseq. (Teixeira, 2012).

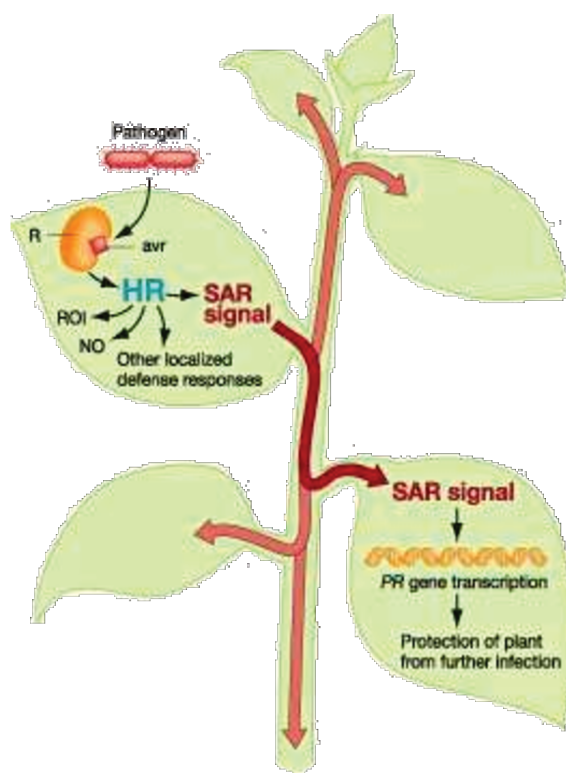


Figura 4. Desenho esquemático representando o reconhecimento de proteínas AVR do patógeno por proteínas R do hospedeiro. A interação resulta em uma HR (hypersensitive response), para conter o patógeno no local da infecção, e ativa uma cascata de sinalização, que resulta em SAR (*Systemic acquired resistance*) e ativam a produção de PRs (proteínas relacionadas a patogênese). Figura extraída de http://wps.prenhall.com/esm_freeman_biosci_1/7/1955/500507.cw/index.html

1.4- Proteínas Efetoras de patogenicidade

Proteínas efetoras de patogenicidade são codificadas por genes AVR_s (Stergiopoulos e de Wit, 2009), secretadas por patógenos dentro do organismo hospedeiro, com o objetivo de aumentar seu poder de infecção (Duda and Rajem 2010). Essas proteínas agem modificando a estrutura das células do hospedeiro, controlando processos celulares, metabolismo e função, e interferindo na sinalização de defesa de planta suprimindo PTI. Alguns efetores podem ser secretados por apressórios antes da penetração do patógeno na célula hospedeira, entrando nas células vegetais, atingindo diversos compartimentos celulares, e outros se movem até células vizinhas para prepará-las antes da invasão (Giraldo e Valetín, 2013). Essas moléculas são capazes, portanto, de reprogramar a defesa da planta e, em plantas resistentes, os efetores são diretamente ou indiretamente, reconhecidos por proteínas de resistência que residem dentro e fora da célula da planta ou em membranas plasmáticas.

Os efetores de fungos hemibiotróficos geralmente possuem sequência sinalizadora para secreção, e são em sua maioria proteínas ricas em cisteínas. Esses resíduos de cisteínas estão envolvidos na formação de pontes dissulfeto que promovem a estabilidade da proteína no ambiente hostil rico em protease no apoplasto do hospedeiro (Stergiopoulos e de Wit, 2009). São proteínas pequenas, e sua expressão é específica durante a invasão celular em seu estágio biotrófico (de Long et al., 2011; Rafiq et al., 2012, Djamei e Kahmann, 2012). Em oomicetes, efetores possuem motivos de aminoácidos associados à translocação do hospedeiro (RXLR, onde X representa qualquer aminoácido, e CRN), muito próximo ao peptídeo sinal. Outras ainda apresentam motivos RSIDELD, CHXC e Y/F/WXC (Kale et al., 2010; Godfrey et al., 2010; Zuccaro et al., 2011; Kemen et al., 2011), em todos os casos relacionados penetração dessas moléculas nas células do hospedeiro, e em muitos casos através de haustórios. A

Maioria dos efetores são proteínas pequenas que perderam esses motivos, impossibilitando a predição de função dessas proteínas (Bozkurt et al., 2012). Além disso, possuem uma baixa conservação nas sequências que codificam essas proteínas.

Efetores fúngicos podem ser grosseiramente agrupados em efetores extracelulares, que são secretados para o apoplasto ou xilema da planta hospedeira, e efetores citoplasmáticos, que são translocados para dentro da célula vegetal, ou utilizam-se de estruturas físicas como haustórios para penetrar nas células, (Stergiopoulos e de Wit, 2009). Um exemplo de um efector não secretado por fungo vem de *M. grisea*, mas é sugerido que ele está envolvido na biossíntese de um metabólito secundário desconhecido o qual é secretado (Bohnert et al, 2004). Efetores apoplásticos de diversos patógenos podem ser usados para identificação de alvos de defesa de plantas (Doehlemann and Hemestberger, 2013, Tomma et al., 2012) . Exemplos desses efetores são inibidores de proteases, efetores ligantes de quitina e inibidores de peroxidases.

A versão atual do genoma de *M. perniciosa* codifica 35 proteínas candidatas a efetores secretados nomeadas *MpCSEPs* (*Moniliophthora perniciosa* candidate secreted effector proteins) que são definidas como proteínas secretadas que não possuem homólogos caracterizados em outros organismos sequenciados. A maioria desses genes são exclusivamente e altamente expressos durante a interação biotrófica do fungo com o cacaueiro. Em geral, esses genes codificam para proteínas pequenas ricas em cisteínas (tamanho médio de 179 aminoácidos, % Cys= 4,3%), que são características típicas de efetores de virulência (Stergiopoulos and de Wit, 2009).

1.5- Proteínas do tipo taumatina

TLPs (*Thaumatococcus-like protein*) são polipeptídeos de aproximadamente 200 aminoácidos (21 a 26 KDa), que possuem grande similaridade a sequência de Taumatina (Velazhahan et al. 1999). A taumatina é uma proteína monomérica de sabor doce que

contém oito pontes dissulfeto (Grenier et al. 1999), isolada do fruto da planta africana *Thaumatococcus danielli* (Van der Wel e Loewe 1972).

Proteínas similares a taumatinas (TLP, *thaumatin-like protein*) possuem domínios conservados que formam um *acidic cleft* em sua estrutura terciária, o qual está relacionado à atividade enzimática das TLPs. Algumas TLPs possuem deleções de peptídeos em um de seus domínios que culminam na ausência de um *acidic cleft* típico. Essas proteínas são conhecidas como *Small TLPs* (sTLPs) encontradas em cereais, coníferas e fungos, possuindo peso molecular de aproximadamente 16-17 KDa (Petre et al. 2011).

Proteínas similares a taumatinas são o produto de uma grande família gênica de alta complexidade que estão envolvidas em defesa de plantas contra patógenos, e também em um amplo espectro de desenvolvimento em fungos, plantas e animais (Liu et al., 2010). Essas proteínas pertencem à classe de proteínas de defesa de plantas denominadas PR-5. Embora essas proteínas sejam comercializadas como adoçantes, em sua maioria as TLPs não apresentam sabor adocicado (Benko-Iseppon *et al.*, 2003). Segundo Kaneko e Kitabatake (2001), a característica doce apresentada pela taumatina tem uma forte relação com seus resíduos de lisina. O autor promoveu modificações gradativas em resíduos de lisina por fosfopirodoxilação, e após análises sensoriais, foi verificada a perda do sabor doce. Após o reestabelecimento dos resíduos de lisina pela desfosforilação das fosfopirodoxilações, o sabor doce foi restabelecido. Foi sugerido também que a estrutura secundária da taumatina não está relacionada com o sabor doce, já que alterações em grupos carboxílicos da proteína não diminuíram seu sabor doce (Kaneko e Kitabatake, 2001). Alguns transgênicos com taumatina apresentaram aumento do sabor doce (Bartoszewski et al. 2003; Wittyand e Harvey 1990). Ohta e colaboradores (2011) relacionam domínios ricos em cisteínas de proteínas receptoras de sabor presentes nas proteínas T1R2 e T1R3 em humanos, sendo responsáveis pela

interação entre elas, e somente no caso dessa interação é possível detectar o sabor doce de taumatina.

As TLPs podem ser induzidas por vários fatores de estresse, como infecção por fungos (Pritsch *et al.*, 2000), ferimentos (Ruperti *et al.*, 2002) e estresse por seca (Jung *et al.*, 2005). Estudos demonstram que as TLPs de plantas possuem atividade antifúngica (Hu e Reddy, 1997, Ye *et al.*, 1999, Lam e Ng, 2001). Liu *et al.* (2010) listou TLPs de animais, plantas e fungos que também demonstraram propriedade antifúngica. Apesar de as TLPs de plantas demonstrarem capacidade de degradar β -1,3-glucanas, polímeros compostos de resíduos de glicose característicos de parede celular de fungos, essa atividade nem sempre está relacionada a capacidade antifúngica de TLPs (van Loon *et al.* 2006). Permeabilização de membranas (Virges *et al.* 1991), mecanismos de indução de apoptose (Narasimhan *et al.* 2005) e inibição de enzimas como tripsina, α -amilase e xilanase foram descritas como atividades que podem ocasionar uma ação antifúngica (Schimoler *et al.* 2001; Fierens *et al.* 2007). Estudos mostram que algumas plantas transgênicas expressando constitutivamente taumatina mostraram um aumento na resistência a fungos (Datta *et al.*, 1999; Fagoaga *et al.*, 2001; Kalpana *et al.*, 2006).

Uma TLP que apresenta atividade inibidora de xilanase (TLXI) de trigo foi descrita por Fierens e colaboradores (2007). Os autores atribuem essa nova atividade de uma TLP a um resíduo de histidina presente na posição 22 da proteína, já que ao sofrer mutação nesse resíduo de histidina, a proteína resultante não foi capaz de formar complexo com GH11- xilanase, o que ocorre normalmente na proteína não mutada (Rombouts *et al.*, 2008). Além disso, foi verificado que TLXI é um inibidor não competitivo (não se liga ao sítio ativo da enzima inibida) e que é capaz de se ligar ao substrato da reação (Fierens *et al.*, 2008). Xilanases são glicosidases (O-glycoside hydrolases) que catalisam a hidrólise das ligações 1,4-b-D-xilosídicas em xilanas.

Xilanases estão envolvidas na produção de xilose, uma fonte de carbono essencial para o metabolismo celular e na infecção de patógenos, sendo produzidas por bactérias, fungos e insetos (Collins *et al.*, 2005; Belien *et al.*, 2006). Brito *et al.*, (2006) verificaram que a xilanase xyn11A produzida por *Botrytis cinerea* é importante para a virulência desse fungo necrotrófico. Xilanases vegetais foram relacionadas ao amadurecimento de frutos e na germinação de sementes (Manenoi e Paull, 2007; Caspers *et al.*, 2001). Além disso, *A. thaliana* possui xilanases expressas no sistema vascular (Suzuki *et al.*, 2002). Outras enzimas hidrolíticas produzidas pelos vegetais são as β 1-3 glucanases (PR-2). Essas enzimas clivam a parede celular fúngica inibindo o crescimento do patógeno. Além disso, essa clivagem gera açúcares livres (beta-glucanas) que são sinalizadores de defesa vegetal (York *et al.*, 2003). Estudos cristalográficos estão sendo utilizados para determinar a estrutura tridimensional dessas proteínas.

Alguns estudos demonstram que TLPs exibem atividade anticongelante, como as AFPs (*antifreeze proteins*). Isso foi demonstrado através do isolamento de TLPs de centeio durante aclimação ao frio (Hiilovaara-Teijo *et al.*, 1999; Yu e Griffith, 1999). Outros experimentos demonstraram que uma TLP da larva do besouro *Dendroides canadensis* agiria como realçadora da atividade de AFPs, interagindo com estas, formando complexos maiores que as AFPs sozinhas (Wang e Duman, 2006).

TLPs estão amplamente descritas em plantas, e estão sendo descobertas em muitos organismos, além de plantas (Shatters *et al.* 2006), incluindo nematódeos (Kitajima and Sato 1999), insetos (Brandazza *et al.* 2004) e fungos (Grenier *et al.* 2000; Sakamoto *et al.* 2006). Porém, pouco é descrito sobre a atuação de TLPs em fungos. Grenier *et al.* (2000) foi o primeiro a reportar a presença de TLPs em fungos basidiomicetos. Em estudos mais recentes, TLG1, uma TLP do cogumelo shitaque (*Lentinula edodes*) foi caracterizado (Sakamoto *et al.* 2006). Essas proteínas não foram detectadas em micélio vegetativo ou em ascos novos, mas sua expressão aumenta em corpo de frutificação

durante a difusão de esporos. TLG1 possui atividade β -1,3-glucanase e demonstra atividade degradativa em lentinan, um componente da parede celular de *L. edodes* compreendendo ligações β -1,3 beta-glucanas com ramificações β -1,6. O autor sugere que essas atividades podem estar relacionadas à lise da parede das células do fungo para que ocorra a formação dos primórdios de basidiocarpo.

2- OBJETIVOS GERAIS

A presente tese teve como objetivo geral a caracterização de duas famílias de genes relacionados à patogenicidade em *M. perniciosa*: *MpTLPs* (*Moniliophthora perniciosa* *Thaumatococcus*- *like* *Protein*) e *MpCSEPs* (*Moniliophthora perniciosa* *Candidate Secreted Effectors*).

A tese será apresentada em dois capítulos. O primeiro capítulo aborda estudos relacionados a proteínas similares a taumatinas em *M. perniciosa* (*MpTLPs*). O segundo capítulo aborda estudos referentes à *MpCSEP* (*Moniliophthora perniciosa* *Candidate secreted effectors*). Em linhas gerais, o trabalho realizado durante este doutoramento seguiu o seguinte organograma.

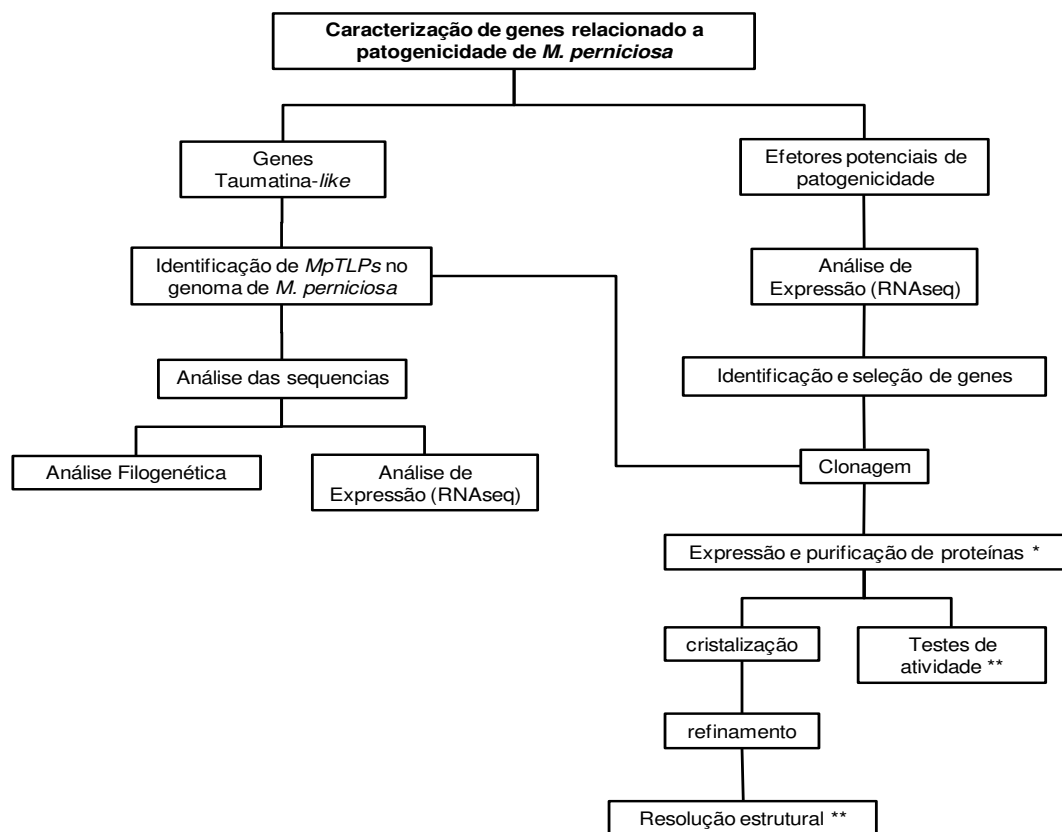


Figura 5: Cronograma do trabalho abordado na presente tese.

* - Etapa limitante para a caracterização das MpTLPs e algumas MpCSEPs.

** - Etapas que serão continuadas em próximos trabalhos.

3- CAPÍTULO I: Caracterização da família gênica Taumatina- like de *M. perniciosa*.

3.1- Apresentação

Esse capítulo refere-se ao estudo da família gênica de Taumatinas de *M. Perniciosa*. Esse projeto começou com o objetivo da caracterização gênica, molecular e funcional de cada membro da família *MpTLPs*. O estudo da família gênica que englobou análise filogenética e expressão gênica das *MpTLPs* resultou no artigo "The large family of thaumatin-like genes in the fungal pathogen *Moniliophthora perniciosa* and its putative roles in witches broom disease of cacao", submetido recentemente à revista Molecular Genetics and Genomics. A segunda parte do presente capítulo aborda a clonagem, expressão e purificação de *MpTLPs*.

3.2- Justificativas

Taumatinas são proteínas descritas como PR5 (relacionadas à patogenicidade classe cinco) em plantas, sendo descritas como proteínas de defesas de plantas e induzidas sob condições de estresse e pela exposição de plantas a patógenos. São proteínas típicas de defesa de planta, podendo apresentar diversas atividades. Entre elas, atividade antifúngica, β - glucanólítica e inibidora de xilanase. Estudos recentes descrevem TLPs em outros organismos, incluindo fungos basidiomicetes. Pouco é descrito sobre a atuação dessas proteínas durante o ciclo de vida de fungos. O Estudo da família gênica *MpTLP* se faz importante por tratar de um tipo de proteína que possui um amplo espectro de atividades, e ainda, por se tratar de uma família gênica com muitos representantes (13 no total). Consideramos ainda a possibilidade de que membros da mesma família podem desempenhar funções diferentes durante o desenvolvimento da doença.

3.3- Objetivo específico

O objetivo desse trabalho foi a identificação e análise de sequências de genes similares a taumatinas (TLPs) em *M. pernicioso*, bem como a caracterização molecular e funcional das proteínas *MpTLPs*.

3.4- Artigo submetido à revista *Molecular Genetics and Genomics*:

The large family of thaumatin-like genes in the fungal pathogen *Moniliophthora perniciosa* and its putative roles in witches broom disease of cacao

Sulamita de Freitas Franco¹, Marcelo Falsarella Carazzolle^{1#}, Paulo José Pereira Lima Teixeira^{1#}, Osvaldo Reis¹, Gonçalo Amarante Guimarães Pereira^{1*}, Jorge Maurício Costa Mondego²

1 - Laboratório de Genômica e Expressão, Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, UNICAMP, Campinas, SP, Brasil

2 - Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Instituto Agronômico, Campinas, SP, Brasil

*- Corresponding author: goncalo@unicamp.br

- These authors contributed equally to this paper

Keywords: Thaumatins, *Moniliophthora perniciosa*, cacao, gene family, gene expansion, RNAseq.

Abstract

Moniliophthora perniciosa is a hemibiotrophic basidiomycete pathogen that is the etiological agent of Witches Broom Disease (WBD) of cacao. The inspection of *M. perniciosa* genome draft resulted in the annotation of thirteen genes encoding Thaumatin-like proteins (TLPs), which are part of a wide gene family present in plants, fungi and animals. In plants, these proteins are known to act as fungal inhibitors, but more recent reports in basidiomycete indicate an involvement in mushroom cell wall remodeling. Based on the analysis of fungal genomes, *Moniliophthora perniciosa* was found to be the fungus species with largest number of TLPs (13 genes), with six of them being are organized in three gene clusters, indicating events of gene duplication. Phylogenetic data revealed a large number of TLPs in basidiomycete when comparing with ascomycete, supporting a participation in specific traits of mushroom-forming species. We found a protein (MpTLP12) containing structure traits such as the small protein length (similar to small TLPs), the absence of amino acid residues composing the REDDD motif and the lack of the conserved acidic cleft of TLPs. By inspecting *M. perniciosa* RNAseq atlas we found that each *TLP* present specific expression profiles but most of them are transcribed during late stages of WBD, when infected stems are dried ('dry brooms') and opportunistic fungus are detected in this tissue. We discuss the possible contribution of TLPs as components for the combat against fungal contenders, in fungal cell wall development and plant cell wall degradation.

Introduction

Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora is a basidiomycete pathogen that causes Witches Broom Disease (WBD) of cacao (*Theobroma cacao* L.), (Aime and Phillips-Mora 2005), one of the most devastating phytopathological problems in the Americas. The disease is responsible for leading to serious economical problems in

producing regions with losses reaching from 50% to 90% of production (Meinhardt *et al.* 2008; Purdy and Schmidt, 1996). This fungus has a hemibiotrophic lifestyle, and presents two morphologically distinct mycelia. The disease begins when basidiospore comes into contact with healthy meristematic cacao tissues, flower cushions or developing fruit (Evans, 1980). Basidiospore germination originates monokaryotic biotrophic (parasitic) mycelia that slowly grow in the intercellular space. During this stage of the interaction, hypertrophy and hyperplasia of the meristematic tissues is observed in an anomalous structure named as 'green broom' (Evans, 1978, Penman *et al.* 2000, Silva, 1999). After two to three months, the fungus becomes dikaryotic and clamp connections. These necrotrophic mycelia invade the plant cell and necrosis occurs in the infected tissues causing the "dry broom" aspect (Evans and Bastos, 1980; Frias *et al.* 1991). After alternate dry and wet periods, the fungus produces basidiomata that release basidiospores, reinitiating the cycle. In fruits, *M. pernicioso* also cause necrosis of the infected tissues (Baker and Holliday 1957).

In 2000, the Witches' broom Genome Project was initiated aiming to decode the *M. pernicioso* genome. Since then, many reports have been elucidating the mechanisms by which the fungus infects the cacao, and also the plant biochemical modifications related to the WBD (Scarpari *et al.* 2005; Rio *et al.* 2008; Zaparoli *et al.* 2011; de Oliveira *et al.* 2012; Teixeira *et al.* 2012; Barsottini *et al.* 2013). The molecular pathways of the trigger that makes the fungus change from biotrophic to a necrotrophic mycelia are not completely elucidated; however a recent study pinpoint that an alternative oxidase (AOX) plays a role in the biotrophic development of *M. pernicioso* and regulates the transition from parasitic to necrotrophic stage (Thomazella *et al.* 2012).

Pathogenicity related proteins (PR proteins) are plant proteins expressed during hypersensitive response (HR) and Systemic acquired resistance (SAR) involved in the combat against pathogens (Van loon *et al.* 2006). PRs are classified into 17 families,

which include PR-5 proteins, also known as Thaumatin-like proteins (TLPs; Stinzi et al. 1993). TLPs are polypeptides of approximately 200 amino acid residues (21 to 26 kDa) that share sequence similarity with thaumatin (Velazhahan et al. 1999), a monomeric sweet-taste protein that contains eight disulfide bonds (Grenier et al. 1999), originally isolated in the fruit of the African plant *Thaumatococcus danielli* (Van der Wel and Loewe 1972). These proteins have conserved domains that form an acidic cleft in their tertiary structure. This cleft is related to the enzymatic activity of TLPs (Liu et al. 2010). Small TLPs (sTLPs) found in cereals, conifers and fungi have molecular mass of 16-17 kDa and have a peptide deletion that abolishes the formation of typical acidic cleft (Petre et al. 2011).

Up-regulation of TLPs has been described in many higher plants infected by pathogens, including bacteria, oomycetes and fungi (Pritsch et al. 2000; Liu et al. 2010, Mukherjee et al. 2010). Subcellular Localization studies revealed that plant pathogen-inducible TLPs are secreted to the apoplast (Islam et al. 2009; Wang et al. 2010). Despite being more studied in plants, TLPs have been discovered in a wide range of organisms (Shatters et al. 2006), including nematodes (Kitajima and Sato 1999), insects (Brandazza et al. 2004) and fungi (Grenier et al. 2000; Sakamoto et al. 2006). In a review Liu et al (2010) lists TLPs from animals, plant and fungi that have been shown to exhibit antifungal activity. Even though plant TLPs were found to degrade β -1,3-glucans, polymers consisting of glucose residues characteristic of fungal cell walls, this activity is not always related to their anti-fungal capacity (van Loon et al. 2006). Membrane permeabilization (Virges et al. 1991), apoptosis-inducing mechanism (Narasimhan et al. 2005) and inhibition of enzymes such as trypsin, α -amylase and xylanase were described as activities that can entail to antifungal properties (Schimoler et al. 2001; Fierens et al. 2007).

Grenier et al. (2000) were the first to report the presence of TLPs in basidiomycete fungi. More recently, TLG1, a TLP of the shitake mushroom (*Lentinula edodes*) was

characterized (Sakamoto et al. 2006). This protein was not detected in vegetative mycelium or young and fresh mushrooms, but its expression increases following mushroom harvest and spore diffusion. TLG1 exhibited β -1,3-glucanase activity and demonstrated degrading activity against lentinan, a *L. edodes* cell wall component comprising β -1,3 beta-glucan with β -1,6 branching.

This manuscript reports the identification of a TLP gene family in the *M. perniciosa* genome (*MpTLPs*). The annotation of publicly available fungal genomes revealed that TLPs presence seems to be pervasive in basidiomycete fungal genomes. We also performed a phylogenetical analysis, evaluated structural features of MpTLPs through protein modelling, and verified their *in-planta* expression profile as well as in *in vitro* stages of *M. perniciosa* development. Based on these results, we provide insights that reinforce the participation of these proteins in avoidance of fungal competitors and in fungal cell wall development and suggest a possible contribution in plant cell wall degradation and plant defense suppression.

Materials and methods

***M. perniciosa* Genome and transcriptome analyses**

One TLP gene (*MpTLP1*) was identified during the annotation of Sanger *M. perniciosa* EST libraries (Rincones et al. 2008). Using the *MpTLP1* sequence as bait, tBLASTN searches were performed in a draft assembly of the *M. perniciosa* genome obtained by Sanger methodology (Mondego et al. 2008). This search led to the identification of 3 additional TLP genes with high similarity to *MpTLP1*. With the emergence of new DNA sequence technologies, the *M. perniciosa* genome assembly was improved with a dataset from 454-pyrosequencing and Illumina sequencing and additional BLAST searches revealed 9 new TLP sequences (<http://www.lge.ibi.unicamp.br/vassoura/bib>). The complete open reading frames (ORFs) of

these genes were predicted by Augustus software (Stanke et al. 2004) and intron-exon boundaries were confirmed with alignment with RNAseq libraries.

Search for TLP in public genomic databases and sequence analyses

Forty six fungal genome sequences available on the Joint Genome Institute (JGI, <http://www.jgi.doe.gov>) in the Broad Institute (<http://www.broadinstitute.org>) and in the NCBI protein websites were mined using MpTLP sequences and a maize thaumatin (zeamatin; NP_001105356.1) sequences (Roberts and Selitrennikof, 1990) as baits in BLAST searches. Similar sequences to *MpTLP* were retrieved from filtered models databases from JGI and Broad (Cut-off E-value was $< 1 \text{ E}^{-10}$). *TLP* sequences from plants, arthropods and nematodes present in NCBI database were individually examined and amino acid sequences with a complete thaumatin family domain were retained for further analyses. For all amino acid sequence comparisons, the procedure described by Petre *et al.* (2012) was applied, limiting the TLP domain as N-x-C-x (3)-V/I-W and Y-x-I/V-x-F-C-x in the N- and C-terminal ends, respectively. Only proteins covering almost 95% of the mature TLP domain were considered.

Phylogenetic analyses

Sequence alignment was carried out using CLUSTAL W 1.8 (Thompson et al. 1994) and edited in the GeneDoc platform (Nicholas and Nicholas 1997). The weighing matrix used was BLOSUM62 and the alignments generated were manually adjusted following the domain signatures and cysteine position described in TLPs. The phylogenetic tree was inferred by using the program MEGA version 5 (Tamura et al. 2007). The tree was constructed as a consensus of 2,000 bootstrap replicates by means of the Neighbour-Joining tree inference with dist-p for amino acid substitutions and pairwise deletion option for handling alignment gaps.

Protein Modelling

The fold recognition-based method was implemented using the I-TASSER metasever (Roy et al. 2010), which constructed structure models for MpTLP1 and MpTLP12 using folds of the most similar proteins deposited in the PDB (Protein Data Bank) database (<http://www.rcsb.org/pdb>). The main templates were based on the structure of *Prunus avium* thaumatin (2ahnA). The modeled structures were displayed using the software PyMOL (DeLano et al. 2002). PDB2PQR software (Dolinsky et al. 2004; Dolinsky et al. 2007; <http://www.poissonboltzmann.org/pdb2pqr>) was applied to prepare structures for continuum electrostatics calculations which were visualized through APBS (Adaptive Poisson-Boltzmann Solver) visualization (Baker 2001; <http://www.poissonboltzmann.org/apbs>).

Biological material

The basidiomata of *M. perniciosa* were produced in the laboratory according to procedures previously described (Pires et al. 2009). Basidiospores were harvested in the following solution: 16% glycerol, 0.01 M MES and 0.01% Tween, pH 6.1 (Frias et al. 1995). *Moniliophthora perniciosa* isolate BP10 and *Theobroma cacao* cv. “Comum” (Forasteiro Genotype) were used to perform the infection experiments. Plantlets (three month old) were inoculated with 30 µL of a basidiospores suspension (1×10^5 spores mL⁻¹) following the procedures described by Frias et al. (1995). Plantlets were kept in a greenhouse under controlled conditions of temperature (26°C) and humidity (>80%), and were collected 30 days after inoculation (green brooms), 47 days after inoculation (initial necrosis; necrosis 1), 66 days after inoculation (advanced necrosis; necrosis 2) and 111 days after inoculation (dry brooms). Fruits from *T. cacao* cv. “Comum” trees presenting symptoms of *M. perniciosa* infection were collected in a cacao farm located in Ilhéus (14°47'20"S, 39°02'56"W), Brazil. Fruits displaying the disease early stage symptom (disordered fruit ripening), partial necrosis, and complete rot were collected. All samples were frozen in liquid nitrogen prior to RNA extraction.

RNA isolation and RNA-seq libraries preparation and analysis

Samples were ground to a fine powder in liquid nitrogen using a pestle and mortar; except basidiospores, which walls were broken by vortexing the sample in RNA extraction buffer (Buffer RLT, RNeasy Plant Mini Kit) and 200 mg glass beads (0.4–0.6 mm, Sigma-Aldrich, St. Louis, MO, EUA). RNA isolation was performed using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Procedures of RNAseq libraries preparation are described elsewhere (Barsottini et al. 2013). These results are part of WBD Transcriptome Atlas (Teixeira et al. unpublished data).

Results

Genome Survey and TLPs annotation

Data mining of *M. pernicios*a Sanger-sequenced genome draft, fungal EST libraries assessment, next generation genome sequencing and RNAseq data from the WBD transcriptome atlas allowed the identification of thirteen TLPs in this fungus. Overall, these members are homogeneous in size and gene structure, with the exception of *MpTLP11* and *MpTLP12* (Table 1). For instance, *MpTLP1* coding sequence (CDS) has 771 bp while *MpTLP11* and *MpTLP12* have 558 pb and 498 bp, respectively (Table 1). This difference in CDS length reflects in protein size and molecular weight, being *MpTLP11* and *MpTLP12* the smallest MpTLPs, resembling small TLPs (sTLPs) from puccinales fungi (Petre *et al.* 2012). We detected three gene clusters of *MpTLPs* in the genome of *M. pernicios*a. Each cluster contain two MpTLPs *in tandem* (*MpTLP1* and *MpTLP2*; *MpTLP3* and *MpTLP4*; *MpTLP11* and *MpTLP12*; Fig.1). Curiously, a β -1,3 glucan synthase gene is located upstream the *MpTLP1* gene (Fig. 1).

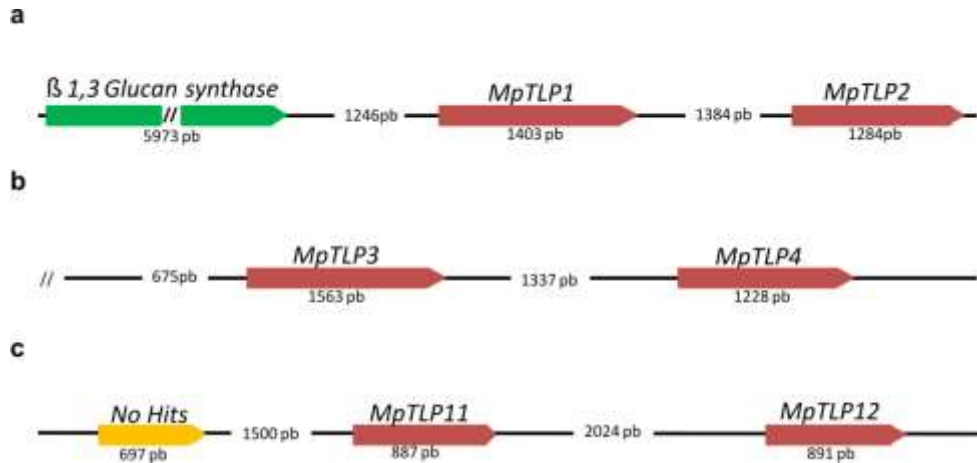


Figure 1: Genomic organization of the *MpTLPs* gene clusters found in *M. pernicioso*. (a) *MpTLP1* and *MpTLP2* are located in the vicinity of a β -1,3-glucan synthase. (b) *MpTLP3* and *MpTLP4* are located in the edge of a supercontig. (c) *MpTLP11* and *MpTLP12* are positioned near a *no hits* (orphan) gene. Lengths of genes (in bp) are depicted below gene representations. Distances of intergenic regions (in bp) are shown below continuous lines. Introns are not indicated in the figure.

As reported for other thaumatins, *MpTLPs* are rich in Cys residues ranging from 11 (*MpTLP11*) to 20 (*MpTLP10*). The majority of the *MpTLPs* has the isoelectric point from 3.9 to 5.55. *MpTLP12* is the only member with a basic PI of 7.98 (table 1). TargetP prediction revealed, hydrophobic signal peptide sequences were predicted with the thirteen *MpTLPs* (NN score 0.80) indicating that these proteins are secreted (data not shown). All *MpTLPs* have a Thaumatin domain (Interpro ID, IPR001938) limited by the signatures N-x-C-x-(3)-V/I-W and Y-x-I/V-x-F-C-x in the N- and C-terminal, respectively. Five highly conserved amino acids (REDDD), which are positioned in an acidic cleft, were detected in most of *MpTLPs*. However, alignment of *MpTLPs* with other thaumatins shown that *MpTLP11* and *MpTLP12* did not contain the totality of these amino acids, replacing second Asp of REDDD motif by a Gln and lacking the Asp/Glu present in the C-terminal of *TLPs* (Fig. 2).

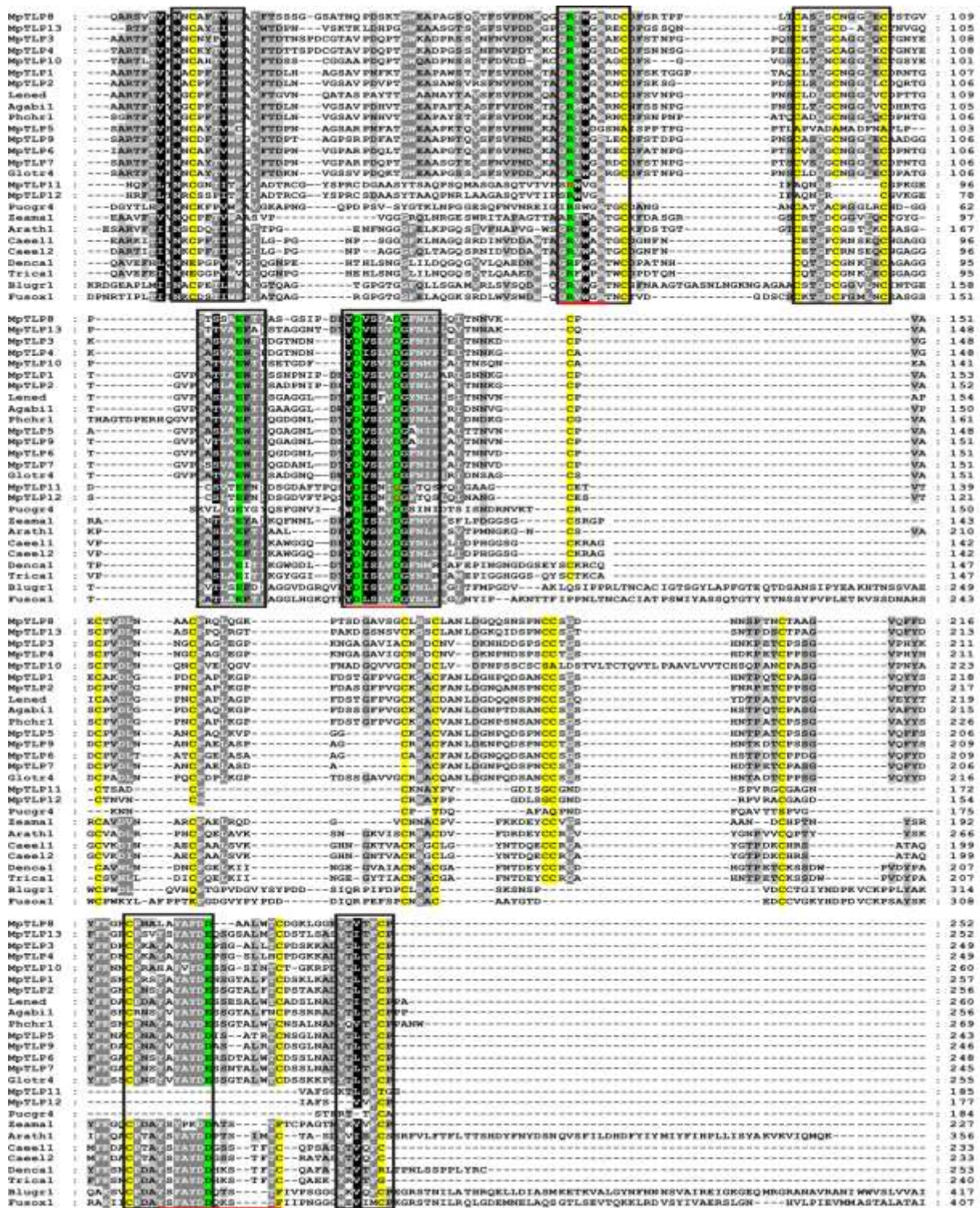


Figure 2: Sequence alignment of MpTLPs with TLPs of representative organisms. Black background, 100% of conservation between amino acids; Light grey background and black letters, conservation between amino acids of 99-80%; Dark grey background and white letters, conservation between amino acids of 79-60%. Conserved cysteines are highlighted in yellow. Amino acid residues from REDDD motif are highlighted in green. Amino acid substitutions in REDDD motif are shown in red. Conserved domains are indicated by boxes. Amino acids forming the bottom of the acidic cleft are underlined in red. The alignment was constructed with Clustal W2 and edited with Gene Doc.

Table 1: Characteristics of the thirteen *MpTLP* genes identified in the *M. pernicioso* genome.

Gene	<i>M. pernicioso</i> Database ID	Genbank ID***	Gene size (bp)	CDS size (bp)	Number of Introns	Protein size (aa)	Cysteine number	BlastP first Hit (NCBI)*	PI/ MW (kDA)**
<i>MpTLP1</i>	MP04497		1403	771	11	257	16	EKV44020.1- hypothetical protein (3e-130) <i>Agaricus bisporus</i>	5.55 / 27.2
<i>MpTLP2</i>	MP04510		1284	771	10	257	16	ELU36916.1- thaumatin-like protein (6e-134) <i>Rhizoctonia solani</i>	4.56 / 27.1
<i>MpTLP3</i>	MP08014		1563	750	8	250	18	EJD01931.1- thaumatin-like protein (5e-94) <i>Fomitiporia mediterranea</i>	4.80 / 26.4
<i>MpTLP4</i>	MP08020		1228	750	8	250	18	EIM87365.1- Osmotin thaumatin-like protein (5e-92) <i>Stereum hirsutum</i>	4.90 / 26.3
<i>MpTLP5</i>	MP07273		1254	732	7	244	13	EKV44094.1- hypothetical protein (1e-92) <i>Agaricus bisporus</i>	4.59 / 25.4
<i>MpTLP6</i>	MP01417		1279	747	7	249	16	EJD04853.1- thaumatin-like protein (5e-121) <i>Fomitiporia mediterranea</i>	3.99 / 26.0
<i>MpTLP7</i>	MP07692		1252	726	8	241	14	EJD04867- 1thaumatin-like protein (7e-80) <i>Fomitiporia mediterranea</i>	3.90 / 25.4
<i>MpTLP8</i>	MP04575		1193	759	7	253	16	EJD04867.1- thaumatin-like protein (2e-90) <i>Fomitiporia mediterranea</i>	4.52 / 26.3
<i>MpTLP9</i>	MP07276		1239	741	8	247	16	EKV44094.1- hypothetical protein (9e-112) <i>Agaricus bisporus</i>	4.00 / 25.7
<i>MpTLP10</i>	MP08739		1485	783	7	261	20	EIW57348.1- thaumatin-like protein (1e-71) <i>Trametes versicolor</i>	4.22 / 27.3
<i>MpTLP11</i>	MP09043		887	558	6	186	11	XP_003034800.1- hypothetical protein (8e- 68) <i>Schizophyllum commune</i>	4.69 / 19.2
<i>MpTLP12</i>	MP09053		891	498	7	164	12	EKV41735.1- hypothetical protein (8e-68) <i>Agaricus bisporus</i>	7.98 / 18.5
<i>MpTLP13</i>	MP12084		1098	759	8	253	16	EMD36872.1-hypothetical protein (2e-93) <i>Ceriporiopsis subvermispora</i>	4.49 / 26.5

* E-values are shown in parentheses.

** Theoretical values from ExPaSy tool (http://web.expasy.org/compute_pi/)

*** These IDS will be released as soon as possible

Evaluation of TLPs in other fungi

Data mining of the forty six fungal genomes, available on the Joint Genome Institute (JGI, <http://www.jgi.doe.gov>) and the Broad Institute (<http://www.broadinstitute.org>) were carried out using MpTLPs and zeamatin sequences as baits in BLASTP in order to retrieve fungal TLPs. *TLPs* were found in all basidiomycetes analyzed (24 species), except in the corn smut pathogen *Ustilago maydis*. *TLPs* are found, in general, in large gene families in basidiomycete. The search here reported identified only one *TLP* per organism in the analyzed ascomycete species (22 species, Table 2).

Protein alignment and Phylogenetic analyses

To evaluate the phylogeny of MpTLPs, the sequences of fungal TLPs retrieved from fungal genome databases were aligned with TLPs from nine arthropodes, from the nematode *C. elegans* and from seven plants. In general, TLPs follow the phylogenetical relationship of analyzed species (Fig. 3). Besides the small TLPs in puccianales fungus *Melampsora larici-populina* and *Puccinia graminis* f.sp. tritici previously described (Petre et al. 2011), it was also identified small TLPs in the agaricomycetes *Coniophora puteana*, *Schizophyllum commune*, *Stereum hirsutum* and *Dichomitus squalens* (sTLPs; wavy line in Fig. 3). In the present analysis, 'small TLPs' MpTLP11 and MpTLP12 are positioned close to sTLPs from puccianales, however their phylogenetical origin seems to be paraphyletic (Fig. 3). The pairs of *MpTLPs* present in genome clusters (MpTLP1 and MpTLP2, MpTLP3 and MpTLP4, MpTLP11 and MpTLP12) are located in the same clades, indicating recent events of duplication. Other clades of MpTLPs are composed by proteins that are not located *in tandem* in the *M. perniciosa* genome (MpTLP5, 6, 7, and 9; MpTLP13 and 8; Fig. 3).

Table 2: Number of TLPs genes in fungal species with different lifestyles.

Basidiomycota	Code	Database	TLP number	Lifestyle
<i>Agaricus bisporus</i> var <i>bisporus</i> (H97) v2.0	Agabi	JGI	3	Saprotrophic
<i>Auricularia delicata</i> SS-5 v1.0	Aurde	JGI	9	Saprotrophic (white rot)
<i>Coniophora puteana</i> v1.0	Conpu	JGI	10	Saprotrophic (brown rot)
<i>Coprinopsis cinerea</i>	Copci	Broad	5	Saprotrophic
<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	Cryne	JGI	3	Human pathogen
<i>Dichomitus squalens</i> v1.0	Dicsq	JGI	4	Saprotrophic (white rot)
<i>Fomitiporia mediterranea</i> v1.0	Fomme	JGI	8	Necrotrophic and saprotrophic (white rot)
<i>Fomitopsis pinicola</i> FP-58527 SS1 v3.0	Fompi	JGI	5	Saprotrophic (brown rot)
<i>Gloeophyllum trabeum</i> v1.0	Glotr	JGI	5	Saprotrophic (brown rot)
<i>Heterobasidion annosum</i> v2.0	Hetan	JGI	4	Hemibiotrophic (root rot) or Biotrophic - saprotrophic
<i>Laccaria bicolor</i> v2.0	Lacbi	JGI	2	Mutualistic ectomycorrhizal
<i>Lentinula edodes</i> *	Lened	NCBI	1	Saprotrophic
<i>Melampsora laricis-populina</i> v1.0	Mella	JGI	3	Biotrophic (leaf rust)
<i>Moniliophthora perniciosa</i>	-		13	Hemibiotrophic (witches broom)
<i>Phanerochaete chrysosporium</i> v2.0	Phchr	JGI	4	Saprotrophic (white rot)
<i>Piriformospora indica</i> DSM 11827	Pirin	JGI	2	Mutualistic endomycorrhizal
<i>Pleurotus ostreatus</i> PC15 v2.0	Pleos	JGI	3	Saprotrophic (white rot)
<i>Postia placenta</i> MAD 698-R v1.0	Pospl	JGI	2	Saprotrophic (brown rot)
<i>Puccinia graminis</i>	Pucgr	Broad	7	Biotrophic (leaf rust)
<i>Rhizoctonia solani</i> AG1-A	Rhiso	NCBI	7	Necrotrophic (root rot, damping-off)
<i>Schizophyllum commune</i> H4-8 v2.0	Schco	JGI	9	Saprotrophic (white rot)
<i>Serpula lacrymans</i> S7.9 v2.0	Serla	JGI	5	Saprotrophic (brown rot)
<i>Stereum hirsutum</i> FP-91666 SS1 v1.0	Stehi	JGI	6	Saprotrophic (white rot)
<i>Trametes versicolor</i> v1.0	Trave	JGI	7	Saprotrophic (white rot)
<i>Ustilago maydis</i>	Ustma	Broad	0	Biotrophic (smut)
<i>Wolfiporia cocos</i> MD-104 SS10 v1.0	Wolco	JGI	4	Saprotrophic (brown rot)
Ascomycota				
<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14	Blugr	NCBI	1	Biotrophic (powdery mildew)
<i>Botryotinia fuckeliana</i> v1.0	Botfu	Broad	1	Necrotrophic (gray mold)
<i>Colletotrichum graminicola</i> M1.001	Colgr	Broad	1	Hemibiotrophic (anthracnose, leaf blight)
<i>Colletotrichum higginsianum</i>	Colhi	Broad	1	Hemibiotrophic (anthracnose)
<i>Fusarium graminearum</i> v1.0	Fusgr	Broad	1	Necrotrophic (root rot)
<i>Fusarium oxysporum</i> v1.0	Fusox	Broad	1	Necrotrophic (root rot)
<i>Gaeumannomyces graminis</i>	Gaugr	Broad	1	Necrotrophic (root rot)
<i>Leptosphaeria maculans</i>	Lepma	JGI	1	Complex lifestyle
<i>Macrophomina phaseolina</i> MS6	Macph	NCBI	1	Necrotrophic (stem rot, root rot)
<i>Magnaporthe oryzae</i>	Magor	Broad	1	Hemibiotrophic (rice blast)
<i>Metarhizium acridum</i>	Metac	NCBI	1	Entomopathogen
<i>Metarhizium anisopliae</i> ARSEF 23	Metan	NCBI	1	Entomopathogen
<i>Nectria haematococca</i> v2.0	Necha	JGI	1	Necrotrophic (root rot)
<i>Neurospora crassa</i> OR74A v2.0	Neucr	Broad	1	Saprotrophic
<i>Phaeosphaeria nodorum</i>	Phano	Broad	1	Necrotrophic (glume blotch)
<i>Podospira anserina</i>	Podan	NCBI	1	Saprotrophic
<i>Pseudocercospora fijiensis</i> v2.0	Psefi	JGI	1	Necrotrophic (black sigatoka)
<i>Pyrenophora tritici-repentis</i>	Pyrtr	JGI	1	Necrotrophic (tan spot)
<i>Sclerotinia sclerotiorum</i> v1.0	Scisc	JGI	1	Necrotrophic (white mold)
<i>Talaromyces marneffei</i> ATCC 18224	Talma	JGI	1	Human pathogen
<i>Trichoderma reesei</i> v2.0	Trire	JGI	1	Saprotrophic
<i>Zymoseptoria tritici</i> v2.0	Zymtr	JGI	1	Necrotrophic (septoria leaf blotch)

* Genome not sequenced yet

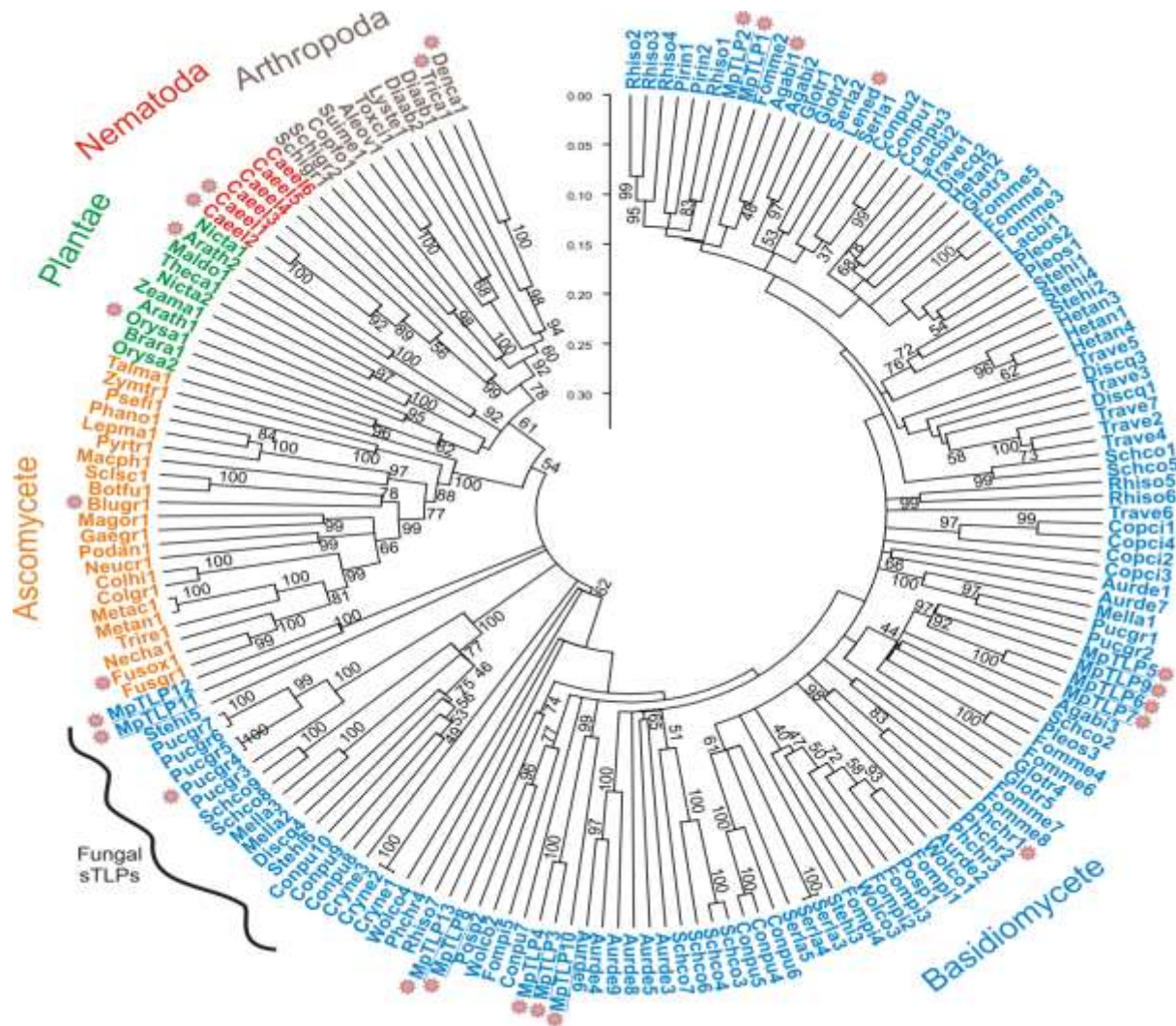


Figure 3: Phylogenetic analysis focusing in fungal TLPs. The phylogenetic tree was constructed as a consensus of 2,000 bootstrap replicates using Neighbor-Joining and dist-p parameters using MEGA5 software. Stars indicate sequences used for alignment shown in figure 2. Bootstrap values > 40 are shown on the tree. Scale indicates amino acid substitutions per site. Codes for fungal TLPs (basidiomycete and ascomycete) are described in table 2. Codes for non-fungal sequences can be found in Additional file 1.

Expression profile of MpTLPs genes throughout *M. perniciosa* development

MpTLPs expression patterns during cacao infection in stems (witches broom, Fig 4A), fruits (Fig. 4B) and *in vitro* grown fungi (Fig. 4c) were assessed by RNA-seq data analysis from WBD Transcriptome Atlas (Teixeira et al. unpublished data). In *in vitro* cultures, *MpTLP12* has the highest expression in basidiospores and germinating basidiospores.

MpTLP8 expression level is relatively equal throughout developmental stages analyzed, while *MpTLP1* was expressed almost exclusively in basidiospores (Fig. 4a). Green house infected plants analysis shows *MpTLP2* presenting the highest expression among TLPs increasing its transcription during WBD, with in the highest expression in the dry broom (1400 RPKM; Fig. 4b). In fact, most of TLPs have their highest expression at the dry broom stage, with emphasis to *MpTLP4*, *MpTLP7* and *MpTLP13*. *MpTLP12* was specifically expressed at the necrosis 2 stage (Fig. 4b). By cheking infected fruit-field data, we detected that *MpTLP1* and *MpTLP2* were the most expressed genes in rotten fruits (Fig. 4c).

Protein structure modeling

In order to compare the tertiary structure of the *MpTLP2* and *MpTLP12*, which are *MpTLPs* with distinct expression profiles and secondary structure features, fold prediction I-TASSER metaserver was used to generate homology models (Fig. 5). *MpTLP1* Ribbon stick (Fig. 5a) and electrostatic surface visualization (Fig. 5c) indicate the presence of the TLPs acidic cleft, while *MpTLP12* seems to lack this tertiary domain (Fig 5b,d). REDDD sites are present in *MpTLP1*, despite silght different positioning as expected in relation to the acidic cleft. For instance, *MpTLP1* Glu₂₃₄ apparently is oriented outside of this cleft (Fig. 5a). *MpTLP12* present a very distinct structure, which is consequence of the absence of the C-terminal region that makes part of the bottom of the acidic cleft (Fig. 5d). Residues from the REDDD site are not found in *MpTLP12* which lacks the C-terminal Asp/Glu and replaces the hypothetical Asp/Glu in position 118 by a Gln (Fig. 5b).

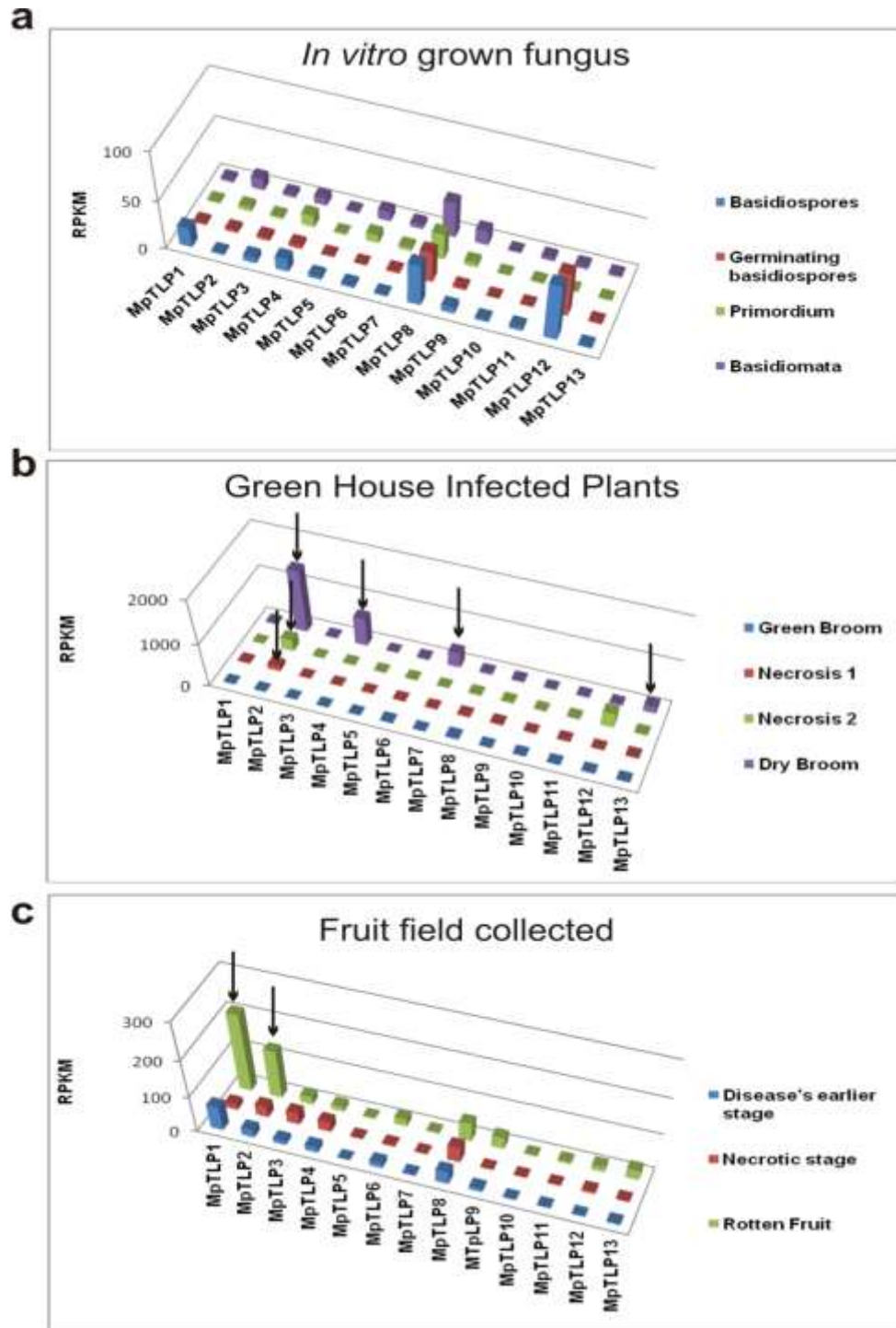


Figure 4: Differential expression of MpTLPs throughout the stages of fungal development and infection. (a) *In vitro* grown fungus. *MpTLP8* and *MpTLP12* have the highest expression among *MpTLPs*. *MpTLP12* was expressed only in basidiospores (germinated and germinated) (b) Green house infected plants (WBD). *MpTLP2* has an increasing of expression throughout WBD, culminating in dry broom stage. *MpTLP4*, *MpTLP7*, *MpTLP13* are also highly expressed in dry brooms. *MpTLP12* is specifically expressed in necrosis 2 stage of WBD. (c) Fruit field collected. *MpTLP1* and *MpTLP2* are predominantly expressed in rotten fruits. Arrows indicate genes with RPKM (reads per kilobase of exon model per million mapped reads) > 100.

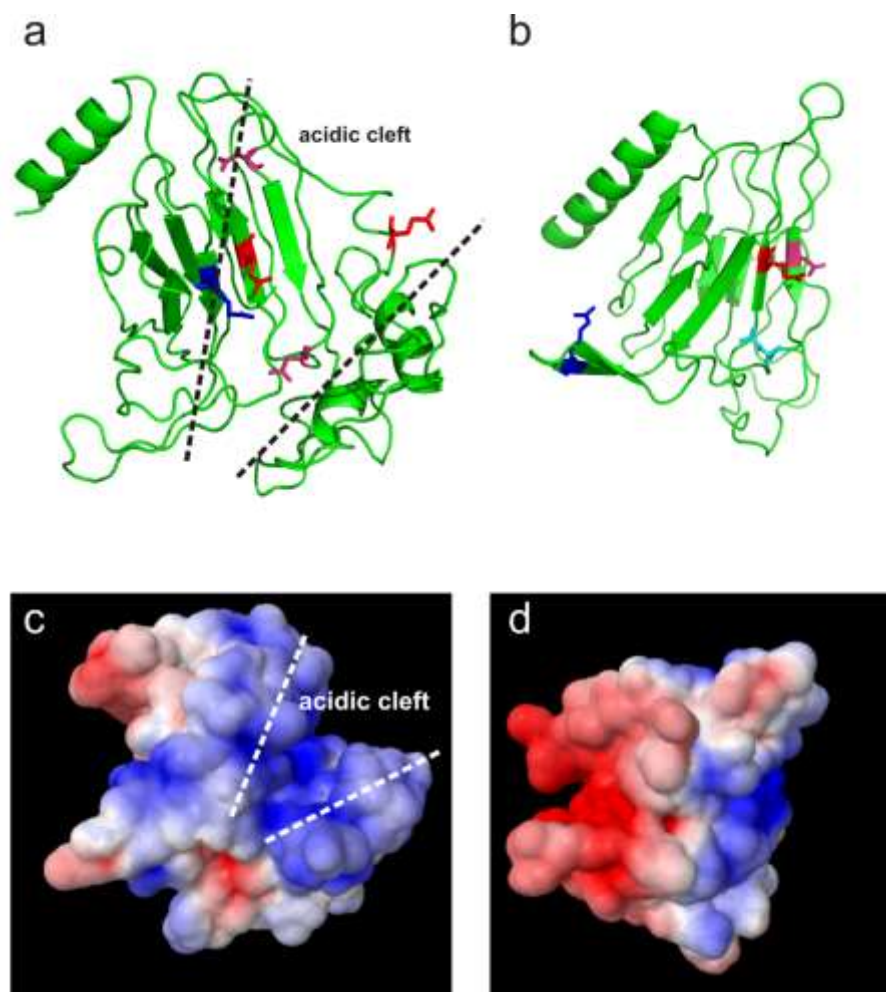


Figure 5: 3D molecular model of MpTLP1 and MpTLP12. Ribbon-sticks representations of MpTLP1 (a) and MpTLP12 (b) indicating amino acids positioned in REDDD motif sites depicted in figure 2. The putative residues in MpTLP1 are highlighted in dark blue (arginines; Arg), red (glutamic acids; Glu), and magenta (aspartic acids; Asp). Note the partial absence of some of these residues and the substitution of an Asp by a Gln (glutamine; light blue) in MpTLP12. Dashed black lines indicate acidic cleft limits. Electrostatic potential distribution on molecular surfaces are shown for MpTLP1 (c) and MpTLP12 (d). Dashed white lines indicate acidic cleft limits. Note the complete absence of acidic cleft in MpTLP12 (d).

DISCUSSION

The TLPs in *M. perniciosus*

Thaumatococcal-like proteins (TLPs) were at first studied in plants where they have been characterized as important proteins in defense against pathogens, most especially fungi (Kitajima and Sato 1999, van Loon et al. 2006). Nevertheless, in the last years many reports have been demonstrating the importance of TLPs in non-plant organisms such as

arthropods (Altincicek et al. 2013), nematodes (Kandasamy et al. 2012) and fungi (Sakamoto et al. 2006). We report here the identification of TLPs in the important cacao pathogen *Moniliophthora perniciosa* (MpTLPs), which causes the destructive Witches broom disease of cacao (WBD).

After thorough genome and transcriptome annotation, we identified a TLP gene family in *M. perniciosa* containing thirteen members. Six of them are organized in three gene clusters, each one containing two TLPs (Fig. 1). Phylogeny data showed that MpTLPs present in genome clusters (MpTLP1 and MpTLP2, MpTLP3 and MpTLP4, MpTLP11 and MpTLP12) are located in the same clades, indicating recent events of duplication. Additionally, a β -1,3 glucan synthase was detected upstream from MpTLP1 (Fig. 1). This data is interesting, since TLPs are known to be β -1,3 glucanases. The genomic proximity of genes involved in the same pathway is common in fungi, such as the case of genes of secondary metabolism (Dhingra et al. 2013).

M. perniciosa has the larger number (13) of genes encoding TLPs in comparison to other fungal species analyzed (Table 2), suggesting important roles for these proteins in WBD causative agent. *M. perniciosa* is considered a hemibiotrophic pathogen. The biotrophic stage of *M. perniciosa* is uncommon among hemibiotrophs, lasting for 2 to 3 months in the plant's infected tissue. During this stage, *M. perniciosa* inhabits a modified stem called 'green broom'. Through a not completely understood mechanism, the fungus shifts its lifestyle to necrotrophy and invades cells resulting in necrosis. After periods of wet and dry, the necrotrophic mycelia give rise to the basidiomatas that produce the infective basidiospores. In this scenario, the fungus has to cope with different antagonistic pressures. First, with the defense mechanisms of the plant and second, with competitors and opportunistic fungi those aim to colonize the 'brooms'. Therefore, TLPs in *M. perniciosa* may act in different stages of fungus development.

The expression of some MpTLPs is increased in the late stages of infection of *M. pernicioso* in stems and pods (Fig. 4). At this point of stems' infection, the necrotrophic mycelia have already invaded the host plant cells, destroying the infected tissues and causing the extensive degeneration symptoms of the dry broom phase. Remarkably, the inspection of the WBD transcriptome Atlas revealed the presence of sequences from other opportunistic fungi in the dry broom stems (Teixeira et al. unpublished results). Similarly, growth of other fungal species is often observed in dry brooms under field conditions. These fungi may benefit from the dead cacao tissue and colonize it in order to obtain feed resources. In this regard, *M. pernicioso* interacts with other fungal species during its necrotrophic stage of development. Martin *et al.* (2007) describes that glucanases secreted by filamentous fungi can degrade fungal cell walls. Therefore, TLPs, which were found to have β -glucanase activity, can be involved in negative fungal–fungal interactions. The increase of MpTLPs transcripts (i.e. *MpTLP1*, *MpTLP2*, *MpTLP4*, *MpTLP7*, *MpTLP13*) during this late stage of the disease may indicate the action of these proteins in the avoidance of fungi competitors.

L. edodes thaumatin-like protein TLG1 degrades lentinan, a β -1,3-glucan found in its own cell wall (Sakamoto *et al.* 2006). It was suggested that TLG1 is involved in cell wall degradation during senescence following harvest and spore diffusion (Sakamoto *et al.* 2006). Despite *TLG1* expression has not been observed in mycelia and growing fruiting bodies (Sakamoto et al. 2006), it is possible that other thaumatins can help in cell wall remodeling during the progression of fungus development. Noteworthy, in dry brooms stage occur the formation of thick-layers of mycelia that lead to possible hyphal fusions preceding basidiomata production (Pires et al. 2009). Therefore, we suggest that MpTLPs expressed in rotten tissues (*MpTLP1*, *MpTLP2*, *MpTLP4*, *MpTLP7*, *MpTLP13*) may participate in hyphal remodeling. Despite *MpTLP* expression in *in-vitro* grown basidiomata

are not as remarkable as *in planta* conditions, *MpTLP8* can also be involved in reproductive structures cell wall development.

Small TLPs in *M. pernicios*: MpTLP12 specific traits

We report the identification of small TLP genes (sTLPs) in *M. pernicios* (MpTLP11 and MpTLP12) and in other agaricales *Coniophora puteana*, *Schizophyllum commune*, *Stereum hirsutum* and *Dichomitus squalens* (Fig. 3). Small TLPs have been identified in the pucciniales *M. larici-populina* and *Puccinia graminis* f.sp. tritici and also in plants (Petre et al. 2011). However, *M. pernicios* sTLPs are not in the same clades as sTLPs, probably arising independently from them. In special, *MpTLP12* has interesting characteristics. First of all, its expression profile is very distinct from the other *MpTLPs*. While most of these genes are expressed during last stages of *in planta* interaction (dry brooms, rotten fruits), *MpTLP12* is specifically expressed in the advanced necrosis 2 stage (66 d.a.i.) and *in vitro* produced in basidiospores (Fig. 4b). Furthermore, MpTLP12 has a very distinct structure resulted from the almost complete deletion of the C-terminal region that make part of the bottom of the acidic cleft (Fig. 2, Fig. 5b, d), and have amino acid substitutions and deletions in the REDDD motif (Fig. 2, Fig. 5b, d). The only sTLP enzymatically characterized to date is TLX1, a sTLP from wheat that was shown to have xylanase inhibitor activity (Fierens et al. 2007; Vandermarliere et al. 2010). As in TLX1, TLP acidic cleft is absent in MpTLP12 (Fig. 5b. d), suggesting a distinct function in relation to other MpTLPs and 'classical' TLPs. We hypothesize that in the advance necrosis 2 stage MpTLP12 may anticipate the defense against aggressive competitors. Interestingly, the rust pathogen *P. graminis* has a large number of sTLPs (Petre et al. 2011). Perhaps, this class of TLPs can be involved in infective mechanisms of basidiomycete pathogens. On the other hand, in spores, MpTLP12 could modulate germination by acting in specific glucans located in spore cell wall.

Hypotheses for TLPs expansion in basidiomycetes

The data mining of fungal genomes indicate a prevalence of TLPs in basidiomycete in comparison to ascomycete (Table 1, Fig. 3). Therefore, TLP 'expansion' in basidiomycete indicates a functional diversification of these proteins, and their importance for mushroom forming species. Due to the larger number of TLPs in basidiomycetes in contrast to ascomycetes and the evidences of the participation of TLG1 in basidiomata fate (Sakamoto et al. 2006), it is likely that TLPs are indeed involved in mushroom development. On the other hand, as discussed above, TLPs can be involved in negative fungal–fungal interactions. The antagonism between fungi can be extreme especially in a competitive environment for food resources such as dense forests, where basidiomycetes are the main carbon decomposers. Therefore, the expansion of TLPs in basidiomycete can be related to the combat against contenders for the food search. Another hypothesis for TLP expansion in basidiomycetes is the action of these proteins in the degradation of plant cell wall. Callose is the main plant polysaccharide containing β -1,3-linkages, and is degraded by β -1,3-glucanases (Chen and Kim 2009). Despite there is no evidence that TLPs can degrade plant β -1,3-glucans, the large families of TLPs found in wood decomposers, and the lack of enzymatic analyses of TLPs of basidiomycetes, led us to hypothesize a putative action in plant cell wall.

ACKNOWLEDGEMENTS

The authors thank Mariana Ribeiro and Dr. Daniel Trindade for help in protein modeling, Dr. Celso Fiori for revision of the manuscript and Dr. Francisco Javier Medrano for hints and suggestions during the first stage of TLP project. SFF received a fellowship (2009/50810-3) from São Paulo Research Foundation (FAPESP). This work was supported by a grant (2009/50119-9) from São Paulo Research Foundation (FAPESP).

REFERENCES

Aime MC, Phillips-Mora W (2005). The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia*. 97: 1012-1022.

Altincicek B, Elashry A, Guz N, Grundler FM, Vilcinskas A, Dehne HW (2013) Next generation sequencing based transcriptome analysis of septic-injury responsive genes in the beetle *Tribolium castaneum*. *PLoS One*. 8(1): e52004.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98: 10037-10041.

Baker RED, Holliday P (1957) Witches' broom disease of cocoa (*Marasmius perniciosus* Stahel). *Phytopathological Papers* 2 : 42.

Barsottini MR, de Oliveira JF, Adamoski D, Teixeira PJ, do Prado PF, Tiezzi HO, Sforça ML, Cassago A, Portugal RV, de Oliveira PS, de M Zeri AC, Dias SM, Pereira GA, Ambrosio AL (2013) Functional diversification of cerato-platanins in *Moniliophthora perniciosus* as seen by differential expression and protein function specialization. *Mol Plant Microbe Interact*. 26:1281-1293.

Brandazza A, Angeli S, Tegoni M, Cambillau C, Pelosi P (2004) Plant stress proteins of the thaumatin-like family discovered in animals. *FEBS Lett* 572:3–7.

Broad Institute. [<http://www.broadinstitute.org/>].

Chen XY, Kim JY. (2009) Callose synthesis in higher plants (2009) *Plant Signal Behav*. 4:489-492.

de Oliveira BV, Teixeira GS, Reis O, Barau JG, Teixeira PJ, do Rio MC, Domingues RR, Meinhardt LW, Paes Leme AF, Rincones J, Pereira GA (2012). A potential role for an extracellular methanol oxidase secreted by *Moniliophthora perniciosus* in Witches' broom disease in cacao. *Fungal Genet. Biol.* 49: 922-932.

DeLano WL (2002) The PyMOL Molecular Graphics System. <http://www.pymol.org>.

Dhingra S, Lind AL, Lin HC, Tang Y, Rokas A, Calvo AM (2013). The Fumagillin gene cluster, an example of hundreds of genes under veA control in *Aspergillus fumigatus*. *PLoS One* ;8(10):e77147

Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* 35: 522-5.

Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 32.W665-W667.

Emanuelsson O, Brunak S, von HG, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953–971.

Evans HC (1978) Witches' broom disease of cocoa (*Crinipellis pernicios*) in Ecuador. *Ann Appl Biol* 89: 185–192.

Evans HC (1980) Pleomorphism in *Crinipellis pernicios*, causal agent of witches' broom disease of cocoa. *Trans Br Mycol Soc* 74: 515-523.

Evans HC, Bastos CN (1980) Basidiospore germination as a means of assessing resistance to *Crinipellis pernicios* (witches' broom disease) in cocoa cultivars. *Trans Br Mycol Soc* 74: 525-536.

Fierens E, Rombouts S, Gebruers K, Goesaert H, Brijs K, Beaugrand J, Volckaert G, Van Campenhout S, Proost P, Courtin CM, Delcour JA (2007) TLXI, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family. *Biochem J* 403:583-591.

Frias G, Purdy L, Schmidt R (1991) Infection biology of *Crinipellis pernicios* on vegetative flushes of cacao. *Plant Disease*. 75: 552-556.

Frias GA, Purdy LH, Schmidt RA (1995) An inoculation method for evaluating resistance of cacao to *Crinipellis pernicios*. *Plant Disease* 79: 787–791.

Grenier J, Potvin C, Asselin A (2000) Some fungi express b-1, 3- glucanases similar to thaumatin-like proteins. *Mycologia* 92:841–848.

Grenier J, Potvin C, Trudel J, Asselin A (1999) Some thaumatin-like proteins hydrolyse polymeric b-1,3-glucans. *Plant J* 19: 473-480.

Islam MA (2009) Ultrastructural studies of *Phellinus sulphurascens* infection of Douglas-fir roots and immunolocalization of host pathogenesis-related proteins. *Mycol Res* 113: 700–712.

Joint Genome Institute. [<http://www.jgi.doe.gov/>].

Kandasamy S, Khan W, Evans F, Critchley AT, Prithiviraj B (2012) Tasco®: a product of *Ascophyllum nodosum* enhances immune response of *Caenorhabditis elegans* against *Pseudomonas aeruginosa* infection. *Mar Drugs*. 1:84-105.

Kitajima S, Sato F (1999) Plant pathogenesis-related proteins: molecular mechanisms of gene expression and protein function. *J Biochem (Tokyo)* 125:1–8.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10:R25.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. (2007) ClustalW and ClustalX version 2. *Bioinformatics* 23: 2947-2948.

Liu JJ, Sturrock R, Ekramoddoullah AK (2010) The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant Cell Rep* 29(5):419-436.

Liu J-J, Zamani A, Ekramoddoullah AKM (2010) Expression profiling of a complex thaumatin-like protein family in western white pine. *Planta* 231:637–651.

Martin K, McDougall BM, McIlroy S, Chen J, Seviour RJ (2007) Biochemistry and molecular biology of exocellular fungal beta-(1,3)- and beta-(1,6)-glucanases. *FEMS Microbiol Rev* 31(2):168-192.

Meinhardt LW, Rincones J, Bailey BA, Aime MC, Griffith GW, Zhang D, Pereira GA (2008) *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe. *Mol Plant Pathol* 9:577-588.

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H et al (2008). A genome survey of *Moniliophthora perniciosa* gives new insights into witches' broom disease of cacao. *BMC Genomics* 9: 548.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNASeq. *Nat Methods* 5:621-628.

Mukherjee AK, Carp MJ, Zuchman R, Ziv T, Horwitz BA and Gepstein S (2010) Proteomics of the response of *Arabidopsis thaliana* to infection with *Alternaria brassicicola*. *J Proteomics* 73:709-720.

Narasimhan ML, Coca MA, Jin J, Yamauchi T, Ito Y, Kadowaki T, Kim KK, Pardo JM, Damsz B, Hasegawa PM, Yun DJ, Bressan RA (2005) Osmotin is a homolog of mammalian adiponectin and controls apoptosis in yeast through a homolog of mammalian adiponectin receptor. *Mol Cell* 17: 171-180.

Newman MA, Sundelin T, Nielsen JT, Erbs G (2012) MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Front Plant Sci* 16;4:139.

Nicholas, K.B., Nicholas H.B. Jr., Deerfield, D.W. II. 1997 GeneDoc: Analysis and Visualization of Genetic Variation, *EMBNEW.NEWS* 4:14

Penman D, Britton G, Hardwick K, Collin H, Isaac S (2000) Chitin as a measure of biomass of *Crinipellis perniciosa*, causal agent of witches' broom disease of *Theobroma cacao*. *Mycol Res* 104: 671-675.

Petre B, Major I, Rouhier N, Duplessis S (2011) Genome-wide analysis of eukaryote thaumatin like proteins (TLPs) with an emphasis on poplar. *BMC Plant Biology* 11:33.

Pires AB, Gramacho KP, Silva DC, Goes-Neto A, Silva MM, Muniz-Sobrinho JS, Porto RF, Villela-Dias C, Brendel M, Cascardo JC, et al (2009) Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes. *BMC Microbiol* 9: 158.

Pritsch C, Muehlbauer GJ, Bushnell WR, Somers DA, Vance CP (2000) Fungal development and induction of defense response genes during early infection of wheat spikes by *Fusarium graminearum*. Mol Plant Microbe Interact 13: 159-169.

Purdy LH, Schmidt RA (1996) Status of cacao Witches' Broom: biology, epidemiology, and management. Annu Rev Phytopathol. 34:573-594.

Rincones J, Scarpari LM, Carazzolle MF, Mondego JM, Formighieri EF, Barau JG, Costa GG, Carraro DM, Brentani HP, Vilas-Boas LA et al (2008) Differential gene expression between the biotrophic-like and saprotrophic mycelia of the witches' broom pathogen *Moniliophthora perniciosa*. Mol Plant Microbe Interact 21: 891–908.

Rio MC, de Oliveira BV, de Tomazella DP, Silva JA, Pereira GA (2008) Production of calcium oxalate crystals by the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom disease of Cacao. Curr Microbiol. 2008 Apr;56(4):363-370.

Roberts WK, Selitrennikoff CP (1990) Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. J Gen Microbiol 136:1771–1778.

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725–738.

Sakamoto Y, Watanabe H, Nagai M, Nakade K, Takahashi M, Sato T (2006) *Lentinula edodes* tlg1 encodes a thaumatin-like protein that is involved in Lentinan degradation and fruiting body senescence. Plant Physiol 141:793–801.

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GA (2005) Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa*. J Exp Bot 56: 865-877.

Schimoler-O'Rourke R, Richardson M and Selitrennikoff CP (2001) Zeamatin inhibits trypsin and α -amylase activities. Appl Environ Microbiol 67:2365-2366.

Shatters RG Jr, Boykin LM, Lapointe SL, Hunter WB, Weathersbee AA 3rd (2006) Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa. J Mol Evol 63:12–29.

Silva SDVM, Matsuoka K (1999) Histologia da Interação *Crinipellis perniciosa* em Cacaueiros Suscetível e Resistente à Vassoura de Bruxa. Fitopatologia Brasileira 24: 54-59.

Stanke M, Steinkamp R, Waack S, Morgenstern B (2004) AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res 32: W309–W312.

Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M, Fritig B (1993) Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie. 75(8):687-706.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28, 2731-2739.

Teixeira PJ, Thomazella DP, Vidal RO, do Prado PF, Reis O, et al (2012) The fungal pathogen *Moniliophthora perniciosa* has genes similar to plant PR-1 that are highly expressed during its interaction with cacao. *PLoS ONE* 7(9): e45929.

Thomazella DP, Teixeira PJ, Oliveira HC, Saviani EE, Rincones J, Toni IM, Reis O, Garcia O, Meinhardt LW, Salgado I, et al (2012) The hemibiotrophic cacao pathogen *Moniliophthora perniciosa* depends on a mitochondrial alternative oxidase for biotrophic development. *New Phytol* 194: 1025-1034.

Vandermarliere E, Lammens W, Schoepe J, Rombouts S, Fierens E, Gebruers K, Volckaert G, Rabijns A, Delcour JA, Strelkov SV (2010) Courtin CM. Crystal structure of the noncompetitive xylanase inhibitor TLXI, member of the small thaumatin-like protein family. *Proteins*. 1;78(10):2391-2394.

van der Wel H, Loewe K (1972) Isolation and characterization of thaumatin I and II, the sweet-tasting proteins from *Thaumatococcus daniellii* Benth. *Eur J Biochem* 31:221–225.

Van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44: 135-162.

Velazhahan R, Datta SK, Muthukrishnan S (1999) The PR-5 family: thaumatin-like proteins. In *Pathogenesis-Related Proteins in Plants*. CRC Press, Boca Raton, FL, 107–129.

Vigers AJ, Roberts WK, Selitrennikoff CP (1991) A new family of plant antifungal proteins. *Mol Plant Microbe Interact* 4:315-323.

Wang X, Tang C, Deng L, Cai G, Liu X, Liu B, Han Q, Buchenauer H, Wei G, Han D, Huang L, Kang Z (2010) Characterization of a pathogenesis-related thaumatin-like protein gene TaPR5 from wheat induced by stripe rust fungus. *Physiol Plantarum* 139: 27–38.

Zaparoli G, Barsottini MR, de Oliveira JF, Dyszy F, Teixeira PJ, Barau JG, Garcia O, Costa-Filho AJ, Ambrosio AL, Pereira GA, et al (2011) The crystal structure of necrosis- and ethylene-inducing protein 2 from the causal agent of cacao's Witches' Broom disease reveals key elements for its activity. *Biochemistry* 50: 9901-9910.

**Additional file 1: Codes for non-fungal species depicted in phylogeny tree (Fig.3).
NCBI accession numbers are on parenthesis.**

Arthropodes

Denca – *Dendroides canadensis* (AAY45729.1)

Trica – *Tribolium castaneum* (Trica1: XP_975175.1; Trica2: XP_975166.1)

Diaab - *Diaprepes abbreviatus* (Diaab1: AAW56443.1; Diaab2: AAW56442.1)

Lyste - *Lysiphlebus testaceipes* (AAW56445.1)

Toxci - *Toxoptera citricida* (AAW56444.1)

Aleov - *Aleuroglyphus ovatus* (ABU97465.1)

Suime - *Suidasia medanensis* (ABU97476.1)

Schigr - *Schistocerca gregária* (Schigr1: AAR97603.1; Schigr2: AAS83110.1)

Copfo - *Coptotermes formosanus* (AFZ78845.1)

Nematodes

Caeel - *Caenorhabditis elegans* (Caeel1: NP_502360.1; Caeel2: NP_502362.1; Caeel3: NP_502361.1; Caeel4: NP_507263.1; Caeel5: NP_500748.2; Caeel6: NP_500747.1)

Plants

Arath – *Arabidopsis thaliana* (Arath1: NP_177893.1; Arath2: NP_195579.2)

Nicta – *Nicotiana tabacum* (Nicta1: CAA06927.1; Nicta2 /osmotin: CAA46623.1)

Theca – *Theobroma cacao* (AAV34889.2)

Zeama – *Zea mays* (zeamatin: NP_001105356.1)

Orysa – *Oryza sativa* (Orysa1: CAB53479.1; NP_001058421.1)

Brara – *Brassica rapa* (AAB95118.1)

Maldo – *Malus domestica* (AAC36740.1)

3.5- Materiais e Métodos

3.5.1- Clonagem, expressão e purificação de *MpTLPs*

O projeto se iniciou com o objetivo principal de purificar MpTLPs seguida de sua caracterização molecular e testes de atividades enzimáticas. Vários métodos de obtenção de proteína foram testados, porém, nenhum resultou em proteína em seu estado solúvel. Ao longo do trabalho, testamos variadas técnicas de obtenção de proteína, começando por com o uso do plasmídeo pET28a e pSW202 que envia a proteína para o periplasma *E. coli* BL21DE3 e BL21C41, até testes de expressão em *Pichia Pastoris*. Esse último teve como limitante o fato de que as leveduras transformadas não se mantinham viáveis para expressão, e morriam durante o crescimento. Esses testes foram realizados durante o trabalho de mestrado com a proteína *MpTLP1*.

Com o surgimento de novas *MpTLPs*, decidimos continuar o trabalho explorando outras proteínas e outras formas de expressão heteróloga. Esse trabalho foi incluído em uma parceria do LGE com o LNBio, dentro do projeto “*SMOLBNET 2.0: Estudos estruturais de proteínas chave para as doenças fúngicas do cacau Vassoura de bruxa e Monilíase – Desenvolvimento de estratégias de controle e entendimento de modelos de patogenicidade*”. Essa parceria foi muito importante, pois tivemos acesso a toda estrutura de um laboratório de expressão e purificação de proteína, somados a *expertise* dos coordenadores do Projeto: André Luis Ambrósio e Sandra Martha Gomes Dias. Além disso, pudemos contar com a experiência e ensinamentos de Daniel Maragno Trindade para a técnica de expressão de proteínas utilizando o sistema de Baculovírus em Células de Insetos.

3.5.2- Clonagem dos gene *MpTLPs*:

3.5.2.1- Organismos hospedeiros

As células competentes de *E. coli* utilizadas nas subclonagens eram da cepa DH10B e GM2163. As cepas de *E. coli* utilizadas para expressão da proteína recombinante foram BL21-C41, ORIGAMI DE 03 e Shuffle (New England Biolabs).

Para produção de vírus e expressão em células de insetos foram utilizadas *E. coli* Dh10BAC (Invitrogen) para recombinação da construção contida no plasmídeo de entrada pFastBac Dual para o Bacmídeo utilizado para a infecção de células de inseto; célula de inseto SF9 (Gibco) para a produção de vírus e HightFive (Invitrogen) para ensaios de produção de proteínas.

3.5.2.2- Plasmídeos

Para subclonagem dos fragmentos amplificados por PCR foi utilizado o plasmídeo pGEM-T Easy (Promega, EUA). Para a superexpressão de proteínas recombinantes em bactéria foram usados os plasmídeos pET28a e pETSUMO (Novagen, EUA), e pFastBac HTa (Invitrogen, EUA) e pFastBac Dual (Invitrogen, EUA). Esse último foi submetido a modificações para a inserção de um gene repórter EGFP, resultando no plasmídeo batizado de pFastBacEGFP, e em um segundo momento uma inserção de uma sequência de cauda de histidina, tornando-se pFastBacEGFP C-HT.

3.5.2.3- Amplificação e subclonagem dos genes *MpTLPs*.

Utilizando as sequências codificantes geradas a partir do software Augustus gene finder, pares de oligonucleotídeos foram desenhado para amplificar os genes a partir de cDNA do fungo, adicionando as sequências reconhecidas pelos sítios de restrição *EcoRI* nos oligonucleotídeos direto e *HindIII* nos reversos (tabela 1). Os genes *MpTLPs* foram amplificados, os fragmentos resultantes foram clonados em pGEM-T Easy e transformado em *E. coli DH10b* através da técnica de eletroporação. Os transformantes foram

selecionados em meio de cultura seletivo, contendo antibiótico (ampicilina 100µg/ml) e X-gal (técnica de α -complementação). Esses clones foram crescidos em meio LB contendo ampicilina (100µg/ml) e em seguida submetidos à extração de DNA plasmidial por lise alcalina (Sambrook et al., 1996).

3.5.2.4- Digestão de *MpTLPs* e clonagem em vetores de expressão

Os plasmídeos obtidos foram digeridos com as enzimas de restrições *EcoRI* e *HindIII*. As digestões foram aplicadas em gel de agarose 0,8% e os fragmentos referentes às sequências dos genes *MpTLPs* foram purificadas do gel através do Kit PureLink (Invitrogen, EUA), e utilizadas para ligação aos plasmídeos de expressão de proteína. Após a ligação, os plasmídeos foram novamente transformados em *E.coli* DH10B e os transformantes foram selecionados através de plaqueamento em meio de cultura contendo antibiótico (canamicina 25µg/ml para construções em pET28a e pETSUMO; ampicilina 100µg/ml para pFast Bac HTa,) e posteriormente através da técnica de PCR. Confirmados os clones, uma nova purificação de DNA plasmidial por lise alcalina foi realizada, e assim obtivemos o gene ligado nos plasmídeos de expressão.

3.5.2.5- Transformação

A construção pET28a-*MpTLP1* foi inserida em *E.coli* BL21-C41 através de eletroporação. As construções pETSUMO- *MpTLP* (1, 2, 4 e 7) foram inseridas em *E.coli* *Origami* e *E.coli* *Shuffle* e pela técnica de transformação por choque térmico. No primeiro caso, os transformantes foram plaqueados em meio de cultura contendo canamicina (25µg/ml) e confirmados através de PCR. As construções pETSUMO- *MpTLP* (1, 2, 4 e 7) em *Origami* foram plaqueadas em meio de cultura contendo canamicina (25 µg/ml) que corresponde a resistência do plasmídeo, estreptomicina 25 µg/ml e tetraciclina 5 µg/ml, que correspondem a resistência da célula. Os transformantes de pETSUMO- *MpTLP* em *E.coli* *Shuffle* foram plaqueados em meio de cultura contendo canamicina (25µg/ml)

ampicilina (100µ/ml) e cloranfenicol (33µg/ml). Para evitar erro os transformantes foram confirmados através da técnica de PCR.

3.5.3- Teste de indução da proteína recombinante em pequena escala

Para a construção pET28a- *MpTLP1*, como se trata da continuidade do trabalho de mestrado da aluna em que esses testes já foram feitos, uma nova tentativa de indução em pequena escala foi feita. Utilizamos um volume de 50ml de meio LB+canamicina para inocular 0,8ml de um pré inóculo crescido *overnight*. A cultura foi incubada em temperatura de 37°C a 250 RPM até atingir OD₆₀₀ = 0,7. Foi retirado 1 ml dessa cultura (amostra sem indução) e, adicionado IPTG (concentração final 0,3 mM) e Etanol 2%. Foi incubado overnight a temperatura de 22°C e agitação de 150rpm. As células colhidas da indução foram precipitadas por centrifugação, seguida de lise celular. As células foram sonicadas sete vezes por quinze segundos cada vez, e centrifugadas a 13000 RPM durante 10 minutos a 4°C. O sobrenadante equivalente à fração solúvel das proteínas totais foi coletado, e o sedimento foi ressuspendido em 5 ml de tampão de coluna. Essa amostra corresponde à fração insolúvel de proteínas totais e analisado em gel SDS-PAGE.

Para as construções pETSUMO- *MpTLP* (1, 2, 4 e 7), uma colônia foi inoculada em 3 ml de meio LB contendo antibióticos, e incubados a 37°C durante cerca de oito horas. Nesse ponto foi coletado 300µl (amostra sem indução) e adicionado 0,2mM de IPTG. Os inóculos foram então incubados a 18°C, 150 RPM *overnight*. As células foram coletadas por centrifugação e ressuspendidas em tampão contendo 50mM Tris- HCL pH 7,5; 100mM NaCl, lisozima, DNase e PMSF. Em seguida as células foram sonicadas e centrifugadas conforme descrito anteriormente, e as frações geradas foram analisadas em gel SDS- PAGE.

3.5.4- Clonagem de *MpTLP1* em vetores de expressão para células de inseto.

Para clonagem em pFastBacEGFP, a sequência de *MpTLP1* que foi inserida em pFast Bac HTa, foi digerida com as enzimas *RSRII* e *HindIII*, e ligada em pFastBacEGFP igualmente digerido. Para a clonagem de *MpTLP1* contendo peptídeo sinal, o produto de PCR foi digerido com as enzimas de restrição *RsrI* e *HindIII*, e posteriormente ligado ao plasmídeo pFastBacEGFP C-HT igualmente digerido.

3.5.4.1-Construção de pFastBacEGFP a partir de pFastBac Dual.

O plasmídeo pFastBac Dual possui duas regiões de clonagem. No sítio regulado pelo promotor do gene do gene p10 (Pp10), foi inserida a sequência codificadora da versão mais brilhante (enhanced) da proteína fluorescente verde (GFP), EGFP a partir de digestão do plasmídeo pEGP-C3 (Clontech) obtendo assim o plasmídeo pFastBacEGFP.

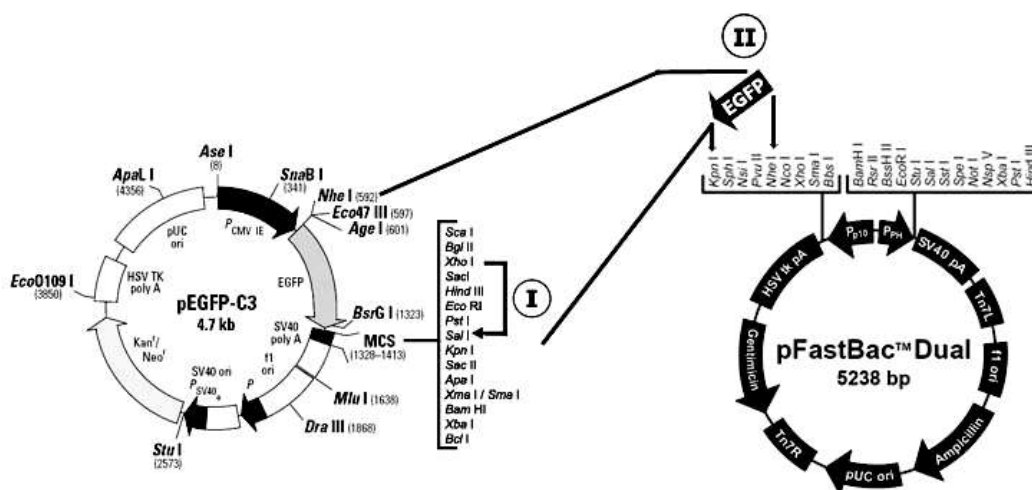


Figura 6. Mapa da criação do vetor pFastBacEGFP. Mapa mostrando a subclonagem da sequência codificadora da versão mais brilhante da proteína verde fluorescente (EGFP), depois de retirada parte do poli-linker do vetor pEGFP-C3 (I), no poli-linker do vetor pFastBac™Dual sob controle do promotor do gene p10 (Pp10) (II),

resistência a tetraciclina. Após o bacmídeo receber a construção, ele será utilizado na transfecção em células de inseto. Para evitar resultados falso-positivos, transformantes apresentando coloração branca foram transferidos para novas placas contendo a mesma composição e reincubados nas mesmas condições.

As cinco colônias positivas foram inoculadas em LB contendo os antibióticos citados acima e as células foram utilizadas para extração de bacmídeo através do mesmo protocolo utilizado para extração de DNA plasmidial por lise alcalina (Sambrook et al., 1996). A recombinação foi confirmada através de PCR de acordo com recomendações do fabricante, utilizados primers M13 Forward e M13 Reverse.

3.5.4.4- Transfecção de células de inseto.

Em uma placa de 6 poços foram semeadas 9×10^5 células de inseto SF9, em meio de crescimento TC100 (Sigma Aldrich) contendo 50u/ml de penicilina e 50u/ml de streptomicina. As células foram incubadas a 27°C por uma hora para a adesão das células placa. Enquanto as células aderiram, foi preparado em um tubo 100µl de meio de crescimento sem BSA e 1µg de Bacmídeo. Paralelamente, 6 ul de Cellfectin® (Invitrogen) foram diluídos em 100µl do mesmo meio. Foi misturado 105µl de cada mistura e incubado a temperatura ambiente por 30 minutos. Durante a incubação do complexo DNA: lipídeo, as células de inseto semeadas em placas foram lavadas duas vezes com meio de cultura sem suplementação por BSA. Logo após, o complexo DNA: lipídeo foi adicionado às células e incubado a 27°C por cinco horas. Após esse período, o complexo é retirado cuidadosamente das células e são adicionados 2 ml de meio de cultura TC100 completo. As células foram incubadas a 27°C durante 72 horas em incubadora úmida. O estoque viral P1 foi purificado através da sedimentação das células por centrifugação a 500g durante 5 minutos e o sobrenadante foi submetido à filtração com membrana de 22µm. Esse estoque foi amplificado através da produção de vírus estoque P2.

3.5.45- Titulação do vírus de MpTLP1

A titulação de vírus P1 foi feita através da metodologia de Plaque Assay viral, conforme protocolo descrito no manual “Bac-to-Bac® Baculovirus Expression System”. Basicamente, se trata de um protocolo onde placas contendo o mesmo número de células são submetidas à infecção por diluição seriada de vírus e após um período de incubação é possível contar o número de colônias infectadas pelo vírus, chegando assim ao título do estoque viral. Porém, não foi utilizado overlay de Neutral-Red agarose, que foi substituída por overlay de agarose comum e o resultado foi determinado através da visualização das placas infectadas em microscopia de fluorescência. O Título foi determinado através da contagem de colônias fluorescente em verde devido à presença de EGFP na construção.

O Título final é dado pela equação:

Título (unidades formadoras de colônia/ml) = número de células fluorescentes x (1/ ml de inóculo/placa)

Após a produção do vírus estoque P2, a metodologia de *plaque assay* foi repetida para titular o novo estoque.

3.5.4.6- Produção de estoque de vírus P2 de MpTLP1

A partir da determinação do título do vírus P1 (3,6 pfu/ ml) foi montado um experimento para amplificação do vírus estoque. A quantia de 1×10^5 células Sf9/ml foram semeadas em quatro garrafas de cultura de células de 175 ml, em 25 ml de meio TC100 completo. Após a adesão das células, foi adicionado 486µl de vírus P1 em cada placa, e em seguida as placas foram incubadas a 27°C por 48 horas. Após a incubação, as culturas foram coletadas e centrifugadas durante 5 minutos a 500g. O sobrenadante foi coletado, filtrado e estocado a 4°C protegido da luz. O vírus estoque P2 também foi submetido à titulação pelo método de *plaque assay* descrito anteriormente.

3.5.4.7- Ensaio de indução em pequena escala. Teste tempo de incubação X MOI

Um ensaio de expressão com variação de tempo de incubação e concentração de vírus (MOI = números de partículas virais por célula) foi realizado a fim de determinar a melhor condição para os ensaios de produção de proteína. Em quatro erlenmeyers de 125 ml foram adicionados 50 ml de meio de cultura *Express Five* (Gibco) com a concentração de 1×10^6 células *High Five* /ml. Quantidades crescentes de vírus P2 de MpTLP2 (título = 3×10^7) para 4 MOI crescentes (0,1; 0,5; 1 e 5) foram adicionadas aos erlenmeyers que em seguida foram incubados a 26°C a 140 RPM. Amostras de 1 ml foram sendo coletadas a cada 24 horas até 72 horas de incubação. As células coletadas foram centrifugadas durante 5 minutos a 500g e em seguida foram lisadas em 40µl de PBS contendo 300 mM de NaCl e 1% do detergente NP40 (Sigma Aldrich), com incubação a temperatura ambiente por 30 minutos e em seguida incubação no gelo durante 10 minutos. Logo após, as células foram submetidas à centrifugação a 20000g durante 10 minutos. O sobrenadante foi separado, e amostras do pellet e sobrenadante das amostras foram preparadas e submetidas a gel SDS-PAGE. Geis das amostras das frações solúveis e insolúveis foram feitos em duplicata para verificação da presença da proteína pela técnica de Western-Blot. O Anticorpo utilizado para a detecção da proteína heteróloga foi o anticorpo Anti-his tag (anti-*mouse*).

3.5.5- Ensaio de produção de proteína em larga escala

Após a determinação de tempo de incubação e MOI ideais para a produção de proteína o ensaio de produção de MpTLP1, um ensaio em larga escala foi determinado. Uma quantidade total de 1 litro de meio de cultura *Express Five* contendo 1×10^9 células *High Five* foi dividido em 2 erlenmeyer de 3 litros. Foi adicionado 13,2 ml do vírus estoque P2 de MpTLP1 a cada erlenmeyer (MOI= 1). Os frascos foram incubados a 26°C a 140

RPM durante 48 horas. Após esse período, as células foram sedimentadas por centrifugação a 500g durante 20 minutos. As células foram lisadas em 30 ml de PBS contendo 1 ml de PMSF (Fenilmetilsulfonilflúor, inibidor de protease), 1,25 mg de DNase, 300 mM de NaCl e 1% do detergente NP40 com incubação a temperatura ambiente por 30 minutos e em seguida incubação no gelo durante 10 minutos. Logo após, as células foram submetidas à centrifugação a 20000g durante 10 minutos. O sobrenadante foi submetido à purificação em coluna de troca afinidade de cobalto.

3.5.6- Purificação de MpTLP1

Dois ml de resina de cobalto Talam (Clontech) foram empacotados em coluna de 10 ml. A resina foi lavada com três volumes de PBS, e em seguida o lisado celular foi incubado a resina a 4°C durante *over night*. Após incubação, a mistura lisado mais resina foi recolocado em coluna e o material que não foi ligado à coluna foi coletado. Em seguida a coluna foi lavada com 10 ml de PBS + 10 mM de imidazol, 10ml de PBS+ 100 mM de imidazol e 10ml de PBS + 500 mM de imidazol. As eluições foram coletadas e observadas em Gel SDS-PAGE, que foi feito em duplicata para que o conteúdo um deles fosse transferido para membrana de nitrocelulose e realizado a detecção de proteína através da técnica de *Western-Blot*. O anticorpo utilizado para a detecção da proteína heteróloga foi o anticorpo Anti-his tag (anti-*mouse*).

3.6- Resultados e Discussão

3.6.1- Amplificação e clonagem de MpTLPs

Os pares de *primers* para clonagem foram testados em gradiente de temperatura para amplificação a partir de cDNA do fungo, e todos os pares de *primers* geraram um amplicom entre 500 a 750 pb.

Determinada a melhor temperatura de amplificação, partimos para amplificação de *MpTLPs* a partir do cDNA, obtido através de RNA extraído da fase biotrófica do fungo

utilizando a enzima de alta fidelidade taq Platinum® (Invitrogen, USA). Apenas *MpTLP3* não gerou fragmento no PCR. A clonagem de *MpTLP1* em pETSUMO foi feita a partir da digestão de uma construção já utilizada pela aluna durante seu trabalho de mestrado, onde *MpTLP1* foi clonada no vetor de expressão *pSW202*, utilizando os mesmos sítios de restrições utilizados no presente trabalho (figura 8).

Tabela 1: Oligonucleotídeos utilizados na amplificação e clonagem de *MpTLPs*. Sequência grifada nos oligos direto: sítio de restrição *EcoRI*; Sequência grifada nos oligos reversos: sítio de restrição *HindIII*. As sequências marcadas com *correspondem a oligonucleotídeos utilizados para a obtenção de *MpTLP1* e *MpTLP4* juntamente com sua sequência sinalizadora de secreção para nova tentativa de expressão em células de inserto.

	Direto (5'-3')	Reverso (5'-3')
<i>MpTLP1</i>	GAATTCGCTCGTAAC TTCACGGTCT *ATCCC GGTCCGATGAACTTTTCGTCGCAAGTG	AAGCTTTCAGGGACAGAATGTCAAAGTAT *TCGACAAGCTTGGGACAGAATGTCAAAGTATAATC
<i>MpTLP2</i>	GAATTCCGTACATTTACCGTATACAACGCA	AAGCTTTCAAGGGCAGAATGTCAGAGTAT
<i>MpTLP3</i>	GAATTCCGTACTTTCACTGTGAAGAACAAC	AAGCTTTCAGGACAGAAAGTAAGCGTATAAT
<i>MpTLP4</i>	GAATTCCGTACTTTCACTGTAAAGAACAAC TG *GTCGCGGATCCATGCAATTATTCTCTATCACTCT T	AAGCTTTCAGGACAGAAAGTAAGCGTGTA *CGACAAGCTTGCAATGAAGAAGGGAACCACTGGG
<i>MpTLP5</i>	GAATTCCGCACGTTTACAGTCAGAAAC	AAGCTTTTAAGGGCAGAAGGTCAACGT
<i>MpTLP6</i>	GAATTCCGTACGTTACAGTCAGGAACA	AAGCTTTTAAGGGCAGAAGGTAAGAGTGTAGTC
<i>MpTLP7</i>	GAATTCCGTACGTTACAGTAAGGAACAAC TG	AAGCTTTCAGGGCAGAAGGTAAGAGTG
<i>MpTLP8</i>	GAATTCGCTCTGTAACAGTCAAGAATAACTG	AAGCTTTCATGGACAGAACGTGACC
<i>MpTLP9</i>	GAATTCCGCACGTTTACAGTCAGGAAC	AAGCTTTTAGGGGCAGAAGGTCAAGG

Os fragmentos de PCR gerados nessa etapa passaram por um processo de adelinção de suas extremidades para que fosse possível sua subclonagem em pGEM-T Easy (Promega- USA). Para isso, foi adicionado ao produto de PCR 0,2mM α ATP, 5 unidades de Taq DNA polimerase, Tampão da enzima 1X e 0,3mM de MgCl, para um volume final de 10ul de reação, que é mantida a 70°C por 30 minutos. Após a ligação, clones positivos foram confirmados por PCR e esses clones foram utilizados para inoculo,

extração de DNA plasmidial e esse plamídeo foi digerido com as enzimas *EcoRI* e *HindIII*, (figuras 8 e 9) para posterior ligação em vetor de expressão pET SUMO e pFASTBAC HTa, que foram submetidos ao mesmo tratamento enzimático para sua linearização. Para pFASTBAC HTa técnica resultou em clones positivos de *MpTLP1*, *MpTLP4* e *MpTLP9*. Para pETSUMO, os clones positivos foram *MpTLP1*, *MpTLP2*, *MpTLP4* e *MpTLP7*.

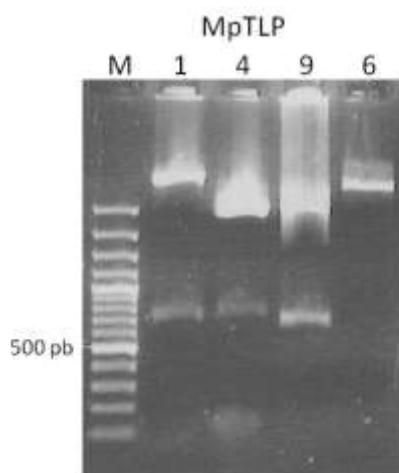


Figura 8: Digestão de pSW202-MpTLP1 e pGEM- MpTLP 4, 6 e 9. Os fragmentos gerados nesse experimento foram utilizados para ligação em vetor de expressão de proteína pETSUMO.

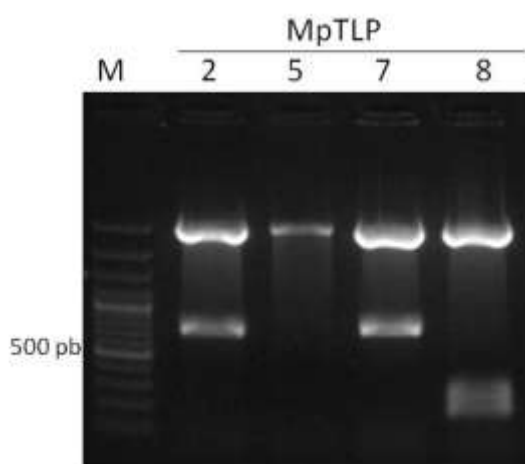


Figura 9: Digestão de pGEM- MpTLPs (2, 5, 7 e 8) com enzimas de restrição *EcoRI* e *HindIII*. Os fragmentos gerados nesses experimentos foram utilizados para ligação em vetor de expressão pETSUMO.

As construções positivas em pETSUMO foram purificadas e utilizadas para transformação em célula de expressão *E. coli Origami B* (DE3), e seguimos com experimento de expressão de proteína em pequena escala. Foi dada continuidade as clonagens de *MpTLP* 3, 5, 6, 8 e 9.

3.6.2- Ensaios de expressão de proteínas MpTLPs

Dando continuidade as tentativas de expressão da proteína MpTLP1, uma nova tentativa foi conduzida com as construções utilizada no trabalho de mestrado, pET28a-MpTLP1 em *E.coli* BL21 C41 foi submetida a um ensaio de expressão utilizando, além de IPTG, Etanol, com o intuito de facilitar a formação das pontes dissulfeto da proteína, porém, o experimento não resultou em proteína expressa na fase solúvel, apenas uma grande quantidade de proteína na fração insolúvel.

Um teste de expressão em pequena escala de pETSUMO-MpTLP (1, 2,4 e 7) foi feito e o resultado analisado em gel SDS- PAGE (fig 10). É possível observar uma banda diferencialmente expressa do tamanho esperado de aproximadamente 40KDa, já que MpTLPs possuem aproximadamente 25KDa, somado com a cauda sumo de aproximadamente 12KDa e cauda de histidina, de 1kDa, resultando em uma proteína de aproximadamente 40KDa. Essa banda foi observada nas frações solúveis das amostras induzidas. Experimentos para confirmação da presença das MpTLPs foram conduzidos através de indução em larga escala, seguido de purificação da proteína utilizando coluna de afinidade His-tag Ni_NTA (GE Healthcare, EUA). *Western- blot* com utilização de anticorpo Antis- His tag foi realizado para a confirmação da presença das proteínas.

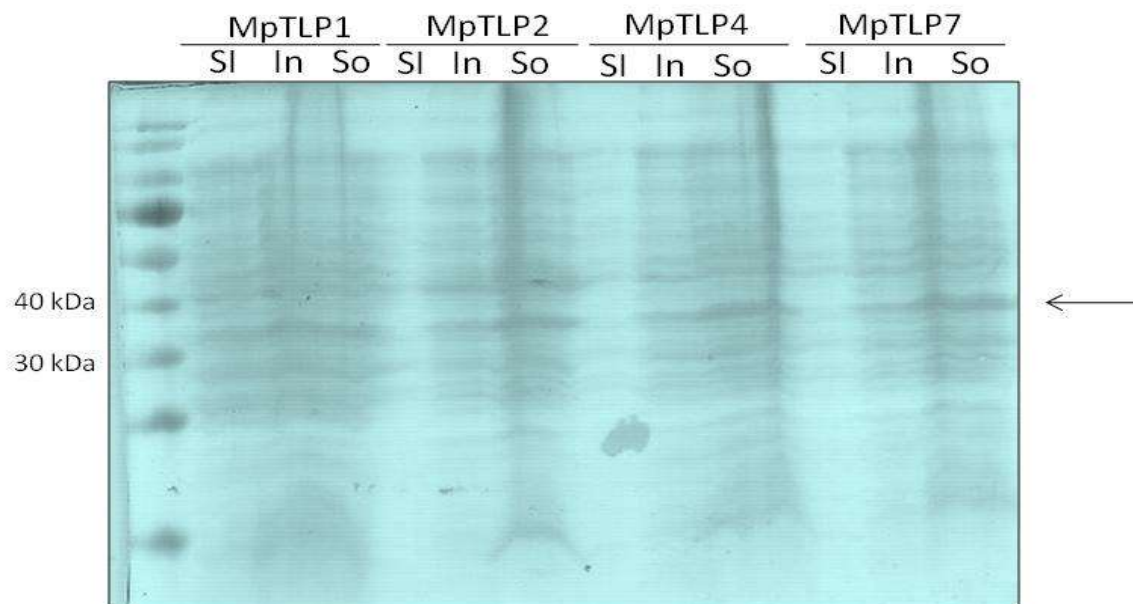


Figura 10: Experimento de indução em pequena escala de MpTLPs . SDS-PAGE contendo proteína total de *E. coli* Origami DE 3 transformadas com pETSUMO- MpTLP1, pETSUMO- MpTLP2, pETSUMO- MpTLP4 e pETSUMO- MpTLP7. A expressão de proteína foi induzida com 0,2 mM de IPTG e o conteúdo celular foi coletado após 16 horas de indução. Seta indica banda aparentemente diferencialmente expressa. SI: sem indução, In: fração insolúvel da indução, So: fração solúvel da indução.

3.6.3- Ensaio de indução em pequena escala de MpTLP1 e MpTLP4 em Shuffle

Diante dos resultados obtidos de expressão de MpTLPs em *E. coli*, os quais resultaram em produção de proteínas em seu estado insolúvel, uma nova cepa de *E. coli* foi testada. A *E. Coli Shuffle*[®] foi escolhida por se tratar de uma bactéria que foi criada para que as formações de pontes dissulfeto sejam formadas em seu citoplasma. Para que isso seja possível, essa bactéria tem deleções de genes tioredoxina redutase e glutathione redutase, que são responsáveis por manterem o ambiente citoplasmático em condições redutoras, para promover a formação de pontes dissulfeto. Essa deleção permite que essas formações ocorram. Além disso, ela possui uma versão de pontes dissulfeto isomerase, para corrigir oxidações erradas de maneira a permitir *folds* corretos.

Apesar de todas essas características favoráveis a produção de MpTLPs, que são proteínas ricas em pontes dissulfeto, ensaios de superexpressão de proteínas em pequena escala demonstraram que o sistema de expressão não foi adequado para superexpressão dessa família de proteína, produzindo novamente proteína em seu estado insolúvel. No ensaio de indução em pequena escala, utilizando-se de inóculos de três colônias de *Shuffle@* contendo a construção pETSUMO- MpTLP1, foi possível identificar que a proteína está sendo expressa (figura 11).

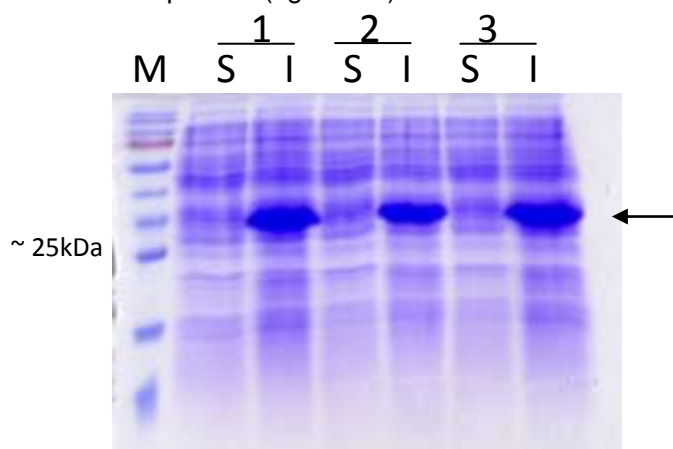


Figura 11: Expressão em pequena escala de MpTLP1- *Shuffle@*. 3 colônias de *Shuffle@* (1, 2 e 3) testadas. SI= não induzido; I = e horas de indução por IPTG.

Porém, quando submetemos as células à lise podemos ver que a proteína está presente predominantemente nas frações insolúveis do ensaio (figura 12).

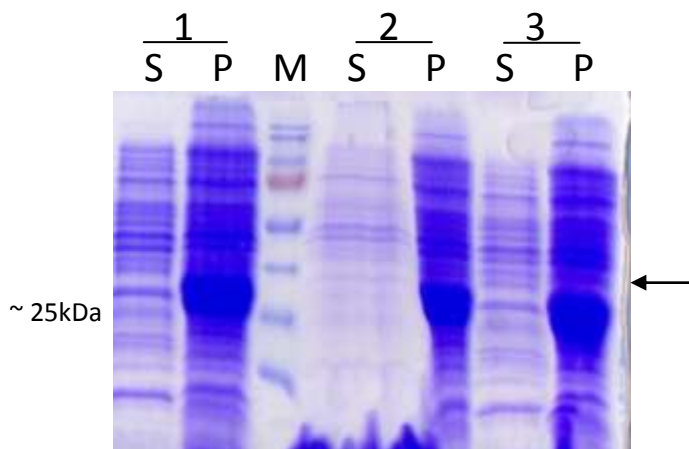


Figura 12: Teste de solubilidade de MpTLP1 em *Shuffle@*. Tres colonias (1, 2 e 3), após serem submetidas a lise celular. P= pellet (fração insolúvel); S = sobrenadante (fração solúvel).

Para que não houvesse dúvidas de que o pH do tampão de lise pudesse ser o causador da insolubilidade da proteína, o teste foi repetido, substituindo o tampão de lise por tampões Tris pH 8.0, HEPES pH 7.0 e MES pH 6.0, contendo as mesmas concentrações de NaCl e lisozima. Todos os experimentos resultaram no mesmo perfil de expressão visualizado na figura acima.

Diante dos resultados não satisfatórios de expressão de proteínas MpTLPs solúveis em sistema heterólogo utilizando *E.coli*, decidimos por mudar o sistema de expressão para o sistema de baculovírus em células de inseto. As células de insetos, por se tratar de um organismo mais complexo quando comparado a bactérias, e possui mecanismos pós traducionais próximo a fungos. Também foi verificada a produção de proteínas ricas em pontes dissulfeto pelo pós doutorando Daniel Maragno Trindade, do laboratório nacional de Biociências (LNBio). Dr. Daniel conduziu um trabalho com a proteína Stanioalcina 1 humana (Trindade *et al*, 2009), a qual possui cinco pontes dissulfeto em sua estrutura, e uma delas é formada entre moléculas, formando assim um dímero. Diante dos bons resultados obtidos nesse trabalho, decidimos tentar utilizar esse sistema, sob sua orientação, utilizando as instalações do Laboratório Nacional de Biociências.

3.6.4- Clonagem de MpTLP1 em pFastBacEGFP

MpTLP1-pFastBac HTA foi digerida com as enzimas de restrição RSRII e HindIII, Dessa forma, o gene MpTLP1 é retirado do plasmídeo juntamente com a sequência referente à cauda de histidina na posição N- terminal. Essa sequência é importante, pois além de facilitar purificação da proteína, possibilita sua detecção através de Western- Blot utilizando anticorpo anti-his tag, já que não possuímos anticorpo anti- MpTLP1. O MpTLP1+histag foi inserido no plasmídeo pFastBacEGFP pela ligação ao plasmídeo igualmente digerido.

A presença de EGFP no plasmídeo é de suma importância para a visualização da transfecção de células de insetos e também em experimentos de titulação de estoque de vírus, já que é extremamente difícil a observação da formação de placas de lise no ensaio de *plaque assay* como sugere o protocolo. Após ligação, o plasmídeo foi transformado em *E. coli* DH10b, os transformantes foram submetidos à extração de DNA plasmídial por lise alcalina e então a presença do gene e sua fase de leitura foram confirmados através de sequenciamento no equipamento 3500 Genetic Analyzer (Applied Biosystems) utilizando primers que flanqueiam a região de clonagem do plasmídeo.

3.6.5- Confirmação da transposição

A recombinação foi confirmada através de PCR de acordo com recomendações do fabricante, utilizados primers M13 Forward e M13 Reverse. Esses primers flanqueiam a região *attTn7* de modo a resultar na amplificação de um fragmento de aproximadamente 300pb se o PCR for a partir de bacmídeo vazio. Se o PCR for a partir do bacmídeo com transposição por pFastBacDual (lembrando que estamos utilizando esse plasmídeo, porém com a adição do gene *EGFP*) resultará em um fragmento de aproximadamente 2500pb + o tamanho da construção (MpTLP1 + EGFP $\approx$ 1900 pb). Portanto, é esperada uma banda de aproximadamente 4500pb. Devido à necessidade de amplificação de um fragmento grande, o fabricante do sistema de expressão em baculovirus Bac-to-Bac® (Invitrogen) recomenda o uso de Taq Platinum® (Invitrogen). Todas as recomendações do fabricante foram seguidas e obtivemos 3 clones positivos (figura 13)

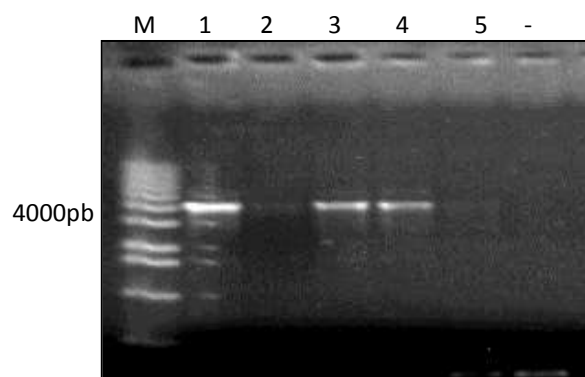


Figura 13: Confirmação de transposição da construção em pFastBacEGFP para bacmídeo. (Colônias 1 a 5). É possível observar a presença de uma banda de aproximadamente 4000pb, referente ao tamanho de amplificação esperado. (Controle negativo = -)

Após a confirmação de transposição o bacmídeo contendo MpTLP1+ histag foi utilizado para transfecção em células de inseto como já descrito anteriormente. Foi possível visualizar células com fluorescência verde através da microscopia de fluorescência em placas contendo células transfectadas. Após a produção do estoque viral P1, este foi quantificado e sua quantificação se deu através da metodologia de *plaque assay* (Bac-to-Bac® Manual, 2004) através da contagem de células apresentando fluorescência verde (devido à presença de *EGFP*) 48 horas após a infecção viral. A quantificação de P1 foi de $3,6 \times 10^6$ ufc/ml. Para produzir o estoque de vírus P2 foi utilizado MOI= 1. A produção de P2 resultou em um montante de 100 ml de vírus a concentração de 3×10^7 ufc/ml, determinado da mesma forma que o título de P1. Após a titulação, um experimento para determinação de tempo de incubação x MOI de vírus utilizada foi executado. O Resultado do experimento podem ser observados nas figuras 14 e 15.

Podemos observar no primeiro caso, em gel referente a amostras da fração insolúvel (figura 14) que temos muita marcação com anticorpo anti-his na altura de aproximadamente 25 KDa, tamanho aproximado de MpTLP1, o que indica uma grande quantidade de MpTLP1 sendo produzida de forma insolúvel em 48 horas de incubação. Podemos perceber também algumas marcações abaixo do tamanho de 25kDa nas amostras referentes a 72 horas de incubação ainda em amostras da fração insolúvel, o

que pode indicar que nesse tempo ocorre ações de algumas proteases que quebram MpTLP1 degradam MpTLP1 resultando em proteínas de menores tamanhos com afinidade ao anticorpo.

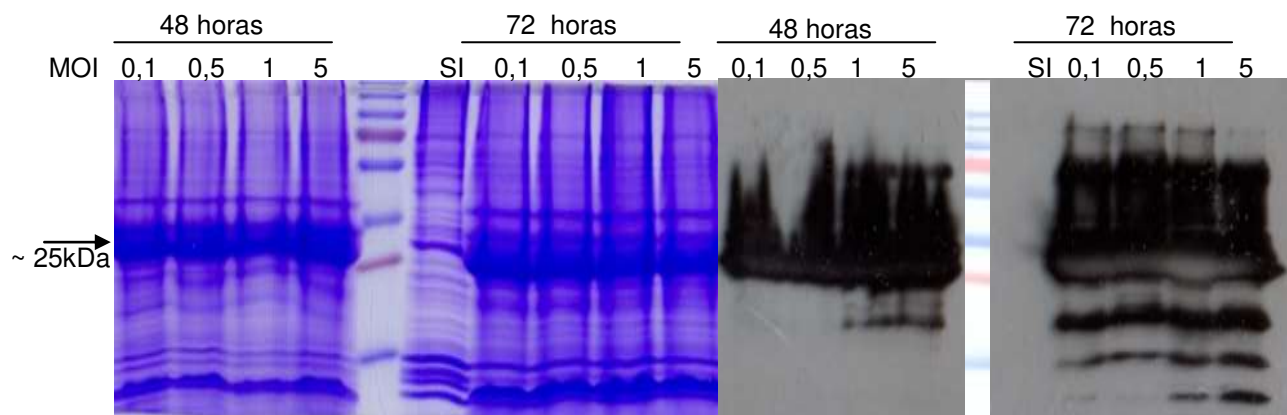


Figura 14: Tempo de incubação X MOI - MpTLP1. Frações insolúveis. Gel SDS-PAGE 12% (esquerda) e Western Blot (esquerda) ambos contendo amostras de testes de indução em quantidade crescente de MOI viral (0,1 a 5) em 48 horas de incubação e 72 horas de incubação referentes a frações insolúveis do experimento. Seta mostra bandas do tamanho esperado. A hibridação mostra a presença da proteína formando que possivelmente está formando dímeros e também sofrendo degradação

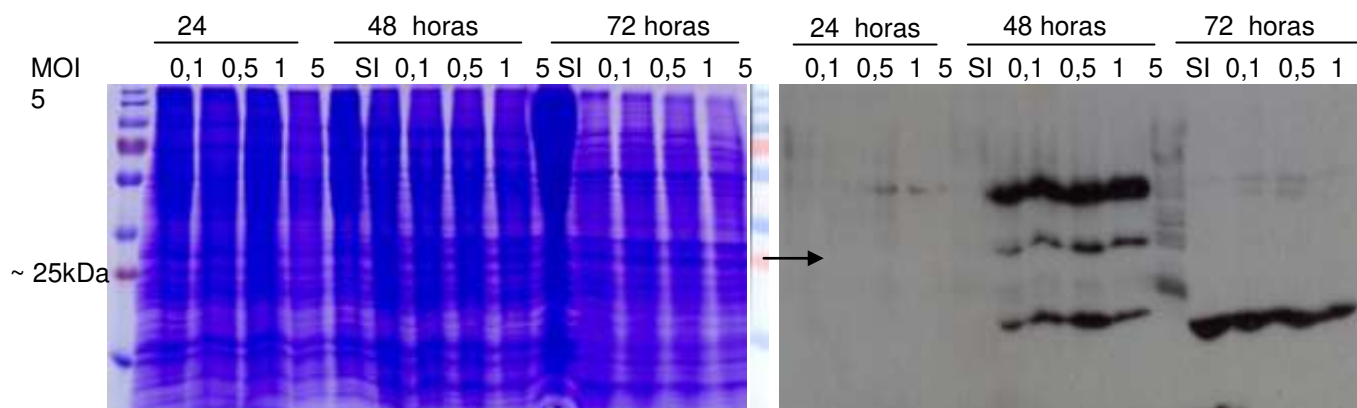


Figura 15: Tempo de incubação X MOI de MpTLP1. Frações solúveis. Gel SDS-PAGE 12% (esquerda) e Western Blot (esquerda) ambos contendo amostras de testes de indução em quantidade crescente de MOI viral (0,1 a 5) em 24, 48 horas de incubação e 72. Nessas últimas foram adicionadas amostras sem vírus (SI). As amostras são referentes a frações solúveis do experimento. Seta mostra banda tamanho da banda esperada. Em 48 horas foi obtido proteína possivelmente formando dímeros, e também sofrendo degradação, enquanto em 72 horas a proteína sofre degradação completa.

Nas imagens correspondentes as amostras obtidas das frações solúveis do experimento (figura 15), podemos observar que em 24 horas de incubação não é possível identificar a presença da proteína, ao passo que em 48 horas de incubação é possível observar 3 bandas definidas. Isso pode estar ocorrendo devido a possíveis formações indesejáveis de dímeros de MpTLP1, e também, a possível ocorrência da degradação da proteína. Em 72 horas de incubação, só é possível visualizar em Western- blot a possível degradação de MpTLP1.

A partir desse experimento, foi determinado que a melhor condição de expressão de MpTLP1- histag é em 48 horas de incubação utilizando MOI viral = 1, pois nesse caso, apesar de haver bandas de degradação e dimerização, foi a condição em que melhor se observa a presença de MpTLP1- histag em maior quantidade.

3.6.6- Experimento de expressão de MpTLP1-histag em larga escala

O Experimento foi efetuado conforme já descrito em materiais e métodos do presente documento, bem como sua purificação. As amostras referentes ao lisado e lavagens com 10mM, 100mM e 500mM de Imidazol, bem como uma amostra correspondente a fração insolúvel do experimento foram corridas em gel SDS-Page, bem como submetidas a teste de detecção de MpTLP1 através de Western- blot. Como mostra a figura 15.

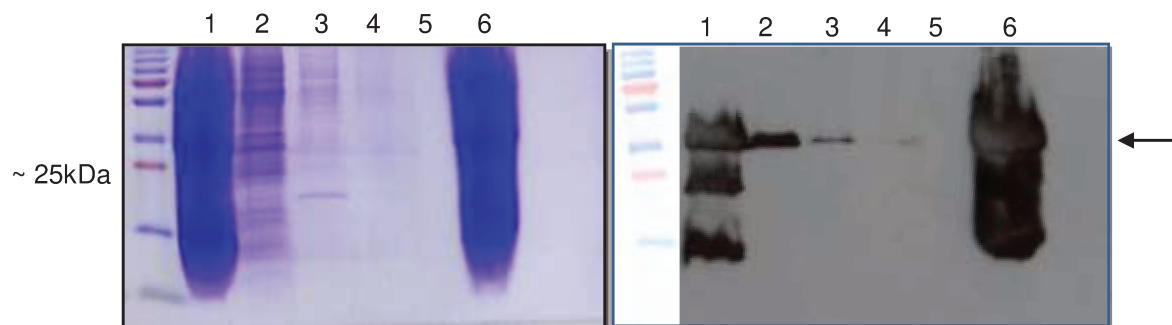


Figura 16: Indução em larga escala MpTLP1. Gel SDS-PAGE 12% (esquerda) e Western Blot (esquerda) ambos contendo amostras de testes de indução em larga escala, frações de purificação em coluna de afinidade: 1- lisado, 2 -lavagem a 10mM de Imidazol, 3- Lavagem a 100mM de imidazol, 4- lavagem a 500mM de imidazol. 5- vazio, 6: fração insolúvel do experimento.

Na imagem de Western-blot podemos identificar a presença de MpTLP1 predominantemente na fração insolúvel do experimento. Também é possível notar que MpTLP1 presente na fração solúvel se liga fracamente a resina de cobalto, e se desliga em sua maioria quando submetida à lavagem em baixa concentração de Imidazol (10mM). Esse experimento foi repetido e obtivemos o mesmo resultado.

A predominância de MpTLP1 nas frações insolúveis dos experimentos em células de inseto, nos levou a pensar que provavelmente essa proteína ainda não está sendo produzida no *fold* correto. E isso pode ter sido gerado pela falta da sequência sinalizadora de secreção da proteína no momento de sua amplificação e clonagem. Isso resultaria na falta de encaminhamento da proteína para determinados compartimentos celulares onde ocorre a formação de pontes dissulfeto, formando assim uma proteína truncada armazenada dentro de vesículas presente no citoplasma celular. Essa hipótese é reforçada pelo fato de que o grupo o qual trabalha com o sistema de expressão em baculovírus frequentemente amplifica o gene juntamente com sua sequência sinalizadora, e que este grupo vem tendo resultados excelentes na produção de proteínas heterólogas através desse sistema. Diante disso, uma nova estratégia de clonagem foi adotada para a obtenção de *MpTLP1*+ peptídeo sinal + *histag* na mesma construção.

3.6.7- Clonagem de MpTLP1+ peptídeo sinal + histag em pFastBacEGFP C-HT

Novos oligonucleotídeos foram desenhados para a amplificação de *MpTLP1*+ peptídeo sinal e *MpTLP4*+ peptídeo sinal a partir de cDNA de *M. perniciososa*. Os primers possuem, além da sequência dos sítios de restrição *RsrI* e *HindIII* para *MpTLP1* e *EcoRI* e *HindIII* para *MpTLP4*, uma sequência extra para que seja possível fazer a digestão direta do produto de PCR sem a necessidade de subclonagem em pGEM-T Easy, como vem sendo feito durante todo o projeto.

Apenas o PCR 2 de *MpTLP1* resultou na amplificação de uma banda no tamanho esperado (~800pb), porém, em todos os casos é possível observar a presença de uma banda de tamanho maior que o esperado. Isso se dá pela contaminação de DNA genômico no cDNA utilizado para a amplificação desses genes. A banda referente à *MpTLP1* foi purificada do gel e em seguida digerida conforme descrito em materiais e métodos, assim como o plasmídeo pFastBacEGFP C-HT (figura 17).

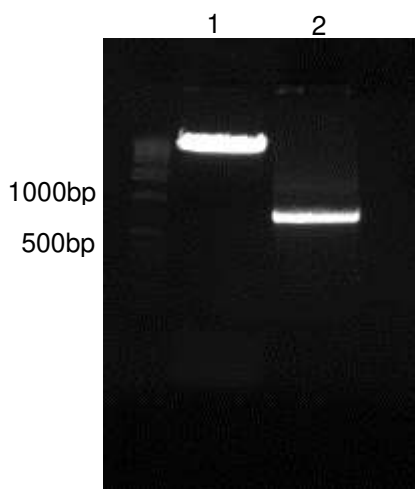


Figura 17: Digestão de pFastBacEGFP C-HT (1) e *MpTLP1*+ peptídeo sinal (2)

A utilização de um plasmídeo contendo cauda de histidina se faz necessário, pois é a única forma de detecção da presença de proteína através de hibridação com anticorpo anti-his pela técnica de Western- blot. Nós não possuímos um anticorpo próprio de *MpTLP1*. Além disso, o uso da cauda em purificação em coluna de afinidade simplifica o processo de obtenção de proteína pura.

Após a ligação de *MpTLP1*+ peptídeo sinal em pFastBacEGFP C-HT, houve a transformação em *E.coli* DH10B seguida de plaqueamento em meio seletivo contendo ampicilina 100mg/ml. Colônias positivas foram selecionadas, seguida de purificação de DNA plasmídial e o plasmídeo possuindo a construção foi confirmado através de sequenciamento no equipamento 3500 Genetic Analyzer (Applied Biosystems).

Esse plasmídeo foi utilizado para transformação em DH10BAC™ (Invitrogen), e todo o processo de seleção de transformantes, obtenção de vírus e mini testes de produção de proteína foi conduzido, porém não obtivemos resultados satisfatórios.

Após todo esforço para a obtenção de MpTLPs de forma solúvel não apresentarem resultados satisfatório, e considerando o prazo avançado do período de doutoramento, essas técnicas foram descontinuadas do trabalho. Consideramos que exploramos muitas técnicas de obtenção de MpTLPs solúveis desde o trabalho de mestrado de aluna, onde foram testados diferentes plasmídeos de expressão em *E. coli*, e também diferentes organismos de expressão como levedura e células de inseto durante o período total de trabalho.

4- CAPÍTULO II: Estudo de genes candidatos a efetores secretados de patogenicidade de *M. pernicioso* (*MpCSEPs*)

Análise de transcriptoma de vassoura verde feito durante o trabalho de doutorado do aluno Paulo Teixeira revelou alguns genes candidatos a efetores de patogenicidade em *M. pernicioso*. A versão atualizada do genoma do fungo conta com 193 proteínas candidatas a efetores secretados (CSEP), que foram definidas como proteínas que são secretadas e que não possuem homólogos caracterizados em outros organismos sequenciados até então (Pedersen *et al.*, 2012).

Grande parte dos genes mais altamente expressos pelo fungo na vassoura verde codifica uma *CSEP*, sugerindo uma importante participação dessa proteína durante a infecção. Um total de 35 *CSEP* com valor de RPKM >50 foi identificado em plantas infectadas. Em geral, esses genes codificam proteínas pequenas e ricas em cisteínas que são características típicas de efetores de virulência (Stergiopoulos and de Wit, 2009). Essas proteínas possuem as mesmas características de proteínas codificadas pelos genes AVR de patógenos (Van den Ackerveken *et al.*, 1993). Quando analisamos o transcriptoma do fungo foi possível notar que pelo menos 35 dessas *CSEPS* são preferencialmente ou até exclusivamente expressos durante a interação do fungo em fase biotrófica com o cacau, podendo ser mais uma evidência de que esses genes codificam efetores potenciais.

No momento do início desse estudo, a análise de transcriptoma do fungo estava em andamento, e gerou, inicialmente, 16 sequências *CSEP*. A caracterização dessas proteínas foi inserida no presente trabalho de doutoramento juntamente com a caracterização das *MpTLPs* no projeto “SMOLBNET 2.0: Estudos estruturais de proteínas chave para as doenças fúngicas do cacau Vassoura de bruxa e Monilíase – Desenvolvimento de estratégias de controle e entendimento de modelos de patogenicidade” foi conduzido no LNBio - Laboratório Nacional de Biociências.

4.1- Justificativa

Sabemos que a interação *T. cacao*- *Moniliophthora perniciosa* é benéfica apenas ao fungo, sendo, portanto, uma interação compatível (quando a doença ocorre). A planta, apesar de ativar seu sistema de defesa, não se mostra eficaz no combate da doença (Teixeira et al, 2013) e isso pode estar relacionado à produção de efetores de patogenicidade que podem ser capazes de neutralizar o sistema de defesa da planta. Pensando nisso, faz-se importante a caracterização das proteínas que nesse trabalho serão nomeadas MpCSEP (*Moniliophthora perniciosa* Candidate secreted effectors).

4.2- Objetivos

Este capítulo refere-se à caracterização de MpCSEPs através de clonagem e superexpressão dessas proteínas em sistema heterólogo (*E. coli*) seguido de purificação, obtenção de cristais para resolução estrutural destas proteínas. Este estudo, assim como o estudo discutido no capítulo 1, fez parte do projeto “SMOLBNET 2.0: Estudos estruturais de proteínas chave para as doenças fúngicas do cacau Vassoura de bruxa e Monilíase – Desenvolvimento de estratégias de controle e entendimento de modelos de patogenicidade”.

4.3- Materiais e Métodos

4.3.1- Análise de expressão gênica de *MpCSEPs*

As bibliotecas e amostras analisadas foram as mesmas utilizadas no artigo do capítulo 1.

4.3.2- Clonagem dos genes *MpCSEPs*

4.3.2.1 - Organismos hospedeiros

As células competentes de *E. coli* utilizadas nas subclonagens eram da cepa Top10 (Invitrogen). A cepa de *E. coli* utilizadas para expressão de proteínas recombinantes Origami(DE3) (Novagen).

4.3.2.2 - Plasmídeos

Para subclonagem dos fragmentos amplificados por PCR foi utilizado o plasmídeo pGEM-T Easy (Promega, EUA). Para a superexpressão de proteínas recombinantes em bactéria foram usados os plasmídeos pET28a e pETSUMO (Novagen, EUA),

4.3.2.3- Amplificação e subclonagem dos genes *MpCSEPs*.

Utilizando as sequências codificantes geradas a partir do software Augustus gene finder, pares de oligonucleotídeos foram desenhados para amplificar os genes a partir de cDNA de vassoura verde, adicionando as sequências reconhecidas pelos sítios de restrição *Bam*HI nos oligonucleotídeos direto e *Hind*III nos reversos. Os genes *MpCSEPs* foram amplificados utilizando esses pares de primers e os fragmentos obtidos foram clonados em pGEM-T Easy e transformado em *E. coli* Top10 através da técnica de choque térmico. Os transformantes foram selecionados em meio de cultura seletivo contendo antibiótico (ampicilina 100µg/ml) e X-gal (técnica de α -complementação). Esses clones foram crescidos em meio LB contendo ampicilina (100µg/ml) e em seguida submetidos à extração de DNA plasmidial por lise alcalina (Sambrook et al., 1996).

4.3.2.3- Digestão das *MpCSEPs* e clonagem em vetores de expressão em *E.coli*.

Os plasmídeos obtidos foram digeridos com as enzimas de restrições *Bam*HI e *Hind*III. As digestões foram aplicadas em gel de agarose 0,8% e os fragmentos referentes

às sequências dos genes *MpCSEPs* foram purificadas do gel através do Kit PureLink (Invitrogen, EUA), e utilizadas para ligação aos plasmídeos de expressão de proteína PetSUMO. Após a ligação, os plasmídeos foram novamente transformados em *E.coli* Top10 e os transformantes foram selecionados através de plaqueamento em meio de cultura contendo antibiótico (canamicina 25µg/ml) e posteriormente através da técnica de PCR. Confirmados os clones, uma nova purificação de DNA plasmidial por lise alcalina foi realizada, e assim obtivemos o gene ligado nos plasmídeos de expressão. O plasmídeo purificado foi utilizado para confirmação da construção através de sequenciamento no equipamento 3500 Genetic Analyzer (Applied Biosystems).

As construções pETSUMO- *MpCSEPs* foram inseridas em *E.coli* Origami DE3 pela técnica de transformação por choque térmico. Os transformantes foram plaqueados em meio de cultura contendo canamicina (25µg/ml) streptomicina (25mg/ml) e tetraciclina (5mg/ml).

4.3.3- Teste de indução e solubilidade

Um teste de expressão das proteínas clonadas foi feito em 100 ml de LB, seguido de teste de solubilidade:

Foi adicionado 600ul de pré inoculo (25µg/ml de canamicina, 25mg/ml de streptomicina e 5mg/ml de tetraciclina+ 1 colônia de Origami transformada, crescimento ON) em 2 erlenmeyers contendo 100 ml de meio de cultura com antibióticos já citados. Foram incubados a 37 °C por aproximadamente 6 horas (OD= 0,8) 200 RPM. Uma alíquota de 300ul foi retirada da cultura nesse ponto (amostra não induzida). Após isso, um dos erlenmeyers foi transferido para um shaker a 18 °C e mantido até a temperatura estabilizar, logo em seguida foi adicionado 0,2mM de IPTG no meio de cultura a 18°C e 37°C, 200 RPM, e foram mantidos nessas condições durante a noite. A cultura foi então centrifugada a 5000 RPM 4 °C, o meio de cultura foi descartado e as células foram

submetidas à lise com tampão contendo 50mM Tris-HCl pH8.5, 150mM NaCl, Glicerol 10%, 15mg Lisozima, 4mg DOC, 25ul DNase, com 1 hora de incubação em gelo. O lisado foi centrifugado por 13000 RPM por 45 min (4 °C) e o sobrenadante foi recolhido, o que corresponde à fração solúvel do extrato celular.

Adicionamos ao extrato clarificado 100ul de resina de níquel (Talon- Clontech), incubando durante 1 hora em geladeira e logo depois o material foi centrifugado, o sobrenadante foi descartado e a resina sofreu 2 lavagens seguida de centrifugação com tampão contendo 50mM Tris-HCl pH8.5 e 150mM NaCl. Todas as fases desse procedimento foram avaliadas em gel SDS-PAGE 15%.

4.3.4- Expressão e purificação em grande escala.

Expressão em larga escala foi realizada em erlenmeyers contendo 1l de LB + antibióticos nas concentrações já descritas anteriormente (total de 8 litros por proteína). O pré inóculo agora é realizado em 50 ml de meio, e 6 ml/litro de indução é utilizado para inóculo. A indução foi feita com a mesma concentração de IPTG, e a temperatura escolhida para indução foi 18 °C nas mesmas condições de agitação e mesmo tempo. O tampão de lise utilizado manteve-se aos componentes e concentrações descritas anteriormente. As células foram sedimentadas a 8000 RPM durante 20 minutos em centrífuga refrigerada, o meio de cultura foi descartado e as células foram submetidas à lise, adicionando a cada litro de indução 80 ml de tampão de lise (50mM Tris-HCl pH8.5, 150mM NaCl, Glicerol 10%, 138mg Lisozima, 10mg DOC, 1ul DNase/ litro de indução), e incubando em gelo por 1 hora.

Após lise, o extrato total celular é centrifugado a 18000 RPM, e o sobrenadante, fração correspondente a porção celular solúvel é submetida à purificação. Para a purificação, 1 ml resina de Níquel (Talon- Clontech) para cada litro de indução é empacotada em coluna, equilibrada com tampão de purificação 50mM Tris-HCl pH8.5 e

150mM NaCl, e o extrato total celular é adicionado à coluna. Logo após, a resina contendo a proteína aderida é lavada com 10X o volume de tampão, e incubada ON a 4°C sob agitação, com 20 U de ULP (Ub1 Specific Protease 1 ou Sumo protease) para a clivagem da cauda Sumo, separando a proteína da cauda de histidina que continua aderida ao metal da resina. Logo após, a proteína é eluída da resina.

Após a purificação por afinidade, a proteína obtida foi submetida à purificação por exclusão por tamanho para a completa retirada de qualquer contaminante que ainda existia na amostra. Essa purificação foi feita em coluna HiloLoad 16/60 Superdex 75 (GE Life Sciences) no equipamento FPLC Akta Purifier (GE Life Sciences). A coluna foi equilibrada com tampão de purificação já mencionado e a solução contendo a proteína foi concentrada até o volume de 2 ml (volume do loop utilizado no equipamento). A proteína purificada é concentrada até 20mg/ml para ser utilizada em experimento de cristalização.

Para MpCSEP14, antes da purificação por exclusão de tamanho, foi adicionada uma etapa de purificação com a utilização da coluna de troca iônica Mono-Q 5/50 (GE), que é equilibrada com tampão 50mM Tris-HCl pH8.5 e 35mM NaCl.

4.3.5- Cristalização

Para obtenção de cristais, as proteínas foram concentradas a 20mg/ml em concentradores Centricon 10kDa (Millipore). Após purificação, as proteínas são submetidas a experimentos de cristalização no laboratório ROBOLAB - Laboratório Automatizado de Cristalização de Macromoléculas (LNBio). Inicialmente, os experimentos montados são com a utilização de kits iniciais de cristalização.

A proporção gota/reagente foi de 1:1 e o volume utilizado de solução do poço foi de 1/3. Para a montagem do experimento, proteínas e reagentes são pipetados e misturados automaticamente pelo equipamento Matrix Maker (Emerald Biosystems), e armazenados a temperatura controlada no equipamento Rock Imager 1000 (Formulatrix).

Esse equipamento gera e registra imagens do conteúdo da placa o que facilita o acompanhamento do experimento.

Tabela 2: Kits iniciais de cristalização

Kits	Características
Crystal Screen HT	Matriz esparsa original desenvolvido por Jankaric & Kim (1991)
JSCG+ Suite	Matriz esparsa otimizada desenvolvida por JSCG
PACT Suite	Varredura sistemática de pH e íons versus PEGs (JSCG)
Precipitant Synergy	Combinação sinérgica de agentes precipitantes mecanicamente distintos
SaltRx	Kit baseado em variações de sais e pH
Wizard I&II	Matriz esparsa da Emerald

4.3.5.1- Refinamento dos cristais

A partir das condições iniciais as quais resultaram em cristais de proteína, selecionamos o melhor cristal para refiná-lo. Para isso, as condições iniciais de formação de cristais sofreram variações em concentração de precipitante e pH. Para MpCSEP5, foram escolhidas duas condições iniciais a serem variadas: condição A: 25% PEG 3350; 10mM Bis-Tris pH 5.5; 200mM NaCl. Condição B: 20% PEG 6000; 10mM MÊS pH 6; 200mM Cloreto de Amônio. As variações na condição A foram feitas da seguinte maneira: Cloreto de Sódio (200mM a 50mM E a H) nas linhas, e concentrações de PEG3350 nas colunas (19% a 29% 1 a 6); Concentração de PEG3350 de 30% a 5% (linha E, 6 a 12), Faixa de pH (4.5 a 7 de 0,5 em 0,5 ponto) linha F, 7 a 12; Adicionar glicerol 3% em 3% (de 0 a 18%); Trocar Cloreto de sódio por Cloreto de lítio (50mM a 300mM) D 7a 12.

As variantes da condição B foram as seguintes: Concentração de PEG3350 de 30% a 5% (linha A, 7 a 12); Faixa de pH (5 a 7.5 de 0,5 em 0,5 ponto) linha B, 7 a 12; Adicionar glicerol 3% em 3% (até 18%); Trocar Cloreto de amônio por Cloreto de lítio (50mM a 300mM) D 7a 12.

Para MPCSEP14, as condições iniciais para formação de cristais foram: Tri-sodium citrate 0,1M pH5,6; 2-propanol 20%; PEG 4000 20%, e as variações para refinamento foram as seguintes: pH de 4,4 a 6,8 (de 0,2 em 0,2); 2-propanol: de 10% a 34% de 2 % em 2%; PEG4000: de 10% a 34% de 2% em 2%, aumento da concentração de PEG4000 e diminuir a de 2-propanol nas mesmas janelas citadas, variação crescente de PEG4000 em 25% de 2- propanol; Variação crescente de PEG4000 em 15% de 2-propanol.

Além disso, os *Kits Silver Bullets* e *Additive Screen* (Rampton Research) foram utilizados na tentativa de melhora dos cristais de MpCSEP5 em placas de gota pendente de 48 poços preparadas manualmente e mantidas a 18°C. O volume total de solução no poço foi de 200ul, e o volume total da gota foi de 2ul, para os dois kits, sendo 1ul de solução de proteína, 0,7% de tampão do poço e 0,3ul de aditivo, para Kit Additive Screen, e 1ul de solução de proteína, 0,5ul de solução de poço e 0,5ul do reagente nas gotas para o kit Silver Bullets.

Adicionalmente, a técnica de *Seeding* também foi empregada. Essa técnica consiste em maceramento de um cristal, cuja formação foi dada na condição inicial testada em placa montada mão e em seguida, espalhar as sementes desse cristal em uma gota contendo reagente+ proteína com a utilização de um fio de pelo de cavalo, em condições de variação a escolher, nesse caso, as variações foram dadas na solução precipitante e proporção proteína: solução. Foi realizada gota sentada em condição cuja formação de cristal foi observada em placa montada mecanicamente: PEG3350 21%, Bis-Tris pH 5.5 10mM, NaCl 50mM. 700ul dessa solução foi colocada em cada pocinho, e para montagem da gota, variamos a proporção de solução proteica: solução de poço, nas proporções de 1:1, 2:1, 3:1, 1:2, 1:3, com volume final da gota= 2ul. O gradiente de concentração de PEG 3350 foi de 14% a 20%. Após montagem das gotas, a placa permaneceu aclimatando a 18°C durante 24 horas, para então as sementes de cristal ser

espalhadas na gota. A diferença das técnicas que envolvem gota sentada e gota pendente pode ser visualizada na figura 18.

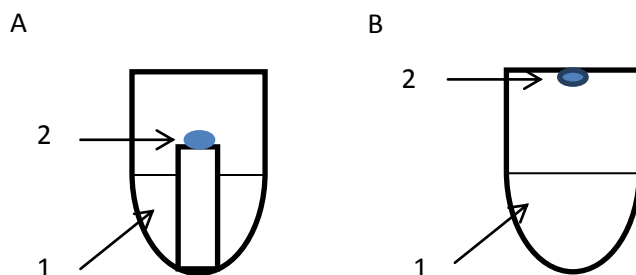


Figura 18: Desenho esquemático de poços de placas de gota sentada (A) e gota pendente (B). 1- solução do poço, 1- gota contendo solução proteica + solução de poço, podendo conter ainda algum outro reagente ou aditivo.

4.3.6- Dicroísmo Circular (CD)

Os experimentos de dicroísmo circular foram realizados no espectropolarímetro JASCO J-810, com a temperatura controlada a 20 °C através de um sistema interno de controle de temperatura (*Peltier type control system* PFD 425S, JASCO). Os espectros de dicroísmo circular no UV distante (250-190 nm) foram obtidos utilizando-se uma cubeta de quartzo de 0,1 cm de caminho ótico, e 0,2 mg/mL de proteína solução. Todos os espectros de dicroísmo circular foram corrigidos com a subtração dos espectros dos tampões utilizados e os dados gerados com a subtração desses espectros foram convertidos em elipticidade molar por resíduo, $[\theta]$, dado em graus $\text{cm}^2 \text{dmol}^{-1}$, calculado de acordo com a equação:

$$[\theta] = (\phi \text{ MRW}) / (10 c l)$$

onde ϕ é a elipticidade (em graus), MRW é o peso médio do resíduo, c é a concentração da proteína em g/mL e l é o caminho óptico em cm.

4.3.7- Espalhamento dinâmico de luz (DLS)

Os experimentos de espalhamento dinâmico de luz foram realizados com o equipamento DynaPro-99-E-15 (Protein Solutions). Uma concentração de 0,2 mg/mL de proteína foi utilizada para os experimentos. A temperatura foi mantida a 25 °C durante todo o experimento. Foram realizadas 100 medidas com o objetivo de melhorar a estatística dos resultados. Os dados foram tratados com o programa fornecido pelo fabricante.

4.4- Resultados

4.4.1- Mineração das sequências *MpCSEPs*

Em análise de transcriptoma realizado durante o Doutorado do aluno Paulo Teixeira, foi observada a existência de proteínas desconhecidas que apresentavam uma alta expressão gênica exclusivamente, ou preferencialmente em bibliotecas referentes à interação planta-patógeno, não apresentando expressão em bibliotecas de estruturas reprodutivas do fungo ou mesmo em fungo cultivado in vitro na maioria dos casos. Essas bibliotecas foram descritas no capítulo 1 desta tese (artigo submetido). Nessa primeira análise, foram selecionados 16 genes.

A partir dessa análise, as sequências correspondentes a esses genes foram utilizadas em tblastn e verificou-se ausência ou, quando existia, uma baixíssima similaridade com alguma região de algum gene. As sequências desses genes desconhecidos foram então traduzidas e se tornaram candidatas ao estudo e caracterização das mesmas após confirmação da existência de peptídeo sinal em todas elas através de análise pelo programa *SignalP* 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>).

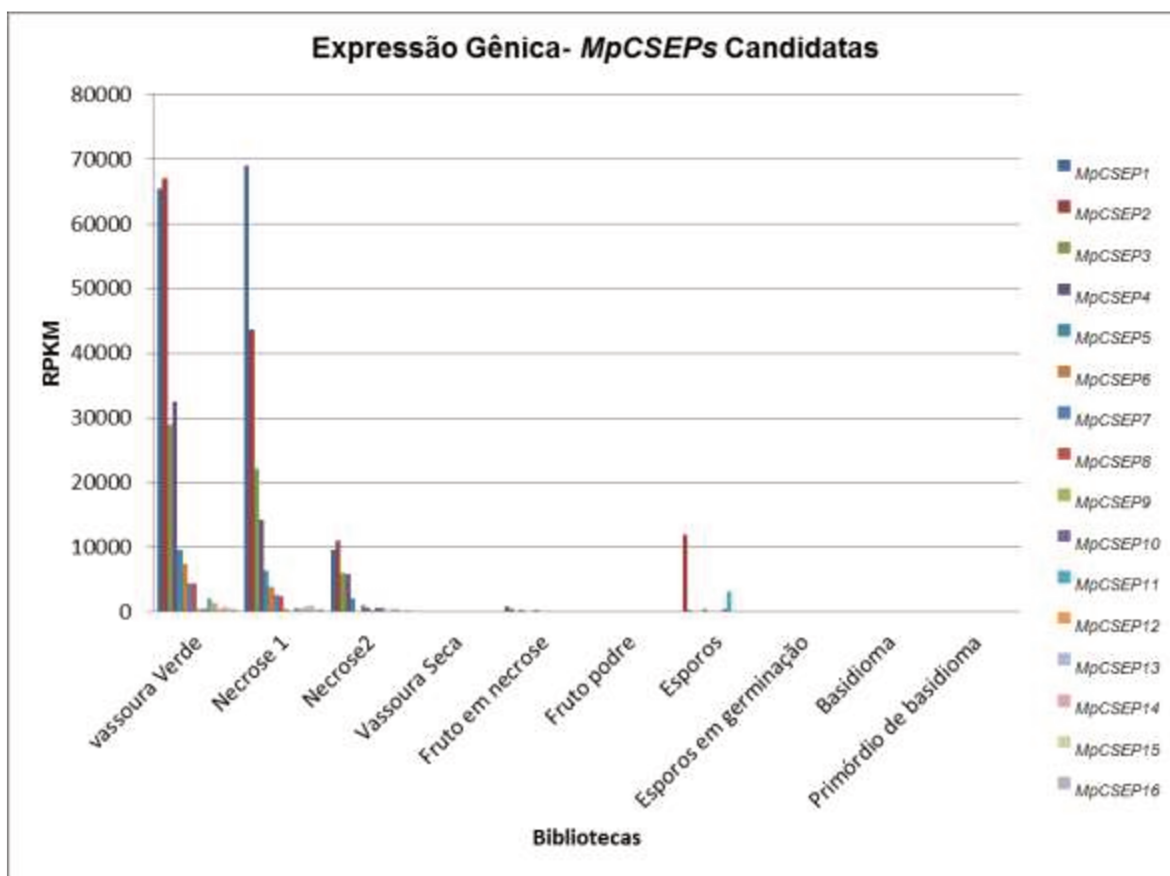


Figura 19 Expressão gênica de candidatas a efetores. Eixo Y: Valores de RPKM, eixo X: Bibliotecas analisadas.

A primeira triagem de seleção de genes entre esses candidatos ao estudo foi realizada a partir da exclusão das proteínas que apresentavam uma grande quantidade de cisteínas em suas sequências, devido à complexidade de expressão heteróloga desse tipo de proteína em sistema de *E. coli*. Esta seleção excluiu 5 genes, restando 11 genes em nosso estudo. Entre estes, o gene *MpSCEP 8* foi excluído, pois não conseguimos uma sequência de qualidade em nossos bancos de dados para a confecção de primers para amplificação dos mesmos.

Tabela 3: Características dos genes *MpCSEPs*

Nome do Gene	Tamanho CDS (pb)	número de introns	Número de Cisteínas	Tamanho da proteína (aa)	Peptídeo sinal	Peptídeo de membrana
<i>MpCSEP1</i>	738	2	5	246	sim	não
<i>MpCSEP2</i>	555	3	9	185	sim	não
<i>MpCSEP3</i>	984	3	15	328	sim	não
<i>MpCSEP4</i>	810	2	5	270	sim	não
<i>MpCSEP5</i>	690	2	0	230	sim	sim
<i>MpCSEP6</i>	594	3	9	183	sim	não
<i>MpCSEP7</i>	930	2	6	310	sim	não
<i>MpCSEP8</i>	246	1	2	82	sim	não
<i>MpCSEP9</i>	324	2	6	108	sim	não
<i>MpCSEP10</i>	600	2	3	200	sim	não
<i>MpCSEP11</i>	570	5	10	190	sim	não
<i>MpCSEP12</i>	711	2	2	237	sim	sim
<i>MpCSEP13</i>	717	4	3	239	sim	sim
<i>MpCSEP14</i>	477	2	3	159	sim	não
<i>MpCSEP15</i>	309	1	6	103	sim	não
<i>MpCSEP16</i>	441	2	1	147	sim	não

4.4.2- Amplificação e clonagem de *MpCSEPs*

A partir da seleção descrita anteriormente, foram desenhados primers para amplificação das 11 sequências codantes restantes, contendo sítios de enzimas de restrições *Bam**HI* no primer direto e *Hind**III* no primer reverso, para posterior clonagem em vetor de expressão heteróloga de proteína. A sequência que corresponde ao peptídeo sinal de cada gene foram excluídas em nossa estratégia de clonagem, assim como regiões transmembrana, detectada pelo programa InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>). Os *primers* para amplificação de cada gene estão dispostos na tabela a seguir.

Tabela 4: Primer para amplificação e clonagem dos genes *MpCSEPs*. Em realce sequências reconhecidas pelas enzimas de restrições *Bam*HI para os primers diretos, e *Hind*III, para os primers reversos.

	<i>Direto (5'-3')</i>	<i>Reverso (5'-3')</i>
<i>MpCSEP1</i>	ATCCC GGATCC ATCCCTACCCCCGTAATGTA	TCGACA AAGCTT CTAGTTGTGGAAGAACGTAGAGGG
<i>MpCSEP4</i>	ATCCC GGATCC CTTCCAACCCAGCTCCCAA	TCGACA AAGCTT CTACAACGAGATCGACACTCCGAA
<i>MpCSEP5</i>	ATCCC GGATCC CTCCCGAGCCCAAAGGTTGA	TCGACA AAGCTT TCACAAATCGACTGAAAGGCCA
<i>MpCSEP7</i>	ATCCC GGATCC CTTCCAACCCAGTTCCCAA	TCGACA AAGCTT TTAAAAGGAGATCACCCTC
<i>MpCSEP9</i>	ATCCC GGATCC ACACCAACCAAGACACCAAA	TCGACA AAGCTT TCAAATATTCACGTCCCAGC
<i>MpCSEP10</i>	ATCCC GGATCC CTTCCAGCCCAGCTGCT	TCGACA AAGCTT TTACAACGAAATCGACACTCCC3
<i>MpCSEP12</i>	ATCCC GGATCC CTTCCAGTCCAGCTGCCC	TCGACA AAGCTT TTATTGAGCATGGATCTCTTCAGAG
<i>MpCSEP13</i>	ATCCC GGATCC TTCAATATCACTGGCCCCGT	TCGACA AAGCTT TTACAATGCACCGGTGGAGAG
<i>MpCSEP14</i>	ATCCC GGATCC TCTCTCGTCTCCCGTGATGGT	TCGACA AAGCTT TTACTTGAGGGCGATAGTACACTTG
<i>MpCSEP15</i>	ATCCC GGATCC GCAAAAACCTTGTGTGCGT	TCGACA AAGCTT CTAGGCGGCCCATTCACACA
<i>MpCSEP16</i>	ATCCC GGATCC GCCAGCATAACTGAGCCTCA	TCGACA AAGCTT TTACGCCTTGACAACATTGAC

Para a amplificação desses genes, foi utilizado cDNA obtido a partir de RNA extraído de plantas doentes. Testes de amplificação foram feitos utilizando RNA das fases da doença com 30 e 45 dias após inoculação, necrose 1, necrose 2, vassoura seca e micélios de *M. perniciosa* cultivado *in vitro* na presença de extrato de cacau em meio de cultura. Grande parte dos genes foram amplificados na condição de necrose 1, portanto esse cDNA foi utilizado para amplificação e clonagem desses genes. Porém, os genes *MpCSEPs* 7, e 15 não apresentaram produto de amplificação e foram descontinuadas do presente estudo.

Os demais genes foram subclonados em pGEMT Easy (Promega) e transformado em células TOP10 para propagação e plaqueados em meio de seleção. Colônias positivas foram então inoculadas em meio de cultura líquido contendo Ampicilina (100ug/ml) e logo após, essas células foram submetidas à lise alcalina (Sambrook, 2001) para purificação dos plasmídeos. Esses plasmídeos foram então digeridos utilizando as enzimas de restrição com sequências de reconhecimento inseridas pelos primers e o

produto de digestão foi purificado de gel de agarose 1% pelo kit Wizard® SV Gel and PCR Clean-Up System (Promega).



Figura 20: Amplificação de genes *MpCSEPs* a partir de cDNA de cacau doente na fase de necrose 1. Os genes *MpCSEP7* e *MpCSEP15* não apresentaram produto amplificado.

Paralelamente, o vetor para expressão heteróloga de proteína PETSumo também foi submetido aos mesmos procedimentos descritos neste parágrafo e foi utilizado para a ligação dos genes. Os produtos dessas ligações foram transformados em *E. coli* TOP10 para a multiplicação da construção, que foram novamente purificadas através de lise alcalina e utilizadas para transformação em células para expressão heteróloga de proteína Origami. O gene *MpCSEP4* não apresentou produto de digestão, enquanto o gene *MpCSEP13* não apresentou colônia positiva após transformação em Origami, e foram descontinuados do estudo.

4.4.3- Teste de expressão, solubilidade e purificação das *MpCSEPs*

As células contendo o vetor de expressão + gene foram selecionadas por plaqueamento em meio de cultura seletivo, e colônias foram submetidas a PCR para confirmação da presença do inserto. Confirmado o inserto, as colônias foram inoculadas em um pré- inóculo de 2 ml e o mesmo foi utilizado para um inóculo em 100 ml de LB contendo antibióticos conforme descrito em materiais e métodos. A indução de todas as proteínas testadas foi feita em duas temperaturas, 30°C e 18°C para verificar se há

diferença na quantidade de proteína produzida bem como em seu estado de solubilidade (figura 21).

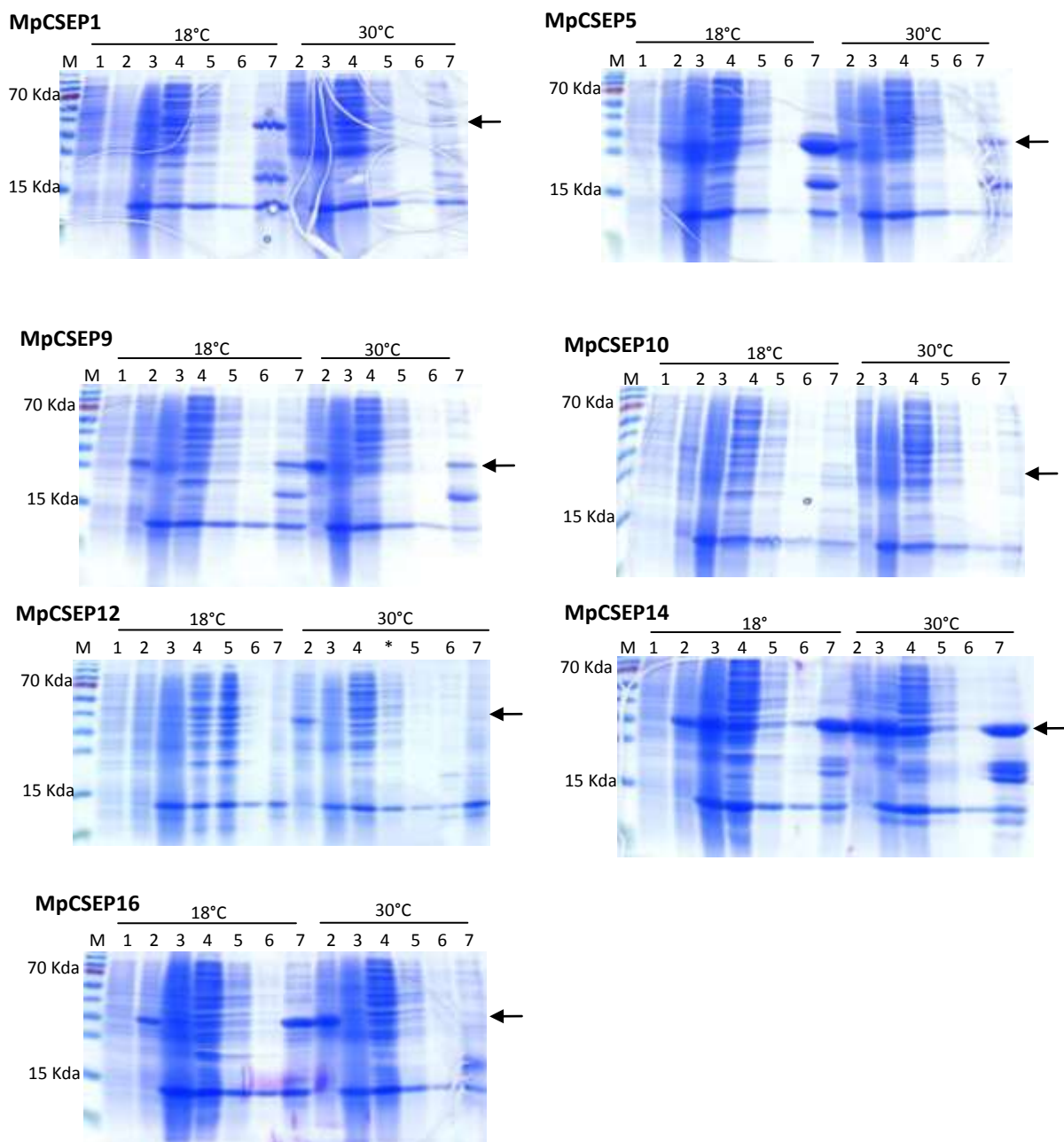


Figura 21: Teste de expressão e solubilidade das proteínas MpCSEP1, MpCSEP5, MpCSEP9, MpCSEP12, MpCSEP14, MpCSEP16 em 18°C e 37°C. M- Prestained protein ladder 10-170, Fermentas; 1- Sem indução; 2- Induzido; 3- Extrato total; 4- Flowthrough; 5- lavagem 1; 6- lavagem 2; 7- Resina+ Proteína , *- repetição da amostra anterior. Setas indicam o tamanho esperado das bandas referentes a proteína + cauda de histidina + cauda SUMO

Alíquotas das culturas antes da indução com IPTG foram retiradas e após indução as células foram sedimentadas e lisadas. O lisado foi submetido a uma purificação para verificar se a proteína, quando produzida, permanecia na fração solúvel do lisado celular.

Foi possível observar a presença de banda diferencialmente expressa na amostra 7, correspondente à proteína ligada na resina de afinidade à cauda de histidina (níquel-TALON) para as proteínas MpCSEP1, MpCSEP5, MpCSEP9, MpCSEP14 e MpCSEP16, e a temperatura que melhor produziu proteína na fração solúvel foi a de 18°C. A partir daqui, todas as induções em larga escala foram feitas sob esta temperatura. Os genes que não apresentaram proteína expressa foram descontinuados desse trabalho.

As proteínas que se mostraram solúveis no teste anterior foram então expressas em 1l de LB, para checar se elas permaneciam solúveis em todos os passos da purificação por afinidade. Nesta etapa, observamos que MpCSEP1, após incubação de proteína ligada à resina com ULP, sofria precipitação de uma parte da proteína (figura 22). Para MpCSEP9 e MpCSEP16 a digestão com a enzima ULP para liberação da proteína não ocorria de forma satisfatória, mesmo testando outras concentrações da enzima, resultando em uma quantidade quase imperceptível em gel da proteína pura (figura 23). MpCSEP5 e MpCSEP14 foram as que melhor se comportaram e resultaram em uma quantidade satisfatória de proteína ao final da purificação por afinidade (Figura 24).

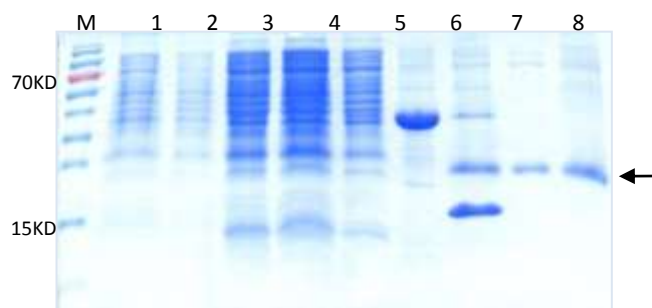


Figura 22: Expressão e purificação de MpCSEP1: M- Prestained protein ladder 10-170, Fermentas; 1- Sem indução; 2- Induzido; 3- Extrato total; 4- Flowtought; 5- lavagem 1; 6- Resina+ Proteína, 7- MpCSEP1 + resina + ULP; 8- MpCSEP1 eluída; 9- precipitado formado durante a incubação da proteína ligada à resina+ ULP. Gel SDS-PAGE 12,5%. Seta indica proteína esperada.

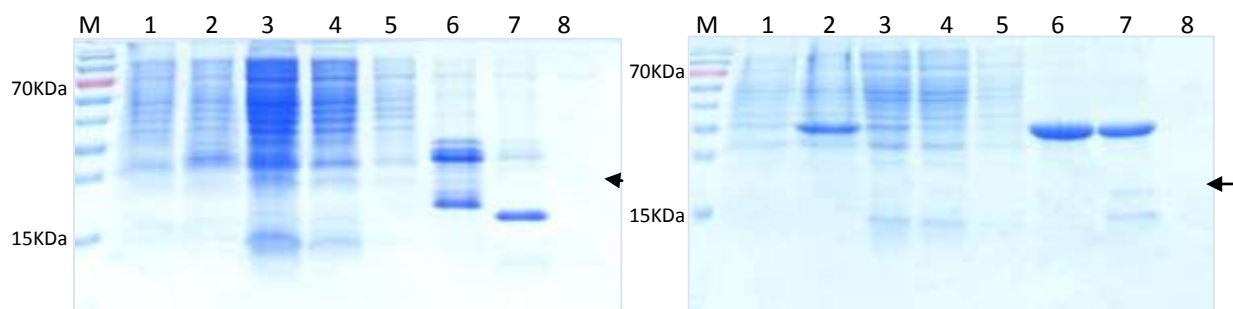


Figura 23: Expressão e purificação de MpCSEP9 (esquerda) e MpCSEP16 (direita): Gel SDS- PAGE 12,5%. M- Prestained protein ladder 10-170, Fermentas; 1- Sem indução; 2- Induzido; 3- Extrato total; 4- Flowtought; 5- lavagem 1; 6- Resina+ Proteína , 7- MpTLP1 + resina + ULP; 8- Proteína eluída. Gel SDS-PAGE 12,5%. Seta indica proteína esperada.

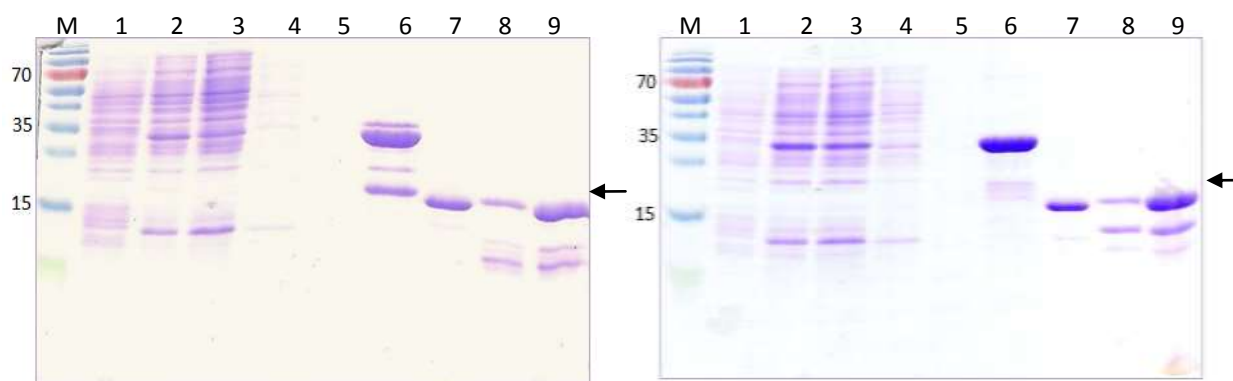


Figura 24: Expressão e purificação de MpCSEP5 (esquerda) e MpCSEP14 (direita): M- Prestained protein ladder 10-170, Fermentas; 1- Sem indução; 2- Extrato total; 3- lavagem1, 5- lavagem 2; 6- Resina+ Proteína, 7- Proteína + resina + ULP; 8- Proteína eluída; 9- Resina pós digestão com ULP. Gel SDS-PAGE 15%. Seta indica proteína esperada.

Focamos o estudo nas proteínas MpCSEP5 e MpCSEP14 (Figura 24) por se mostrarem mais promissoras aos passos de purificação. A partir daqui, indução foi feita em 8 litros de LB, seguindo os passos de purificação mostrados ate agora. Para obtenção da proteína mais pura, adicionamos cromatografia por exclusão de tamanho para retirada de outras moléculas que ainda estão presentes mesmo após a purificação por afinidade. Resultado dessa purificação mostrou a permanência de um contaminante que se mostrava por uma banda imediatamente acima da banda correspondente a proteína de trabalho. Para corrigir isso, um ajuste na purificação por afinidade, com o aumento da

quantidade de resina de afinidade foi adotado. Essa modificação resolveu o problema para MpCSEP5 (Figura 25), mas para MpCSEP14 acrescentamos ainda uma purificação de troca iônica utilizando coluna MonoQ, o que minimizou o problema (Figura 26).

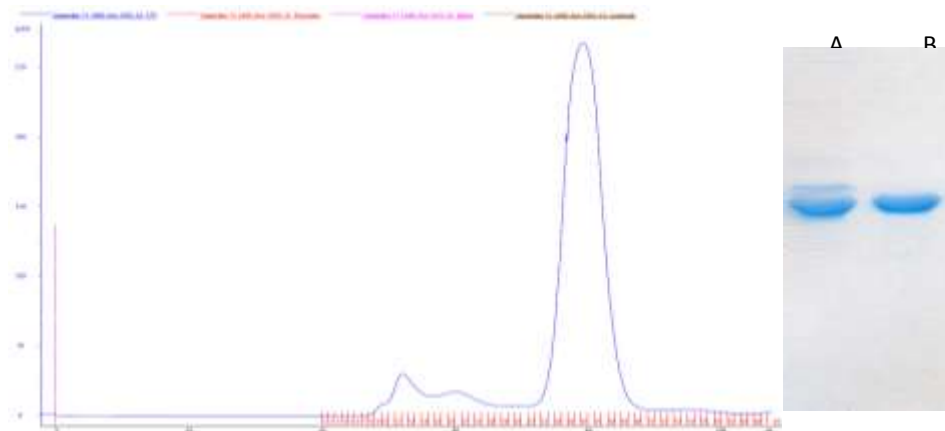


Figura 25: Purificação de MpCSEP5. Curva gerada durante purificação por exclusão de tamanho (Hiload 16/60 Superdex 75) pelo software UNICOR no equipamento AKTA Purifier (esquerda). Gel SDS-PAGE 15%. A: Produto da purificação de MpCSEP5 pelo protocolo original; B- Produto de purificação de MpCSEP5 através do protocolo otimizado.

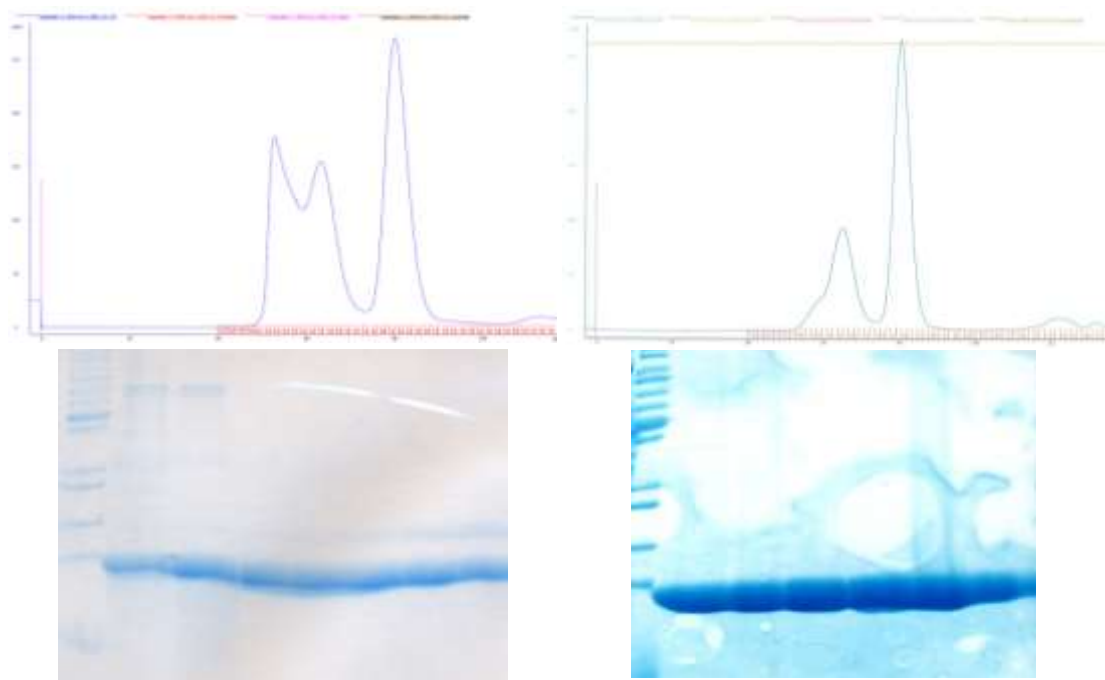


Figura 26: Purificação de MpCSEP14. Curva gerada durante purificação por exclusão de tamanho (Hiload 16/60 Superdex 75) pelo software UNICOR no equipamento AKTA Purifier (acima). Gel SDS-PAGE 15% do produto da purificação da proteína (abaixo). Esquerda: purificação por exclusão de tamanho após afinidade. Direita: purificação com a adição da etapa de purificação por troca iônica (Mono-Q 5/50, GE).

4.4.4- Cristalização

As proteínas puras foram utilizadas para técnicas que visam à formação de cristais para resolução de sua estrutura tridimensional. Foram testados ao todo seis kits iniciais conforme as práticas adotadas pelo laboratório. A concentração inicial para testes de cristalização das proteínas utilizadas foi de 20mg/ml.

MpCSEP5 mostrou-se promissora apresentando cristais em algumas das condições testadas nesses kits, e a partir dessas condições, placas para o refinamento do cristal foram confeccionadas em robô com as variações citadas em materiais e métodos. Foi possível observar uma pequena melhora no aspecto do cristal. Essa proteína foi ainda testada com o uso dos kits *Additive Screen* e *Silver Bullets* (*Hampton Research*) em placas feitas a mão e ainda pela técnica de *seeding* (descrito materiais e métodos).

Utilizamos a técnica de semear núcleos de cristalização (*seeding*) na tentativa de melhorar o perfil anisotrópico dos cristais. Essa técnica consiste em transferir microcristais oriundos a partir de um pequeno cristal triturado (uma semente) para uma gota. Nela, a concentração dos reagentes está em uma condição onde não ocorre nucleação de cristal, somente crescimento (fase metaestável). Para que aconteça o crescimento do cristal é necessário semear núcleos de cristalização. Esta ultima não apresentou resultados satisfatórios. Os cristais formados pelas demais técnicas possuem um perfil oco, com aspecto de “canudos” o que desfavorece a difração (Figura 27). Esses cristais estão sendo submetidos a mais refinamentos para a obtenção de cristais mais robustos.

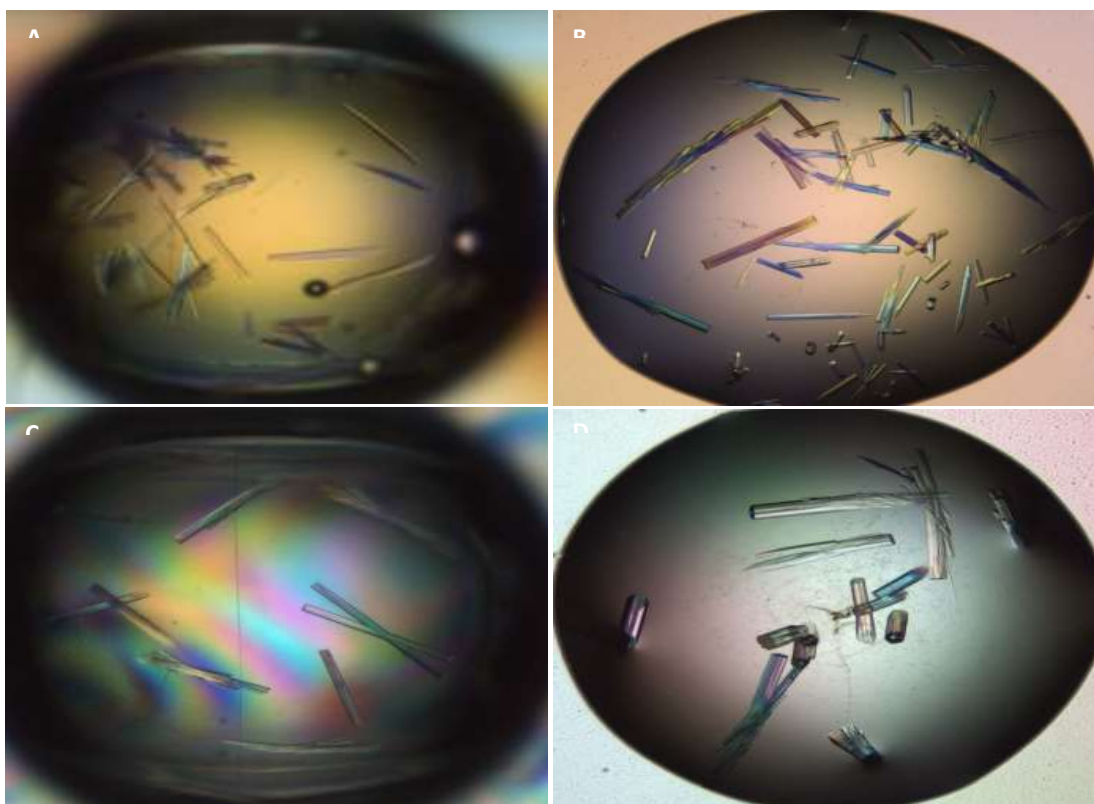


Figura 27: Cristais de MpCSEP5. Imagens geradas pelo equipamento Rock Imager 1000 de visualização de cristais. **(A)** Condição de kit *Crystal Screen I* (Hampton Research). **(B)** Condição de refinamento com kit *Silver Bullet*. **(C)** Condição de refinamento com kit *Additive Screen*. **(D)** Refinamento por variação de condição inicial (pH e concentração de sal).

Tentativas de obtenção de cristais de MpCSEP14 também foram efetuadas com o uso dos mesmos kits iniciais, utilizando a mesma concentração inicial de proteína em solução (20mg/ml), que não gerou nenhum cristal. Testamos as mesmas condições utilizando o dobro de proteína em solução (40mg/ml) e obtivemos formação de alguns cristais, (figura 28) os quais também foram submetidos a variações de pH, concentração de precipitante, e outras modificações para obter cristais de melhor qualidade.

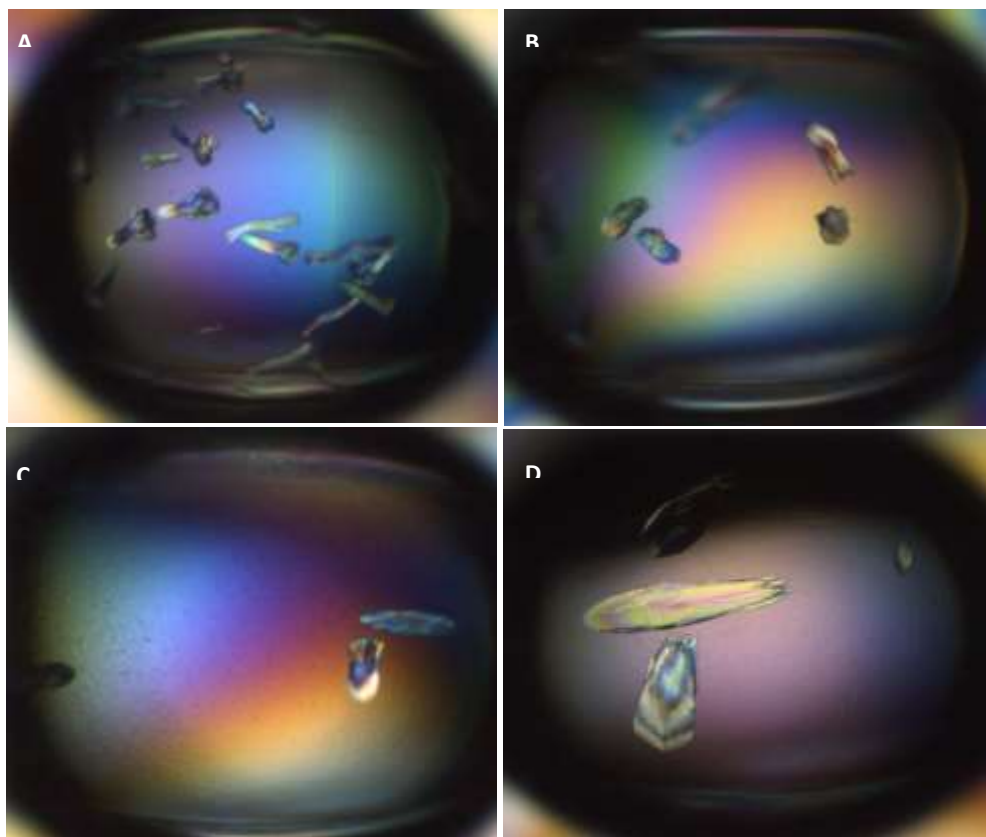


Figura 28: Cristais de MpCSEP14. Imagens geradas pelo equipamento Rock Imager 1000 de visualização de cristais. **(A)** Condição de kit *Crystal Screen* (Hampton Research). **(B)** Condição de refinamento com kit *Silver Bullet*. **(C)** Condição de refinamento com kit *Additive Screen*. **(D)** Refinamento por variação de condição inicial (pH e concentração de sal).

4.4.5- Dicroísmo Circular e Espalhamento de Luz Dinâmico

Experimento de dicroísmo circular (CD) de MpCSEP5 gerou uma curva de absorvância, que mostra que a proteína apresenta predominantemente folhas β em sua estrutura (mínimo em 210 nm), enquanto a curva gerada para MpCSEP14 indicam que sua estrutura secundária é predominantemente α hélice (figura 29). Dados de Espalhamento de luz Dinâmico (DLS) indicaram que MpCSEP5 em solução de trabalho apresenta um peso molecular aproximado de 21 KDa (figura 30) e aproximadamente 29KDa para MpCSEP5. As sequências das proteínas recombinantes indicam um peso molecular de 15KDa para MpCSEP5 e 13KDa para MpCSEP14. Analisando esses

dados, podemos dizer que na solução de trabalho, MpCSEP5 está em sua forma predominantemente monomérica na solução de trabalho a 25°C, enquanto MpCSEP14 apresenta-se em forma de dímero nas mesmas condições.

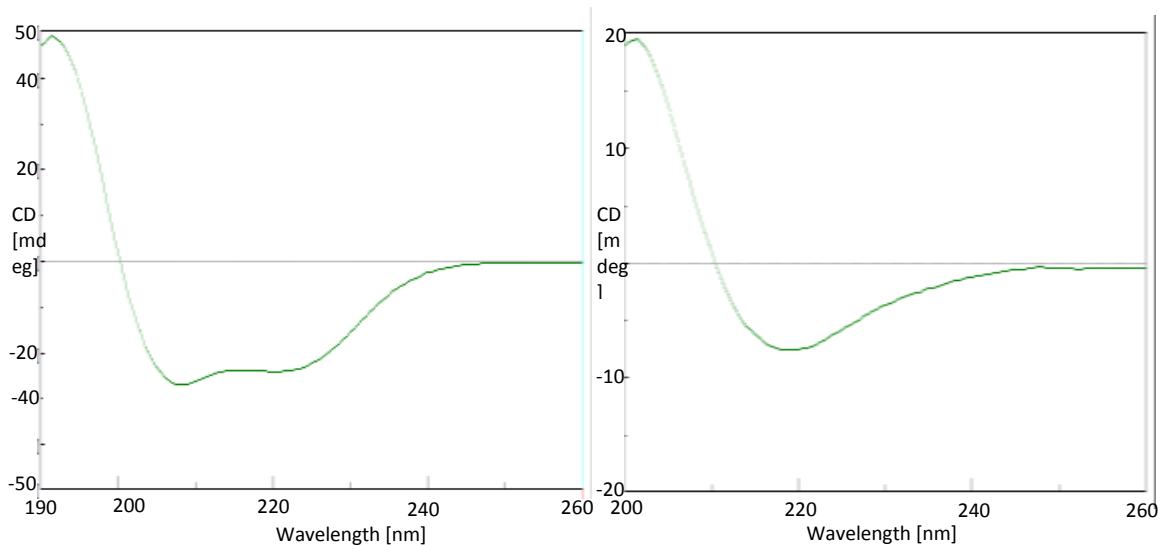


Figura 29: Dicroísmo circular. MpCSEP5 (esquerda) e MpCSEP14 (direita).

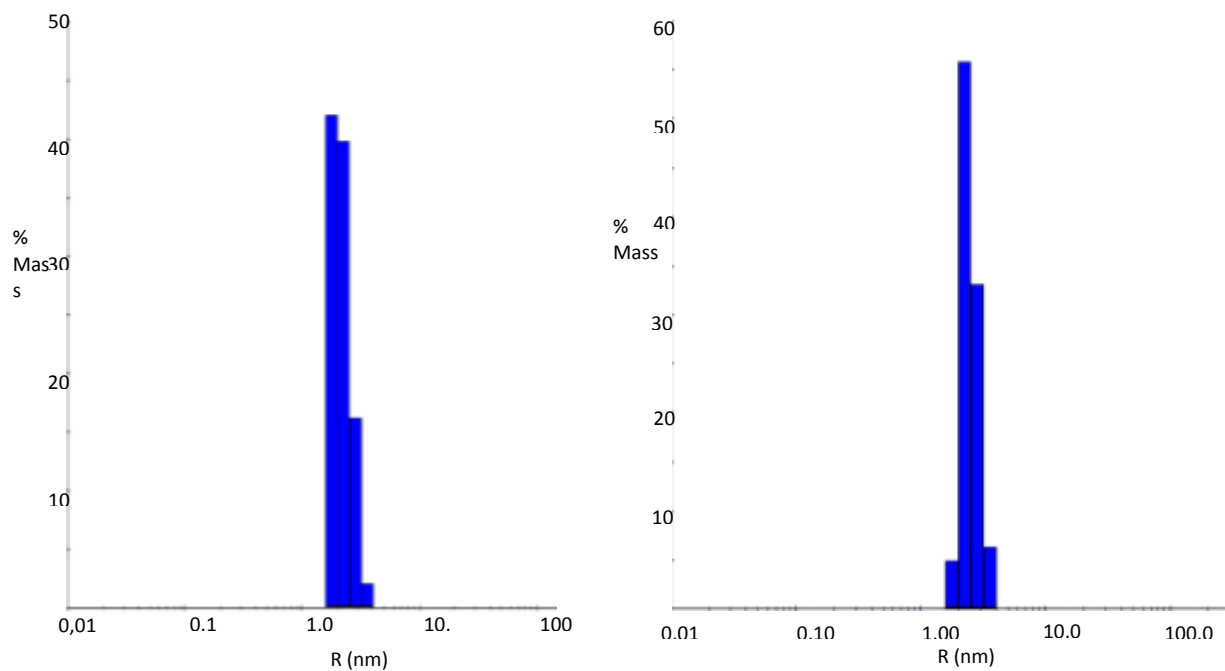


Figura 30: Espalhamento de luz dinâmico. MpCSEP5 (esquerda) e MpCSEP14 (direita).

4.5- Discussão

Muitos fungos fitopatógenos são capazes de reprogramar o metabolismo celular e a defesa de plantas pela secreção de proteínas pequenas conhecidas como efetores. Considerando que ocorre uma interação compatível entre *M. pernicioso* e cacau, podemos assumir que a resposta de defesa da planta não foi eficiente para prevenir o estabelecimento do micélio biotrófico de *M. pernicioso* na planta. De acordo com a definição proposta por Jones e Dangl (2006), a resposta de defesa verificada em vassoura verde pode ser classificada como “defesa basal”, resposta de defesa que é inativada por patógenos virulentos em hospedeiros susceptíveis. Esse tipo de defesa é geralmente consequência de proteínas efetoras produzidas pelo patógeno com o objetivo de neutralizar o sistema imune do hospedeiro. Embora nenhuma proteína efetora tenha sido caracterizada em *M. pernicioso* até o momento, análise de RNA-seq aponta alguns candidatos a efetores que são altamente e em sua maioria exclusivamente expressos em vassoura verde (Teixeira *et al.*, 2013). Esses genes se enquadram na descrição de efetores de patogenicidade de Stergiopoulos e de Wit (2009) são proteínas pequenas, secretadas e em sua maioria possuem um alto índice de cisteínas. Esses genes estão entre os mais expressos durante a interação *T. cacao*- *M.pernicioso*. (Teixeira *et al.*, 2013). A alta expressão desses genes durante a interação planta patógeno, somados ao fato de serem genes desconhecidos, o fracasso do sistema de defesa da planta e as características das proteínas codificadas por eles tornaram esses genes alvos para nosso estudo. Essas proteínas foram incluídas no último ano do projeto “SMOLBNET 2.0:

Estudos estruturais de proteínas chave para as doenças fúngicas do cacau Vassoura de bruxa e Monilíase – Desenvolvimento de estratégias de controle e entendimento de modelos de patogenicidade”. Por sua potencial importância no sucesso de infecção do fungo, e pela descontinuidade da caracterização de MpTLPs no projeto devido a ausência de proteínas solúveis. A inclusão dessas proteínas a esse projeto fez-

se importante por se tratar de uma troca de aprendizado e acesso à estrutura do laboratório LNBIO. As etapas de análise de genes, desenho de primers clonagem e sequenciamento para confirmação de clones foram feitas no Laboratório de Genômica e Expressão (UNICAMP), enquanto a parte de expressão, purificação e cristalização foram feitas no Laboratório Nacional de Biociências (LNBio- CNPEM). A cristalografia foi utilizada como uma ferramenta de estudo na tentativa de obtenção das estruturas de MpCSEPs o que poderia facilitar na descoberta da função biológica dessas proteínas. Os cristais obtidos até o momento foram testados em linha de luz MX1 e MX2, mas não apresentou difração e os processos de refinamento deverão ser continuados.

A tabela a seguir sumariza o trabalho o trabalho realizado com as MpCSEPs.

Tabela 5: Resumo das etapas que foram seguidas nesse capítulo. Em vermelho, as etapas limitantes para cada proteína; em verde as etapas concluídas e em azul estão as etapas que ainda podem ser abordadas em trabalhos futuros.

Nome do gene	Seleção por n° de cisteína	Amplificação	Clonagem	Expressão	Purificação	Cristalização	Resolução estrutural	Teste atividade
<i>MpCSEP1</i>								
<i>MpCSEP2</i>								
<i>MpCSEP3</i>								
<i>MpCSEP4</i>								
<i>MpCSEP5</i>								
<i>MpCSEP6</i>								
<i>MpCSEP7</i>								
<i>MpCSEP8*</i>								
<i>MpCSEP9</i>								
<i>MpCSEP10</i>								
<i>MpCSEP11</i>								
<i>MpCSEP12</i>								
<i>MpCSEP13</i>								
<i>MpCSEP14</i>								
<i>MpCSEP15</i>								
<i>MpCSEP16</i>								

4.6-Perspectivas

Apesar de os estudos das proteínas MpCSEP 1, 9, 13 e 16 terem sido descontinuados no presente trabalho, os estudos das mesmas serão retomados com ajustes na expressão e purificação dessas proteínas. Estas proteínas foram excluídas do presente estudo apenas por uma questão de fluxo de trabalho, e não está relacionado a impossibilidade de trabalhar com as mesmas.

Paralelamente, esforços estão sendo direcionados para testes de atividades das proteínas. Primeiramente, protocolos de medição de ROS (espécies reativas de oxigênio) através de medição de fluorescência em discos foliares estão sendo padronizados. A produção ROS pelos discos foliares pode ser indicativa de que o sistema imune da planta estaria ativo. Sua medição, portanto, pode ser importante para avaliarmos a diferença de liberação de ROS em discos foliares em controles tratados com flagelina, tratamento de discos com a MpCSEPs e tratamento de discos com controle (flagelina, utilizado como elicitador) somado a adição de MpSCEPs. A medição de liberação de ROS é medida em fluorímetro (Nicaise *et al.*, 2013), na presença de luminol e peróxido de hidrogênio. Esse teste pode nos dar indícios de que as MpCSEPs atuam na proteção ou até mesmo no ataque celular, caso haja diminuição ou aumento da fluorescência na presença ou não de proteínas indutoras de defesa de plantas, indicado pela diminuição ou aumento de ROS produzidos pelo tecido vegetal. Até o momento, não foi possível estabelecer essa técnica em nosso laboratório, visto que ainda não conseguimos padronizar os controles. Outros testes que estão sendo preparados são testes “*in planta*” de atividades de MpCSEPs. Um experimento mais detalhado nos permitiria medir a diferença na expressão gênica de genes marcadores de defesa de plantas: glucanase, gene da via do etileno ACS sintase, quitinase, peroxidase, fator de transcrição PR1 (NPR1) e gene PR1. Plântulas de cacau saudáveis seriam desafiadas com um controle (flagelina), MpCSEPs e uma combinação de controle somado a Flagelina. Esses genes foram identificados em *T. cacao* e primers para

Real Time-PCR foram desenhados. Medições de expressão de genes marcadores para defesa de planta poderiam nos indicar se está havendo ou não uma repressão da defesa de planta. Pretendemos continuar o estudo com as proteínas MpCSEPs através de orientações de novos alunos de Iniciação Científica ou Mestrado, para dar continuidade aos testes de atividade das proteínas, bem como a resolução estrutural das mesmas. A resolução da estrutura é uma ferramenta que poderá nos orientar a possíveis ligantes e inferir possíveis atividades dessas proteínas.

5- Referências bibliográficas:

- Agrios G N.** *Plant pathology*. San Diego, Academic Press, 1997, 635
- Aime MC, Phillips-Mora W.** The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia*. 2005; 97:1012-22
- Bartoszewski G, Niedziela A, Szwacka M, Niemirowicz-Szczytt K.** Modification of tomato taste in transgenic plants carrying a thaumatin gene from *Thaumatococcus daniellii* Benth. *Plant Breed*.2003. 122: 347–351.
- Bozkurt, T. O., Schornack, S., Banfield, M. J. & Kamoun, S.** Oomycetes, effectors, and all that jazz. *Curr. Opin. Plant Biol.* **15**, 483–492 (2012)
- Bastos CN.** Avaliação de Fungicidas Sistêmicos no Controle da Vassoura-de-Bruxa do Cacauiero. *Agrotropica*. 1989; 1: 128-132.
- Bastos C.N. and H.C. Evans,** A New Pathotype of *Crinipellis-Perniciosa* (Witches Broom Disease) on Solanaceous Hosts. *Plant Pathology*, 1985. 34(2): p. 306-312.
- Beliën T, Van Campenhout S, Robben J, Volckaert G.** Microbial endoxylanases: effective weapons to breach the plant cell-wall barrier or, rather, triggers of plant defense systems? *Molecular Plant Microbe Interactions*. 2006 19:1072-1081.
- Benko-Iseppon A M, Winter P, Huettel B, Staginnus C, Muehlbauer FJ, Kahl G.** Molecular markers closely linked to *Fusarium* resistance genes in chickpea show significant alignments to pathogenesis-related genes located on Arabidopsis chromosomes 1 and 5. *Theoretical and Applied Genetics*. 2003; 107: 379–386.
- Bennett, S.** 2004. Solexa Ltd. *Pharmacogenomics* (5): 433-438.
- Brandazza A. Angeli S. Tegoni M. Cambillau C. Pelos P.** Plant stress proteins of the thaumatin-like family discovered in animals. *FEBS Letters*. 2004; 572: 3–7.
- Brito N, Espino JJ, González C.** The endo-beta-1,4-xylanase xyn11A is required for virulence in *Botrytis cinerea*. *Molecular Plant Microbe Interactions*. 2006; 19:25-32.
- Caspers MP, Lok F, Sinjorgo KM, van Zeijl MJ, Nielsen KA, Cameron-Mills V.** Synthesis, processing and export of cytoplasmic endo-beta-1,4-xylanase from barley aleurone during germination. *Plant Journal*. 2001; 26:191-204.
- Ceita GO, Macedo JNA, Santos TB , Alemanno L, Gesteira AS, Micheli F, Mariano AC, Gramacho KP, Silva DC, Meinhardt LW, Pereira GAG, Cascado JCM.** Involvement of calcium oxalate degradation during programmed cell death in *Theobroma cacao* tissues triggered by the hemibiotrophic fungus *Moniliophthora perniciosa*. *Plant Science*. 2007; 173. 106 -117.
- Chaves, F.C. and T.J. Gianfagna,** *Necrotrophic phase of Monihophthora perniciosa causes salicylic acid accumulation in infected stems of cacao*. *Physiological and Molecular Plant Pathology*, 2006. 69(1-3): p. 104-108.
- Colditz F, Niehaus K, Krajinski F.** Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins and in an increased tolerance upon infection with the oomycete *Aphanomyces utoiches*. *Planta*. 2007; 226:57–71
- Collins T, Gerday C, Feller G.** Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews*. 2005; 29:3-23.
- Christensen AB, Cho BH, Naesby M, Gregersen PL, Brandt J, Madri-Ordena~na K, Collinge DB, Thordal-Christensen H.** The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins. *Mol Plant Pathol* (2002) 3:135–144
- de Jonge, R., Bolton, M. D. & Thomma, B. P. H. J.** How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. *Curr. Opin. Plant Biol.* **14**, 400–406 (2011).
- Dabydeen, S. and T.N. Sreenivasan,** *Occurrence, Anatomy, and Morphology of Blisters on Cocoa Seedlings Inoculated with Crinipellis-Perniciosa*. *Plant Disease*, 1989. **73**(6): p. 481-484.
- Dangl JL, Jones JD.** Plant pathogens and integrated defence responses to infection. *Nature*. 2001 Jun 14:826-33.
- Daniell S, Mellits KH, Faus I, Connerton I.** Refolding the sweettasting protein thaumatin II from insoluble inclusion bodies synthesized in *Escherichia coli*. *Food Chemistry*. 2000; 7: 105–

110.

Datta K, Velazhahan R, Oliva N, Onal M, Mew T, Kush G S, Muthukrishnan S, Datta S K. Over-expression of the cloned rice thaumatin-like protein (PR-5) gene in transgenic rice plants enhances environmental friendly resistance to *Rhizoctonia solani* causing sheath blight disease. *Theor. Appl. Genet.* 1999; 98:1138–1145

Deacon, L.J., et al., *Diversity and function of decomposer fungi from a grassland soil.* *Soil Biology & Biochemistry*, 2006. **38**(1): p. 7-20.

Djamei, A. & Kahmann, R. *Ustilago maydis*: dissecting the molecular interface between pathogen and plant. *PLoS Pathog.* **8**, e1002955 (2012).

Evans HC e Bastos CN. Uma Avaliação do ciclo de vida da Vassoura-de-bruxa (*Crinipellis pernicioso*) do Cacaueiro. *Fitopatologia Brasileira*. 1979; 4: p. 104.

Evans HC e Prior C. Cocoa pod diseases: Causal agents and control. *Outlook on Agriculture*. 1987; 16: 35-41.

Evans, H. C. Pod rot of cacao caused by *Moniliophthora (Monilia) roreri*. *Phytopathological Papers* 1981. **24**:1-44

Evans, H. C., Stalpers, J. A., Sampson, R. A., and Benny, G. L. (). On the taxonomy of *Monilia roreri*, an important pathogen of *Theobroma cacao* in South America. *Canadian Journal of Botany*. 1978. **56**:2528-2532

Evans, H.C. Cacao diseases - the trilogy revisited. 2007. *Phytopathology* **96**:S138

Felix G, Duram JD, Volko S, Boller T. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant Journal*. 18 (3):265:76. 1999

Fierens E, Gebruers K, Courtin CM, Delcour JA. Xylanase inhibitors bind to nonstarch polysaccharides. *Journal of Agricultural and Food Chemistry*. 2008; 23:564-70

Fierens E, Rombouts S, Gebruers K, Goesaert H, Brijs K, Beaugrand J, Volckaert G, Van Campenhout S, Proost P, Courtin CM, Delcour JA. TLXI, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family. *Biochemical Journal*. 2007; 403:583-91.

Fagoaga C, Rodrigo I, Conejero V, Hinarejos C, Tuset J.J., Arnau J, Pina J.A, Navarro L, Pena L. Increased tolerance to *Phytophthora citrophthora* in transgenic orange plants constitutively expressing a tomato pathogenesis related protein PR-5. *Mol Breeding*. 2001. 7:175–185.

Franco SF, Teixeira PJ, Thomazella DPT, Medrano FJ, Mondego JMC, Pereira, GAG, Caracterização de um gene similar a Taumatina expresso pelo fungo causador da Vassoura de Bruxa, *Moniliophthora pernicioso*. In *CD-ROM dos trabalhos apresentados durante o 54º Congresso Brasileiro de Genética*, 2008.

Fulton, R. H. The cacao disease trilogy: black pod, monilia pod rot, and witches'-broom. *Plant Disease* 1989, **73**:601-603.

Garcia O, Macedo JN, Zapparoli G, Tiburcio R, Micheli F, Gesteira A, Meinhardt LW, Rincones J, Alvim F, Schiavinato M A, Pereira GAG, Cascardo JCM. Identification and characterization of novel Necrosis and Ethylene inducing Proteins (NEPs) in the hemibiotrophic basidiomycete *Crinipellis pernicioso*, the causal agent of witches' broom of cocoa. *Mycological Research*. 2007; 111: 443-55.

Gesteira, A.S., et al., *Comparative analysis of expressed genes from cacao meristems infected by Moniliophthora pernicioso.* *Ann Bot*, 2007. 100(1): p. 129-40.

Gebruers K, Debyser W, Goesaert H, Proost P, Van Damme J, Delcour JA. *Triticum aestivum* L. endoxylanase inhibitor (TAXI) consists of two inhibitors, TAXI I and TAXI II, with different specificities. *Biochemistry Journal*. 2001 353:239-244.

Ghosh R, Chakrabarti C. Crystallization and preliminary X-ray diffraction studies of NP24-I, an isoform of a thaumatin-like protein from ripe tomato fruits. *Acta Crystallogr Section F Structural Biology and Crystallization Communications*. 2005; 61:806-7.

Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, Thordal-Christensen H. Powdery_mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics*. 2010 May 20;11:31

Godoy, G., et al., *Use of Mutants to Demonstrate the Role of Oxalic-Acid in Pathogenicity of Sclerotinia-Sclerotiorum on Phaseolus-Vulgaris.* *Physiological and Molecular Plant Pathology*, 1990. **37**(3): p. 179-191.

Greenstein S, Shadkchan Y, Jadoun J, Sharon C, Markovich S, Osherov N. Analysis of the *Aspergillus nidulans* thaumatin-like *cetA* gene and evidence for transcriptional repression of *pyr4* expression in the *cetA*-disrupted strain. *Fungal Genetics and Biology*. 2006; 43:42-53.

Grenier J, Potvin C, Trudel J, Asselin A. Some thaumatin-like proteins hydrolyse polymeric β -1,3-glucans. *Plant Journal*, 1999; 19: 473-480.

Hagiwara A, Yoshino H, Sano M, Kawabe M, Tamano S, Sakaue K, Nakamura M, Tada M, Imaida K, Shirai T. Thirteen-week feeding study of thaumatin (a natural proteinaceous sweetener), sterilized by electron beam irradiation, in *Sprague-Dawley* rats. *Food and Chemical Toxicology*. 2005; 43:1297-302.

Hedger JN, Pickering V, Aragundi J. Variability of populations of the witch's broom disease of cocoa *Crinipellis perniciosa*. *Transactions of the British Mycological Society*. 1987; 88:533-546.

Hensel, L.L., et al., Developmental and Age-Related Processes That Influence the Longevity and Senescence of Photosynthetic Tissues in *Arabidopsis*. *Plant Cell*, 1993. 5(5): p. 553-564.

Hiilovaara-Teijo M, Hannukkala A, Griffith M, Yu XM, Pihakaski-Maunsbach K, Snow-mold- induce apoplastc proteins in winter rye leaves lack antifreeze activity. *Plant Physiology*. 1999; 121: 665-674.

Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R. & Kamoun, S. Emerging concepts in effector biology of plant-associated organisms. *Mol. Plant Microbe Interact*. **22**, 115–122 (2009)

Hu X, Reddy AS. Cloning and expression of a PR5-like protein from *Arabidopsis*: inhibition of fungal growth by bacterially expressed protein. *Plant Molecular Biology*. 1997; 34:949–959.

Jones, Jonathan D G, and Jeffery L Dangl. 2006. The plant immune system. *Nature* 444, no. 7117 (November): 323-9. doi:10.1038/nature05286.

Jung YC, Lee HJ, Yum SS, Soh WY, Cho DY, Auh CK, Lee TK, Soh HC, Kim YS, Lee SC. Drought-inducible-but ABA-independent-thaumatin-like protein from carrot (*Daucus carota* L.). *Plant Cell Reports*. 2005; 24:366-73.

Kale SD, Gu B, Capelluto DG, et al. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell*. 2010;142:284–295

Kalpna K, Maruthasalam S, Rajesh T, Poovannan K, Kumar K K, Kokiladevi E, Raja J A J, Sudhakar D, Velazhahan R, Samiyappan R, Balasubramanian P. Engineering sheath blight resistance in elite indica rice cultivars using genes encoding defense proteins. *Plant Science*. 2005. 117. 203-215.

Kaneko R, Kitabatake N. Structure-sweetness relationship in thaumatin: importance of lysine residues. *Chemical Senses*. 2001; 26:167-77.

Kemen E, Gardiner A, Schultz-Larsen T, Kemen A C, Balmuth A L, Robert-Seilaniantz A, Bailey K, Holub E B, Studholme, D J, MacLean D and Jones J D. (2011) *Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of Arabidopsis thaliana*. *PLoS Biology*, Vol.9 (No.7).

Kilaru, A., B.A. Bailey, and K.H. Hasenstein, *Moniliophthora perniciosa* produces hormones and alters endogenous auxin and salicylic acid in infected cocoa leaves. *FEMS Microbiol Lett*, 2007. 274(2): p. 238-44.

Knogge W. Fungal Infection of plants. *The Plant Cell*. 1996; 8:1711-1722.

Lam SK, Ng TB. First simultaneous isolation of a ribosome inactivating protein and an antifungal protein from a mushroom (*Lyophyllum shimeji*) together with evidence of synergism of their antifungal effects. *Archives of biochemistry and biophysics*. 2001; 171-280.

Leone P, Menu-Bouaouiche L, Peumans WJ, Payan F, Barre A, Roussel A, Van Damme EJ, Rougé P. Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å. *Biochimie*. 2006; 88:45-52

Lima EL, Barros EV, Araújo JG, Thiébaud JT. Frequência da Poda Fitossanitária no Controle da Vassoura-de-bruxa do Cacaueiro. *Agrotropica*, 1991; 3: 75-80.

Li, D., et al., A repressor complex governs the integration of flowering signals in *Arabidopsis*. *Dev Cell*, 2008. 15(1): p. 110-20.

Liu J J, Sturrock R, Ekramoddoullah A K. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant Cell Rep*. 2010 ;29(5):419-36.

Mackintosh C A, Lewis J, Radmer L E, Shin S, Heinen S J, Smith L A, Wyckoff M N, Dill-Macky R, Evans C K, Kravchenko S, Baldridge G D , Zeyen R J, Muehlbauer G J. Overexpression of defense response genes in transgenic wheat enhances resistance to *Fusarium* head blight. *Plant Cell Rep.* 2007. 26: 479–488.

Manenoi A, Paull RE. Papaya fruit softening, endoxylanase gene expression, protein and activity. *Plant Physiology.* 2007; 131:470-80

Marathe R, Dinesh-Kumar SP. Plant defense: one post, multiple guards?! *Molecular Cell.* 2003 Feb;11:284-6.

Maria S., Magdalena K., Anita O., Magdalena E. K. e Stefan M. Variable properties of transgenic cucumber plants containing the thaumatin II gene from *Thaumatococcus daniellii*. *Acta Physiology Plantarum*, 2002. 24, 173-185.

Masuda T, Tamaki S. Kaneko R, Wada R, Fujita Y, Mehta A, Kitabatake N. Cloning, expression and characterization of recombinant sweet-protein thaumatin II using the methylotrophic yeast *Pichia pastoris*. *Biotechnology and Bioengineering* 2004; 30:761-9.

Meinhardt LW, Bellato CM, Rincones J, Azevedo RA, Pereira GAG. Development of an in vitro germination and growth method for the biotrophic phase of *Crinipellis pernicioso*, the causal agent of witches' broom disease of *Theobroma cacao*. *Current Microbiology*, 2006; 52: 191- 196

Mondego JMC, Carazzolle M F, Costa GGL, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, García O, Thomazella DPT, Oliveira BV, Pires ABL, Rio MCS, Araújo MRR, Moraes MH, Castro LAB, Gramacho KP , Gonçalves MS, Moura Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JCM , Pereira GAG . A genome survey of *Moniliophthora Perrnicioso* gives new insights into Witches' Broom Disease of cacao. *BMC Genomics* 2008, 9:548.

Narasimhan ML, Coca MA. Jin J, Yamauchi T, Ito Y, Kadowaki T, Kim KK, Pardo JM, Damsz B, Hasegawa PM, Yun DJ. Bressan RA. Osmotin is a homolog of mammalian adiponectin and controls apoptosis in yeast through a homolog of mammalian adiponectin receptor. *Molecular Cell.* 2005; 17: 171-180.

Nicaise V, Joe A, Jeong BR, Korneli C, Boutrot F, Westedt I, Staiger D, Alfano JR, Zipfel C.

Pseudomonas HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. EMBO J. 2013 Feb 12;32(5):701-12.

Nimchuk Z, Eulgem T, Holt BF 3rd, Dangl JL. Recognition and response in the plant immune system. *Annual Review of Genetics* . 2003; 37:579-609.

Pedersen C, Ver Loren van Themaat E, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu X, Maekawa T, Weßling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD. Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics.* 2012 Dec 11;13:694

Pereira JL, Ram A, Figueiredo J M, de Almeida L C. La primera aparición de la "escoba de bruja" en la principal área productora de cacao del Brasil. *Turrialba.* 1989: 459-461.

Pinheiro MM, Sandroni M, Lummerzheim M, Oliveira DE. A defesa das plantas contra as doenças.. *Ciência Hoje* 147. *Universidade Federal do Rio de Janeiro.* 1999. Disponível em: <http://cienciahoje.uol.com.br/materia/resources/files/chmais/pass/ch147/planta.rtf>

Pires, J. L. L., E. D. M. N. Resistência à Vassoura-de-Bruja na espécie *Theobroma cacao*. *Fitopatologia Brasileira.* 1995, 20: 303.

Pirovani P C, Cascardo JCM and Pereira GAG. Proteínas do fluido apoplástico. In *Workshop da Vassoura-de-bruxa*. Ilhéus: UESC and UNICAMP. 2005

Pritsch C. Muehlbauer GJ. Bushnell WR. Somers DA. Vance CP. Fungal development and induction of defense response genes during early infection of wheat spikes by *Fusarium graminearum*. *Molecular Plant-Microbe Interactions.* 2000; 13: 159-69.

Prusty, R., P. Grisafi, and G.R. Fink, *The plant hormone indoleacetic acid induces invasive growth in Saccharomyces cerevisiae.* *Proc Natl Acad Sci U S A*, 2004. 101(12): p. 4153-7.

Purdy LH. Schmidt RA. Status of cacao witches' broom: biology, epidemiology, and management. *Annual review of phytopathology*, 1996; 34: 573-94

Quirino, B.F., et al., *Molecular aspects of leaf senescence*. Trends in Plant Science, 2000. 5(7): p. 278-282.

Robinette, D. and A.G. Matthysse, *Inhibition by Agrobacterium-Tumefaciens and Pseudomonas-Savastanoi of Development of the Hypersensitive Response Elicited by Pseudomonas-Syringae Pv Phaseolicola*. Journal of Bacteriology, 1990. 172(10): p. 5742-5749.

Rafiqi, M., Ellis, J. G., Ludowici, V. A., Hardham, A. R. & Dodds, P. N. Challenges and progress towards understanding the role of effectors in plant-fungal interactions. *Curr. Opin. Plant Biol.* 15, 477–482 (2012).

Ram, A.; Valle, R. R.; Gardini, E. A. Monília do cacaueiro. São Paulo: Fundação Cargill, 2004. 36.

Rajam M V, Chandola N, Goud P S, Singh D, Kashyap V, Choudhary M L, Sihachakr D. Thaumatin gene confers resistance to fungal pathogens as well as tolerance to abiotic stresses in transgenic tobacco plants. *Biol.Plant.* 2007. 51:135–141.

Rincones J, Mazotti GD, Griffith GW, Pomela A, Figueira A, Leal Jr GA, Queiroz MV, Pereira JF, Azevedo RA, Pereira GAG, Meinhardt LW, Variabilidade Genética e Polimorfismos Cromossômicos de *C. pernicioso* Isolado de Diversos Hospedeiros na América do Sul. *Mycological Research*. 2006; 110:821-32.

Rincones J, Meinhardt LW, Vidal BC, Pereira GAG. Electrophoretic karyotype analysis of *Crinipellis pernicioso*, the causal agent of witches' broom disease of *Theobroma cacao*. *Mycological Research*. 2003; 107: 452-458

Rincones J, Scarpari LM, Carazzolle M, Mondego JMC, Formighieri E, Barau JG, Costa GGL, Carraro DM, Brentani HP, Vilas-Boas L, Oliveira B V, Sabha M, DIAS R, Cascardo J, Azevedo RA. Differential gene expression between the biotrophic-like and the saprotrophic mycelia of the Witches' broom pathogen *Moniliophthora pernicioso*. *Molecular Plant-Microbe Interactions*. 2008; 21:891-908

Rio MC, Oliveira BV, Thomazella DP, Silva JA, Pereira GA. Production of calcium oxalate crystals by the basidiomycete *Moniliophthora pernicioso*, the causal agent of witches' broom disease of Cacao. *Current Microbiology*. 2008; 56:363-70

Rombouts S, Fierens E, Vandermarliere E, Voet A, Gebruers K, Beaugrand J, Courtin CM, Delcour JA, de Maeyer M, Rabijns A, Van Campenhout S, Volckaert G. His22 of TLXI plays a critical role in the inhibition of glycoside hydrolase family 11 xylanases. *Journal of Enzyme Inhibition and Medicinal Chemistry*. 2008; 12:1.

Ruperti B. Cattivelli L. Pagni S. Ramina A. Ethylene-responsive genes are differentially regulated during abscission, organ senescence and wounding in peach (*Prunus persica*). *Journal of Experimental Botany*. 2002; 53: 429-37.

Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD. Systemic Acquired Resistance. *Plant Cell*. 1996; 8:1809-1819.

Sakamoto Y, Watanabe H, Nagai M, Nakade K, Takahashi M, Sato T. *Lentinula edodes tlg1* encodes a thaumatin-like protein that is involved in lentinan degradation and fruiting body senescence. *Plant Physiology*. 2006; 141(2):793-801.

Sambrook JRDW. Molecular cloning: a laboratory manual. New York, U.S.A.: 2001; Cold Spring Harbor Laboratory, Cold Spring Harbor.

Santos, R.X., et al., Carbon source-dependent variation of acquired mutagen resistance of *Moniliophthora pernicioso*: similarities in natural and artificial systems. *Fungal Genet Biol*, 2008. 45(6): p. 851-60.

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GAG. Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis pernicioso*. *Journal of Experimental Botany*, 2005; 56: 865-77

Schestibratov K A, Dolgov S V. Transgenic strawberry plants expressing a thaumatin II gene demonstrate enhanced resistance to *Botrytis cinerea*. *Sci. Hortic.* 2005. 106:177–189

Silva SDVM. e Matsuoka K. Histologia da Interação *Crinipellis pernicioso* em Cacaueiros Suscetível e Resistente à Vassoura-de-Bruxa. *Fitopatologia Brasileira*. 1999. 24: p. 54-59.

Stergiopoulos I, de Wit PJ. Fungal effector proteins. *Annual Review of Phytopathology Vol. 47: 233-263, 2009.*

Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M, Fritig B. Plant 'pathogenesis-related' proteins and their role in defense against pathogens. *Biochimie*. 1993. 75(8): p.687-706.

Suzuki M, Kato A, Nagata N, Komeda Y. A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. *Plant Cell Physiology*. 2002; 43:759-67

Taiz, L., Fisiologia vegetal. 2004.

Teixeira, PJPLT. Construção de um atlas transcriptômico para a doença da Vassoura de Bruxa. Tese de doutorado. UNICAMP, 2013.

Teixeira, PJPLT; Thomazella, DPT; Reis, O; Prado, PFV; Rio, MC; Costa GGL; Vidal, RO; Mondego, JMC; Piotr, M; Pereira, GAG. High-resolution transcript profiling of the biotrophic interaction between *Theobroma cacao* and the fungal pathogen *Moniliophthora perniciosa*. Artigo submetido à Plant Cell, 2013.

Teixeira, PJPLT; Thomazella; Vidal, RO; Prado, PFV; Reis, O, Baroni, RM; Franco SF; JMC; Piotr, Pereira, GAG; Mondego, JMC. The Fungal Pathogen *Moniliophthora perniciosa* Has Genes Similar to Plant PR-1 That Are Highly Expressed during Its Interaction with Cacao. [PLoS One](https://doi.org/10.1371/journal.pone.0145929). 2012;7(9):e45929.

Thomma, B. P., Nurnberger, T. & Joosten, M. H. Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell* 23, 4–15 (2011).

Thomazella DP, Teixeira PJ, Oliveira HC, Saviani EE, Rincones J, Toni IM, Reis O, Garcia O, Meinhardt LW, Salgado I, Pereira GA. **The hemibiotrophic cacao pathogen *Moniliophthora perniciosa* depends on a mitochondrial alternative oxidase for biotrophic development.** *New Phytologist*, 2012.

Van Damme EJ, Charels D, Menu-Bouaouiche L, Proost P, Barre A, Rouge P, Peumans WJ. Biochemical, molecular and structural analysis of multiple thaumatin-like proteins from the elderberry tree (*Sambucus nigra* L.). *Planta*. 2002; 214: 853-62.

Van den Ackerveken GF, Van Kan JA, Joosten MH, Muisers JM, Verbakel HM, De Wit PJ. Characterization of two putative pathogenicity genes of the fungal tomato pathogen *Cladosporium fulvum*. *Molecular Plant-Microbe Interactions*. 1993; 6:210-5.

van Loon, L.C., Rep, M. & Pieterse, C.M. Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol*, 2006. 44: p. 135-162.

Wang H, Ng TB. Eryngin, a novel antifungal peptide from fruiting bodies of the edible mushroom *Pleurotus eryngii*. *Peptides*. 2004; 25: 1-5.

Wang L, Duman JG. A thaumatin-like protein from larvae of the beetle *Dendroides canadensis* enhances the activity of antifreeze proteins. *Biochemistry*. 2006; 31;45: 1278-1284.

Witty M ,Harvey W J. Sensory evaluation of transgenic *Solanum tuberosum* producing r-thaumatin II. *New Zealand Journal of Crop and Horticultural Science*.1990; 18:77–80

Ye XY, Wang HX, Ng TB. First Chromatographic Isolation of an Antifungal Thaumatin-like Protein from *French Bean* Legumes and Demonstration of Its Antifungal Activity, *Biochemical and Biophysical Research Communications*. 1999; 263, 130–134,

York WS, Qin Q, Rose JK. Proteinaceous inhibitors of endo-beta-glucanases. *Biochimica et Biophysica Acta* . 2004; 12;1696:223-33

Yu XM, Griffith M. Antifreeze proteins in winter rye leaves form oligomeric complexes. *Plant Physiology*. 1999; 119: 1361-70.

Zaparoli G, Garcia O, Medrano FJ, Tiburcio R, Costa GL, Pereira GAG: Identification of a second family of genes in *Moniliophthora perniciosa*, the causal agent of Witches' Broom disease in cacao, encoding necrosis inducing proteins similar to cerato platanins. *Mycological Research*. 2009.113:61-72.

Zipfel, Cyril. 2008. Pattern-recognition receptors in plant innate immunity. *Current opinion in immunology* 20, no. 1 (February): 10-6. doi:10.1016/j.coi.2007.11.003.

Zuccaro A, Lahrmann U, Guldener U, Langen G, Pfiffi S, Biedenkopf D, Wong P, Samans B, Grimm C, Basiewicz M, Murat C, Martin F, Kogel KH. Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont *Piriformospora indica*. *Plos Pathog*. 2011; Oct;7(10):e1002290.

Anexo 1: Sequências CDS e de aminoácidos de *MpTLPs*

> seq1 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP1) gene, complete CDS

```
ATGAAACTTTTCGTCGCAAGTGCTTCTCTCGCTGGGTTGTCCACCGTTGCAGCTCGTACCTTCACGGTCTATAACGCGTGT
CCGTTTACCATCTGGCCAGCTCTATTTACGGATCTCCATGCTGGATCCGCTGTTTCTAACTTCAAGACCGGCTGGGAAGCT
CCCGCATGGTCCACGGTAACGTTCTCGGTGCCAGACAACTGGACGGCAGGAAGGATCTGGGCCAGACGTAACCTGCGACTTC
TCCAAGACCGGAGGTCCCACTGCTCAGTGCCTCACTGGCGGGTGCAATGGAGGCCTCGAATGCGACAGAAACACAGGCACC
GGTGTGCTCCCGCAACGATAGCGGAGTGGAATTTGTCGAGTAACCCCAATATTCCCGATAACTACGATGTCTCGCTTGTG
GACGGTTACAATCTTCTGCCAGGATCTCTAACAACAAAGGGTGCCCGTTGCCGAATGTGCAAAGGATCTAGGACCAGAC
TGTCTGCTCCCTTGAAAGGCCCATTTGACTCGACGGGATTCCCCGTGGGTTGCAAGAGTGCTTGCTTTGCAAACCTGGAC
GGCCACCTCAGGACTCGGCCAATTGTGTTCTGGCAGCCATAACACCCCGCAAACCTGCCCTGCCAGCGGCGTTCAATAC
TATAGTTACTTCAAGTCAAATTGTCCAAGGTCTTACGCCTACGCCTACGACGAGAATAGTGGCACTGCTTTGTTCACTTGC
GACTCGAACTCAAGGCTGATTATACTTTGACATTTGTCCCTGA
```

>

```
MKLFVASASLAGLSTVAARTFTVYNACPFTIWPALFTDLHAGSAVPNFKTGWEAPAWSTVTFVSPDNWTAGRIWARRNCDF
SKTGGPTAQCLTGCGNGGLECDRNTGTGVPPATIAEWLSSNPNI PDNYDVS LVDGYNLPARISNNKGCPVAECAKDLGPD
CPAPLKGPFDFSTGFPVGCKSACFANLDGHPQDSANCCSGSHNTPQTCFASGVQYYSYFKSNCPRSYAYAYDENS TALFTC
DSKLKADYTLTFCP
```

>seq2 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP2) gene, complete CDS

```
ATGAAGTTGTTCCGCCCTCTCTCACAGGCTCCGCCTTCTATCTAGCGTTGCTGCTCGTACATTTACCGTATACAACGCA
TGCCCGTTACAGATCTGGCCTGCAGTGTTACCGACCTCAACGTTGGCTCCGCCGTTCCCGACGTGCCTACCGGATGGGAG
GCCTCAGCATGGAGTGTGCGGTCTTTCAACGTCCCGGATAATTGGACTGCGGGAAGGATTTGGGGAAGGAAGAATTGCGAT
TTCAGCAAATCTGGCCCCGATTCTTGTCTTTCCGGCGGGTGCAATGGCGGTCTCCTTTGTGACCAGCGCACTGGAACAGGT
GTTCTCTCTGTATCACTTGCGGAATGGACGCTCTCTGCTGATCCCAATATCCCAGATCACTATGACGCTCTCACTTGTGAC
GGATAACAACCTGCCTATGAGGATCACCAACAACAAGGGATGCCCGTGGCTGATTGCCAGTCGACTTGGGTCCCAATTGC
CGTCCCCCATTAAGGAGCCCTTCGATGCTTCTGGCTTCCCGTAGGATGCAAATCCGCATGCTTTGCCAACCTAGACGGA
AACCAAGCGGAATCGCCTAAGTCTGCTGTTCTGTTGATTTTAACAGGCCGAGACTTGCCTCCAGCGGCGTCCAATTCTAC
GACTATTTCAAGGGTAAGTGCCTCACTTATGCTTACGCATACGACGAGAGCAGCGGCACCGCTTTGTTCACTTGCCTT
TCGACTGCCAAGGCCGACTATACTCTGACATTCTGCCCTTGA
```

>

```
MKLFAASLTGSAFLSSVAARTFTVYNACPFTIWPVAFVFDLNVGSAVPDVP TGWEASAWSVRSFNVDPDNWTAGRIWGRKNCD
FSKSGPDSCLSGGCNGGLLCDQRTGTGVPPVSLAEWTLADPNIPDH YDVS LVDGYNLPMRITNNKGCPVADCPVDLGPNC
PAPLKGPFDFASGFPVGCKSACFANLDGNQANSNPCCSGDFNRPETCPPSGVQFYDYFKGNCPNSYAYAYDESSG TALFTCP
STAKADYTLTFCP*
```

>seq3 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP3) gene, complete CDS

```
ATGCAACTCTTCTCCGTCACCATTTCTCGCTGTTGCCTCTTCCGCAGCCGCTCGTACTTTCACTGTGAAGAACAACCTGCAAT
TACCTGTCTGGCCAGCAATCTTACCGATACCAATAGTCCCAGCTGCGGAAGTGTGTACCCGACCAACCAACTGGCTGG
AAAGCTGATCCACGCACTCGGTCAACTTCAATGTTCCCGACAACCTGGAATGTGGCCGATCTTGGGGTCGAGCCGAATGC
GACTTTTCGACTAACCAGGTCTCAATCGTGCAAGACCGGTGGATGCGCTGGCGGATTGCAATGTACCCGAAATTACGAG
AAGCCCCGCTCTGTTGCCGAATGGACTTTTGATGGAACAAATGACAACCTACGATGTTTCCCTCGTCGATGGATTCAACATT
CCCCTGGAGATCACGAACAACAAAGACTGTCCTGTAGGATCTTGGCCGGTCGATTTGAATAACGGATGTCCCGCTGGCTTG
GAAGGTCCTAAGAATGGAGCAGGTGCAGTCATTGCATGCAACAGCGACTGTAACGTTGACAAGAACCATGATGACTCCCCA
AGCTGCTGCACTGGAAGTCACAATAAACCCGAGACCTGTCCTTCTAGTGGCGTTCCGCACTATAAACTTTAAGGATAAA
TGTCCCAAGGCCTACGCCTTTGCCTACGACGAACCCAGTGGCGCTCTCCTTACCTGTCCAGATAGCAAGAAGGCTGATTAT
ACGCTTACTTTCTGTCTTGA
```

>

```
MQLFSVITILAVASSAAARTFTVKNNCNYPVWP AIFTD TNSPDCGTAVPDQPTGWKADPRSSVNFNVPDNWKCGR IWGRAEC
DFSTNPGPQSKTGGCAGGLQCTGNYEKPASVAEWF DGTNDNYDVS LVDGFNI PLEITNNKDCPVGSCPVDLNNGCPAGL
EGPKNGAGAVIACNSDCNVDKNHDDS PSCCTGSHNKPETCPSSGVPHYKYFKDKCPKAYAFAYDEPSGALLT CPDSK KADY
TLTFCP*
```

>seq4 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP4) gene, complete CDS

```
ATGCAATTATTCTCTATCACTCTTCTCGCTGTGCTTCTTCCGCAGCTGCCCGTACTTTCACTGTAAAGAACAACCTGCAAG
TACACTGTTTGGCCTGCGATTTTACCGACACCACGAGCCCCGACTGTGGAACAGCTGTACCCGACCAACCAACTGGCTGG
AAAGCTGATCCAGGTAGTTCTGTCAATTTCAACGTTCTTGACACCTGGAATGCGGGCGTATGTGGGGTCGACGCGATTGC
GACTTTTCCAACAACCTCAGGCCCGAGTCATGCGGGACTGGTGGTTGTGCTGGCGGATTGAAATGTACTGGAACCTATGAG
```

AAGCCCCGCTCTGTTGCTGAGTGGACTTTTGACGGCACTAACGACAACTATGATATCTCTCTTGTGCGATGGGTTCAATGTT
 CCTCTGGAGATTACGAACAACAAGGGTTGCGCGGTGGGATCTTGCCCTGTTGATCTGAATAACGGATGTCCCGCTGGCTTG
 GAAGGCCCAAAGAATGGAGCAGGCGCAGTCATCGGATGCAACAGCGACTGTAACGTCGACAAGAACCCTAATGATTTCCCGG
 AGCTGCTGCACCGGAAGTCACGATAAACCTGAGACATGCCCTCCTAGCGGCGTTCCGCACTACCCTACTTCAAGGATAAA
 TGCCCCAAGGCTTATGCTTATGCGTACGACGAGCCAGTGTTCCCTTCTTCATTGCCAGATGGCAAGAAGGCTGATTAC
 ACGCTTACTTTCTGTCCTTGA

>
 MQLFSITLLAVASSATARTFTVKNNCKYTVWPAIFTDTTSPDCGTAVPDQPTGWKADPGSSVNFNVPDWTWKGRMWGRRDC
 DFSNNSGPESCGTGGCAGGLKCTGNYEKPASVAEWTFDGTNDNYDISLVDGFNVPLEITNNKGCavgscpVDLNNGPCAGL
 EGPKNAGAVIGCNSDCNVNDKNPNDSpscctGSHDKPETCPPSGVPHYHYFKDKCPKAYAYAYDEPSGSLHCPDGKKADY
 TLTFCP*

>seq5 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP5) gene,
 complete CDS

ATGAAAGGATTCTTTCGCCCTTCGCATTGCTCAGCGTCGCTCGACTGTCAGTCTCGCAGCTTTACAGTCAGAAACAACTGT
 GCTTACACCGTTTGGTGCATGTTACCGACCCCTAACGCTGGTTCCGCCCCGCCAAATTTTGCAACCGGGTGGGAAGCCGCA
 CCAAGACTCAAGTGTCTTCTCCGTTCCAAACAATTGGAAGGCCGTCGCATCTGGGACGGCTCCAATGCAATTTCTCCA
 CCAACCCCGGGCCCACTCTTGCGCCAGTGGCGGATGCAATGGCGGACTTCAATGCGCCGCTACCGCGGGTGTTCACCG
 GCCTCGCTTGTGAATGGACTCTTCAAGGTGCAGGAAACCTTGACTATTATGATGTCTCGCTCGTTCGATGGCGCTAACATT
 CCTATGGCGATCACTACCAACGTGAACTGCCCCGTCGCTGACTGCCCGGTGCACTTGAATGCAAACCTGCCCGCTCAGCTC
 AAGGTACCCGGAGGATGTAAATCGGCTTGTGTGGCCAACTTGACGGAACCCGCAAGATTCCCCAACTGCTGCAGCGGG
 TCTCACACACCCAGCTACCTGCCCTTCCAGTGGTGTCAATTCTTCAGCTACTTCAAAAATGCCGCCCAACGCTTAT
 GCTTATGCTTACGACGACATCTCAGCTACGAGGACCTGCAACTCTGGTCTGAACGCCGACTACACGTTGACCTTCTGCCCT
 TAA

>
 MKGFFAFALLSVASTVSARTFTVRNNCAYTVWCMFTDPNAGSARPNFATGWEAAPKTQVSFSVPNNWKAGRIWDGSNAISP
 PTPGPTLAPVADAMADFNAPLPAGVPPASLAEWTLQAGNLDYYDVSLVDGANIPMAITTNVNCVPVADCPVDLNNCPAQL
 KVPGGCKSACVANLDGNPQDSPNCCSGSHNTPATCPSSGVQFFSYFKNACPNAYAYYDDISATRTCNSSLNADYTLTFCP
 *

>seq6 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP6) gene,
 complete CDS

ATGAAAGGACTCTTTGCTCTTGTGCCTATTGCTCTCGCTCGACTGTCATCGCGCTACGTTACAGTCAGGAACAACTGT
 GCCTTTACTGTCTGGCCCGGATTGTTCACTGACCCCAACGTCGGCCCCGCTCGCCAGATCAACTGACTGGATGGGAAGCA
 GCGCCTGGTACTCAGGTATCTTTTTTCACTACCAACGATTGGAAGCGGGAAGAATCTGGGGCCGAAAAGAATGCGACTTT
 GCCACCAACCCCGGACCACTTCTGTGTCAGCGGTGGTTGCAACGGAGGCCTCGAATGCGATCCAAACACGGGAACAGGC
 GTACCCCCCGCAAGTATAGCAGAATGGACCCCTTCAAGGGGATGGCAACCTCGATTACTATGATGTCTCCCTCGTCGACGGT
 TTCAACATTTCCCATGGCAATTACCAACAATGTAGACTGCCCGGTGGCCGACTGTCCGGTGCAGTTGACCGCTACCTGCCCC
 GGAGAGCTTGCATCAGCAGCAGGATGTGCATCAGCCTGCTTTGCAAACCTGGACGGAACACAGCAAGATTCCGCTAACTGC
 TGCAGTGGGTGCGACAGCACCCCTGATACCTGCCCCCGGACGGTGTCAATTCTATGACTTCTTCAAGGGTGCCTGTCCC
 AATTCCTACGCTTATGCATACGACGAGCGCAGCGACACTGCTCTCTGGACATGCGACTCCTCTTTGAACGCCGACTACACT
 CTTACCTTCTGCCCTTAA

>
 MKGLFALVPIALASTVIARTFTVRNNCAFTVWPGLFTDPNVGPARPDQLTGWEAAPGTQVSFSVPNDWKAGRIWGRKECDF
 ATNPGPTSCVSGGCNGGLECDPNTGTGVPPASIAEWTLQGDGNLDYYDVSLVDGFNIPMAITNNVDCPVADCPVDLTATCP
 GELASAAGCASACFANLDGNQQDSANCCSGSHSTPDTPPDGVQFYDFFKGACPNAYAYYDERSDTALWTCSSSLNADYT
 LTFCP*

>seq7 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP7) gene,
 complete CDS

ATGAAAGGACTCTTTGCTATCATACCTCTCGCCGTCGCTCGGCTGTCAGCGCACGTACGTTACAGTAAGGAACAACTGT
 GCCTATACCGTTTGGCCTGGATTGTTACCGACCCCTAACGTTGGCCCCGCGCTCCGATCAACCGACTGGATGGGAAGCC
 CGCTCTGGTACTGAGGTGTCTTTCAATGTACCAATGATTGGAAGCAGGAAGAATCTGGGGCCGAAAGAGATTGTGACTTT
 TCTACTAACCCAGGACCGACTTCTTGTGTCAGCGGTGGTTGCAACGGAGGCCTCGAATGCGATCCAAACACGGGAACAGGT
 GTGCCACCTTCGAGTGTGGCAGAATGGACTCTTCAGGGAGATGCCAACCTCGATTATTATGATGTTTCCCTTGTGCGATGGG
 TTCAACATTTCTATGGCAATTACTAACAACGTGGACTGCCAGTGGCTGACTGTCCAGTCGACCTGAACGCAAACCTGCCCCA
 GCAGAGCTCGCATCAGATGCAGCAAACCTGGACGGGAACCCGCAAGATTCTCCTAACTGCTGCAGTGGGTACACGATACC
 CCCGAAACCTGCCAGCCAGTGGCGTTCAATTTTACGATTTCTTCAAGGGTGCCTGCCCAACTCCTATGCTTACGCATAC
 GACGAGAGCAGCAACACTGCTCTCTGGACTTGCGACTCTTCTTTGAATGCCGATTACACTCTTACCTTCTGCCCTTGA

>
 MKGLFAIIPLAVASAVSARTFTVRNNCAYTVWPGLFTDPNVGPARPDQPTGWEAASGTEVSFNPNDWKAGRIWGRRDCDF
 STNPGPTSCVSGGCNGGLECDPNTGTGVPPSSVAEWTLQGDANLDYYDVSLVDGFNIPMAITNNVDCPVADCPVDLNNCP

AELASDAGELSQH*TR*SESLSKPVK*SACFANLDGNPQDSPNCCSGSHDTPETCPASGVQFYDFFKGACPNSYAYAYDES
SNTALWTCDDSSLNADYTLTFCP*

>seq8 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP8) gene,
complete CDS

ATGGATCGGTTATTTCGTCATTCTTTTACTTTCAACCACCATCTTGCTCGCTGCTCAAGCTCGCTCTGTAACAGTCAAGAAT
AACTGTGCTTTTACTGTATGGCCAGCAATTTTCACTTCTTCGCTCGGCGGAAGTGCACGAACCAACCAGATTCGAAGACT
GGATGGGAAGCTCCCGCAGGCTCACAAGTTACTTTCTCAGTACCAGACAACCTGGCAGGGTGGCCGAATTTGGGGTCGTAGA
GATTGCGATTTCTCCAGGACACCTCCTTTGACCTGTGCCAGTGGATCTTGTAACGGTGGCCTCGAATGTACATCTACCGGC
GTACCACCAACAAGTTCAGCAGAATTCACCCCTTGTCATCAGGTTCTATACCAGATTTCTATGACGTGTCACTCGCGGACGGT
TTCAACCTCCCCATACAAATCACCAATAACGTCAAATGTCCAGTTGCTGAATGCACCGTCGACTTGAATGCAGCTTGCCCA
AGACAGCTTCAGGGTAAACCCACGTCCGATGGGGCAGTCTCGGGCTGCCTGAGCTCGTGTCTTGCCCAATTTGGATGGACAA
CAATCGAATTCTCCCACTGCTGCTCTGGTGACAACAACAGTCCGACTAACTGTACAGCAGCTGGTGTCCAATTCTTTGAT
TACTTTAAGGGTAGATGCCCCGAATGCACCTTGCTTATGCGTTTGATGATGCTGCTCTTTGGACATGTGATGGAAAGTTGGG
GGAAATTACACGGTCAGTTCTGTCCATGA

>

MDRLFVILLSTTILLAAQARSVTVKNNCAFTVWPAIFTSSSGGSATNPQDSKTGWEAPAGSQVTFVSPDNWQGGRIWGRR
DCDFSRTPPLTCASGSCNGGLECTSTGVPPTSSAEFTLASGSI PDFYDVSLADGFNLPIQITNNVKCPVAECTVDLNAACP
RQLQGKPTSDGAVSGLSSCLANLDGQQSNSPNCCSGDNNSPNTCTAAGVQFFDYFKGRCPNALAYAFDDAALWTCDGKLG
GNYTVTFCP*

>seq9 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP9) gene,
complete CDS

ATGAAAGGACTTTTCGCTTTCGCGTTGCTCAGCGTCGCCTCGACTGCCAGTGTCTCGCACGTTTACAGTCAGGAACAACCTGC
GACTTCACCGTTTGGCCTGGAATGTACACCGACCCTACCGCTGGTCCCTCCCGCCAGATTTTGCAACTGGGTGGGAAGCT
GCGCCTAATACTCAAGTGTCCTTCTCCGTTCCAAACGATTGGAAGGCCGGTTCGCATCTGGGGACGGCTCGAATGCGATTTC
TCCACTGATCCCGGTCCCAATTCCTGCGCTAGTGGCGGGTGCAATGGTGGACTGGAATGCGCCGCTGATGGCGGTACTGGT
GTTCCACCGGTTACCCCTTGCTGAATGGACTCTTCAAGGTGCAGGAAACCTTGACAATTACGATGTTTCTATCGTCGATGGC
GCTAATATTCTATGGCGGTCACTAATAACGTGAACTGCCCCGTGCTGACTGCCCCGTCGACTTGAATGCAAACCTGCCCC
GCTGAGCTTGCATACCCGCCGGATGCAGATCGGCTTGCTTTGCCAACCTGGATGAAAACCCGCAAGATTCCCCCAACTGC
TGTACCGGGTCTCACAACTGAAGATACCTGCCCTTCCAGCGGTGTTCAAGTTCTACAGTTACTTCAAGGACGCTTGTCCC
AACGCCATATGTCTACGCTTACGACGACGCCTCTGCTTTGCGGACCTGCGATTCTGGTTTGAATGCCGACTACACCTTGACC
TTCTGCCCTAA

>

MKGLFAFALLSVASTASARTFTVRNNCDFTVWPGMYTDPTAGPSRPDFATGWEAAPNTQVSFSVPNDWKAGRIWGRLECDF
STDGPNNSCASGGCNGGLECAADGGTGVPVTLAEWTLQGAGNLNDYDVSIVDGANIPMAVTNNVNCVPADCPVDLNANCP
AELASPAGCRSACFANLDENPQDSPNCTGSHNTEDTCPSSGVQFYSYFKDACPNAYVYAYDDASALRTCDSGLNADYTLT
FCP*

>seq10 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP10)
gene, complete CDS

ATGCATTTGCCGACTGTTCTTATTGTTGCTGTGCGCTCCACAGCAACTGCACGTACTTTGACGGTTAAGAACAACTGTGCT
CATACGGTCTGGCCCGCATCTTTACCGACTCTAGCTGTGGAGGCGCCGCGCCAGATCAACCAACCGGCTGGCAAGCTGAT
CCAAACAGCAGTCTCACCCTCGATGTAGATGATTGGAGATGCGGTTCGATCTGGGGTCGTGCAGGCTGTGATTTTAGTGGC
GTAGGTTCTTGTCTGACAGGAACTGCGAAGGAGGGCTTGAGTGTACCGGAAGCTACGAGCCGCCTGCAACTGTTGCAGAA
TGGACGTTACGCGAGACTGGGGACTTTTACGATGTTTCTGTGATTGACGGGTTCAATATGCCAATGGCTATCACTAATAGC
CAGAACTGTGCAGAAGCCTCATGCCCTGTCGATCTAAATCAGAACTGCCCCGTAGAACTTCAAGGTGTTTTCAACGCCGAC
GGACAGGTGCTGCGGTGTAATAGCGACTGCCTCGTTGATCCCAACCAAGCTCGTGTCTTTGCTCTGCCCTCGACTCCACA
GTATTGACTTGTACACAGGTGACTCTGCCAGCTGCTGTACTGGTAGTCACATGTACAGCCAACCAGCAAATTGCCAGCT
AGTGGCGTCCCGAACTATGCTTACTTTAAAAATAACTGTCCGAGGGCACATGCGTTGCTTTATGACGAATCCAGTGTTCT
ATCAACACCTGCACTGGTAAAGACCCGATTATACTCTTACATTCTGTCCTTGA

>

MHLPTVLIVAVASTATARTLTVKNNCAHTVWPAIFTDSSCGGAAPDQPTGWQADPNSSLTFDVEDWRGRIWGRAGCDFSG
VGSCLTGNCGGLECTGSYEPPATVAEWTTFSETGDFYDVSVIDGFNMPMAITNSQNCAEASCPVDLNQNCPELVQGVFNAD
GQVVGNSDCLVDPNPSSSCSALDSTVLCTQVTLPAAVLVVTCHSQPANCPASGVPNYAYFKNNCPRAHAFVYDESSGS
INTCTGKRPDYTLTFCP*

>seq11 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLP11)
gene, complete CDS

ATGCTGAACATGAGATTTGCTCAGTTCTCGTTCTTCCACCTTTCTTTGAGCCCGCTGCTTGTGCTAGCTGACCACCAATTC
ACCTTCACCAACCGATGCGGCAACACTATCACCCAGTAATTGCTGACACCAGGTGTGGATATAGTCTCGCTGTGATGGC
GCCGCTAGCTACACTTCCGCTCAACCCAGTCAGATGGCTTCGGGAGCATCCCAACGGTTACCGTTCCATCTAACTGGGTC
GGCCGCATCTTTGCGCAAAACGGAAGCTGTGGTCCCAAAGGTGAAGATTGCTCGGTTACTGAGTTTAATCTGGATTCCGGC

GATGCCTTCACCCCCAGTCGTATGACATTTCAAACATTCAAGGATTACACAATCGTTCCAAATAGGCGCTGCCGGTTGC
GAAACTGTAAC TTGTACCA GTGCTGATTGTGGGTGCAAAAATGCATATCCGGTCGGAGATATCAGCGGTTGTGGGAACGAC
AGCCCTGTCAGGGGATGTGGTGCAGGAAATGTTGCTTTTAGCGGTAAAGACTCTTTCTTTCACCGGATCCTAA

>

MLNMRFAQFSFFTLSPLLVLADHQFTLTNRCGNTITPVIADTRCGYSPRCDGAASYTSAQPSQMASGASQTVTVPSNWW
GRIFAQNGSCGPKGEDCSVTEFNLD SGDAFT PQSYDISNIQGFTQSFQIGAAGCETVTCT SADCGCKNAYPVGDISGCGND
SPVRGCGAGNVA FSGKTL SFTGS*

>seq12 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTLp12)
gene, complete CDS

ATGAAACTCTTGCAGGCCTCCCTTCTCGTTGTGGCTCTCAGCCCAATGCTCTCCCTGGCTGCTCATCGGTTACCCCTCACA
AATAGATGTTTCGAGTCCAATCACGCCTATCATTGCGGATACCCGGTGTGGATATAGTCCTCGCTGCAGCGATGCCGCTAGC
TACACTGCGGCACAACCCAATCGTCTCGCAGCCGGAGCTTCGCAAACAGTTACCATCCCTTCCCGATGGGTGCGACGTATT
TTTGGCCAGAATGGACGATGTGGCGTCAAAGGAGAAAGTTGTTTCATTGACCGAGTTCAATCTTGACTCGGGGGACGTCTTC
ACTCCGCAATCCTATGATATCTCGAACATCCAGGGCTTCACGCAGTCACTTCAGATCAACGCCAACGGGTGTGAAAGTGTT
ACCTGCACGAACGTCAACTGTGGTTGTAGGAGCGCTATCCCCCTGGGGATCTTAGCGGGTGCAGGAAATGACAGGCCCGTA
CGAGCATGTGGTGCAGGAGATATCGCTTTCAGCGTGGTATTCTGCCCTTGA

>

MLSLAAHRFTLTNRCSSPITPIIADTRCGYSPRCDGAASYTAAQPNRLAAGASQTVTIPSRWVGRIFAQNGRCGVKGESCS
LTEFNLD SGDVFT PQSYDISNIQGFTQSLQINANGCESVTCTNVNCGCRSAYPPGDL SGCNDRPVRACGAGDIAFSVVFC
P*

>seq13 [organism=Moniliophthora perniciosa] Thaumatin-like protein (MpTL13) gene,
complete CDS

ATGATGTGGACACTATTACTGGTATTGGCCGCGGCGAATAGCAGAGGGCTTGCAGCTCGCACCTTTACAGTGACCAACAAT
TGTGCATACACGATATGGCCTGCTATCTGGACGGACCCGAACGTCTCCAAAACAAAAGTAGATCATCCAGGAGGATGGGAG
GCGGCCTCCGGAACGAGCGTTTCATTCTCAGTGCCTGATGATTGGGGACCTGGACGAATATGGGGTCGCCGAGAATGTGAT
TTTGGGAGCTCTCAAACGGAACCTTGATCAGTGGTGGTTGCGATGCCTTGAGTGTACGAATGTGGGACAACCTCCTACT
ACTGTTGCCGAGTTCGCTTTGTCTACGGCTGGTGGTAATACGGATTACTATGACGTTTCACTAGTCGACGGGTTCAATCTT
CCTTTGCAAGTCACCAACAATGCCAAATGTCAAGTGGCGTCTGTCTGTTGACCTGAACGCCTCATGTCCACCACAACTT
CGAGGTACTCCCGCCAAAGATGGATCCAACCTCCGTATGTAAAAGCTCCTGCCTTGCAAACCTTGACGGAAGCAGATCGAT
TCTCCCAATTGTTGCTCGGGAACGAGTAACACACCCGACAGCTGCACTCCAGCTGGAGTTCAGTTTTATGACTTTTTCAAG
GGAAGATGTCCCTCAGTATATAGCTACGCGTATGACGAGCAGAGCGGTTCTGCCTTGATGAGCTGCGATTCAACACTTTCA
GCGTCGTATACCATCACTTTCTGTCCATAGATTA

>

MMWTL LLLVLAAANSRGLAARTFTVTNNCAYTIWPAIWTDPNVSKTKLDHPGGWEAASGTSVSFSVPDDWGPGRIGRRECD
FGSSQNGTCISGGCDALECTNVGQPPPTVAEFALSTAGGNTDYDVS LVDGFNLPLQVTNNAKQVASC PVDLNASCPPQL
RGTPAKDGSNSVCKSSCLANLDGKQIDSPNCCSGTSNTPDSCTPAGVQFYDFFKGRCP SVYSYAYDEQSGSALMSCDSTLS
ASYTITFCP*

Anexo 2: Sequências CDS e de aminoácidos das *MpCSEPs* estudadas.

>MpCSEP1

ATGCAATCCTGGCGCTCTCTTTTGTCAATCTTTACTTTCTCTCTGCCCTTGCGACACTCTCGCATGCTATCCCTACCCCC
CGTAATGTAGCTGTGCAAGCTCGTGCTCCCCACGGCATCGACGGATATTCGGCAACCATGACGGTTCCTTCGGCCATGAC
AGCTCCTTCGGCAAAGGTGGCTCTGGCCGTGGCGGCCATAGCTCCCTCGACCATCATAATGGTGATGAGTACGGCGGATAC
TGCGACGACGTGTGCCACCAGGCCAAGCAGCACAGCCTTTACGATATCATCGTCGACATCAAGGTCAAAATCACGGAAATT
GCGCAAACCGTCAAGACGTATGAGACCTGTGACGAGGTAACGTCCGACCTCACTTCGACGAGATCAGAAAGCACTTCGAG
ACCGGCATCCAAACCATCACGACTTACATTAGTGCTGAACTTGATATCACCATCATCTTGGGCGGACATAGCTCTGACGAG
ATCAAAGCTGTTCTCTGCGATCTCCTGCATATCATCGTCGAACTCATCCAGTTCATTGCGCAGACCAGCATCCAACTGCT
GTCCAAATCGTCTTCTCGCAGTTAAAAGACCTGATCGTCTGCATTGGCGGGTTCATCCAGGTTATCACCGTCGTCGTGACC
GAATTGCGAGTTACGATATTCGACGATCTTGACTTCTTCTCAAGATCTGCAATACCTACGGTGTTGATCCCTCTACGTTCT
TTCCACAACCTAG

>

MQSWRSLLSIFTFLSALATLSHAIPTRNVAVEARAPHGIDGYSNGHDSFGHDSSFSGKGGSGRGHSSLDHNGDEYGGY
CDDVCHQAKQHSLYDIIVDIKVKITEIAQTVKTYETDEGNVRPHFDEIRKHFETGIQTITTYISAEIDITIILGGHSSDE
IKAVLCDLLHIIIVELIQFIAQTSIQTAQIVFSQLKDLIVCIGGFIQVITVVVTELRVTIFDDLDFFKICNTYGVDPSTF
FHN*

>MpCSEP4

ATGCAGTCCCGCTCGCTCTCGTTCTTGTCCGTCTTGCTCCTCGTCTTCACGCAGCTTACGTCTGCTCTTCCCAACCCAGCT
CCCCAAGAGCTCGTGGTTCGTGCTGCTGAAGCCAAAGCCGCTGCTGTCAATATCGCAGTGCGCGACAATCCCCCGCCCG
TCCCCGGGCAACGGAAGGTGCGGCAACGGTAATGATTTCGCACGACAATGGGCGGATGCGCAAGATCGACGATGTGCGTGAA
AATGACGATTTCAACAACAACGACAACAACAACGACAACAACCCGCCGTCCCCCTGGCCGCAACAACAGGTGCGACGAA
AACGACTTCGATTGTTGGAGGACGGTCGATTGGAATCGTTGCGATCAAACCTGTTGGGACATCAACGGTTCCAACCTTCATC
ATCTATATCCGAGCTCAAATCCTTGTGAGATCAATTACCTCATCCAATCTGGCTCGGACATCGACTACAACACTCTCAGT
AGTCACGTCAATAATATCCGAGGTTACATTGACCAACTCGCGATCGGACAAGCAGCTATGGCAATCAGTGGTCCCAATAC
AACGGCGGATATATCCGTAATCAAATCGTAGACATCGTCGATGGACTCAACCGGATGATCGATCATTATCCTCCAATGTC
AATTCGAATGTTTACGGGGCGCCTTGACGCCATTGTGCCAGGATGTGGGGTCTTGGTTCAGTACAGCCAATAACGCCGGG
TTTGGCATTTTCGAGCACGATCTCGAGCATTAGTCAATTCACGCAATTGTTCGGCCTCCAAGTTCGGAGTGTGATCTCGTTG
TAG

>

MQSRSLSFLSVLLLVFTQLTSALPNPAPKELVVRAAEAKAAAVNIAVRDNSPPSPGNRSGNGNDSHDNGMRKIDDVGE
NDDFNNDNNNDNNPPLPPGRNNRCDENDFDCWRVTDWNRCDQTCWDINGSNFIIYIRAQILVEINYLIQSGSDIDYNTLS
SHVNNIRGYIDQTRDRTSSYGNQWSQYNGGYIRNQIVDIVDGLNRMIDHLLSSNVNSNVYGGALQPLCQDVGSWFSTANNAG
FGISSTISSISQFTQLSASKFGVSISL*

>MpCSEP5

ATGCAAACCTGGCGCACAAATCTTTTCGCTCTTTTTCTCTCTGTTACCCCTGAACTTGACAGGTCTCTGCCCTCCCGAGCCCA
AAGGTTGACGGTGGCTTGTTGCTCGTGCTCCTTCTGCCCCGGGTGAACGCCCTTAACATACGCGCAGACAAGCACGCG
CACGCTGTCCTTGCTGATGTTACAGCCAAGATTAATAATTCGAGTTGGTGCTATCCTAGGCGCCAGCGACCCTGATGAAGTC
CAGGTGCAGGTGCACATCGACGCCATCAAAGCACAGATCAAAGTGGCCCTAGGAGCGCTACAGGCCGACGTTGACGCCAAT
GTCGACCTTGATGTTATCCTCGGGGACATGGCTTAGACGACTATGCTAACCTTATTAACGACATTGCCAGCTCATCGCG
CAGCTTGTC AACCGCTTCGAAGGCGCCGACGGTATTCTATCCCTGCTAGCGGATTTATCCAAAGACTTTGATGGCTATTTG
AAGTTGGCCGTTGGCCTTTCAGTCGATTGGGTAATATCCTTATTGGGTTGGGAATTAAAGTCTTCTCAACGTGGGTGGT
CTGTTGGGGATCTCGCTTGGGGTCGGCATCAACCTAGGTTG**TAG**

>MQTWRTILSLFFFLFTLNLQVSALPSPKVDGGLVARAPSAPGLNALNIRADKHAHAVLADVHAKIKIRVGAILGASDPDE
VQVQVHIDAIKAQIKVALGALQADVDANVDLDVILGGHGLDDYANLINDIAQLIAQLVNRFEAGDGLISLLADLSKDFDGY
LKLAVGLSVDLGNILIGLGKIVFLNVGGLLGISLGVGINLGL

> MpCSEP7

ATGCAGTCCCGCCGCTCGCTCTCGTTCTTGTCCGTCTTGCTCCTCCTCCTCACGCAGCTTACGTCTGCTCTTCCCAACCCA
GTTCCCAACGCTCTCGAGGCTCGTGTTGCCGGACCCAAAGCCGTTGCTGGCAATCTCGCAGTGCGCGCGGACGATAGCTAT
GGCGGTAGTGGCGGTAGTGGCGGTAATACCAATCAAAACACGCCGCCGACCACCCCGTCCAAGTCCAAGTCCAAGGAGGAA
ACCAAAAATGACGGTGGAAGCCCGGCTCGGAGGATGGCAAGGTCGACGAGGACAAGTCGTTGTCCACCTCTAGGCTCAAG
GAGGACAATACTGAGGATGGTACCAAGAAAAACAAGCTGCTGTCCGATTCCAAGGAGGAAAATGATGACGGTCACGGCTCG
TCCAAGTATGACAATGACGACATTAACAACTCCGAAGATTGCGACGAGTGTGAGAACAAAGTCCACATGGGTCAAACATTGC
GATAGCAGCTGTCAGAACGACGGGAAGAGCTTCATTAACATATGTTCAAGACGGTATTGGTTATGAGTTCGGCCTCATCAAG
CGATATGGCCCCGGAATCAGCTCTGACGAGTGCAATAAGCACATCAACAAGATCCAAAATTACATCTGGCAAATCAACCAA

CAGACGGGCACCCCTTGGCAGTGTGTGGTTCCAAGGCAATGGCGCTCAATACCATGGTCAAGTCGATGGCGTCGTCGGTCTT
ATCAACCAGATCATCCAGTATTTCTCCAACGCCAATATCGATGTCAGCGTTTATGGGAGCGTCCTGCCGCAATTGTGCAGC
GCAACGGGTTCTTGGATTAGAGGTCCAATAACTTCGGGCTGGGACTTACGGCCTCGATCTCGAACAGCATCCAAACCTTC
GTGCAGCAGTCGTCCTAGCTTTGGAGTGGTGATCTCCTTT**TAA**

>

MQSRRSLSFSLVLLLLLTQLTSALPNPVPNALEARVAGPKAVAGNLAVRADDSYGSGSGGNTNQNTPPTTPSKSKSKEE
TKNDGGKPGSEDEGKVEDKSLSTSRLEKEDNTEDGTTKKNLLSDSKEENDDGHGSSKYDNDINNSEDCECENKSTWVKHC
DSSCQNDGKSFINYVDGIGYEFGLIKRYGPGISSDECNKHINKIQNYIWQINQQTGTLGSVWFQNGAQYHGQVDGVVGL
INQIIQYFSNANIDVSVYGSVLPQLCSATGSWIQRSNNFGLGLTASISNSIQTFVQQSSSFGVVISF*

>MpCSEP9

ATGCTCGGCAGTCGCATACTCCTCATCGGGCTTTTCTTTCATCCAGGCATCTCTTGCACACCAACCAAGACACCAAAACAT
GTCGACTGTGCTTGTCTTGATGGATTGAACCATTGGGCGGGCAGAAATACACAAATCGCATGCAAGGTAGTAGCGGGAGAT
TACAAGAATGACAAGGATGCCGTTCTGTTGGTTTGACATTGGAGATGGTCATGGAACGTGAACGTGCAGAATGACTAAAGCA
GTATCAAAGAACGTTGACTGGAACCTCGTTTGCAACCAATGGGGCGAGCGCCATCGGGCCAATGCTGGGACGTGAATATT
TGA

>

MLGSRIILLIGLFFIQASLATPTKTPKHVDCRCLDGLNHWAGENTQIACKVVAGDYKNDKDAVRWFDIGDGHGTVTCRMTKA
VSKNVDWNFVCNQWGGAPSGQCWDVNI*

>MPCSEP10

ATGCAGTCCCGACTCTCAATCTCGCTCTTGTCCGCCTTGCTCCTCTTGTTCACGCAGCTTGCTACTGCTCTTCCCAGCCCA
GCTGCTCGAGAGGTTGAGGCTCGTGTGCCGACGTCGAGGCCCGAGAATGGGGCTATGACAATAACGAATGGGACCACCAT
AGTGCTTGGGATATCCTCGTCCAAGTGCAGGCCGATATCTCTGTTGAGATCGATGCTATTAAGAGCTACCAGGAGCTCACC
TATGACGATTGCTCCAAACACGTCACCGTAATTCGAAGCCGCATCACGCATATCTACGAGACCATTACGGCGAAATACGAT
GCCGGTTCTAGCTGGGATGATATCTTCGACGGACACGATTCTGATGAGATCCGTGCTCAGATCGTTATCATCGTCAACGTC
ATCACGGACCTGCTCAACTATTGCTACGACATCGGAATCAACATTTACATCTCCATCCTGACAGACTTGTGTAACGACATT
AATGTTTGGTTGCAATTGTCTGCTCGCGTCGTTGTGGGGCTTTACAGCTTCGCTCTCAGTTAGTATCGGTTGCTTCGTGGGC
TTGTCTCCAACCTGGGAGTGTTCGATTTCGTT**TAA**

>

MQSRLSISLLSALLLFTQLATALPSPAAREVEARAADVEAREWGYDNNEWDHHSABDILVQVQADISVEIDAISYQELT
YDDCSKHVTVIRSRITHIYETIQAKYDAGSSWDDIFDGHDSDEIRAQIVIIIVNVITDLLNYCYDIGINIYISILTDLCNDI
NVWLQLSARVVVGLSASLSVSIQSVGLSSNLGVSISL*

>MpCSEP12

ATGCAGTCCCGTCGCTCAATCTCGCTCTTGTCCGCTTGCTTTTCTCTTTGCGCAGCTTGTTTCTGCCCTTCCCAGTCCA
GCTGCCCGACAGCTCGAGGCTCGTATTGCTGAGCCCGCTGCTGTGCGAGCCGTGTCGCAAGGATGGAACTATGGCGATAAG
AAATTGAACGGGGACCAGTATGGCGATGTAAACTGGACGGGGACAATGACTGGGACTACGGGAATAGCAATAACGAACAT
TATGGCGATGAATCGGAACACCATGAACGCAATGCATGGGAGGTTCTCCTCCAAGCTCACCTCGATATCTCTGCTGAGGTC
GATATTATCAAGAGCTACCAGCAACTACCTACGACGATTGCTCCAAACACATCGTCGTAATTCGAAACCACATCACGCGT
ATCCACGAGACCATTACGGTGAAATACCAGGCTGGTACTAGCTGGACTGATATCTTCGATGGACACGACTCTGAAGAGATC
CATGCTCAAATTGTTATCATCATCAATGTCAATTACGGACCTGTTCAACTACTTCTACAGCATCGGAATCAGTCTCTACATC
TCCCTCCTGGTGGACTTGTGCAACGACATTAGCGTTTGGTTGCAGTACTCCTCTACCGTTGTGGTGGGGCTTTTCGGCCTCG
CTCTCTGTTAGTGTTAGCACGTTCTGCTGGGTCTGTCTCCAACCTGGGAGTGTCGATCTCGTT**TAA**

>

MQSRRSISLLSVLLFLFAQLVLSALPSPAARQLEARIAEPAAVDAVSQGWNYGDKKLNGDQYGDVKLDGDNDWDYGNSSNEH
YGDESEHHERNAWEVLLQAHLDISAEDIIKSYQQLTYDDCSKHIIVIRNHIETRIHETIQVKYQAGTSWTDIFDGHDSSEI
HAQIVIIINVITDLFNYPYSIGISLYISLLVDLCNDISVWLQYSSTVVVGLSASLSVSVSTFVGLSSNLGVSISL*

>MpCSEP13

ATGGCTCGTTCTGTTTTAAACCTGTTCTCGTTTGGCGATTCTCTCCGCGTTTTACTCTTCTGCTTTCAATATCACTGGC
CCCGTCTTACCAGTGAATGCCCGACAAACGTTCTGTCAACTGGACATGGTTTGAAGGAGACCCGACGAACATTGTG
GTCAACCTTTTCATGAACAACGATATCGATGTTTGCACCTTTGGACGGGCATATCCGTAACCTCGAAAGGAATCCGGCAC
TATATTGACGGTATGCAAGAAGCGGATGGGAGGAAGTTTGAAATTTTTCTTTCAAAGCGAATGACACTGGGCAATATATC
CTCTGCGCTATGAGGATATCACTCCACCAAATGGCACCTCAACTTTCGCTCCGTCGCAAACTCTACTGCCTTCAATGTA
ACAAATACGACTACGACGATCGAAATGCCACCTCCTTTACACCTACCCCAACAGCGTCGAAACCCCAAGTACACCGACT
GGGAGTGCCGTTTCGACGGCTGGACACAACACACCCAGCAACACACCCGAGAACAAATCAAAGCTCTCCACCGGTGCATTG

GCTGGAGCGATTGTTGGAGGTGTCTTATTTGCAATTGTCTTGATAGCTTTTATATTATTTTTCTTCGCCGGGCGTCTCTTC
TTCCCTTGATGGTTACAGCAAGGAAGAAGTTTTACAGTAAAGTTTCGCATGCTCCTCCAAAGGACAAG**TAA**

>

MARSVLNLFFVCAFLSAFYSSAFNITGPVLTTVNAPDNVPVNWTWFEGDPTNIVVNLFMNNVSDVCTFGRAYPVTSKGIRH
YIDGMQEADGRKFGNFSFKANDTGQYILCAYEDITPPNGTLNFRSVANSTAFNVTNTTTTIEMPPFTPTPTASKPPSTPT
GSAVSTAGHNTPSNTPENNSKLSTGALAGAIVGGVLFIVLIAFILFFFAGRLFFPVMVTARKKFYSKVSHAPPKDK*

>MpCSEP14

ATGCTCGGCCTTTCAACCTTTTTCTTTGCCCTCGCAGTCGCTACCTCTGCCTCTGCTTCTCTCGTCTCCCGTGATGGTGGC
GATAGCATGAAAATCAATGTTTCAGGTCCCACTTTCAACGTTAGCCATGGTAACGCAGGCCTCTCGGCTTCTGGCGTCCTC
AATCCCTTTGGCCAGATTGGGATTGGGTGTGGTGTAGCTATGATGCCGCCAGCTCCGGCAAAGACGCTCAGTATGGTGGC
GGTCTCAATGCAGGAACACCCAGCTACGGTCTAGGTGGTGGATTTCGTATTTCCGACAAGAATATGACTGTTGGATTTGGT
ATTGGCGCACCGAAAAACGTTTCTGCAGATACCAATGTCGTTATCACCAAAGAGGGAAAAATCGATTTCACCTTTGTGTCC
ACCCAACCTATCACCTGTGCCGATACCACTATTGGTGGGAAGAAAGGCGTCAAGTGTACTATCGCCTCCAAG**TAA**

>

MLGLSTFFFALAVATSASASLVSRDGGDSMKINVSGPTFNVSHGNAGLSASGVLPFGQIGIGCGVSYDAASSGKDAQYGG
GLNAGTPSYGLGGGFAISDKNMTVGFIGAPKNVSADTNVITKEGKIDFTFVSTQPITCADTTIGGKKGVKCTIASK

>MpCSEP15

ATGCGCCTAACTCCTATCCTTCTCATCGGGCTTGCCTTTGTCCAAGCTTCTATCGCCGCAAAAACCTTGTGTGCGTATGT
CATTCGAGTCAAAAGACCCGGGATGGGTGCAAACAGATGGAAGCGGACTCGAAGAAGAATTCAAGACCCGGTGTGTGTTCAA
TATGACAACTCCAAGGGCAAAACTTCCGCTGTGTGCCGCATCTCTGCACACTACGTAACCTACCTAAGAAAGAAACAGAAT
GTTGGTTGGACCACCTGGTGGCTGAACCAGTGCCACCTAAAAATGGGCTGTGTGAATGGGCCGCC**TAG**

>

MRLTPILLIGLAFVQASIAAKTLVCVCHSSQKTRDGCKQMEADSKKNSRPGVVQYDNSKGKTSAVCRISAHYVTYLRKKQN
VGWTTWWLNQCHPKNGLCEWAA

>MpCSEP16

ATGAAGTTCACCACCGCTCTCCTTTCTGCTCTCACATTCTCAGCAATCTCTTTTCGCTCAGCAAGCCAGCATAACTGAGCCT
CAGGATGGGGCAAATATTGAGGTTGGATGTCAAGTCAAGATCCGTGTTGAAACAAGACAAATAGCATTCATCGATTGCCCC
GATGCGAAGATCACCCATCTCGGCCTCGTGCTCGATATTTAGCCAAGGAGTCCAACACCAACCAGAATATCACACAGATT
CTGTACACAGGAGCCTACGAGCCCAAGGCTCCTGATTACTACTCAAATCTAACGGAGGAGTTTACGGTGACCATTCCAGCC
GACTTTAGAGCTGGTCCGGCAGTGCTTAGGTCGTACTACTTTGGTGTAGTAGAGGTGGAGGCTTTGGATCGTGTGGCGTAC
AACACCTCGAGCGTCAACGTCAATGTTGTCAAGGCG**TAA**

>

MKFTTALLSALTFSAISFAQQASITEPQDGANIEVGQVKIRVETRQIAFIDSPDAKITHLGLVLDISAKESNTNQNTQI
LYTGAYEPKAPDYYSNLTEEFTVTIPADFRAGPAVLRSYYFGVVEVEALDRVAYNTSSVNVNVVKA*

The Fungal Pathogen *Moniliophthora perniciosa* Has Genes Similar to Plant PR-1 That Are Highly Expressed during Its Interaction with Cacao

Paulo J.P.L. Teixeira^{1,2}, Daniela P.T. Thomazella^{1,2}, Ramon O. Vidal^{1,2}, Paula F.V. do Prado¹, Osvaldo Reis¹, Renata M. Baroni^{1,3}, Sulamita F. Franco¹, Piotr Mieczkowski⁴, Gonçalo A.G. Pereira^{1,2}, Jorge M.C. Mondego³

1 Departamento de Genética, Evolução e Bioagentes, Universidade Estadual de Campinas, Campinas, São Paulo, Brazil, **2** Laboratório Nacional de Biotecnologia, Campinas, São Paulo, Brazil, **3** Centro de Pesquisa e Desenvolvimento em Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, Campinas, São Paulo, Brazil, **4** Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America

Abstract

The widespread SCP/TAPS superfamily (SCP/Tpx-1/Ag5/PR-1/Sc7) has multiple biological functions, including roles in the immune response of plants and animals, development of male reproductive tract in mammals, venom activity in insects and reptiles and host invasion by parasitic worms. Plant Pathogenesis Related 1 (PR-1) proteins belong to this superfamily and have been characterized as markers of induced defense against pathogens. This work presents the characterization of eleven genes homologous to plant PR-1 genes, designated as MpPR-1, which were identified in the genome of *Moniliophthora perniciosa*, a basidiomycete fungus responsible for causing the devastating witches' broom disease in cacao. We describe gene structure, protein alignment and modeling analyses of the MpPR-1 family. Additionally, the expression profiles of MpPR-1 genes were assessed by qPCR in different stages throughout the fungal life cycle. A specific expression pattern was verified for each member of the MpPR-1 family in the conditions analyzed. Interestingly, some of them were highly and specifically expressed during the interaction of the fungus with cacao, suggesting a role for the MpPR-1 proteins in the infective process of this pathogen. Hypothetical functions assigned to members of the MpPR-1 family include neutralization of plant defenses, antimicrobial activity to avoid competitors and fruiting body physiology. This study provides strong evidence on the importance of PR-1-like genes for fungal virulence on plants.

Citation: Teixeira PJL, Thomazella DPT, Vidal RO, Prado PFV, Reis O, et al. (2012) The Fungal Pathogen *Moniliophthora perniciosa* Has Genes Similar to Plant PR-1 That Are Highly Expressed during Its Interaction with Cacao. PLOS ONE 7(9): e45929. doi:10.1371/journal.pone.0045929

Editor: Dee A. Carter, University of Sydney, Australia

Received: March 21, 2012; **Accepted:** August 27, 2012; **Published:** September 20, 2012

Copyright: © 2012 Teixeira et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by funds from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), 2006/53553-3, 2007/50062-0, 2009/51018-1 and 2009/50119-9. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: goncalo@unicamp.br

† These authors contributed equally to this work.

Introduction

The basidiomycete fungus *Moniliophthora perniciosa* is the causative agent of witches' broom disease (WBD) in cacao. This devastating disease is responsible for large losses in cacao plantations in the Americas and is a potential threat to other cacao-growing areas throughout the world [1,2]. *M. perniciosa* displays a hemibiotrophic lifestyle, with sequential biotrophic (infective) and necrotrophic stages in the plant. These two mycelial stages are morphologically distinct: whereas the biotrophic mycelium is monokaryotic, the necrotrophic stage is dikaryotic and presents clamp connections for nuclei transfer.

The disease cycle initiates when fungal basidiospores infect meristematic tissues of cacao—such as shoots, fruits and floral cushions—where they germinate and develop as biotrophic monokaryotic hyphae. *M. perniciosa* does not use any specialized infection structure to enter the plant (i.e. apressorium), as observed for the majority of biotrophic and hemibiotrophic fungi [3]. This fungus enters the host tissues through stomata or wounds and colonizes the plant apoplast as thick monokaryotic hyphae. In this

stage of the disease, the parasitic fungus causes drastic morphophysiological alterations in the host, resulting in the formation of hyperplastic and hypertrophic stems, known as green brooms. During the disease progression, the pathogen switches to its necrotrophic dikaryotic stage, which parallels the death of the infected plant tissue. In this phase of WBD, known as dry broom, *M. perniciosa* colonizes the dead plant and can be found in the intracellular spaces of cacao. After alternating wet and dry periods, the fungus produces basidiomata that release basidiospores, reinitiating the disease cycle [1,2].

During recent years, efforts have been directed to develop a solution to control this disease. In 2000, the WBD genome initiative (www.lge.ibi.unicamp.br/vamoura) was launched and, since then, it has supported several molecular and biochemical studies involving both the pathogen and the plant [4–9]. With the recent technological advances in the area of DNA sequencing, transcriptomes representing a variety of growth and developmental conditions of *M. perniciosa*—including transcriptomes of the fungus developing in plants—were sequenced using the RNA-seq technology. As a result, a comprehensive database named WBD

Transcriptome Atlas has been constructed, and has contributed important information on the molecular basis of the *M. perniciosa*-cacao interaction (Teixeira *et al.*, manuscript in preparation).

The establishment of a disease process depends on the ability of the pathogen to overcome or neutralize plant defenses and then initiate a parasitic relationship with its host. However, to halt pathogenic colonization, plants have developed an arsenal of defense responses, which include induction of pathogenesis-related (PR) genes [10], production of secondary metabolites as well as reinforcement of cell walls. Also, usually triggered by the recognition of a pathogen attack, plants produce highly toxic radicals, such as nitric oxide and reactive oxygen species, which can lead to the establishment of a local cell death (the hypersensitive response, HR). Among the induced pathogenesis-related genes, PR-1s have been frequently identified and used as markers of plant defense responses [10]. Notably, they were shown to have microbicide activity against oomycetes and fungi [11–14].

PR-1 proteins are members of a superfamily named SCP/TAPS (Sperm-Coating Protein/Tipx-1/Ag5/PR-1/Sc7) or CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1). This superfamily has members throughout the eukaryotic kingdom, suggesting an important role for these proteins in the biology of eukaryotes [15,16]. Thus far, only a single report has shown the existence of enzymatic activity for a SCP/TAPS protein [17]. The protein Tex31 of the predatory marine mollusk *Conus textile* showed serine-proteolytic activity against a specific pro-peptide precursor of a venom toxin [17]. In addition, structural analyses indicated that four highly conserved amino acids (two histidines and two glutamates) form the putative catalytic site of SCP/TAPS proteins [16–22]. Although the existence of biochemical activity has not been shown for any other SCP/TAPS proteins, they are associated with various biological processes, such as male reproductive tract development [23,24], immune responses in plants and animals [25], venom activity of reptiles and insects [26–29] and host invasion by parasites [30–32].

In fungal species, SCP/TAPS proteins have been studied in *Saccharomyces cerevisiae*, in which they are highly expressed under nutrient starvation conditions [33]. In the basidiomycete *Schizophyllum commune*, SCP/TAPS proteins have been associated with fruiting body formation [34]. Interestingly, in the ascomycetes *Candida albicans* and *Ficarium oxysporum*, deletion of a SCP/TAPS gene impaired virulence on animals, indicating a role for this class of genes in fungal pathogenicity [35,36]. Considering that PR-1 proteins are widespread markers of the induced defense response in plants, what would be the function of their homologs in a plant pathogenic fungus, such as *M. perniciosa*?

This article describes the identification of a SCP/TAPS family in the *M. perniciosa* genome, the analysis of structural features of these genes, and their expression profile throughout *M. perniciosa* development. *M. perniciosa* SCP/TAPS proteins were modeled, and some structural differences were revealed among them. Based on these results, we present a hypothetical model in which SCP/TAPS proteins play a role in *M. perniciosa* pathogenicity by interfering with the defense system of cacao plants.

Results

Characterization of the PR-1 gene family in *M. perniciosa*

Annotation of a genome draft of *M. perniciosa* [7], and inspection of fungal EST libraries [6,8] identified four PR-1-like genes in this pathogen (*MpPR-1a* to *MpPR-1d*). Later, improvements in the genome assembly obtained with next generation sequencing data (unpublished data) allowed the identification of seven additional members of the *MpPR-1* family (*MpPR-1e* to *MpPR-1k*), totaling

eleven PR-1-like genes in *M. perniciosa*. These members are very heterogeneous in size and gene structure, with coding sequences (CDS) ranging from 447 to 1,152 nucleotides and intron composition varying between two to five introns. Three of these genes (*MpPR-1c*, *MpPR-1d* and *MpPR-1j*) are organized in tandem. Sequence details of these eleven genes are shown in Table 1, and their respective structures (exon-intron positions) are depicted in figure S1.

Hydrophobic signal peptide sequences predicted with the TargetP program [37] were identified in all eleven *MpPR-1* sequences (NN score >0.80), strongly suggesting that these proteins are secreted. Additionally, all *MpPR-1* proteins showed a single SCP/TAPS domain (InterPro ID IPR014044), as predicted by the InterProScan server [38]. These domains were approximately 130 amino acids in length and ranged between 34% and 82% of the total amino acid sequence of an individual *MpPR-1* protein (Fig. 1A). In addition to the SCP/TAPS domain, *MpPR-1b* and *MpPR-1g* present N-terminal and C-terminal extensions, respectively. No other InterProScan predicted domain was identified in these extensions or in any of the *MpPR-1* proteins described. Interestingly, a careful manual inspection revealed that the C-terminal extension of *MpPR-1g* is rich in residues of lysine (K) and glutamic acid (E) that are mostly organized in alternating positions, resulting in the formation of a “KEKE” motif [39] (Fig. 2). The *MpPR-1b* N-terminal extension is also a low complexity region, being rich in serine, threonine and proline residues. However, no described motif could be recognized.

Alignment of the amino acid sequences encoded by all *MpPR-1* genes revealed significant similarity only over the SCP/TAPS domains (Fig. 1B). The amino acids proposed to form the putative catalytic site of SCP/TAPS proteins (two histidines and two glutamic acids, shown in red) were identified in six *MpPR-1s* (*MpPR-1b*, *MpPR-1c*, *MpPR-1d*, *MpPR-1e*, *MpPR-1h* and *MpPR-1j*). In contrast, *MpPR-1a* presented two, *MpPR-1f* had only one and *MpPR-1g*, *MpPR-1i* and *MpPR-1k* had none of the four conserved residues (Fig. 1B).

Protein structure modeling

To explore the tertiary structural characteristics within the SCP/TAPS domains of the *MpPR-1* family members, we created homology models using the fold prediction metasever I-TASSER (Fig. 3). The derived models indicated that the *MpPR-1* SCP/TAPS domains adopt the α - β - α sandwich conformation, which is common to all superfamily members across the species studied [16,18] (Fig. 3A). Furthermore, all *MpPR-1* proteins possess the large cleft proposed to constitute the SCP/TAPS active site. As mentioned above, six *MpPR-1* proteins (*MpPR-1b*, *MpPR-1c*, *MpPR-1d*, *MpPR-1e*, *MpPR-1h* and *MpPR-1j*) have the four conserved residues of the putative catalytic site. These residues are localized within this cleft (Fig. 3B) and are found in the same direction of the orthologous amino acids identified in previous SCP/TAPS crystal structures [18,21,22,25]. These residues are lacking (*MpPR-1g*, *MpPR-1i* and *MpPR-1k*) or partially absent (*MpPR-1a* and *MpPR-1f*) in other members of the *MpPR-1* family (Fig. 3C), indicating some diversification in their mode of action.

Remarkably, *MpPR-1b* and *MpPR-1g* contain two modules: the SCP/TAPS domain and either N-terminal (*MpPR-1b*) or C-terminal (*MpPR-1g*) extensions. Protein models indicate that such extensions are structurally organized and have α -helix conformations (Fig. 3A). These additional regions possibly confer a different activity or regulation to these proteins.

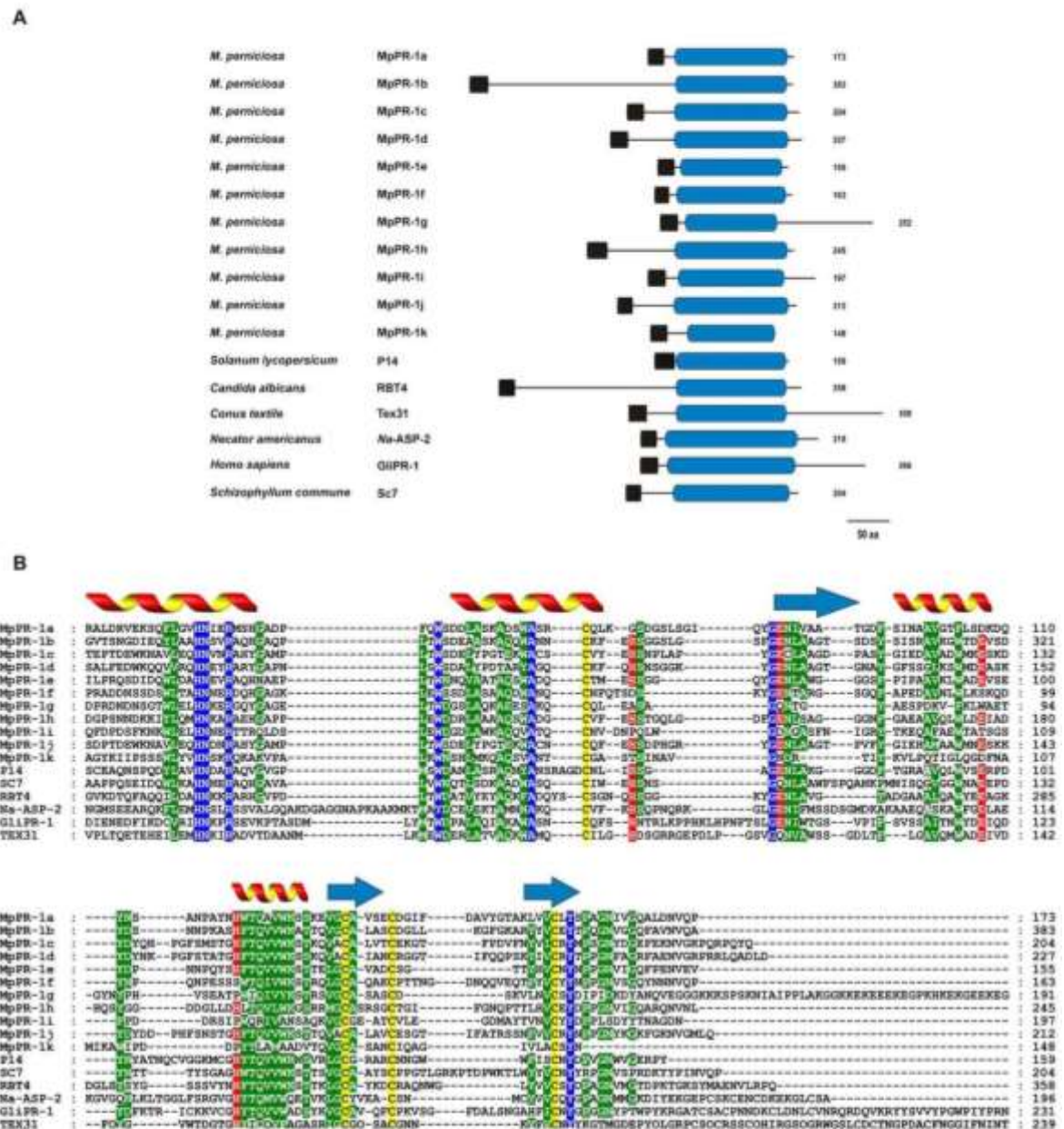


Figure 1. Comparison of MpPR-1 and SCP/TAPS proteins of representative organisms. (A) Domain arrangement of SCP/TAPS proteins. Hydrophobic signal peptides are shown in black and SCP/TAPS domains are represented in blue. The numbers on the right show the size of each protein. Large N-terminal and C-terminal expansions are observed in MpPR-1b and MpPR-1g, respectively. (B) Alignment of the conserved domain of SCP/TAPS proteins. In general, the SCP/TAPS superfamily members show similarities only over the SCP/TAPS domain. Conserved residues (100% of identity) are shown in blue and semi-conserved residues (at least 60% of identity) in green. Putative active site residues are highlighted in red and cysteines in yellow. Secondary structure elements are shown above the alignment (arrow: β -sheets; helix α -helices). P14, tomato PR-1 (GenBank P04284); RBT4, repressed by TUP1 from *Candida albicans* (GenBank AAG09789); Tex31, SCP/TAPS from the mollusk *Conus textile* (GenBank CAD36507); Na-ASP-2, *Necator americanus* secreted protein (GenBank AAP41952); GliPR-1, human glioma PR-1 protein (GenBank P48060); Sc7, SCP/TAPS from the basidiomycete *Schizophyllum commune* (GenBank P35794).

doi:10.1371/journal.pone.0045929.g001

Table 1. Characteristics of the eleven MpPR-7 genes identified in the *M. perniciosa* genome.

Gene name	GenBank accession number	CDS size (bp)	Number of introns	Protein size (aa)	Signal peptide	BlastX First Hit (Swissprot)*	BlastP First Hit (NCBI-NC)*
MpPR-7a	JN620340	522	2	173	Yes (NN score = 0.894)	SC14 (1e-19) Schizophyllum commune	XP_001876593.1 - predicted protein (1e-54) Laccaria bicolor
MpPR-7b	JN620341	1152	5	383	Yes (NN score = 0.844)	SC7 (3e-23) Saccharomyces cerevisiae	XP_001873270.1 - predicted protein (2e-63) Laccaria bicolor
MpPR-7c	JN620342	615	4	204	Yes (NN score = 0.898)	PRY1 (1e-28) Saccharomyces cerevisiae	XP_001889714.1 - predicted protein (2e-48) Laccaria bicolor
MpPR-7d	JN620343	684	5	227	Yes (NN score = 0.951)	PRY1 (3e-31) Saccharomyces cerevisiae	XP_001889714.1 - predicted protein (8e-66) Laccaria bicolor
MpPR-7e	JN620344	468	3	155	Yes (NN score = 0.919)	PR1C (4e-35) Nicotiana glauca	XP_003038858.1 - hypothetical protein (2e-45) Schizophyllum commune
MpPR-7f	JN620345	492	3	163	Yes (NN score = 0.947)	SC7 (4e-25) Schizophyllum commune	CCA68.48 - related to PRY1 (8e-42) Puccinia striiformis
MpPR-7g	JN620346	759	4	252	Yes (NN score = 0.905)	SC7 (4e-13) Schizophyllum commune	EFY9592.1 - hypothetical protein (4e-15) Mitsurinaea anisopliae
MpPR-7h	JN620347	738	5	245	Yes (NN score = 0.804)	PR1B (5e-24) Nicotiana glauca	XP_001828886.1 - hypothetical protein (3e-41) Cupressus nirens
MpPR-7i	JN620348	498	3	165	Yes (NN score = 0.940)	SC7 (4e-05) Schizophyllum commune	EGO02028.1 - hypothetical protein (3e-10) Serpula lacrymans
MpPR-7j	JN620349	639	4	212	Yes (NN score = 0.918)	PRY3 (8e-30) Saccharomyces cerevisiae	XP_001889714.1 - predicted protein (2e-54) Laccaria bicolor
MpPR-7k	JN620350	447	2	148	Yes (NN score = 0.944)	P14 (2e-03) Solanum lycopersicum	XP_002578075.1 - venom allergen 21 (4e-09) Schistosoma mansoni

*e values are shown in parentheses.
doi:10.1371/journal.pone.0045929.t001

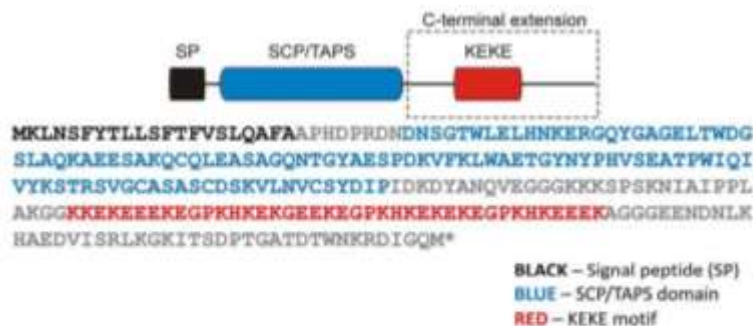


Figure 2. Domains identified in the MpPR-1g protein. In addition to the SCP/TAPS domain, this protein has a KEKE motif in its C-terminal extension. This motif is known to mediate the interaction with other proteins or ions.
doi:10.1371/journal.pone.0045929.g002

Expression profile of MpPR-1 genes throughout *M. perniciosa* development

The expression pattern of MpPR-1 family members was assessed by quantitative real time PCR (qPCR) throughout the different developmental stages of *M. perniciosa*. As shown in Figure 4, each MpPR-1 gene showed a specific expression profile: expression of MpPR-1a, MpPR-1b, MpPR-1d and MpPR-1i genes was relatively uniform, with nearly similar expression values for most conditions

analyzed. In contrast, MpPR-1e was up-regulated in the dikaryotic hyphae, whereas MpPR-1j was exclusively expressed in mushrooms (basidiomata). Strikingly, five MpPR-1 genes (MpPR-1c, MpPR-1f, MpPR-1g, MpPR-1h, and MpPR-1k) were highly expressed during the biotrophic interaction of the fungus with the cacao plant (green broom stage of WBD). These genes were poorly expressed in the necrotrophic hyphae (dry broom stage of

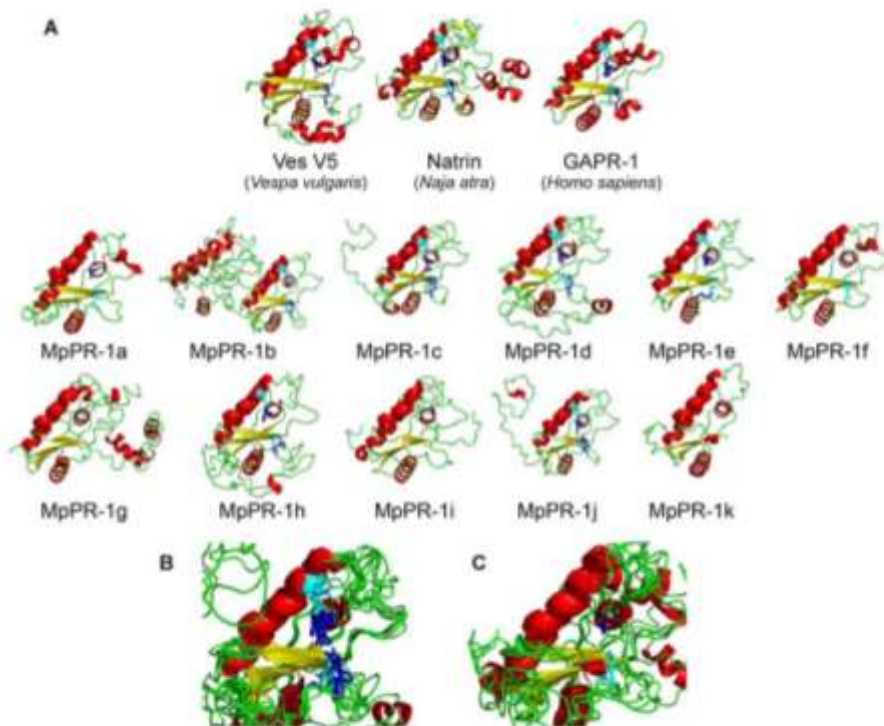


Figure 3. Homology modeling of MpPR-1 proteins. (A) Ribbon stick representation showing the folding of eleven MpPR-1 proteins and three SCP/TAPS proteins used to obtain these models. The putative residues forming the catalytic site are highlighted in dark blue (histidines) and light blue (glutamic acids). Note the presence of an additional protein module in MpPR-1b and MpPR-1g. These modules respectively correspond to the N-terminal and C-terminal extensions observed in these proteins. (B) MpPR-1b, MpPR-1c, MpPR-1d, MpPR-1e, MpPR-1h and MpPR-1j have the four putative active site residues of the SCP/TAPS domain. (C) These residues are partially or completely absent in MpPR-1a, MpPR-1f, MpPR-1g, MpPR-1i and MpPR-1k.
doi:10.1371/journal.pone.0045929.g003

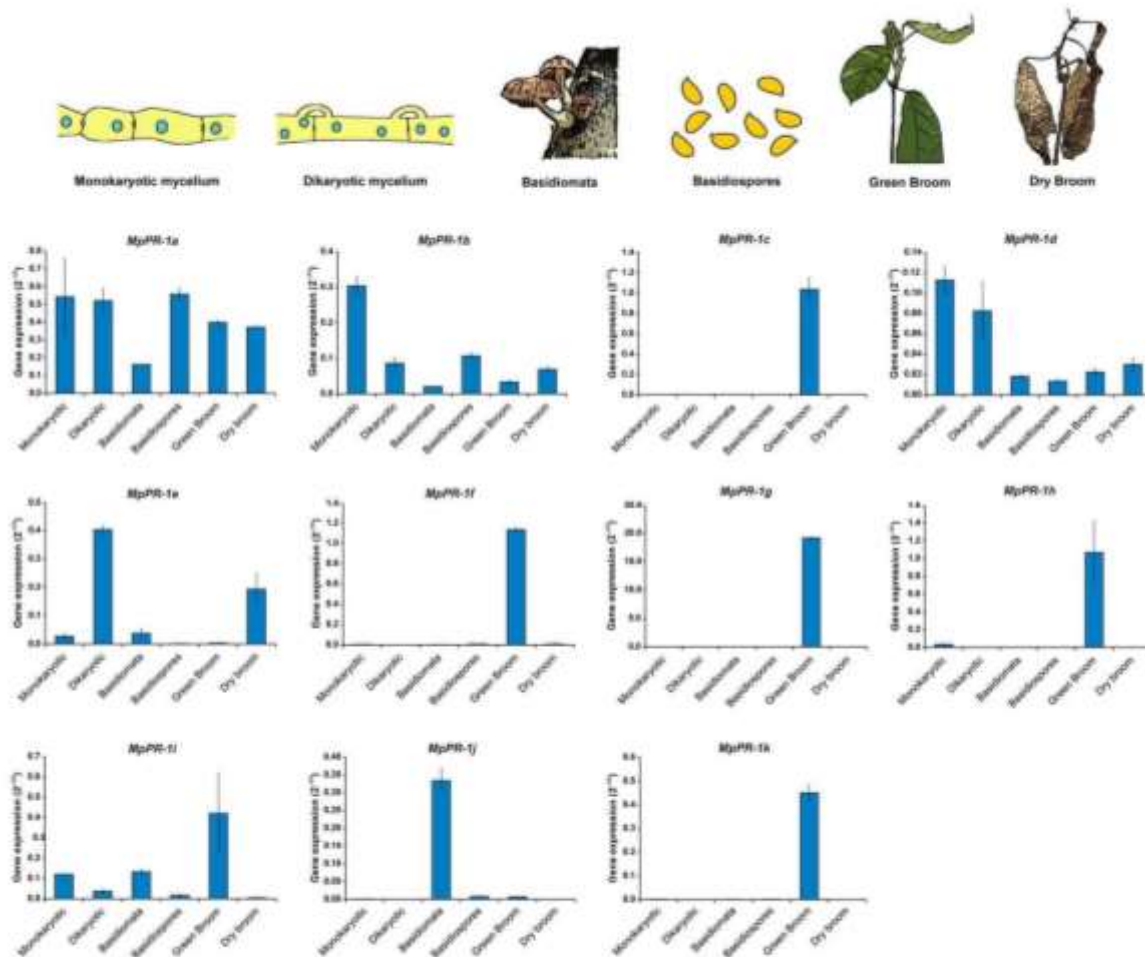


Figure 4. Transcriptional profile of *MpPR-1* family members throughout the *M. perniciosa* life cycle. Each *MpPR-1* gene has a distinct expression profile during fungal development. "Monokaryotic" and "Dikaryotic" hyphae represent the two mycelial stages (biotrophic and necrotrophic) grown under *in vitro* conditions. "Green broom" and "dry broom" correspond to the biotrophic and necrotrophic stages of *M. perniciosa*, respectively, during its interaction with cacao. Analyses were performed by qPCR and the *M. perniciosa* β -actin gene was used as endogenous control to normalize data. Error bars represent standard deviations determined with two biological replicates. Representative drawings of the conditions analyzed are shown on the top.
doi:10.1371/journal.pone.0045929.g004

WBD) and in the *ex planta* conditions, suggesting a specific role for the encoded proteins in fungal pathogenicity.

Characterization of an *MpPR-1* cluster

As mentioned above, genes *MpPR-1c*, *MpPR-1d* and *MpPR-1j* are arranged in tandem over a region of approximately 5 kbp in the *M. perniciosa* genome (Fig. 5). This gene cluster points to the occurrence of gene duplication events during the evolution of the *MpPR-1* family. Indeed, the proteins encoded by these three genes are more similar to each other than to other *MpPR-1* members (data not shown). In accordance, these three genes have very similar structures, with a minor difference in *MpPR-1d*, which has an additional intron and a mini-exon following exon 2 (Fig. S1). Importantly, genes *MpPR-1c* and *MpPR-1j* are also highly similar at the nucleotide level (84% identity), indicating a recent event of gene duplication. Despite their similarity, these genes have distinct expression profiles: whereas *MpPR-1j* is highly expressed in

basidiomata, *MpPR-1c* is mainly expressed during cacao infection (green broom stage of WBD) (Figs. 4 and 5).

Discussion

In this study, we identified a family of genes encoding proteins of the SCP/TAPS superfamily in the plant pathogen *Moniliophthora perniciosa*. SCP/TAPS proteins are found in a vast number of organisms, including plants, insects, mammals, fungi, mollusks and worms. Plant Pathogenesis-related proteins (PR-1) belong to this superfamily and are known to accumulate after pathogen invasion [10]. Despite being broadly spread, evidence for the importance of fungal SCP/TAPSs in plant-pathogen interactions has not yet been described.

M. perniciosa has a larger number (11) of SCP/TAPS genes encoding PR-1-like secreted proteins than other fungal species analyzed to date (Table S1). Although these proteins have a single

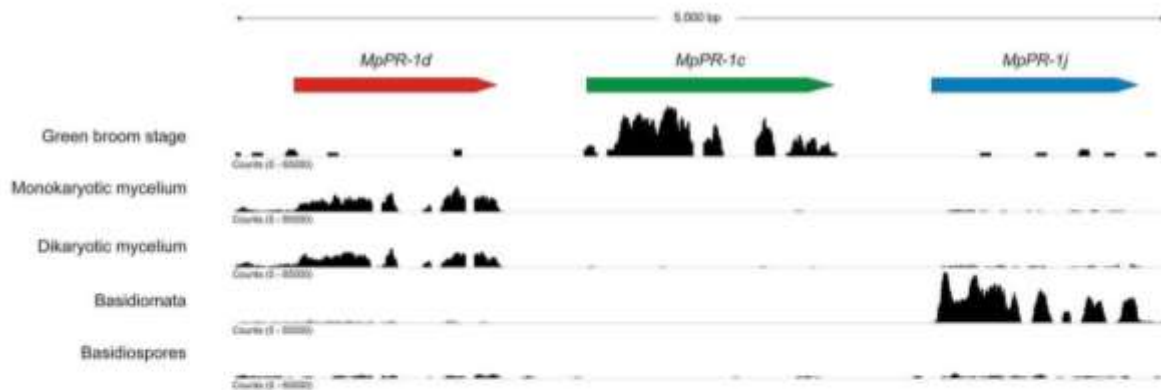


Figure 5. Genomic organization and transcriptional profile of the *MpPR-1* gene cluster found in *M. perniciosa*. The *MpPR-1c*, *MpPR-1d* and *MpPR-1j* genes are arranged in tandem over a region of approximately 5 kbp. Analysis of the WBD RNA-seq Atlas shows the expression profile of these *MpPR-1* genes in different conditions (green broom – *in planta* development of the biotrophic monokaryotic hyphae; monokaryotic mycelium; dikaryotic mycelium; basidiomata and basidiospores). Data were visualized using the Integrative Genomics Viewer [62]. The black coverage plot shows cumulative RNA-seq read coverage along the transcripts in all different conditions. Note that these genes were named according to the order they were identified in the fungal genome, and the nomenclature does not necessarily reflect their relative localization in the genome. doi:10.1371/journal.pone.0045929.g005

SCP/TAPS domain, they are very divergent in sequence (Fig. 1B). Some of them show extensions in their N-terminal (*MpPR-1b*) or C-terminal (*MpPR-1g*) regions (Fig. 1A and Fig. 3A). Also, six *MpPR-1*s (*MpPR-1b*, *MpPR-1c*, *MpPR-1d*, *MpPR-1e*, *MpPR-1h* and *MpPR-1j*) have all four conserved amino acids (two histidines and two glutamates) of the active site proposed for SCP/TAPS proteins, whereas the other five proteins (*MpPR-1a*, *MpPR-1f*, *MpPR-1g*, *MpPR-1i* and *MpPR-1k*) lack the catalytic tetrad (Fig. 1B and Fig. 3). However, the basic signatures of SCP/TAPS domains remain conserved.

The high number of genes and the variation in protein sequence of *MpPR-1* members may reflect a process of diversification of this gene family in *M. perniciosa*. In accordance, the cluster including three *MpPR-1* members (Fig. 5) indicates the occurrence of recent events of gene duplication, which are central for the generation of genetic variability. In this regard, functional diversity may occur within this family, and different *MpPR-1*s likely play a particular role in *M. perniciosa* biology.

Although widely distributed among several species, the functions of SCP/TAPS proteins are still uncertain. Their broad distribution indicates that they play a role in a plethora of biological processes. In the mushroom-forming basidiomycete *Schizophyllum commune*, SCP/TAPS proteins were identified in the basidiomata, being involved in the formation of pseudo-parenchymous tissue of this reproductive structure [34]. In a similar way, *MpPR-1j* is exclusively expressed in the basidiomata of *M. perniciosa*, suggesting a role for this isoform in the physiology/metabolism of fruiting bodies. In contrast, we also identified SCP/TAPS genes in the genomes of some basidiomycetes that do not produce mushrooms (e.g., *Ustilago maydis*, *Puccinia graminis*, *Melampsora spp* and *Cryptococcus spp*) (data not shown). Therefore, it is likely that SCP/TAPS proteins have functions in basidiomycetes other than fruiting body development/physiology.

In response to pathogen invasion, plants typically produce PR-1 proteins, which have antimicrobial activity and, consequently, are able to inhibit the development of fungi and oomycetes [11–13]. Considering that, we hypothesize that some *MpPR-1*s could play a role in limiting the growth of other microbial competitors (e.g., oomycetes from the genus *Phytophthora*, responsible for causing black pod rot in cacao), thus favoring *M. perniciosa* colonization

during WBD progression. Functional experiments are needed to confirm the existence of antimicrobial activity in any of the *MpPR-1* proteins identified.

Gene expression analyses revealed that five *MpPR-1* genes (*MpPR-1c*, *MpPR-1f*, *MpPR-1g*, *MpPR-1h* and *MpPR-1k*) are strikingly up-regulated in the green broom stage of WBD, when the fungus grows biotrophically within the plant tissues (Fig. 4). Remarkably, inspection of the WBD RNA-seq Transcriptome Atlas (Teixeira *et al.*, manuscript in preparation) revealed that *MpPR-1g* and *MpPR-1h* are among the most highly expressed genes of *M. perniciosa* during its biotrophic interaction with cacao. Moreover, in addition to the green broom stage, *MpPR-1f* and *MpPR-1h* are notably expressed in germinating basidiospores (Fig. S2), a critical stage for the establishment of infection. None of these genes were significantly expressed in non-germinating spores or in the dry broom stage of WBD (Fig. 4 and Fig. S2), strongly indicating a major role for the encoded proteins in the infective (biotrophic) stage of *M. perniciosa*. Similarly, SCP/TAPS genes identified in some animal parasitic worms (*Schistosoma mansoni*, *Brugia malayi*, *Necator americanus* and *Ancylostoma caninum*) are highly expressed in the infective stage and are considered important pathogenicity factors [30,31,40–42]. In these parasites, SCP/TAPSs are supposed to contribute to their virulence by modulating the host immune response [15,16].

Recently, a SCP/TAPS protein in the plant-parasitic nematode *Globadara rastrochionis* (Gr-VAP1) was shown to function as an effector by interacting with the tomato cysteine protease Rcr3, which is also a target of the Avr2 effector from the fungus *Cladosporium fulvum* [43]. In addition, SCP/TAPS proteins have been identified in other plant infecting nematodes (e.g. *Heterodera glycines*, *Meloidogyne incognita* and *Bursaphelenchus xylophilus*), and these are thought to be required for the establishment of parasitism [44–48]. Considering the expression pattern of some SCP/TAPS genes in *M. perniciosa* and the functions ascribed for the encoded proteins in other pathogenic organisms, it is plausible that some *MpPR-1*s play a role in the *M. perniciosa*-cacao interaction and may be candidate effectors of this fungal pathogen. Accordingly, a recent study that aimed at the identification of putative effectors in *Melampsora larici-populina* and *Puccinia graminis* reported the enrichment of SCP/TAPSs in the predicted secretome of these rust fungi

[49]. A fungal SCP/TAPS gene was also identified in EST libraries produced from rye infected with the ascomycete pathogen *Claviceps purpurea* [50], suggesting that these genes might also be important in other plant-fungus interactions.

In recent years, a role for SCP/TAPSs as virulence factors has emerged in many organisms. It is likely that these proteins converged as pathogenicity mechanisms in distinct pathogens/parasites from either plants or animals. Whereas the function of fungal SCP/TAPSs as virulence factors in plant pathogens remains to be confirmed, previous studies demonstrated that these proteins are required for fungal virulence on animals (e.g. *C. albicans* and *F. oxysporum*) [35,36]. The ascomycete *F. oxysporum* is a multi-host pathogen that is able to infect both plants and animals. Previous work by Prados-Roales *et al.* verified that *ffr1*, one of the six SCP/TAPS genes from this pathogen, is required for fungal virulence on animals but not on plants [36]. Given that the *F. oxysporum* genome contains five other SCP/TAPS genes, the absence of a phenotype on plants can be explained by the occurrence of functional redundancy in this gene family. Notably, there is evidence that *ffr1* is part of a gene family that has expanded in *F. oxysporum* and in other plant pathogenic Sordariomycetes [36].

Although the precise activity of SCP/TAPSs is currently unknown, Prados-Roales *et al.* presented the first genetic evidence for a biological function of the proposed active site of SCP/TAPS proteins [36]. The authors demonstrated that the integrity of the active site is required for *F. oxysporum* virulence on animals. In *M. perniciosa*, six MpPR-1s (MpPR-1b, MpPR-1c, MpPR-1d, MpPR-1e, MpPR-1h and MpPR-1j) contain all four amino acids of the proposed active site. In contrast, the other five proteins do not have the complete catalytic tetrad (Fig. 1B). In this regard, it is possible that *M. perniciosa* PR-1s have distinct mode of actions. For instance, whereas those proteins with the complete catalytic tetrad can function as enzymes, the other PR-1s may act as inhibitors.

Concomitantly to the up-regulation of some MpPR-1 genes in *planta*, we identified a cacao PR-1 gene over-expressed in the green broom stage of WBD (Fig. S3). Based on these findings, we suggest that some MpPR-1s could act as competitive inhibitors of the plant PR-1, modulating the cacao immune response. It has already been shown that the SCP/TAPS protein *Na-ASP-2* of the hookworm *Necator americanus* has a high structural similarity to chemokines, and this protein is proposed to be an antagonistic ligand of receptors that activate the immune system of the vertebrate host [51]. Furthermore, NIF (Neutrophil inhibitory factor), a SCP/TAPS protein from *Angylostoma caninum*, interferes with the host immune system by interacting with neutrophil receptors [32]. Confirmation of this interesting mechanism in the *M. perniciosa*-cacao interaction may be of primary relevance to the understanding of many other plant diseases and will shed light on our understanding of PR-1 functions.

Among the MpPR-1 genes that are highly expressed in *planta*, MpPR-1g is the only one with a C-terminal extension in addition to the SCP/TAPS domain (Fig. 1A and Fig. 3A). This additional region is rich in lysine (K) and glutamic acid (E) residues, which are mostly organized in alternating positions, resulting in the formation of a "KEKE" motif [39] (Fig. 2). This motif is known to mediate protein-protein associations [39,52] and is also able to bind divalent ions, such as calcium and zinc [53]. Calcium is an important regulator of many cellular processes, including plant defense responses [54]. In this regard, this additional module may be important in determining the mode of action of MpPR-1g. Whether this protein interacts with other proteins, particularly cacao proteins, and/or interferes with the plant calcium signaling during infection should be the object of future studies.

Overall, this study presents important evidence on the role of fungal SCP/TAPSs in the context of a plant-pathogen interaction. Although the precise function of each MpPR-1 family member is currently unknown, the information provided in our study suggests they have potential roles in some important biological processes, such as fruiting body metabolism, spore penetration and modulation of the host defense response. As a consequence, our results may inform the study of the role of PR-1-encoding genes in other organisms, particularly phytopathogens. Further studies concerning the *M. perniciosa* PR-1 gene family will focus on the characterization of this interesting family in terms of fungal development and roles in the *M. perniciosa* interaction with cacao.

Materials and Methods

Biological material

Isolate CP02 of *Moniliophthora perniciosa* (Stahel) Almeida & Phillips-Mora [55], was used to perform the experiments. Under *in vitro* conditions, the fungus can only be maintained as a dikaryotic mycelium, and all other developmental stages (basidiomata, basidiospores and monokaryotic mycelium) are obtained from the dikaryotic stage. The reproductive structures (basidiomata) were produced in laboratory according to the protocol described by Pires *et al.* [8]. Fresh basidiomata were used to collect basidiospores according to Frias *et al.* [56].

Basidiospores suspensions were utilized for the *in vitro* production of the monokaryotic mycelia. For this purpose, approximately 3.75×10^5 basidiospores were inoculated in 125 ml Erlenmeyer flasks containing 50 ml liquid medium (LMCpL+), as described by Meinhardt *et al.* [57]. Liquid cultures were maintained at 28°C and incubated under agitation at 120 rpm. Dikaryotic mycelium was inoculated in the same medium and maintained under the same conditions. Both mycelia were collected 7 days post inoculation to perform the experiments.

Theobroma cacao L. cv. "Comum" was used to perform the infection experiments. Three-month-old plantlets were inoculated with 30 µL of a basidiospore suspension (1×10^5 spores mL⁻¹) according to the procedure described by Frias *et al.* [56]. Plantlets were kept in a greenhouse under controlled conditions of temperature (26°C) and humidity (>80%). Green brooms (biotrophic stage) and dry brooms (necrotrophic stage) were collected 30 and 105 days post inoculation, respectively.

Sequence analysis

Inspection of the *M. perniciosa* genome led to the identification of eleven genes encoding proteins similar to plant pathogenesis-related proteins 1 (PR-1). These genes were named MpPR-1a to MpPR-1k according to the order they were discovered. The complete open reading frames (ORFs) of these genes were predicted using the program Augustus [58] and confirmed by cDNA sequencing. These sequences have been submitted to GenBank with the accession numbers JN620340 to JN620350. Blast searches were performed using the NCBI-NR and Swissprot databases. Domain prediction of the encoded proteins was performed using the InterProScan server [38] and the presence of a signal peptide for secretion was predicted using the software TargetP 1.1 [37].

Total RNA extraction and cDNA synthesis

With the exception of basidiospores, samples were ground to a fine powder in liquid nitrogen using a pestle and mortar. Basidiospores walls were broken by vortexing the sample in RNA extraction buffer (Buffer RLT, RNeasy Plant Mini Kit) and 200 mg glass beads (0.4–0.6 µm, Sigma-Aldrich, St. Louis, MO,

EUA). RNA isolation was performed using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. RNA was treated with DNase I AmpGrade (Invitrogen, Carlsbad, CA, USA) and its concentration was accessed using the ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA). cDNA was synthesized from 1 µg total RNA using the SuperScript II Reverse Transcriptase (Invitrogen), according to the manufacturer's instructions.

Gene expression assays

Quantitative real time PCR (qPCR) was performed on a StepOne Plus Real Time PCR System (Applied Biosystems, Foster City, CA, USA) using Sybr Green I dye for the detection of PCR products. Each reaction contained 8 µl SYBR Green PCR Master Mix (Applied Biosystems), 250 nM each primer and 50 ng cDNA template in a final volume of 16 µl. No-template reactions were included as negative controls for each set of primers used. The thermal cycling conditions were 94°C for 10 min, followed by 40 cycles of 94°C for 15 s, 53°C for 30 s and 60°C for 1 min, with fluorescence detection at the end of each cycle. In addition, a melting curve analysis was performed to verify the amplification of a single product per reaction. All reactions were conducted in technical triplicates using two independent biological replicates of each sample. The *M. perniciosa* β -actin gene was used to normalize data and expression levels are presented as $2^{-\Delta\Delta C_T}$. Primers used in this assay are shown in Table S2.

Protein structure modeling

The fold recognition-based method was implemented using the I-TASSER server [59], which constructed structure models for each MpPR-1 protein using folds of the most similar proteins deposited in the PDB (Protein Data Bank) database (<http://www.rcsb.org/pdb>). The main templates were based on the structure of three proteins: i) Natrin (PDB 1xta), a component of the venom of the snake *Naja atra*; ii) GAPP-1 (PDB 1amb), a SCP/TAPS protein associated with the membrane of the human Golgi system; and iii) Ves V5 (PDB 1gru), present in the venom of the wasp *Vespula vulgaris*. The modeled structures were validated by analyzing the Ramachandran plots generated by PROCHECK [60], and the models were displayed using the software PyMOL [61].

Supporting Information

Figure S1 Structure of the MpPR-1 genes. Exons are represented by boxes, while introns are shown as lines. Exons are colored to highlight the regions encoding important protein

features: predicted signal peptides (black), SCP/TAPS domain (blue) and the remaining ORF (gray). (TIF)

Figure S2 Expression levels of MpPR-1 genes in germinating and non-germinating basidiospores. MpPR-1f and MpPR-1b are highly expressed in germinating basidiospores, supporting a role for the encoded proteins in the establishment of witches' broom disease. Data are part of the WBD Transcriptome Atlas and were obtained by RNA-seq sequencing. Gene expression values are given in Reads Per Kilobase of exon model per Million mapped reads (RPKM). (TIF)

Figure S3 Gene expression levels of a cacao PR-1 (ID CGD0027635) in infected and healthy plants. Similar to some MpPR-1 genes, a cacao PR-1 (TcPR-1) is up-regulated in the green broom stage of WBD. The analysis was performed by qPCR and the *T. cacao* α -tubulin gene (ID CGD0029727) was used as endogenous control to normalize data. Gene IDs refer to the Cacao Genome Database (<http://www.cacaogenomedb.org>). The qPCR assay was conducted as described in the Material and Methods section and primers used in the experiment were: TcPR-1_F: 5' ACCTTATGCGGAGAACCTTG 3', TcPR-1_R: 5' GGAGTAATCATAGTCGGCCTTC 3', TcTub_F: 5' AC-CAATCTTAACCGCCTTGTCT 3' and TcTub_R: 5' GTTAGTCTGGAACCTCAGTCACAT 3'. (TIF)

Table S1 Number of SCP/TAPS genes in fungal species with different lifestyles. Numbers correspond to the genes coding proteins with the InterPro ID IPR014044. (DOC)

Table S2 Primers used for quantitative real time PCR analyses of *M. perniciosa* PR-1 genes. (DOC)

Acknowledgments

We thank Dr. Halley Caixeta de Oliveira (State University of Campinas) for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: PJPLT JMC. Performed the experiments: PJPLT DPTT ROV PFVP OR RMB SFF JMC. Analyzed the data: PJPLT DPTT ROV JMC. Contributed reagents/materials/analysis tools: PM. Wrote the paper: PJPLT DPTT JMC GAGP.

References

- Meinhart LW, Rincosón J, Bailey BA, Aime MC, Griffin GW, et al. (2006) *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe? Mol Plant Pathol 9: 377–389.
- Purdy LH, Schenck RA (1996) STATUS OF CACAO WITCHES' BROOM: biology, epidemiology, and management. Annu Rev Phytopathol 34: 573–594.
- Perfect SE, Green JK (2001) Infection structures of biotrophic and hemibiotrophic fungal plant pathogens. Mol Plant Pathol 2: 104–108.
- Scarpini LM, Meinhart LW, Mazzareira P, Pomella AW, Schiavastato MA, et al. (2009) Biochemical changes during the development of witches' broom: the most important disease of cacao in Brazil caused by *Moniliophthora perniciosa*. J Exp Bot 36: 865–877.
- García O, Macedo JA, Tiburcio R, Zaparoli G, Rincosón J, et al. (2007) Characterization of necrosis and ethylene-inducing proteins (NEP) in the biotrophic phase of *Moniliophthora perniciosa*, the causal agent of witches' broom in *Theobroma cacao*. Mycol Res 111: 443–452.
- Rincosón J, Scarpini LM, Garasolis MF, Mondago JM, Formighieri EF, et al. (2008) Differential gene expression between the biotrophic-like and saprotrophic mycelia of the witches' broom pathogen *Moniliophthora perniciosa*. Mol Plant Microbe Interact 21: 891–908.
- Mondago JM, Garasolis MF, Costa GC, Formighieri EF, Parizzi LF, et al. (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. BMC Genomics 9: 540.
- Pires AB, Gramacho KP, Silva DC, Gues-Neto A, Silva MM, et al. (2009) Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes. BMC Microbiol 9: 158.
- Thomazella DP, Teixeira PJ, Oliveira HC, Saviani EE, Rincosón J, et al. (2012) The hemibiotrophic cacao pathogen *Moniliophthora perniciosa* depends on a mitochondrial alternative oxidase for biotrophic development. New Phytol 194: 1025–1034.
- Van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. Annu Rev Phytopathol 44: 135–162.
- Rauscher M, Adam AL, Wirtz S, Guggenheim R, Mendgen K, et al. (1999) PR-1 protein inhibits the differentiation of rust infection hyphae in leaves of acquired resistant broad bean. Plant J 19: 625–633.
- Niderman T, Genot E, Bruyere T, Gees R, Stütz A, et al. (1995) Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. Plant Physiol 108: 17–27.

13. Kiba A, Nishihara M, Nakatsuma T, Yamamura S (2007) Pathogenesis-related protein 1 homologue is an antifungal protein in *Wassila japonica* leaves and confers resistance to *Botrytis cinerea* in transgenic tobacco. *PLANT BIOTECHNOLOGY* 24: 247–254.
14. Zhu F, Xu M, Wang S, Jia S, Zhang P, et al. (2012) Prokaryotic expression of pathogenesis related protein 1 gene from *Nectria haemaphysalis* antifungal activity and preparation of its polyclonal antibody. *Biotechnol Lett*.
15. Cantacessi C, Campbell BE, Visser A, Goldhof P, Nolan MJ, et al. (2009) A portrait of the "SCP/TAPS" proteins of eukaryotes: developing a framework for fundamental research and biotechnological outcomes. *Biotechnol Adv* 27: 376–388.
16. Gibbs GM, Roelants K, O'Bryan MK (2008) The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins: roles in reproduction, cancer, and immune defense. *Endocr Rev* 29: 865–897.
17. Milne TJ, Abbenante G, Tyndall JD, Halliday J, Lewis RJ (2003) Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J Biol Chem* 278: 31105–31110.
18. Fernandez C, Seyssere T, Bruyere T, Ramage P, Mosinger E, et al. (1997) NMR solution structure of the pathogenesis-related protein P14a. *J Mol Biol* 266: 576–593.
19. Henriksen A, King TP, Mirza O, Monsalve RI, Meno K, et al. (2001) Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. *Proteins* 45: 438–448.
20. Serrano RL, Kuhn A, Hendricks A, Helms JB, Sinning L, et al. (2004) Structural analysis of the human Golgi-associated plant pathogenesis related protein GAFR-1 implicates dimerization as a regulatory mechanism. *J Mol Biol* 339: 173–183.
21. Guo M, Teng M, Niu L, Liu Q, Huang Q, et al. (2006) Crystal structure of the cysteine-rich secretory protein scorpion reveals that the cysteine-rich domain has a K⁺ channel inhibitor-like fold. *J Biol Chem* 281: 12405–12412.
22. Shikamoto Y, Suto K, Yamazaki Y, Morita T, Mizuno H (2005) Crystal structure of a GRIIP family Ca²⁺-channel blocker derived from snake venom. *J Mol Biol* 350: 735–743.
23. Mizuki N, Sarapat DE, Garcia-Sanz JA, Kanahara M (1992) The mouse male germ cell-specific gene Tps-1: molecular structure, mode of expression in spermatogenesis, and sequence similarity to two non-mammalian genes. *Mamm Genome* 3: 274–280.
24. Gibbs GM, Ota G, Reddy T, Koppers AJ, Martinez-Lopez P, et al. (2011) Cysteine-rich secretory protein 4 is an inhibitor of transient receptor potential TRP with a role in establishing sperm function. *Proc Natl Acad Sci U S A* 108: 7034–7039.
25. Seyssere T, Fernandez C, Murrenthalers C, Wuthrich K (1998) Structure comparison of human glioma pathogenesis-related protein GPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system. *Proc Natl Acad Sci U S A* 95: 2262–2266.
26. Lu G, Villalba M, Coscia MR, Hoffman DR, King TP (1993) Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. *J Immunol* 150: 2823–2830.
27. King TP, Spangfort MD (2000) Structure and biology of stinging insect venom allergens. *Int Arch Allergy Immunol* 123: 99–106.
28. Mochales-Morales J, Martin BM, Posada LD (1990) Isolation and characterization of heliothermine, a novel toxin from *Heliothis virescens* (Mexican bearded lizard) venom. *Toxicon* 28: 299–309.
29. Wang J, Shen B, Guo M, Lou X, Duan Y, et al. (2005) Blocking effect and crystal structure of natriin toxin, a cysteine-rich secretory protein from *Naja atra* venom that targets the BKCa channel. *Biochemistry* 44: 10145–10152.
30. Hawdon JM, Jones BF, Hoffman DR, Hotes PJ (1996) Cloning and characterization of Ancylostoma-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae. *J Biol Chem* 271: 6672–6678.
31. Chalmers IW, McArdle AJ, Coulson RM, Wagner MA, Schmid R, et al. (2006) Developmentally regulated expression, alternative splicing and distinct subgroups in members of the *Solenostoma nasutum* venom allergen-like (SeNAL) gene family. *BMC Genomics* 9: 89.
32. Moyle M, Foster DL, McGrath DE, Brown SM, Laroche Y, et al. (1994) A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18. *J Biol Chem* 269: 10008–10015.
33. Zúñiga S, Bodinovic J, Garcia-Gastalero JM, Jon na, Balente JP, et al. (1999) Deletion of 24 open reading frames from chromosome XI from *Solenostoma nasutum* and phenotypic effects of the deletions. *Gene* 233: 141–150.
34. Schuren FH, Agardsdottir SA, Kothe EM, Scheer JM, Wesels JG (1993) The Sc7/Sc14 gene family of *Schistosoma mansoni* codes for extracellular proteins specifically expressed during fruit-body formation. *J Gen Microbiol* 139: 2003–2009.
35. Braun BR, Head WS, Wang MX, Johnson AD (2000) Identification and characterization of TUP1-regulated genes in *Candida albicans*. *Genetics* 156: 31–44.
36. Pradon-Rossier RC, Roldan-Rodriguez R, Serena G, Lopez-Berges MS, Guarro J, et al. (2012) A PR-1-like Protein of *Fusarium oxysporum* Functions in Virulence on Mammalian Hosts. *J Biol Chem* 287: 21970–21979.
37. Emanuelson O, Brunak S, von HG, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953–971.
38. Zdobnov EM, Apweiler R (2001) InterProScan: an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847–848.
39. Realini C, Rogers SW, Rechsteiner M (1994) KEKE motifs: proposed roles in protein-protein association and presentation of peptides by MHC class I receptors. *FEBS Lett* 348: 109–111.
40. Murray J, Gregory WF, Gomez-Ibarra N, Atmadja AK, Maizels RM (2001) Expression and immune recognition of *Brugia malayi* VAL-1, a homologue of veridip venom allergens and *Ancylostoma* secreted proteins. *Mol Biochem Parasitol* 118: 89–96.
41. Cantacessi C, Hofmann A, Young ND, Broder U, Hall RS, et al. (2012) Insights into SCP/TAPS Proteins of Liver Flukes Based on Large-Scale Bioinformatic Analyses of Sequence Datasets. *PLoS One* 7: e31164.
42. Del Valle A, Jones BF, Harrison LM, Chadderton RC, Cappella M (2000) Isolation and molecular cloning of a secreted hookworm platelet inhibitor from adult *Ancylostoma ceylanicum*. *Mol Biochem Parasitol* 129: 167–177.
43. Lozano Torres JL, Wilbers RH, Gasconski P, Bozhov J, Finkers-Torczak A, et al. (2012) Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proc Natl Acad Sci U S A* 109: 10119–10124.
44. Gao B, Allen R, Maier T, Davis EL, Baum TJ, et al. (2001) Molecular characterization and expression of two venom allergen-like protein genes in *Heterotis niloticus*. *Int J Parasitol* 31: 1617–1623.
45. Lozano J, Smart G (2011) Survival of Plant-parasitic Nematodes inside the Host. In: Perry RN, Wharton DA, editors. *Molecular and Physiological Basis of Nematode Survival*. Oxfordshire: CAB, 20–62.
46. Ding X, Shields J, Allen R, Hussey RS (2000) Molecular cloning and characterization of a venom allergen Ag3-like cDNA from *Meloidae isopoda*. *Int J Parasitol* 30: 77–81.
47. Wang X, Li H, Hu Y, Fu P, Xu J (2007) Molecular cloning and analysis of a new venom allergen-like protein gene from the root-knot nematode *Meloidae isopoda*. *Exp Parasitol* 117: 133–140.
48. Kang JS, Koh YH, Moon YS, Lee SH (2012) Molecular properties of a venom allergen-like protein suggest a parasitic function in the pinewood nematode *Bursaphelenchus xylophilus*. *Int J Parasitol* 42: 63–70.
49. Saunders DG, Win J, Cano LM, Seabrook TJ, Kamoun S, et al. (2012) Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PLoS One* 7: e29647.
50. Oeser B, Beaumont F, Haarmann T, Lorenz N, Nathues E, et al. (2009) Expressed sequence tags from the flower pathogen *Claviceps purpurea*. *Mol Plant Pathol* 10: 665–684.
51. Asajo OA, Good G, Dhar K, Loukas A, Zhan B, et al. (2005) X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Nippostrongylus brasiliensis*, and a vaccine antigen for human hookworm infection. *J Mol Biol* 346: 801–814.
52. Kobayashi YM, Alsalhani BA, Jones LR (2000) Localization and characterization of the calyculin-binding domain of triadlin 1. Evidence for a charged beta-strand in mediating the protein-protein interaction. *J Biol Chem* 275: 17639–17646.
53. Realini C, Rechsteiner M (1995) A proteasome activator subunit binds calcium. *J Biol Chem* 270: 29664–29667.
54. Reddy AS, Ali GS, Celestik H, Day IS (2011) Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. *Plant Cell* 23: 2010–2022.
55. Aime MC, Phillips-Mora W (2006) The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia* 97: 1012–1022.
56. Prias GA, Purdy LH, Schimidt RA (1995) An inoculation method for evaluating resistance of cacao to *Crinipella perniciosa*. *Plant Disease* 79: 787–791.
57. Meinhardt LW, Bellato CM, Rincomes J, Azevedo RA, Caceres JC, et al. (2006) In vivo production of biotrophic-like cultures of *Crinipella perniciosa*, the causal agent of witches' broom disease of *Theobroma cacao*. *Curr Microbiol* 52: 191–196.
58. Stanke M, Steinke R, Waack S, Morgenstern B (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 32: W309–W312.
59. Roy A, Kuchikura A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5: 725–738.
60. Leskowitz RA, MacArthur MW, Moss DS, Thornton JM (1993) Procheck: A Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* 26: 283–291.
61. DeLano WL (2002) The PyMOL Molecular Graphics System. <http://www.pymol.org>.
62. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, et al. (2011) Integrative genomics viewer. *Nat Biotechnol* 29: 24–26.