UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

RICARDO AUGUSTO TIBURCIO

PADRÕES DE EVOLUÇÃO GÊNICA E GENÔMICA EM *MONILIOPHTHORA PERNICIOSA* (BASIDIOMICOTA, AGARICALES), AGENTE CAUSAL DA VASSOURA DE BRUXA

Tese apresentada ao Instituto de Biologia para obtenção do Título de Doutor em Genética e Biologia Molecular, na área de Genética de Microrganismos.

Este exemplar corresponde à redação final da tese defendida pelo(a) candidato (a) Ricardo Augusto Tiburcio e aprovada pela Comissão Julgadora.

ORIENTADOR: PROF. DR. GONÇALO A. GUIMARÃES PEREIRA

CAMPINAS, 2010

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

T434p

Tiburcio, Ricardo Augusto
Padrões de evolução gênica e genômica em *Moniliophthora perniciosa* (Basidiomicota, Agaricales), agente causal da vassoura de bruxa / Ricardo Augusto Tiburcio. – Campinas, SP: [s.n.], 2010.

Orientador: Gonçalo Amarante Guimarães Pereira.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Vassoura-de-bruxa (Fitopatologia). 2. *Moniliophthora perniciosa*. 3. Cacau. 4. Transferência gênica horizontal. 5. Genética - Evolução. I. Pereira, Gonçalo Amarante Guimarães. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

(rcdt/ib)

Título em inglês: Paterns of genic and genomic evolution in *Moniliophthora perniciosa* (Basidomicota, Ggaricales) causal agent of witche's broom disease.
Palavras-chave em inglês: Witche's broom disease; *Moniliophthora perniciosa*; Cacao; Horizontal gene transfer; Evolutionary genetics.
Área de concentração: Genética de Microorganismos.
Titulação: Doutor em Genética e Biologia Molecular.
Banca examinadora: Gonçalo Amarante Guimarães Pereira, Ana Maria Lima de Azeredo-Espin, Marcelo Menossi, Aristóteles Góes-Neto, Antônio Rossi Filho.
Data da defesa: 12/02/2010.
Programa de Pós-Graduação: Genética e Biologia Molecular.

Campinas, 12 de fevereiro de 2010

Banca examinadora

Prof. Dr. Gonçalo Amarante Guimarães Pereira (Orientador) _____ Assinatura

Profa. Dra. Ana Maria Lima de Azeredo-Espin _____ Assinatura

Prof. Dr, Marcelo Menossi _____ Assinatura

Prof. Dr, Antônio Rossi Filho _____ Assinatura

Prof. Dr. Aristóteles Goés-Neto _____ Assinatura

Prof. Dr. Sérgio Furtado dos Reis _____ Assinatura

Prof. Dr. Antônio Vargas de Oliveira Figueira _____ Assinatura

Prof. Dr. Diogo Meyer _____ Assinatura

iii

AGRADECIMENTOS

Ao Professor Doutor Gonçalo Amarante Guimarães Pereira, meu orientador, pelo apoio e incentivo constantes, pelo exemplo profissional e por me permitir compartilhar dos recursos do Laboratório de Genômica e Expressão (LGE) durante o desenvolvimento deste trabalho.

Aos membros da banca e da pré-banca, pela disponibilidade e pelos comentários e sugestões que tanto ajudaram na finalização deste trabalho.

Às agência financiadoras dos projetos genomas de *M. perniciosa* e *M. roreri*.

Aos (muitos) colegas do LGE, atuais e passados, pelo apoio e companheirismo e por compartilharem seu conhecimento e sua competência.

A minha família, por todos estes anos de constante apoio e estímulo enquanto eu cuidava do sonho de fazer ciência.

Resumo

Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora é um fungo basidiomiceto hemibiotrófico causador da vassoura de bruxa no cacau (*Theobroma cacao* L.). Esta doença é caracterizada pela produção de ramos hiperplásicos e hipertróficos e frutas partenogênicos durante a fase biotrófica, seguido por necrose dos tecidos infectados durante a fase necrotrófica. Esta é a principal doença da produção de cacau na América do Sul e no Caribe, tendo causado uma diminuição de mais de 65% na produção no sul da Bahia, a principal região produtora do país.

Vários genes relacionados à patogenicidade foram encontrados em *M. perniciosa* como parte do Projeto Genoma da Vassoura de Bruxa. No entanto, os padrões de evolução destes genes, que poderiam ajudar a compreender a evolução da fitopatogenicidade dessa espécie, ainda não foram estudados. Também não foram estudados as mudanças ocorridas no genoma desta espécie durante a sua evolução. Estas mudanças são particularmente relevantes uma vez que a maioria das espécies estreitamente relacionadas com *M. perniciosa* não é fitopatogênica. Desta forma a compreensão da evolução deste modo de vida na espécie poderia ajudar a entender a evolução de fungos fitopatógenos em geral.

Este trabalho tem como objetivo ampliar o conhecimento sobre o surgimento da patogenicidade em *M. perniciosa* analisando a evolução dos genes potencialmente relacionados a ela, aumentando assim o conhecimento sobre a vassoura de bruxa e ajudando na busca por uma maneira de controlar a doença. Para isso, foram estudados os padrões de evolução de diversos genes ligados ao processo de necrose característica da vassoura de bruxa e genes candidatos a terem sido obtidas por transferência horizontal de genes (HGT). Os resultados indicam que a evolução da patogenicidade em *M. perniciosa* envolveu tanto à aquisição de novos genes quanto a utilização para novas funções de outros já presentes nos ancestrais saprotróficos. Em particular, a busca de genes horizontalmente transferidos encontrou dois genes potencialmente envolvidos na doença que ainda não tinham sido

estudados, sugerindo que a análise evolutiva a nível genômico é um instrumento relevante para a escolha de genes candidatos para análises em bancada.

Palavras chave

Vassoura de bruxa, *Moniliophthora perniciosa*, Cacau, Transferência gênica horizontal, Evolução de genes

Abstract

Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora is a hemibiotrophic basidiomycete fungus that causes witches' broom disease in cocoa (*Theobroma cacao* L.). This disease is characterized by the production of hyperplasic and hypertrophic stems and parthenogenetic fruits during biotrophic life-phasee, followed by necrosis of infected tissues during necrotrophic life-phase. This is the main disease affecting the cocoa production in South America and the Caribbean, causing a decrease of more than 65% in the production in southern Bahia, the main producing region.

Several genes related to pathogenicity were found in *M. perniciosa* as part of the Witches' Broom Genome Project. However the patterns of evolution of these genes, that could help to understand the evolution of plant pathogenicity in this species, were not yet studied. Also the changes occurred in this species at a genomic level during its evolution were not studied. These changes are particularly relevant since the majority of species closely related to *M. perniciosa* is not phytopathogenic and so the understanding of this evolution would help to understand the evolution of plant pathogens in general.

This work attempts to increase the knowledge about the emergence of pathogenicity in *M. perniciosa* studding the evolution of genes potentially related to it, therefore increasing the knowledge about the witches' broom and helping in the search for a way to control the disease. To do that, it were studied the evolutionary patterns of several genes related with the necrosis process characteristic of witches' broom and genes candidate to have been

obtained by horizontal gene transfer (HGT). These later were detected by a genome level screening. The results indicate that the evolution of plant pathogenicity in *M. perniciosa* involved both the acquisition of new genes and the use of genes already present in the saprotrophic ancestral in new functions. In particular, the search for horizontally transferred genes found two genes potentially involved in the disease which were not yet studied, suggesting that the evolutionary analysis at the genomic level is a relevant tool to the choice of genes candidate to laboratory analysis.

KEY-WORDS

Witches' broom disease, *Moniliophthora perniciosa*, Cacao, Horizontal gene transfer, Gene evolution.

Sumário

INTRODUÇÃO GERAL.

A vassoura de bruxa é a principal praga a afetar a produção de cacau (*Theobroma cacao* L.) na América do Sul na atualidade. Esta doença é causada pelo fungo basidiomiceto *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora (Agaricales, Marasmiaceae), endêmico da região amazônica, onde sua presença endêmica já inviabilizava historicamente a produção cacaueira em larga escala. Seu aparecimento no sul da Bahia, a principal região produtora de cacau no país, causou uma queda de 400 para 130 mil toneladas por ano na produção de cacaueira no Brasil entre o fim dos anos 80 e início dos anos 90, (Santos Filho *et al.* 1998; de Arruda *et al.* 2003; Griffith *et al.* 2003).

O *M. perniciosa* é um fitopatógeno hemibiotrófico. O ciclo infeccioso se inicia quando seus basidiósporos penetram na planta, aparentemente através da epiderme, e germinam crescendo pelos espaços intercelulares (o apoplasto) e avançando em direção aos meristemas. Nesta fase biotrófica a produção de hifas é pequena e estas são monocarióticas (Evans 1980). Ao atingir os meristemas o fungo provoca perda da dominância apical, com consequente formação de vários ramos axilares hipertróficos e hiperplásicos. Os ramos produzidos exibem grande desenvolvimento e vigor e seu conjunto é chamado de “vassoura verde”.

Após o crescimento dos ramos da vassoura verde inicia-se o processo de necrose destes, que avança culminando na produção da “vassoura seca” que corresponde ao conjunto de ramos da vassoura verde necrosados. Nesta fase necrotrófica o fungo prolifera e invade o interior das células necrosadas, produzindo grande quantidade de hifas nas quais são encontradas grampos de conexão entre células adjacentes, indicadores do processo de dicariotização. O mecanismo de disparo da mudança para a fase necrotrófica ainda não é completamente conhecido, mas já se sabe que *in vitro* é possível manter as hifas na fase biotrófica por meio da adição de glicerol como única fonte de carbono ou da manutenção do fungo em baixas concentrações de

nutrientes (Meinhardt *et al.* 2006). Pungartnik *et al.* (2009) sugeriram que o balanço entre a disponibilidade de glicerol e peróxido pode controlar esta transição de fase. Uma vez necrosado o ramo pode persistir ligado a planta por um período de até três anos, durante os quais são produzidos basidiomas e esporos, geralmente após uma alternância entre períodos úmidos e secos.

Alem dos ramos o *M. perniciosa* pode infectar também flores e frutos. Nestes últimos a presença do fungo na fase biotrófica leva a partenocarpia (produção de frutos sem sementes) e portanto sem valor comercial, pois o principal produto do cacau é a sua semente, matéria prima importante na indústria alimentícia. Nas flores a presença do fungo pode levar a produção de ramos vegetativos ou ao desenvolvimento anormal destas. Flores e frutos infectados também podem produzir basidiomas e esporos. Independente do tecido infectado o maior sintoma da presença do fungo é a necrose de tecidos da planta, sendo este também o maior dano gerado pela vassoura de bruxa (Purdy & Smith 1996; Silva *et al.* 2002).

Graças ao sequenciamento do genoma do *M. perniciosa* (Mondego *et al.* 2008) muitos genes com possíveis papéis na infecção inicial do cacaueiro pelos esporos; na hipertrofia, hiperplasia e perda de dominância apical dos meristemas dos ramos infectados; no metabolismo energético e resistência a processos oxidativos e na proteção da vassoura seca contra a entrada de outros saprófagos foram encontrados. Alguns destes genes vêm sendo estudados em detalhes na busca de uma melhor compreensão do seu papel no ciclo de vida do fungo, visando uma possível futura aplicação no controle da vassoura de bruxa. Paralelamente, vêm sendo estudadas também as mudanças histológicas, fisiológicas e bioquímicas que ocorrem nos tecidos infectados do cacau, tanto na fase biotrófica quanto na necrotrófica. Com base nos resultados obtidos até o momento já se sabe que a fase biotrófica, dependente dos nutrientes cedidos pela planta, aparentemente é a responsável tanto pela a hipertrofia quanto pela indução da necrose características da vassoura de bruxa e que ambos os processos estão

relacionados com a tomada de nutrientes pelo fungo. Recentemente as mudanças na bioquímica dos tecidos do cacau infectados pelo *M. perniciosa* foram descritas por Scarpari *et al.* (2005), que detectaram alterações condizentes com resposta ao stress, como aumento da concentração de etileno e de glicerol durante a fase biotrófica da infecção. O aumento de glicerol inclusive sugere que o fungo dispara o processo de hipertrofia do ramo como uma forma de sequestrar nutrientes da planta durante a fase biotrófica, para utilizá-los na fase necrotrófica para produção de basidiomas e esporos.

Na fase necrotrófica os processos de nutrição do *M. perniciosa* ainda não são conhecidos, mas já se sabe que a necrose em si parece estar relacionada com processos de degradação de cristais de oxalato de cálcio acumulados pelas variedades de cacau suscetíveis ao fungo, sendo que variedades da planta que não acumulam estes cristais são resistentes ao *M. perniciosa*. O oxalato de cálcio degradado parece estar envolvido na inicialização de um processo de morte celular programada via produção de espécies reativas de oxigênio (Ceita *et al.* 2005) e liberação de íons cálcio, que se ligam a pectina presente na parede celular tornando-a mais acessível às pectinases do fungo (Rio *et al.* 2008). Também já foram detectadas, e estudadas, duas famílias de proteínas relacionadas a necrose causada pelo fungo nos ramos e frutos infectados. A primeira delas é formada por proteínas indutoras de necrose e produção de etileno (NEPs) (Garcia *et al.* 2007), já descritas como elicitores de necrose em fungos e oomicetos fitopatogênicos, enquanto a segunda é formada por cerato-plataninas (Zaparoli *et al.* 2009), que normalmente não são relacionadas a necrose.

Alem das informações obtidas pelo estudo de genes individuais, muitas outras informações a respeito do *M. perniciosa* e da vassoura de bruxa vendo sendo obtidas a partir dos dados do projeto genoma deste fungo. Por exemplo Rincones *et al.* (2008) analisaram a expressão diferencial de genes entre as fases biotrófica e necrotrófica e encontraram uma série de proteínas

relacionadas com fitopatogenicidade preferencialmente expressas na fase biotrófica, incluindo glioxal oxidases, quitina sintases, proteinases e lipases. Os autores sugeriram ainda que a aquisição de metabólitos de carbono e nitrogênio pode estar relacionada a transição de fase. Adicionalmente a análise do genoma completo de *M. perniciosa* (Mondego *et al.* 2008) revelou a presença de famílias de proteínas que podem estar relacionadas com detoxicação e combate a resposta contra patógenos produzida pelas plantas infectadas, como citocromo p450 oxigenases, transportadores de membrana e enzimas antioxidativas. O mesmo trabalho mostrou ainda a presença de muitas proteínas de degradação de parede celular, num repertório similar ao de *Fusarium graminearum* e *Magnaporthe grisea*, dois ascomicetos fitopatogênicos bem estudados.

Embora já tenham sido obtidos grandes avanços na compreensão da genética por trás dos mecanismos bioquímico e fisiológicos usados pelo *M. perniciosa* durante a infecção e necrose do cacaueiro, ainda se conhece pouco sobre a evolução destes. *M. perniciosa* é um fungo da família Marasmiaceae e sabe-se que nesta família apenas umas poucas espécies aparentadas a ele, como *M. roreri* (que também infecta o cacau) e algumas outras do gênero *Crinipellis*, são fitopatógenos, sendo que a maioria das espécies da família é saprotrófica (Aime & Phillips-Mora 2005). É de se esperar portanto que a evolução da fitopatogenicidade em *M. perniciosa* tenha envolvido não apenas a readaptação para novas funções ligadas a fitopatogênese de genes já existentes nos ancestrais saprotróficos mas também a obtenção de novos genes ligados a este novo modo de vida. Em particular, os genes relacionados a patogenicidade descritos acima devem apresentar padrões de evolução diferenciados quando comparados aos demais genes da espécie, ou aos mesmos genes em espécies não patogênicas aparentadas, incluindo seleção natural diferencial e ocorrência de duplicação e expansão gênica. É possível também que alguns daqueles genes tenham sido obtidos por transferência gênica horizontal (HGT) a partir

de outros patógenos do cacau, como oomicetos ou fungos ascomicetos. Transferências deste tipo foram relacionadas com a evolução de novos modos de vida em fungos por vários autores (Richards *et al.* 2006; Wenzl *et al.* 2006; Sanders 2006; Stukenbrock and McDonald 2008; Oliver and Solomon 2008, Khaldi *et al.* 2008).

Tendo em vista as hipóteses expostas acima a respeito da evolução da fitopatogenicidade em *M. perniciosa*, este trabalho procurou avaliar quais daqueles padrões de evolução podem ser encontrados em genes relevantes na espécie. Para tanto foram construídas filogenias para os genes das proteínas indutoras de necrose (NEPs), cerato-plataninas (MpCPs) e oxidase alternativa (AOX), todos eles sabidamente relacionados a processos fisiológicos importantes na doença. A partir destas filogenias buscou-se por padrões que pudessem ser relacionados a eventos seletivos; duplicações gênicas e formação de parálogos e famílias gênicas ou ainda ocorrência de HGT ao longo da história evolutiva destes genes. Em paralelo foram realizados estudos evolutivos no genoma mitocondrial do fungo, visando verificar a possibilidade de que a evolução desta organela, ou de um plasmídio mitocondrial encontrado nela, possa se relacionar com processos de senescência observados “in vitro” para o *M. perniciosa*. Por fim foi realizada uma busca por outros possíveis genes oriundos por HGT entre os genes preditos pelo sequenciamento do genoma de *M. perniciosa* a fim de se avaliar a importância deste processo na evolução da fitopatogênese na espécie. Espera-se que os resultados obtidos nestas análises possam incrementar a compreensão da evolução da fitopatogenicidade em *M. perniciosa*, auxiliando tanto na escolha de genes e processos bioquímicos e fisiológicos candidatos a alvo no combate à vassoura de bruxa quanto no estabelecimento de modelos gerais de evolução desta característica em fungos.

Os resultados obtidos ao longo do trabalho serão apresentados a seguir em cinco capítulos. O primeiro e o segundo relatam as análises realizadas

com as proteínas relacionadas ao processo de necrose ocorrido na fase final da vassoura de bruxa: NEPs (capítulo I) e cerato-plataninas (capítulo II). O capítulo III relata o estudo realizado com a enzima da via de oxidação alternativa (AOX) encontrada neste fungo. A análise do genoma mitocondrial de *M. perniciosa* e do plasmídio encontrado ligado a este genoma é mostrada no capítulo IV. Por fim o capítulo V trás os resultados obtidos na busca por genes oriundos de HGT no genoma de *M. perniciosa*. O material dos capítulos I, II e IV foi anteriormente publicado como parte de artigos de pessoal do Laboratório de Genômica e Expressão, enquanto o do capítulo III foi apresentado na forma de painel no 57º Congresso Brasileiro de Genética. Por fim o do capítulo V foi publicado no Jornal of Molecular Evolution, sendo apresentado aqui na sua versão final.

CAPÍTULO I – ANÁLISE FILOGENÉTICA DAS PROTEÍNAS DE NECROSE DE *MONILIOPHTHORA PERNICIOSA*.

INTRODUÇÃO E OBJETIVOS.

Durante a análise inicial das sequências genômicas de *M. perniciosa* foram encontradas duas cópias de uma proteína indutora de necrose e etileno (NEPs). Estas proteínas têm sido encontradas em diversos fitopatógenos em anos recentes, incluindo ascomicetos, oomicetos e bactérias e formam uma nova classe de elicitores: proteínas reconhecidas pelo sistema de defesa da planta (Gijzen & Nurnberger 2006; Pemberton & Salmond 2004). Em plantas a morte celular programada (PCD) induzida em resposta à presença de elicitores visa impedir a progressão do patógeno para alem da área inicial de infecção. No entanto patógenos com uma fase de vida saprotrófica, como o *M. perniciosa*, podem usar o próprio tecido necrosado para sua nutrição e produção de basidiomas e esporos, completando assim seu ciclo de vida.

Os mecanismos pelos quais as NEPs induzem a necrose ainda não são conhecidos, mas há evidências tanto da produção de etileno, um hormônio vegetal envolvido em processos de maturação e senescência, quanto de ânions super-óxido, ligados a processo de morte celular por dano oxidativo (Fellbrich *et al.* 2002). Também foi relatada a expressão do gene PR1, ligado a resistência dependente de ácido salicílico, bem como a indução de outros genes relacionados a PCD (Veit *et al.* 2001), resposta a stress e defesas contra patógenos (Bailey *et al.* 2005; Keates *et al.* 2003; Verica *et al.* 2004). Recentemente foi sugerido ainda que as NEPs podem ser não apenas elicitores, mas também toxinas, capazes de realizar por si só a quebra da membrana plasmática em células de dicotiledôneas. Em todos os casos a proteína parece agir fora da célula vegetal, sugerindo a existência de receptores de membrana para ela. O fato de que diferentes patógenos e hospedeiros apresentam diferentes mecanismos de resposta a esta família de proteínas sugere que elas possam agir de mais de uma maneira, o que

também é sugerido pela alta proporção de fitopatógenos com mais de uma cópia funcional desta proteína (Gijzen & Nurnberger 2006; Tyler *et al.* 2006). Em *M. perniciosa* foi detectado que as duas cópias da NEP (NEP1 e NEP2) tem diferentes padrões de oligomerização e são expressos diferencialmente entre as fases de vida do fungo (Garcia *et al.* 2007).

Neste trabalho é apresentada uma filogenia produzida para todas as NEPs descritas até o momento, incluindo as duas cópias da proteína encontradas em *M. perniciosa*. A partir desta filogenia foram avaliados os padrões de duplicação desta família de proteínas, a fim de se verificar se há evidência para relacionar as diferentes formas de ação da mesma com sua evolução. Também foi avaliada a possibilidade de que a presença das NEPs em grupos não diretamente relacionados de eucariotos e procariotos possa ser explicada por transferência gênica horizontal (HGT). Os resultados destas análises foram publicados como parte de artigo "Characterization of necrosis and ethylene-inducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom in *Theobroma cacao"* (Garcia *et al.* 2007).

MATERIAL E MÉTODOS.

As sequências de aminoácidos usadas para análise filogenética das NEPs foram localizadas na base de dados não redundante (nr) do GenBank por meio de BLASTp e tBLASTx (Altschul *et al.* 1990; McGinnis & Madden 2004), usando como "query" a região codificante da sequência da NEP1 de *M. perniciosa*. Foram obtidas 31 sequências oriundas de 20 espécies de fungos, oomicetos e bactérias, indicadas na tabela 1 do anexo. Estas sequências foram alinhadas pelo programa ClustalX (Thompson 1997), com o uso das opções defaut para proteínas, exceto pelo uso da matriz de transição PAM (Dayhoff *et al.* 1972). Trechos pouco conservados ou com excesso de in/dels foram manualmente removidos usando-se o editor de alinhamentos do programa MEGA 3 (Kumar *et al.* 2004). Também foram removidas as regiões correspondentes aos peptídios sinais, identificados por meio do

programa Signal P3 (Bendtsen *et al.* 2004) obtendo-se no final cerca de 330 resíduos de aminoácidos confiavelmente alinhados. A filogenia foi construída com o programa MEGA 3 empregando algoritmo de Neigbor Joining e a matriz de transição PAM e permitindo variação contínua nas taxas de mutação entre sítios modelada por uma aproximação da distribuição gama com parâmetro alfa fixado em 1,0. O suporte dos ramos foi verificado com o uso de 1000 bootstraps.

RESULTADOS.

A figura 1 mostra a filogenia obtida para as proteínas indutoras de necrose de fungos, oomicetos e bactérias. Nela pode ser notada a existência de dois parálogos principais desta proteína, sendo que algumas espécies de fungos ascomicetos (*Aspergillus nidulans* e *A. fumigatus*, *Magnaporthe grisea* e *Gibberella zeae*) apresentam ambos. As sequências destes dois parálogos são claramente distintas, sendo que as sequências do tipo I apresentam duas inserções ausentes nas demais, uma delas terminal e outra próxima a posição 250 do alinhamento (figura 2). Já os oomicetos, embora em geral apresentem mais de uma NEP por espécie, só possuem sequências do segundo parálogo (tipo II, a direita da figura 6). É de se notar que sequências de bactérias do gênero *Bacillus* estão posicionadas entre os oomicetos, que portanto não formam um grupo monofilético. Entre as bactérias por sinal há espécies tanto possuidoras de um quanto de outro parálogo, mas não de ambos. Em eucariotos os parálogos têm um peptídeo sinal, duas cisteínas conservadas e um heptâmero altamente conservado (GHRHWDE) próximo ao meio da proteína.

As sequências de *M. perniciosa* pertencem ambas ao tipo II da proteína. Estão agrupadas como irmãs e situadas como grupo externo em relação as sequências do mesmo tipo obtidas de ascomicetos, posição esperada no caso de herança vertical desta proteína nos fungos. No entanto o ramo responsável por este grupamento asco/basidio é pouco sustentado (210/1000

bootstraps), sendo que a mesma falta de sustentação é notada em outros ramos desta filogenia.

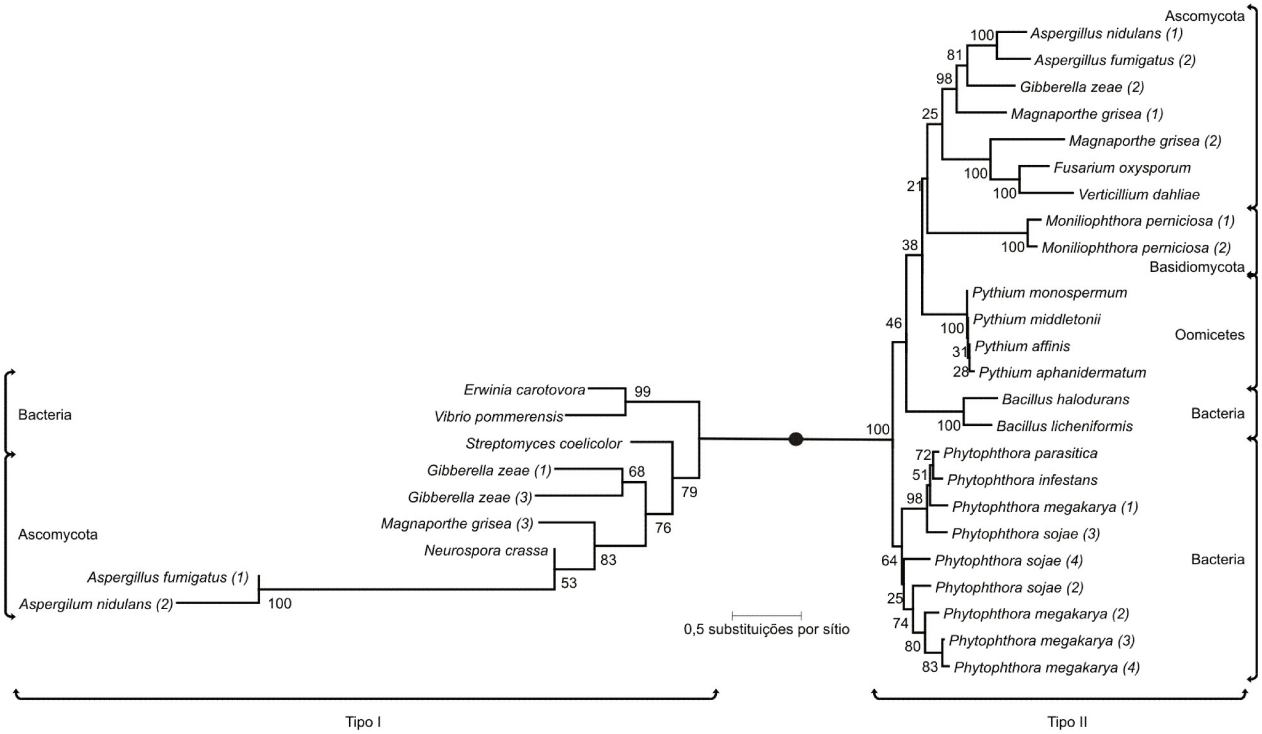


Figura 1 Filogenia enraizada das NEPs de fungos, oomicetos e bactérias. A raiz da árvore e provável ramo de separação entre parálogos do tipo I (a esquerda) e tipo II (a direita) está indicado por um círculo preenchido. Os valores próximos aos ramos internos são bootstraps.

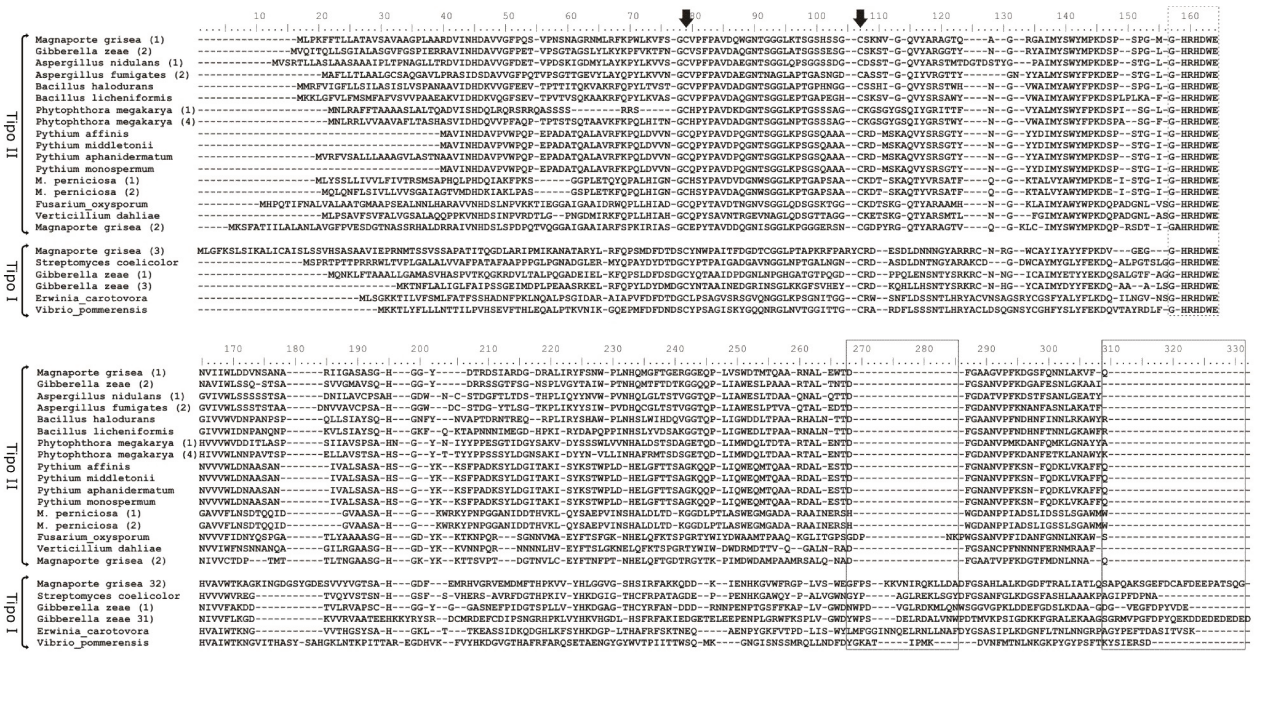


Figura 2 sequências das NEPs de *M. perniciosa* e outras espécies selecionadas. Os resíduos conservados de cisteína são indicados por setas. O heptâmero conservado GHRHDWE está circundado por uma linha pontilhada. As inserções presentes nas sequências do tipo I estão circundadas por linhas contínuas.

DISCUSSÃO.

Embora não se trate de uma proteína extremamente variável, a filogenia obtida para a NEP apresentou uma série de ramos com baixa sustentação. Esta baixa sustentação, a mistura de organismos de vários grupos em pontos distintos da filogenia e a presença de sequências de uma mesma espécie em pontos distintos desta pode indicar que os fatores de necrose deste tipo tem uma existência antiga, desde o ancestral comum de procarioto e eucariotos, e passaram por processos complexos de duplicação e perda gênica nas linhagens de fungos, bactérias e oomicetos. Por isso o sinal filogenético desta proteína é fraco, dificultando a recuperação de sua história evolutiva. Este problema torna-se ainda mais forte na medida em que a proteína é pequena, em geral com menos de 250 resíduos de aminoácido, o que por si só induz a obtenção de ramos com baixos bootstraps. Na filogenia obtida não foi possível posicionar a raiz da árvore em nenhum ponto que agrupasse as três espécies de bactérias entre si e a parte das demais espécies usadas na análise, sugerindo que já havia pelo menos duas versões do gene nos procariotos, as quais podem ter sido mantidas em alguns organismos como *Gibberella sp.*, enquanto outros mantiveram apenas uma das cópias. Com a evolução, novas cópias dessas versões específicas podem ter aparecido independentemente dentro de linhagens de fungos ou oomicetos. O sequenciamento completo de genomas é necessário para reforçar ou eliminar essa hipótese.

Uma explicação alternativa para esta hipótese envolve a obtenção destas proteínas via HGT tanto por *M. perniciosa* quanto pelas bactérias. Estas últimas apareceram dispersas pela filogenia. Esta dispersão e a presença de NEPs apenas em umas poucas espécies bacterianas, mesmo havendo centenas de genomas completos deste grupo no GenBank, torna

difícil postular uma explicação para a evolução das NEPs bacterianas que não envolva pelo menos algumas ocorrências de HGT. Quanto a *M. perniciosa*, embora em princípio seu posicionamento na filogenia como grupo irmão dos ascomicetos seja o esperado por transferência vertical, o ramo que une estes dois grupos de fungo é muito pouco sustentado. Alem disso a presença de oomicetos próximos a *M. perniciosa* e sem formar um grupo monofilético, torna impossível descartar a hipótese da obtenção das NEPs em *M. perniciosa* via HGT a partir de oomicetos. Esta transferência seria facilitada pela existência de patógenos do cacaueiro pertencentes a este grupo, como por exemplo *Phytophthora. megakaria* e *P. palmivora* (Evans 2007)[1].

[1] No capítulo V será descrito como a busca por genes horizontalmente transferidos em *M. perniciosa* confirmou a origem das NEPs desta espécie por esta via. A mesma análise indicou ainda que as NEPs se originaram em ascomicetos, tendo os demais grupos obtido esta proteína em vários eventos de HGT.

CAPÍTULO II – ANÁLISE FILOGENÉTICA DAS CERATO-PLATANINAS DE *MONILIOPHTHORA PERNICIOSA*.

INTRODUÇÃO E OBJETIVOS.

A busca por genes na sequência do genoma de *M. perniciosa* revelou uma segunda família de proteínas secretadas relacionadas a fitotoxicidade e fitopatogenia, as cerato-plataninas. Cerato-plataninas são pequenas proteínas secretadas (cerca de 12 kDa) com abundância de resíduos hidrofóbicos, incluindo 4 cisteínas conservadas formadoras de duas pontes de sulfito (Pazzagli *et al.* 2006). A família foi descobertas recentemente em *Ceratocystis fimbriata*, um patógeno de plátanus (*Platanus acerifolia*) e tem sido relatada em vários outros fungos fitopatógenos e patógenos animais. Suas funções ainda não foram bem caracterizadas, mas já se sabe que entre os primeiros a proteína se localiza na parede celular do fungo e é capaz de disparar a morte de células vegetais tanto em plantas hospedeiras quanto não-hospedeiras (Pazzagli *et al.* 1999; Boddi *et al.* 2004). Proteínas da mesma família também estão descritas como alergênicos de Aspergillus fumigatus (Rementeria *et al.* 2005) e *Coccidioides immitis* (Pan & Cole 1995) em humanos.

Em *M. perniciosa* foi observado que pelo menos uma das 5 ORFs de cerato-platanina presentes no genoma produz uma proteína capaz de atuar como indutor de necrose tanto em folhas de cacau quanto de tabaco. Mais especificamente esta cerato-platanina parece têm um efeito sinergístico com as proteínas de necrose descritas no capítulo anterior, sendo que a ação conjunta de ambas induz um padrão de necrose no cacaueiro que é mais similar ao observado durante a vassoura de bruxa do que o produzido por qualquer das duas proteínas separadamente. Foi percebido também que a proteína é expressa preferencialmente na fase biotrófica (Zaparoli *et al.* 2009).

Neste trabalho é apresentada uma filogenia produzida para todas cerato-plataninas descritas até o momento, incluindo três das cinco cópias da

proteína encontradas em *M. perniciosa* (as com ORF completa). A partir desta filogenia foram avaliados os padrões de duplicação desta família de proteínas, a fim de se verificar se há evidência para relacionar as diferentes formas de ação da mesma com sua evolução. Também foi avaliada a possibilidade de que a ampla distribuição das cerato-plataninas, bem como sua presença em *M. perniciosa*, pode ser explicada por transferência gênica horizontal (HGT). Os resultados destas análises foram publicados como parte de artigo "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 " (Zaparoli *et al.* 2009).

MATERIAL E MÉTODOS.

As sequências usadas para análise filogenética das cerato-plataninas homólogas as encontradas em *M. perniciosa* foram localizadas na base de dados não redundante (nr) do GenBank por meio de tBLASTx, (Altschul *et al.* 1990; McGinnis & Madden 2004) usando como "query" a região codificante da sequência da cerato-platanina 1 desta espécie (MpCP1). Adicionalmente foram buscadas sequências de cerato-plataninas nos genomas dos basidiomicetos *Laccaria bicolor* (http://genome.jgi-psf.org/ Lacbi1/Lacbi1.info.html), *Phanerochaete chrysosporium* (http:// genome.jgi-psf.org/whiterot1/whiterot1.info.html) e *Coprinus cinereus* (http://www.broad.mit.edu/annotation/genome/ coprinus_cinereus/Info.html), usando uma instalação local do BLAST. Foram obtidas 30 sequências oriundas de 20 espécies de fungos ascomicetos e basidiomicetos, indicadas na tabela 2 do anexo. Estas sequências foram alinhadas pelo programa ClustalX (Thompson 1997), com o uso das opções "defaut" deste para proteínas, exceto pelo uso da matriz de transição PAM (Dayhoff *et al.* 1972). Trechos pouco conservados ou com excesso de in/dels foram manualmente removidos usando-se o editor de alinhamentos do programa MEGA 3 (Kumar *et al.* 2004). Também foram removidas as sequências das ORFs 4 e 5, que não estavam completas. Por fim foram removidas as regiões

correspondentes aos peptídios sinais, identificados por meio do programa Signal P3 (Bendtsen *et al.* 2004) obtendo-se 110 resíduos de aminoácidos confiavelmente alinhados. A filogenia foi construída com o programa PhyML (Guindon & Gascuel 2003) usando-se as sequências de nucleotídeos correspondentes as regiões alinhadas das proteínas. Nesta construção foi usado o algoritmo de Máxima Verossimilhança e a matriz de transição GTR, escolhida pelo programa MrAIX.pl (Nylander 2004), permitindo variação contínua nas taxas de mutação entre sítios modelada por uma aproximação da distribuição gama com parâmetro alfa fixado em 2,0. O suporte dos ramos foi verificado com o uso de 1000 bootstraps.

RESULTADOS E DISCUSSÃO.

A figura 3 mostra a filogenia obtida para as cerato-plataninas disponíveis no GenBank. Pode-se observar que poucos ramos estão bem sustentados, uma provável consequência do pequeno tamanho da proteína. A divisão dos basidiomicetos em três linhas, cada uma das quais ligada a um grupo diferente de ascomiceto, sugere a existência de mais de um parálogo ancestral nesta família, mas o pequeno suporte dos ramos basais de cada parálogo sugerido e a ausência de diferentes parálogos em uma mesma espécie dificulta a validação desta hipótese. Pode-se sugerir ainda que se tratem de três genes com função distinta, mas novamente ainda não há clareza suficiente a respeito das funções e modos de ação das cerato-plataninas para que se possa avaliar esta hipótese. Ao contrário do observado para as NEPs (figura 2) não há nenhuma diferença estrutural marcante entre as sequências destes três possíveis parálogos das cerato-plataninas (figura 4). De modo geral as cerato-plataninas são bem conservadas, especialmente nos 60 primeiros aminoácidos após o peptídeo sinal. Uma hipótese alternativa para explicar a separação das cerato-plataninas de basidiomicetos em três grupos é a ocorrência de transferência gênica horizontal nos basidiomicetos *P. chrysosporium*, *C. cinereus* e nos ascomicetos do gênero *Coccidioides*. Neste caso as sequências

remanescentes se posicionariam de forma próxima ao esperado em caso de herança estritamente vertical do gene.

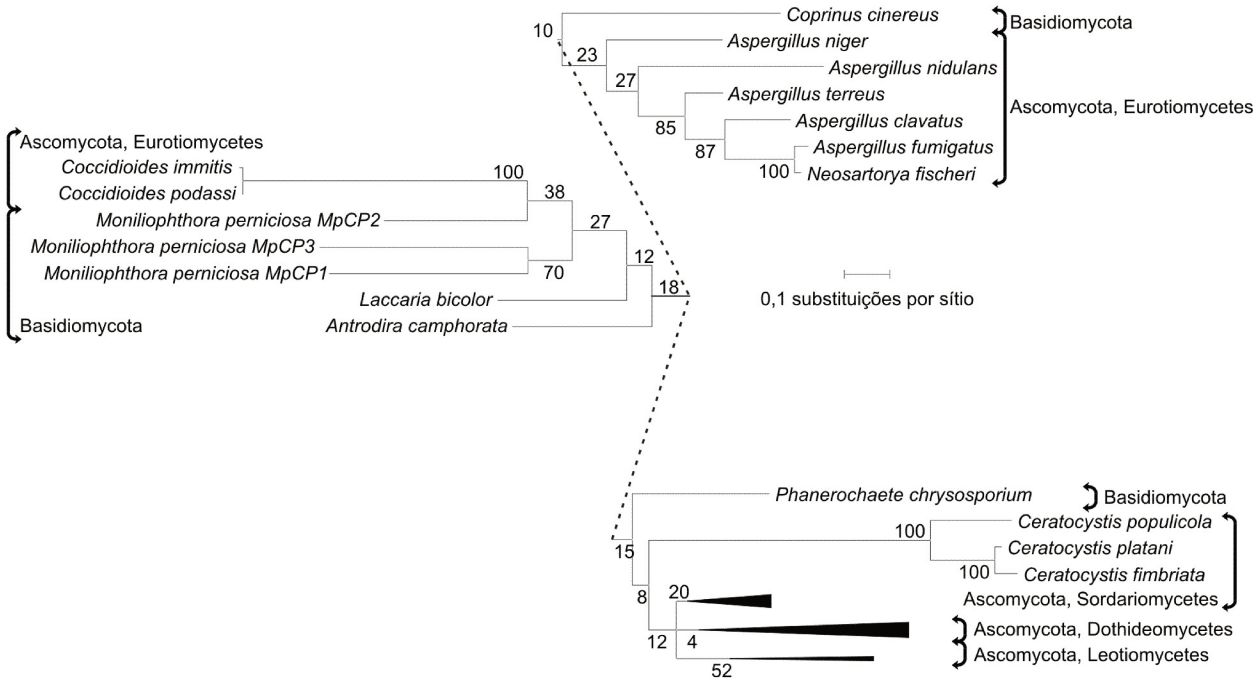


Figura 3 Filogenia não enraizada das cerato-plataninas disponíveis no GenBank. As linhas pontilhadas são usadas apenas para conectar os ramos basais de cada um dos três parálogos postulado para as cerato-plataninas, não fazendo parte dos mesmos. Os valores próximos aos ramos internos são bootstraps. Parte das espécies de ascomicetos são representadas agrupadas pela família ou ordem.

As cerato-plataninas de *M. perniciosa* estão agrupadas com as do ascomiceto *Coccidioides*. Em particular a sequência da MpCP2 é monofilética com as deste gênero. No entanto este agrupamento é pouco sustentado, e pode ser resultado apenas de uma elevada divergência entre as cerato-plataninas da espécie. Embora estas proteínas pareçam ter sido duplicadas no próprio *M. perniciosa* ou em algum ancestral próprio (os demais basidiomicetos do mesmo parálogo não tem duplicações) elas apresentam um auto grau de divergência entre si, comparável por exemplo ao observado entre as cinco espécies distintas de *Aspergillus* e maior do que o observado entre as três de *Ceratocystis*. Esta divergência elevada sugere que estas proteínas possam ter sofrido pressão seletiva diversificadora, o que seria compatível com possíveis alterações funcionais estre as diferentes cópias da mesma. Neste contexto vale ressaltar que apenas a MpCP1 foi avaliada

quanto a habilidade de induzir necrose, mas não se espera que esta seja a função original destas proteínas entre os basidiomicetos, dada a ausência de necrotróficos entre as demais espécies do grupo incluídas na filogenia. Em todo caso será necessário avaliar as características funcionais das demais cerato-plataninas de *M. perniciosa*, bem como de eventuais cópias encontradas em *M. roreri*, antes de ter um quadro mais claro da evolução destas no gênero.

```
                                         10        20        30        40        50        60        70        80
                                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Moniliophtora perniciosa 1       ------------MKSIAIFTPILILLTISAGAVKLSYDEAYDNPSSSLLSVTCSDGENGLYPK-YRTFGDLPGFPCIGGSS
Moniliophtora perniciosa 2       --------------MKFTTTIIALALAASTGAVQLRFDNTYDNASGSMNTVACSTGANGLSQR-FPTFGSVPTFPHIGASS
Moniliophtora perniciosa 3       ----------------MKFIAAVALLATSAVAVQLQYDPVYDNADQSFGTVACSDGPNGMLTKGYSTFGSVPSY--VGAVD
Coccidioides_immitis             -----MKFSLLSAIAAAVFVPFTSATPLASTTD-LSYDTHYDDPSLALSGVTCSDGDNGMITKGYNTAGEIPNYPHVGGAF
Coccidioides_posadasii           -----MKFSLLSAIAAAVFVPFTSATPLASTAD-LSYDTHYDDPSLPLSGVTCSDGDNGMITKGYNTAGEIPNYPHVGGAF
Laccaria bicolor                 --------------MKYLSSLLVFAATCLAQTFQVSYDTKYDDGSTSLATVACSDGTNGLLTKGFTTFGSLPKFPYIGGSF
Antrodia camphorata              -------------------------------GVNVTYDPFFDNPNNSLSYVACSDGTNGLLTKGYTTLGSLPDFPYIGGAY
Coprinus cinereus                ---------------MKFSAVFALLLSAYASAQTASYDFTYDNGGSSLTSVACSDGPNGLITRGYSTFGSLPGFPRIGGVP
Aspergillus nidulans             -MKLATLLPAFAAAALAAPSNPLVARQTSGTS-TLSYDTKYDVADTSLTTVACSDGDYGLITEGYTTFGSLPTFARIGGAP
Phanerochaete chrysosporium      ------MKFFAAVTALVFAASAALAAPSTTTTVTVSYDETYDNANGSLATVACSDGPNGMLTKGFTTFGSLPHFPNIGGAA
Ceratocystis populicola          ------------------MKFSILPMISSAMAVS-SYDPVYGNPGQSMNTVACSNGKFGLAAK-YPTLGDLPAFPNVGGIP
Ceratocystis platani             ------------------MKFSILPMIASAMAVS-SYDPIYAAD-LSMGSVACSNGDHGLMAQ-YPTLGEVPGFPNVGGIP
Ceratocystis fimbriata           ------------------MKFSILPMIASAMAVS-SYDPIYASG-LSMGSVACSNGDHGLMDQ-YPTLGQVPGFPNVGGIP

                                         90       100       110       120       130       140       150       160
                                 ...|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
Moniliophtora perniciosa 1       DIAGYNS-PNCGSCYQLTYSSAHTTPKSIYMVAIDRSAEGFTASKQAMDDLTNKRAEELGTVNVDVRKVDFSRCERQS--
Moniliophtora perniciosa 2       DIGGFNS-PACGNCYTISFTFQGVTR-SINLVAIDHAGNGFNVAQAAMDELTNGNAVALGTIDVQSQQVARSVCGL----
Moniliophtora perniciosa 3       TITGWNSE-SCGTCYQITWS---GTGKTIHVVGVDVAGNGFNVGQRAMDDLTNGQAVALGNIDVTATLVDKSACRL----
Coccidioides_immitis             TVETWNS-PNCGKCYKVTYNA----K-TIFLTAIDHSNSGFNIAKKSMDVLTNGRAEELGRIKVTYEEVASSLCGLK---
Coccidioides_posadasii           TVETWNS-PNCGKCYKVTYNA----K-TIFLTAIDHSNSGFNIAKKSMDVLTNGRAEELGRIKVTYEEVASSLCGLK---
Laccaria bicolor                 AIAGYNS-PDCGTCWELTYNNGAGVVKSINVLAIDVAKPGFNIGLVAMNNLTGGQAVHLGRVQATVKQVNASICGLKV--
Antrodia camphorata              AIAGWNS-PSCGTCWELTYNN-----VSINILGIDTAA-GFNIALTAMNVLTNNAAVDLGEVDAAAIQVDSSVCGL----
Coprinus cinereus                AVTGWNS-PACGSCWELTYTPPNGAKRSINIIAVDVGRGGFNLAHAALNDLTNGQATQLGRVNVAARQVAPSACG-----
Aspergillus nidulans             TIPGWNS-PNCGSCYQITYTSATGQTNSIIMTAIDTAPGGFNVGQQAMDLLTGGRAAELGRVDITWTQVGREQCGLTPRP
Phanerochaete chrysosporium      VIPGFNSD-QCGTCWKLTYN---G--ASVNVIAIDHAQAGFNIALGAMNTLTNNNAVFLGRVNAQAEQVAASECGL----
Ceratocystis populicola          DIPDWNNDQNCGTCWKVTV--PN--RNPIFILGVDSGKGNFNVSPQAFTELTGST-DAGRVNDIQYEKVSADNCKMPVA-
Ceratocystis platani             DIAGWDS-PSCGTCWKVTI--PNG--NSIFIRGVDSGRGGFNVNPTAFTKLVGST-EAGRVDNVNYVQVDLSNCINGAN-
Ceratocystis fimbriata           DIAGWNS-PNCGTCWKVTI--PNG--NSIFIRGVDSGKGAFNVNPDAFVKLVGSL-EAGRVENVNYVQVDLSNCINGAN-
```

Figura 4 Alinhamento das sequências de aminoácidos de cerato-plataninas selecionadas. A seta preta indica o resíduo conservado de cisteína.

CAPÍTULO III – ANÁLISE FILOGENÉTICA E EVOLUTIVA DA OXIDASE ALTERNATIVA DE *MONILIOPHTHORA PERNICIOSA*.

INTRODUÇÃO E OBJETIVOS.

Oxidases Alternativas (AOX) são enzimas localizadas na membrana interna da mitocôndria e capazes de transferir 4 elétrons diretamente da ubiquinona para o oxigênio molecular, produzindo uma molécula de água. Embora tenha sido estudada principalmente em plantas, a AOX tem uma distribuição bastante ampla, estando presente em vários grupos de organismos eucariotos, dentre eles fungos, animais e protistas, bem como em alguns procariotos. A via alternativa de fluxo de elétrons estabelecida pelas AOXs não resulta na produção de NADH, FADH ou no bombeamento de prótons para o espaço entre as membranas da mitocôndria, e portanto não é usada para produção de ATP. Contudo é capaz de produzir calor, função frequentemente relatada em plantas superiores (Watling *et al.* 2008), e também evitar a morte celular por inibidores da cadeia respiratória, que parece ser uma das suas principais funções em fitopatógenos (Joseph-Horne *et al.* 2001). Enzimas desta família podem ainda estar relacionadas à regulação da produção de energia (ATP) e à manutenção do metabolismo redox da célula (Vanlerberghe & McIntosh 1997). Em angiospermas, as AOXs são codificadas por uma família multi-gênica, tradicionalmente dividida em duas subfamílias, AOX1 e AOX2, sendo que a AOX2 é encontrada apenas em dicotiledôneas, enquanto a AOX1 aparece também em monocotiledôneas. Esta última é geralmente expressa em resposta a condições de stress, enquanto a AOX2 pode se apresentar como constitutiva ou pode ser diferencialmente expressa de acordo com as fases de desenvolvimento, como descrito em algumas espécies de dicotiledôneas (Karpova *et al.* 2002).

A detecção de uma oxidase alternativa expressa no genoma de *M. perniciosa* levou ao estabelecimento de um projeto visando verificar seu possível papel como elemento de resistência a resposta contra patógenos que se espera seja desencadeada no cacau quando da infecção pelo fungo.

Com este objetivo o presente trabalho buscou verificar, por meio de um estudo filogenético, se a estrutura de subfamílias gênicas encontrada nas AOXs de angiospermas tem paralelo em fungos. Além disso, foram estudados aspectos da evolução destas proteínas visando principalmente identificar a ocorrência de seleção natural diferencial entre as duas subfamílias de AOXs de angiospermas, bem como entre as AOXs de angiospermas e fungos e de fungos entre si, caso sejam detectadas subfamílias neste grupo. Por fim foram analisados os padrões de seleção em cada sítio da proteína, em busca de possíveis regiões de importância seletiva. Os resultados aqui descritos foram apresentados no 54º Congresso Brasileiro de Genética (Thomazella *et al.* 2008).

MATERIAL E MÉTODOS.

As sequências usadas para análise filogenética da oxidase alternativa foram localizadas por meio de tBLASTx contra a base de dado não redundante (nr) do GenBank, usando como “query” a região codificante da sequência de AOX de *M. perniciosa*. Foram obtidas 92 sequências oriundas de 54 espécies distintas de fungos e plantas (incluindo algas verdes), indicadas na tabela 3 do anexo. Também foram obtidas sequencias de AOX em oomicetos, animais e bactérias, mas estas não foram incluídas na filogenia. Devido a ampla variedade de sequências utilizadas estas foram inicialmente alinhadas por grupo (fungos e plantas) e a seguir os grupos foram alinhados uns aos outros. Todos os alinhamentos foram feitos pelo programa ClustalX (Thompson 1997), com o uso das opções “defaut” deste para proteínas, exceto pelo uso da matriz de transição JTT (Jones *et al.* 1992). Trechos pouco conservados ou com excesso de in/dels foram manualmente removidos usando-se o editor de alinhamentos do programa MEGA4 (Tamura *et al* 2007) obtendo-se no final cerca de 250 resíduos de aminoácidos confiavelmente alinhados. Apenas as regiões codificantes, obtidas manualmente a partir das anotações de cada sequência no GenBank, foram utilizadas na análise. As filogenias foram construídas por meio do algoritmo

de Máxima Verossimilhança com o uso do pacote PhyML (Guindon & Gascuel 2003) usando a matriz de transição JTT e permitindo variação nas taxas de mutação entre sítios modelada por uma aproximação discreta, com 4 classes, da distribuição gama. O suporte dos ramos foi verificado com o uso de 1000 bootstraps.

A partir das regiões comuns as AOX de fungos e plantas foram analisadas as razões entre as taxas de mutação não sinônima e sinônima (dN/dS) para cada sítio nos dois grupos separadamente, a fim de se verificar a ocorrência de sítios sob seleção. Adicionalmente verificou-se a ocorrência de pressões seletivas distintas entre as regiões descritas como conservadas e não conservadas em plantas (Vanlerberghe & McIntosh 1997) e entre as regiões equivalentes em fungos. Finalmente foram verificadas a existência de diferentes pressões seletivas entre as AOX1 e 2 de plantas, que já foram descritas como diferencialmente expressas, (Finnegan *et al.* 1997) e entre as AOX de plantas e de fungos. As análises de seleção foram realizadas por meio do programa HyPhy (Kosakovsky *et al.* 2005).

RESULTADOS.

Filogenia das AOXs de fungos e plantas.

A filogenia obtida (figura 5) confirma a subdivisão das AOXs de angiospermas em dois grupos, AOX1 e AOX2, já relatada na literatura, sendo que as AOX do tipo 2 são exclusivas de dicotiledôneas, enquanto as do tipo 1 aparecem também em monocotiledôneas. Esta divisão parece não se aplicar as demais plantas terrestres e às algas verdes, embora *Chlamydomonas reinhardtii* tenha duas cópias distintas do gene. Em fungos não foi encontrada este tipo de divisão em subfamílias. Embora algumas espécies deste grupo, como *Aspergillus nidulans* e *Coprinopsis cinerea*, também tenham mais de uma cópia do gene, na maioria dos casos em que isso ocorre as duas cópias estão agrupadas como grupos irmãos. Apenas as cópias de *Neurospora crassa* aparecem suficientemente apartadas para sugerirem

parálogos não espécie-específicos, mas nenhuma das outras sequências agrupadas com estas duas está duplicada.

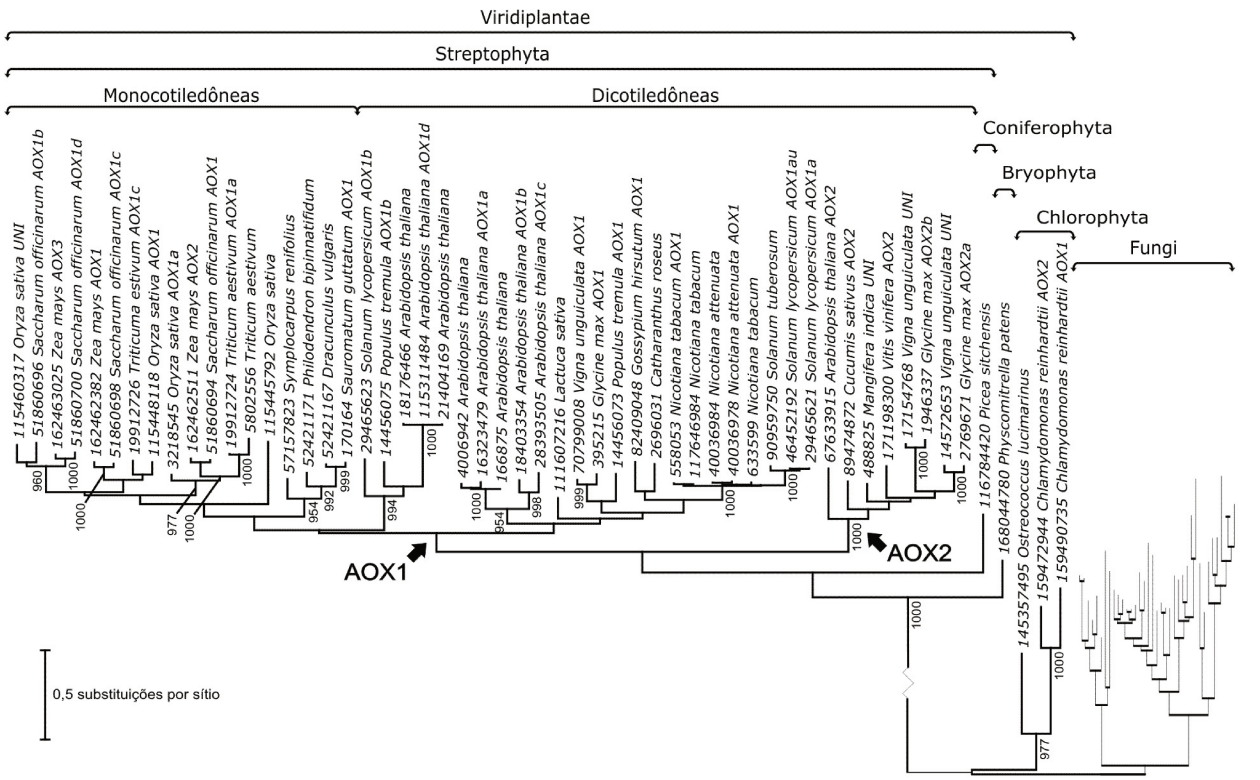
Alguns exemplos de sequências de AOX de plantas e fungos são mostradas na figura 6. As de angiospermas possuem duas cisteínas conservadas e, eventualmente, uma terceira, enquanto que as de fungos possuem apenas uma, em uma posição diferente das de plantas. *Chlamydomonas reinhardtii* e *Ostreococus lucimarinus* apresentam as mesmas cisteínas que os fungos, enquanto que a briófita *Physcomitrella patens* apresenta uma das cisteínas conservadas de angiospermas, mas não a outra, e a gimnosperma *Picea sitchensis* não apresenta nenhuma das duas cisteínas de angiospermas, mas apresenta a terceira.

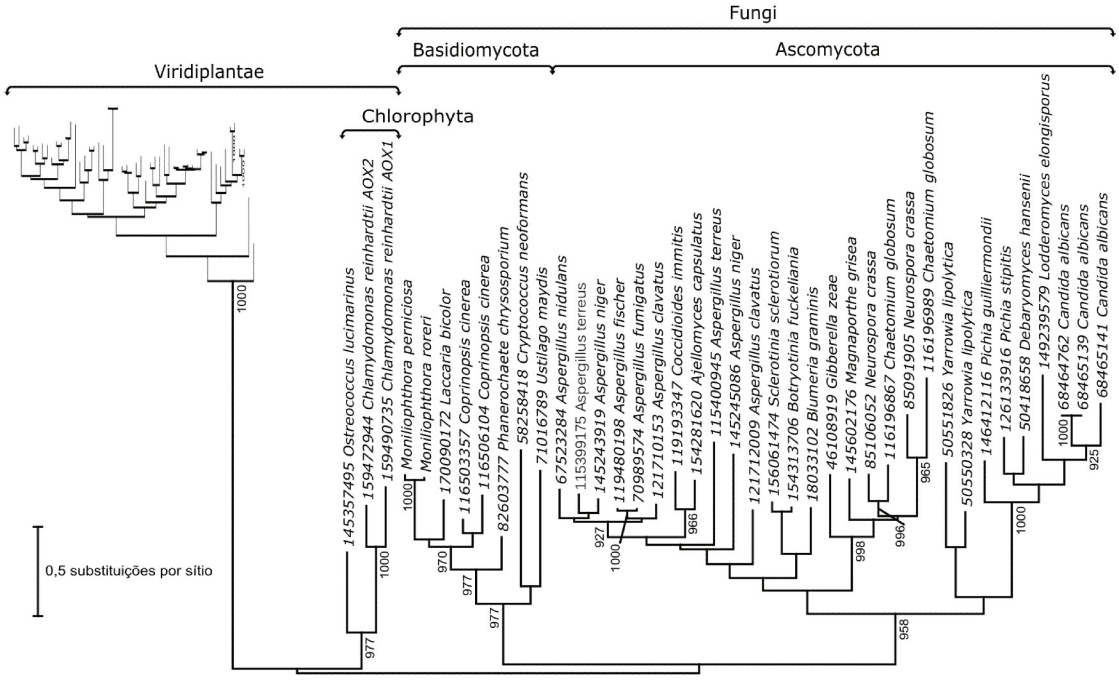
Análises de seleção em AOX de fungos e plantas.

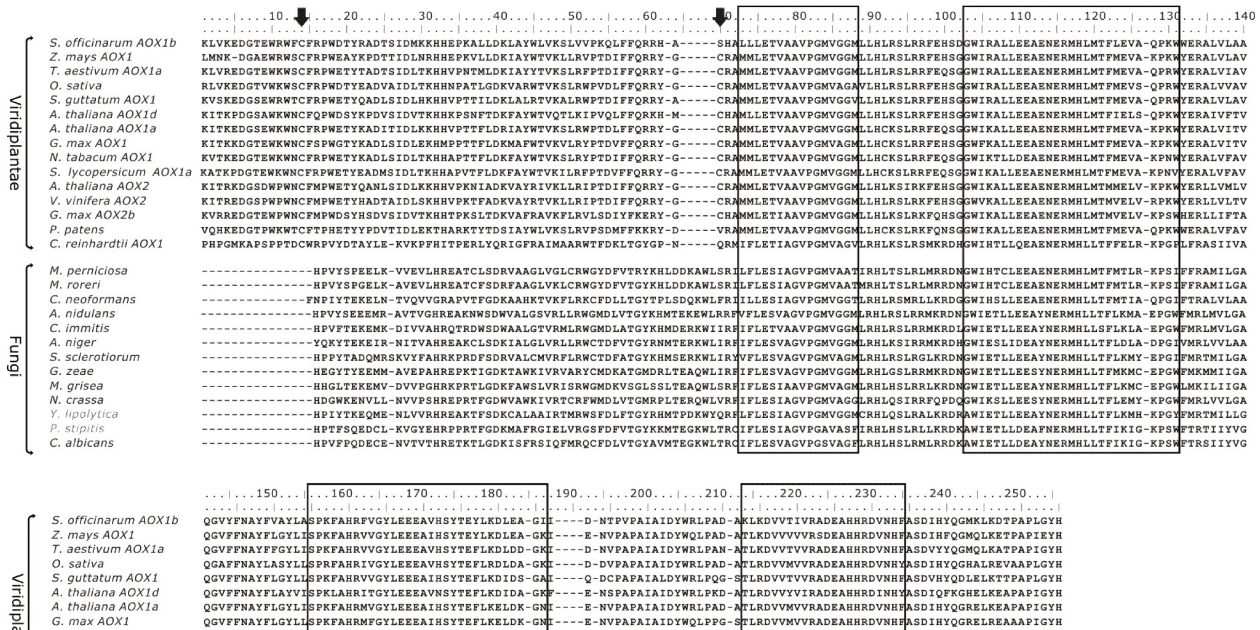
Tanto em fungos quanto em plantas todos os sítios analisados mostraram uma razão dN/dS significativamente menor que 1 (para $p<0,05$), sugerindo que estão sob seleção negativa (estabilizadora). Além disso, as regiões descritas como conservadas nas sequências de plantas (C1 a C4 na figura 6) tem uma razão dN/dS significativamente menor que as demais ($p<0,001$). Em fungos as regiões homólogas a elas também são significativamente mais conservadas que as demais ($p<0,001$).

Embora tanto plantas quanto fungos apresentem razões dN/dS menores que 1 para as sequências como um todo, estes últimos têm esta razão significativamente menor que as plantas ($p<0,001$), tanto dentro quanto fora das regiões conservadas (dN/dS=0,0342486 e 0,0762592 respectivamente para os fungos e dN/dS=0,0516574 e 0,110282 para as plantas), sugerindo que a seleção negativa (estabilizadora) atua mais fortemente em fungos. Por outro lado não foram encontradas diferenças significativas entre as razões dN/dS das AOX1 e 2 de plantas ($p=0,11$).

Figura 5 (na próxima página) Filogenia por máxima verossimilhança das sequências de AOX de plantas e fungos disponíveis no NCBI. Os valores de bootstraps maiores que 900/1000 são mostrados próximos aos respectivos ramos. Os ramos de origem das AOX1 e 2 estão indicados por setas. Os grupos taxonômicos indicados foram obtidos no NCBI. A raiz da árvore foi posicionada entre fungos e viridiplantae (plantas terrestres e clorófitas).


Viridiplantae
Streptophyta
Monocotiledôneas
Dicotiledôneas
Coniferophyta
Bryophyta
Chlorophyta
Fungi
115460317 Oryza sativa UNI
51860696 Saccharum officinarum AOX1b
162463025 Zea mays AOX3
518860700 Saccharum officinarum AOX1d
162462382 Zea mays AOX1
51860698 Saccharum officinarum AOX1c
19912726 Triticuma estivum AOX1c
115448118 Oryza sativa AOX1
3218545 Oryza sativa AOX1a
162462511 Zea mays AOX2
51860694 Saccharum officinarum AOX1
19912724 Triticum aestivum AOX1a
5802556 Triticum aestivum
115445792 Oryza sativa
57157823 Symplocarpus renifolius
52421171 Philodendron bipinnatifidum
52421167 Dracunculus vulgaris
170164 Sauromatum guttatum AOX1
29465623 Solanum lycopersicum AOX1b
14456075 Populus tremula AOX1b
18176466 Arabidopsis thaliana
115311484 Arabidopsis thaliana AOX1d
21404169 Arabidopsis thaliana
4006942 Arabidopsis thaliana
16323479 Arabidopsis thaliana AOX1a
166875 Arabidopsis thaliana
18403354 Arabidopsis thaliana AOX1b
28393505 Arabidopsis thaliana AOX1c
111607216 Lactuca sativa
70799008 Vigna unguiculata AOX1
395215 Glycine max AOX1
14456073 Populus tremula AOX1
82409048 Gossypium hirsutum AOX1
2696031 Catharanthus roseus
558053 Nicotiana tabacum AOX1
11764984 Nicotiana tabacum
40036984 Nicotiana attenuata
40036978 Nicotiana attenuata AOX1
633599 Nicotiana tabacum
90959750 Solanum tuberosum
46452192 Solanum lycopersicum AOX1au
29465621 Solanum lycopersicum AOX1a
67633915 Arabidopsis thaliana AOX2
89474872 Cucumis sativus AOX2
488825 Mangifera indica UNI
171198300 Vitis vinifera AOX2
17154768 Vigna unguiculata UNI
1946337 Glycine max AOX2b
14572653 Vigna unguiculata UNI
2769671 Glycine max AOX2a
116784420 Picea sitchensis
168044780 Physcomitrella patens
145357495 Ostreococcus lucimarinus
159472944 Chlamydomonas reinhardtii AOX2
159490735 Chlamydomonas reinhardtii AOX1
AOX1
AOX2
0,5 substituições por sítio


Fungi
Basidiomycota
Ascomycota
Viridiplantae
Chlorophyta
145357495 Ostreococcus lucimarinus
159472944 Chlamydomonas reinhardtii AOX2
159490735 Chlamydomonas reinhardtii AOX1
Moniliophthora perniciosa
170090172 Laccaria bicolor
116503357 Coprinopsis cinerea
116506104 Coprinopsis cinerea
82603777 Phanerochaete chrysosporium
58258418 Cryptococcus neoformans
71016789 Ustilago maydis
67523284 Aspergillus nidulans
115399175 Aspergillus terreus
145243919 Aspergillus niger
119480198 Aspergillus fischer
70989574 Aspergillus fumigatus
121710153 Aspergillus clavatus
119193347 Coccidioides immitis
154281620 Ajellomyces capsulatus
115400945 Aspergillus terreus
145245086 Aspergillus niger
121712009 Aspergillus clavatus
156061474 Sclerotinia sclerotiorum
154313706 Botryotinia fuckeliania
18033102 Blumeria graminis
46108919 Gibberella zeae
145602176 Magnaporthe grisea
85106052 Neurospora crassa
116196867 Chaetomium globosum
85091905 Neurospora crassa
116196989 Chaetomium globosum
50551826 Yarrowia lipolytica
50550328 Yarrowia lipolytica
146412116 Pichia guilliermondii
126133916 Pichia stipitis
50418658 Debaryomyces hansenii
149239579 Lodderomyces elongisporus
68464762 Candida albicans
68465139 Candida albicans
68465141 Candida albicans
0,5 substituições por sítio



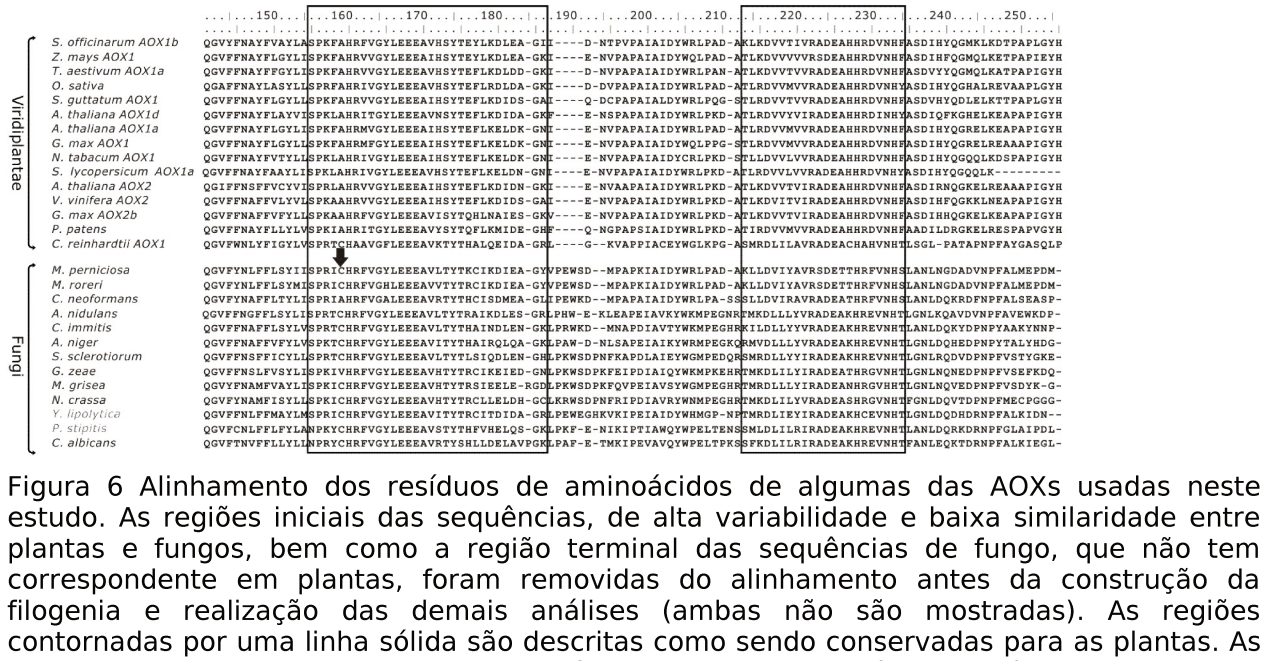


Figura 6 Alinhamento dos resíduos de aminoácidos de algumas das AOXs usadas neste estudo. As regiões iniciais das sequências, de alta variabilidade e baixa similaridade entre plantas e fungos, bem como a região terminal das sequências de fungo, que não tem correspondente em plantas, foram removidas do alinhamento antes da construção da filogenia e realização das demais análises (ambas não são mostradas). As regiões contornadas por uma linha sólida são descritas como sendo conservadas para as plantas. As setas pretas marcam a posição dos resíduos de histidina e ácido glutâmico ligantes do átomo de ferro do sítio ativo.

DISCUSSÃO.

As cisteínas conservadas nas AOXs de angiospermas aparentemente participam de um processo de dimerização que acontece nestas (Millenaar & Lambers, 2003). Esta dimerização por sua vez está ligada à regulação da AOX em plantas, que ocorre pela inativação da proteína na forma de dímero, na qual o sítio alostérico fica indisponível para o regulador (alfa-ceto ácidos). Em fungos a presença de apenas uma cisteína parece impedir esta dimerização, e a consequente regulação dela decorrente. Neste grupo a

enzima monomérica é regulada aparentemente por nucleotídeos purínicos atuando em um sítio alostérico ainda não conhecido (Joseph-Horne *et al.* 2001). Esta diferença sugere que as oxidases alternativas destes dois grupos podem ter mecanismos regulatórios distintas.

As regiões conservadas da AOX descritas em plantas estão relacionadas ao sítio ativo da proteína, que é formado por um centro binuclear contendo átomos de ferro (Moore *et al.* 1995; Siedow *et al.* 1995). O fato de que fungos apresentem seleção negativa (estabilizadora) mais forte nas mesmas regiões sugere que elas também tenham importância como sítio ativo neste grupo. Embora esta divisão em regiões conservadas e não conservadas não tenha sido descrita para o grupo o mesmo padrão de conservação das AOX de plantas é notado em fungos (figura 9), bem como de seleção. Isso sugere mecanismos moleculares de funcionamento da AOX similares, embora as funções fisiológicas nestes dois grupos sejam distintas.

A variação entre as sequências de fungos é maior que entre as das angiospermas. Este fato é condizente com a maior antiguidade da divergência entre os grupos de fungos do que entre os de angiospermas. Paradoxalmente a razão dN/dS é menor em fungos do que em plantas, sugerindo que entre os fungos as AOXs não devam apresentar tantas funções distintas quanto as de plantas. A ausência de subfamílias amplamente distribuídas em fungos também indica uma maior conservação da função desta no grupo.

CAPÍTULO IV – ANÁLISE FILOGENÉTICA E EVOLUTIVA DO GENOMA E DO PLASMÍDIO MITOCONDRIAL DE *MONILIOPHTHORA PERNICIOSA*.

INTRODUÇÃO E OBJETIVOS.

O genoma mitocondrial de *M. perniciosa* foi alvo de uma série de análises em separado dentro do projeto genoma deste fungo devido, entre outros fatores, a possibilidade que esta organela tenha importância em vários processos metabólicos relacionados ao ciclo infeccioso da vassoura de bruxa. Esta importância foi sugerida inicialmente pela detecção de senescência em culturas de *M. perniciosa*, um processo que, em fungos, pode estar relacionado com o comportamento de plasmídios mitocondriais (Osiewacz *et al.* 1989). Outra evidência desta importância é a ocorrência de oxidases alternativas (AOXs) na espécie. Estas proteínas estabelecem vias alternativas para o fluxo de elétrons na cadeia respiratória e podem estar relacionadas, entre outros, com processos de resistência do fungo ao stress oxidativo resultante da resposta contra patógenos por parte da planta (Vanlerberghe & McIntosh 1997).

As análises do genoma mitocondrial de *M. perniciosa* revelaram que este apresenta uma série de peculiaridades em relação aos de outros fungos. Entre estas destacam-se o seu grande tamanho (é o maior genoma mitocondrial de fungo descrito até o momento), a presença de extensas regiões intergênicas (nas quais foram detectadas numerosas ORFs) e a existência de um plasmídeo mitocondrial do tipo invertron permanentemente incorporado ao genoma. Estes plasmídios se caracterizam por serem lineares e formados por um gene para RNA polimerase e um gene para DNA polimerase em orientações opostas e flanqueados por repetições invertidas (Sakaguchi 1990).

Neste trabalho foram avaliados os padrões de evolução dos plasmídios mitocondriais do tipo invertron, homólogos ao encontrado em *M. perniciosa,* a fim de se verificar se estes estão subdivididos em mais de um grupo e se a transferência gênica horizontal teve influência na evolução destes a nível do

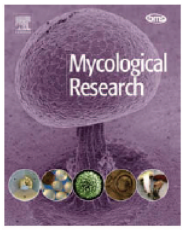
grupo como um todo, como já foi observado entre de espécies próximas (Bok *et al.* 1999; Griffiths 1998). Também foi construída uma filogenia dos genomas mitocondriais de fungos disponíveis no GenBank, usando-se as sequências de citocromo c oxidase. Por fim foi avaliada a possível ocorrência de conservação da ordem gênica (sintenia) entre os genomas mitocondriais de basidiomicetos disponíveis atualmente. Os resultados destas análises foram publicados como parte do artigo "The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stable integrated plasmid" (Formighieri *et al.* 2008). O artigo em questão é apresentado a seguir na sua íntegra. As referências bibliográficas são apresentadas juntamente com o artigo e não foram repetidas na seção de Bibliografia apresentada no final desta tese.

MYCOLOGICAL RESEARCH 112 (2008) 1136–1152

ELSEVIER

British Mycological Society promoting fungal science

journal homepage: www.elsevier.com/locate/mycres



The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stable integrated plasmid

Eduardo F. FORMIGHIERI[a], Ricardo A. TIBURCIO[a], Eduardo D. ARMAS[b], Francisco J. MEDRANO[a], Hugo SHIMO[a], Nicolas CARELS[c], Aristóteles GÓES-NETO[d], Carolina COTOMACCI[a], Marcelo F. CARAZZOLLE[a], Naiara SARDINHA-PINTO[a], Daniela P. T. THOMAZELLA[a], Johana RINCONES[a], Luciano DIGIAMPIETRI[e], Dirce M. CARRARO[f], Ana M. AZEREDO-ESPIN[g], Sérgio F. REIS[h], Ana C. DECKMANN[a], Karina GRAMACHO[i], Marilda S. GONÇALVES[j], José P. MOURA NETO[j], Luciana V. BARBOSA[k], Lyndel W. MEINHARDT[l], Júlio C. M. CASCARDO[m], Gonçalo A. G. PEREIRA[a,]*

[a]*Laboratório de Genômica e Expressão – Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, 13083-970, Campinas – SP, Brazil*
[b]*Laboratório de Ecotoxicologia, Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, 13400-970, Piracicaba – SP, Brazil*
[c]*Laboratório de Bioinformática da Universidade Estadual de Santa Cruz, 45650-000, Ilhéus – BA, Brazil*
[d]*Laboratório de Pesquisa em Microbiologia (LAPEM), Departamento de Ciências Biológicas, Universidade Estadual de Feira de Santana (UEFS), 44031-460, Feira de Santana – BA, Brazil*
[e]*Instituto de Computação, Universidade Estadual de Campinas, 13084-971, Campinas – SP, Brazil*
[f]*Ludwig Institute For Cancer Research, 01509-010, São Paulo – SP, Brazil*
[g]*Departamento de Genética e Evolução e Laboratório de Genética Animal, Centro de Biologia Molecular e Engenharia Genética – CBMEG, Universidade Estadual de Campinas, 13035-875, Campinas – SP, Brazil*
[h]*Departamento de Parasitologia, Instituto de Biologia, Universidade Estadual de Campinas, 13083-970, Campinas – SP, Brazil*
[i]*CEPLAC/CEPEC/SEFIT, 45600-970, Itabuna – BA, Brazil*
[j]*Laboratório de Biologia Molecular – Faculdade de Farmácia, Universidade Federal da Bahia, 40170-290, Salvador – BA, Brazil*
[k]*Laboratório de Biologia Molecular – Departamento de Biologia Geral, Instituto de Biologia, Universidade Federal da Bahia, 40170-290, Salvador – BA, Brazil*
[l]*Sustainable Perennial Crops Laboratory, USDA-ARS, 10300 Baltimore, Av., BARC-W, Beltsville, MD 20740, USA*
[m]*Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, 45650-000, Ilhéus – BA, Brazil*

ARTICLE INFO

Article history:
Received 29 October 2007
Received in revised form
19 March 2008
Accepted 24 April 2008
Corresponding Editor:
Gareth W. Griffith

ABSTRACT

We present here the sequence of the mitochondrial genome of the basidiomycete phytopathogenic hemibiotrophic fungus *Moniliophthora perniciosa*, causal agent of the Witches' Broom Disease in *Theobroma cacao*. The DNA is a circular molecule of 109 103 base pairs, with 31.9 % GC, and is the largest sequenced so far. This size is due essentially to the presence of numerous non-conserved hypothetical ORFs. It contains the 14 genes coding for proteins involved in the oxidative phosphorylation, the two rRNA genes, one ORF coding for a ribosomal protein (*rps3*), and a set of 26 tRNA genes that recognize codons for all

* Corresponding author. Tel.: +55 19 35216650.
E-mail address: goncalo@unicamp.br

doi:10.1016/j.mycres.2008.04.014

Keywords:
Basidiomycota
Codon usage
Crinipellis perniciosa
Genomics
Marasmiaceae
Witches' broom disease

amino acids. Seven homing endonucleases are located inside introns. Except *atp8*, all conserved known genes are in the same orientation. Phylogenetic analysis based on the *cox* genes agrees with the commonly accepted fungal taxonomy. An uncommon feature of this mitochondrial genome is the presence of a region that contains a set of four, relatively small, nested, inverted repeats enclosing two genes coding for polymerases with an invertron-type structure and three conserved hypothetical genes interpreted as the stable integration of a mitochondrial linear plasmid. The integration of this plasmid seems to be a recent evolutionary event that could have implications in fungal biology. This sequence is available under GenBank accession number AY376688.

Introduction

Moniliophthora perniciosa is the causal agent of witches' broom disease (WBD) in cacao (*Theobroma cacao*). This fungus is a basidiomycete previously classified as *Crinipellis perniciosa*, and recently, it was reclassified due to its close relation to *Moniliophthora roreri* (Aime & Phillips-Mora 2005; Evans *et al.* 2002), the causal agent of frosty pod rot (FPR) in cacao. These two represent the two most devastating cacao diseases in the Americas. *M. perniciosa* presents a hemibiotrophic life style, with differentiated mycelia in each phase. The biotrophic mycelia are mononucleated, grow very slowly and colonize the intercellular spaces (apoplast) of the infected plant tissue. This phase is difficult to obtain in the laboratory, but recently we have characterized the conditions for its *ex planta* growth (Meinhardt *et al.* 2006). The saprophytic mycelia are dikaryotic with clamp connections and grow vigorously in axenic cultures.

Due to the economical importance, our group is involved in the study of the complex biology of the hemibiotrophic fungus *M. perniciosa* (Garcia *et al.* 2007; Meinhardt *et al.* 2006; Rincones *et al.* 2003; Rincones *et al.* 2006; Rincones *et al.* in press; Rio *et al.* 2008; Scarpari *et al.* 2005). Included in this investigation is a genome project (http://www.lge.ibi.unicamp.br/vassoura), and during the course of the sequencing we obtained the complete mitochondrial genome.

Besides *M. perniciosa*, 49 fungal mitochondrial genomes have been completely sequenced and annotated to date (http://ww.ncbi.nlm.nih.gov/genomes/ORGANELLES/futax_short.html). These genomes encompass species from all phyla. Despite the importance of the *Basidiomycota*, only five mitochondrial genomes from this group have been studied.

Mitochondria are generally accepted as vestigial endosymbiotic proteobacterial ancestors (Gray *et al.* 1999; Lang *et al.* 1999), and they can evolve faster than nuclear DNA (Burger *et al.* 2003; Burger & Lang 2003; Nosek & Tomaska 2003). The relationship between mitochondria and the nucleus has evolved over time and several mitochondrial genes encoding proteins for mitochondrial metabolism have been transferred to the nuclear genome (Adams & Palmer 2003; Burger *et al.* 2003; Dimmer & Rapaport 2008). Fungal mitochondrial genomes are smaller than those of plants, but larger than animal mitochondrial genomes (Burger *et al.* 2003). The gene content of the typical fungal mitochondrial genome is greatly reduced (*ca* 30–40 genes) in comparison to the genomes of bacteria and the mitochondrial genomes of other eukaryotes. Some conserved characteristics of mitochondrial genomes are the high A + T content, lack of methylation, conservation of the gene function, and high copy number (Campbell *et al.* 1999).

Fungal mitochondrial genomes vary greatly in size, show different topology, and even in the use of the genetic code among species (Gray *et al.* 1999). The number of genes can fluctuate enormously from three to 67 for protein encoding genes and from seven to 27 for tRNAs (Adams & Palmer 2003). They usually contain 14 genes encoding hydrophobic subunits of respiratory chain complexes, two genes for the large and small ribosomal subunits and a set of tRNAs (Gray *et al.* 1999). Differences in length and organization of the intergenic regions, differences in intron content (from 0 to 30) and size (ranging from 0.15 to 4 kb), unidentified ORFs, and the presence of plasmids contribute significantly to the size variability of the genomes among species (Bertrand 2000; Burger *et al.* 2003; Clark-Walker 1992; Hur *et al.* 1997). Also, there is no apparent correlation between size and gene content in the mitochondrial genomes (Burger *et al.* 2003).

The shortage of reliable morphological and ultrastructural characters is a challenge in the taxonomic classification of fungi. However, in recent years the use of mitochondrial sequence data of proteins involved in essential processes conserved throughout evolution have been used routinely to build phylogenetic trees for the characterization of the origin and evolution of fungal organisms (Bullerwell *et al.* 2003; Bullerwell & Lang 2005; Seif *et al.* 2005). Therefore, more complete mtDNA sequences from *Basidiomycota* may provide insights to the origin and evolution of fungal mitochondria.

Plasmids have been found in many different fungi (Griffiths 1995; Kempken 1995a; Meinhardt *et al.* 1990, 1997). All of the plasmids found in these fungi are mitochondrial with the possible exception of the *Alternaria* plasmids (Shepherd 1992). Plasmids are of two types, circular and linear. In *Neurospora* spp., both circular and linear types are found. The plasmids of *Absidia glauca* are circular (Haenfler *et al.* 1992) and all the remaining examples are linear. Some features are common to almost all of the mitochondrial linear plasmids, namely: (1) terminal inverted repeats; (2) proteins covalently attached to the 5′ termini; and (3) open reading frames (ORFs) for viral-like RNA and DNA polymerases (Meinhardt *et al.* 1990, 1997). Thus, these linear plasmids in filamentous fungi are thought to be descended from a common ancestor. However, horizontal transfer of a linear plasmid in *Ascomycota* is also shown (Kempken 1995b).

Although in most cases no function has been assigned for these elements, in some species of filamentous fungi their

presence and activity interferes with the growth of mycelia (Bertrand *et al.* 1985, 1986). In *Neurospora* spp. insertion of the plasmid DNA into the mitochondrial genome results to the disruption of several genes leading to senescence and death due to defective respiration (Bertrand 2000; Griffiths 1992, 1998; Rieck *et al.* 1982). Conversely, in *Podospora anserina* strain AL2 and in the slime mould *Physarum polycephalum*, which are species that normally senesce, insertion of the plasmids into the mtDNA has been correlated with an increased life-span (Hermanns *et al.* 1994; Maas *et al.* 2007; Nakagawa *et al.* 1998). Among *Basidiomycota*, invertron-like DNA plasmids have been completely sequenced in *Agaricus bitorquis* (Robison & Horgen 1999), *Flammulina velutipes* (Nakai *et al.* 2000), and *Pleurotus ostreatus* (Kim *et al.* 2000), but no specific function has been associated with their presence in these hosts.

The present work is a comprehensive study of the *M. perniciosa* mitochondrial genome. We describe its organization, gene content, and order, together with a comparative and phylogenetic analysis with other sequenced mitochondrial genomes. We provide evidence for the presence of potentially new mitochondrial genes and for the evolution of the mitochondrial genome by the stable integration of a linear mitochondrial plasmid.

Material and methods

Fungal isolate, library construction and sequence assembly

Total DNA was extracted from strain CP02 of *Moniliophthora perniciosa*, as described previously (Rincones *et al.* 2003). We observed a relation between the integrity of the isolated DNA and the age of the culture used for the extraction (the older the plate, the worst the quality of the DNA). For that reason, a total of 50 genomic libraries were constructed using DNA isolated from young and old mycelia. Fragments ranging in size from 1–2 and 2–4 kbp were obtained from total DNA by sonication or nebulization (Surzcki 2000), and then cloned into the *Sma*I site of pUC18 or pCR4Blunt (TOPO Shotgun Subcloning kit, Invitrogen - Life Technologies, São Paulo, Brasil). Each library corresponded to independent cloning events using DNA obtained from individually growing cultures of CP02. The inserts were sequenced and analysed following the whole genome shotgun (WGS) approach (Venter *et al.* 1998).

Reads were assembled using the software package Phred/Phrap/Consed (Ewing & Green 1998; Ewing *et al.* 1998; Gordon *et al.* 1998) and the accuracy of the assembly was confirmed with the program CAP3 (Huang & Madan 1999).

All isolates used in this study can be obtained from the Laboratório de Genômica e Expressão at UNICAMP, as well as further information concerning the collection data.

Gene annotation

For the sake of clarity, we have defined the ORFs found in the mitochondrial genome as follows: (1) conserved known genes, as those normally present in mitochondrial genomes with known function; (2) conserved hypothetical ORFs, as those ORFs present also in other mitochondrial genomes, but without a defined function; and (3) non-conserved hypothetical ORFs, as those ORFs that do not overlap with conserved known genes or conserved hypothetical ORFs, and coding for products with minimum size compatible with the smallest conserved known genes (*atp8* has 52 aa, *atp9* has 73 aa and *nad4L* has 87 aa).

Conserved known genes have been named as their counterparts in other genomes. Intronic conserved hypothetical ORFs have been named as *oi#gene*. Intrinic non-conserved hypothetical ORFs have been named as *oBi#gene*. Conserved hypothetical ORFs located in the plasmid region have been named *hypP#*. Non-conserved hypothetical ORFs have been named as *hyp#*. Here # represents a number and gene is the name of the gene where it its located.

Conserved known genes and conserved hypothetical ORFs were detected by similarity searches for known genes and ORFs present in other fungal mitochondrial genomes using BLAST (Altschul *et al.* 1990). Non-conserved hypothetical ORFs were detected using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/).

Regions coding for tRNAs were identified with the program tRNAscan-SE (Lowe & Eddy 1997). The rRNAs sequences were identified by comparison with homologous *Basidiomycota* counterparts. Repeats were analysed using Reputer (Kurtz *et al.* 2001) and Tandem Repeats Finder (Benson 1999).

Exon/intron boundaries of conserved known genes were identified by comparison and alignment, using BLAST and CLUSTAL (Higgins 1994), with the genes from other fungal mitochondrial genomes.

PCR amplification

Amplification of selected regions of the mitochondrial genome was carried out by PCR. Primers P1 (5′-GCA GGGAAGGGATATATAGG-3′), P2 (5′-TTTGAGAGAGCATCAAAT CC-3′), P3 (5′-TTTTGAGAG AGCATCAAATCC-3′), and P4 (5′-AA AGAACTGAAATCCGAGG-3′) were designed to mach regions outside (P1 and P4) and inside (P2 and P3) of the inverted repeats that we did use to define the boundaries of the inserted plasmid (see Results). Total DNA from isolates in Table 2 purified as described above was used for the PCR experiments.

GC content, GC skew, and codon usage

GC content and GC skew (Grigoriev 1998) were calculated using the program Analyse seq (http://ludwig-sun2.unil.ch/~vioannid/TP_module5/18-03/freak2.html). Local GC content was calculated with sliding windows over 5000 bp and 500 bp with a pace of 500 and 50 bp, respectively. Codon usage was analysed using the program CODONW (Peden 1999).

Principal component analysis (PCA)

A matrix presenting the relative use of each codon for each ORF was evaluated by PCA in order to search for patterns of codon usage (Joliffe 1986). These data were submitted to analysis by the method of singular value decomposition of non-centred and non-scaled data matrix by means of the R software (R Development Core Team 2007). The number of principal components to be retained was determined by

scree-plot and the percentage of accumulated variation explained by the components.

Phylogenetic analysis

Protein sequences were aligned using ClustalW (Thompson *et al.* 1994) with default options, except for the use of the JTT substitution matrix (Jones *et al.* 1992). Terminal and internal poorly aligned regions were removed from the alignment before the phylogenetic analysis. Phylogenetic trees were constructed with the program PhyML (Guindon & Gascuel 2003) using the ML method, with JTT substitution matrix and gamma distribution. The value of the shape parameter α used was optimized by the program PhyML. Branch support was tested by BS analysis with 100 resamplings (Felsenstein 1985).

Genes that code for the mitochondrial cytochrome c oxidase (*cox1*, *cox2* and *cox3*), properly annotated, from 50 fungi with complete sequence for the mitochondrial genome were used to build the phylogenetic tree. The optimized value of the shape parameter α was 1.444 for this set of sequences.

Also, the genes *dpoB* and *rpo* from the mitochondrial plasmid from ten species that present a complete invertron structure (both genes in the opposite orientation flanked by inverted repeats) (Griffiths 1995) were used to build a phylogenetic tree. A total of 13 complete plasmids were used in this analysis. Three of the species used present more than one copy of the plasmid (*Neurospora crassa*, *Pleurotus ostreatus*, and *Flammulina velutipes*). Sequences with similarity to *dpoB* or *rpo* not organized with an invertron structure were not included in this analysis. The optimized value of the shape parameter α was 1.549 for this set of sequences.

Comparative genomics

The comparison of the gene content and order of all known mitochondrial genomes of *Basidiomycota* was carried out by visual inspection due to the small number of genomes used, six. All conserved known genes coding for proteins and the rRNAs were designed as boxes linearly as they appear in the genome beginning with the large rRNA (*ml*) in a clockwise direction.

Results

Assembly of the mitochondrial genome

After the initial assembly of 124 565 reads produced in the sequencing of all the libraries a large contig enclosing 5448 reads was generated. This contig was identified as the mitochondrial genome of the fungus by comparison to other fungal mitochondrial genomes. Inclusion of additional reads from new sequencing rounds and a careful assembly resulted in a final contig containing 6920 reads. The consensus sequence corresponded to a circular DNA of 109 103 bp without gaps, regions of low consensus quality, or regions of high-quality discrepancy (Phrap assembly). This corresponded to an approximate coverage of 31 times. The complete mtDNA sequence is available at GenBank (accession number AY376688).

Although most fungal mitochondrial genomes are circular (Torriani *et al.* 2008; Wang *et al.* 2008; Zivanovic *et al.* 2005), some linear genomes have been reported (Forget *et al.* 2002; Rycovska *et al.* 2004). Due to the programs we use not being able to assemble circular DNAs we have used a different strategy in order to check whether the mitochondrial genome of *M. perniciosa* is linear or circular. The strategy consisted in the masking of unique regions (minimum length of 750 bp) of the genome and new assemblies of all the reads were performed with the exception of the masked region. This was repeated masking different regions independently. If the mitochondrial genome is linear this strategy should render at least two different contigs representing the regions on both sides of the masked region. The result was always one contig comprising all the genome except the masked region; thus indicating that the mitochondrial genome is circular.

Conserved known genes

Fig 1 shows the map of the 119 possible coding regions present in the mitochondrial genome. Seventeen genes were classified as conserved known genes, coding for proteins were identified by comparison with other fungal mitochondrial genomes where the function of those have been described. Fourteen of those are involved in the oxidative and energy metabolism; seven subunits of the NADH dependent dehydrogenases, three cytochrome c oxidases, three F0 subunits of the ATP synthases and one cytochrome b. One gene, *rps3*, is involved in ribosome assembly (Bullerwell *et al.* 2000). The last two code for polymerases, a DNA-dependent RNA polymerase (*rpo*) and a DNA-directed DNA polymerase (*dpoB*).

Two rRNAs and 26 tRNAs were identified in the mitochondrial genome. These tRNAs are grouped into regions of the genome presenting a relatively high GC content (Fig 3). The location of the tRNAs is shown at the most inner circle in Fig 1, and Table 1 list all the tRNAs anti-codons with its corresponding amino acid.

Conserved and non-conserved hypothetical ORFs

Twelve additional putative ORFs have been defined as conserved hypothetical ORFs based on the fact of the existence of similar sequences in other fungal mitochondrial genomes. Three of them are present between a set of inverted repeats together with the genes *dpoB* and *rpo*, and the other nine are intronic ORFs located inside the genes that code for *cox1*, *cob*, *nad4* and *nad5* (Fig 1).

The analysis of the remainder of the sequence with the program ORF Finder located 59 possible non-conserved hypothetical ORFs and three non-conserved intronic ORFs located inside the genes *nad5*, *cob* and *cox2* (Fig 1). One copy of the LAGLIDADG motif was found in one hypothetical ORF, hyp15.

Introns

Twelve introns have been identified: *cox1* has six introns, *cox2* and *cob* have two introns each, and both *nad4* and *nad5* have one intron each (Fig 1). All the six introns in *cox1*, one of the introns in *cob*, and the intron in *nad4* and *nad5* contain one conserved hypothetical ORF each. Besides these conserved

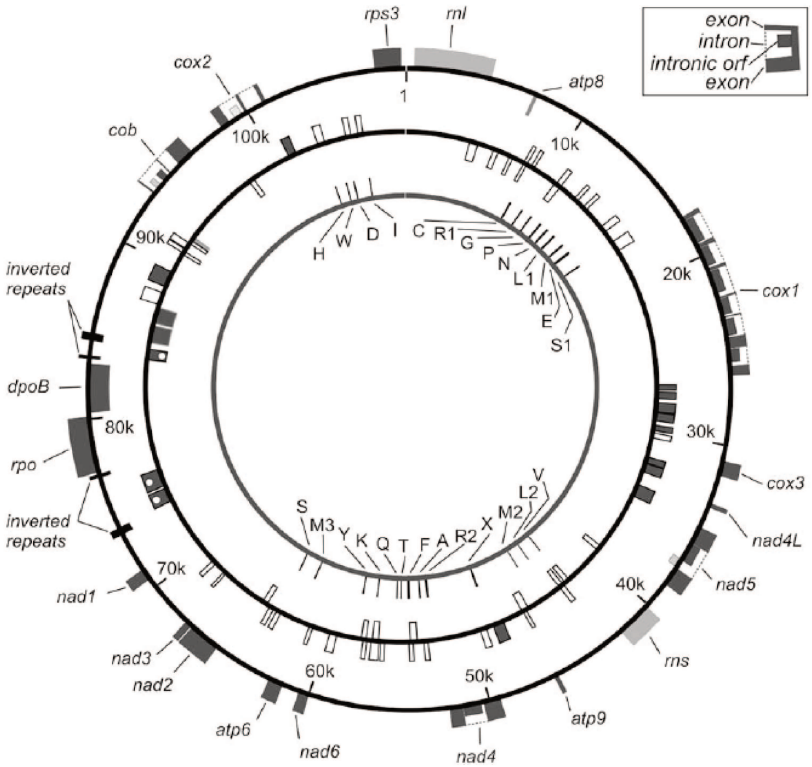


Fig 1 – Genetic map of the mitochondrial genome from *Moniliophthora perniciosa*. At the outermost circle the conserved known genes (protein coding genes as black boxes and rRNAs in grey, the name of each gene is indicated) and the intronic ORFs (conserved hypothetical in black and non-conserved hypothetical in grey) are represented (small boxes inside the gene). All other conserved and non-conserved hypothetical ORFs are represented at the middle circle (genes with the same codon usage as the conserved known genes as black boxes, the rest of the hypothetical genes are in grey). The three conserved hypothetical ORFs present in the plasmid are indicated by a white dot inside the box. The tRNAs, labelled by their one letter code corresponding amino acid, are represented at the innermost circle (X represents the undefined anti-codon). Genes outside the circle are oriented clockwise and those inside are anticlockwise.

hypothetical ORFs, *cob*, *cox2*, and *nad5* contain one non-conserved hypothetical ORF each.

Two different motifs belonging to two types of homing endonucleases have been found in eight of the intronic conserved hypothetical ORFs. Two copies of the LAGLIDADG motif were found in six of them (*oi1cox1*, *oi2cox1*, *oi4cox1*, *oi5cox1*, *oi1nad5*, and *oi1nad4*), and *oi1cob* presented only one of such motifs (Heath *et al.* 1997). Inside the intronic ORF *oi6cox1* one GIY-YIG motif was found (van Roey *et al.* 2001).

Codon usage and ORF orientation

The preferential codon usage was calculated from the exon sequences of the conserved known genes (Table 1). At least one tRNA for every amino acid has been identified, and for one of them it was not possible to determine the anti-codon. Most of the codons can be translated with tRNAs present in the mitochondrial genome by the wobble rule (Jukes 1984). There are two codons, AUA and UGA, without the corresponding tRNA.

When the PCA was applied to all the potential protein-coding genes (91 known and hypothetical ORFs) a clear pattern could be identified (Fig 2). The genes segregated into two different groups (Fig 2, left panel). One first group (A) was formed up by 45 non-conserved hypothetical ORFs, and the three intronic non-conserved hypothetical ORFs: *oBi1nad5*, *oBi1cob* and *oBi1cox2*; and a second group (B) made up by the remaining 43 ORFs; including all conserved known genes (17), nine conserved hypothetical ORFs, and 14 non-conserved hypothetical ORFs.

A significant number of ORFs belonging to group A are oriented in the anticlockwise direction (18 out of 48). The remaining 30 ORFs show a clockwise orientation. Most of the ORFs (38 out of 43) in group B are oriented in the clockwise direction; these include most the known conserved genes (15 out of 17), all of the intronic conserved ORFs (9), two conserved hypothetical ORFs and 12 non-conserved hypothetical ORFs. It is likely that these non-conserved, hypothetical ORFs represent true genes due to the fact that they have the same preferential codon usage and the same clockwise orientation as the conserved known genes. Only five ORFs are found in the

Table 1 – Codon usage in the conserved known genes from the mitochondrial genome

Amino acid	Codon	E	Anti-codon	Amino acid	Codon	E	Anti-codon	Amino acid	Codon	E	Anti-codon	Amino acid	Codon	E	Anti-codon
Phe	UUU	334	(g)aa	**Ser**	UCU	167	(u)ga	**Tyr**	UAU	217	(g)ua	**Cys**	UGU	32	(g)ca
	UUC	96	**GAA**		UCC	16	(u)ga		UAC	39	**GUA**		UGC	4	**GCA**
Leu	UUA	545	**UAA**		UCA	140	**UGA**	**Ter**	UAA	25	*	**Trp**	UGA	41	-
	UUG	52	(u)aa		UCG	9	(u)ga		UAG	14	*		UGG	21	**CCA**
	CUU	79	(u)ag	**Pro**	CCU	109	(u)gg	**His**	CAU	64	(g)ug	**Arg**	CGU	1	(u)cg
	CUC	5	(u)ag		CCC	3	(u)gg		CAC	36	**GUG**		CGC	1	(u)cg
	CUA	72	**UAG**		CCA	47	**UGG**	**Gln**	CAA	111	**UUG**		CGA	10	**UCG**
	CUG	21	(u)ag		CCG	3	(u)gg		CAG	14	(u)ug		CGG	1	(u)cg
Ile	AUU	270	(g)au	**Thr**	ACU	135	(u)gu	**Asn**	AAU	235	(g)uu	**Ser**	AGU	114	(g)cg
	AUC	44	**GAU**		ACC	5	(u)gu		AAC	34	**GUU**		AGC	19	**GCU**
	AUA	286	-		ACA	118	**UGU**	**Lys**	AAA	184	**UUU**	**Arg**	AGA	78	**UCU**
Met	AUG	102	**CAU**		ACG	4	(u)gu		AAG	18	(u)uu		AGG	1	(u)cu
Val	GUU	135	(u)ac	**Ala**	GCU	164	(u)gc	**Asp**	GAU	119	(g)uc	**Gly**	GGU	136	(u)cc
	GUC	5	(u)ac		GCC	13	(u)gc		GAC	20	**GUC**		GGC	9	(u)cc
	GUA	160	**UAC**		GCA	91	**UGC**	**Glu**	GAA	111	**UUC**		GGA	138	**UCC**
	GUG	20	(u)ac		GCG	13	(u)gc		GAG	16	(u)uc		GGG	12	(u)cc

The tRNAs that can potentially match codons by the wobble rule are indicated by small caps in parentheses. Missing tRNAs are indicated by a dash, and the stop codons are indicated by asterisks.

anticlockwise direction; two conserved known genes (*atp8* and *dpoB*), one conserved hypothetical ORF (*hypP3*), and two non-conserved hypothetical ORFs. These three hypothetical ORFs are located immediately upstream of the gene *dpoB*. Also, all the tRNAs and the two rRNAs are in the clockwise orientation.

GC content

The mitochondrial genome presents a relatively low GC content (31.9 %) in comparison to that of the genomic DNA (47.7 %), calculated using all the non-mitochondrial contigs. Comparing the GC average for positions one (GC1), two (GC2), and three (GC3) of conserved protein coding genes and for all mtDNA (GCm), we found that GC3 was significantly biased towards lower values (supplementary data), which has been observed in protein coding genes in some mitochondrial genomes (Stewart & Beckenbach 2006).

The GC content and the cumulative GC skew of the mitochondrial genome are shown in Fig 3. The GC content (Fig 3, blue and black lines) presents some variation along the mitochondrial genome with three well-defined significant regions. It begins with a content between 35 and 40 %, maintaining this

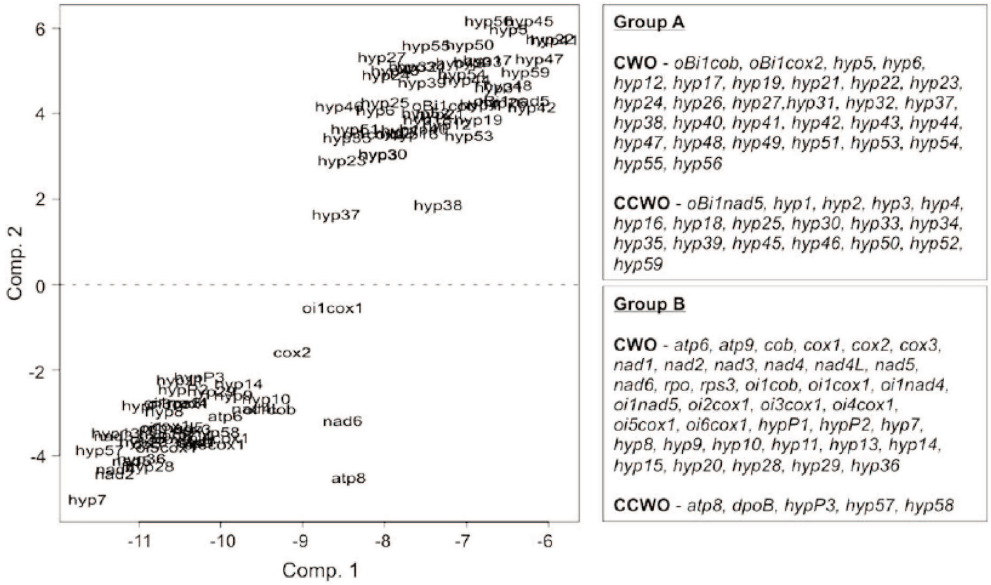


Fig 2 – PCA of codon usage. Left panel shows a plot of the sequence distribution according to codon usage (two-component solution accounting for 78.1 % of the total variance; 65.4 and 12.7 % from components one and two, respectively). Right panel shows the relations of ORFs classified by group and orientation.

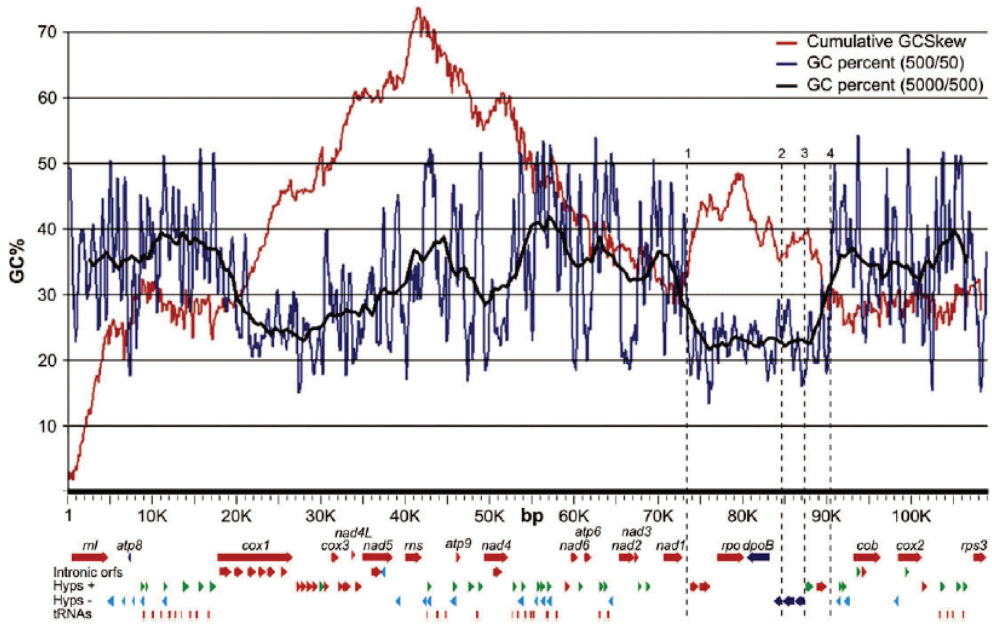


Fig 3 – GC content and GC skew along the mitochondrial genome. Blue and black lines represent the local GC content calculated with windows of 500 and 5000 bp, respectively. Red line represents the cumulative GC skew. The position and orientation of the ORFs is indicated at the bottom of the figure. Dashed lines 1 and 2 represent the position of the large internal repeats, and lines 3 and 4 represent the two groups of non-conserved hypothetical ORFs (*hyp58–hyp57* and *hyp28–hyp29*) with opposite orientation and inside the region of low GC content. The position of the ORFs is represented at the bottom of the figure.

average up till the 19 kbp position where it begins to drop down to around 24 % at the 28 kbp position; then it begins to raise up to around the previous level at the 45 kbp position where it drops again to a minimum at the 50 kbp position just under the 30 % content; then it goes up to its maximum local content just over 40 % between the 55 to the 59 kbp position; then the content goes back to the 35 % average content up to the 71 kbp position where it drops down, for the last time, to its lower value (around 23 %) at the 76 kbp position; it remains at this low value up till the 88 kbp position where it raises again up to an average content of 35 % at the 92 kbp position, and remains around this value till the end.

The cumulative GC skew (Fig 3, red line) raises up to a maximum at the 42 kbp position and then drops down up to the 72 kbp position where it goes up again until generating a new local maximum between the previous position and the 90 kbp position; after this position levels up and its relatively constant up till the end of the sequence.

The most significant feature of this analysis comes from the combined observation of the GC content and the GC skew. The region located between the 73 kbp and the 90 kbp positions (Fig 3, between dotted lines 1 and 4) is characterized by the lowest GC content and a change in the direction of the cumulative GC skew (red line). Also conspicuous is the significant reduction in the noise of the profile of the GC content calculated with the smaller window (500 bp) of this region with respect to the rest of the genome (blue line). This kind of pattern has been interpreted as regions of genome rearrangement (Grigoriev 1998). This region contains nine ORFs and four sets of inverted repeats. It begins with one set of inverted repeats 347 bp long (Fig 3, dotted line 1), continues with two conserved hypothetical ORFs (*hypP1* and *hypP2*), then a set of inverted repeats 130 bp long followed by *rpo* and *dpoB*, then another set of inverted repeats 130 bp long followed by one conserved hypothetical ORF (*hypP3*), then the last set of inverted repeats (Fig 3, dotted line 2) appears followed by two non-conserved hypothetical ORFs (*hyp58* and *hyp57*) with a anticlockwise orientation (Fig 3, dotted line 3), and then the region ends with two more non-conserved, hypothetical ORFs (*hyp28* and *hyp29*) with a clockwise orientation (Fig 3, dotted line 4). The region delimited by the two bigger sets of inverted repeats that contains the genes *rpo* and *dpoB*, together with three conserved hypothetical ORFs (Fig 3, dotted lines 1 and 2), resembles the structure of a linear plasmid (Griffiths 1995).

The analysis of the GC skew in mammals (Touchon *et al.* 2005) and bacteria (Arakawa *et al.* 2007) has proved useful in the determination of putative origin and termini of replication, indicated by a change in the polarity of this index. Despite not having any experimental data about the location of the origin of replication, we are proposing that the putative origin of replication of the mitochondrial genome of *Moniliophthora perniciosa* might be located around position 1 where we have a minimum for the GC skew index (Fig 3).

Integrated plasmid

The analysis of the sequence showed the presence of two genes, *rpo* and *dpoB*, that code for a DNA-dependent RNA

polymerase and a DNA-directed DNA polymerase, respectively. These two genes are not typical mitochondrial genes. They are found in plasmids, and these plasmids can be free in the linear form or integrated into the mitochondrial genome (Griffiths 1995, 1998).

Four sets of inverted repeats were found in the region surrounding these two genes, two small, 130 bp long, and two large, 347 bp long. The organization of this region is described above and shown in Fig 4. Invertron-type plasmids are characterized by large terminal inverted repeats of around 1000 bp (Griffiths 1998), as also found in transposable elements, bacteriophages, and adenoviruses (Sakaguchi 1990). We are proposing the existence of an invertron-type plasmid integrated in the *Moniliophthora* perniciosa mitochondrial genome in the region delimited by the large inverted repeats (347 bp).

The plasmid region presents three conserved, hypothetical ORFs; two of them (hypP1 and hypP2) are located before rpo and have the same orientation as this polymerase. The other ORF, hypP3, is located after dpoB and it has the same orientation as this gene, but opposite to the first three genes (Fig 4A). These two sets of conserved, hypothetical ORFs are located between the large and the small inverted repeats. Four additional, non-conserved, hypothetical ORFs (hyp58, hyp57, hyp28, and hyp29) are located after the last large inverted repeat. These ORFs are inside the region of low GC content and with a significant deviation of the GC skew (Fig 3, between dotted lines 2 and 4). The first two (hyp58 and hyp57) have the same orientation as dpoB, and the last two (hyp28 and hyp29) have the opposite orientation. A comparison of the protein sequence of these hypothetical ORFs showed that *hypP1*, *hypP3*, and *hyp58* are significantly similar (BlastX e-values: *hypP1* and *hyp58* – 1e-19 and *hypP1* and *hypP3* – 2e-8); as it happens with *hyp57* and *hyp28* (BlastX e-values: *hyp57* and *hyp28* – 7e-15). The orientation and similarity among these sequences seems to indicate the integration and duplication of the genes present at this location.

In order to check the relationship between the plasmid and the mitochondrial genome, we followed two different strategies. The first one consisted in a careful analysis of the assembly of the sequencing data, and the second one involved an analysis of the sequence by PCR. The analysis of the assembly of the sequencing data did not show any overrepresentation of reads along the whole mtDNA assembly, and more specifically at the plasmid region. In the shotgun strategy used here an overrepresentation of reads is an indication of the presence of a different number of copies of some part of the DNA. This should indicate the presence of independent entities, like plasmids. This is the case of the mitochondrial genome that is overrepresented with respect to the genomic DNA, 31 *versus* 2.5 times, respectively. Another indication of the existence of a separate plasmid would be the abrupt interruption of the reads at the ends of free plasmids, but this situation was not detected. This *in silico* analysis is a strong indication that the plasmid did not exist in an independent form. Therefore, it may be stably integrated in the genome.

It is important to note that genomic libraries were constructed from mycelia at different ages, with young and old mycelia. It has been reported that some linear plasmids play a role in fungal aging, normally called senescence, by jumping over the mtDNA presenting different insertion points in the sequence, disrupting this way essential genes (D'Souza *et al.* 2005; Fox & Kennell 2001). If the aging of *M. perniciosa* mycelia is caused by a similar process, we should be able to detected different combinations of mitochondrial and plasmid DNAs

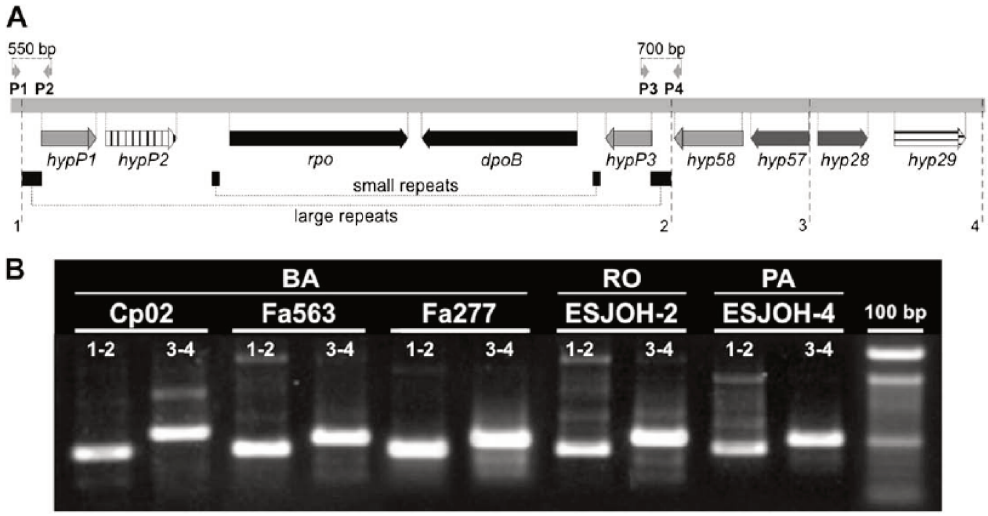


Fig 4 – (A) Genetic map of the plasmid and adjacent regions. Black boxes represent the two polymerases. Conserved hypothetical ORFs, *hypP1* and *hypP3*, together with non-conserved hypothetical ORFs *hyp58* that show significant similarity (BlastX e-values: *hypP1* and *hyp58* – 1e-19 and *hypP1* and *hypP3* – 2e-8) are represented by light grey boxes. Non-conserved hypothetical ORFs, *hyp57* and *hyp28*, also show significant similarity (BlastX e-values: *hyp57* and *hyp28* – 7e-15) are represented by dark grey boxes. Inverted repeats are represented by the small black boxes connected by a line according to its size. The position of the primers P1, P2, P3, and P4 are indicated by the boxes on top, as well as the expected size of the amplification product. (B) Analysis of the boundaries of the plasmid region by PCR using primers P1–P2 (1–2) and P3–P4 (3–4). Origin of the samples: BA, Bahia; RO, Roraima; PA, Para. Dashed lines indicate the same regions as in Fig 3.

by misalignments in the assembly. However, such inconsistencies, which would be detected by 'high-quality discrepancies' (Ewing & Green 1998; Ewing *et al.* 1998), have not been detected and the assembly consistently produced the same consensus sequence. Therefore, we concluded that the plasmid was stably integrated in a defined position in the mtDNA of this isolate.

An experimental approach was also devised in order to further check the integration of the plasmid DNA in other isolates. Four different primers were designed, two of them just outside the plasmid sequence representing sequences belonging to the mtDNA, and the other two inside the boundaries of the region considered as being the plasmid representing sequences at regions outside the larger external inverted repeats. The region to be amplified by these primers encompasses the external repeats that define the boundaries of the plasmid. A representation of the position of the primers and the expected size of the amplified fragment is shown in Fig 4A. In the case of the plasmid being integrated, fragments corresponding to the integration boundaries will be amplified. The results of the PCR analysis performed with five isolates from different regions of Brazil are shown in Fig 4B. The PCR analysis was repeated with further 23 isolates of *M. perniciosa*. The nomination and origin of all isolates used in this study are summarized in Table 2. In all isolates we detected the amplification of the specific region for the combination of primers P1–P2 and P3–P4 indicating the integration of the plasmid into the mitochondrial genome. No positive result with any of the isolates was obtained with the combination of primers P1–P4, which should produce a 1002 bp fragment if the region between the external inverted repeats were not present. Thus, only in the absence of the integrated plasmid will the last combination of primers (P1–P4) render a positive result. The distance between the locations of these primers with the plasmid integrated (11 972 bp) makes it technically impossible to obtain an amplification product in the conditions used here.

Table 2 – Isolates of *Moniliophthora perniciosa* with indication of the municipality of collection and institution responsible for this collection (repository) used in the PCR experiments

Isolate	Local of collection	Repository
CP02	Itabuna, BA, Brazil	UESC[a]
CP09	Ilhéus, BA, Brazil	CEPLAC[b]
Belmonte	Belmonte, BA, Brazil	UFB[c]
Ilhéus	Ilhéus, BA, Brazil	UFB[c]
Santo Amaro	Santo Amaro, BA, Brazil	UFB
FA42, FA276, FA277, FA278, FA562, FA563	Itabuna BA, Brazil	FAC[d]
FA281	Aiquara, BA, Brazil	FAC
FA287	Inema, BA, Brazil	FAC[d]
FA293	Gandu, BA, Brazil	FAC
FA300	Ibirataia, BA, Brazil	FAC
FA311	Itagiba, BA, Brazil	FAC
FA317	Ilhéus, BA, Brazil	FAC
BP10	Itapebi, BA, Brazil	FAC
FA551	Tabatinga, AM, Brazil	FAC
ESJOH-1	Marituba, PA, Brazil	ESALQ[e]
ESJOH-2	Ouro Preto, RO, Brazil	ESALQ
ESJOH-3	Belém, PA, Brazil	ESALQ
ESJOH-4	Altamira, PA, Brazil	ESALQ
ESJOH-5	Medicilandia, PA, Brazil	ESALQ
ESJOH-6	Ariquemes, RO, Brazil	ESALQ
ESJOH-7	Manaus, AM, Brazil	ESALQ
ESJOH-8	Ji-Paraná, RO, Brazil	ESALQ
ESJOH-9	Alta Floresta, MT, Brazil	ESALQ

a UESC (Universidade Estadual de Santa Cruz) Ilhéus, Bahia, collected by Júlio Cascardo.
b CEPLAC (Comissão Executiva do Plano de Lavoura do Cacau) Ilhéus, Bahia, collected by Karina Gramacho.
c UFB (Universidade Federal de Brasilia) Brasilia, collected by Maricília Arruda.
d FAC (Fazenda Almirante Cacau) Ilhéus, Bahia, collected by Alan Pomella.
e ESALQ - CENA (Centro de Energia Nuclear na Agricultura, in Escola Superior de Agricultura "Luiz de Queiroz") Piracicaba, São Paulo, collected by Paulo Albuquerque (ERJOH – CEPLAC).

Comparative genomics

The inspection of the mitochondrial genomes of *Moniliophthora perniciosa* together with those from other *Basidiomycota* species did not show significant overall synteny of the protein coding genes (Fig 5). However, a more detailed inspection showed that *nad* genes are present as groups (nad2–nad3, nad4L–nad5, and nad4–nad6); there is a tendency to stay together on the genome in the three members of the class *Agaricomycetes* (*Schizophyllum commune*, *Pleurotus ostreatus*, and *M. perniciosa*).

We compared the mitochondrial genomes of *M. perniciosa* with that of *Podospora anserina* because they are two of the biggest, over 100 kb, described so far (Fig 5). Interestingly, the main reason for the large size of the *P. anserina* mitochondrial genome is the presence of numerous introns (Cummings *et al.* 1990), whereas the mitochondrial genome from *M. perniciosa* posses many non-conserved hypothetical ORFs.

Phylogeny

Phylogenetic trees were built using the three cytochrome oxidase genes (*cox1*, *cox2*, and *cox3*) from 50 fungal species independently. The trees obtained for the individual genes were essentially the same, without significant differences. Due to this a new phylogenetic tree was built using the three *cox* genes concatenated. This tree showed the same topology as those built with independent genes, but presented higher BS support for most of the branches, with the exception of some basal branches. Fig 6 shows the rooted phylogenetic tree for this analysis of the three *cox* genes concatenated. The results largely agree with fungal taxonomy and with recent fungal phylogenies (James *et al.* 2006; Kouvelis *et al.* 2004).

The unrooted phylogenetic tree built with the polymerase genes from the plasmid using all complete fungal invertron-like plasmids, including kalilo-like plasmids, is shown in Fig 7. There is a clear separation of the plasmid sequences into two major groups. One of them was made up by the kalilo-like plasmids (Fig 7, right), as described by Sakaguchi (1990) and the other by the non-kalilo-like linear plasmids (Fig 7, left). This separation has no correlation with the phylogenetic separation between *Basidiomycota* and *Ascomycota* obtained with the genes encoding proteins of the respiratory

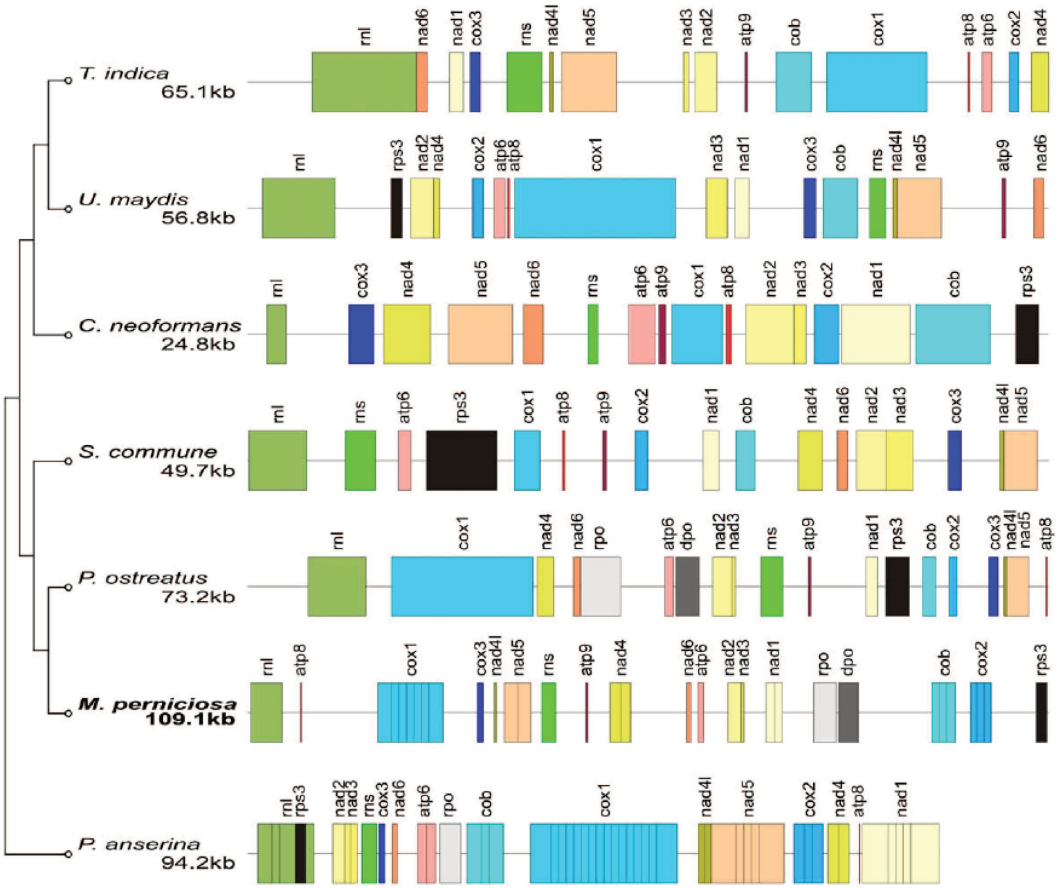


Fig 5 – Comparison of the gene content and order of the known mitochondrial genomes from *Basidiomycota* and *Podospora anserina*. Genetic maps of the mitochondrial genomes from *Tilletia indica* (NC_009880), *Ustilago maydis* (NC_008368), *Cryptococcus neoformans* (NC_004336), *Schizophyllum commune* (NC_003049), *Pleurotus ostreatus* (NC_009905) and *Moniliophthora perniciosa* (NC_005927). Gene sizes and total mitochondrial genome length are drawn to mach the genome of *M. perniciosa*. The mitochondrial genome of *Podospora anserina* (NC_001329) is represented at the bottom drawn proportionally to the genome of *M. perniciosa*, for size comparison.

chain (Fig 6). In this analysis, *M. perniciosa* genes grouped together with the kalilo-like plasmids from other fungi. There are other types of plasmids with a circular topology that preset only one of the polymerases (Griffiths 1995). These plasmids were not used in the phylogenetic analysis. Although they present sequence similarity, they represent non-homologous genes from different types of plasmids with different evolutionary origin (Griffiths 1995).

Morphology of Moniliophthora perniciosa *in vitro*

The germination of *Moniliophthora perniciosa* spores in rich media plates results in the growth of mycelia that presents a smooth, white aspect (Fig 8, left) with a continuous growth rate. After several months and plate passes, around ten, the mycelia begin to change its aspect becoming more rugged with a wave-like growth appearance (Fig 8, right) and the synthesis of a brown pigmentation. We consider the mycelia presented in the left panel of Fig 8 as young, and that in the right panel as old. This change appears to be irreversible, no regression to the young phenotype has been observed.

The yield of the DNA isolation depends greatly with the age of the mycelia used. Old mycelia render low amounts of material and present a relatively high degree of degradation. In contrast, young mycelia did not seem to show this problem. Another significant fact was that young mycelia had a higher mtDNA content with respect to genomic DNA than old mycelia, as estimated by the number of reads belonging to each class. For this reason we obtained many libraries encompassing different ages of the fungal culture. The consensus sequence obtained from the assembly of the reads belonging to young and old mycelia libraries did not show any difference.

Discussion

The mitochondrial genome of *Moniliophthora perniciosa* (109 103 bp) is the largest fungal mitochondrial genome

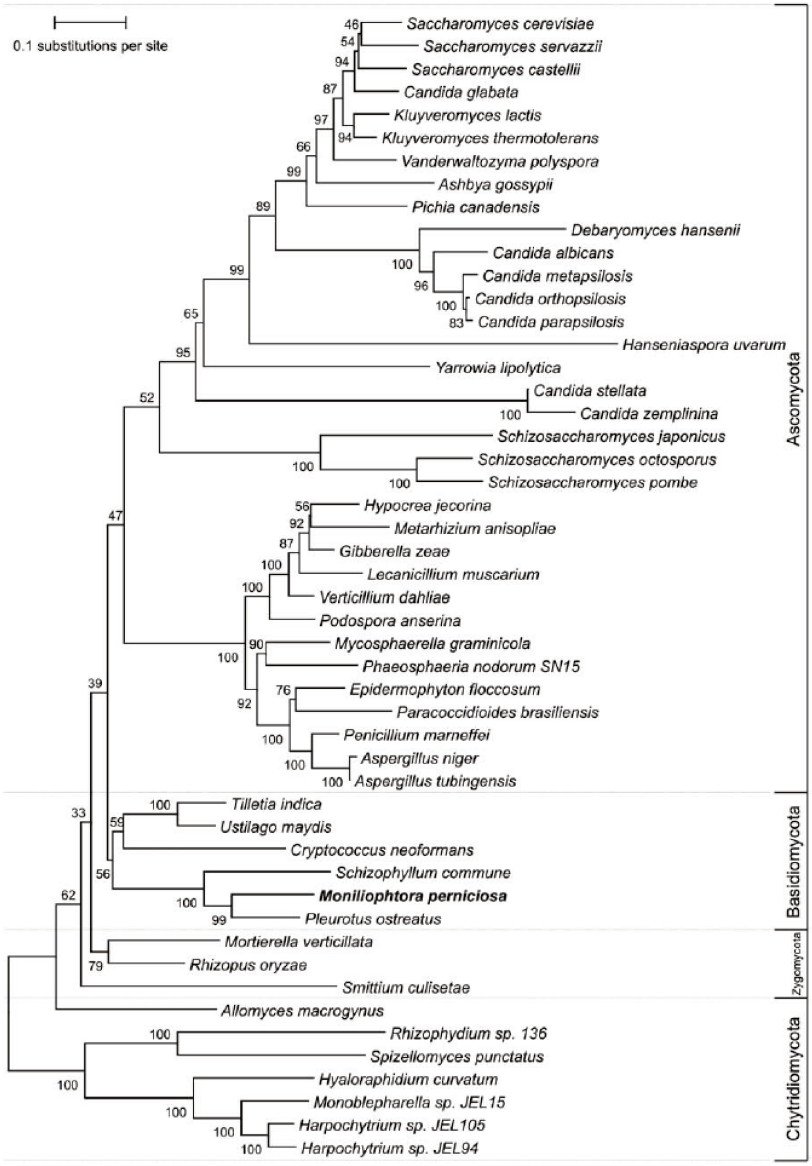


Fig 6 – Phylogenic tree built from the aligned portions of concatenated protein sequences of cox1, cox2, and cox3. GenBank sequences used from: *Allomyces macrogynus* (NC_001715), *Ashbya gossypii* (NC_005789), *Aspergillus niger* (NC_007445), *A. oryzae* (NC_008282), *A. tubingensis* (NC_007597), *Candida albicans* (NC_002653), *C. glabrata* (NC_004691), *C. metapsilosis* (NC_006971), *C. orthopsilosis* (NC_006972), *C. parapsilosis* (NC_005253), *C. zemplinina* (NC_005972), *C. neoformans* (NC_004336), *Debaryomyces hansenii* (NC_010166), *Epidermophyton floccosum* (NC_007394) *Gibberella zeae* (NC_009493), *Hanseniaspora uvarum* (NC_007780), *Harpochytrium* sp. JEL105 (NC_004623), *Harpochytrium* sp. JEL94 (NC_004760), *Hyaloraphidium curvatum* (NC_003048), *Hypocrea jecorina* (NC_003388), *Kluyveromyces lactis* (NC_006077), *K. thermotoloerans* (NC_006626), *Lecanicillium muscarium* (NC_004514), *Metarhizium anisopliae* (NC_008068), *Moniliophthora perniciosa* (NC_005927), *Monoblepharella* sp. JEL15 (NC_004624), *Mortierella verticillata* (NC_006838), *Mycosphaerella graminicola* (NC_010222), *Paracoccidioides brasiliensis* (NC_007935), *Penicillium marneffei* (NC_005256), *Phaeosphaeria nodorum* SN15 (NC_009746), *Pichia canadensis* (NC_001762), *Pleurotus ostreatus* (NC_009905), *Podospora anserina* (NC_001329), *Rhizophydium* sp. 136 (NC_003053), *Rhizopus oryzae* (NC_006836), *Saccharomyces castellii* (NC_003920), *S. cerevisiae* (NC_001224), *S. servazzii* (NC_004918), *Schizophyllum commune* (NC_003049), *Schizosaccharomyces japonicus* (NC_004332), *S. octosporus* (NC_004312), *S. pombe* (NC_001326), *Smittium culisetae* (NC_006837), *Spizellomyces punctatus* (NC_003052), *Tilletia indica* (NC_009880), *Ustilago maydis* (NC_008368), *Vanderwaltozyma polyspora* (NC_009638), *Verticillium dahliae* (NC_003060) and *Yarrowia lipolytica* (NC_002659).

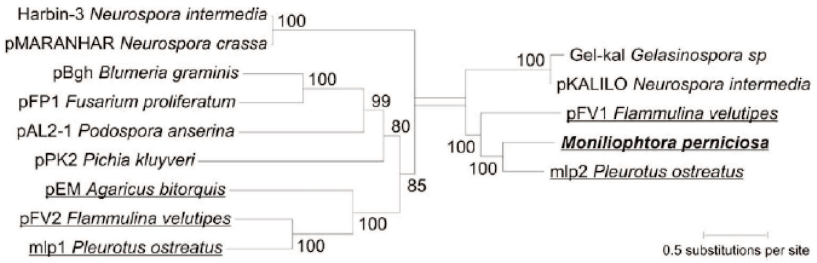


Fig 7 – Phylogenetic tree of the genes encoding polymerases from invertron–like mitochondrial plasmids. GenBank sequences used from: pEM in *Agaricus bitorquis* (AF096908 and X63075), pBgh in *Blumeria graminis* (NC_004935), pFV1 in *Flammulina velutipes* (AB028633), pFV2 in *F. velutipes* (AB028634), pFP1 in *Fusarium proliferatum* (NC_010425), Gel-kal in *Gelasinospora* sp. G114 (L40494), *Moniliophthora perniciosa* (NC_005927), pMARANHAR in *Neurospora crassa* (X55361), Harbin-3 in *N. intermedia* (NC_000843), pKALILO in *N. intermedia* (X52106), pPK2 in *Pichia kluyveri* (Y11606), mlp1 in *Pleurotus ostreatus* (NC_002135), mlp2 in *P. ostreatus* (AF355103) and pAL2-1 in *P. anserina* (X74137). Double line branch indicates a major separation of sequences in two groups: kalilo-like (right) and non-kalilo-like (left). *Basidiomycota* are underlined.

sequenced so far (http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/fu.html), the sixth among *Basidiomycota*, and the third of the class *Agaricomycetes*. To the best of our knowledge, the two members of the genus *Moniliophthora*, *M. perniciosa* and *M. roreri*, are the only ones from this taxonomic class with hemibiotrophic life-styles that also infect the aerial parts of plants. Therefore, the analysis of the sequence adds information to the knowledge of the fungal mitochondrial genomes, more specifically to the *Basidiomycota*, with potential to unveil specific information of this kind of organism.

This genome presents a circular topology, as have most of the fungal mitochondrial genomes sequenced so far. Only five, out of 50, have been described as linear. Also, it is A + T biased, like most other fungi. The tendency of A or T predominance becomes evident by the A/T frequency at the third codon position of the protein coding genes, 88.7 %, which is nearly a neutral position that could normally be occupied by any of the four bases without changing the amino acid identity. An explanation for this bias in the mitochondrial genome proposed by Xia (1996) suggest that, in order to maximize transcriptional efficiency, codon usage will be adapted to the most abundant nucleotide. Due to the high abundance of ATP over the other triphosphate nucleotides, codons ending in A are used more frequently.

All the 14 genes coding for the inner mitochondrial membrane proteins involved in the respiratory chain and the two genes for the rRNAs (*ml* and *ms*) are present in the *M. perniciosa* mitochondrial genome. Also, it contains an ancient gene coding for a ribosomal protein (*rps3*) present in the mitochondrial genome of some *Ascomycota*, *Basidiomycota*, *Zygomycota*, and a few *Chytridiomycota* (Bullerwell *et al.* 2000). These genes make up the standard set of genes found in other fungi.

The size of the fungal mitochondrial genome varies widely, the smallest reported is around 11 kbp from *Hanseniaspora uvarum* (Pramateftaki *et al.* 2006) and the largest is around 121 kbp from *Suillus grisellus* (Bruns & Palmer 1989); and most of them are smaller than 60 kbp (Bullerwell & Lang 2005). It is generally accepted that any variation in mitochondrial genome size is 'due mainly to the amount of noncoding material, because all mitochondrial DNA code for the same thing: some components of the electron transport chain, some structural RNAs of the mitochondrial ribosomes, and a range of mitochondrial tRNA' (Deacon 2006). The increased size of two of the biggest genomes appears to have different origin (Fig 5). Whereas *P. anserina* has larger sizes and numbers of intronic regions, *M. perniciosa* has bigger intergenic regions between the known conserved genes. In these intergenic regions we were able to find a set of conserved and non-conserved hypothetical ORFs (Fig 1). The existence of the plasmid region also contributes to the increased size of this genome. This region, between the large inverted repeats, represents around 10 % of the total genome.

The analysis of the codon usage by PCA showed that a group of 17 of those ORFs present a similar pattern to that of the known conserved genes (group B in Fig 2). Moreover, 14 of these ORFs are present a clockwise orientation, as most of the known conserved genes. Although little is known regarding transcription and RNA metabolism in non-model organisms, the presence of mitochondrial genes in the same orientation is an indication of the presence of one transcription unit, as transcription in mitochondria is universally accepted as polycistronic (Schäfer *et al.* 2005). Also, these ORFs and the known conserved genes are located in genomic regions not populated by tRNAs (Fig 1); so the protein coding regions are not superimposed by the tRNA coding regions, thus avoiding transcriptional problems. Conversely, the ORF coding for the tRNA without a defined anti-codon is located in a region populated by conserved genes (Fig 1, X at the innermost circle); this might indicate that the ORF is not real and does not code for anything. Transcription in mitochondria generates big transcripts that are processed with the tRNA regions indicating breaking points (tRNA punctuation model) in order to generate the final mRNAs (Ojala *et al.* 1981).

It is known that there are a large number of proteins in mitochondria, most of them encoded by nuclear genes (Dimmer & Rapaport 2008). Moreover, it is believed that most of these genes migrated from the mitochondrial genome

to the nucleus. Therefore, it is possible that in fungi, which present a large variability in the size and structure of their mitochondrial genomes, some functional genes still remain in the mitochondrial genome contributing to the mitochondrial proteome.

With this information, together with the data presented here, we propose the existence of 17 potential new genes in the mitochondrial genome of *M. perniciosa* coding for proteins without known functions in this fungus. If this assumption is true, it is probable that the sequencing of new fungal mitochondrial genomes will lead to the identification of homologues.

Phylogeny

Concatenated sequences have been used successfully in phylogenetic analysis based on mitochondrial genes (Woo *et al.* 2003; Kouvelis *et al.* 2004). The phylogenetic analysis of the known fungal mitochondrial genomes is in accordance with recent studies using mitochondrial genes (Woo *et al.* 2003; Kouvelis *et al.* 2004), even considering the limited number of available complete genomes of *Basidiomycota*. Also, it concurs with the same type of analysis performed using genomic genes (James *et al.* 2006).

The phylogenetic tree built with the polymerase genes present in the plasmid region show a completely different picture. The species segregates into two groups according to the type of plasmid, kalilo-like and non-kalilo-like (Fig 7, right and left, respectively). Both groups have members of the *Ascomycota* and *Basidiomycota* not following the phylogenetic analysis using the *cox* mitochondrial genes (Woo *et al.* 2003; Kouvelis *et al.* 2004) or genomic genes (James *et al.* 2006). The most significant finding is the segregation of the plasmids in two groups. This segregation suggests an ancient duplication of these plasmids, before the separation of *Ascomycota* and *Basidiomycota*, following a separate evolutionary path independently of the host.

Synteny

Fungal mitochondrial genomes do not show an overall high degree of synteny (Kouvelis *et al.* 2004), in contrast to the metazoan mitochondrial genomes that show the same gene content and order (Boore 1999). Significant synteny has been described for species belonging to the classes *Sordariomycetes* (Kouvelis *et al.* 2004) and *Eurotiomycetes* (Cardoso *et al.* 2007). These data indicate that, at least at the level of class, some fungal mitochondrial genomes from *Ascomycota* show significant synteny. In the case of *Basidiomycota*, the number of sequenced mitochondrial genomes is significantly lower, but out of six genomes three of them belong to the class *Agaricomycetes*. The analysis of the content and order of genes did not show an overall synteny in this case, in contrast to what happens in the case of the *Sordariomycetes* and *Eurotiomycetes*, which have clear synteny. Although we have very limited data, they seem to indicate that some groups of *Basidiomycota* are more diverse than some groups of *Ascomycota*.

Conversely, we have recently sequenced the complete mitochondrial genome of *Moniliophthora roreri*, a species very close to *M. perniciosa*. A preliminary analysis showed that, in this case, all known conserved genes present the same order in the two genomes (result not shown) indicating a high degree of synteny. Taken together, our data indicate that synteny is not a general feature of the *Basidiomycota* mitochondrial genome and that possibly gene order is mixed during speciation.

Mitochondrial plasmid

Many filamentous fungi have mitochondrial plasmids (Griffiths 1995). These plasmids do not appear to have any phenotypic effect on their hosts with a few exceptions (Bertrand 2000; Griffiths 1992). These include the mitochondrial plasmids pKALILO of *Neurospora intermedia* (Bertrand *et al.* 1985, 1986), pMARANHAR of *N. crassa* (Court *et al.* 1991), and pAL2-1 of *Podospora anserina* (Maas *et al.* 2004). These plasmids have a typical invertron structure (with long inverted repeats) and affect the life span of their hosts. *Neurospora* strains normally do not senesce, but those carrying one of the mitochondrial plasmids pKALILO or pMARANHAR die in two to three weeks. This is associated with mitochondrial instability and insertions of the plasmid into the mitochondrial genome (Bertrand *et al.* 1985, 1986; Myers *et al.* 1989; Rieck *et al.* 1982), and for that these mitochondrial plasmids have been named senescence plasmids (Griffiths 1992). In the naturally senescent species *P. anserina*, insertion of pAL2-1 in strain AL2 is correlated with an increased life-span, and therefore, this mitochondrial plasmid has been named life-span-prolonging plasmid (Hermanns *et al.* 1994). However, this life-span extension requires more than the presence of pAL2-1, as the majority of natural isolates containing pAL2-1 homologous plasmids do not show this effect (van der Gaag *et al.* 1998). This process seems to be a combination of the insertion of the plasmid and a complex set of mitochondrial DNA rearrangements, including two insertions of the plasmid and a 3.6 kb deletion of the region between the *cytb* and *atp6* genes (Osiewacz *et al.* 1989; Hermanns & Osiewacz 1992). This process may have stabilized the plasmid sequence as an integral part of a mitochondrial genome, which thereby may have acquired new functions.

We detected possible sequence rearrangements in the neighbourhood of the plasmid region of the *Moniliophthora perniciosa* mitochondrial genome. Typically, the length of inverted repeats found in linear plasmids is around 1000 bp (Griffiths 1998). The plasmid region in *M. perniciosa* shows four sets of inverted repeats, two smaller (130 bp) and two larger (347 bp), and three conserved hypothetical ORFs located between the smaller and the larger repeats (Fig 4A). It seems that these ORFs were inserted inside the inverted repeats disrupting them, creating the new four smaller inverted repeats. Also, four non-conserved, hypothetical ORFs are located just after the last of the larger inverted repeats in the same region of the genome that is defined by its lower GC content and a change in direction of the GC skew (Fig 3). It is interesting to note that two of the conserved hypothetical ORFs, *hypP1* and *hypP3*, present a significant degree of similarity and they are located at the inner part of the extremities of the region delimited by the larger inverted repeats, their coding region starts just inside the inverted repeats, and they present opposite orientation. Also, the non-conserved hypothetical ORFs

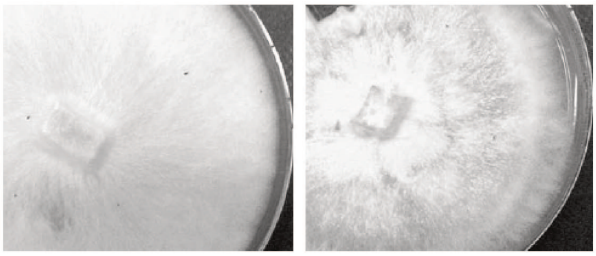


Fig 8 – Macroscopic morphology of young (left) and old (right) cultures of *Moniliophthora perniciosa* grown in rich media.

hyp58 shows significant similarity with *hypP1* and *hypP3*. There are two more non-conserved hypothetical ORFs, *hyp57* and *hyp28*, with opposite orientation and significant similarity. The existence of ORFs between the inverted repeats is an indication of an integration–recombination event that took place in this region. We believe that this event disrupted the typical inverted repeats present in linear plasmids, and as a consequence, the plasmid became stably integrated and remained part of the genome. Also, the non-conserved, hypothetical ORFs with significant similarity and opposite orientation indicate the occurrence of such integration–recombination events at this region of the genome.

Some symptoms that characterize fungal ageing are a decrease of growth rate and fertility, a reduction in the formation of aerial hyphae, and an increase in pigmentation. Except for the fertility, which has not been studied, all of the other symptoms have been observed in *M. perniciosa* growing in the laboratory (Fig 8). We are currently studying this process in order to see whether it can be characterized as senescence. Ageing in *Neurospora* is linked to the process of integration–excision of the linear plasmids (Osiewacz *et al.* 1989; Hermanns & Osiewacz 1992). Our sequencing data showed that the plasmid in *M. perniciosa* is stably integrated, thus, the origin of the ageing process observed in this fungal species seems to be different from that provoked by the linear plasmids in *Neurospora*.

The plasmid exists in all the isolates of *M. perniciosa* analysed indicating that it is an integral part of its mitochondrial genome. The GC content and GC skew indicate that the integration of the plasmid is a recent event and that the sequence is not adapted to the rest of the mitochondrial genome, thus maintaining some of its original features. It would be interesting to know the behaviour of the species before the integration of this element in its mitochondrial genome.

In our survey of many isolates of *M. perniciosa* from diverse geographical origins, we consistently found the plasmid integrated in the mitochondrial genome and concluded that this is a feature of the species, mitochondrial genome. Remarkably, a preliminary analysis of the mitochondrial genome of the close species *M. roreri* showed a high degree of similarity between both genomes (a similarity of 57 % for the comparison of *M. roreri* with *M. perniciosa*) but, curiously, the plasmid region was not present (not shown).

Although speculative, we believe that the insertion could confer some evolutionary advantage to the species. If this was not the case, the insertion would have suffered a negative selection or, as happens in *P. anserina*, we would have two types of isolates: one without the integrated plasmid and one with it in the mitochondrial genome.

In summary, we described in this work the largest basidiomycete mitochondrial genome sequenced so far. This genome has some remarkable features, such as the presence of putative new mitochondrial genes and the stable integration of a complete linear plasmid sequence. We are currently investigating how these elements can contribute to the physiology of this fungus.

Acknowledgements

This research was supported by the Brazilian agencies CNPq (research fellowship to N. C.), Capes, CNPq Regional Genoma Program, SEAGRIm and FAPESP (No. 02/09280-1).

REFERENCES

Adams KL, Palmer JD, 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution* **29**: 380–395.

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.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

Arakawa K, Saito R, Tomita M, 2007. Noise-reduction filtering for accurate detection of replication termini in bacterial genomes. *FEBS Letters* **581**: 253–258.

Benson G, 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**: 573–580.

Bertrand H, 2000. Role of mitochondrial DNA in the senescence and hypovirulence of fungi and potential for plant disease control. *Annual Review of Phytopathology* **38**: 397–422.

Bertrand H, Chan BS, Griffiths AJ, 1985. Insertion of a foreign nucleotide sequence into mitochondrial DNA causes senescence in *Neurospora intermedia*. *Cell* **41**: 877–884.

Bertrand H, Griffiths AJ, Court DA, Cheng CK, 1986. An extrachromosomal plasmid is the etiological precursor of kalDNA insertion sequences in the mitochondrial chromosome of senescent *Neurospora*. *Cell* **47**: 829–837.

Boore JL, 1999. Animal mitochondrial genomes. *Nucleic Acids Research* **27**: 1767–1780.

Bruns TD, Palmer JD, 1989. Evolution of mushroom mitochondrial DNA: *Suillus* and related genera. *Journal of Molecular Evolution* **28**: 349–362.

Bullerwell CE, Burger G, Lang BF, 2000. A novel motif for identifying *rps3* homologs in fungal mitochondrial genomes. *Trends in Biochemical Sciences* **25**: 363–365.

Bullerwell CE, Forget L, Lang BF, 2003. Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Research* **31**: 1614–1623.

Bullerwell CE, Lang BF, 2005. Fungal evolution: the case of the vanishing mitochondrion. *Current Opinion in Microbiology* **8**: 362–369.

Burger G, Gray MW, Lang BF, 2003. Mitochondrial genomes: anything goes. *Trends in Genetics* **19**: 709–716.

Burger G, Lang BF, 2003. Parallels in genome evolution in mitochondria and bacterial symbionts. *IUBMB Life* **55**: 205–212.

Campbell A, Mrazek J, Karlin S, 1999. Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA* **96**: 9184–9189.

Cardoso MAG, Tambor JHM, Nobrega FG, 2007. The mitochondrial genome from the thermal dimorphic fungus *Paracoccidioides brasiliensis*. *Yeast* **24**: 607–616.

Clark-Walker GD, 1992. Evolution of mitochondrial genomes in fungi. In: Welstenholme DR, Jeon KW (eds), *Mitochondrial Genomes*. Academic Press, San Diego, pp. 89–127.

Court DA, Griffiths AJF, Kraus SR, Russel PJ, Bertrand H, 1991. A new senescence inducing mitochondrial linear plasmid in field-isolated *Neurospora crassa* strains from India. *Current Genetics* **19**: 129–137.

Cummings DJ, McNally KL, Domenico JM, Matsuura ET, 1990. The complete DNA sequence of the mitochondrial genome of *Podospora anserina*. *Current Genetics* **17**: 375–402.

Deacon JW, 2006. *Fungal Biology*, 4th edn. Blackwell Publishing, Oxford.

Dimmer KS, Rapaport D, 2008. Proteomic view of mitochondrial function. *Genome Biology*. doi:10.1186/gb-2008-9-2-209.

D'Souza AD, Sultana S, Maheshwari R, 2005. Characterization and prevalence of a circular mitochondrial plasmid in senescence-prone isolates of *Neurospora intermedia*. *Current Genetics* **47**: 182–193.

Evans HC, Holmes KA, Phillips W, Wilkinson MJ, 2002. What's in a name? *Crinipellis*, the final resting place for the frosty pod rot pathogen of cocoa? *Mycologist* **16**: 148–152.

Ewing B, Green P, 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* **8**: 186–194.

Ewing B, Hillier L, Wendl MC, Green P, 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8**: 175–185.

Felsenstein J, 1985. Confidence-limits on phylogenies — an approach using the bootstrap. *Evolution* **39**: 783–791.

Forget L, Ustinova J, Wang Z, Huss VAR, Lang BF, 2002. *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. *Molecular Biology and Evolution* **19**: 310–319.

Fox AN, Kennell JC, 2001. Association between variant plasmid formation and senescence in retroplasmid-containing strains of *Neurospora* spp. *Current Genetics* **39**: 92–100.

Garcia O, Joci N, Macedo J, Tibúrcio R, Zaparoli G, Rincones J, Bittencourt L, Ceita G, Micheli F, Gesteira A, Mariano A, Schiavinato M, Medrano FJ, Meinhardt L, Pereira GAG, Cascardo J, 2007. Characterization of necrosis and ethylene-inducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom disease in *Theobroma cacao*. *Mycological Research* **111**: 443–455.

Gordon D, Abajian C, Green P, 1998. Consed: a graphical tool for sequence finishing. *Genome Research* **8**: 195–202.

Gray MW, Burger G, Lang BF, 1999. Mitochondrial evolution. *Science* **283**: 1476–1481.

Griffiths AJF, 1992. Fungal senescence. *Annual Review of Genetics* **26**: 351–372.

Griffiths AJF, 1995. Natural plasmids of filamentous fungi. *Microbiological Reviews* **59**: 673–685.

Griffiths AJF, 1998. The kalilo family of fungal plasmids. *Botanical Bulletin of Academia Sinica* **39**: 147–152.

Grigoriev A, 1998. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Research* **26**: 2286–2290.

Guindon S, Gascuel O, 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696–704.

Haenfler J, Teepe H, Weigl C, Kruft V, Lurz R, Woestemeyer J, 1992. Circular extra-chromosomal DNA codes for a surface protein in the (positive) mating type of the zygomycete *Absidia glauca*. *Current Genetetics* **22**: 319–325.

Heath PJ, Stephens KM, Monnat Jr RJ, Stoddard BL, 1997. The structure of I-CreI, a group I intron-encoded homing endonuclease. *Nature Structural Biology* **4**: 468–476.

Hermanns J, Asseburg A, Osiewacz HD, 1994. Evidence for a life span-prolonging effect of a linear plasmid in a longevity mutant of *Podospora anserina*. *Molecular and General Genetics* **243**: 297–307.

Hermanns J, Osiewacz HD, 1992. The linear mitochondrial plasmid pAL2-1 of a long-lived *Podospora anserina* mutant is an invertron encoding a DNA and RNA polymerase. *Current Genetics* **22**: 491–500.

Higgins DG, 1994. CLUSTAL W: multiple alignment of DNA and protein sequences. *Methods in Molecular Biology (Clifton, N.J.)* **25**: 307–318.

Huang X, Madan A, 1999. CAP3: A DNA sequence assembly program. *Genome Research* **9**: 868–877.

Hur M, Geese WJ, Warig RB, 1997. Self-splicing activity of the mitochondrial group-I introns from *Aspergillus nidulans* and related introns from other species. *Current Genetics* **32**: 399–407.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schussler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lucking R, Budel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R, 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443**: 818–822.

Joliffe IT, 1986. *Principal Component Analysis*. Springer-Verlag, New York.

Jones DT, Taylor WR, Thornton JM, 1992. The rapid generation of mutation data matrices from protein sequences. *CABIOS* **8**: 275–282.

Jukes TH, 1984. The wobble rule. *Journal of Applied Biochemistry* **6**: 115–116.

Kempken F, 1995a. Plasmid DNA in mycelial fungi. In: Kuck U (ed), *The Mycota. Part II. Genetics and Biotechnology*, vol. II. Springer-Verlag, Heidelberg, pp. 169–187.

Kempken F, 1995b. Horizontal transfer of a mitochondrial plasmid. *Molecular and General Genetics* **248**: 89–94.

Kim EK, Jeong JH, Youn HS, Koo YB, Roe JH, 2000. The terminal protein of a linear mitochondrial plasmid is encoded in the N-terminus of the DNA polymerase gene in white-rot fungus *Pleurotus ostreatus*. *Current Genetics* **38**: 283–290.

Kouvelis VN, Ghikas DV, Typas MA, 2004. The analysis of the complete mitochondrial genome of *Lecanicillium muscarium* (synonym *Verticillium lecanii*) suggests a minimum common gene organization in mtDNAs of Sordariomycetes: phylogenetic implications. *Fungal Genetics and Biology* **41**: 930–940.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R, 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29**: 4633–4642.

Lang BF, Gray MW, Burger G, 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annual Review of Genetics* **33**: 351–397.

Lowe TM, Eddy SR, 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**: 955–964.

Maas MFPM, de Boer HJ, Debets AJM, Hoekstra RF, 2004. The mitochondrial plasmid pAL2-1 reduces calorie restriction mediated life span extension in the filamentous fungus *Podospora anserina*. *Fungal Genetics and Biology* **41**: 865–871.

Maas MFPM, Sellem CH, Hoekstra RF, Debets AJ, Sainsard-Chanet A, 2007. Integration of a pAL2-1 homologous mitochondrial plasmid associated with life span extension in *Podospora anserina*. *Fungal Genetics and Biology* **44**: 659–671.

Meinhardt F, Kempken F, Kamper J, Esser K, 1990. Linear plasmids among eukaryotes: fundamentals and application. *Current Genetics* **17**: 89–95.

Meinhardt F, Schaffrath R, Larsen M, 1997. Microbial linear plasmids. *Applied Microbiology and Biotechnology* **47**: 329–336.

Meinhardt LW, Bellato CM, Rincones J, Azevedo RA, Cascardo JC, Pereira GAG, 2006. *In vitro* production of biotrophic-like cultures of *Crinipellis perniciosa*, the causal agent of witches' broom disease of *Theobroma cacao*. *Current Microbiology* **52**: 191–196.

Myers CJ, Griffiths AJF, Bertrand H, 1989. Linear kalilo DNA is a *Neurospora* mitochondrial plasmid that integrates into the mitochondrial DNA. *Molecular and General Genetics* **220**: 113–120.

Nakagawa CC, Jones EP, Miller DL, 1998. Mitochondrial DNA rearrangements associated with mF plasmid integration and plasmodial longevity in *Physarum polycephalum*. *Current Genetics* **33**: 178–187.

Nakai R, Sen K, Kurosawa S, Shibai H, 2000. Basidiomycetous fungus *Flammulina velutipes* harbors two linear mitochondrial plasmids encoding DNA and RNA polymerases. *FEMS Microbiology Letters* **190**: 99–102.

Nosek J, Tomaska L, 2003. Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Current Genetics* **44**: 73–84.

Ojala D, Montoya J, Attardi G, 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**: 470–474.

Osiewacz HD, Hermanns J, Marcou D, Triffi M, Esser K, 1989. Mitochondrial DNA rearrangements are correlated with a delayed amplification of the mobile intron (plDNA) in a long-lived mutant of *Podospora anserina*. *Mutation Research* **219**: 9–15.

Peden JF, 1999. Analysis of codon usage. PhD Thesis. University of Nottingham, Nottingham, pp. 226.

Pramateftaki PV, Kouvelis VN, Lanaridis P, Typas MA, 2006. The mitochondrial genome of the wine yeast *Hanseniaspora uvarum*: a unique genome organization among yeast/fungal counterparts. *FEMS Yeast Research* **6**: 77–90.

R Development Core Team, 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna.

Rieck A, Griffiths AJ, Bertrand H, 1982. Mitochondrial variants of *Neurospora intermedia* from nature. *Canadian Journal of Genetics and Cytology* **24**: 741–759.

Rincones J, Mazotti GD, Griffith GW, Pomela A, Figueira A, Leal Jr GA, Queiroz MV, Pereira JF, Azevedo RA, Pereira GAG, Meinhardt LW, 2006. Genetic variability and chromosome-length polymorphisms of the witches' broom pathogen *Crinipellis perniciosa* from various plant hosts in South America. *Mycological Research* **110**: 821–832.

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

Rincones J, Scarpari LM, Carazzolle MF, Mondego JMC, Formighieri EF, Barau JG, Costa GGL, Carraro DM, Vilas-Boas LA, De Oliveira BV, Sabha M, Dias R, Cascardo JM, Azevedo RA, Meinhardt LW, Pereira GAG. Differential gene expression between the biotrophic-like and saprotrophic mycelia of the Witches' broom pathogen *Moniliophthora perniciosa*. *Molecular Plant–Microbe Interactions*, in press.

Rio MCS, de Oliveira BV, de Tomazella DPT, da Silva JAF, Pereira GAG, 2008. Production of calcium oxalate crystals by the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao. *Current Microbiology* **56**: 363–370.

Robison MM, Horgen PA, 1999. Widespread distribution of low-copy-number variants of mitochondrial plasmid pEM in the genus *Agaricus*. *Fungal Genetics and Biology* **26**: 62–70.

Rycovska A, Valach M, Tomaska L, Bolotin-Fukuhara M, Nosek J, 2004. Linear versus circular mitochondrial genomes: intraspecies variability of mitochondrial genome architecture in *Candida parapsilosis*. *Microbiology* **150**: 1571–1580.

Sakaguchi K, 1990. Invertrons, a class of structurally and functionally related genetic elements that includes linear DNA plasmids, transposable elements, and genomes of adeno-type viruses. *Microbiological Reviews* **54**: 66–74.

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GAG, 2005. Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa*. *Journal of Experimental Botany* **56**: 865–877.

Schäfer B, Hansen M, Lang BF, 2005. Transcription and RNA-processing in fission yeast mitochondria. *RNA* **11**: 785–795.

Seif E, Leigh J, Liu Y, Roewer I, Forget L, Lang BF, 2005. Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms. *Nucleic Acids Research* **33**: 734–744.

Shepherd HS, 1992. Linear, non-mitochondrial plasmids of *Alternaria alternata*. *Current Genetics* **21**: 169–172.

Stewart JB, Beckenbach A, 2006. Insect mitochondrial genomics 2: the complete mitochondrial genome sequence of a giant stonefly, *Pteronarcys princes*, asymmetric directional mutation bias, and conserved plecopteran A + T-region elements. *Genome* **49**: 815–824.

Surzcki S, 2000. *Basic Methods in Molecular Biology*. Springer-Verlag, New York.

Thompson JD, Higgins DG, Gibson TJ, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673–4680.

Torriani SF, Goodwin SB, Kema GH, Pangilinan JL, McDonald BA, 2008. Intraspecific comparison and annotation of two complete mitochondrial genome sequences from the plant pathogenic fungus *Mycosphaerella graminicola*. *Fungal Genetics and Biology* **45**: 628–637.

Touchon M, Nicolay S, Audit B, Brodie of Brodie E-B, d'Aubenton-Carafa Y, Arneodo A, Thermes C, 2005. Replication-associated strand asymmetries in mammalian genomes: toward detection of replication origins. *Proceedings of the National Academy of Sciences, USA* **102**: 9837–9841.

van der Gaag M, Debets AJM, Osiewacz HD, Hoekstra RF, 1998. The dynamics of pAL2-1 homologous linear plasmids in *Podospora anserina*. *Molecular and General Genetics* **258**: 521–529.

van Roey P, Waddling CA, Fox KM, Belfort M, Derbyshire V, 2001. Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. *The EMBO Journal* **20**: 3631–3637.

Venter JC, Adams MD, Sutton GG, Kerlavage AR, Smith HO, Hunkapiller M, 1998. Shotgun sequencing of the human genome. *Science* **280**: 1540–1542.

Wang Y, Zeng F, Hon CC, Zhang Y, Leung FC, 2008. The mitochondrial genome of the basidiomycete fungus *Pleurotus ostreatus* (oyster mushroom). *FEMS Microbiology Letters* **280**: 34–41.

Woo PCY, Zhen HJ, Cai JJ, Yu J, Lau SKP, Wang J, Teng JLL, Wong SSY, Tse RH, Chen R, Yang HM, Liu B, Yuen KY, 2003. The mitochondrial genome of the thermal dimorphic fungus *Penicillium marneffei* is more closely related to those of molds than yeasts. *FEBS Letters* **555**: 469–477.

Xia X, 1996. Maximizing transcription efficienty causes codon usage bias. *Genetics* **144**: 1309–1320.

Zivanovic Y, Wincker P, Vacherie B, Bolotin-Fukuhara M, Fukuhara H, 2005. Complete nucleotide sequence of the mitochondrial DNA from *Kluyveromyces lactis*. *FEMS Yeast Research* **5**: 315–322.

Capítulo V – Genes acquired by horizontal transfer are potentially involved in the evolution of phytopathogenicity in *Moniliophthora perniciosa* and *Moniliophthora* roreri, two of the major pathogens of cacao.

Neste trabalho foram buscados genes obtidos por transferência gênica horizontal no genoma de *M. perniciosa* e de *M. roreri*. Os resultados foram publicados no artigo "Genes Acquired by Horizontal Transfer Are Potentially Involved in the Evolution of Phytopathogenicity in *Moniliophthora perniciosa* and *Moniliophthora roreri*, Two of the Major Pathogens of Cacao " (Tiburcio et al. 2010), apresentado a seguir na sua íntegra. As referências bibliográficas são apresentadas juntamente com o artigo e não foram repetidas na seção de Bibliografia apresentada no final desta tese.

J Mol Evol (2010) 70:85–97
DOI 10.1007/s00239-009-9311-9

Genes Acquired by Horizontal Transfer Are Potentially Involved in the Evolution of Phytopathogenicity in *Moniliophthora perniciosa* and *Moniliophthora roreri*, Two of the Major Pathogens of Cacao

Ricardo Augusto Tiburcio · Gustavo Gilson Lacerda Costa · Marcelo Falsarella Carazzolle · Jorge Maurício Costa Mondego · Stephen C. Schuster · John E. Carlson · Mark J. Guiltinan · Bryan A. Bailey · Piotr Mieczkowski · Lyndel W. Meinhardt · Gonçalo Amarante Guimarães Pereira

Received: 13 August 2009 / Accepted: 3 December 2009 / Published online: 23 December 2009

Abstract *Moniliophthora perniciosa* and *Moniliophthora roreri* are phytopathogenic basidiomycete species that infect cacao causing two important diseases in this crop: "Witches' Broom" and "Frosty Pod Rot", respectively. The ability of species from this genus (*Moniliophthora*) to cause disease is exceptional in the family Marasmiaceae. Species in closely related genera including, *Marasmius*, *Crinipellis,* and *Chaetocalathus,* are mainly saprotrophs and are not known to cause disease. In this study, the possibility that this phytopathogenic lifestyle has been acquired by horizontal gene transfer (HGT) was investigated. A stringent genome comparison pipeline was used to identify potential genes that have been obtained by *Moniliophthora* through HGT. This search led to the identification of three genes: a metallo-dependent hydrolase (MDH), a mannitol phosphate dehydrogenase (MPDH), and a family of necrosis-inducing proteins (NEPs). Phylogenetic analysis of these genes suggests that *Moniliophthora* acquired NEPs from oomycetes, MDH from actinobacteria and MPDH from firmicutes. Based on the known gene functions and on previous studies of *M. perniciosa* infection and development, a correlation between

R. A. Tiburcio · G. G. L. Costa · M. F. Carazzolle · G. A. G. Pereira (✉)
Laboratório de Genômica e Expressão, Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas, 13083-970 Campinas, SP, Brazil
e-mail: goncalo@unicamp.br
URL: http://www.lge.ibi.unicamp.br

R. A. Tiburcio
e-mail: tiburcio@lge.ibi.unicamp.br

M. F. Carazzolle
Centro Nacional de Processamento de Alto Desempenho em São Paulo, Universidade Estadual de Campinas, CP 6141, Campinas, SP 13083-970, Brazil

J. M. C. Mondego
Centro de Pesquisa e Desenvolvimento em Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, Campinas, SP 13083-970, Brazil

J. M. C. Mondego
Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP 13083-970, Brazil

S. C. Schuster
Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, PA 16802, USA

J. E. Carlson
School of Forest Resources, Pennsylvania State University, University Park, PA 16802, USA

M. J. Guiltinan
Department of Horticulture, Pennsylvania State University, University Park, PA 16802, USA

B. A. Bailey · L. W. Meinhardt
Sustainable Perennial Crops Laboratory, USDA-ARS, 10300 Baltimore Av., BARC-W, Beltsville, MD 20740, USA

P. Mieczkowski
Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Mary Ellen Jones, Room 921, 27599-3280 Chapel Hill, NC, USA

gene acquisition and the evolution of the phytopathogenic genus *Moniliophthora* can be postulated.

Keywords *Moniliophthora perniciosa* · *Moniliophthora roreri* · Witches' broom disease · Frosty pod rot · Cacao · Horizontal gene transfer

Introduction

Moniliophthora perniciosa (Stahel), Aime & Phillips-Mora and *Moniliophthora roreri* H. C. Evans, Stalpers, Samson & Benny are phytopathogenic fungi of the order Agaricales (Basidiomycete) that infect cacao (*Theobroma cacao* L.). *M. perniciosa* is a hemibiotrophic species that causes Witches' broom disease (WBD), which has severely impacted cacao production in Brazil after its introduction into the southern region of Bahia, the main cacao-producing region in the country (Pereira et al. 1996; Meinhardt et al. 2008). WBD of cacao affects all cacao-growing regions of South America (Meinhardt et al. 2008). *M. roreri* causes frosty pod rot (FPR) disease. This fungus is extremely harmful to cacao production in countries of northwestern South America, such as Colombia, Peru, Ecuador and Venezuela, as well as in Central America and the Caribbean (Phillips-Mora 2003).

The spores of *M. perniciosa* and *M. roreri* are able to infect cacao fruits, first inducing parthenocarpy and then necrosis during the late stages of infection. However, only the spores of *M. perniciosa* are able to infect meristems, first causing hyperplasy and hypertrophy (the so called green brooms). During this stage of the disease, the fungus exists as monokaryotic mycelia, developing inside the apoplast and displaying a biotrophic lifestyle. Eight to twelve weeks after the initial infection, this homothallic fungus becomes dikaryotic, with mycelia showing clamp connections and displaying a saprotrophic lifestyle. The dikaryotic hyphae spread through the broom causing necrosis and death of the infected tissues. These necrotic hypertrophic branches are known as "dry brooms". Following alternating wet and dry periods, the dikaryotic hyphae produce basidiomes that release spores, thus completing the *M. perniciosa* life cycle (Almeida et al. 1997).

On the other hand, *M. roreri* does not produce basidiomes. Its spores are produced inside conidia-like structures that cover cacao fruits forming a thick white layer. When released, the spores form dense clouds of white aerial spores that drift with the air currents. Apparently, *M. roreri* does not form clamp connections and presents only monokaryotic mycelia throughout its development (Griffith et al. 2003).

The phytopathogenic characteristics of *M. perniciosa* and *M. roreri* are exceptions amongst related basiomycete species, which belongs to the genera *Marasmius*, *Crinipellis*, and *Chaetocalathus* and are mainly saprotrophic litter fungi (Aime and Phillips-Mora 2005). The phylogenetic analyses showed that, despite their contrasting morphologies, *M. perniciosa* and *M. roreri* are sister species, which suggest that their phytopathogenicity has evolved only once, in a common ancestor (Aime and Phillips-Mora 2005). It is also proposed by these authors that the Section *Iopodinae* within the genus *Crinipellis*, which contains 11 of the 75 *Crinipellis* species, and include all the other phytopathogens of this genus, can be congeneric with the genus *Moniliophthora*. *Moniliophthora* contains only *M. perniciosa*, *M. roreri*, and an unnamed asymptomatic and presumably symbiotic endophyte of grasses, isolated from New Mexico (Aime and Phillips-Mora 2005).

Since 2001, when the WBD genome project was initiated (http://www.lge.ibi.unicamp.br/vassoura), several articles have analyzed possible pathogenicity and developmental mechanisms of *M. perniciosa*. For example, a biochemical characterization of the interaction between cacao and *M. perniciosa* revealed metabolic alterations in infected plants, including the increase in malondialdehyde and glycerol contents and the production of ethylene and oxalate, a phytopathogenic factor that could scavenge calcium ions that bind to pectin polymers, thus making plant cell walls more accessible to fungal pectinases (Scarpari et al. 2005; Rio et al. 2008). It was also found that *M. perniciosa* produces proteins that induce necrosis in cacao, such as necrosis and ethylene-producing proteins (NEPs) and cerato-platanins (Garcia et al. 2007; Zaparoli et al. 2009). Furthermore, Rincones et al. (2008) analyzed the differential gene expression between biotrophic and saprotrophic mycelia, indicating that the biotrophic fungus expresses a series of proteins related to pathogenicity (i.e., glyoxal oxidase, chitin synthase class V, proteinases, and lipases). Additionally, a genome survey of *M. perniciosa* shows that this fungus has a myriad of cytochrome P450 monooxygenases, efflux transporters, and anti-oxidative enzymes that could be important to detoxification during the progression of WBD (Mondego et al. 2008). This arsenal of plant cell wall degrading enzymes is similar to that found in the hemibiotrophic ascomycetes *Fusarium graminearum* and *Magnaporthe grisea* (Mondego et al. 2008). On the other hand, the majority of the studies performed with *M. roreri* are descriptive, lacking the biochemical basis needed to understand FPR. In view of these data, a comparative genomic initiative was launched in order to use the available information from *M. perniciosa* and WBD to expand the knowledge base of *M. roreri* and FPR (http://www.lge.ibi.unicamp.br/roreri). Large drafts of both genomes are in final assembly and are currently being used for functional analyses and evolutionary comparisons.

The genetic and biochemical characteristics of *M. perniciosa* found so far indicate that the evolution of their

phytopathogenicity involved both changes in the functions of individual enzymes (i.e., the use of cerato-platanins to induce necrosis) as well as changes at the genomic level (i.e., the expansion of the cytochrome P450 monooxygenases, carboxyesterases, lipases, and deuterolysins protein families). On the other hand, the rarity of phytopathogenicity in species from genera related to *Moniliophthora* suggests that the evolution of this trait in *M. roreri* and *M. perniciosa* may have involved more than the adaptation of ancestral genes and gene families. It is probable that new genes related to phytopathogenicity were obtained by these two species during evolution through horizontal gene transfer (HGT). Many HGT events have been reported in fungi, both influencing the evolution of the main lineages resulting in changes to one or a few related species. Pathogenicity is particularly important since it has been reported that HGT events have been associated with its evolution in several different fungi (Richards et al. 2006; Wenzl et al. 2005; Sanders 2006; Stukenbrock and McDonald 2008; Oliver and Solomon 2008, Khaldi et al. 2008).

This study reports the identification of three coding genes possibly acquired by HGT in the genomes of *M. perniciosa* and *M. roreri*. These genes are reported in literature as being related to fungal phytopathogenicity, leading us to suggest that HGT is involved in the evolution of this trait in the *Moniliophthora*.

Materials and Methods

HGT Candidates Search

Gene candidates for HGT in *M. perniciosa* were identified by screening the 14,501 coding genes predicted in a draft of the genome sequence using the gene predictor program Augustus (Stanke and Waack 2003), according to the procedures described by Mondego et al. 2008. The genome sequence has 43 million base pairs (Mbp) distributed in 10,216 contigs and was assembled with reads from Sanger sequencing (1.1× coverage), GS-FLX 454 (5× coverage), and Illumina/Solexa (44× coverage) technologies.

The 14,501 *M. perniciosa* predicted genes were submitted to a local BLASTp search (Altschul et al. 1990) against a protein database containing all basidiomycete sequences available in the GenBank non-redundant protein database (nr), except for those of *M. perniciosa* previously published, that were manually excluded. Predicted genes with e-values $\leq 1e^{-10}$ were removed from further analyses and the remaining were submitted to local tBLASTn searches against a database with DNA sequences from basidiomycetes available from the GenBank nucleotide database (nt). As in the previous BLASTp search, e-values $\leq 1e^{-10}$ excluded those genes from the list of potential HGT candidates. These two BLAST searches were used to find and remove *M. perniciosa* predicted genes with homology to basidiomycetes. The *M. perniciosa* predicted genes without homology to basidiomycetes were submitted to a local BLASTp search against the GenBank nr database. Those with the first hit against a fungi sequence (with an e-value $\leq 1e^{-10}$) and those without hits were also excluded as HGT candidates. These search parameters left only the predicted genes that had no homology to basidiomycetes and those with the highest homology to non-fungal species for consideration as HGT candidates, and these were submitted to phylogenetic analysis. The excluded basidiomycete protein and DNA sequences included the genomes of *Coprinus cinereus*, *Cryptococcus neoformans* (Loftus et al. 2005), *Laccaria bicolor* (Martin et al. 2008), *Ustilago maydis* (Kamper et al. 2006), and *Phanerochaete chrysosporium* (Martinez et al. 2004); all with more than 10× coverage. All the sequences used from GenBank were downloaded in October 2008.

Phylogenetic Analysis of HGT Candidates

Protein sequence alignments of HGT candidates and their homologous sequences available at GenBank nr database were made using ClustalX (Thompson et al. 1997), with default options, and manually edited to remove large gaps and regions that could not be reliably aligned. Bayesian phylogenies were constructed using MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003) and Maximum Likelihood phylogenies were estimated using PHYML 3.0.1 (Guindon and Gascuel 2003). Both phylogenic analyses were implemented using a WAG model (Whelan and Goldman 2001) and a gamma distribution with four rate categories. The proportion of invariant sites was estimated by maximizing the likelihood of the phylogeny. For Bayesian phylogenies, two independent Metropolis-coupled Monte Carlo Markov Chain (MCMCMC) runs, each with one cold and three heated chains (heat parameter = 0.2), were analyzed for one million generations. Ten thousand trees were saved in each run (every 100th tree was saved) and the phylogenetic trees were summarized using 15,000 samples (burning parameter = 2,500), after confirmation that these samples were in the plateau phase of the run. For Maximum Likelihood phylogenies, 1,000 bootstraps were performed to achieve branch support values. To be considered as an indicative of a potential HGT event, the branch grouping *Moniliophthora* sequences with a distant related group must have Bayesian posterior probability ≥90% and ML bootstrap support ≥80%.

Each phylogeny suggestive of HGT was compared with another one of the same sequences constructed with the

imposition of a sister group relationship among *Moniliophthora* and ascomycetes sequences. This constrained phylogeny was constructed by MRBAYES 3.1.2 using the same model described above. The comparison between the ancestral and the constrained phylogeny was done using the Bayes Factors (Kass and Raftery 1995). Twice the difference in the harmonic mean of log-likelihoods from each MCMCMC run (the Bayes Factor) was used to determine relative support for alternative hypotheses. Values for the Bayes Factors of >10 indicate a strong support for the less constrained tree over the a priori imposed constraint (Kass and Raftery, 1995).

Gene Loss and Duplication

The evaluation of gene loss and duplication in eukaryotes under the hypothesis of no HGT were made using a simplified topology of the eukaryotes phylogeny with the following ten lines: *Moniliophthora* (I), the five basidiomycete species with complete genome sequences mentioned above (II–VI); ascomycetes (VII); animals (VIII), land plants and green and red algae (IX, named as viridiplantae), and alveolates, kinetoplastids, and oomycetes (X, named as stramenopiles and alveolates). This phylogeny was manually obtained by merging relevant branches of the eukaryotes phylogeny proposed by Van de Peer et al. (2000) and from the fungi phylogeny proposed by James et al. (2006). The presence and absence of each putative horizontally transferred gene in each of these eukaryote lines were manually inferred from the results of BLASTp.

Further Analyses

Sequences with significant homologies to the putative horizontally transferred genes of *M. perniciosa* were found in the *M. roreri* genome draft sequence using tBLASTn. This draft was produced with reads obtained by Illumina/Solexa (33× coverage) and GS-FLX 454 single ends (15× coverage) technologies and comprise 53 Mbp assembled into 8,256 contigs (unpublished data). Similar sequences to these genes were also found in a *M. perniciosa* EST database containing 1.4 Mbp in 2,474 sequences, including contigs and singlets, which can be used to check for gene expression of the putative HGT genes.

GC contents of *M. perniciosa* putative horizontally transferred genes were compared with those of other predicted coding genes from the same species, to find extreme values of GC that may indicate a horizontal origin for the gene. GC contents of *M. perniciosa* predicted coding genes were obtained using a PERL script. The rank and the average GC content were obtained using Microsoft Excel™.

Results

HGT Candidates Search

This study used a pipeline with two filters to search for *M. perniciosa* predicted coding genes that are candidates obtained by HGT as summarized in Fig. 1. In the first filter, the 14,501 predicted coding genes of *M. perniciosa* were compared with basidiomycete sequences available in GenBank, which excluded 10,612 genes as possible HGT candidates since they have homologs in this group (Fig. 1a). This filter used BLASTp and tBLASTn to detect homologies with both annotated and non-annotated proteins.

In the second filter, the 3,889 predicted genes without homologous sequences in basidiomycetes (Fig. 1b) were submitted to a BLASTp search against nr database, producing three categories of results. The first category consists of 3,654 no-hit genes, which can be either unknown genes, some of which may be exclusive to *M. perniciosa*, or closely related basidiomycetes, or erroneous predictions (Fig. 1c). The second category had 198 genes whose first-hit are fungal sequences (all of them from ascomycetes, Fig. 1d). These two categories were excluded as HGT candidates. The third category comprised the remaining 37 predicted genes with high homology to non-fungal species and only these genes were considered reliable HGT candidates (Fig. 1e). Nine out of these 37 HGT candidates have only one hit in nr (Fig. 1f). Among these, 2 have their highest homology to animal sequences, one to protozoan sequences and six to bacterial sequences (Table 1). These nine candidates are considered possible HGT genes, but the

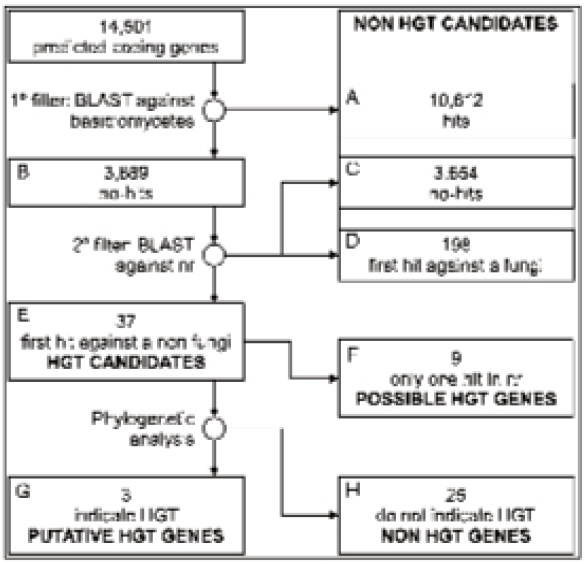


Fig. 1 Overview of the methodology used to search for HGT genes in *Moniliophthora* and of the results of each search step

Table 1 Predicted genes of *M. perniciosa* without homologs in basidiomycetes and with only one hit in the GenBank nr database

Gene	Hit organism	Hit accession number	Hit definition	*E* value
1	*Nitrosomonas eutropha C91*	YP_743769.1	Hypothetical protein	$7e^{-52}$
2				$5e^{-51}$
3	*Enterobacter* sp. 638	YP_001176630.1	Hypothetical protein	$2e^{-26}$
4				$3e^{-21}$
5	*Ralstonia eutropha JMP134*	YP_294602.1	Twin-arginine translocation pathway signal	$5e^{-18}$
6	*Gloeobacter violaceus*	NP_924911.1	WD-repeat protein	$1e^{-11}$
7	*Homo sapiens*	XP_001130133.2	Hypothetical protein	$5e^{-11}$
8	*Caenorhabditis briggsae*	CAP21100.1	Hypothetical protein	$9e^{-11}$
9	*Tetrahymena thermophila SB210*	XP_001028914.1	DDE superfamily endonuclease containing protein	$1e^{-10}$

Genes 1 and 2 have the same hit, and the same with genes 3 and 4

limited number of similar sequences hindered further phylogenetic analysis to validate this hypothesis.

The 28 HGT candidates with more than one hit in nr database produced 25 phylogenies that did not support the HGT hypothesis (Fig. 1h). The remaining three candidates fulfilled three criteria used here as indicative of HGT: (1) grouping with sequences from a non-related t group with the gene on a well-supported branch; (2) when the candidate gene is included as a sister group of the sequences from the closest gene group (i.e., ascomycetes for all of them) produces a significantly less supported phylogeny; and (3) the alternative hypothesis of gene evolution by loss and duplication produces a large number of these events. These three candidates are analyzed in detail, as follows.

Metallo-Dependent Hydrolase

The first putative horizontally transferred gene found in *M. perniciosa* is a metallo-dependent hydrolase (MDH) similar to imidazolonepropionase and other related amidohydrolases (COG1228, e-value $4e^{-40}$ at CDD search). *M. perniciosa* MDH GC content is 52.98 and 7.5% of the predicted genes of this species have GC content higher than this value. An MDH homolog was also found in the *M. roreri* genome.

Figure 2a shows the phylogeny of MDH. Some of its internal branches have low posterior probabilities and there are bacterial groups divided by these branches. In addition, some distinct bacterial lines were grouped together into well-supported branches. Ascomycete sequences are split into two well-supported groups, both related with proteobacteria. Most ascomycete species have only one paralog, but *Penicillium marneffei* and *Neurospora crassa* have both, and the same occurs with some proteobacteria (i.e., *Sphingomonas* sp. and *Caulobacter* sp.). The *M. perniciosa* MDH groups with actinobacteria from the genera *Salinispora* and *Mycobacterium* in another well-supported branch. The comparison between this phylogeny and the other with ascomycetes sequences constrained as a sister group of *M. perniciosa* produced a Bayes factor of 1,118.24, which constitutes very strong evidence in favor of the original unconstrained phylogeny (Table 2).

The other major eukaryotic lines do not have homologs to the MDH gene. The hypothesis suggesting that presence or absence of MDH paralogs in eukaryotes was caused by gene loss and duplication rather than HGT requires a minimum of 17 independent events, which is an improbable recurrence (Fig. 2b). An explanation using HGT would require only one event of transference to *Moniliophthora* (from actinobacteria) and eight gene losses in eukaryotes (Fig. 2c) or, alternatively, three independent HGT: one for *Moniliophthora* (from actinobacteria) and two for ascomycetes (both from proteobacteria).

Necrosis-Inducing Proteins (NEPs)

M. perniciosa NEPs (pfam05630, e-value $4e^{-44}$ from CDD search) were also found as having been acquired by HGT. *M. perniciosa* has three copies of this gene, while *M. roreri* has two. *M. perniciosa* NEPs GC contents are 54.26% for NEP1, 55.16% for NEP2, and 56.57% for NEP3. Only 3, 1.6, and 0.7% of the predicted genes of *M. perniciosa* have higher GC contents than those three NEP genes, respectively.

Figure 3a shows the NEPs phylogeny. *M. perniciosa* sequences are grouped with oomycete; while ascomycete sequences are split into two groups, one related to actinobacteria and the other to firmicutes and proteobacteria. Many ascomycetes from the genera *Botrytis*, *Botryotinia*,

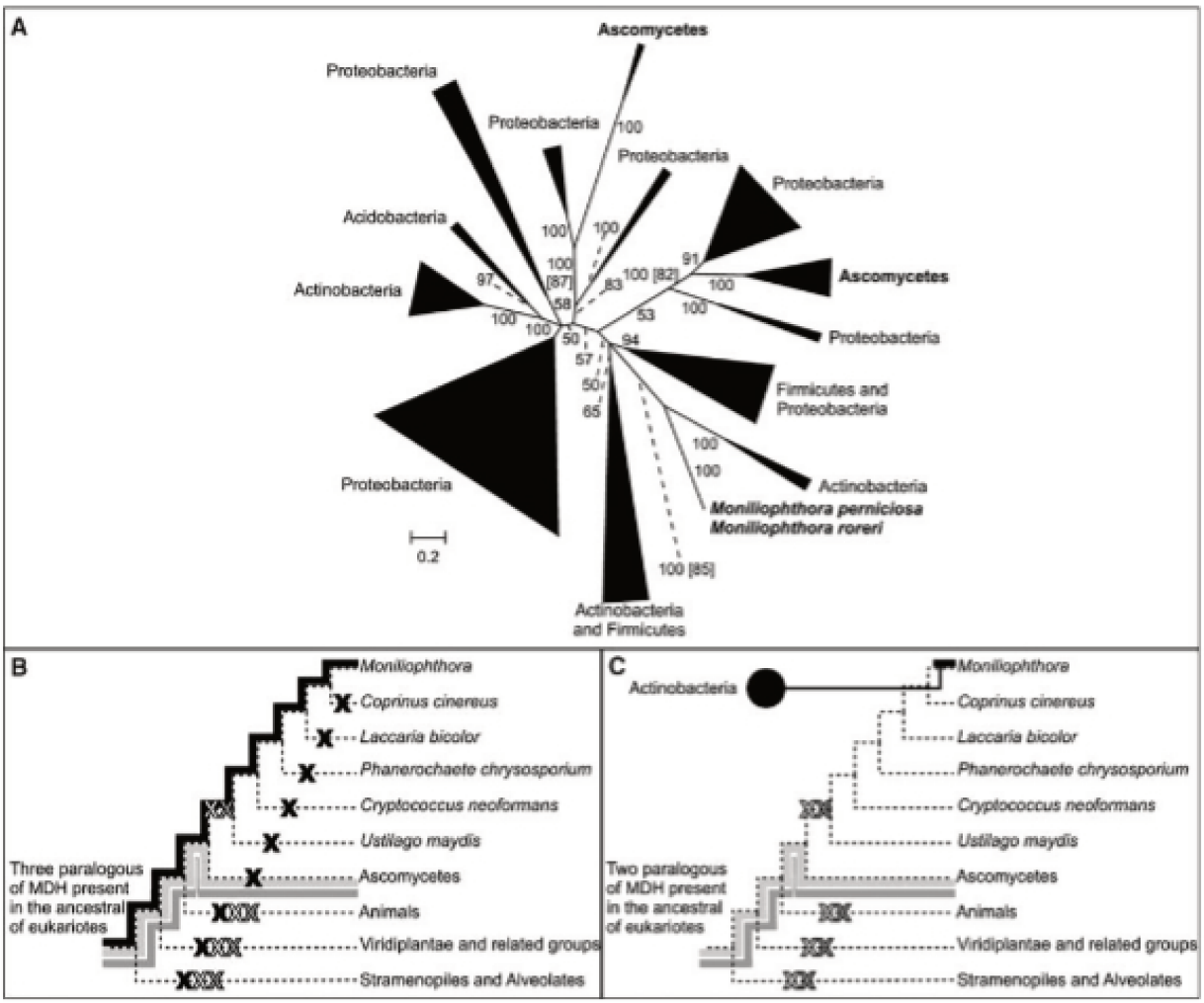


Fig. 2 **a** Phylogeny of metallo-dependent hydrolase (MDH). Values near internal branches (or linked to them by *dotted lines*) are posterior probabilities. Bootstrap values are shown in *brackets* for branches discussed in the text. The width of each triangle is proportional to the number of sequences included in the group and their vertices are positioned at the root of the group. Groups with names in *bold face* are eukaryotes. All other groups are bacteria. Two alternative hypotheses to explain the evolution of MDH in eukaryotes are shown in (**b**) and (**c**). **b** Evolution of eukaryotic MDH by only gene loss and duplication. Homologous forms of MPDH present in *Moniliophthora* and in ascomycetes are represented over the topology of the eukaryotic phylogeny (*dotted tree*) by the *black*, *dark*, and *light gray continuous lines*, respectively. Each gene loss is indicated by an **X** of the same color for the lost paralogs. Since *Moniliophthora* sequences do not group with ascomycetes sequences in phylogeny (**a**) and these are spliced in two groups, three paralogs and a minimum of 17 independent gene losses are necessary to support this hypothesis. The three paralogs must be present in a eukaryotic ancestor, since they also exist in bacteria. **c** Evolution of eukaryotic MDH by means of only HGT. One event of HGT from actinobacteria to *Moniliophthora* and eight gene losses are required to support this hypothesis. *Colors* are as in (**b**)

and *Aspergillus* have the two forms of NEPs. The comparison between this phylogeny and the other with ascomycetes sequences constrained as a sister group of *M. perniciosa* produced a Bayes factor of 779.74, which constitutes very strong evidence in favor of the original unconstrained phylogeny (Table 2).

To explain the absence of NEPs in the other eukaryotes lines, the separation of ascomycete sequences into two groups and the absence of a direct relationship between ascomycetes and *Moniliophthora* sequences without invoking HGT, requires at least three ancestral gene forms and 16 gene losses (Fig. 3b). An explanation for the presence of these genes through HGT would require only two events of this kind in eukaryotes, one from ascomycetes to oomycetes and the other from oomycetes to *Moniliophthora* (Fig. 3c).

Table 2 Results of the Bayes Factor analyses of putative horizontally transferred gene phylogenies

Gene	Log likelihood		Bayes factor
	Obtained phylogeny	Constrained phylogeny	
MPDH	−50,648.84	−51,420.30	1,542.92
MDH	−148,788.30	−149,347.42	1,118.24
NEP	−22,408.81	−22,798.68	779.74

The constrained phylogenies are those that are constructed with the imposition of a sister group relationship among *Moniliophthora* and ascomycetes sequences. The obtained phylogenies are the original phylogenies obtained without any model-constrain. Bayes factor are calculated as 2× (log-likelihood of obtained phylogeny − log-likelihood of constrained phylogeny). Values >10 are considered as very strong evidence against model 1 (the original phylogeny, in this case)

Mannitol 1-Phosphate 5-Dehydrogenase

The third putative horizontally transferred gene found in *M. perniciosa* encodes a mannitol 1-phosphate 5-dehydrogenase (MPDH; EC number 1.1.1.17; e-value $1e^{-118}$), an enzyme that converts fructose 6-phosphate to mannitol 1-phosphate in a reversible reaction. GC content of *M. perniciosa* MPDH is 50%, very close to the average for this species (49.85%). A homolog for this gene was found in the genome of *M. roreri*.

Figure 4a shows the phylogeny of MPDH. The *M. perniciosa* gene is grouped with genes from *Clostridium* sp., a genus that belongs to the firmicutes bacterial group, while the ascomycete sequences are grouped with actinobacteria sequences. The attempt to constrain *M. perniciosa* sequence as a sister group from the ascomycetes produced a phylogeny with a marginal log-likelihood of −51,420.30, while the unconstrained phylogeny has a marginal log-likelihood of −50,648.84. Twice the difference between the marginal likelihoods of the two phylogenies (their Bayes factor) amounts to 1,542.92, which constitutes very strong evidence in favor of the unconstrained phylogeny (Table 2).

The proteobacteria MPDH sequences are separated into two groups, suggesting the existence of at least two paralogs for this gene. The two most probable branches that separate MPDH genes are indicated in Fig. 2a. Ascomycetes and *Moniliophthora* do not share the same paralogs, irrespective of the exact branch that separate them, and no other eukaryote lines have any paralogs of this gene. The hypothesis suggesting that this pattern of presence/absence of MPDH genes in eukaryotes was generated only by gene loss and duplication (without HGT) would require a minimum of 13 independent losses, which is an improbable recurrence (Fig. 4b). On the other hand, an explanation for the presence of this gene through HGT would require only two independent HGT: one from firmicutes to *Moniliophthora* and the other from actinobacteria to ascomycetes (Fig. 4c).

Discussion

The Genus *Moniliophthora* Likely Contains More than the Three Identified Horizontally Transferred Genes

HGT events are usually screened for by phylogenetic analysis or by anomalous gene signatures. Methods based on phylogenetic analysis are believed to be the most robust if gene phylogenies can be reconstructed with reliability (Poptsova and Gogarten 2007). However, it is necessary to reduce the number of potential HGT candidates before performing phylogenetic analysis since computing resources required to reconstruct the phylogeny of all genes found in a complete genome is usually very extensive. In this study, the detection of genes obtained by HGT was done by means of phylogenetic analysis of candidates selected from the predicted genes of *M. perniciosa*.

The first filter used to reduce the number of HGT candidates in *M. perniciosa* genome was to discard genes with homologs in other basidiomycetes, to remove those obtained by vertical inheritance. Although this filter correctly removed the vertically inherited genes, it could have also removed three classes of horizontally transferred genes: (I) those from ancestral HGT events that occurred in basidiomycetes; (II) those that are transferred to other basidiomycetes in independent events (multiple HGT); and (III) those from gene replacements (the substitution in *Moniliophthora* of a gene that is also present in other basidiomycetes by a homologous transfer from a more distant related species). This latter type of exclusion may have a significant effect in this study, since a previous study (Mondego et al. 2008) suggests that some gene families are expanded in *Moniliophthora*, and gene replacement could be included in this expansion.

In the second filter used for HGT candidate selection, the predicted genes of *M. perniciosa* without homologs in basidiomycetes were used to search the GenBank nr database, and those showing highest similarity to that of ascomycete species were excluded (Fig. 1d). This filter removes genes present at the root of the fungal phylogeny but not in the five completely sequenced species of basidiomycete. However, this filter could also remove genes transferred from ascomycetes to *Moniliophthora*. The failure to detect HGT from ascomycetes could be an important limitation in this study, especially since another cacao pathogen, *Ceratocystis cacaofunesta*, belongs to this group, and could be a potential gene donor. This limitation could be avoided with the construction of comparative

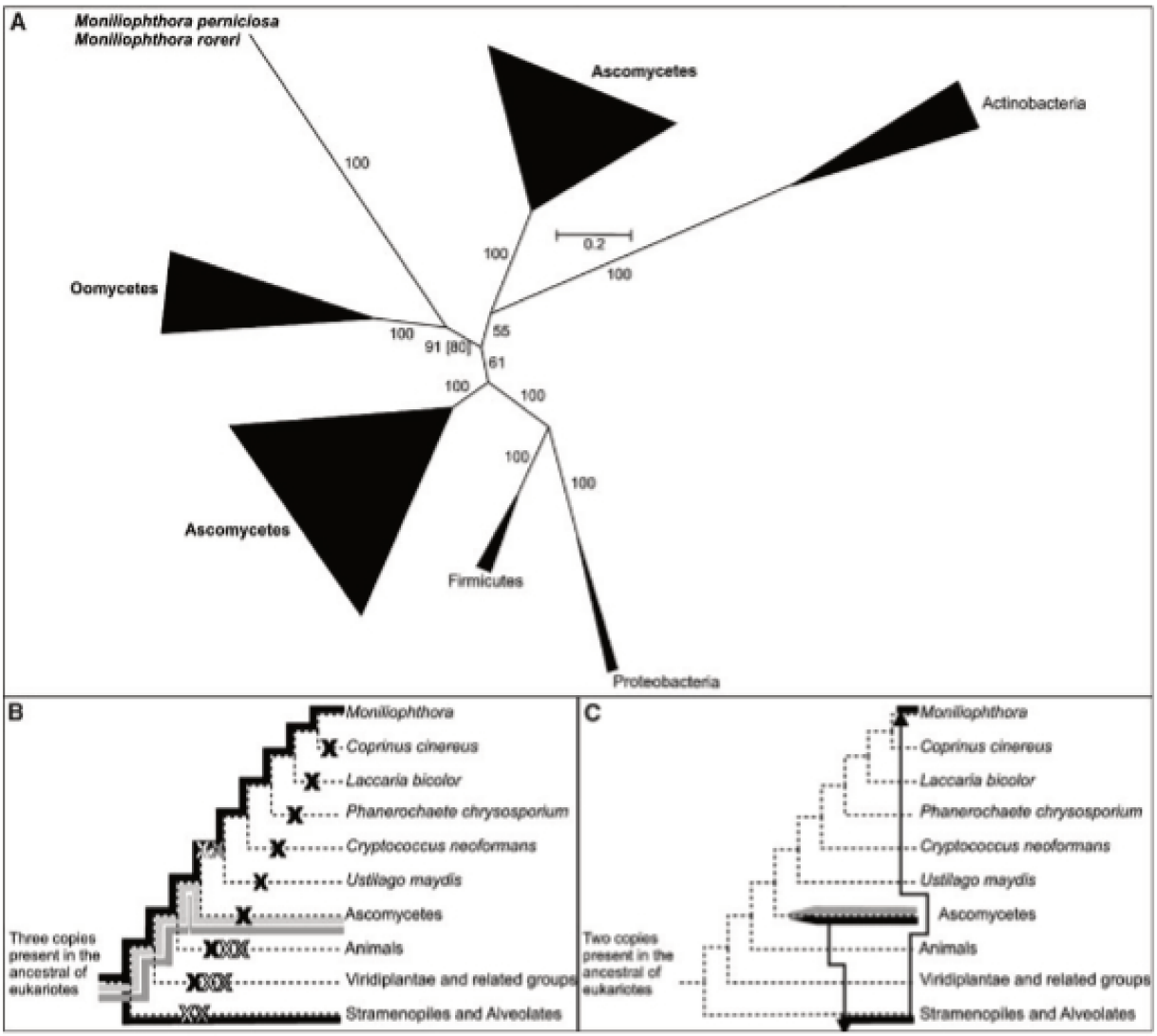


Fig. 3 **a** Phylogeny of necrosis ethylene-inducing protein (NEP). Values near internal branches are posterior probability. Bootstrap values are shown in *brackets* for some relevant branches discussed in the text. The width of each triangle is proportional to the number of sequences included in the group and their vertices are positioned in the root of the group. Groups with names in *bold face* are eukaryotes. All other groups are bacteria. Two alternative hypotheses to explain the evolution of NEP in eukaryotes are shown in (**b**) and (**c**). **b** Evolution of eukaryotic NEP by only gene loss and duplication. Homologous forms of NEP present in *Moniliophthora*/oomycetes and in ascomycetes are represented over the topology of the eukaryotic phylogeny (*dotted tree*) by the *black*, *dark*, and *light gray continuous lines*, respectively. Each gene loss is indicated by an **X** of the same color for the lost paralogs. Since *Moniliophthora* sequences do not group with ascomycetes sequences in phylogeny (**a**) and these are spliced into two groups, three paralogs and a minimum of 16 independent gene losses are necessary to support this hypothesis. The three paralogs must be present in a eukaryotic ancestor, since two of them also exist in bacteria (those from ascomycetes) and one of them (that from *Moniliophthora* and oomycetes) has no direct relationship with any other. **c** Evolution of eukaryotic MDH by means of only HGT. One event of HGT from oomycetes to *Moniliophthora* and another from ascomycetes to oomycetes are required to support this hypothesis. *Colors* are as in (**b**). In this hypothesis, the sequences of oomycetes can be related to one of the paralogs of ascomycetes

phylogenies for all the genes with similarity to ascomycetes. Unfortunately the computational requirements needed to implement all these phylogenies are too great at this time, at least for Bayesian or ML approaches, since there are at least 170 genes within the *M. perniciosa genome* with more than one hit in the nr.

J Mol Evol (2010) 70:85–97

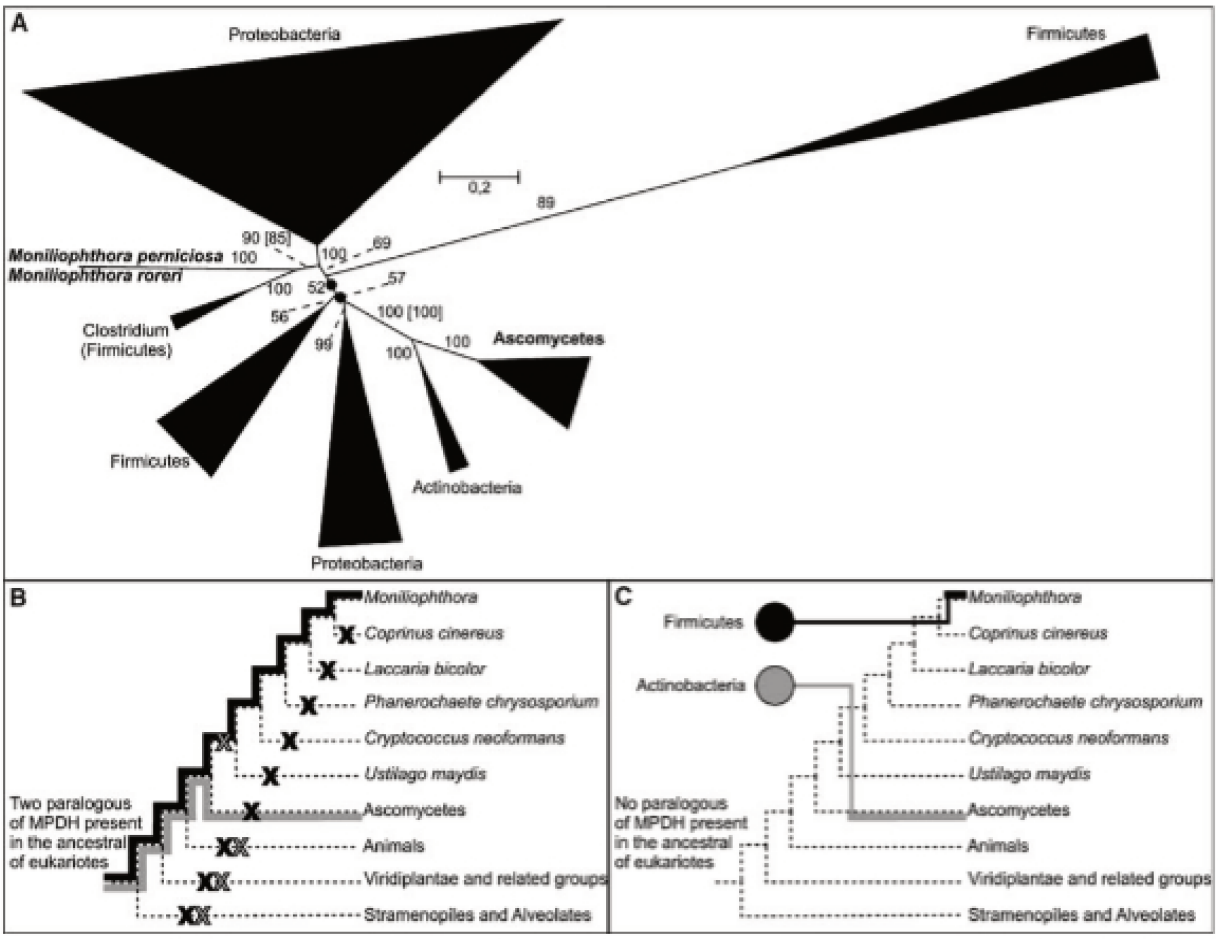


Fig. 4 **a** Phylogeny of mannitol phosphate dehydrogenase (MPDH). Values near internal branches (or linked to them by *dotted lines*) are posterior probabilities. Bootstrap values are shown in *brackets* for some relevant branches discussed in the text. The width of each triangle is proportional to the number of sequences from the group and their vertices are positioned at the root of the group. *Solid circles* mark the two most probable branches that could separate paralog forms of MPDH. Groups with names in *bold face* are eukaryotes. All other groups are bacteria. Two alternative hypotheses to explain the evolution of MPDH in eukaryotes are shown in (**b**) and (**c**). **b** Evolution of eukaryotic MPDH by only gene loss and duplication. Homologous forms of MPDH present in *Moniliophthora* and in ascomycetes are represented over the topology of the eukaryotic phylogeny (*dotted tree*) by the *black* and the *gray continuous lines*, respectively. Each gene loss is indicated by an **X** of the same color for the lost paralogs. Since *Moniliophthora* sequences do not group with ascomycetes sequences in phylogeny (**a**) and the other lines of eukaryotes do not have this gene, two paralogs and a minimum of 13 independent gene losses are necessary to support this hypothesis. Both paralogs must be present in a eukaryotic ancestor, since they also exist in bacteria. **c** Evolution of eukaryotic MPDH by means of only HGT. Two independent events of HGT are required to support this hypothesis, one from firmicutes to *Moniliophthora* and other from actinobacteria to an ancestral ascomycete. *Colors* are as in (**b**)

The Origins of Horizontally Transferred Genes in *Moniliophthora* Are Consistent with Their Evolution from a Saprophytic Ancestor

Phylogenetic analysis indicates that two out of the three horizontally transferred genes found in *Moniliophthora* have originated from bacteria. MDH and MPDH from *Moniliophthora* are related with those from actinobacteria and firmicutes, respectively, and these two groups are common in soil and decomposed organic material (Youssef and Elshahed 2009). The occurrence of soil bacteria among gene donors is consistent with the presumed evolutionary history of phytopathogenic *Moniliophthora* (Aime and Phillips-Mora 2005). Most of the other known fungal species related to *Moniliophthora* are saprotrophic and soils are a common habitat for species with this lifestyle. Thus, the common ancestor of *M. roreri* and *M. perniciosa* could have obtained MDH and MPDH genes in this habitat before evolving into a pathogen. Gene exchanges among free-living soil microorganisms was proposed as being important to HGT in bacteria living in the planktonic and attached states (Massoudieh et al. 2007) and the same

could also be true for soil fungal communities. The enzymes released by saprophytic fungi favor the breakdown of cell walls and membranes that surround organisms, with the consequential release of genetic material. Continuous contact of the fungi with this material could facilitate the uptake of DNA fragments containing complete genes and other elements that could mediate their incorporation in the acceptor's DNA, such as transposons or plasmids (Kempken 1995). HGT mechanisms appear to be more suitable to fungi than others proposed for eukaryotes, such as the establishment of symbiosis with unicellular organisms or the mediation of gene acquisition by the phagotrophic habit; the so called you are what you eat hypothesis (Doolittle 1998). It is interesting to note that ascomycetes apparently obtained their MPDH by HGT from actinobacteria (Fig. 2a). The wide distribution of the MPDH enzyme in ascomycetes and the monophylogeny of their sequences suggest that, unlike *Moniliophthora*, MPDH was transferred to this group at a basal point in their evolution. Considering the variety of possible relationships between the mannitol and the phytopathogenicity in fungi (see discussion below) it is possible to hypothesize that MPDH is important to a pathogenic lifestyle, and to further extrapolate it may be suggested that this ancestral HGT is one of the reasons that there are more phytopathogenic species among ascomycetes than in the basidiomycetes.

Contrary to MDH and MPDH, the NEPs from *Moniliophthora* have homologs in another eukaryotic group, the oomycetes. The most likely NEP donors are species of the *Phytophthora* genus, which also infect cacao, such as *P. palmivora*. The similarity between the GC content of the NEPs from *M. perniciosa* and *P. megakarya* (approximately 56%) shows a good evidence of this origin (Garcia et al. 2007). This transmission may have occurred in necrotic tissues from stems, fruits, or leaves released by a cacao plant infected by *Phytophthora* and invaded by a saprotrophic ancestral species of *Moniliophthora*. The contact between the living cells of the fungus and the dead cells of oomycetes could have resulted in the transfer of DNA fragments containing copies of the NEP genes from the latter to the former species. Another possibility is that the ancestral *Moniliophthora* acquired a phytopathogenic lifestyle before acquiring the NEPs and this transference has occurred in a cacao plant infected by both this ancestor and by a *Phytophthora*. Double infections like this are commonly seen in cacao plantations in Brazil, as reported for *P. palmivora* and *M. perniciosa* by Flament et al. (2001). Since cerato-platanins are present in the genome of *Moniliophthora* and other basidiomycetes and can also promote necrosis in cacao plants (Zaparoli et al. 2009), it is possible that in the early stages of their evolutionary change the ancestral *Moniliophthora* promoted necrosis using only these proteins. In this scenario, NEPs would have been the last horizontally transferred gene obtained by the ancestor of *Moniliophthora*, a feature consistent with the extreme GC content of these genes in *M. perniciosa*. There is evidence that codon-usage and GC content could be adaptive and that horizontally acquired genes tend to adopt values close to the "normal" for the species after some time (Goodarzi et al. 2008).

Ascomycetes have two different paralogs of NEPs, and each one of them is present in a different group of bacteria. Furthermore, unlike MDH and MPDH, NEPs are more common among ascomycetes than in bacteria, and many species among these fungi have both NEPs paralogs, while bacteria, oomycetes, and *Moniliophthora* species have only one. This scenario suggests that NEPs originated in ascomycetes, and were later horizontally transferred to bacteria (in two different events, one for each paralog) and to oomycetes, and from this group to *Moniliophthora* (Fig. 4c).

The Horizontally Transferred Genes Found in *Moniliophthora* Are Related to Phytopathogeny in Fungi

The phylogenies and gene loss analysis of MPDH, MDH, and NEPs indicate that these genes were obtained by HGT in a common ancestor of pathogenic *Moniliophthora*. Among these genes only NEPs were previously analyzed (Garcia et al. 2007). In that article, the NEPs of *M. perniciosa* were concentrated in the apoplast and external surfaces of cacao cell walls during WBD and were able to induce necrosis of infected tissues. *M. roreri* NEPs have not yet been studied, but their strong resemblance with *M. perniciosa* sequences and the occurrence of necrosis in FPR suggest that they are active proteins. NEPs are well characterized in oomycetes, but are also common in phytopathogenic ascomycetes (Jenning et al. 2000; Gijzen and Nurnberger 2006; Pemberton and Salmond 2004). Their presence in *Moniliophthora* suggests that this fungus promotes an infection that is more similar to that of ascomycetes than to the other phytopathogenic basidiomycetes.

Metallo-dependent hydrolases are a large group of proteins that show a conserved metal binding site. It is proposed that the metal ion (or ions) deprotonates a water molecule for a nucleophilic attack on the substrate. Some amidohydrolases were found to be involved in the metabolic reactions of amino acids and nitrogenated bases, which acts on carbon–nitrogen bonds, and in many cases releases ammonia, glutamate, or biurate (Holm and Sander 1997). MDHs from *Moniliophthora* are also similar to those from bacteria that inhabit harsh environments such as *Salinispora arenicola* (marine sediments), *Natranaerobius thermophilus* (sediment of alkaline hypersaline lakes), and *Sphingomonas wittichii* (highly contaminated natural

environments with pesticides and polyhalogenated aromatic compounds) Based on the hydrolytic activity of MDHs and on their presence in bacteria that thrive in extreme conditions, the MDH from *Moniliophthora* may function with hydrolytic detoxifying capacity and/or in nitrogen metabolism, an hypothesis that could be experimentally tested.

In contrast to MDH, MPDH is strongly related with phytopathogeny in fungi and is part of the ascomycete mannitol cycle. In this cycle, fructose is phosphorylated to fructose 6-phosphate by hexokinase (EC 2.7.1.1). Then fructose 6-phosphate is converted to mannitol 1-phosphate by MPDH (EC 1.1.1.17). The posterior removal of the phosphate group by a phosphatase (EC 3.1.3.22) produces inorganic phosphate and mannitol. The cycle is closed by MADH (EC 1.1.1.138), which catalyzes the reversible conversion of mannitol to fructose. According to Dijksterhuis and De Vries (2006), mannitol (a polyol or sugar alcohol) is one of the compatible solutes used by fungi in response to a variety of stress conditions. Polyols are considered transport and storage forms of carbon that are less accessible to cellular metabolism than glucose, and can be accumulated in high amounts without interfering with normal cellular proteins (Noiraud et al. 2001). Mannitol has been linked to thermotolerance in the entomopathogenic fungus *Beauveria bassiana* (Liu et al. 2008), to sporulation in the necrotrophic fungus *Stagonospora nodorum* (Solomon et al. 2006) and to carbohydrate storage and as a reactive oxygen species (ROS) scavenger in *Uromyces fabae*, a biotrophic basidiomycete pathogen of *Vicia faba* (Voegele et al. 2005). Mutation of the MPDH gene in *Alternaria alternata* produced mutants that were less virulent on tobacco than the wild-type fungus, whereas MADH mutants have little effect on virulence (Vélëz et al. 2008). Interestingly, the double mutant had the most severe decrease in virulence (Vélëz et al. 2008).

MPDH seems to be absent from most basidiomycetes (Hult et al. 1980), indicating that they do not have a mannitol cycle and produce mannitol only by direct reduction of fructose, which is catalyzed by MADH. In *U. fabae*, a large increase in the concentration of mannitol during the biotrophic phase of infection was reported and the mannitol produced at this stage is used as a carbohydrate source in spores (Voegele et al. 2005). Curiously, a tBLASTn search did not detect a gene similar to MADHs of *U. fabae* or *Agaricus bisporus* in *M. perniciosa* or *M. roreri* genomes or the *M. perniciosa* ESTs sequences. The presence of MPDH and the absence of MADH in their genome indicate that these two species may have a mannitol metabolism more similar to that of ascomycetes than to basidiomycetes. Even though the ability of *M. perniciosa* and *M. roreri* to use mannitol is unknown, the data cited above allow the creation of functional hypotheses about the role of MPDH during WBD and FPR that can be tested with the analysis of the mannitol content in green and dry brooms and infected cacao fruits, differential expression of MPDH during the life cycle of both species and, possibly, immunolocalization assays of the MPDH enzyme.

Final Considerations

In summary, this study indicates that some genes present in *Moniliophthora* species could have been acquired by horizontal transfer. It is remarkable that two of them—MPDH and NEPs—have been previously reported as involved in phytopathogenicity. Taken together, these results suggest that the ability of these fungal species to cause disease in cacao could have been acquired, or at least improved, by HGT.

Many other genes and gene families from the *M. perniciosa* genome were described as related to the hemibiotrophic and pathogenic lifestyle of *M. perniciosa* (Mondego et al. 2008), but all these are ancestral genes and many are present in other non-phytopathogenic basidiomycetes, suggesting that HGT has acted as an important step in this evolutionary change. It should be stressed that each one of the horizontally transferred genes came from a different donor, including prokaryotes and eukaryotes. The great biodiversity present in tropical forest habitats such as the Amazon Basin (the suggested center of origin for both *Moniliophthora* species) and Bahia Atlantic Rain Forest (the most productive cacao growing region in Brazil) increase the chance of interactions between microorganisms and could consequently increase the probability of HGT. The diversity of microorganisms in these habitats is still poorly understood, and there is a high probability that many "new genes" may be found there. Therefore, some of the 3,654 predicted genes in *M. perniciosa* that do not have homology in GenBank could have been acquired by HGT, but the homologous sequences have not yet been found.

Due to their close evolutionary relationship, the same horizontally transferred genes were found in both *M. roreri* and *M. perniciosa*, which suggests that HGT happened before speciation of this genus. The *M. roreri* lifecycle during the disease process has not been completely unraveled. While this fungus shows a biotrophic phase during disease development, no identifiable fungal morphological changes have yet been found. The presence of the same horizontally transferred genes could help develop hypotheses for understanding the hemibiotrophic life style of *M. roreri*. Finally, the study of the role of HGT in phytopathogenicity emergence could contribute to the selection of possible interesting genes for further experimental analysis focusing on acquiring a better understanding of the emergence of phytopathogenic mechanisms

and of the lifestyles of these fungi, which could lead to the development of new control strategies against these pathogens.

Acknowledgments We thank Dr. Joana Rincones for review and commentaries about the article. We are indebted to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Secretaria de Agricultura, Irrigação e Reforma Agrária do Estado da Bahia (SEAGRI), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), and Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support.

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

Almeida OC, Chiacchio FPB, Rocha HM (1997) Sobrevivência de Crinipellis perniciosa (Stahel) Singer em vassouras secas de cacaueiros (*Theobroma cacao* L.) do estado da Bahia. Agrotrópica 9:23–28

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Dijksterhuis J, De Vries RP (2006) Compatible solutes and fungal development. Biochem J 399(2):e3–e5

Doolittle WF (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Flament MH, Kebe I, Clement D, Pieretti I, Risterucci AM, N'Goran JA, Cilas C, Despreaux D, Lanaud C (2001) Genetic mapping of resistance factors to Phytophthora palmivora in cocoa. Genome 44:79–85

Garcia O, Macedo JA, Tiburcio R, Zaparoli G, Rincones J, Bittencourt LM, Ceita GO, Micheli F, Gesteira A, Mariano AC, Schiavinato MA, Medrano FJ, Meinhardt LW, Pereira GAG, Cascardo JC (2007) Characterization of necrosis and ethylene inducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom in Theobroma cacao. Mycol Res 111:443–455

Gijzen M, Nurnberger T (2006) Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. Phytochem 67:1800–1807

Goodarzi H, Torabi N, Najafabadi HS, Archetti M (2008) Amino acid and codon usage profiles: adaptive changes in the frequency of amino acids and codons. Gene 407(1–2):30–41

Griffith GW, Nicholson J, Nenninger A, Birch RN, Hedger JN (2003) Witches' brooms and frosty pods: two major pathogens of cacao. New Zeal J Bot 41(3):423–435

Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52(5):696–704

Holm L, Sander C (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. Proteins 28(1):72–82

Hult K, Veide A, Gatenbeck S (1980) The distribution of the NADPH-regenerating mannitol cycle among fungal species. Arch Microbiol 128(2):253–255

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung G, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schüßler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher MP, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lücking R, Büdel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006) Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature 443:818–822

Jenning JC, Apel-Birkhold PC, Bailey BA, Anderson JD (2000) Induction of ethylene biosynthesis and necrosis in weed leaves by a *Fusarium oxysporum* protein. Weed Sci 48:7–14

Kamper J, Kahmann R, Bolker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Muller O, Perlin MH, Wosten HA, de Vries R, Ruiz-Herrera J, Reynaga-Pena CG, Snetselaar K, McCann M, Perez-Martin J, Feldbrugge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, Gonzalez-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Greilinger D, Munch K, Rossel N, Scherer M, Vranes M, Ladendorf O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho EC, Cahill MJ, Boyce KJ, Klose J, Klosterman SJ, Deelstra HJ, Ortiz-Castellanos L, Li W, Sanchez-Alonso P, Schreier PH, Hauser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schluter T, Margolis J, Platt D, Swimmer C, Gnirke A, Chen F, Vysotskaia V, Mannhaupt G, Guldener U, Munsterkotter M, Haase D, Oesterheld M, Mewes HW, Mauceli EW, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BW (2006) Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. Nature 444(7115):97–101

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795

Kempken F (1995) Horizontal transfer of a mitochondrial plasmid. Mol Gen Genet 2248:89–94

Khaldi N, Collemare J, Lebrun MH, Wolfe KH (2008) Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol 9:R18

Liu Q, Ying SH, Feng MG, Jiang XH (2008) Physiological implication of intracellular trehalose and mannitol changes in response of entomopathogenic fungus *Beauveria bassiana* to thermal stress. Antonie Van Leeuwenhoek 95(1):65–75

Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, Allen JE, Bosdet IE, Brent MR, Chiu R, Doering TL, Donlin MJ, D'Souza CA, Fox DS, Grinberg V, Fu J, Fukushima M, Haas BJ, Huang JC, Janbon G, Jones SJ, Koo HL, Krzywinski MI, Kwon-Chung JK, Lengeler KB, Maiti R, Marra MA, Marra RE, Mathewson CA, Mitchell TG, Pertea M, Riggs FR, Salzberg SL, Schein JE, Shvartsbeyn A, Shin H, Shumway M, Specht CA, Suh BB, Tenney A, Utterback TR, Wickes BL, Wortman JR, Wye NH, Kronstad JW, Lodge JK, Heitman J, Davis RW, Fraser CM, Hyman RW (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science 307(5713):1321–1324

Martin F, Aerts A, Ahrén D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J,

Yin T, Chalot M, Henrissat B, Kües U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of Laccaria bicolor provides insights into mycorrhizal symbiosis. Nature 452(7183):88–92

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, Coutinho PM, Henrissat B, Berka R, Cullen D, Rokhsar D (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22(6):695–700

Massoudieh A, Mathew A, Lambertini E, Nelson KE, Ginn TR (2007) Horizontal gene transfer on surfaces in natural porous media: conjugation and kinetics. Vadose Zone J 6:306–315

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(5):577–588

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, García O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araújo MR, de Moraes MH, Castro LA, Gramacho KP, Gonçalves MS, Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' broom disease of cacao. BMC Genomics 9:548

Noiraud N, Maurousset L, Lemoine R (2001) Transport of polyols in higher plants. Plant Physiol Biochem 39(9):717–728

Oliver RP, Solomon PS (2008) Recent fungal diseases of crop plants: is lateral gene transfer a common theme? Mol Plant Microbe Interact 21(3):287–293

Pemberton CL, Salmond GPC (2004) The Nep1-like proteins—a growing family of microbial elicitors of plant necrosis. Mol Plant Pathol 5:353–359

Pereira JL, de Almeida LCC, Santos SM (1996) Witches' broom disease of cocoa in Bahia: attempts at eradication and containment. Crop Prot 15(8):743–752

Phillips-Mora W (2003) Origin, biogeography, genetic diversity and taxonomic affinities of the cacao fungus *Moniliophthora roreri* as determined using molecular, phytopathological and morphophysiological evidence, doctoral dissertation. University of Reading, Reading, UK, 349 pp

Poptsova MS, Gogarten JP (2007) The power of phylogenetic approaches to detect horizontally transferred genes. BMC Evol Biol 7:45

Richards TA, Dacks JB, Jenkinson JM, Thornton CR, Talbot NJ (2006) Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. Curr Biol 16:1857–1864

Rincones J, Scarpari LM, Carazzolle MF, Mondego JM, Formighieri EF, Barau JG, Costa GG, Carraro DM, Brentani HP, Vilas-Boas LA, de Oliveira BV, Sabha M, Dias R, Cascardo JM, Azevedo RA, Meinhardt LW, Pereira GA (2008) Differential gene expression between the biotrophic-like and saprotrophic mycelia of the witches' broom pathogen *Moniliophthora perniciosa*. Mol Plant Microbe Interact 21(7):891–908

Rio MCS, Oliveira BV, Thomazella DP, Fracassi da Silva JA, Pereira GAG (2008) Production of calcium oxalate crystals by the basidiomycete *Moniliophthora perniciosa*, the causal agent of the Witches' broom disease of cacao. Curr Microbiol 56(4):363–370

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Sanders IR (2006) Rapid disease emergence through horizontal gene transfer between eukaryotes. Trends Ecol Evol 21(12):656–658

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 cacao in Brazil caused by *Crinipellis perniciosa*. J Exp Bot 56:865–877

Solomon PS, Waters ODC, Jorgens CI, Lowe RGT, Rechberger J, Trengove RD, Oliver RP (2006) Mannitol is required for asexual sporulation in the wheat pathogen *Stagonospora nodorum* (glume blotch). Biochem J 399:231–239

Stanke M, Waack S (2003) Gene prediction with a hidden-markov model and a new intron submodel. Bioinformatics 19(Suppl. 2):215–225

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. Annu Rev Phytopathol 46:75–100

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The clustalx windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Van de Peer Y, Baldauf SL, Doolittle WF, Meyer A (2000) An updated and comprehensive rRNA phylogeny of (crown) eukaryotes based on rate-calibrated evolutionary distances. J Mol Evol 51:565–576

Vélëz H, Glassbrook NJ, Daub ME (2008) Mannitol biosynthesis is required for plant pathogenicity by Alternaria alternata. FEMS Microbiol Lett 285(1):122–129

Voegele RT, Hahn M, Lohaus G, Link T, Heiser I, Mendgen K (2005) Possible roles for mannitol and mannitol dehydrogenase in the biotrophic plant pathogen *Uromyces fabae*. Plant Physiol 137(1):190–198

Wenzl P, Wong L, Kwang-Won K, Jefferson RA (2005) A functional screen identifies lateral transfer of beta-glucuronidase (gus) from bacteria to fungi. Mol Biol Evol 22:308–316

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691–699

Youssef NH, Elshahed MS (2009) Diversity rankings among bacterial lineages in soil. ISME J 3(3):305–313

Zaparoli G, Cabrera OG, Medrano FJ, Tiburcio RA, Lacerda G, Pereira GAG (2009) 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. Mycol Res 113:61–72

CONSIDERAÇÕES FINAIS.

As análises realizadas com o genoma de *M. perniciosa* e com genes individuais desta espécie permitiram uma melhor compreensão da evolução de alguns aspectos da fitopatogenia apresentada por esta espécie. Em particular ficou claro que a transferência gênica horizontal teve um papel relevante neste processo, ao mesmo tempo em que genes já presentes no repertório ancestral da espécie também puderam ser reaproveitados. Em alguns casos tais genes passaram por modificações mais extensas, como no caso da ceratoplatanina, que sofreu duplicações e divergência em relação as cópias ancestrais, com possível alteração de função em algumas destas cópias. Em outros casos o gene manteve os padrões de evolução típicos da família, como ocorreu com a AOX, indicando que mesmo uma mudança aparentemente grande no modo de vida de uma espécie pode ser feita aproveitando-se genes ancestrais em sua forma original. Dois dos três genes adquiridos horizontalmente por *M. perniciosa* ainda não haviam sido alvo de uma avaliação criteriosa em bancada, sendo que um deles, o da manitol fosfato desidrogenase (MPDH), tem sido descrito como relacionado a fitopatogenia em fungos ascomicetos. Este achado sugere que a busca por genes com padrões de evolução particulares pode ser uma ferramenta valiosa no estabelecimento de genes de interesse para este tipo de avaliação.

Análises “in silico” de dados de sequências genômicas têm produzido uma grande quantidade de informações sobre as características de genomas individuais. No entanto a pouca disponibilidade de espécies completamente sequenciadas entre os eucariotos e a concentração destas em apenas alguns grupos e modos de vida tem limitado a produção de análises de caráter comparativo, entre as quais se incluem as de cunho evolutivo. Para o *M. perniciosa* esta limitação se revelou particularmente restritiva, uma vez o número de espécies de basidiomicetos sequenciadas não chega a uma

dezena, e nenhuma delas é proximamente aparentada. A recente emergência de novas tecnologias de sequenciamento mais rápidas e baratas deverá mudar em breve o quadro descrito acima, permitindo a transformação da genômica de eucariotos de uma ciência descritiva para uma comparativa. Estas comparações certamente trarão uma nova visão das alterações ocorridas nos genomas eventos de especiação e de como estas alterações se relacionam com eventuais mudanças nos modos de vida das novas espécies. Em um aspecto mais prático será possível também escolher com mais confiabilidade genes e processos bioquímicos e fisiológicos a serem trabalhados em bancada conforme cada objetivo buscado pelo pesquisador. Eventualmente o sequenciamento genômico de espécies próximas não patogênicas do gênero *Crinipellis* permitirá o estabelecimento de um modelo mais completo das mudanças gênicas e genômicas ocorridas durante o estabelecimento do modo de vida patogênico no gênero *Moniliophthora*.

BIBLIOGRAFIA

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.

Almeida OC, Chiacchio FPB, Rocha HM (1997) Sobrevivência de *Crinipellis perniciosa* (Stahel) Singer em vassouras secas de cacaueiros (*Theobroma cacao* L.) do estado da Bahia. Agrotrópica 9: 23-28.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) "Basic local alignment search tool." J. Mol. Biol. 215: 403-410.

Bailey BA (1995) Purification of a protein from culture filtrates of *Fusarium-oxysporum* that induces ethylene and necrosis in leaves of erythroxylum-coca. Phytopathology 85: 1250-1255.

de Arruda MCC, Miller RNG, Ferreira MASV, Felipe MSS (2003) Comparison of *Crinipellis perniciosa* isolates from Brazil by ERIC repetitive element sequence-based PCR genomic fingerprinting. Plant Pathology, 52: 236-244.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. Journal of Molecular Biology 340: 783-795.

Ceita GO, Macêdo JNA, Santos TB, Alemanno L, Gesteira AS, Micheli F, Mariano AC, Gramacho KP, Silva DC, Pereira GAG, Cascardo JCM (2005). Susceptibility of *Theobroma cacao* to *Crinipellis perniciosa*: a programmed cell death triggered by calcium oxalate degradation. In: XII International Congress on Molecular Plant-Microbe Interactions, Mérida. Proceedings of the XII International Congress on Molecular Plant-Microbe Interactions. v. P10-13: 194-194.

Dayhoff MO, Eck RV, Park CM (1972) A model of evolutionary change in proteins. In: Dayhoff MO (ed), Atlas of Protein Sequence and Structure, Vol. 5. National Biomedical Research Foundation, Washington, D.C., pp. 89-99.

Dijksterhuis J, De Vries RP (2006) Compatible solutes and fungal development. Biochem J 399(2): e3-e5.

Doolittle WF (1998). You are what you eat: A gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14: 307-311.

Evans HC (1980) Pleomorphism in *Crinipellis perniciosa*, causal agent of witches broom disease of cocoa. Transactions of the British Mycological Society, 74: 515-523.

Evans HC (2007) Cacao diseases – The Trilogy Revisited. Phytopathology 97(12): 1640-1643.

Fellbrich G, Romanski A, Varet A, Blume B, Brunner F, Engelhardt S, Felix G, Kemmerling B, Krzymowska M, Nurnberger T (2002) NPP1, a *Phytophthora*-associated trigger of plant defense in parsley and arabidopsis. The Plant Journal 32: 375-390.

Finnegan PM, Whelan J, Millar AH, Zhang Q, Smith MK, Wiskich JT, Day DA (1997) Differential expression of the multigene family encoding the soybean mitochondrial alternative oxidase. Plant Physiol. 114(2): 455-66.

Flament MH, Kebe I, Clement D, Pieretti I, Risterucci AM, N'Goran JA, Cilas C, Despreaux D, Lanaud C (2001) Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. Genome 44: 79–85.

Formighieri EF, Tiburcio RA, Armas ED, Medrano FJ, Shimo H, Carels N, Góes-Neto A, Cotomacci C, Carazzolle MF, Sardinha-Pinto N, Thomazella DP, Rincones J, Digiampietri L, Carraro DM, Azeredo-Espin AM, Reis SF, Deckmann AC, Gramacho K, Gonçalves MS, Moura Neto JP, Barbosa LV, Meinhardt LW, Cascardo JC, Pereira GA. The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stable integrated plasmid. Mycol Res. 2008 Oct;112(Pt 10):1136-52.

Garcia O, Macedo JA, Tiburcio R, Zaparoli G, Rincones J, Bittencourt LM, Ceita GO, Micheli F, Gesteira A, Mariano AC, Schiavinato MA, Medrano FJ, Meinhardt LW, Pereira GAG, Cascardo JC (2007) Characterization of necrosis and ethyleneinducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom in Theobroma cacao. Mycol Res 111: 443-455.

Gijzen M, Nurnberger T (2006) Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. Phytochemistry 67: 1800-1807.

Goodarzi H, Torabi N, Najafabadi HS, Archetti M, 2008. Amino acid and codon usage profiles: adaptive changes in the frequency of amino acids and codons. Gene 407(1-2): 30-41.

Griffith GW, Nicholson J, Nenninger A, Birch RN, Hedger JN (2003) Witches' brooms and frosty pods: two major pathogens of cacao. New Zealand Journal of Botany, 41: 423-435.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology, 52(5): 696-704.

Holm L, Sander C (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. Proteins 28(1): 72-82.

Hult K, Veide A, Gatenbeck S (1980) The Distribution of the NADPH-regenerating Mannitol Cycle among Fungal Species. Archives of Microbiology 128(2): 253-255.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung G, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schüßler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher MP,

Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lücking R, Büdel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006) Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature 443: 818-822.

Jenning JC, Apel-Birkhold PC, Bailey BA, Anderson, JD (2000) Induction of ethylene biosynthesis and necrosis in weed leaves by a *Fusarium oxysporum* protein. Weed Sci 48:7-14.

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences." Computer Applications in the Biosciences 8(3): 275-282.

Joseph-Horne T, Derek W, Wood PM (2001) Fungal respiration: a fusion of standard and alternative components. Biochimica et Biophysica Acta 1504: 179-195.

Kamper J, Kahmann R, Bolker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Muller O, Perlin MH, Wosten HA, de Vries R, Ruiz-Herrera J, Reynaga-Pena CG, Snetselaar K, McCann M, Perez- Martin J, Feldbrugge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, Gonzalez-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Greilinger D, Munch K, Rossel N, Scherer M, Vranes M, Ladendorf O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho EC, Cahill MJ, Boyce KJ, Klose J, Klosterman SJ, Deelstra HJ, Ortiz-Castellanos L, Li W, Sanchez- Alonso P, Schreier PH, Hauser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schluter T, Margolis J, Platt D, Swimmer C, Gnirke A, Chen F, Vysotskaia V, Mannhaupt G, Guldener U, Munsterkotter M, Haase D, Oesterheld M, Mewes HW, Mauceli EW,

DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BW (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444(7115): 97-101.

Karpova OV, Kuzmin EV,Elthon TE, Newtona KJ (2002) Differential Expression of Alternative Oxidase Genes in Maize Mitochondrial Mutants . The Plant Cell, 14: 3271-3284.

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90: 773-795.

Keates SE, Kostman TA, Anderson JD, Bailey BA (2003) Altered gene expression in three plant species in response to treatment with Nep1, a fungal protein that causes necrosis. Plant Physiology 132: 1610-1622.

Khaldi N, Collemare J, Lebrun MH, Wolfe KH (2008) Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol 9: R18.

Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21: 676-679.

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150-163.

Liu Q, Ying SH, Feng MG, Jiang XH (2008) Physiological implication of intracellular trehalose and mannitol changes in response of entomopathogenic fungus *Beauveria bassiana* to thermal stress. Antonie Van Leeuwenhoek 95(1): 65-75.

Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, Allen JE, Bosdet IE, Brent MR, Chiu R, Doering TL, Donlin MJ, D'Souza CA, Fox DS, Grinberg V, Fu J, Fukushima M, Haas BJ, Huang JC, Janbon G, Jones SJ, Koo HL, Krzywinski MI, Kwon-Chung JK, Lengeler KB, Maiti R, Marra MA, Marra RE, Mathewson CA, Mitchell TG, Pertea M, Riggs FR, Salzberg SL, Schein JE, Shvartsbeyn A, Shin H,

Shumway M, Specht CA, Suh BB, Tenney A, Utterback TR, Wickes BL, Wortman JR, Wye NH, Kronstad JW, Lodge JK, Heitman J, Davis RW, Fraser CM, Hyman RW (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science 307(5713): 1321-1324.

Martin F, Aerts A, Ahrén D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kües U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452(7183): 88-92.

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, Coutinho PM, Henrissat B, Berka R, Cullen D, Rokhsar D (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22(6): 695-700.

Massoudieh A, Mathew A, Lambertini E, Nelson KE, Ginn TR (2007) Horizontal Gene Transfer on Surfaces in Natural Porous Media: Conjugation and Kinetics. Vadose Zone J 6: 306-315.

McGinnis S, Madden TL (2004) "BLAST: at the core of a powerful and diverse set of sequence analysis tools." Nucleic Acids Res. 32: W20-W25.

Meinhardt LW, Bellato CM, Rincones J, Azevedo RA, Cascardo JC, Pereira GAG (2006) In vitro production of biotrophiclike cultures of *Crinipellis perniciosa*,

the causal agent of witches' broom disease of *Theobroma cacao*. Current Microbiology 52: 191-196.

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(5):577-88.

Millenaar FF, Lambers H (2003). The alternative oxidase; in vivo regulation and function. Plant Biology. 5: 2-15

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, Garcia O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araujo MR, de Moraes MH, Castro LA, Gramacho KP, Goncalves MS, Moura Neto JP, Goes Neto A, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA. A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. BMC Genomics. 2008 Nov 18;9(1): 548.

Moore AL, Umbach AL, Siedow JN. 1995. Structure-function relationships of the alternative oxidase of plant mitochondria: a model of the active site. J. Bioenerg. Biomembr. 27: 367-77 .

Noiraud N, Maurousset L, Lemoine R (2001) Transport of polyols in higher plants. Plant Physiol and Biochem 39(9): 717-728.

Nylander JAA (2004) MrAIC.pl Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden.

Oliver RP, Solomon PS (2008) Recent fungal diseases of crop plants: is lateral gene transfer a common theme? Mol Plant Microbe Interact 21(3): 287-93.

Ottmanna C, Luberackia B, Kufnerc I, Kocha W, Brunnerc F, Weyandb M, Mattinend L, Pirhonend M, Anderluhe G, Seitza HU, Nurnbergerc T, Oeckinga C (2009) A common toxin fold mediates microbial attack and plant defense . PNAS 106(25): 10359-10364.

Pan SC, Cole GT (1995) Molecular and biochemical characterization of a *Coccidioides immitis*-specific antigen. Infection and Immunity 63: 3994-4002 .

Pazzagli L, Cappugi G, Manao G, Camici G, Santini A, Scala A (1999) Purification, characterization, and amino acid sequence of cerato-platanin, a new phytotoxic protein from *Ceratocystis fimbriata* f. sp. platani. Journal of Biological Chemistry 274: 24959-24964.

Pazzagli L, Pantera B, Carresi L, Zoppi C, Pertinhez TA, Spisni A, Tegli S, Scala A, Cappugi G (2006) Cerato-platanin, the first member of a new fungal protein family: cloning, expression, and characterization. Cell Biochemistry and Biophysics 44: 512-521.

Pemberton CL, Salmond GPC (2004) The Nep1-like proteins – a growing family of microbial elicitors of plant necrosis. Mol Plant Pathol 5: 353–359.

Pereira JL, de Almeida LCC, Santos SM (1996) Witches' broom disease of cocoa in Bahia: Attempts at eradication and containment. Crop Protection 15(8): 743-752.

Phillips-Mora W (2003) Origin, biogeography, genetic diversity and taxonomic affinities of the cacao fungus *Moniliophthora roreri* as determined using molecular, phytopathological and morpho-physiological evidence [Doctoral dissertation]. Reading, UK: Univ. Reading. 349 p.

Poptsova MS, Gogarten JP (2007) The power of phylogenetic approaches to detect horizontally transferred genes. BMC Evol Biol 7: 45.

Pungartnik C, Melo SC, Basso TS, Macena W, Cascardo JC, Brendel M (2009) Reactive Oxygen Species and Autophagy Play a Role in Survival and Differentiation of the Phytopathogen *Moniliophthora* perniciosa. Fungal Genet Biol. Mar 24. [Epub ahead of print]

Pemberton CL, Salmond GPC (2004) The Nep1-like proteins - a growing family of microbial elicitors of plant necrosis. Molecular Plant Pathology 5: 353-359.

Purdy LH, Schimidt RAA (1996) Status of cacao witches' broom: biology, epidemiology and management. Annu. Rev. Phytopathology, 34: 537-594.

Rementeria A, Lopez-Molina N, Ludwig A, Vivanco AB, Bikandi J, Ponton J, Garaizar J, 2005. Genes and molecules involved in *Aspergillus fumigatus* virulence. Revista Iberoamericana de Micologia 22: 1-23.

Richards TA, Dacks JB, Jenkinson JM, Thornton CR, Talbot NJ (2006) Evolution of Filamentous Plant Pathogens: Gene Exchange across Eukaryotic Kingdoms. Curr Biol 16: 1857-1864.

Rincones J, Scarpari LM, Carazzolle MF, Mondego JM, Formighieri EF, Barau JG, Costa GG, Carraro DM, Brentani HP, Vilas-Boas LA, de Oliveira BV, Sabha M, Dias R, Cascardo JM, Azevedo RA, Meinhardt LW, Pereira GA (2008) Differential gene expression between the biotrophic-like and saprotrophic mycelia of the witches' broom pathogen *Moniliophthora perniciosa*. Mol Plant Microbe Interact 21(7): 891-908.

Rio MCS, Oliveira BV, Thomazella DP, Fracassi da Silva JA, Pereira GAG (2008) Production of calcium oxalate crystals by the basidiomycete *Moniliophthora perniciosa*, the causal agent of the Witches' Broom Disease of cacao. Curr Microbiol 56(4): 363-370.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Sanders IR (2006) Rapid disease emergence through horizontal gene transfer between eukaryotes. Trends Ecol Evol 21(12): 656-658.

Santos Filho LP, Freire ES, Carzola IM (1998) Estimativas de perda de produção de cacau causadas por vassoura-de-bruxa (*Crinipellis perniciosa*, Aime & Phillips-Mora) na Bahia. Agrotrópica, 10: 127-130.

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 cacao in Brazil caused by *Crinipellis perniciosa*. J Exp Bot 56: 865-877.

Siedow JN, Umbach AL, Moore AL. 1995. The active site of the cyanide resistant oxidase from plant mitochondria contains a binuclear iron center. FEBS Lett. 362: 10-14 .

Silva SD, Luz EDMN, Almeida OC, Gramacho K, Bezerra JL (2002) Re-descrição da sintomatologia causada por *Crinipellis perniciosa* em cacaueiro. Agrotrópica 14: 1-23.

Solomon PS, Waters ODC, Jorgens CI, Lowe RGT, Rechberger J, Trengove RD, Oliver RP (2006) Mannitol is required for asexual sporulation in the wheat pathogen Stagonospora nodorum (glume blotch). Biochem J 399: 231-239.

Stanke M, Waack S (2003) Gene prediction with a hidden-markov model and a new intron submodel. Bioinformatics 19 Suppl. 2: 215-225.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. Annu Rev Phytopathol 46: 75-100.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Thompson JD, Gibson TJ, Plewniak, F, Jeanmougin, F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 24: 4876-4882.

Thomazella DPT, Tiburcio RA, Rincones J, Toni IM, Pereira GAG (2008) Filogenia e Evolução das Oxidase Alternativas de Plantas e Fungos. In: 54º Congresso Brasileiro de Genética, 2008, Salvador. Livro de Resumos do 54º Congresso Brasileiro de Genética.

Tiburcio RA, Costa GGL, Carazzolle MF, Mondego JMC, Schuster SC, Carlson JE, Guiltinan MJ, Bailey BA, Mieczkowski P, Meinhardt •LW, Pereira GAG (2010) Genes Acquired by Horizontal Transfer Are Potentially Involved in the Evolution of Phytopathogenicity in *Moniliophthora perniciosa* and *Moniliophthora roreri*, Two of the Major Pathogens of Cacao . J Mol Evol, 70:85-97.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee MK, McDonald WH, Medina M, Meijer HJ, Nordberg EK, Maclean DJ, Ospina-Giraldo MD, Morris PF, Phuntumart V, Putnam NH, Rash S, Rose JK, Sakihama Y, Salamov AA, Savidor A, Scheuring CF, Smith BM, Sobral BW, Terry A, Torto-Alalibo TA, Win J, Xu Z, Zhang H, Grigoriev IV, Rokhsar DS, Boore JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313: 1261-1266.

Van de Peer Y, Baldauf SL, Doolittle WF, Meyer A (2000) An updated and comprehensive rRNA phylogeny of (crown) eukaryotes based on rate-calibrated evolutionary distances. J Mol Evol 51: 565-576.

Vanlerberghe GC, McIntosh L (1997) Alternative Oxidase. From Gene to Function - Annu. Rev. Plant Physiol. Plant Mol. Biol. 48: 703-734.

Veit S, Worle JM, Nurnberger T, Koch W, Seitz HU (2001) A novel protein elicitor (PaNie) from *Pythium aphanidermatum* induces multiple defense responses in carrot, arabidopsis, and tobacco. Plant Physiology 127: 832-841.

Vélëz H, Glassbrook NJ, Daub ME (2008) Mannitol biosynthesis is required for plant pathogenicity by *Alternaria alternata*. FEMS Microbiol Lett 285(1): 122-129.

Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ (2004) Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. Plant Cell Reports 23: 404-413.

Voegele RT, Hahn M, Lohaus G, Link T, Heiser I, Mendgen K (2005) Possible roles for mannitol and mannitol dehydrogenase in the biotrophic plant pathogen Uromyces fabae. Plant Physiol 137(1): 190-198.

Watling JR, Grant NM, Miller RE, Robinson SA (2008) Mechanisms of thermoregulation in plants. Plant Signal Behav. 3(8): 595-597.

Wenzl P, Wong L, Kwang-Won K, Jefferson RA (2005) A functional screen identifies lateral transfer of beta-glucuronidase (gus) from bacteria to fungi. Mol Biol Evol 22: 308-316.

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18: 691-699.

Youssef NH, Elshahed MS (2009) Diversity rankings among bacterial lineages in soil. ISME J. 3(3): 305-13.

Zaparoli G, Cabrera OG, Medrano FJ, Tiburcio RA, Lacerda G, Pereira GAG (2009) 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. Mycol Res 113: 61-72.

Anexos.

Tabela 1 Sequências usadas na construção da filogenia das NEPs.

Ascomicetos	
Espécie	Número de acesso
Aspergillus fumigatus (1)	709855450
Aspergillus fumigatus (2)	709860792
Aspergillus nidulans (1)	675255061
Aspergillus nidulans (2)	67525754
Fusarium oxysporum	26971315
Gibberella zeae (1)	461232781
Gibberella zeae (2)	461151034
Gibberella zeae (3)	461268180
Magnaporthe grisea (1)	399470906
Magnaporthe grisea (2)	399698446
Magnaporthe grisea (3)	399684781
Neurospora crassa	850937787
Verticillium dahliae	427423729
Basidiomicetos	
Espécie	Número de acesso
Moniliophthora perniciosa (1)	EF109894
Moniliophthora perniciosa (2)	EF109895
Oomicetos	
Espécie	Número de acesso
Phytophthora infestans	898886007
Phytophthora megakarya (1)	599392616
Phytophthora megakarya (2)	599392655
Phytophthora megakarya (3)	599392552
Phytophthora megakarya (4)	599392539
Phytophthora parasítica	133464738
Phytophthora sojae (1)	126986719
Phytophthora sojae (2)	213271049
Phytophthora sojae (3)	213271039
Pythium affinis	371834056
Pythium aphanidermatum	58343786
Pythium middletonii	371834035
Pythium monospermum	371834015
Bactérias	
Espécie	Número de acesso
Bacillus halodurans	47118318
Bacillus licheniformis	561609849
Erwinia carotovora	496094915
Streptomyces coelicolor	244189616

Vibrio pommeresis	14331088

Tabela 2 Sequências usadas na construção da filogenia das cerato-plataninas.

Ascomicetos	
Espécie	Número de acesso
Aspergillus clavatus	GI:121715869
Aspergillus fumigatus	GI:71001827
Aspergillus nidulans	GI:67540205
Aspergillus niger	GI:145231638
Aspergillus terreus	GI:115384119
Botrytis cinerea	GI:5828302
Ceratocystis fimbriata	GI:121624696
Ceratocystis platani	GI:121624715
Ceratocystis populicola	GI:121624693
Chaetomium globosum	GI:116198414
Coccidioides immitis	GI:119193607
Coccidioides posadasii	GI:25528648
Cochliobolus lunatus	GI:37003430
Gibberella pulicaris	GI:56181486
Gibberella zeae	GI:46137392
Gibberella zeae	GI:46139380
Hypocrea atroviridis	GI:92112037
Hypocrea viride	GI:118402815
Hypocrea virens	GI:73672052
Hypocrea virens	GI:93009003
Leptosphaeria maculans	GI:27085248
Magnaporthe grisea	GI:39940863
Neosartorya fischeri	GI:119481376
Neurospora crassa	GI:85091040
Phaeosphaeria nodorum	GI:3329508
Trichoderma atroviride	GI:118402803
Basidiomicetos	
Espécie	Número de acesso
Antrodia camphorata	*GI:47156029*
Coprinus cinereus	*BLAST Local*
Laccaria bicolor	*BLAST Local*
Moniliophthora perniciosa MpCP1	*EU250339*
Moniliophthora perniciosa MpCP2	*EU250341*
Moniliophthora perniciosa MpCP3	*EU250343*
Phanerochaete chrysosporium	*BLAST Local*

Tabela 3 Sequências usadas na construção da filogenia das oxidases alternativas.

Ascomicetos

Espécie	Número de acesso
Ajellomyces capsulatus	154281620
Aspergillus clavatus	121710153
Aspergillus clavatus	121712009
Aspergillus fischer	119480198
Aspergillus fumigatus	70989574
Aspergillus nidulans	67523284
Aspergillus niger	145243919
Aspergillus niger	145245086
Aspergillus terreus	115399175
Aspergillus terreus	115400945
Blumeria graminis	18033102
Botryotinia fuckeliania	154313706
Candida albicans	68464762
Candida albicans	68465139
Candida albicans	68465141
Chaetomium globosum	116196867
Chaetomium globosum	116196989
Coccidioides immitis	119193347
Cryptococcus neoformans	58258418
Debaryomyces hansenii	50418658
Gibberella zeae	46108919
Lodderomyces elongisporus	149239579
Magnaporthe grisea	145602176
Neurospora crassa	85106052
Neurospora crassa	85091905
Pichia guilliermondii	146412116
Pichia stipitis	126133916
Sclerotinia sclerotiorum	156061474
Yarrowia lipolytica	50551826
Yarrowia lipolytica	50550328
Basidiomicetos	
Espécie	Número de acesso
Coprinopsis cinerea	116503357
Coprinopsis cinerea	116506104
Laccaria bicolor	170090172
Moniliophthora perniciosa	
Moniliophthora roreri	
Phanerochaete chrysosporium	82603777
Ustilago maydis	71016789
Monocotiledôneas	
Espécie	Número de acesso
Dracunculus vulgaris	*52421167*

Oryza sativa	*115445792*
Oryza sativa AOX1	*115448118*
Oryza sativa AOX1a	*3218545*
Oryza sativa UNI	*115460317*
Philodendron bipinnatifidum	*52421171*
Saccharum officinarum AOX1	*51860694*
Saccharum officinarum AOX1b	*51860696*
Saccharum officinarum AOX1c	*51860698*
Saccharum officinarum AOX1d	*51860700*
Sauromatum guttatum AOX1	*170164*
Symplocarpus renifolius	*57157823*
Triticum aestivum	*5802556*
Triticum aestivum AOX1a	*19912724*
Triticum aestivum AOX1c	*19912726*
Zea mays AOX1	*162462382*
Zea mays AOX2	*162462511*
Zea mays AOX3	*162463025*
Dicotiledôneas	
Espécie	Número de acesso
Arabidopsis thaliana	18176466
Arabidopsis thaliana	21404169
Arabidopsis thaliana	4006942
Arabidopsis thaliana	166875
Arabidopsis thaliana AOX1a	16323479
Arabidopsis thaliana AOX1b	18403354
Arabidopsis thaliana AOX1c	28393505
Arabidopsis thaliana AOX1d	115311484
Arabidopsis thaliana AOX2	67633915
Catharanthus roseus	2696031
Cucumis sativus AOX2	89474872
Glycine max AOX1	395215
Glycine max AOX2a	2769671
Glycine max AOX2b	1946337
Gossypium hirsutum AOX1	82409048
Lactuca sativa	111607216
Mangifera indica UNI	488825
Nicotiana attenuata	40036984
Nicotiana attenuata AOX1	40036978
Nicotiana tabacum	117646984
Nicotiana tabacum	633599
Nicotiana tabacum AOX1	558053
Populus tremula AOX1	14456073
Populus tremula AOX1b	14456075
Solanum lycopersicum AOX1a	29465621

Solanum lycopersicum AOX1au	46452192
Solanum lycopersicum AOX1b	29465623
Solanum tuberosum	90959750
Vigna unguiculata AOX1	70799008
Vigna unguiculata UNI	17154768
Vigna unguiculata UNI	14572653
Vitis vinifera AOX2	171198300
Gimnospermas	
Espécie	Número de acesso
Picea sitchensis	116784420
Briófitas	
Espécie	Número de acesso
Physcomitrella patens	168044780
Clorófitas	
Espécie	Número de acesso
Chlamydomonas reinhardtii AOX1	159490735
Chlamydomonas reinhardtii AOX2	159472944
Ostreococcus lucimarinus	145357495