

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

Acássia Benjamim Leal Pires

**“Morfogênese de basidiomatas e
identificação de genes relacionados à
frutificação em *Moniliophthora perniciosa*”**

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
ACÁSSIA BENJAMIM
LEAL PIRES
e aprovada pela Comissão Julgadora.

Tese apresentada ao Instituto de
Biologia para obtenção do título de
Doutor em Genética e Biologia Molecular
na área de Genética de
Microorganismos.

Orientador: Dr. Gonçalo Amarante Guimarães Pereira

Co-orientador: Dr. Júlio César de Mattos Cascardo

Campinas 2009

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP**

P665m

Pires, Acássia Benjamim Leal

Morfogênese de basidiomatas e identificação de genes relacionados à frutificação em *Moniliophthora perniciosa* / Acássia Benjamim Leal Pires. – Campinas, SP: [s.n.], 2009.

Orientadores: Gonçalo Amarante Guimarães Pereira, Júlio César de Mattos Cascardo.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Basidiomatas - Desenvolvimento. 3. *Moniliophthora perniciosa*. 4. Genes. 5. Microorganismos - Morfologia. I. Pereira, Gonçalo Amarante Guimarães. II. Cascardo, Júlio César de Mattos. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

(rcdt/ib)

Título em inglês: Morphogenesis of basidiomatas and identification of genes related to fruiting of *Moniliophthora perniciosa*.

Palavras-chave em inglês: Basidiomatas - Development; *Moniliophthora perniciosa*; Genes; Microorganisms - Morphology.

Área de concentração: Genética de Microorganismos.

Titulação: Doutora em Genética e Biologia Molecular.

Banca examinadora: Wanderley Dias da Silveira, Antônio Vargas de Oliveira Figueira, Sérgio Hermínio Brommonschenkel, Jörg Kobarg, Fábio Trindade Maranhão Costa.

Data da defesa: 30/06/2009.

Programa de Pós-Graduação: Genética e Biologia Molecular.

Campinas, 25 de junho de 2009

BANCA EXAMINADORA

Prof(a). Dr . Gonalo A. Guimares Pereira (Orientador)

Assinatura

Profa. Dr Wanderley Dias Silveira

Assinatura

Prof(a). Dr .Marcelo Brocchi

Assinatura

Prof(a). Dr Antnio Vargas de Oliveira Figueira

Assinatura

Prof. Dr Srgio Hermnio Brommonschenkel

Assinatura

Prof. Dr Jrge Konarg

Assinatura

Prof. Dr. Fbio Trindade Maranho Costa

Assinatura

Prof. Dr . Carlos Augusto Colombo

Assinatura

*Dedico**A Deus, pai querido**A Jesus, irmão e mestre**A Neuza Benjamim, mãe e suporte**A Adriana, irmã incentivadora**A Frank, esposo e amigo**e ao nosso querido filho João Vítor*

AGRADECIMENTOS

Agradeço ao meu orientador, Dr. Gonçalo Pereira, pela orientação, pelo incentivo e por ter facilitado a execução dos experimentos na Bahia.

Agradeço ao Dr. Júlio Cascardo, pela co-orientação mais próxima, paciência, amizade e apoio.

Agradeço à CAPES, porque criou o PQI (Programa de Qualificação Institucional), que garantiu a execução deste trabalho.

Agradeço ao colega Dr. Ronan Xavier Correia, porque coordenou muito bem esse programa PQI, na UESC.

Agradeço à Universidade Estadual de Santa Cruz, pelo incentivo à minha capacitação.

Agradeço à FAPESB, pelo apoio financeiro com o auxílio ao projeto de doutorado.

Agradeço ao CNPq, ao PADCT e à FINEP, pelos fomentos recebidos direta e indiretamente.

Agradeço aos colegas Cristina Caribé, Fátima Alvim e Carlos Priminho Pirovani, pelo encorajamento, ajuda mútua, apoio e amizade.

Agradeço aos meus estudantes de graduação, pela paciência e incentivo e aos meus IC Mylene, Ondina, Graziella, Lucianna, Jairo e Ricardo, pelo apoio no Lab e pela amizade.

Agradeço aos colegas dos Laboratórios de Biologia Molecular, Genômica e Proteômica da UESC, especialmente Maíza, Dayana, Dahyanne, Eliana, Cristiano, Braz, Stênio, Robson, Fabiana, Virgínia e Fabienne, pela amizade e incentivo.

Agradeço aos colegas do Laboratório de Genômica e Expressão da Unicamp, pelo apoio e incentivos, especialmente Maricene, Leandra, Ana Deck, Odalys, Johana, Ana Paula, Edu e Marcelo e ao apoio fundamental de Eliane.

Agradeço a Fábio, Fabiana, Ari e Marta, que me hospedaram muitas vezes em Campinas, repartindo seus bens materiais e espirituais.

Agradeço a meus irmãos na fé, especialmente Eduardo Boulhosa e Cris, Edson e Rosângela, Augusto e Lirian, Davi e Meire, Elias e Cali, Diana, Adriana Halla, Samira, Daiane, Marina, Priscila, Gueu, Nazinho, Kika, Eliene, e a todos do grupo, pelas orações e incentivos.

Agradeço à minha mãe, Adriana, Tia Nilde, Tia Neide, Tia Jó, Tio Carlos, Rebeca, Carlson, Jane e Gideão, Flávio, Maica e Tico, Vinícius e Dileuza, Fabrício e Selminha, D. Edna, S. Chico, Candinho, Sueli e Edvaldo, Taline e Tâmara, pelo apoio, orações e amizade.

Agradeço a Joice e Suelene, pelas quatro mãos que me apoiaram quando estive ausente em casa.

Agradeço especialmente a Frank e João Vítor, por compreenderem minhas ausências e me ajudarem nos momentos mais difíceis.

Agradeço a todos os que, embora não citados aqui, me apoiaram para a conclusão deste trabalho.

ÍNDICE

Resumo	viii
<i>Abstract</i>	x
Introdução geral	1
Objetivos	16
Capítulo I	17
Capítulo II	68
Capítulo III	105
Conclusões finais	130
Referências Bibliográficas	134
Anexo I	138
Anexo II	149
Anexo III	153
Anexo IV	155

RESUMO

Moniliophthora perniciosa é um basidiomiceto hemibiotrófico patogênico ao cacauero e outras espécies vegetais. Sua ação na planta consiste em provocar superbrotamentos, epinastia, hiperplasia e, por fim, necrose de tecidos, que passam a ser a matéria prima para a produção de seus esporos, no interior de estruturas chamadas basidiomas. A proposta deste trabalho foi entender como o fungo se prepara para a produção de basidiomas e que genes podem estar relacionados a esse processo. Para isso, foram feitas análises morfológicas do micélio anterior à emergência de basidiomas com o auxílio de estereomicroscopia, microscopia de luz e microscopia de varredura. As modificações apresentadas pelo micélio antes de sua frutificação mostram um comportamento similar ao de outros *Agaricales*. Modificações e aspecto das hifas, e a formação de uma camada superficial a partir de fusão de hifas, foram correlacionadas à emergência de primórdios. As fases de nódulo hifal, agregado, primórdio inicial e primórdio diferenciado foram detectadas. A análise morfológica permitiu, também, inferir aspectos morfogenéticos. Para a análise dos genes envolvidos foi feita uma biblioteca de cDNA representativa de toda a fase de frutificação do fungo. Os 2759 ESTs (Expressed Sequence Tags) foram analisados *in silico* para a busca de genes candidatos. A expressão gênica foi avaliada para alguns ESTs selecionados. Um macroarranjo foi feito com 192 clones e hibridizado com cDNAs obtidos do micélio em momentos anteriores à formação dos basidiomas. Adicionalmente, doze genes candidatos foram analisados por RT-qPCR. A análise *in silico* da biblioteca possibilitou a identificação de 1533 unigenes dos 2759 cromatogramas produzidos. Destes, 981 apresentaram similaridade a genes com funções conhecidas e 1374 foram presentes também no genoma de *M. perniciosa*, indicando que algumas das seqüências sem similaridade nos bancos de dados analisados podem ser genes ainda não descritos em outras espécies. Foram identificados prováveis genes que codificam para hemolisinas,

hidrofobinas, citocromo p450 monoxigenase, ciclofilinas, o-merilstigmato oxigenase, dentre outros, comuns em basidiomicetos durante o desenvolvimento de basidiomatas. A análise da abundância de transcritos mostrou 25 genes que tiveram um pico de expressão no micélio com primórdios e 43 mais abundantes nesta mesma fase e também em basidiomatas maduros. Os doze genes analisados por RT-qPCR mostraram perfis mais detalhados de expressão e indicaram a ativação da produção inicial de basidiomatas ainda na fase de micélio rosa avermelhado, coincidindo com os dados morfológicos. As seqüências da biblioteca foram também utilizadas para a identificação de genes no genoma. Cerca de 14.000 genes foram preditos nesse fungo, sendo que os ESTs gerados neste trabalho contribuíram para essa identificação e também para a determinação do padrão de introns do fungo. A identificação de genes presentes nessa fase do ciclo de vida do *M. pernicioso* abre caminho para estudos genéticos de processos biológicos que levam à frutificação em basidiomicetos. Este é o primeiro estudo morfológico comparando o desenvolvimento inicial de basidiomatas *in vivo* e *in vitro* de *M. pernicioso* e a primeira descrição de genes expressos nessa fase do ciclo de vida fungo.

ABSTRACT

Moniliophthora perniciosa is a hemibiotrophic basidiomycete that infects cocoa and other plant species. The symptoms in the infected plants range from secondary growth, epinasty, hyperplasia and finally necrosis of tissues. The dead tissues are the raw material to produce spores within structures called basidiomatas. The purpose of this study was to understand how the fungus is preparing itself to produce the reproductive structures (basidiomatas) and which genes are related to that process. Morphological analysis of mycelium prior to emergence of basidiomatas was conducted by stereomicroscopy, light microscopy and scanning electron microscopy. To analyze the genes involved in basidiomata development, RNA from previous and fructification phases in artificial medium were used to create a representative cDNA library of the fungus. The 2759 ESTs (expressed sequence tags) obtained were analyzed *in silico* searching for candidate genes. The gene expression was evaluated for some selected ESTs. A macroarray analysis was performed with 192 selected clones and hybridized with RNAs extracted from mycelium in different phases during basidiomata formation. In addition, twelve candidate genes were analyzed by RT-qPCR. The morphological changes observed for the mycelium prior fructification show a pattern similar to other members of the order Agaricales. Changes and appearance of hyphae forming a surface layer from the fusion of hyphae were correlated with the emergence of primordia. The stages of nodule hifal, aggregate, initial primordium and differentiated primordium were detected. Morphological analysis has also inferred morphogenetic aspects. The *in silico* analysis of the library enabled the identification of 1,533 unigenes of 2,759 chromatograms produced. Of these 981 showed similarities to genes with known functions, and 1,374 were also present in the genome of *M. perniciosa*, indicating that some of the sequences without similarity in the databases can be analyzed genes not yet described. We identified candidate

genes that encode for hemolysin, hydrophobins, P450 monooxygenase, cyclophilin the meristigмато-oxygenase, among others, also common in fungi during fruiting. The analysis of the abundance of transcripts showed that 25 genes had a peak of expression in mycelium with primordia with 43 and more abundant in this phase and also in mature basidiomatas. The twelve genes analyzed by RT-qPCR showed more detailed profiles of expression and indicated the activation of the initial production of basidiomatas still in reddish pink mycelium, correlates with the morphological data. The sequences of the library were also used to identify genes in the genome. Approximately 14,000 genes were predicted in this fungus, and the ESTs generated here contributed to this identification and also for determining the pattern of the fungus introns. The identification of genes present in this phase of the life cycle of *M. pernicioso* opens for new possibilities to control the spread of the fungus and also for genetic studies of biological processes that lead to fruiting in basidiomycetes. This is the first morphologic study of comparing both *in vivo* and *in vitro* early development of basidiomatas of *M. pernicioso* and the first description of genes expressed at that stage of the life cycle of this fungus.

INTRODUÇÃO GERAL

A ação saprofítica dos fungos é extremamente benéfica para a vida na terra. Eles são grandes recicladores e têm importância fundamental nos ecossistemas florestais. Além disso, são fontes valiosíssimas de compostos químicos, tais como antibióticos, vitaminas e outras drogas (Alexopoulos, 1996). Entretanto, alguns desses decompositores evoluíram para um estilo de vida ambíguo, passando parte do seu ciclo dentro de hospedeiros e parte decompondo a própria espécie que induziu à morte! Este é o caso de fungos hemibiotróficos, como *Moniliophthora perniciosa* (Sthael) Aime e Philips-Mora. Esse patógeno é o causador da doença Vassoura-de-bruxa, que ocorre em cacaueiros (Purdy e Smidth, 1996). O interessante é que ele coexiste com a planta afetada, sobrevivendo por algum tempo no seu interior nos espaços intercelulares, até que, como consequência dessa interação, induz a morte de alguns ou de todos os ramos da planta infectada. Neste momento, o referido patógeno apresenta-se como saprofítico, funcionando como reciclador dessa matéria morta e, ao mesmo tempo, garantindo a sua sobrevivência com a produção de seus propágulos infectivos, os esporos (Meinhardt et al., 2008).

M. perniciosa pertence ao filo Basidiomycota, que tem como característica principal a produção de esporos sexuais denominados basidiósporos em estruturas microscópicas especializadas chamadas basídias. Os basidiósporos são formados por plasmogamia, cariogamia e meiose, sendo os dois últimos de ocorrência nas basídias. Cada basídia produz tipicamente quatro basidiósporos e estes são haplóides, podendo ser uninucleados ou binucleados entre as espécies (Alexopoulos, 1996). Com relação às estruturas microscópicas, a

presença de grampos de conexão e basidiósporos usualmente do tipo balistosporos são características presentes em Basidiomicetos (Carlile et al., 2001).

O gênero *Moniliophthora*, após a reclassificação filogenética, passou a pertencer à família *Marasmiaceae* (Phillips-Mora, 2005). Essa família pertence à ordem *Agaricales*, caracterizada pela produção de estruturas macroscópicas denominadas basidiomatas que abrigam as basídias (Alexopoulos, 1996). Numa recente reclassificação baseada em análises de seqüências de DNA, o parentesco entre espécies dessa ordem foi re-analisado e a família *Marasmiaceae* ficou agrupada num clado chamado *Marasmioid*, ao qual pertencem os fungos gileados, saprofíticos, produtores de esporos brancos, além dos que possuem forma cifelóide, ressupinata e do tipo clube (Matheny et al., 2006). Membros dessa família possuem estipe, estatura marasmióide (tamanho pequeno), presença de pelos dextrinóides no ‘pileipellis’ (Singer, 1942) e produzem basidiomatas com pigmentos da cor púrpura, violeta ou vermelho que não mudam de cor na presença do KOH (Singer, 1976; 1986).

Os três estágios do micélio, que normalmente ocorrem em basidiomicetos (Alexopoulos, 1996), também estão presentes no biotipo C de *M. perniciosa* (Figura 1), que infecta os cacaueiros (Evans, 1978 e Bastos et al., 1988). O micélio primário desenvolve-se a partir da germinação de esporos em tecidos meristemáticos da planta infectada (Aragundi, 1982; Delgado e Cook, 1976; Evans, 1980). Em cacaueiros, esse tipo de micélio é encontrado nos tecidos com hiperplasia e hipertrofia, denominados ‘vassouras vegetativas’ ou ‘vassouras verdes’, que podem ser terminais, axilares ou de almofada floral. Também é encontrado em todos os outros tecidos infectados pelo fungo na fase inicial da doença (Silva et al., 2002; Silva e Matsuoka, 1999). As hifas desse micélio são monocarióticas, têm espessura entre 5 e 8 um, não possuem grampo de conexão e ocupam os espaços intercelulares (Evans, 1980). Quando esporos são colocados em água para germinar, a transição de micélio primário para secundário

ocorre rapidamente e isso só pode ser retardado se esses esporos forem mantidos em solução contendo glicerol como fonte de carbono e nenhuma fonte de nitrogênio (Meinhardt et al., 2006).

O estágio secundário do micélio ocorre quando as hifas formam grampos de conexão e se tornam dicarióticas (Figura 1). Elas tornam-se mais finas (entre 1 a 3 μm) e invadem as células vegetais. Esse micélio pode ser mantido *in vitro* por muito tempo, desde que a cultura seja renovada periodicamente (Rocha, 1983), mas não representa o fungo no estágio em que induz os sintomas na planta, por esse motivo análises sobre a interação do fungo com a planta devem ser feitas no micélio primário. O micélio de *M. pernicioso* parece ser homotático primário (Griffith e Hedger, 1994), porque foram observados grampos de conexão em culturas monospóricas (Baker e Cowdy, 1943), dicarionização a partir de hifas monocarióticas derivadas de culturas monospóricas (Delgado e Cook, 1976) e produção de basidiomatas a partir de culturas monospóricas (Purdy, 1983).

O estágio terciário do micélio de basidiomicetos é representado por tecidos organizados e especializados chamados basidiomatas (Figura 1), que ocorrem em espécies fúngicas mais complexas (Alexopoulos, 1996). Essas estruturas são multicelulares e seu desenvolvimento requer processos de diferenciação, em alguns momentos semelhantes aos de outros eucariotos mais complexos (Moore, 1998). Entretanto, embora as hifas possuam parede celular, a sua morfogênese segue um padrão diferente das células vegetais, por possuírem modo de crescimento apical e ramificação apenas na direção perpendicular do eixo principal. Assim, para a proliferação celular, que ocorre por ramificação de hifas e por conseqüente aumento de ápices, é necessário um controle da direção de crescimento das hifas (Moore, 1998).

Os basidiomatas originam-se de uma única hifa ou de várias hifas que se ramificam para formar um aglomerado compacto. As estruturas formadas a partir desses aglomerados

mostram diferenciação celular e adesão entre si. Deste modo, são células agrupadas ordenadamente, porém independentes, isto é, não possuem contato lateral como plasmodesmatas, junções conectantes ou outras estruturas semelhantes que ocorrem em plantas e animais (Moore, 1998). Isso sugere que os sinais para a diferenciação devem ser propagados pelo ambiente extracelular, enquanto em plantas ocorre com frequência a comunicação de sinais entre células.

O desenvolvimento de basidiomatas em *Coprinopsis cinereus* ocorre a partir de uma hifa que se ramifica intensamente e forma um pequeno aglomerado chamado ‘nó hifal’ (Kues, 2000). Ao longo do micélio secundário apto para o desenvolvimento de basidiomatas, vários desses ‘nós hifais’ podem ser percebidos. Esses nós progridem para o segundo estágio do desenvolvimento, com a formação de um aglomerado maior de hifas compactas sem forma definida. As hifas desses aglomerados são globosas e infladas. A diferenciação ocorre nesse estágio, e a estrutura é chamada de ‘primórdio bolha’. As hifas crescem em paralelo, em direção perpendicular ao micélio original, e com adesão entre si. Grupos de hifas se diferenciam, originando tecidos que formarão o píleo e a estipe. A meiose, para formação dos basidiósporos também ocorre nesse estágio (Figura 2). A seguir, o primórdio diferenciado sofre apenas alongamento celular para o seu desenvolvimento até basidiomatas maduros (Kües, 2000).

O desenvolvimento dessas estruturas em outros basidiomicetos segue padrões semelhantes. Em *Laccaria* spp., o estágio inicial de basidiomatas é caracterizado pela presença de pequenos anéis de hifas (Figura 3). Hifas adjacentes a esses anéis se ramificam e se tornam com aparência muito irregular, formando posteriormente um aglomerado que contém muitos ápices voltados para o centro da massa micelial. O núcleo central desse aglomerado se

desenvolve em um primórdio que se diferencia a seguir e se torna um basidiomata maduro (Massicote et al., 2005).

Em *Mycena stylobates*, o primeiro estágio detectado do desenvolvimento de basidiomatas foi o de aglomerado, abaixo da superfície. Posteriormente, este aglomerado aumenta e emerge (Whalter et al., 2001). A diferenciação, mostrada em detalhes por esses autores, ocorre após a emergência, quando as hifas das laterais do aglomerado começam a se separar das hifas centrais, por um crescimento desigual e encurvamento para o centro. A partir desse momento, as hifas centrais continuam a diferenciação, formando no ápice o píleo e no centro e abaixo, a estipe (Figura 4).

O desenvolvimento de basidiomata também foi observado em *Agaricus bisporus* (Heckmann et al., 1989). Os autores constataram alterações no micélio que antecedem a frutificação e classificaram-no em 3 estágios. No estágio 1, as hifas inicialmente apresentaram-se em forma de filetes, com pequenas ramificações e a seguir foi observada fusão lateral de hifas. Várias hifas aparentaram-se fundidas, dando um aspecto de placas formadas a partir de hifas vazias e aderidas entre si. Um material mucilaginoso foi observado entre hifas individuais. Ocorreu também a presença de hifas ornamentadas com pequenas projeções de material da própria parede celular. O estágio 2 do micélio foi caracterizado pela consolidação de numerosas hifas em estruturas visíveis a olho nu denominadas ‘cordões’, com aparência de placas em alguns locais ou emaranhado de hifas agrupadas em outros. Muitas hifas desses ‘cordões’ apresentavam-se vazias e a direção delas não era uniforme. O terceiro estágio do micélio foi caracterizado pela presença dos primórdios formados a partir da junção de vários desses ‘cordões’.

A ontogenia de basidiomatas envolve diferenciação celular, porém de modo distinto de animais e plantas. A visão superficial da formação dos basidiomatas lembra o desenvolvimento

de brotos florais, com a formação de camadas que se sobrepõem a um núcleo central e o encurvamento de células em determinadas direções. Também lembra o desenvolvimento em animais, com a utilização de sinais propagados entre as células. Entretanto, possui eventos bem distintos que devem ser específicos de fungos. Assim, os genes envolvidos na indução e formação dessas estruturas são interessantes, pois processos semelhantes a plantas ou animais podem estar ocorrendo e comparações podem ser realizadas para aumentar a compreensão de tais eventos. Por outro lado, processos bem diferentes podem estar ocorrendo e isso ajudará a explicar o estilo de vida desses fungos.

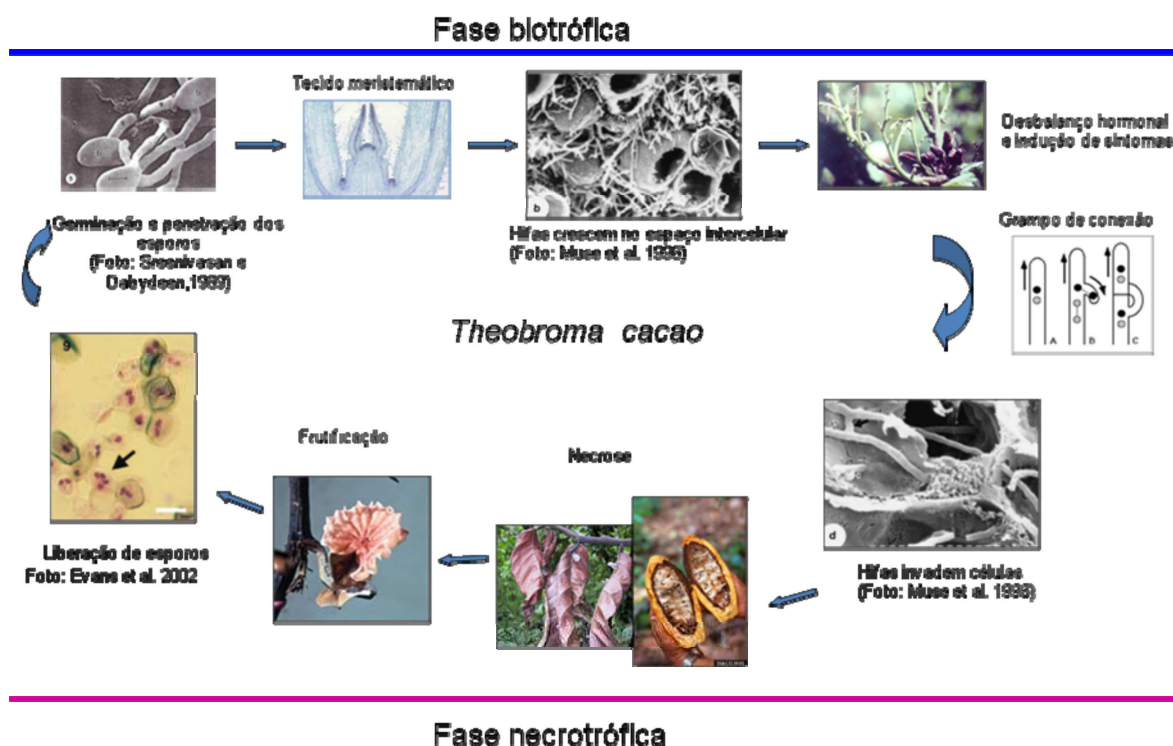


Figura 1. Ciclo de vida do fungo *Moniliophthora perniciosa* biótipo C. Plantas de cacau (*Theobroma cacao*) são infectadas a partir de meristemas. O fungo cresce inicialmente nos espaços intercelulares e posteriormente muda de fase tornando-se dicariótico pela formação de grampes de conexão. Os tecidos necrosados servem de substrato para o desenvolvimento dos basidiomatas onde são formados os basidiósporos. Estes, por sua vez, dão início a um novo ciclo ao germinarem e penetrarem em tecidos meristemáticos.

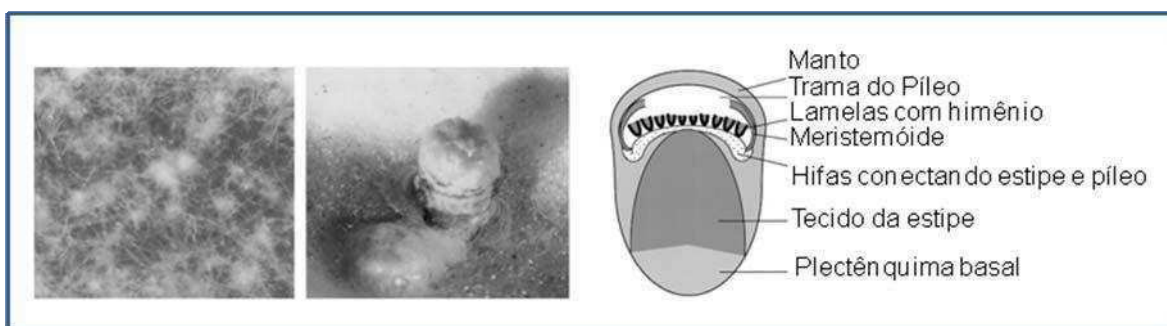


Figura 2. Desenvolvimento inicial de basidiomatas de *C. cinerea*. Da esquerda para a direita: nós hifais, primórdio e representação esquemática de um corte tangencial em um primórdio, mostrando a organização dos tecidos.

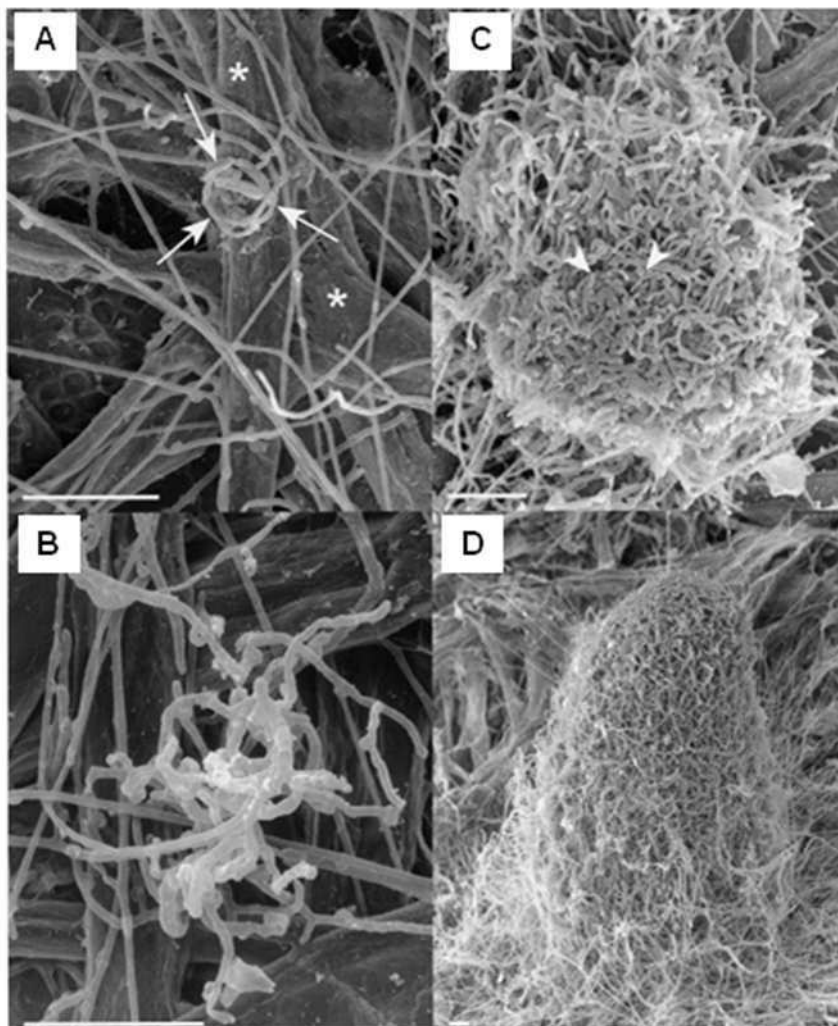


Figura 3. Início do desenvolvimento de basidiomatas em *Laccaria* sp. Inicialmente as hifas assumem a forma anelar (A), seguida de ramificações em redor (B). Em sequência, forma-se um aglomerado compacto (C) com hifas concêntricas (setas) que depois crescem verticalmente a partir do centro (D). Esta figura faz parte do artigo de autoria de Massicote e colaboradores (2005).

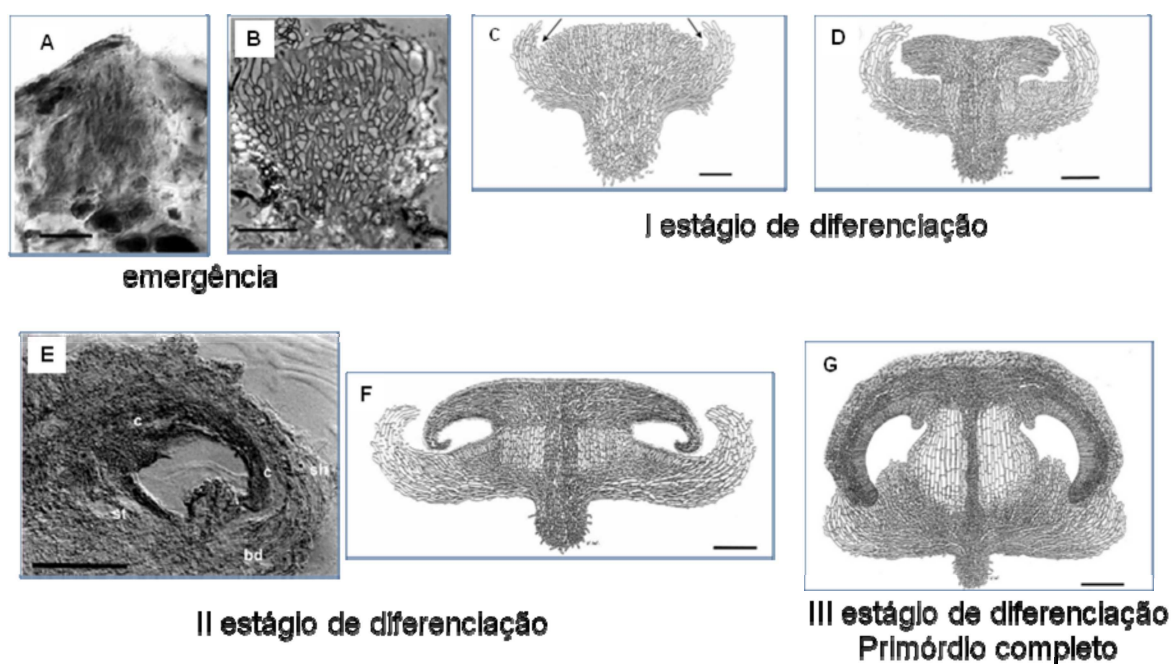


Figura 4. Emergência e estágios de diferenciação de basidiomatas em *M. stylobates*. Inicialmente hifas se aglomeram abaixo da superfície do meio (A) e o aglomerado emerge contendo hifas que crescem em direção ao ápice central (B). A seguir (C-G) podem ser observados estágios de diferenciação do primórdio que culminam na formação da estipe e do píleo. Estas fotos foram descritas no artigo de autoria de Walter e colaboradores (2001).

O primeiro evento genético determinante para a frutificação em basidiomicetos é a dicariorização das hifas e conseqüente transformação do micélio primário em secundário (Kües, 2000). Esse evento é regulado por genes de cruzamento ('mating-types'), que são responsivos a fatores ambientais (Kües, 2000; Bahn et al., 2007). Basidiomicetos homotáticos, como *M. pernicioso*, não possuem hifas diferentes e por isso o cruzamento entre hifas não precede a frutificação. Entretanto, foram detectados genes homólogos aos de cruzamento em outros fungos. Neste caso, é possível que eles atuem na formação dos grampões de conexão para a formação dos dicárions e, de fato, esse evento ocorre antes da frutificação (Mondego et al., 2008).

Uma vez em estágio saprofítico, o micélio frutifica em resposta a fatores ambientais (Kues e Liu, 2000). Em *C. cinerea*, diversos fatores ambientais são responsáveis pela formação de basidiomatas, entre eles luz, concentração de nutrientes do meio, temperatura e umidade (Kües, 2000). Os basidiomatas de *M. pernicioso* desenvolvem-se em tecidos necrosados da planta infectada ricos em micélio secundário. Em condições naturais, são produzidos após dias chuvosos com temperaturas amenas (Evans, 1981). Em tecidos necrosados, como as conhecidas vassouras secas, são produzidos sob períodos alternados de umidade (8 h por dia) e seca (16 horas por dia), sob temperatura de 25 °C e fotoperíodo de 12 horas (Rocha e Wheeler, 1982). Os esporos são unicelulares, hialinos, com parede fina, medindo entre 12 µm e 17 µm e não resistem ao ressecamento e à luz ultravioleta (Baker e Crowdy, 1943). Porém, a liberação dos esporos ocorre normalmente nas madrugadas, sob condições de umidade relativa elevada e temperaturas amenas.

O ciclo de vida do *M. pernicioso* inicia e termina com a produção dos basidiósporos (Figura 1). Milhões são produzidos em cada basidiomata e a sua dispersão é feita pelo vento. Um basidiomata pode permanecer ativo por cinco dias em média e aqueles produzidos na parte

aérea da planta (vassouras secas) são os mais eficientes na propagação da doença (Silva et al., 2002). Também ocorre produção de basidiomatas em outras partes da planta, porém esse fungo é susceptível ao ataque de insetos, de bactérias e outros fungos como o *Trichoderma stromaticum*, e isso ocorre com mais frequência próximo ao solo (Samuels et al., 2000).

Uma das possibilidades do controle da doença no campo seria a eliminação dessas fontes de tecidos aptos à produção de basidiomatas. Na prática, essa alternativa é muito dispendiosa porque envolve uma grande quantidade de mão de obra e não tem a eficiência máxima. O fungo *T. stromaticum* está sendo utilizado como controle biológico (De Souza, et al., 2008), porém essa alternativa ainda é questionável (Castro et al., 2001).

O conhecimento mais aprofundado de como essas estruturas reprodutivas são geradas e quais genes estão envolvidos pode abrir caminho para outras formas de controle da produção de esporos por esse fungo. Isso poderia ser feito pela identificação de fungicidas específicos, ou simplesmente pela alteração da forma de condução da cultura, para inibir determinadas rotas essenciais à produção de basidiomatas. A primeira hipótese, entretanto, é inviável, pois a eficiência da aplicação de fungicidas para fins de controle da produção de basidiomatas é baixa, porém, a segunda hipótese pode ser viável. Por outro lado, o conhecimento sobre a regulação genética que determina a produção de basidiomatas, do ponto de vista científico, pode servir para análises genéticas envolvendo mutantes, para estudos evolutivos, para estudos bioquímicos de compostos secundários e para identificação de novas enzimas ou produtos gênicos de interesses diversos, inclusive biotecnológicos.

Há alguns genes de outros fungos já descritos que são expressos especificamente em basidiomatas ou em fases iniciais de seu desenvolvimento. Foram observados aumento de expressão de genes que codificam para citocromo p450 *CYP4*, septina, ATP sintase subunidade alfa, hidrofobinas, lacases, fator de transcrição MFBA, lecitina ABL, galecninas,

hemolisinas e um fator de transcrição PriA (Kües e Liu, 2000). Uma série de levantamentos de genes expressos na fase de frutificação de basidiomicetos tem sido feita (Ospina-Giraldo, et al., 2000; Lacourt et al., 2002; Lee et al., 2002; Yamada et al., 2006). Nesses levantamentos, além dos genes citados acima, foram detectados genes relacionados a choques térmicos, metabolismo de glicose, proteólise, degradação de parede celular, metabolismo secundário e sinalização mediada pelas rotas de AMPc e AMPK.

Dois genes que codificam para galectinas em *C. cinerea*, Cgl1 e Cgl2, são expressos diferencialmente. A expressão de Cgl2 se inicia em estágios iniciais da formação de basidiomatas, correlaciona com a formação dos nódulos hifais e é reprimida pela luz. O gene Cgl 1 é expresso nos agregados iniciais e é induzida pela luz (Boulianne et al., 2000). Também nesse fungo, um gene que codifica para uma sintase de ácidos graxos ciclopropano é essencial para a produção de basidiomatas e é expresso nos estágios tardios do desenvolvimento de basidiomatas (Liu et al., 2006).

Em *Lentinus edodes*, a frutificação é associada a altos níveis de AMPc e aumento na atividade de adenilato ciclase (Takagi et al., 1988) e à presença de uma proteína de adesão celular com alto peso molecular contendo o motivo Arg-Gly-Asp, a *MFBAC* (Kondoh et al., 1995). Também em *L. edodes*, um gene que codifica para um provável proto-oncogene da família Myb, Le.CDC5, é expresso a partir dos estágios iniciais da formação de basidiomatas e o seu produto é detectável até os estágios de maturação (Nakazawa et al., 2006). Os transcritos de um gene que codifica para outra proteína de adesão, da família das fascilinas, foram especificamente elevados na presença de primórdios e basidiomatas maduros (Miyazaki et al., 2007). Outro gene foi descrito recentemente nesse fungo, codificando para um fotorreceptor de luz azul PHRA, homólogo a gene que codifica para a proteína ‘white-collar2’. Ele é expresso durante a frutificação, é regulado por luz e regula a síntese de tirosina (Sano et al., 2009).

Em *Pleurotus ostreatus* e *Agrocybe aegerita* foram detectadas hemolisinas associadas à fase de primórdios (Berne et al., 2002; Berne et al., 2007). Essas proteínas, chamadas ostreolisinas e aegerolisinas, foram purificadas e estão sendo caracterizadas. Elas possuem função hemolítica em células de mamíferos (Žužek et al., 2006), têm a capacidade de formar poros (Berne et al., 2005), ligam-se especificamente a lipídios de membrana, especialmente colesterol (Sepčić et al., 2003; 2004) e são moduladas por lisofosfolipídios (Chowdhury et al., 2008). A função exata dessas proteínas na formação de basidiomatas ainda não foi determinada (Berne et al., 2009).

Hidrofobinas fazem parte de uma família de proteínas secretadas também relacionadas à frutificação. Elas formam um filme anfipático na superfície, na interface ar-água. Biologicamente funcionam para adesão e revestimento de hifas em fungos filamentosos e são relacionadas ao crescimento celular, desenvolvimento, infecção e virulência (Wösten e de Vocht, 2000). Há hidrofobinas expressas especificamente durante a formação de basidiomatas, enquanto outras não se expressam nesse período (Albuquerque et al., 2004; Santos, 2005; Ma et al., 2007).

Esse é um estudo pioneiro para a identificação de genes expressos em *M. perniciosus* na fase de transição do micélio secundário para o terciário. Em adição a isso, foi necessário entender como as estruturas se formam, pois ainda não existiam relatos sobre o desenvolvimento inicial de basidiomatas em *M. perniciosus*. A inexistência de uma coleção de mutantes desse fungo não possibilitou análises genéticas diretas, porém foi possível a construção de uma biblioteca de seqüências curtas expressas (ESTs) gerada a partir de amostras do fungo cultivado *in vitro*, sob condições de indução da frutificação. As seqüências dessa biblioteca foram muito úteis para a validação do projeto Genoma *M. perniciosus*, pois representam transcritos que serviram para a identificação de genes e para os trabalhos de

anotação do genoma. A análise da expressão de transcritos selecionados serviu para confirmar algumas hipóteses sobre a fase do desenvolvimento inicial dos basidiomatas e abriu caminho para estudos posteriores que venham a elucidar as rotas bioquímicas relacionadas a esse evento.

OBJETIVOS

Objetivo Geral

Identificar genes de *Moniliophthora perniciosa* responsáveis pela sua frutificação.

Objetivos Específicos

- I. Identificar genes expressos em micélio cultivado em condições propícias à frutificação;
- II. Correlacionar mudanças morfológicas no micélio que resultam na formação dos basidiomatas com os níveis de transcritos obtidos;
- III. Identificar genes presentes no genoma de *M. perniciosa* correspondentes aos transcritos caracterizados.

Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes

Acássia BL Pires^a, Karina P Gramacho^b, Delmira da Costa Silva^a, Aristóteles Góes-Neto^c, Mylene de Melo Silva^a, Jairo S Muniz-Sobrinho^a, Ricardo F Porto^a, Cristiano Villela-Dias^a, Martin Brendel^a, Júlio CM Cascardo^{a*}, Gonçalo AG Pereira^d

BMC Microbiology 2009, **9**:158 doi:10.1186/1471-2180-9-158

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes

BMC Microbiology 2009, **9**:158 doi:10.1186/1471-2180-9-158

Acassia BL Pires (ableal@uesc.br)
 Karina P Gramacho (karina@cepec.gov.br)
 Delmira C Silva (demilra@uesc.br)
 Aristoteles Goes-Neto (agoesnt@uefs.br)
 Mylene M Silva (mylmelo@yahoo.com.br)
 Jairo S Muniz-Sobrinho (jairo_muniz@hotmail.com)
 Ricardo F Porto (ricfporto@hotmail.com)
 Cristiano Villela-Dias (c.villeladias@gmail.com)
 Martin Brendel (martinbrendel@yahoo.com.br)
 Julio CM Cascardo (cascardo@uesc.br)
 Goncalo GA Pereira (goncalo@unicamp.br)

ISSN 1471-2180

Article type Research article

Submission date 21 November 2008

Acceptance date 4 August 2009

Publication date 4 August 2009

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes

Acássia BL Pires^a, Karina P Gramacho^b, Delmira C Silva^a, Aristóteles Góes-Neto^c, Mylene M Silva^a, Jairo S Muniz-Sobrinho^a, Ricardo F Porto^a, Cristiano Villela-Dias^a, Martin Brendel^a, Júlio CM Cascardo^{a*}, Gonçalo AG Pereira^d

^a Centro de Biotecnologia e Genética, Laboratório de Genômica e Expressão Gênica, Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Rodovia Ilhéus-Itabuna, km 16, 45662-000, Ilhéus-Bahia, Brazil. Phone: + 55 73 3680 5010; Fax: 55 73 3680 5258

^b Centro de Pesquisas do Cacau, CEPLAC, Rodovia Ilhéus-Itabuna, km 22, Ilhéus-Bahia, Brazil.

^c Laboratório de Pesquisa em Microbiologia (LAPEM), Universidade Estadual de Feira de Santana, Av. Transnordestina, s/n, bairro Novo Horizonte, 44036-900, Feira de Santana, Bahia, Brazil.

^d Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, Campinas, São Paulo, Brazil.

ABLP - ableal@uesc.br; KPG - karina@cepec.gov.br; DCS - delmira@uesc.br; AGN arigoesneto@gmail.com; MMS mylmelo@yahoo.com.br; JSMS jairo_muniz@hotmail.com; RFP - ricfporto@gmail.com; CVD - c.villeladias@gmail.com; MB - martinbrendel@yahoo.com.br; JCMC - cascardo@uesc.br; GAGP goncalo@lge.ibi.unicamp.br

*

Corresponding author

Abstract

Background: The hemibiotrophic fungus *Moniliophthora perniciosa* is the causal agent of Witches' broom, a disease of *Theobroma cacao*. The pathogen life cycle ends with the production of basidiocarps in dead tissues of the infected host. This structure generates millions of basidiospores that reinfect young tissues of the same or other plants. A deeper understanding of the mechanisms underlying the sexual phase of this fungus may help develop chemical, biological or genetic strategies to control the disease.

Results: Mycelium was morphologically analyzed prior to emergence of basidiomata by stereomicroscopy, light microscopy and scanning electron microscopy. The morphological changes in the mycelium before fructification show a pattern similar to other members of the order *Agaricales*. Changes and appearance of hyphae forming a surface layer by fusion were correlated with primordia emergence. The stages of hyphal nodules, aggregation, initial primordium and differentiated primordium were detected. The morphological analysis also allowed conclusions on morphogenetic aspects. To analyze the genes involved in basidiomata development, the expression of some selected EST genes from a non-normalized cDNA library, representative of the fruiting stage of *M. perniciosa*, was evaluated. A macroarray analysis was performed with 192 selected clones and hybridized with two distinct RNA pools extracted from mycelium in different phases of basidiomata formation. This analysis showed two groups of up and down-regulated genes in primordial phases of mycelia. Hydrophobin coding, glucose transporter, Rho-GEF, Rheb, extension precursor and cytochrome p450 monooxygenase genes were grouped among the up-regulated. In the down-regulated group relevant genes clustered coding calmodulin, lanosterol 14 alpha demethylase and PIM1. In addition, 12 genes with more detailed expression profiles were analyzed by RT-qPCR. One aegerolysin gene had a peak of expression in mycelium with primordia and a second in basidiomata, confirming their

distinctiveness. The number of transcripts of the gene for plerototolysin B increased in reddish-pink mycelium and indicated an activation of the initial basidiomata production even at this culturing stage. Expression of the glucose transporter gene increased in mycelium after the stress, coinciding with a decrease of adenylate cyclase gene transcription. This indicated that nutrient uptake can be an important signal to trigger fruiting in this fungus.

Conclusions: The identification of genes with increased expression in this phase of the life cycle of *M. pernicioso* opens up new possibilities of controlling fungus spread as well as of genetic studies of biological processes that lead to basidiomycete fruiting. This is the first comparative morphologic study of the early development both *in vivo* and *in vitro* of *M. pernicioso* basidiomata and the first description of genes expressed at this stage of the fungal life cycle.

Background

Moniliophthora pernicioso (Stahel) Aime and Phillip-Mora (2005) [1] is a hemibiotrophic basidiomycete that causes Witches' Broom Disease (WBD) in cocoa (*Theobroma cacao* L). Currently, WBD occurs in South and Central America and can cause crop losses of up to 90% [2]. In Bahia (Brazil), *M. pernicioso* was identified in 1989 [3] and, as a consequence of its spreading, the annual production of cocoa beans dropped from 450,000 to 90,000 tons within 12 years, reducing export values from an all-time high of about US\$ 1 billion to 110 million. During this period nearly 200,000 rural workers lost their jobs, resulting in an intensive migration from farms to urban areas [4].

The fungus infects young meristematic tissues inducing hypertrophy and hyperplasia, loss of apical dominance, and proliferation of axillary shoots. The hypertrophic growth of the -3

infected vegetative meristems (green broom) is the most characteristic symptom of WBD [5].

Basidiomata, in which basidiospores are produced, develop on dead but attached dry brooms of cacao trees in the field, after dry and wet periods. Basidiospores are spread by wind and depend on sufficient moisture for survival. They can only germinate on and infect susceptible cacao tissues (i.e. buds, young leaves, flower cushions, or young pods) if relative humidity levels are near 100%. Shortly after infection the pathogen establishes a biotrophic relationship with the host during which the fungus has an intercellular, biotrophic, monokaryotic growth phase, without clamp connections. Four to six weeks later, the hyphae become dikaryotic, develop clamp connections and the fungus grows saprophytically [5]. A well-characterized concerted series of cell death events [6] causes the green broom to become necrotic, and basidiomata are formed in a favorable environment after 6 weeks or more [7].

Information about morphological development and environment that affect basidiomata and basidiospore production of *M. pernicioso* are important to improve the *in vitro* culture of the pathogen and to study its life cycle. Environmental conditions for basidiomata production have been described by Suarez [8], Rocha [9] and Rocha and Wheeler [10, 11]. An artificial production of basidiomata has been studied by several authors, but an ideal production mode has not yet been achieved. Stahel [12] observed basidiomata development on mycelial mats in agar cultures. Purdy *et al.* [13] and Purdy and Dickstein [14] modified Stahel's methods to produce basidiomata on mycelial mats. Griffith and Hedger [7] improved basidiomata production by using bran-vermiculite medium, a method currently used to produce *M. pernicioso* basidiospores. Later, Niella *et al.* [15] modified medium formulation and Macagnan *et al.* [16] removed vermiculite and the extra layer of cacao powder and CaSO₄ originally used to cover the medium and to reduce the time to fruiting. The difficulty of obtaining axenic cultures and the long cultivation time has hindered more detailed studies on the morphology

and early development of *M. perniciosus* basidiomata.

Several studies of basidiomata development in other basidiomycetes, e.g., *Agaricus bisporus*, *Flammulina velutipes*, *Boletus edulis* [17] as well as mycorrhizal fungi such as *Laccaria* sp. [18] have already been published, complementing research on *Coprinopsis cinerea* and *Schizophyllum commune*, which are models for developmental studies in macroscopic basidiomycota [19].

Basidiomata of *M. perniciosus* produced either in nature [20, 21, 22] or under laboratory conditions [13, 7, 14] have been studied and their morphology was originally described by Stahel [12]. Later, Delgado and Cook [23] showed that the hyphae found in basidiomata are dikaryotic whereas basidia are monokaryotic (i.e. diploid, following karyogamy).

Although the microscopic characteristics and growth patterns of both monokaryotic and dikaryotic mycelia have been described elsewhere [24, 25, 26], there is no microscopic characterization of the pattern of basidiomata development. We provide the first description of primordium development of *M. perniciosus* basidiomata. Based on our observations the development was divided in four stages, similar to those described for *A. bisporus* (17). Together with the sequencing and annotation of the *M. perniciosus* genome [27], detailed morphologic information is important for future research into *M. perniciosus* mutants, complementing genetic studies. Here we describe and histologically compare the development of both *in vivo* and *in vitro*-grown *M. perniciosus* basidiomata and analyze the expression of 192 selected ESTs by macroarray and of 12 ESTs by RT-qPCR.

Results and Discussion

Morphological observations

Observations of dead brooms kept in humid chambers or collected directly from the field showed the presence of a thin mat of saprophytic mycelium on the surface of the

brooms. It was possible to notice color changes and the morphology that preceded basidiomata formation on this mat. The aerial mycelium formed a thick layer with notable color modifications: it was initially white (Figure 1A), then yellow (Figure 1B) and later, reddish pink (Figure 1C). At a later stage, dark-brown to reddish spots appeared until onset of primordium growth (Figure 1E and 1F). The same characteristics were observed in artificial cultivation (Figure 1D), which allowed a monitoring of the morphogenetic stages of *M. pernicioso* basidiomata.

Currently two media are used to produce basidiomata of *M. pernicioso*. The “Griffith medium” [7] contains pieces of bran/vermiculite covered with a casing layer of peat/gypsum, while the “Macagnan medium” [16] contains dry broom material. When plugs of dikaryotic mycelia are transferred from agar culture to either of these two solid media and incubated at 25 °C in Petri dishes, a network of hyphae initiates growth within and on the surface of the solid particles. Once the medium is well-colonized (similar to spawn-running in mushroom cultivation), basidiomata production is induced by opening the dishes, suspending the block of substrate (Figure 1D), and subjecting it to a regime of intermittent watering and a daily photoperiod of 10-12 h light.

When cultured in the “Griffith medium”, mycelial mats of *M. pernicioso* isolate CP03 (see Methods) turned light-yellow four days after exposure to air and water, changing to reddish-pink after a further ten days, finally becoming dark-reddish pink until the onset of basidiomata development, some two to eight weeks later. These color changes were not uniform among parts of mycelial mats, varying according to irrigation intensity. The whitish aerial mycelium remained visible until the end of cultivation on some parts of the mycelial mats. Color changes also occurred in long-term stored mycelia at 25 °C, however, basidiomata formation was never observed. Since mycelium color change was a pre-requisite for primordium formation, we standardized the collections according to their color.

In an examination of the mycelial mats during the 32-day incubation period in Petri dishes, prior to incubation in the wetting/drying chambers, branched and agglomerated hyphae (mycelial cords) were observed fanning out on the surface of the substrate, appearing as long strands (Figure 2A, yellow arrow), with probable hyphal fusion along part of their length (Figure 2A, white arrow). At some points, hyphae were covered in a thin amorphous layer, apparently composed of plant cell wall material (Figure 2A, red arrow), as well as irregularly swollen and ornamented cells (Figure 2A, pink arrow). After exposure to water and air in the wetting/drying chamber, there appeared to be further agglomeration of hyphae into thicker structures, often covered with a layer of amorphous material (Figure 2B) and some raised areas with curved hyphae were also observed (Figure 2C). These changes were concurrent with the formation of yellow, reddish pink and dark-reddish pink pigmentation on the mat surfaces. In contrast, the mycelium on dry brooms already formed a dense layer at the white stage, probably due to the fact that this layer is formed in response to regular irrigation to which the brooms were subjected from the beginning of the experiment (Figure 1A and 1C).

Curved hyphae, leading to a possible hyphal fusion, were observed at this moment and in all distinct stages of the superficial mycelium, a pattern also observed in *Laccaria* spp [18]. Side-by-side hyphal branches evolved to larger plate-like structures in reddish pink mycelium (Figure 2B) and in mycelium forming the primordia apex (Figure 2D). These plate structures were not always continuous and some mycelial strands appeared empty or dry (not shown). A microscopic tissue section of reddish-pink mycelium in air contact revealed a distinctive mycelium layer with a mean thickness of 60 μm (Figure 2E, arrow), as well as internal net patterns of hyphae.

Similar patterns of hyphal growth were reported by Heckman et al. [28] in *A. bisporus* before basidiomata formation [28]. These authors recognized four morphological stages of

mycelium and observed side-by-side hyphal fusions and the formation of hyphal wall ornamentation, which occurred in the first mycelial growth phase [28]. In the second stage, hyphal fusion led to the formation of structures called strands. Microscopic primordia were formed in the third stage in more compact masses, in areas of dense mycelial growth. At the fourth stage, primordia were visible to the unaided eye. Fused and ornamented hyphae as well as strands appeared in *M. pernicioso* before primordium development. Therefore, the process of primordium development of *M. pernicioso* was similar to that observed for *A. bisporus*, except for the formation of an impermeable surface layer in hyphae and the type of hyphal ornamentation only observable in *M. pernicioso*.

The chemical composition of the impermeable surface layer was investigated. No reduced sugars, lipids and phenols were detected (data not shown). If these layers consisted of empty fused hyphae, chitinases were possibly active in this event. Lopes [29] observed an increased expression of chitinases in *M. pernicioso* in the reddish pink mycelium prior to basidiomata formation. It may also be possible that these areas are rich in hydrophobins, a protein required in basidiomata formation in several other fungi that form a thin outer layer on hyphae exposed to the air [30]. These proteins form an amphipathic layer between hydrophilic-hydrophobic interfaces, which protects the hyphae-inducing aerial mycelia [31]. An increased expression of hydrophobin-encoding genes was observed during mycelial mat growth of *M. pernicioso* [32].

Changes in pigmentation of the superficial mycelium of *M. pernicioso* were described by Purdy *et al.* [13] and by Griffith and Hedger [7]. In our experiments, changes in pigmentation were observed in mycelial mats washed in chambers until basidiomata emergence, indicating a correlation with basidiomata formation. The same color of the surface mycelium persists in the primordia, especially in the apices. The appearance of hyphal nodules coincided with the change in pigmentation from yellow to pink of the surface mycelium as described before (Figure 2F), and the primordia

emerged after this color had darkened. Stronger pigmentation was observed on the primordia apex exactly at points of densely aggregated hyphae, which leads us to believe that pigmentation is correlated with hyphal aggregation. The term “hyphal nodules” has been used to describe the initial phases of basidiomata development [19] as well as for the nodules in the regions of the “initials” and in the morphogenesis-directing primordia [33].

Primordia of *M. pernicioso* appeared when the dense mycelial mat showed reddish-pink pigmentation. The first signal of primordial development was probably the appearance of primary hyphal nodules as well as internal local aggregations on dark pink–reddish mycelium (Figure 2F). Thereafter, hyphal interaction led to the formation of compact aggregates that can be considered an undifferentiated stage called initial primordium or secondary hyphal nodule [19] (Figure 3A). Hyphae belonging to such aggregates were short, large and strongly stainable with fuchsin acid, a substance present in Pianezze III solution, used to distinguish fungal from plant tissues (Figure 3A). The primordium emerged from within the surface mycelial layer (Figure 1E) as a well-defined protuberance (Figure 1F) with hyphae similar to those found in the aggregates (Figure 4A). The primordium initial (Figure 1F and Figure 3C) then underwent differentiation to form stipe, pileus (Figure 4B) and lamellae (Figure 4C). Hyphae of the primordium apex were cylindrical, with round apices and parallel growth, bending at the end distal to the pileus (Figure 4D, detail). Stipe hyphae were more compact, flat, growing vertically (Figure 4E). Amorphous material and clamped hyphae were also present on the apical primordium surface (Figure 2D and Figure 4F, respectively).

The various developmental stages of *M. pernicioso* basidiomata formation were very similar to those previously described in detail for *Agaricus* sp. [17], *C. cinerea* [19], *Mycena stylobates* [34] and *Laccaria* spp. [18]. Differentiation in *Agaricus* occurred at the initial stage to produce a bipolar fruiting body primordium [17, 19]. This process appears to be conserved among *Agaricales* with

slight differences between species. It was rather difficult to microscopically observe the hyphal nodule of the mycelial mats grown on “Griffith medium” due to the density of the hyphal layer. However, the primary hyphal nodule stages of *M. perniciosus* basidiomata were inferred from the presence of areas of intense localized ramifying hyphal aggregates in small nodules (Figure 2F). These nodules progressed to a globose aggregate, surrounded by a dense layer of amorphous material, an irregular arrangement of interwoven hyphae on the internal tissue of dry brooms stained green (Figure 3A), which can be considered the initial stage of hyphal aggregation. This hyphal agglomerate is distinguished by acid fuchsin which stains only living tissues [35]. Aggregates found in dark reddish-pink mycelium (Figure 2F) indicated a competent mycelium from which primordia may originate, similar to the aggregates in *Laccaria* sp., which would give rise to basidiomata [18].

Globose aggregates appeared on the surface with a protective layer covering a hyphal bulb (Figure 1E, *). Walther *et al.* [34] described a similar phenomenon in the initial development of *M. stylobates*. The initial formation of this layer can be observed in *M. perniciosus* (Figure 3A, arrow) that later covered the surface of the protuberant area (Figure 1E, *). Then, an initial emerged (Figure 1F and Figure 3C) and differentiated into a primordium, here referred to as the third stage (Figure 3E). It is likely that enzymatic digestion by chitinases [36] occurred in the hyphae of the outer layer, thereby allowing the “initial” to emerge as a dense layer, with amorphous material in the center of the protuberance. Differentiation continued leading to the formation of the lamellae (Figure 3E, arrow and Figure 4C) and later the pileus (Figure 4B). The apical region of initials formed the pileus and the basal region formed the stipe (Figure 4B). At the end of this stage the immature pileus and stipe (Figure 4G) could be seen with lamellae already established (not shown). Lamellae expanded after two to three days (Figure 4H), depending on sufficiently high moisture levels, as already observed for other basidiomycetes [17]. The hymenium was enclosed by incurved margins of the pileus, only

being exposed when the basidiomata matured (Figure 4G and 4H). Finally the stipe elongated and the pileus expanded to expose the hymenium for basidiospore liberation (Figure 4I). Basidiomata maturation was regulated by humidity and not all initial primordia progressed to form basidiomata (not shown).

Primordia emerged from 75 d after the exposure of substrate-grown mycelia to water and light in the humid chamber (Figure 1G). The first basidiomata were observed about 10 d after the first primordium was visible, but undifferentiated primordia were still present on the mat surface when basidiomata appeared. Density of primordia was high, their size not uniform and their production discontinuous, suggesting a programmed induction, as in plant inflorescences. The morphogenesis observed in the initials (Figure 3) resembled that of other Basidiomycota. Hyphae aggregated towards the surface and assumed a vertical position concurrent with an increase in diameter and compartment length (distance between septa) (Figure 3A and Figure 4A, arrow). These hyphae differentiated to form an agglomerate (Figure 3A) where they converged in an apical group (Figure 3B, #) and two lateral groups, growing in towards the bottom (Figure 3B, black square). A parallel bundle of hyphae with an inclination in direction to the center of the agglomerate was also observed (Figure 3B, *). This bundle diminished in length when the central aggregates increased in size; later, a lateral appendix to the primordium was observed (Figure 3D, arrows and *). Lateral groups (Figure 3D, #) increased in prominence during development, and the convergent hyphae at the agglomerate apex became vertically prominent (Figure 3D, black squares).

The lateral groups tended to bend downwards away from the apex (Figure 3C, *). A group of basal hyphae, however, bent upwards, supporting the hyphal extremity that bent downwards (Figure 3C, arrow and 3D, arrow). As the lateral hyphae expanded, the overlapping of these hyphae diminished (Figure 3E, * and 3F, arrows), increasing the space between these hyphal

groups (Figure 3E, arrow). A micrograph of an emerged primordium (Figure 4C) shows a difference in opacity between hyphae, suggesting that a partial digestion led to the spaces between the lamellae. Another freehand section shows the lateral bending of hyphae and the differentiation of the stipe (Figure 4B). This primordium already possessed a differentiated hymenium (not shown).

Studies in *Agaricus* sp. and other edible fungi revealed a hemi-angiocarpous standard developmental stage [17, 19], with a veil covering the primordium. In these fungi, a cluster of parallel and oriented hyphae emerges and forms the stipe and the pileus develops from the apical region. *Laccaria* sp. has a plectenchymal tissue from which the stipe originates, whilst the pileus arises from an apical prosenchymal tissue, as in *Agaricus* [18]. Similar structures were observed in *M. perniciosus* (Figure 3B). However, the development was pseudoangiocarpous since the hymenium was protected by the immature pileus, and no inner veil was present (Figure 4B) [37]. The morphogenetic mechanism was classified as concentrated, based on the description of Reijnders [38] since defined globose primordia with a complex anatomy (Figure 3A) were formed. This is compatible with pileostiptocarpic development because stipe and pileus-originated elements were already present in the primordia at an early stage (Figure 4B).

Genes related to the early development of *M. perniciosus* basidiomata

The molecular basis of cell differentiation that precedes basidiomata formation was recently investigated [17, 19, 39]. Developmentally regulated genes have been identified for some basidiomycetes such as *A. bisporus* [40], *C. cinerea* [19], *Pleurotus ostreatus* [41], among others. Moreover, the rapid increase of fully or partially sequenced genomes and ESTs from fungi already available in databanks allow the *in silico* identification of genes possibly involved in these processes

[42, 43]. However, the understanding of the direct association between these identified genes and their function in the initial development of basidiomata is still incipient. For example, the study of the ESTs of *P. ostreatus* led to the identification of pleurotolysins expressed specifically in the primordial stage. The function of these proteins is being studied, but their role in primordia formation is not yet elucidated [44].

Since the studies in *M. pernicioso* are also in an early stage, the identification of genes related to basidiomata development was a first step to establish a possible correlation between the developmental stages and their expression. The description of morphological changes in mycelium prior to the development of reproductive structures is a key step for subsequent morphogenetic studies and, at this point, helped in the search for genes related to these processes. So far, our contribution has been the analysis of the abundance of transcripts for some selected genes in specific moments during induction of fungal fruiting. Two independent but related tests were carried out. Using 192 genes from a library derived from mycelium in the fructification stage, a reverse Northern analysis, also known as macro array was performed, contrasting the early culturing with the final stage, when the first basidiomata appear. Additionally, a RT-qPCR was performed for 12 genes, analyzing their expression in each of the stages described in the above-described morphological studies.

The development of basidiomycetes such as *C. cinerea*, one of the best-studied to date [19], served as guideline underlying the choice of the genes. In the case of these fungi, fructification seems to occur in genetically pre-conditioned mycelium and in response to nutrient deficiency, as well as to stimuli such as alternating light / dark, humidity and CO₂ concentration [19]. Based on these studies, genes were selected and identified in the available library.

Expression profiles of genes involved in basidiomata development by macroarray

A macroarray analysis was performed with 192 genes encoding putative proteins involved in fruiting, to detect differences in their expression profile between mycelia in white and primordial phases, which would allow their identification as induced or repressed at these two contrasting developmental stages (Figure 5). ESTs were obtained from a full-length cDNA library, previously constructed from mycelia, primordia and mature basidiomata collected during fructification (Pires et al., unpublished data) and selected based on their similarity with known conserved genes. The complete list of the selected genes is shown in Table S1 [see Additional file 1] as well as the fold change values obtained by comparing the results of each spot in the 'white' and 'primordia' stages. A classification based on the likely functions of these gene products was performed as described by Gesteira *et al.* [45], to deepen the understanding of the participation of these genes in the fructification process of *M. pernicioso*. The Table S1 [see Additional file 1] shows also some genes for which the increase of transcripts in the primordial stage compared to the white phase was significant by the Student's t test of means.

The macroarray analyses give us an overview of gene activity during fruiting in *M. pernicioso*. We discriminated 192 genes in two expression patterns: group I, containing up-regulated genes in the white mycelium phase and group II, containing up-regulated genes in the primordia mycelium phase (Figure 5). Some genes are noteworthy because previous descriptions report their participation in the fruiting process of other fungi. In this trial, hydrophobins were represented by four clones and three of them showed increased expression during the primordial stage.

Hydrophobins are cysteine-rich proteins specific for filamentous fungi, capable of generating amphipathic films on the surface of an object [31]. They are related to a broad range of growth and development processes, among them the formation of aerial structures [46]. At least five *M. perniciosa* hydrophobin-encoding genes have been identified [27]. The differences in expression in mycelial mat cultures for basidiomata production were considerable. Unlike four other genes for hydrophobin, one gene was shown to have increased expression in the presence of primordia [32] and two were identified in a compatible *M. perniciosa*-*T. cacao* cDNA library derived from green brooms [45].

Studies in other fungi show that hemolysin expression is specifically increased in the presence of primordia [47], but in this experiment there was no significant increase in the expression of the genes that encode for aegerolysins. Only one gene for pleurotolysin A decreased significantly. On the other hand, genes encoding cytochrome P450 monooxygenase and a heat shock protein had increased expression in the primordial stage, which may indicate the induction of fruiting in response to stress [17]. Cytochrome P450 monooxygenases ('P450s') are a super-family of haem-thiolate proteins that are involved in the metabolism of a wide variety of endogenous and xenobiotic compounds [48]. In *C. cinerea*, the cytochrome P450 similar to CYP64 is most expressed in pilei and seems to be involved in the synthesis route of aflatoxins that seem to be important for fruiting in *Aspergillus* spp. [17].

The appearance of primordia coincided with the decrease of transcripts for calmodulin and increased expression for genes coding for signaling proteins such as RHO1 guanine nucleotide exchange factor (RHO-GEF), RHO GDP-dissociation inhibitor, GTP-binding protein RHEB homolog precursor, indicating that signaling is most likely mediated by fruiting-associated proteins of the Ras family. Additionally, the genes for cellular transport of glucose and gluconate were clearly more

significantly transcribed at the primordial stage [see additional file 1], while a probable transcription factor GAL4 decreased. This indicates that glucose depletion of the medium, which occurs throughout the culture, must be important for fructification and must be related to cAMP signaling [49]. Gene *gtiI*, encoding an inducer of gluconate transport in *Pseudomonas aeruginosa*, controls glucose catabolism, increasing the low-affinity transport system of glucose [50]. The glucose transporter present in this test is rather similar to the high-affinity glucose transporter SNF3, although this has not been confirmed experimentally [51]. Glucose metabolism can be related to fructification [17].

The increase of gene transcripts for vacuolar ATP synthase, phospholipid-transporting ATPase and reductase levodione also indicates that nutrient uptake during the primordial stage serves to form nutrient reserves prior to basidiomata elongation [17]. This is confirmed by the increase of transcripts for several genes of primary and secondary metabolism that may be related to the synthesis of glycerol and lipids. In *C. cinerea* reserves are remobilized and glycogen accumulated in the primordial stage [19].

The expression of three genes related to cell division was significantly higher, two for a 123 kD protein of cell division (Cdc123) and one encoding a suppressor of kinetochore, and one *PIM1* gene was significantly less expressed in the primordial stage. Cdc123 proteins are regulators of eIF2 in *Saccharomyces cerevisiae* and are regulated by nutrient availability [52]. This simultaneous increase indicates the predominance of phase G1 of cell division. As the formation of spores occurs in already differentiated primordia, it is likely that the collected phase contains a larger number of non-differentiated primordia.

There was also a significant increase of six genes of unknown function, one of them showing no similarity with any sequence in the available public data banks.

Expression analyses of genes involved in basidiomata development by RT-qPCR

The gene expression profile obtained by the macroarray in two distinct phases suggested physiological changes in mycelia prior to basidiomata production. However, more detailed analyses should be performed to monitor the expression of key genes (previously described in the literature as involved in basidiomata development). Quantitative PCR is an accurate technique to analyze gene expression. It is 10,000 to 100,000 times more sensitive than RNase protection assays and 1,000 times more sensitive than dot blot hybridization [53]. Therefore, a more detailed RT-qPCR analysis was performed with 12 ESTs in order to observe a possible relationship between transcript levels of all samples collected (Figure 6). RNA was obtained from mycelium samples of all seven developmental stages: white, yellow and reddish pink phases, before and after stress, and during basidiomata formation.

The hypothesis that nutrient depletion might act as a signal for fructification was confirmed since some genes related to primary metabolism and to nutrient uptake were down-regulated when primordia emerged. Conversely, gene expression related to nutrient recycling and stress response increased during this phase, as did the expression of genes directly involved in cell development (Figure 3). The relative expression of the 12 genes in stages that precede fructification helped elucidate the correlation between nutrient depletion and fructification (Figure 6) since the genes *MpRHEB*, *MpRHOI-GEF*, *MpADE*, *MpMBF*, and *MpRAB* putatively involved in signaling are associated with internal perception of the signals triggered by nutrient depletion and other stresses, which was noticeable before the primordia appeared. The putative gene *MpRHEB* is associated with growth regulation probably during nitrogen depletion [54]. Its expression in *M. perniciosus* increased in reddish pink mycelium, immediately before stress and continued at a high level until the beginning of the primordial and basidiomata phases (Figure 6D). The expression of the high-affinity

transporter MpGLU

[51] peaked in this mycelium before stress (Figure 6E), strongly indicating a nutritional deficit, namely low external glucose concentration. Moreover, expression of MpCPR and MpCYP was low during this period (Figure 6G and 6K), indicating a lower basal metabolism [48]. The expression of MpRAB (Figure 6J) may indicate nutrient remobilization, since it is involved in intracellular traffic [55, 56].

During the water stress applied to trigger *in vitro* fructification expression of some genes peaked. Transcripts of putative MpMBF (multi-protein-bridging factor), a co-activator related to tolerance to abiotic stresses in plants [57], increased 2.4-fold (Figure 6I). Other genes with increased expression during this stress period were MpRHO1-GEF (Figure 6H), involved in signaling for the regulation of polarized growth [58] and MpRPL18 (Figure 6L) involved in protein synthesis. Involvement of signalization, probably cAMP-mediated, is likely due the expression of adenylate cyclase that decreased in the yellow and reddish-pink mycelial phases, to return to the original levels observed on white mycelium just after the stress period (Figure 6F). As adenylate cyclase is subject to post-translational regulation, studies of enzymatic activity would be necessary to confirm this hypothesis. The gene *prho/gef* is, therefore, possibly correlated with cAMP pathways. Repression of the glucose transporter coincided with the repression of the adenylate cyclase gene, which also indicates cAMP signaling. In *S. pombe* the glucose levels are regulated by adenylate cyclase [59] and in *Sclerotinia sclerotiorum* the development of reproductive structures is negatively regulated by cAMP [60]

Putative aegerolysins and pleurotolysin B of *M. perniciosa* are differentially expressed during fructification

As described for other fungi, probable hemolysins are highly expressed at the fructification

stages [47, 61]. We identified three putative genes involved in fructification, two more closely related to the identified *AA-PriI* or *PriAs* of *Agrocybe aerogerita* and *P. ostreatus*, respectively, and one more closely related to pleurotolysin B, also identified in *P. ostreatus*. Their different expression profiles suggested that they are different genes (Figure 6A to 6C). The expression of *MpPRIA1* encoding a putative aegerolysin, decreased in the yellow-and reddish pink-mycelium phases, and also before stress, but increased 4.3-fold in mycelia with primordia, and about 90-fold in the basidiomata, compared to the white mycelium stage (Figure 6A). The expression of the putative hemolysin-encoding gene *MpPRIA1* increased 17-fold at the reddish pink mycelium stage, but decreased 11-fold before stress, 4-fold in stressed mycelia, and 47.4-fold in mycelia with primordia. The transcripts of *MpPRIA2* increased 23-fold in basidiomata, but were lower in mycelia with primordia (Figure 6B). The transcripts of gene *MpPLYB*, corresponding to a pleurotolysin B, increased 1.4-fold in the yellow mycelium stage, 15.2-fold in reddish pink mycelia, and remained at high levels in the mycelia before stress (11.7-fold), when stressed (11.2-fold) and in mycelia with primordia (10.1-fold), but decreased in basidiomata, where it was only 1.6 times higher than in white mycelia (Figure 6C).

Hemolysins, already identified in some bacteria and fungi, comprise a cytolytic protein family, whose members appear abundantly during primordia and basidiomata formation [47, 58, 61, 62]. *MpPRIA1* and *MpPRIA2* have homologous regions but seem to correspond to two individual genes whose expression coincides with the morphological differentiation of primary hyphal nodules from primordia. These hemolysins may contribute to the process of hyphal aggregation [61] as their expression occurred, although at low levels, before the appearance of primordia, when hyphae became globose for the formation of the “initials”. This stage coincides with the reddish pink mycelium stage, where hyphal nodules are detectable. The exact function of these proteins remains

unclear, but their involvement in programmed cell death (PCD), as proposed by Kues and Liu [17], seems rather unlikely because ostreolysins have lytic function, acting in cholesterol- and sphingomyelin-containing membranes [63] at a pH between 7 and 8 [64], which is not usually found in fungal cells.

The known fungal hemolysins have some variations in amino acid sequences, but all share the conserved domain aegerolysin (code PF06355 by Pfam database [65]). Aegerolysin *Aa-PriI* from *A. aegerita* has the same molecular weight as the 16 kDa ostreolysin of *P.*

ostreatus and is mainly expressed in the initial stage of primordium formation. PriA (or pleurotolysin or PlyA) of *P. ostreatus* forms a subfamily with the aegerolysin superfamily, which includes the Asp-hemolysins of *Aspergillus fumigatus*, and some hypothetical proteins of *Clostridium bifermentans*, *P. aeruginosa* and *Neurospora crassa*. *P. ostreatus* hemolysin consists of multiple components with isoforms A and B that assemble to a protein complex that leads to the formation of transmembrane pores (diameter 4 nm), specifically allowing lysis of cholesterol and sphingomyelin-containing membranes [63]. Isoform A, called PlyA [17 kDa *PlyA*] has 138 amino acid residues whereas the 59 kDa isoform B polypeptide (PlyB) consists of 538 amino acids.

The two aegerolysin ESTs expressed by *M. perniciosus* constitute two distinct genes (Figures 7 and 8). MpPRIA1 has an ORF of 417 bp with an intron at position 103 whereas the ORF of MpPRIA2 is 406 bp long with an intron at position 134 (data not shown). Both have a conserved aegerolysin domain between residues 4 to 136 (MpPRIA1) and 29 to 135 (MpPRIA2) and can be aligned with a hypothetical protein MPER_11381 (gbEEB90416.1) (Figure 7A) and MPER_04618 (gbEEB96271.1 -not shown) of *M. perniciosus* FA553 and proteins described as aegerolysins of *A. aegerita* (spO42717.1), *P. ostreatus* (PlyA gbAAL57035.1 and ostreolysin -gbAAX21097.1), *A. fumigatus* Af293 (XP_748379.1), *A. fumigatus* (gbBAA03951.1) *Coccidioides immitis* RS (XP_001242288.1) *A.*

niger (XP_001389418.1) (Figure 7A). The evolutionary distance between these putative aegerolysins and above-cited aegerolysin of the Gene Bank database was estimated (Figure 7B). The distances were shorter between MpPRIA1 and MpPRIA2 and aegerolysins of *Pleurotus* and *Agrocybe* than between Mp-PriAs and Asp-hemolysins and ostreolysins of *Aspergillus*. The MpPLYB ORF has 576 bp and two introns (not shown) at positions 211 and 408 corresponding to the genomic DNA of *M. perniciosus* in position 178 to 368 of the sequence deposited in GeneBank (accession no. ABRE01016965). The MpPLYB ORF is more similar to hypothetical proteins of *M. perniciosus* FA553 (gb EEB89936.1) and pleurotolysin B gene described for *P. ostreatus* (gbBAD66667.1) and it can be aligned with proteins described as *Gibberella zeae* PH-1 (XP_390875.1) *A. flavus* NRRL3357 (gbEED49642.1) and *Chaetomium globosum* CBS 148.51 (XP_001227240.1) (Figure 8A). A conserved transmembrane domain MAC/Perforin [PF_01823] occurs between residues 1 and 258. The evolutionary distance between these putative pleurotolysin B and above-cited proteins of the Gene Bank database was estimated (Figure 8B). The distance was shortest between MpPlyB and pleurotolysin B of *Pleurotus*, while the similarity with hypothetical protein MpER_11918 of *M. perniciosus* was highest.

Conclusions

Our analysis of gene expression is an initial approach to correlate gene expression with distinct developmental stages of *M. perniciosus* basidiomata. Gene expression profiles in mycelia before basidiomata induction indicate that the observed morphological changes correlate with induction of genes known to be involved in the development of new macroscopic structures in other fungi. An involvement of a glucose depletion-dependent cell signaling is suggested by the regulation of

adenylate cyclase and glucose transporter genes. However, other up-regulated genes may be responsible for the formation of hyphal nodules, redirecting cytoskeleton modeling, hyphal thickness or nutrient uptake, and most of them may be essential for the maintenance of basidiomata. Our data provide new information about the development of basidiomata in *M. pernicioso* and identify a set of genes probably involved in this process. This information may be useful for further studies towards a more complete understanding of the cell processes and genetic, physiological and environmental controls leading to basidiomata initiation. Once the key genes that determine growth and development of *M. pernicioso* are known, strategies can be provided for an enhanced control of this phytopathogen and for a successful monitoring of witches' broom disease in *T. cacao*.

Methods

Fungal strains and growth conditions

A considerable number of observations of the early primordia development were made in infected brooms collected from cocoa plantations in Itajuípe (14° 40' 43" S, 39° 22' 31" W), Bahia, Brazil. The brooms were kept in a moist chamber and basidiomata formation was induced. Briefly, they were soaked for 1 h in 1% benomyl solution (Sigma Chemical Co., St. Louis, USA), to kill the ascomycete fungi present on the broom surfaces, hung in a chamber (12:12 h light:dark) and sprayed with de-ionized water for 1 min/h for each 24 h period. *M. pernicioso* strain CEPEC 1108 (designated CP03) of the C biotype of *M. pernicioso* was also used for morphological studies. Mycelial starter cultures from the culture collection of the Cocoa Research Center (CEPEC, Ilhéus, Bahia, Brazil) were grown on PDA (Potato Dextrose Agar) for three weeks in the dark, at room temperature. Basidiomata were obtained from mycelial mats, as described by

Griffith and Hedger [7] with the modifications introduced by Niella *et al.* [15]. A solid bran-based medium was prepared (50 g wheat flour; 40 g vermiculite; 6 g $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$, 3 g CaCO_3 and 120 mL distilled water; moisture content 65-70%, pH 7.0-7.5). The mixture was placed in Petri dishes, covered with aluminum foil and autoclaved twice for 90 min (121°C). The cooled medium was inoculated with two 5-mm disc plugs from 1 to 3-week-old mycelium, grown on 2% PDA medium. Cultures were incubated at 25°C in the dark. After mycelia had completely colonized the surface of the bran medium (usually 3-4 weeks), cultures were covered with a 5-mm thick layer (5-10 g per culture), composed of 200 g coarse peat, 50 g CaCO_3 , 50 g vermiculite and 125 mL distilled water (moisture content 70-75%, pH 7.0-7.5). These cultures were incubated for 3 to 4 weeks at 25 °C in the dark and then hung vertically in a broom chamber [14], and maintained at 23 °C $\pm$ 2 °C for 75 d. Irrigation consisted of spraying de-ionized water daily for 7 h with a 12 h period of fluorescent warm white light (65-80W). After 30 d in the chambers, the irrigation was suspended for 7 d, a procedure routinely used to induce fructification.

Microscopic analyses

The preparation of mycelial mat samples for light microscopy was conducted according to standard histological methods [66]. For histological studies of basidiomata development at various stages, samples were fixed after collection by dehydration in a gradient of ethanol/tertiary butyl alcohol series (50 to 100%) for 2 h each, and thermally embedded in paraffin (melting point 56.5 °C; Paraplast plus; Fisher Sci. Co., Pittsburgh, USA). The embedded tissues were radially cut (5 to 14 μm thick) with a rotary microtome. Serial sections were thermally mounted on microscope slides coated with Haupt's adhesive and 4% formalin [67]. The sections were immersed/rinsed three times in 100% xylene and passed through a series of xylene and absolute ethyl alcohol (EtOH) 1:1, absolute ETOH, and 70% ETOH. Some sections were stained with Pianezze III-B stain [68, 69]. This procedure

specifically stained soluble and insoluble proteins red with acid fuchsin and non-living material, i.e. polysaccharides and phenol, green to dark green [35]. Other sections were stained for 1 h with 1% astra blue and then for 1 h with 1% safranin. Macroscopic and tissue analysis was performed in a stereomicroscope (Olympus magnifying glass model SZ2-LGB) and an optical microscope (Olympus model CX41RF), both connected to digital cameras (Olympus model NOC 7070). For histochemical tests, sections of mycelial mats were checked by Fehling' Test [70] to detect reduced sugars, by Sudan III solution [71] to detect lipids and by Floroglucinol Acid solution [72] to detect phenolic compounds. For scanning electron microscopy (SEM), samples were fixed in FAA (5% formaldehyde; 5% acetic acid; 63% ethanol), and dehydrated in increasing acetone solutions (30 to 100%), for 15 min at each concentration. Sections were dried to the critical point, mounted in stubs, and covered with gold before SEM analysis (Model LEO 54X (Zeiss), at the State University of Feira de Santana (Feira de Santana, Bahia, Brazil).

Fungal strains, sampling, growth conditions for molecular analysis and RNA isolation

M. pernicioso strain FA553 (Cp02), sequenced by the WBD Genome Project [27] was used for macroarray and RT-qPCR analyses. Growth conditions were described as above except for some details: the chamber was a glass box (40x30x30 cm) with hooks on the lid underside. Units of mycelial mats were suspended on these hooks and washed aseptically. Temperature and light were as mentioned above. Samples were collected in the different pigmentation phases: white, yellow, reddish-pink, reddish-pink before stress and reddish-pink mycelium after stress (10 d without irrigation); mycelium containing primordia, and basidiomata (Figure 1G). Individual samples of

CP02 were processed using the RNAeasy Plant Midi Kit (Qiagen, Valencia, USA). The RNA samples were qualitatively and quantitatively analyzed by denaturing formaldehyde/agarose gel electrophoresis and optical density was determined [73]. Aliquots of each sample were stored at -80 °C until analysis. Figure 1G summarizes sampling for RNA extractions.

cDNA library construction and analysis of differential gene expression by macroarray

The macroarray membrane was spotted with 192 cDNA clones in duplicate, which were selected from a cDNA library based on their putative role in basidiomata development in other fungi and their involvement in nutrient depletion and cell signaling. For the cDNA library construction, the *M. perniciosus* strain CEPEC 1108 (CP03) was cultured as previously described and mycelium samples in white, yellow, reddish-pink, dark reddish pink and primordium stages, as well as from basidiomata were used to construct a full-length, non-normalized cDNA library. Total RNA was extracted from samples using RNAs in RNA Plant Midi Kit as described by the manufacturer (Qiagen) and after quantification, 1 µg was used to construct the library using DB SMART Creator cDNA library as described by the manufacturer (Clontech). cDNA strands longer than 400 bp were cloned directionally into the pDNR-LIB plasmid. ElectroMAXTM DH10BTM cells (Invitrogen) were transformed and colonies selected and grown in 96-well microtiter plates in LB, 40% glycerol medium containing 30 µg/L chloramphenicol and stored at -80 °C. A total of 2,759 clones were sequenced using a capillary sequencer (Mega Bace 1000, GE Healthcare). After the filtering, trimming, and clustering processes the 1,533 obtained ESTs were evaluated based on functional annotation. The cDNA fragments used to spot the macroarray membrane were amplified by PCR using M13 primers [forward 5'-CAGGAAACAGCTATGAC-3' and reverse 5'-GTAAAACGACGGCCAG-3'] that annealed to the vector pDNR-LIB (Clontech), transferred in duplicate to membranes (Hybond N+, Amersham Biosciences) [73] and fixed using a UV crosslinker

(Spectronics Corporation). For macroarray hybridization, two distinct RNA pools were used: one cDNA mixture of three distinct biological samples from the initial cultivation phases on artificial media (white phase), and another cDNA mixture of three distinct biological samples from the primordial stage. The membrane was hybridized twice with each cDNA pool. Labeling (400 ng of each cDNA pool), pre-hybridization (4 h), hybridization (2.5 h) and signal detection were performed as recommended by the manufacturer of the Alkaphos kit (GE Healthcare). The membranes were exposed to X-Omat (Kodak) film for 2.5 h and the images captured using the Scanner Power Look 1120 UDS (Amersham Biosciences) and analyzed with BZ Scan [74]. The presence or absence of the signal, as well as the intensity, was registered for each individual spot. Global normalization and clustering of the generated intensities, using software Cluster version 3.0 [75]. The default Cluster for normalization was performed eight times, with genes centralized by average. A total clustering of genes was made by the uncentered method (Pearson correlation). This value used in hierarchical clustering represents the average intensity of each gene. Student's t-test, was used after global standardization and before clustering to establish a comparison between means. The values significant at 5% probability and the genes accession numbers are shown in Table S1 [see Additional file 1] together with the fold change values based on the means generated after normalization by Cluster 3.0 software.

Quantitative analyses of reversed transcripts (RT-qPCR)

During the growth period in artificial medium, 12 selected genes were analyzed based on their expression pattern derived from the macroarray. The following genes were selected from the EST data base (www.lge.ibi.unicamp.br/vassoura) encoding the proteins: three putative hemolysins (CP03-EB-001-020-G09-UE.F; CP03-EB-001-008-C10-UE.F; CP03-EB-001024-G03-UE.F), a

putative 60S ribosomal L18 protein (CP03-EB-001-001-E05-UE.F), a putative Rho1/GEF (CP03-EB-001-012-F03-UE.F), a putative Rab (Ras family) (CP03-EB001-020-F11-UE.F), a putative multi-protein-bridging factor (CP03-EB-001-025-E06-UE.F), a putative Ras-GTP-binding protein Rhb1 (CP03-EB-001-005-E11-UE.F), a putative glucose transporter (CP03-EB-001-015-G10-UE.F), a putative cytochrome P450 (CP03-EB-001-025D09-UE.F), a putative adenylate cyclase (CP03-EB-001-025-C05-UE.F), and a putative NADPH-cytochrome P450 reductase (CP03-EB-001-001-B10-UE.F). A putative polyubiquitin (CP03-EB-001-020-H08-UE.F) was used as reference gene. All PCR primers (MWG, Imprint Genetics Corp) were designed using the GeneScript online Real-Time Primer Design tool (<https://www.genscript.com/ssl-bin/app/primer>) [see Additional file 2]. One microgram of total RNA treated with RQ1 DNase I (Invitrogen) was reverse-transcribed using Power Script (Invitrogen) at a final volume of 20 μ L. The primer T_m was set at 59° C to 61° C and the amplicon sizes ranging from 100 to 105 bp. Quantitative PCR was performed using SYBRGreen[®] (Invitrogen) for the detection of fluorescence during amplification, and assays were performed on an ABI PRISM 7500 Sequence Detection System (SDS) coupled to the ABI PRISM 7500 SDS software (Applied Biosystems, Foster City, USA), using standard settings. A 20 μ L RT-PCR reaction consisted of 2 μ L SYBRGreen 1x (Applied Biosciences), 1x PCR buffer, 200 mM dNTPs, 3 mM MgCl₂, ¼ 50X Rox, 200 nM each primer and 10 μ L single-stranded cDNA. The thermal cycling conditions were 50 °C for 2 min, then 94 °C for 10 min, followed by 40 cycles of 94 °C for 45 s, 57 °C for 35 s for annealing, and 72 °C for 35s. A dissociation analysis was conducted after all amplifications to investigate the formation of primer dimers and hairpins. Melting temperatures of the fragments were determined according to the manufacturer's protocol. No-template reactions were included as negative controls in every plate. Sequence Detection Software (Applied Biosystems, Foster City, USA) results were imported into Microsoft Excel for further analysis. Raw expression

levels were calculated from the average of the triplicate ddCT (RQ) values using the standard curve obtained for each primer pair (ABI PRISM 7500 Sequence Detection System User Bulletin #2). A non-parametric *t* test was performed in order to compare the expression values obtained for each gene between the samples.

Molecular analyses of aegerolysin genes

The two putative aegerolysin genes (Mp*PRIA1* and Mp*PRIA2*) and one putative pleurotolysin B (Mp*PLYB*), were analyzed by aligning ESTs and genomic sequences using Clustal W (EBI) [76]. The contigs were screened for conserved domains and for introns using ORFINDER software (NCBI-www.ncbi.nlm.nih.gov/projects/gorf). The amino acid sequences generated from the most likely ORFs were aligned against four sequences available at the UNIPROT database [77] using Multalign [78]. The evolutionary history was inferred using the Neighbor-Joining method [79]. The evolutionary distances were calculated following the Poisson correction method [80] and expressed in units of number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 116 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 [81]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches [82].

Authors' contributions

ABLP –Fungus culturing, RNA extraction, cDNA library construction, microscopy tissue

preparations, macroarray and RT-qPCR analyses, electronic microscopy analyses and manuscript drafting. MMS - Fungus maintenance, RNA extraction and cDNA library construction. KPG – Fungus maintenance, microscopy tissue preparations and manuscript drafting. DCS- microscopy slide preparations and biochemical tests. RFP and JSMF– macroarray construction. CVD – macroarray construction and RT qPCR analyses. AGN – scanning microscopy analyses and manuscript draft preparation. MB- manuscript preparation and result interpretation. JCMC and GAGP – headed and promoted the Project, manuscript elaboration. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by the National Council of Scientific and Technological Development (CNPq) and the Programa de Apoio ao Desenvolvimento Científico e Tecnológico (PADCT).

This work was also supported by FINEP (Grant 01.07.0074-00) and FAPESB (Grant 1431080017116) and is part of the *M. pernicios*a proteomic project.

A.B.L.P. holds a PQI/CAPES fellowship. The Fundação de Apoio à Pesquisa do Estado da Bahia (FAPESB) funded A.B.L.P., C.V.D. and M.B. and the PROIIC program of UESC funded M.M.S.

We thank Antônio Figueira, Raul Valle, John Hammerstone and Gareth W Griffith for critical reading of the manuscript and Braz Tavares da Hora Júnior for introduction to macroarray analysis.

References

1. Aime MC, Phillips-Mora W: **The causal agent of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of *Marasmiaceae*.** *Mycologia* 2005, **97**: 1012-1022.
2. Purdy LH, Schmidt RA: **Status of cacao witches' broom: biology, epidemiology, and management.** *Annu Rev Phytopath* 1996, **34**: 573-594.

3. Pereira JL, Ram A, Figueiredo JM, Almeida LCC: **Primeira ocorrência de vassoura-de-bruxa na principal região produtora de cacau do Brasil.** *Agrotrópica (Brazil)* 1989, **1**: 79-81.
4. Trevizan SDP; Marques M: **Impactos sócio-econômicos da crise do cacau: um estudo de comunidade-caso.** *Agrotrópica (Brazil)* 2002, **14**: 127-136.
5. Meihardt LW, Rincones J, Bailey B, Aime MC, Griffith GW, Zhang D, Pereira G: ***Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe?** *Mol Plant Pathol* 2008, **9** (5), 577–588.
6. Ceita GO, Macedo JNA, Santos TB, Allemano L, Gesteira AS, Micheli F, Mariano AC, Gramacho KP, Silva DC, Meinhardt L, Mazzafera P, Pereira GGA, Cascardo JCM: **Involvement of calcium oxalate degradation during programmed cell death in *Theobroma cacao* tissues triggered by the hemibiotrophic fungus *Moniliophthora perniciosa*.** *Plant Sci (Limerick)* 2007, **173**: 106-117.
7. Griffith GW, Hedger JN: **A novel method for producing basidiocarps of the cocoa pathogen *Crinipellis perniciosa* using a bran-vermiculite medium.** *Europ J Plant Pathol* 1993, **99**: 227-230.
8. Suarez C: **Growth of *Crinipellis perniciosa* (Stahel) Singer *in vivo* and *in vitro*.** PhD. Thesis, University of London, 1977.
9. Rocha HM: **The ecology of *Crinipellis perniciosa* (Stahel) Singer in Witches' broom on cocoa (*Theobroma cacao* L.).** PhD Thesis, University of London, London, England, 1983.
10. Rocha HM, Wheeler BEJ: **The water balance as an important factor in basidiocarp production by *Crinipellis perniciosa*, the causal fungus of cocoa witches' broom.** *Proc 8th Internat. Cocoa Res. Conf. 1981. Cartagena, Columbia: Cocoa Producers Alliance*, 1982, 381–386.
11. Rocha HM, Wheeler BEJ: **Factors influencing the production of basidiocarps and the deposition and germination of basidiospores of *Crinipellis perniciosa*, the causal agent of witches' broom disease on cocoa (*Theobroma cacao*).** *Plant Pathol* 1985, **34**: 319-328.
12. Stahel G: **Contribution to the knowledge of witch broom disease.** Surinam Department of Agriculture Bull. 39. *In: Tropical Agriculture* 1919, **IX**: 167-176.

13. Purdy LH, Trese AT, Aragundi JA: **Proof of pathogenicity of *Crinipellis perniciosus* to *Theobroma cacao* by using basidiospores produced in *in vitro* cultures.** *Theobroma (Brazil)* 1983, **13**: 157-163.
14. Purdy LH, Dickstein ER: **Basidiocarp development on mycelial mats of *Crinipellis perniciosus*.** *Plant Dis* 1990, **74**: 493-96.
15. Niella G, Resende ML, Castro HA, Carvalho GA de, Silva LHCP: **Aperfeiçoamento da metodologia de produção artificial de basidiocarpos de *Crinipellis perniciosus*.** *Fitop Brasileira* 1999, **24**: 523–527.
16. Macagnan D, Romeiro RS, Souza J, Pomella AWW: **Isolation of actinomycetes and endospore-forming bacteria from the cacao pod surface and their activity against the witches' broom and black pod pathogens.** *Phytoparasitica* 2006, **34**: 122-132.
17. Kues U, Liu Y: **Fruiting body production in basidiomycetes.** *Appl Microbiol Biotechnol* 2000, **54**: 141-152.
18. Massicotte HB, Melville LH, Peterson RL: **Building a basidiocarp: a case study of *Laccaria* spp. fruitbodies in the extraradical mycelium of *Pinus* ectomycorrhizas.** *Mycologist* 2005, **19**: 141-149.
19. Kues U: **Life history and developmental processes in the basidiomycete *Coprinus cinereus*.** *Microbiol Mol Biol Rev* 2000, **64**: 316-353.
20. Almeida LC, Bastos CN, Ferreira NP: **Produção de basidiocarpos de *Crinipellis perniciosus* em dois sistemas de cultivo de cacaueiro.** *Fitop Brasileira* 1995, **20**: 60-64.
21. Evans HC, Bastos CN: **Basidiospore germination as a means of assessing resistance to *Crinipellis perniciosus* (Witches' broom disease) in cocoa cultivars.** *Trans Br Mycol Soc* 1980, **89**: 525-536.
22. Evans HC: **Witches' broom disease - A case study.** *Cocoa Growers Bulletin* 1981, **32**: 5-19.
23. Delgado JC, Cook AA: **Nuclear condition of basidia, basidiospores, and mycelium of *Marasmius perniciosus*.** *Canad J Botany* 1976, **54**: 66-72.
24. Muse RB, Collin HA, Isaac S, Hardwick K: **Effects of the fungus *Crinipellis perniciosus*, causal agent of witches' broom disease, on cell and tissue cultures of cocoa (*Theobroma cacao* L.).** *Plant Pathol* 1996, **45**: 145–154.

25. Kilaru A, Hasenstein KH: **Development and pathogenicity of the fungus *Crinipellis pernicios* on interaction with cacao leaves.** *Phytopathology* 2005, **95**: 101-107.
26. Silva SDVM, Matsuoka K: **Histologia da interação *Crinipellis pernicios* em cacaueiros suscetível e resistente à vassoura-de-bruxa.** *Fitop Brasileira* 1999, **24**: 54-59.
27. Mondego JMC, Carazzolle MF, Costa GGL, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro, DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, García O, Thomazzela DPT, Oliveira BV, Pires ABL, Rio MCS, Araújo MRR, Castro LAB, Gramacho KP, Gonçalves MS, Góes-Neto A, Barbosa LV, Guiltinan MJ, Bailey B, Meinhardt L, Cascardo JCM, Pereira GAG: **A genome survey of *Moniliophthora pernicios* gives new insights into Witches' Broom Disease of cacao.** *BCM Genomics* 2008, **9**:548.
28. Heckman CA, Pelok SD, Kimpel SA, Wu LC: **Scanning electron microscope studies on fruitbody primordium formation in *Agaricus bisporus*.** *Mycologia* 1989, **81**: 717-727.
29. Lopes MA: **Estudo molecular de quitinases de *Crinipellis pernicios* (Stahel) Singer.** M. S. Thesis, Universidade Estadual de Santa Cruz, Ilhéus – Bahia, Brazil 2005.
30. Kershaw MJ, Talbot NJ: **Hydrophobins and repellents: proteins with fundamental roles in fungal morphogenesis.** *Fungal Gen Biol* 1998, **23**: 18-33.
31. Wösten HAB, De Vocht ML: **Hydrophobins, the fungal coat unraveled.** *Biochim Biophys Acta* 2000, **1469**: 79-86.
32. Santos SC: **Caracterização de hidrofobinas do fungo *Crinipellis pernicios* (Stahel) Singer, causador da doença Vassoura-de-Bruxa no cacaueiro.** M. S. Thesis, Universidade Estadual de Santa Cruz, Ilhéus – Bahia, Brazil, 2005.
33. Reijnders AFM: **On the origin of specialized trama types in the *Agaricales*.** *Mycol Res* 1993, **97**: 257-268.
34. Walther V, Rexer KH, Kost G: **The ontogeny of the fruit bodies of *Mycena stylobates*.** *Mycol Res* 2001, **105**: 723-733.
35. Fisher DB: **Protein staining of ribboned Epon sections for light microscopy.** *Histochemie* 1968, **16**: 92-96.
36. Lopes MA, Gomes DS, Bello-Koblitz MG, Pirovani CP, Cascardo JCM, Góes-Neto A, Micheli F: **Use of response surface methodology to examine chitinase regulation in the basidiomycete *Moniliophthora pernicios*.** *Mycol Res* 2008, **112**: 399-406.

37. Alexopoulos CJ, Mims CW, Blackwell M: **Introductory Mycology** (4th Ed.). John Wiley and Sons, New York, USA 1996.
38. Reijnders AFM: **Lês Problèmes du développement du carpophore des *Agaricales* et de quelques groupes voisins**. Junk, The Hague, 1963.
39. Busch S, Braus GH: **How to build a fungal fruit body: from uniform cells to specialized tissue**. *Mol. Microb.* 2007, **64**(4):873-876.
40. De Groot PWJ, Schaap PJ, Van Griensven LJLD, Visser J: **Isolation of developmentally regulated genes from the edible mushroom *Agaricus bisporus***. *Microbiology* 1997, **143**: 1993–2001.
41. Lee SH, Kim BG, Kim KJ, Lee JS, Yun DW, Hahn JH, Kim GH, Lee KH, Suh DS, Kwon ST, Lee CS, Yoo YB: **Comparative analysis of sequences expressed during the liquid-cultured mycelia and fruit body stages of *Pleurotus ostreatus***. *Fungal Gen Biol* 2002, **35**: 115-134.
42. Soanes DM, Talbot, NJ: **Comparative genomic analysis of phytopathogenic fungi using expressed sequence tag (EST) collections**. *Mol Plant Pathol* 2006, **7**(1): 61-70.
43. Nowrousian M, Kück U: **Comparative gene expression analysis of fruiting body development in two filamentous fungi**. *FEMS Microbiol Lett* 2006, **257**: 328–335.
44. Berne S, Lah L, Sepčić K: **Aegerolysins: Structure, function, and putative biological role**. *Protein Sc.* 2009, **18** (4):694-706.
45. Gesteira AS, Micheli F, Carels N, Silva AC da, Gramacho KP, Shuster I, Macedo JN, Pereira GAG, Cascardo JM: **Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa***. *Ann Bot* 2007, **100**: 129-140.
46. Wösten HAB: **Hydrophobins: multipurpose proteins**. *Annu Rev Microb* 2001, **55**: 625-646.
47. Vidic I, Berne S, Drobne D, Maček P, Frangež R, Turk T, Štrus J, Sepčić K: **Temporal and spatial expression of ostreolysin during development of the oyster mushroom (*Pleurotus ostreatus*)**. *Mycol Res* 2005, **109**: 377-382.
48. Yadav JS, Doddapaneni H, Subramanian V: **P450ome of the white rot fungus *Phanerochaete chrysosporium*: structure, evolution and regulation of expression of genomic P450 clusters**. *Biochem Soc Trans* 2006, **34**: 1165-1169.

49. Byrne SM, Hoffman CS: **Six *git* genes encode a glucose-induced adenylate cyclase activation pathway in the fission yeast *Schizosaccharomyces pombe*.** *J Cell Sci* 1993, **105**: 1095-1100.
50. Whiting PH, Midgley M, Dawes E: **The regulation of transport of glucose, gluconate and 2-oxogluconate and of glucose catabolism in *Pseudomonas aeruginosa*.** *Biochem J* 1976, **154**: 659-668.
51. Ko CH, Liang H, Gaber RF: **Roles of multiple glucose transporters in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1993, **13(1)**: 638-648.
52. Bieganowski, P, Shilinski K, Tschlis PN, Brenner C: **Cdc123 and checkpoint forkhead associated with RING proteins control the cell cycle by controlling eIF2 γ abundance.** *J. Biol. Chem.* 2004, **279** (43): 44656-44666.
53. Wong ML, Medrano JF: **Real-time PCR for mRNA quantitation.** *BioTechniques* 2005, **39**: 75-85.
54. Mach KE, Furge KA, Albright CF: **Loss of Rhb1, a Rheb-Related GTPase in fission yeast, causes growth arrest with a terminal phenotype similar to that caused by nitrogen starvation.** *Genetics* 2000, **155**: 611-622.
55. Grosshans BL, Ortiz D, Novick P: **Rabs and their effectors: achieving specificity in membrane traffic.** *Proc Natl Acad Sci USA* 2006, **103**: 11821-11827.
56. Schwartz SL, Cao C, Pylypenko O, Rak A, Wandinger-Ness A: **Rab GTPases at a glance.** *J Cell Sci* 2007, **120**: 3905-3910.
57. García P, Tajadura V, García I, Sánchez Y: **Rgf1p is a specific Rho1-gef that coordinates cell polarization with cell wall biogenesis in fission yeast.** *Mol Biol Cell* 2006, **17**: 1620-1631.
58. Berne S, Križaj I, Pohleven F, Turk T, Maček P, Sepčić K: ***Pleurotus* and *Agrocybe* hemolysins, new proteins hypothetically involved in fungal fruiting.** *Biochim Biophys Acta* 2002, **1570**: 153-159.
59. Kao RS, Morreale E, Wang L, Douglas Ivey F, Hoffman CS: ***Schizosaccharomyces pombe* Git1 is a C2-domain protein required for glucose activation of adenylate cyclase.** *Genetics* 2006, **173**: 49-61.

60. Jurick II WM, Rollins JA: **Deletion of the adenylate cyclase (*sacI*) gene affects multiple developmental pathways and pathogenicity in *Sclerotinia sclerotiorum*.** *Fungal Gen Biol* 2007, **44** (6): 521-530.
61. Berne S, Pohleven J, Vidic I, Rebolj K, Pohleven F, Turk T, Maček P, Sonnenberg A, Sepčić K: **Ostreolysin enhances fruiting initiation in the oyster mushroom (*Pleurotus ostreatus*).** *Mycol Res* 2007, **111**: 1431-1436.
62. Fernandez-Espinar MT, Labarère J: **Cloning and sequencing of the *Aa-PriI* gene specifically expressed during fruiting initiation in the edible mushroom *Agrocybe aegerita*, and analysis of the predicted amino-acid sequence.** *Curr Genet* 1997, **32**: 420-424.
63. Sepčić K, Berne S, Rebolj K, Batista U, Plemenitas A, Sentjurs M, Maček P: **Ostreolysin, a pore-forming protein from the oyster mushroom, interacts specifically with membrane cholesterol-rich lipid domains.** *FEBS Lett* 2004, **575**(1-3):81-5.
64. Berne S, Sepčić K, Anderluh G, Turk T, Maček P, Ulrih NP: **Effect of pH on the pore forming activity and conformational stability of Ostreolysin, a lipid raft-binding protein from the edible mushroom *Pleurotus ostreatus*.** *Biochemistry* 2005, **44**: 11137–11147.
65. Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer ELL, Bateman A: **Pfam: clans, web tools and services.** *Nucleic Acid Res* 2006, **34**: D247-D251.
66. Johansen DA: **Plant microtechniques.** McGraw-Hill, New York, New York, USA 1940, 194.
67. Van Cotte W, Fryns-Claessens E: **Plantenanatomie in Practijk.** J. Lier, Belgium: Van In. 1972.
68. Vaughan RE: **A method for the differential staining of fungus and host cells.** *Ann Mol Bot Gard* 1914, **1**: 241-242.
69. Gramacho KP: **Disease resistance and pathogenic variability in the fusiform rust-slash pine pathosystem.** *PhD Thesis, University of Florida, Gainesville* 1999.
70. Purvis MJ, Collier DC, Walls D: **Laboratory techniques in botany.** London, Butterworths 1964, 153.

71. Sass JE: **Botanical microtechnique**. 2nd ed. Ames, The Iowa State College Press 1951, 228.
72. Sambrook J, Russell DW: **Molecular Cloning. A Laboratory Manual**. Third edition, New York: Cold Spring Harbor Laboratory, 2001.
73. Lopez F, Rougemont J, Loriod B, Bourgeois A, Loi L, Bertucci F, Hingamp P, Houlgatte R, Granjeaud S: **Feature extraction and signal processing for nylon DNA microarrays**. *BMC Genomics* 2004, **5**: 38.
74. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns**. *Proc Natl Acad Sci USA* 1998, **95**: 14863-14868.
75. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**. *Nucl Acids Res* 1994, **22**: 4673-4680.
76. Apweiler R, Bairoch A, Wu CH: **Protein sequence databases**. *Curr Opin Chem Biol* 2004, **8**:76-80.
77. Corpet F: **Multiple sequence alignment with hierarchical clustering**. *Nucl Acids Res* 1988, **16**: 10881-10890.
78. Saitou N, Nei M: **The neighbor-joining method: A new method for reconstructing phylogenetic trees**. *Mol Biol Evol* 1987, **4**: 406-425.
79. Zuckerkandl E, Pauling L: **Evolutionary divergence and convergence in proteins**, pp. 97-166. *In* *Evolving Genes and Proteins*, edited by V Bryson and HJ Vogel. Academic Press, NY, 1965.
80. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0**. *Mol Biol Evol* 2007, **24**: 1596-1599.
81. Felsenstein J: **Confidence limits on phylogenies: An approach using the bootstrap**. *Evolution* 1985, **39**:783-791.

Figure legends

Figure 1. Mycelial stages prior to emergence of *M. perniciosus* primordia. A, B, C. Mycelial mat originating from basidiospore germination on dead cocoa branches. D. Mycelial mat cultured on artificial substrate. Mycelium is initially white (A) then turns yellow (B) and changes to reddish pink

(C) (A, B, C; bars= 0.5 cm), and maintains this color during primordial and basidiomata development, both in natural and artificial conditions (D; bar=1.25 cm). E. Globose protuberance covered by mycelial mat (*) and openings for initial sprouting (bar= 1 mm). F. Primordia emergence (bar= 1 mm). G. Schematic representation of the sampling during cultivation for library construction (CP03) and macroarrays and RTqPCR (CP02). Lateral numbers indicate days of cultivation. Box A – time 0, when the Petri dishes were inoculated. Box B – First harvest before hanging the mycelia in moist growth chambers. Box C- Second harvest with yellow mycelia. Box D- Third harvest with pink-reddish mycelium. Box E- Fourth harvest with reddish-pink mycelium before stress. Box F - Fifth harvest with dark pink mycelia (CP03), or reddish-pink after stress (CP02). G- Sixth harvest of primordia and fully-developed basidiomata. The days of cultivation differ due the differences between fungal isolates.

Figure 2. Aspects of hyphal organization before fruiting of *M. perniciosus*. A-D: Scanning electron micrograph shows aerial hyphae. E-F. Section of mycelial mat of the “dark reddish pink” stage on dead cocoa branch, stained with Lugol and Safranin. A: Hyphae of mycelial mat in the white phase (Griffith medium). Note branched hyphae (yellow arrow), hyphal fusion (white arrow), thin layer apparently composed by cell wall materials (red arrow) and hyphae with irregular aspect (pink arrow; bar= 10 μ m). B. Details of external hyphae after some days of exposure of mycelial mat to frequent irrigation. Note impregnated material in superficial hyphae (bar=10 μ m). C. Dark reddish pink mycelia with protuberance on the hyphae surface were they over layer the impregnated material, fanning out in ring shape (bar=20 μ m). D. Amorphous material recovering hyphae in differentiated primordium (bar= 10 μ m). E. An outer layer (arrow) and aggregate aerial hyphae can be seen on the surface (bar= 0.12 mm). F. Hyphal nodule observed in reddish-pink mycelium (bar= 0.04 mm).

Figure 3. Early developmental stages of *M. perniciosus* basidiomata. A. Globose hyphal aggregate (initial primordium) under a superficial layer of mycelial mat (bar= 0.25 mm). B. Schematic drawing of the area marked in A showing the grouping of protective hyphae (*) laterally involving another more compact group (#). At the top another group of converging hyphae grows downwards (black squares) (bar= 0.12 mm). C. Tissue section showing an emerging undifferentiated “initial” (bar= 0.25 mm). D. Schematic drawing of C showing the expansion of marked hyphae presented in Fig. 2B. The arrows indicate the same previous protective layer but the

compact bulb has already overlapped it (bar= 0.25 mm). **E.** Another “initial” in a more advanced developmental state (bar= 0.25 mm). **F.** Schematic drawing of E showing protective hyphae placed in parallel positions and the laterally expanding bulb hyphae (arrows) (bar= 0.25 mm).

Figure 4. Aspect of primordia of *M. pernicioso*. **A.** Section of initial primordium stained with Pianezze III. Note the globose form, the distance between the septa and the pigment impregnated within the hyphal cell wall (arrow; bar= 0.1mm). **B.** Hand-cut section of a primordium showing the separation between the tissue that will develop the pileus (P) and the rod that will develop the stipe (S) (bar= 0.25 mm). **C.** Lamella appeared by digestion in areas of pileus (bar= 0.25 mm). **D-F.** Scanning electron micrograph. **D.** Differentiated primordium with radial growing hyphae in pileus (bar= 100 μ m, on detail bar = 30 μ m). **E.** Densely packed stipe hyphae (bar= 20 μ m). **F.** Clamped hyphae of primordium (bar= 2 μ m). **G.** Primordia extension stage (bar= 1 mm). **H.** Different primordia in extension stage (bar= 0.5 cm). **I.** Basidiomata obtained *in vitro* with exposed lamellae (bar= 1 cm).

Figure 5. Genes expressed differentially in white mycelia and mycelia with primordia A.

Hierarchical clustering illustrating groups of 192 *M. pernicioso* genes coordinately expressed at the moment of fruiting versus white mycelium stage by macroarray assay. The column W represents samples of white mycelium stages and P the primordium stage. For each gene, the medium mRNA levels represented by red or green, indicating up-regulation or down-regulation, respectively. The legend indicates the corresponding values of intensity. Two groups are formed: A = higher gene expression in 'white' mycelium and B = higher expression in mycelium with 'primordia'. On the right are examples of genes evaluated in each group.

Figure 6. RT-qPCR of genes expressed in different phases during the culture of *M.*

***pernicioso* in basidiomata-inducing medium.** In Y axis values of RQ using primers constructed for each gene and in axis X corresponding samples of RNA originated from mycelia in the following stages: 1= cDNA of mycelium white stage, 2= cDNA of yellow mycelium stage, 3= cDNA of reddish pink mycelium stage, 4= cDNA of reddish pink mycelium before stress, 5= cDNA of reddish pink mycelium after stress, 6= cDNA of mycelium with primordia

and 7= cDNA of basidiomata. RQ= relative quantification measured by ddCt. (*) - significant at 5% probability, (**) - significant at 1% probability by the statistical T test.

Figure 7. Comparison between *M. pernicios*a aegerolysins and other fungi. **A-** Alignment for similarity between ORFS of the two probable aegerolysins of *M. pernicios*a (Mp1-PriA and Mp2-PriA) and aegerolysins of *M. pernicios*a FA553 (gbEEB90416.1), *A. aege*rita (spO42717.1), *P. ostreatus* (PriA - gbAAL57035.1 and ostreolysin - gbAAX21097.1), *A. fumigatus* Af293 (XP 748379.1), *A. fumigatus* (gbBAA03951.1) *C. immitis* RS (XP 001242288.1) *A. niger* (XP 001389418.1). Strictly conserved residues are shown in black and similar residues in gray. Consensus symbols: ! is any of IV, \$ is any of LM, % is any of FY, # is any of NDQEBZ. Domain PF06355 (aegerolysin family) is present in Mp1PriA (residues 4-136, score 8.7e-61) and Mp2-PriA (residues 29-135, score 4.2e-34). **B.** Phylogenetic analysis of the probable aegerolysin genes of *M. pernicios*a with above-cited sequences. The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed.

Figure 8. Comparison between *M. pernicios*a pleurotolysin and other fungi. **A-** Alignment for similarity between ORFS of the one probable pleurotolysin B of *M. pernicios*a (Mp-PlyB) and hypothetical proteins of *M. pernicios*a FA553 (gb EEB89936.1), *P. ostreatus* (gb BAD66667.1), *G. zeae* PH-1 (XP 390875.1), *A. flavus* NRRL3357 (gbEED49642.1), *C. globosum* CBS 148.51 (XP 001227240.1). Strictly conserved residues are shown in black and similar residues in gray. Consensus symbols are used similarly as in Fig.7. Domain MAC/Perforin (PF01823) is present in Mp-PlyB (residues 1 to 258, score -35,2). **B.** Phylogenetic analysis of the probable pleurotolysin B gene of *M. pernicios*a with above-cited sequences. The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed.

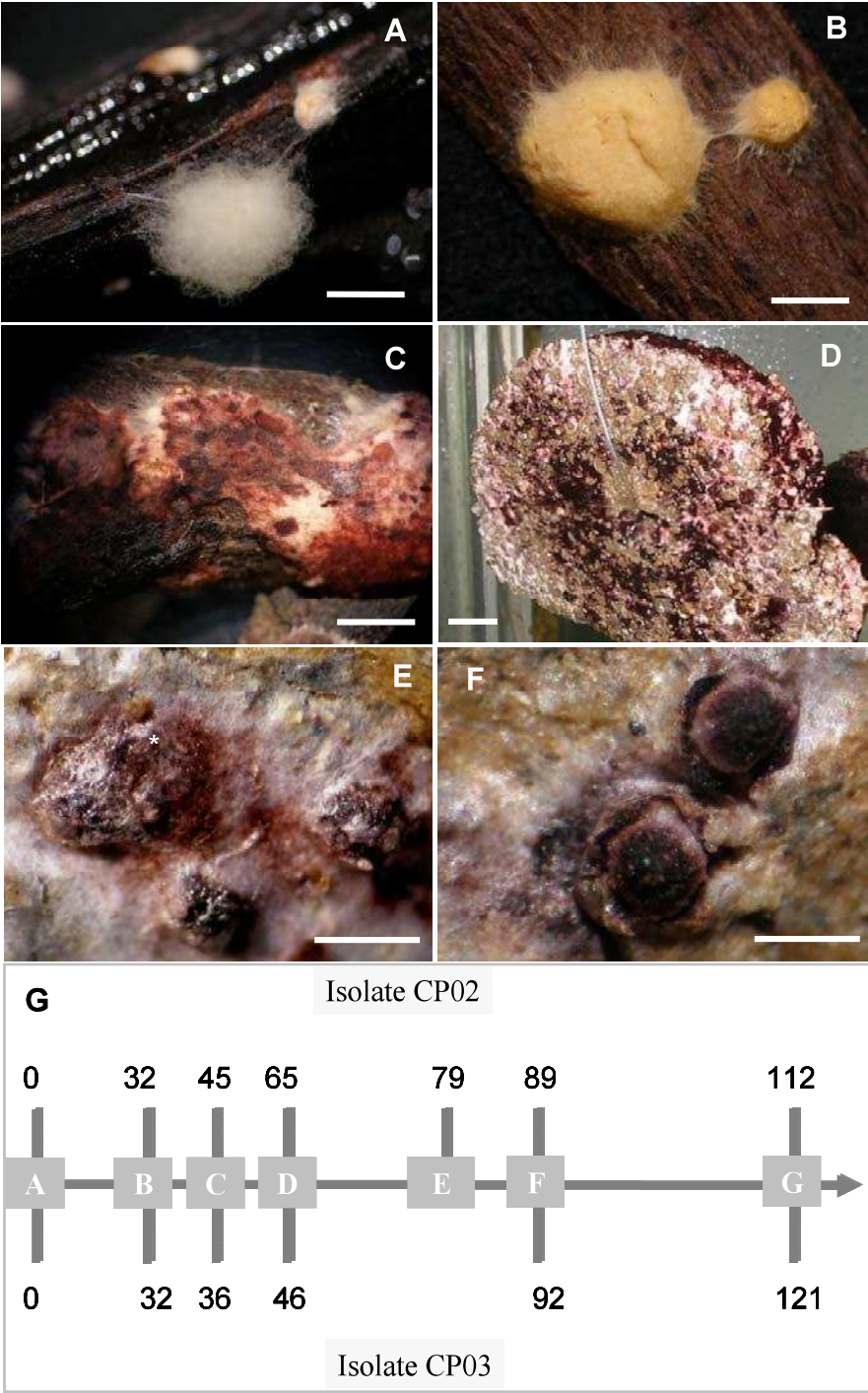


Figure 1

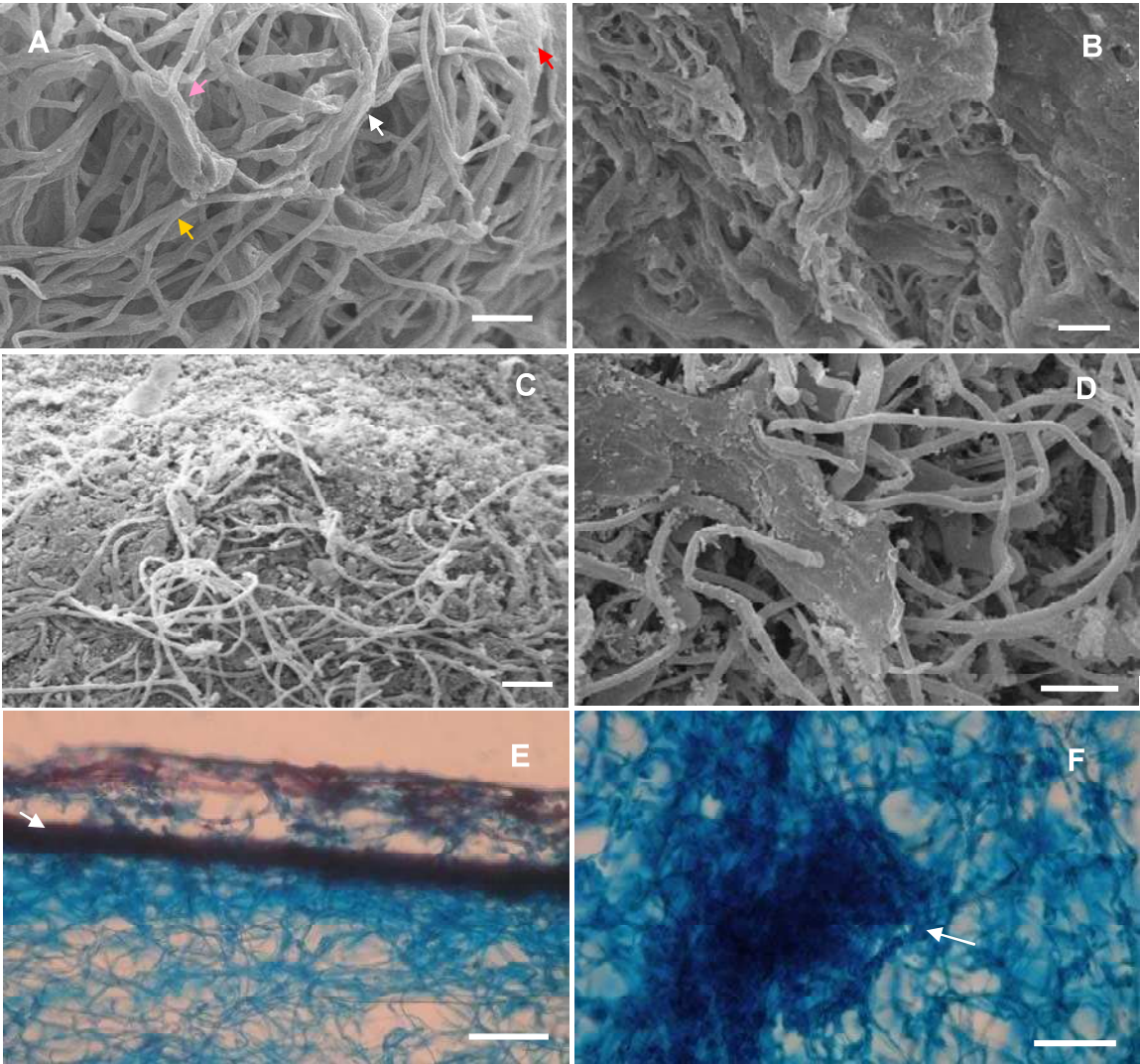


Figure 2

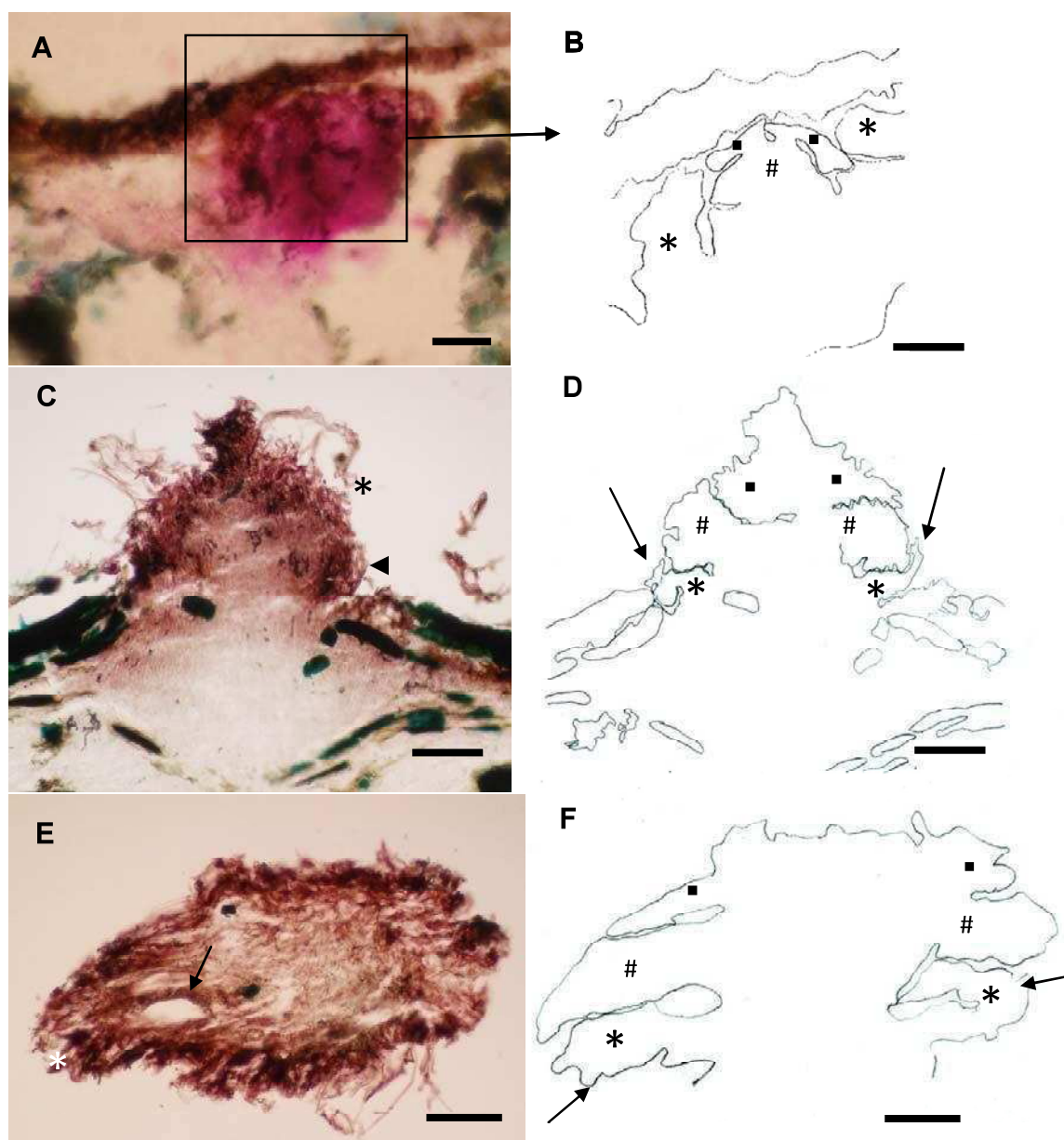


Figure 3

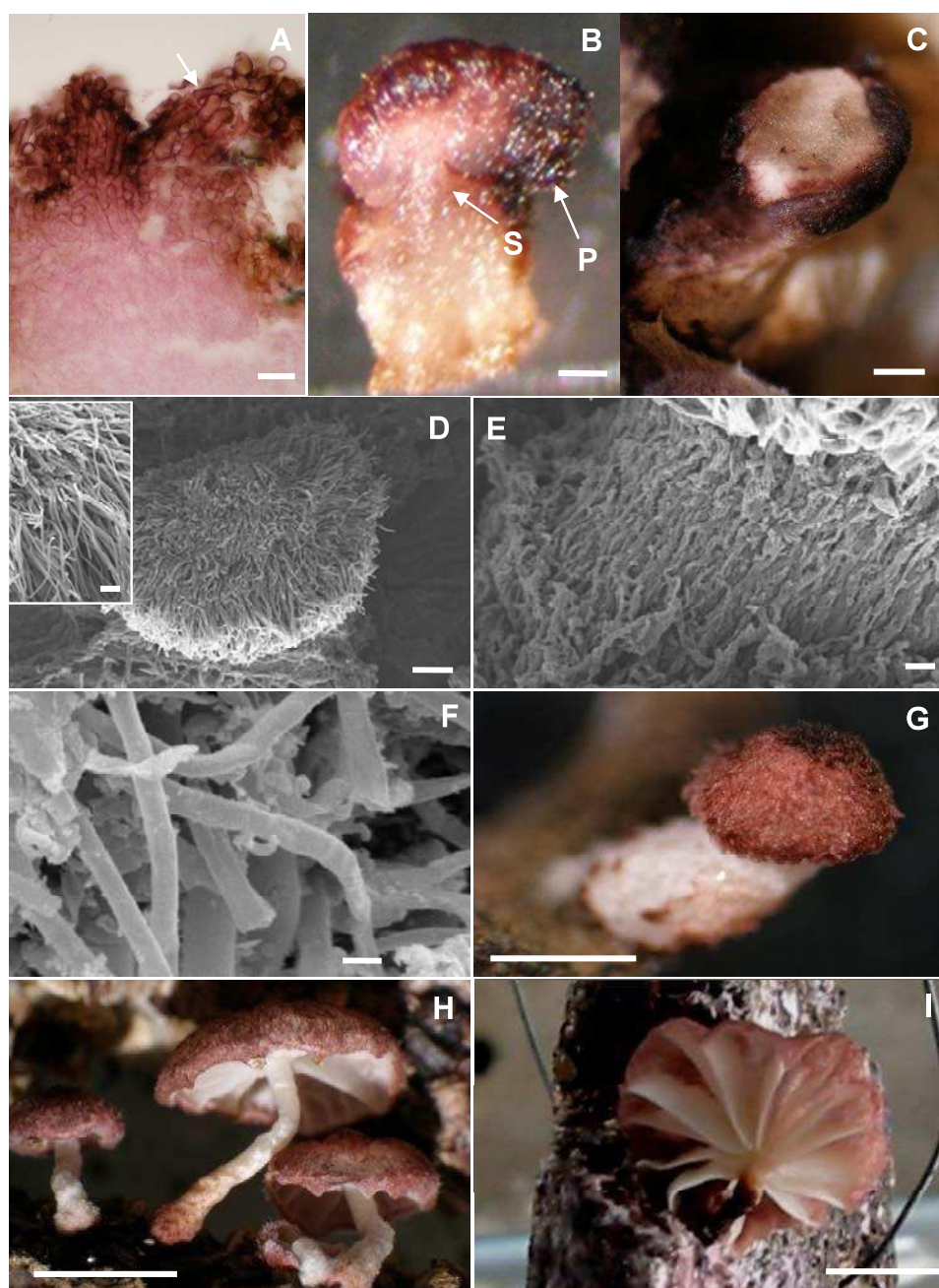


Figure 4

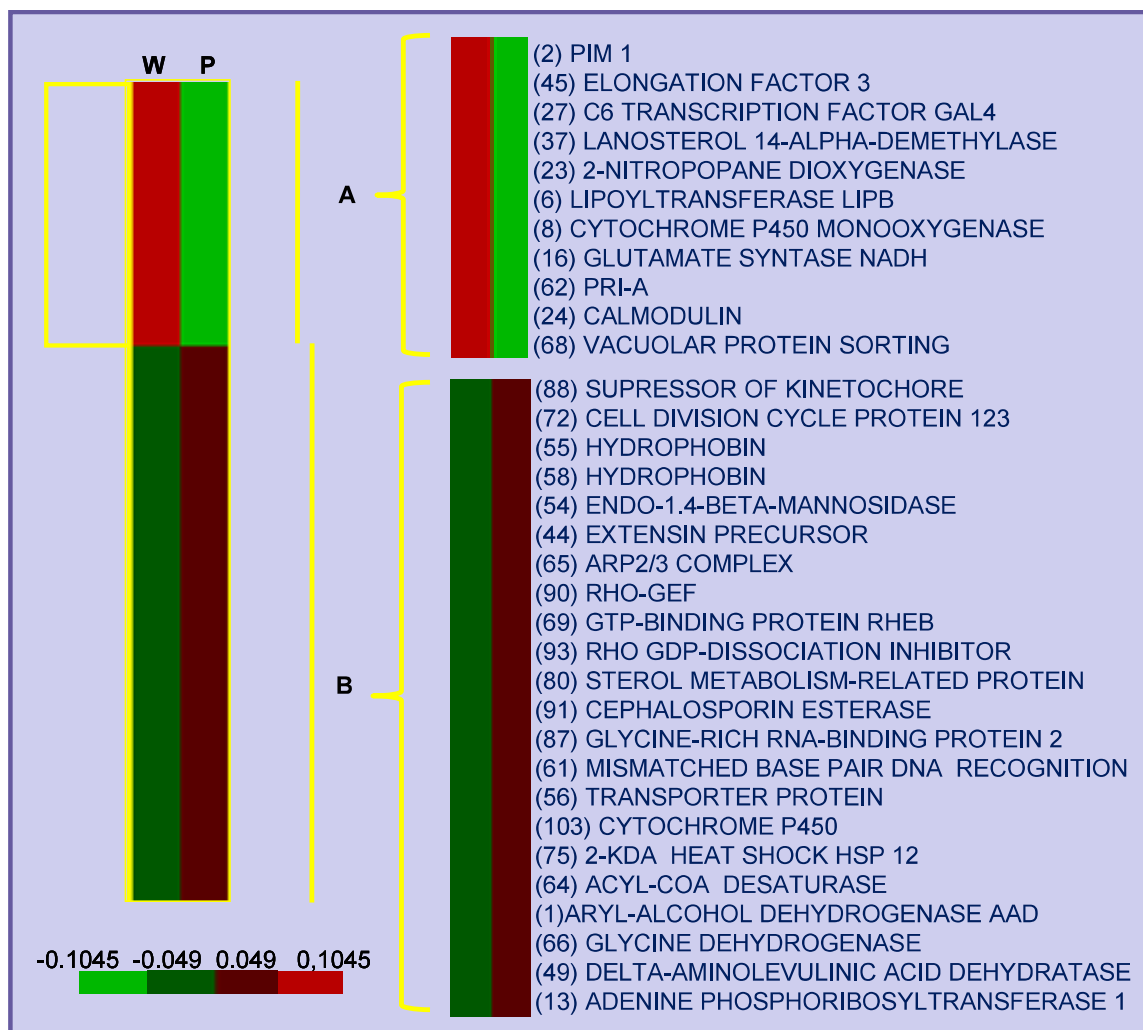


Figure 5

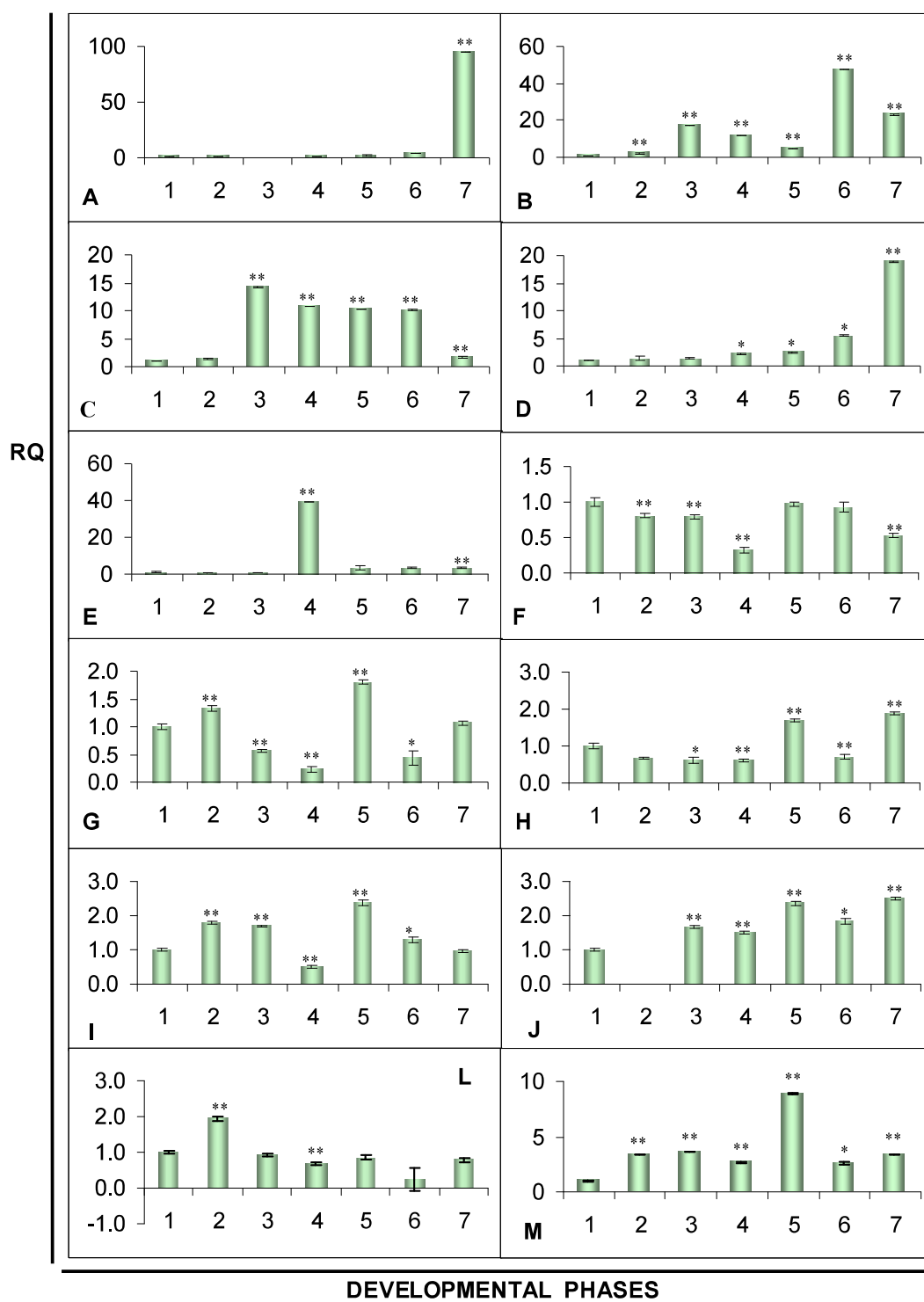
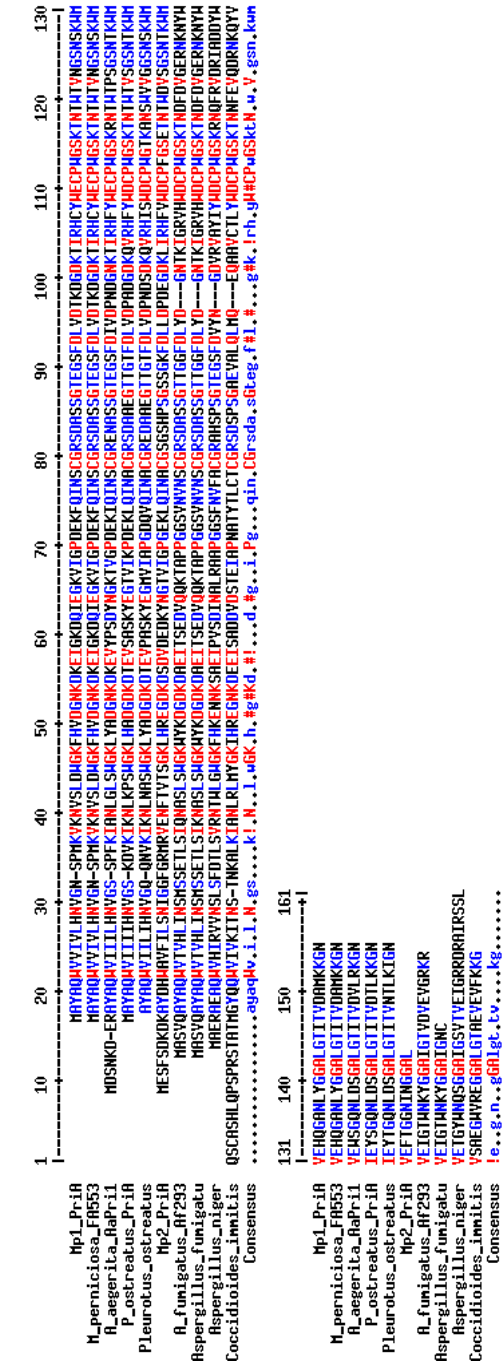


Figure 6

A



B

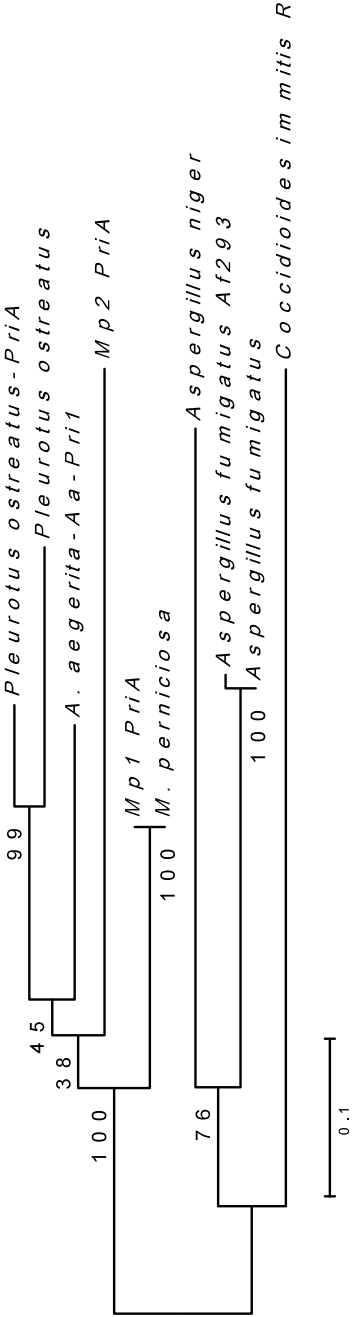
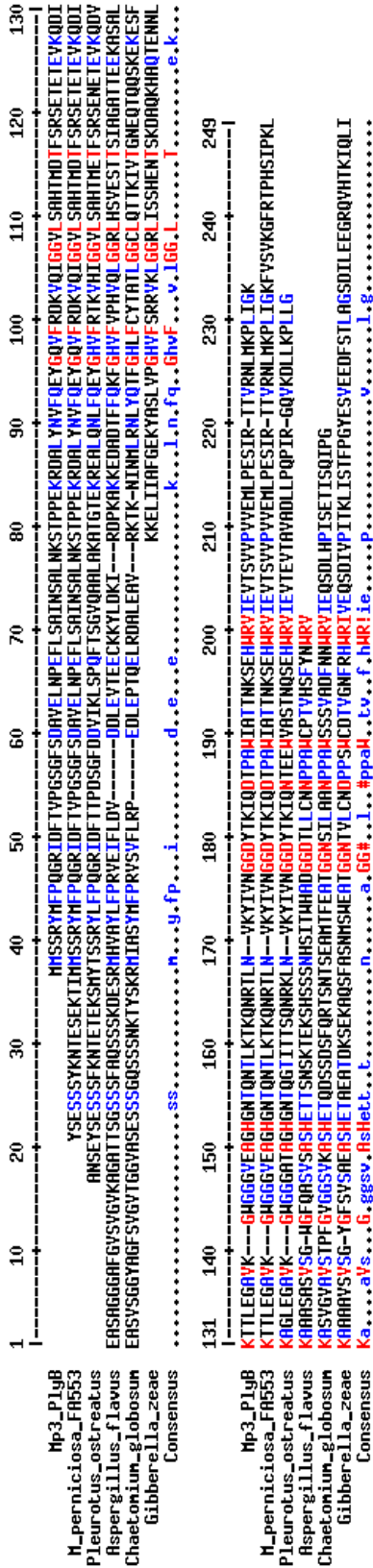


Figure 7

A



B

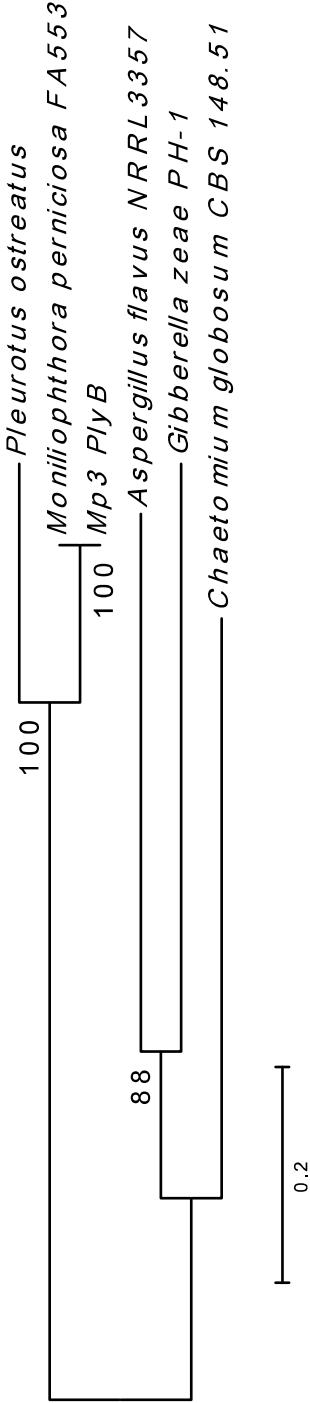


Figure 8

Additional files provided with this submission:

Additional file 1: supplemental table s1.xls, 45K

<http://www.biomedcentral.com/imedia/1518510613295668/supp1.xls>

Additional file 2: supplemental table s2.xls, 20K

<http://www.biomedcentral.com/imedia/4947925629282099/supp2.xls>

**Análise *in silico* de ESTs e abundância de transcritos
presentes no estágio de frutificação de *Moniliophthora*
*perniciosa***

Acássia Benjamim Leal Pires, Marcelo Carazzole, Abelmon Gesteira, Cristiano Villela
Dias, Robson Dias, Ondina Maria Morais Freire, Graziella Dias Rocha, Lucianna
Freitas, Gonçalo Amarante Guimarães Pereira e Júlio César de Mattos Cascardo

Manuscrito em preparação

Análise *in silico* de ESTs e abundância de transcritos presentes no estágio de frutificação de *Moniliophthora perniciosa*

Acássia Benjamim Leal Pires^a, Marcelo Carazzole^b, Abelmon Gesteira^c, Cristiano Villela Dias^a, Robson Dias^a, Ondina Maria Morais Freire^a, Graziella Dias Rocha^a, Lucianna Freitas^a, Júlio César de Mattos Cascardo^a e Gonçalo Amarante Guimarães Pereira^b

^a Centro de Biotecnologia e Genética, Laboratório de Genômica e Expressão Gênica, Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Rodovia Ilhéus-Itabuna, km 16, 45662-000, Ilhéus-Bahia, Brasil.

^b Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, Campinas, São Paulo, Brasil.

^c Empresa Brasileira de Pesquisa Agropecuária – Centro de Mandioca e Fruticultura, Rua da Embrapa, s/n, 44380-000 – Cruz das Almas – BA.

Resumo

Moniliophthora perniciosa é um basidiomiceto hemibiotrófico patogênico ao cacauero e outras espécies vegetais. Sua ação na planta consiste em provocar superbrotamentos, epinastia, hiperplasia e por fim necrose de tecidos, que passam a ser a matéria prima para a produção de estruturas chamadas basidiomas, onde são produzidos seus esporos. A programação genética para a produção dessas estruturas foi estudada por meio da análise de uma biblioteca de cDNA confeccionada a partir de micélio sob condições de indução da frutificação e basidiomas maduros. Foram produzidos 2759 cromatogramas sendo que, destes, 1533 geraram seqüências válidas. Um total de 1374 apresentaram similaridade a seqüências do genoma de *M. perniciosa* e 981 apresentaram similaridade a seqüências do banco de dados NR do GenBank. Foram identificados prováveis genes que codificam para hemolisinas, hidrofobinas, citocromo p450 monoxigenases, ciclofilinas, o-merilstigmato oxigenase, dentre outros, comuns também durante a fase de frutificação em outros fungos. Foram escolhidos 192 ESTs para a análise da sua expressão em seis momentos específicos do micélio antes da frutificação. Destes, 25 tiveram um pico de expressão no micélio com primórdios presentes e 43 foram mais abundantes nesta mesma fase e em basidiomas maduros. Estes transcritos representam genes expressos especificamente em primórdios e basidiomas. A identificação desses genes é importante para compreender a fase reprodutiva do ciclo de vida do *M. perniciosa* e também para estudos genéticos posteriores.

Introdução

Moniliophthora perniciosa (Aime e Phillips-Mora, 2005) é um basidiomiceto hemibiotrófico da família Marasmiaceae (Hibbett et al., 2007) que afeta principalmente cacaueiros, cujos frutos são utilizados para a produção de chocolate. O fungo penetra na planta a partir de hifas originadas de esporos que invadem os espaços intracelulares de tecidos meristemáticos, causa sintomas de epinastia e superbrotamento, e posteriormente torna-se saprofítico, invadindo as células e concluindo seu ciclo de vida com a produção de esporos (Meinhardt et al., 2008). Os galhos resultantes do superbrotamento são conhecidos inicialmente por ‘vassoura verde’ e, após a necrose dos tecidos esses galhos, são chamados de ‘vassoura seca’. Essa doença conhecida por ‘Vassoura-de-bruxa’ causou perdas de até 90% na produção de cacau após a sua introdução no sul da Bahia (Santos Filho et al., 1998) e isso desencadeou uma série de problemas sociais e econômicos nessa região (Trevizan e Marques, 2002). A estimativa de perdas na região Sul da Bahia resultantes da infecção por esse fungo foi de 105 mil toneladas nas safras de 1995/96 e 1996/97 (Santos Filho et al., 1998).

A suscetibilidade dos cacaueiros cultivados no sul da Bahia e as condições climáticas adequadas foram os principais agravantes para a dispersão desse fungo que se dá principalmente com a ajuda do vento e da chuva (Purdy e Schmidt, 1996; Alves et al., 2006). Os basidiomatas, suas estruturas reprodutivas especializadas, embora pequenas (3-5 mm), produzem milhares de esporos (Rocha e Wheeler, 1985). São produzidos abundantemente nas condições climáticas no sul da Bahia, principalmente entre os meses de maio e setembro (Silva et al., 2002). Ocorrem apenas em tecidos já necrosados (Purdy e Dickstein, 1990; Silva et al., 2002), principalmente na parte aérea da planta. Em condições de campo, a produção de basidiocarpos se dá entre 6 a 16 meses após a necrose dos galhos ou frutos (Evans, 1981). Em condições controladas, tecidos necrosados produzem basidiocarpos entre 6 e 12 semanas e

preferencialmente em temperaturas entre 20 e 25°C, 100 $\mu\text{E.m}^{-2}\text{s}^{-1}$ de iluminação e regime hídrico de 8 horas de umidade e 16 horas de seca por dia (Rocha e Wheeler, 1985). O micélio frutifica entre 3 e 8 meses após ser inoculado em meio apropriado e mantido em condições controladas de umidade, luz e temperatura (Griffith e Hedger, 1993; Niella et al., 2000). As condições mais favoráveis à liberação de esporos são temperaturas em torno de 25 °C, 80% de umidade relativa e iluminação em torno de 100 $\mu\text{E.m}^{-2}\text{s}^{-1}$ (Rocha e Wheeler, 1985). Assim, percebe-se que já foram estudadas, em *M. pernicioso*, as condições de umidade, luz e temperatura que permitem a produção de basidiomatas *in vivo* e em condições artificiais, porém ainda não há estudos mostrando como ocorre a regulação genética para a formação dessas estruturas.

A base genética do desenvolvimento de corpos de frutificação tem sido alvo de estudos para alguns basidiomicetos. O uso alimentar ou medicinal de fungos como *Agaricus bisporus*, *Pleurotus ostreatus* e *Coprinopsis cinerea* tem motivado pesquisas para descobrir genes responsáveis pela formação dessas estruturas especializadas, objetivando melhorias em sua produção (Kues e Liu, 2000). Além disso, a compreensão do aumento da complexidade entre os processos de desenvolvimento na escala evolutiva também é uma questão que motiva as pesquisas, uma vez que os fungos são organismos em que já se percebe a tendência de especialização para a formação de tecidos. Isso poderá ajudar a estabelecer modelos de desenvolvimento que são conservados entre eucariotos. Entretanto, o estudo da base genética para a indução da frutificação em fungos fitopatogênicos tem como principal objetivo gerar mais informações que possam ser úteis ao seu controle, ou que possibilitem uma melhoria nas condições de cultivo *in vitro*, e assim facilitem pesquisas visando alternativas de controle. Além disso, o estudo molecular do processo de indução de basidiomatas em fungos

patogênicos poderá também contribuir para a compreensão de processos conservados evolutivamente.

Neste trabalho, foi realizado um levantamento sobre os genes relacionados à frutificação do *M. pernicioso* e uma análise de expressão de 192 candidatos. Uma biblioteca de cDNAs representativa dessa fase de desenvolvimento foi confeccionada e feita a identificação de genes por comparações de seqüências. A seguir, a expressão diferencial de 192 genes foi analisada em cinco momentos específicos desse cultivo.

Material e métodos

1. Isolado e condições de cultivo

O isolado utilizado para a construção da biblioteca foi o CP03 (Micoteca CEPEC-CEPLAC/Brasil n. 1108) obtido a partir de esporos conservados em glicerol a frio (N2 líquido) e crescido inicialmente em meio BDA (Batata Dextrose Agar) a 25°C no escuro. Para a indução de frutificação, 3 discos de micélio foram inoculados em placa de Petri contendo um meio preparado com farelo de trigo, calcário, gesso, vermiculita, terço de cacau e água (Griffith e Hedger, 1993; Niella et al., 1999). Após aproximadamente 30 dias a 25°C, as placas inoculadas foram abertas e o conteúdo (meio entremeado por micélio) foi suspenso por ganchos de metal em uma cabine de vidro (50x30x30 cm) previamente desinfetada com cloro ativo seguido de exposição à luz UV por 40 min. Essa cabine foi aberta em câmara de fluxo laminar duas vezes ao dia para molhar as unidades com água estéril. O fotoperíodo aplicado foi de 12 horas e a temperatura de aproximadamente 25°C. As amostras coletadas foram: 1) micélio no momento da suspensão das placas, quando apresentava a cor branca, 32 dias após a inoculação; 2) micélio com coloração amarela aos 36 dias após a inoculação; 3) micélio com

coloração rosa escuro com 46 dias após inoculação; 4) micélio com coloração rosa escuro, 96 dias após a inoculação, após um corte de sete dias no suprimento de água, 5) micélio com 121 dias após a inoculação, quando alguns basidiomatas já tinham sido coletados e havia a presença de vários primórdios e 6) basidiomatas destacados do meio.

Outro isolado, CP02, o mesmo utilizado para o sequenciamento do genoma, também foi cultivado posteriormente como descrito acima, para extração de RNA necessário para a confecção de um macroarranjo. Porém, as datas de coleta variaram: 1) micélio no momento da suspensão das placas, quando aparentava a cor branca, 32 dias após a inoculação; 2) micélio com coloração amarela aos 45 dias após a inoculação; 3) micélio com coloração rosa escura com 65 dias após inoculação; 4) micélio rosa escuro, 79 dias após a inoculação, imediatamente antes de um corte de dez dias no suprimento de água, 5) micélio com 89 dias após a inoculação, imediatamente após o estresse por falta d'água por 10 dias; 6) micélio com 121 dias após a inoculação, quando alguns basidiomatas já tinham sido coletados e havia a presença de vários primórdios e 7) basidiomatas destacados do meio.

2. Isolamento de RNA e construção da biblioteca

A extração de RNA do isolado CP03 foi feita a partir das amostras congeladas a 80°C negativos utilizando o kit RNeasy Plant Mini de acordo com os procedimentos do fabricante (Qiagen, USA, Valencia, CA). A análise qualitativa das amostras de RNA obtidas foi feita em gel desnaturante com formamida/folmaldeído (Sambrook e Russell, 2001) e a quantificação foi feita em espectrofotômetro UV-visível modelo Cary® 100, da Varian, Palo Alto, CA, USA. Concentrações equivalentes de RNA de cada amostra foram misturadas e uma alíquota de 1µg dessa mistura contendo RNA total foi utilizada para construir uma biblioteca usando o kit DB SMART Creator cDNA library (Clontech). Os cDNAs maiores que 400 pb, selecionados por

filtração em colunas Chroma Spin 4000 (Clontech) foram clonados direcionalmente no plasmídeo pDNR-LIB e utilizados para transformar células competentes (EletroMax™ DH10B™ - Invitrogen). Colônias isoladas foram crescidas em microplacas de 96 poços em LB (meio de cultivo Lurian-Bertani) com 40% de glicerol e $30 \mu\text{g.L}^{-1}$ e estocadas em 80°C negativos.

3. Extração de DNA plasmidial e sequenciamento

DNA plasmidial foi extraído de clones individuais usando o procedimento de lise alcalina (Sambrook and Russell, 2001) adaptado para microplacas de 96 poços. A qualidade e quantidade do DNA plasmidial foi observada em gel de agarose a 0,8% em TBE corado com brometo de etídio (0,25ug/ul). Um total de 2759 clones foram selecionados ao acaso e seqüenciados na extremidade 5' utilizando o kit DyEnamic ET Dye Terminator (MegaBACE – GE Healthcare) e o primer M13 'forward' 5'-TAAAACGACGGCCAGT-3'. As seqüências foram produzidas em seqüenciador MegaBACE 1000 (GE Healthcare).

4. Processamento dos dados e análise computacional

Os cromatogramas obtidos foram depositados no banco de dados do projeto genoma de *M. pernicioso* (Mondego et al., 2008) e nomeados de acordo com o sistema criado para esse projeto no qual são considerados o isolado, o nome da biblioteca, o número da biblioteca, o número da placa e o número do poço na placa. O primeiro tratamento dos dados foi a utilização do algoritmo PHRED (Ewing et al., 1998) que nomeia e faz o controle de qualidade das bases. A seguir o Cross-Match foi utilizado para mascarar a presença de 'polylinker' e vetor. As seqüências foram, então, comparadas ao banco de dados não redundantes do NCBI usando o algoritmo BlastX e também com o banco de dados do Projeto Genoma *M.*

perniciosa (Mondego et al., 2008) usando o algoritmo BlastN (Altschul et al., 1997). As seqüências que foram utilizadas para a filtragem e agrupamento por similaridade foram as que obedeciam aos critérios de qualidade: a) mais de 200 pares de bases com qualidade superior a 20 atribuído pelo PHRED, b) significativa similaridade ao banco de dados não redundantes do NCBI (BlastX com E-value $\leq 1e \times 10^{-05}$), c) significativa similaridade de nucleotídeos com o banco de dados do genoma de *M. perniciosa* (BlastN com E-value $\leq 1e \times 10^{-10}$). Essas seqüências selecionadas foram limpas ('trimmed'), como descrito por Telles and Silva 2001 e agrupadas por similaridade ('clustering') usando o programa CAP3 (Huang e Madan, 1999). Os agrupamentos de seqüências curtas (≤ 100 pb) e com artefatos foram removidos (Telles et al., 2001). Foi feita em seguida uma anotação manual usando comparações com os bancos de dados públicos (NCBI) e um banco de dados de fungos. Os agrupamentos que sobraram foram submetidos novamente a comparações com o banco de dados não redundantes do NCBI (BlastX), o banco de dados do genoma de *M. perniciosa*, de *Laccaria bicolor* e *Ustilago maydis* (BlastN) (Altschul et al., 1997). Foi feita também uma comparação com outro banco de ESTs de *M. perniciosa* (Rincones et al., 2008) obtidos de amostras de fungo cultivados em condições de micélio hemibiotrófico e saprofítico, mas não em estágio de frutificação. Os resultados da busca de similaridade de cada agrupamento foram disponibilizados em um portal que possui acessos para a página de anotação dentro do Projeto Genoma de *M. perniciosa*. Os unigenes foram anotados manualmente e classificados de acordo com Journet e colaboradores (2002).

5. Confecção do macroarranjo

Fragmentos de DNA representativos de 192 genes foram utilizados para construir um macroarranjo. Os fragmentos de DNA foram obtidos por amplificação a partir de minipreps de

clones da biblioteca de cDNA não normalizada, representativa da fase de frutificação. Os critérios utilizados para a seleção destes clones foram: 1) clones que representam genes constitutivos; 2) clones mais freqüentes; 3) clones que representam genes possivelmente envolvidos com o processo de frutificação; 4) clones que representam genes de funções desconhecidas, mas que foram freqüentes na biblioteca.

Os insertos destes clones foram amplificados utilizando o par de ‘primers’ para um fragmento do M13 contido no vetor de clonagem pDNR-LIB (Clontech), M13 ‘Forward’ (5’-CAGGAAACAGCTATGAC-3’) e o M13 ‘Reverse’ (5’-GTAAAACGACGGCCAG-3’). Os componentes da reação de PCR foram: 1X tampão para PCR (Fermentas), 1,2 mM MgCl₂, 0,15 mM dNTPs, 0,1 µM primer ‘forward’, 0,1 µM primer ‘reverse’, 1 unidade de *Taq* (Fermentas) e 6 µL de DNA miniprep diluído 100x, num volume final de 50 µL. As reações foram realizadas em placa de 96 poços e transferidas ao termociclador com o seguinte programa: desnaturação a 94°C por 4 minutos, seguido de 35 ciclos de 40 seg. a 94°C, 40 seg. a 55°C e 90 seg. a 72°C. Para finalizar, a reação foi incubada a 72°C por 7 min. para polimerização final. Os insertos amplificados por PCR foram visualizados em gel de 1% agarose contendo 0,5 µg/ml de brometo de etídio e aplicados diretamente na membrana, sem quantificação, na quantidade total de 7,5 uL do produto amplificado. A desnaturação foi feita com NaOH 0,2M a 37 °C por 15 minutos. Os insertos foram fixados em membrana de nylon (Hybond N+, Amersham Biosciences) utilizando um replicador de 96 canaletas seguindo um protocolo padrão de transferência em solução salina tipo ‘dot blot’ (Sambrook e Russell, 2001). O DNA foi ligado covalentemente à membrana por luz UV, 0,15 J/cm² (UV CrossLinker). A membrana contendo os insertos amplificados foi hibridizada contra cinco amostras de cDNA obtidos das seguintes fases do desenvolvimento do *M. perniciosus* isolado CP02 (mesmo utilizado para o projeto genoma): 1) Micélio crescido em meio para

frutificação, coletado no primeiro dia da exposição do micélio nas câmaras de crescimento (fase branca, 32 dias após o início do cultivo); 2) Micélio rosa avermelhado coletado 45 dias após o início do cultivo; 3) Micélio coletado após 79 dias do início do cultivo, após um estresse de 10 dias sem água; 4) Micélio coletado na fase de produção de primórdios (final do cultivo); 5) Basidiomatas maduros coletados na fase final do cultivo. As amostras de RNA total obtidas de cada uma dessas fases foram utilizadas para síntese da primeira fita de cDNA pela utilização da transcriptase reversa (Invitrogen) e marcação simultânea da fita sintetizada com a fosfatase alcalina seguindo a metodologia recomendada pelo fabricante do kit Alkphos para preparo de sondas e revelação (GE Healthcare). As imagens foram capturadas no Image Scanner Power Look 1120 UDS (Amersham Biosciences) e analisadas pelo programa BZ Scan. Cerca de 400 ng de cada sonda (quantificadas por espectrofotometria) foram marcadas utilizando hexa-oligonucleotídeos (Amersham Bioscience) e a enzima Klenow (Amersham Biosciences), como descrito por Feinberg e Vogelstein (1983). Para a hibridação, foi utilizado um tampão contendo 10% sulfato de dextran (Calbiochem), 2X SSPE, 1% SDS, 0,5% leite em pó desnatado e 0,1 mg/mL DNA de esperma de salmão. As membranas foram pré-hibridadas nesta solução a 55° C por 4 horas. A hibridação foi conduzida a 55°C ‘overnight’; e, posteriormente, as membranas foram lavadas segundo instruções do fabricante e expostas a filmes radiográficos X-OMAT (Kodak) durante 2,5 horas. A presença ou ausência de sinal, assim como sua intensidade, foi registrada para cada ponto com cada uma das sondas. Foi feita uma normalização global das intensidades geradas com o auxílio do *software* Gene Cluster versão 3.0 (Eisen et al., 1998), seguindo o protocolo padrão do programa, ou seja, uma centralização de genes e arranjos pela média repetindo-se 8 vezes cada uma e um agrupamento hierárquico global utilizando a correlação de Pearson. As medidas de diferenças de expressão

foram feitas a partir dos dados obtidos por essa normalização e agrupamento, tomadas em relação à fase ‘branca’ do micélio.

Resultados

Biblioteca de cDNAs representativa da fase de frutificação do fungo

A biblioteca de cDNA construída é representativa do fungo exposto a condições artificiais de indução da frutificação. A intenção foi identificar novos genes e outros relacionados a essa fase do seu ciclo de vida bem como auxiliar na investigação de elementos necessários para a predição de genes das seqüências obtidas no rascunho do Genoma estrutural de *M. pernicioso*. Um total de 2759 fragmentos de seqüências expressas foi gerado e a sua análise encontra-se resumida na Figura 1. Um total de 1.533 unigenes foi identificado, ou seja, 55,6% das seqüências produzidas. Isso representa aproximadamente 11% do total de 14.072 genes identificados no Genoma rascunhado de *M. pernicioso* (Mondego et al., 2008). As seqüências dos unigenes foram anotadas manualmente e serão submetidas ao banco de dados DB EST do GenBank posteriormente. O número de seqüências redundantes pode ser observado na Figura 2. O número máximo de seqüências por unigene composto (‘contig’) foi 14 em um apenas, sendo que 122 unigenes compostos apresentaram 2 seqüências, 31 tiveram 3 seqüências e 28 tiveram entre 4 e 9 seqüências. A redundância foi de 44,4%, indicando que a biblioteca um número representativo de clones da biblioteca foi seqüenciado. A anotação funcional para identificar as prováveis proteínas codificadas foi conduzida utilizando comparações por BLAST nos bancos de dados do NCBI (non redundant BLAST), Gene Ontology (GO), SwissProt e Banco de Dados de Genomas de Fungos. Os genes identificados foram agrupados em categorias funcionais como descrito por Journet e colaboradores (2002).

Uma lista com os unigenes classificados por prováveis funções pode ser vista na Tabela 1 (anexo I). A maior parte das prováveis funções encontradas foi agrupada como “desconhecidas” (Figura 3), representando 24,13% daqueles que tiveram alguma similaridade com seqüências depositadas nos bancos públicos. O metabolismo primário foi associado a 20,24% das seqüências e síntese protéica e processamento a 18,58%. Em seguida, expressão gênica e metabolismo de RNA foram representados em 5,8% das seqüências e o estímulo abiótico e desenvolvimento, 5,21%.

Os dez genes cujas *seqüências* foram mais abundantes na biblioteca estão na Tabela 2, codificando para prováveis subunidades ribossomais menor e maior, duas prováveis hidrofobinas, uma pequena proteína de choque térmico, uma MBF (*multi-protein-bridging factor*), uma mioinositol oxigenase, um proteolípido de membrana, uma ubiquitina e uma aegerolisina. Desses, hidrofobinas e aegerolisinas são relatadas ao desenvolvimento de basidiomatas (Kues e Liu, 2000), a pequena proteína de choque térmico é relatada em resposta a estresses, o coativador transcricional MBF à produção de pigmento via indução de lacase (Walton et al., 2005) e a mioinositol oxigenase à produção de ácido ascórbico (Lorence et al., 2004).

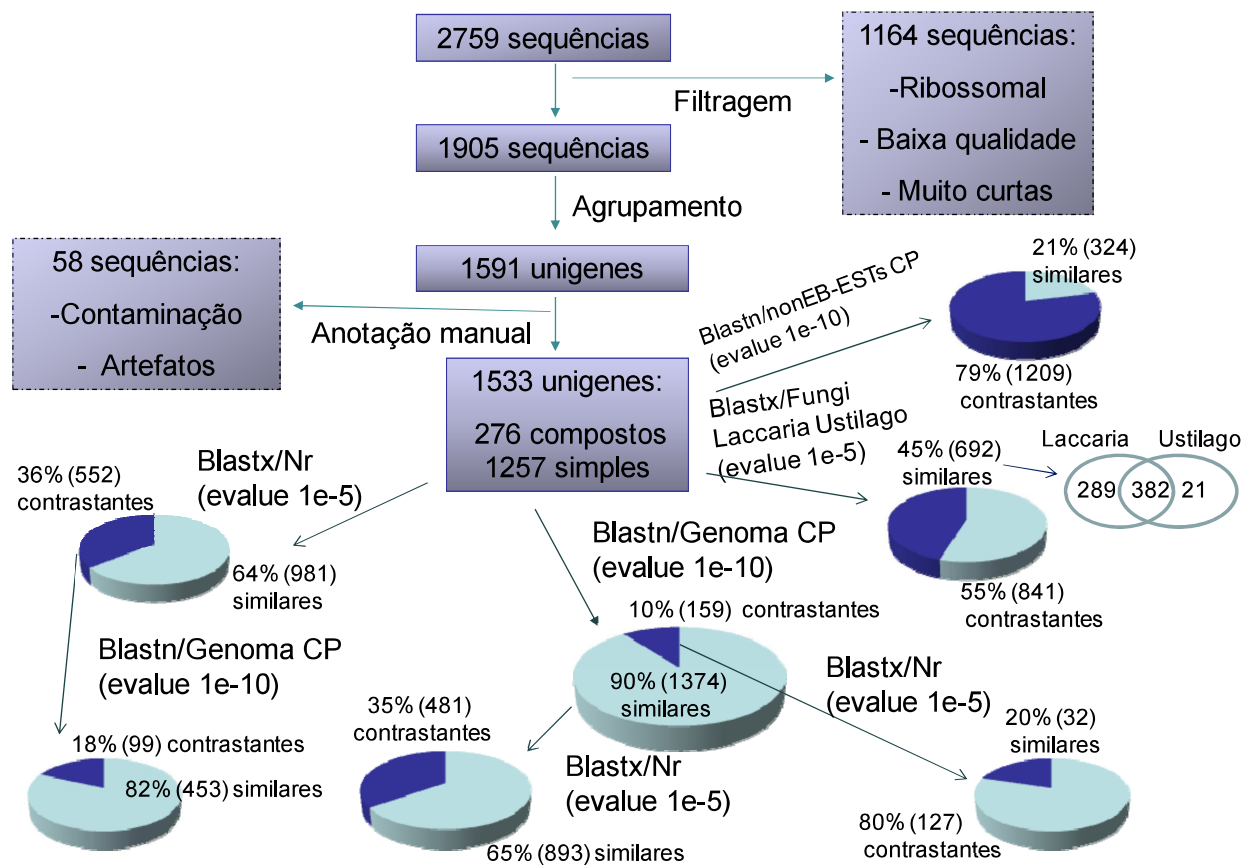


Figura 1 - Análise da biblioteca de cDNA representativa da fase de frutificação de *M. perniciosa*. As 2.759 seqüências representam os cromatogramas submetidos à análise. As setas indicam os processos seguidos para selecionar as seqüências que representam os unigenes. Foram realizadas filtragem, agrupamento, comparação de seqüências pelo BlastX e BlastN e anotação baseada em GO e SwissProt.

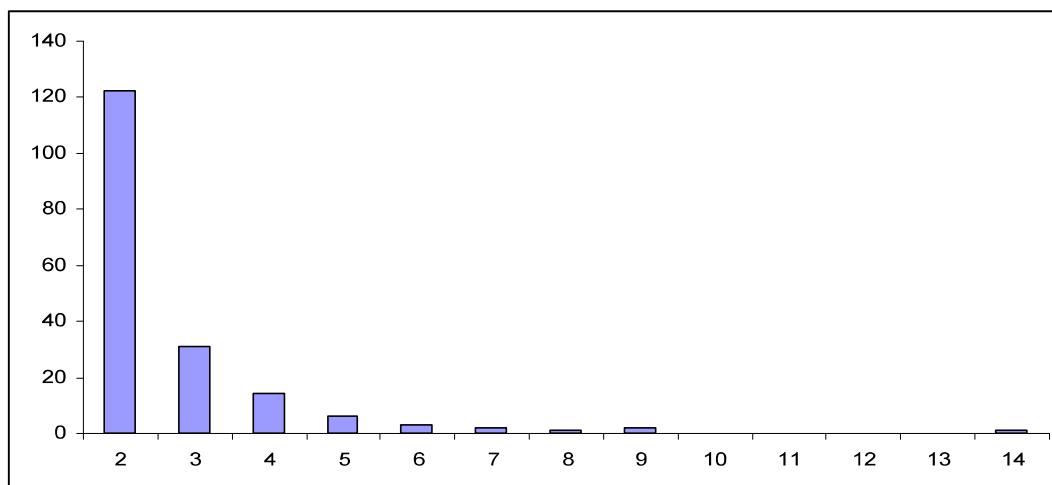


Figura 2 – Distribuição do número de seqüências que formaram os 276 unigenes compostos. A maioria (121) foi formada pelo agrupamento de 2 seqüências e o número máximo de seqüências por agrupamento foi 14.

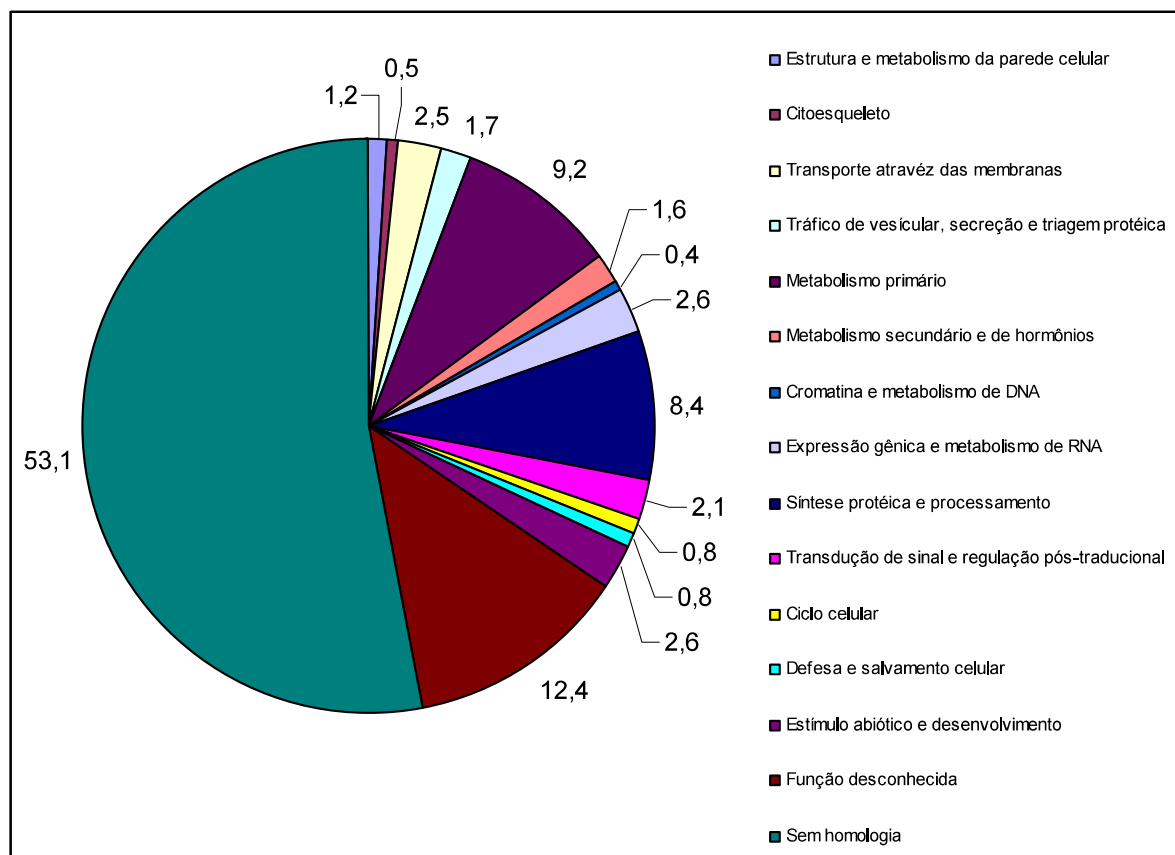


Figura 3 – Distribuição por classes de função dos 981 unigenes que apresentaram similaridades a seqüências do GenBank. A maioria não possui função conhecida ou não apresentou similaridade a nenhuma seqüência do banco de dados.

Tabela 2 – Unigenes que apresentaram maior frequência e suas prováveis funções

Ref. na Biblioteca	N. Sequências	Provável função	Organismo homólogo	Tamanho (pb)	E-value
Contig458	5	Aegerolisina Aa-Pri1	<i>Agrocybe aegerita</i>	518pb	3,00E-56
Contig448	5	Ubiquitina/RPS27a	<i>Aspergillus fumigatus</i>	639pb	1,00E-75
Contig618	5	Pequena proteína de choque térmico	<i>Laccaria bicolor</i>	571pb	2,00E-48
Contig111	5	Subunidade pequena de RNA ribossomal	<i>Lentinula edodes</i>	881pb	1,00E-88
Contig451	5	Proteolípido 3 da membrana plasmática	<i>Neurospora crassa</i>	394pb	6,00E-22
Contig408	5	Proteína expressa	<>	582pb	<>
Contig604	6	Mio-inositol oxigenase	<i>Aspergillus oryzae</i>	837pb	7,00E-77
Contig643	6	MBF (<i>multiprotein-bridging factor1</i>)	<i>Kluyveromyces lactis</i>	643pb	5,00E-37
Contig423	6	Pequena proteína de choque térmico	<i>Laccaria bicolor</i>	602pb	3,00E-40
Contig415	7	Hidrofobina SC3	<i>Coprinopsis cinerea</i>	667pb	3,00E-23
Contig622	7	Proteína expressa	<>	581pb	<>
Contig85	8	Subunidade maior RNA ribossomal mitocondrial	<i>Suillus sinuspaulianus</i>	2051pb	1,00E-135
Contig581	9	Hidrofobina	<i>Coprinopsis cinerea</i>	566pb	2,00E-23
Contig488	9	Proteína expressa	<>	653pb	<>
Contig348	14	Subunidade pequena de RNA ribossomal ou TAR1p	<i>Moniliophthora roreri</i>	390pb	0,0

Outros genes potencialmente relacionados ao desenvolvimento de basidiomas foram identificados por análise *in silico* da biblioteca. Entre os classificados na função ‘estímulo abiótico e desenvolvimento’ (Tabela 1, em anexo), estão, além dos que codificam para aegerolisinas, aqueles que codificam para citocromo P450 monoxigenase, NADPH-citocromo P450 redutase, pleurotolisina B, proteínas de choque térmico e septinas (Berne et al., 2007; Kues e Liu, 2000). Há alguns prováveis fatores de transcrição que são candidatos interessantes para estudos posteriores, BTF3, NFX tipo dedos de zinco, C6 e o CCR4-NOT. Os genes relacionados ao metabolismo primário e secundário podem também estar associados à produção de basidiomatas, uma vez que a formação destas estruturas demanda energia e produção de compostos específicos (Kues, 2000). Alguns envolvidos na produção de ácidos graxos podem ter função na produção de lipídeos específicos de membrana: delta 9 –

desaturase de ácidos graxos, delta 12 – desidrogenase de ácidos graxos, oxigenase de ácidos graxos e sintase de ácidos graxos. Outros do metabolismo de glicose podem ser ativados para a produção de reservas dos basidiomas como trealose fosfato sintase, malato desidrogenase e gliceraldeído 3 fosfato desidrogenase. E outros pertencentes a essas classes de funções podem ser bons candidatos para análises posteriores.

Os transcritos dessa biblioteca foram também comparados *in silico* com um banco de transcritos de *M. perniciosa* obtidos de bibliotecas não normalizadas do fungo em condição saprofítica e semibiotrófica, feitas utilizando o mesmo kit (cDNA SMART Library – Clontech) (Rincones et al., 2008). Foram identificados 1.209 unigenes que não apresentaram similaridades com as seqüências das bibliotecas citadas, sendo assim considerados únicos da fase de frutificação do fungo. Dentre esses, estão os genes codificadores para aegerolisinas, pleorotolisinas B, proteína antígeno1 associada ao esperma, pim1, sorting nexin 3, riboflavina quinase, fator 1 de alternância de guanina da Rho1, inibidor da dissociação de GTP da Rho1, coactosina, proteína de reconhecimento de bases trocadas e DNA cruciforme (*mismatched base pair and cruciform DNA recognition protein*), proteína DIF1, ATPases AAA (*ATPase family associated with various cellular activities*), adenilato ciclase, várias outras de função conservada porém não identificadas (*conserved expressed protein*) e várias seqüências com ORFs não identificadas (*expressed protein*).

Outra análise *in silico* foi feita por BlastN utilizando a seqüência fasta para comparar com o banco de dados do genoma de *Laccaria bicolor* e *Ustilago maydis* (Tabela 3, anexo II). Esses dois fungos são basidiomicetos também, porém *L. bicolor* pertence ao subfilo Agaricomycotina, que produz basidiomatas, enquanto *U. maydis* pertence ao subfilo Ustilacomycotina, os quais não produzem basidiomatas. A análise de similaridades dos unigenes montados a partir da biblioteca com os dados desses genomas mostrou que 692

unigenes foram similares nos dois bancos de dados, sendo que, destes, 382 apresentaram similaridade em ambos, 289 foram específicos a *L. bicolor* e 21 específicos a *U. maydis*. Assim, os unigenes que apresentaram similaridade apenas a genes de *L. bicolor* são potencialmente relacionados à formação de basidiomatas. Entre eles, os mais frequentes foram aqueles que codificam para hidrofobinas, proteínas de choque térmico e proteínas ribossomais. Alguns genes relacionados à regulação do ciclo celular estiveram entre os específicos a *L. bicolor*, especificamente os unigenes que codificam para uma provável proteína do tipo homeodomínio, uma provável proteína de cruzamento sexual (‘mating-type’), uma provável proteína de transformação viral Jun, e duas prováveis proteínas contendo domínios do tipo ‘dedos de zinco’. Outros também parecem ser bons candidatos para análises, os unigenes que codificam para o MBF (‘multiprotein-bridging factor’), para uma o-metil-estigmato-cisteína, para uma lisofosfolipase e uma desidrogenase de ácidos graxos delta.

Análise da expressão diferencial de 192 ESTs por macroarranjo

No intuito de analisar a expressão diferencial de genes candidatos à regulação da diferenciação de micélio secundário do *M. perniciosus* em suas estruturas reprodutivas, foram escolhidos 192 candidatos, identificados na biblioteca acima citada, para uma análise de abundância de transcritos detectada por macroarranjo, utilizando como sondas amostras de RNA coletadas durante o cultivo do fungo em condições de frutificação. Parte dessa análise foi apresentada no artigo publicado recentemente (Pires et al., 2009). Entretanto, agora, aquela análise foi acrescida de mais amostras coletadas nas fases de micélio ‘rosa’, ‘após o estresse’ e ‘basidiomatas’. Os 192 transcritos analisados apresentaram uma expressão diferenciada nos cinco momentos distintos do cultivo do fungo sob condições de indução de frutificação. O agrupamento pelo método de correlação de Pearson dos dados normalizados, feito com o

auxílio do programa Gene Cluster, produziu 6 perfis distintos de expressão (Figura 4). No grupo I, os genes apresentaram um aumento de transcritos quando os primórdios estavam presentes e em basidiomatas. O grupo II foi formado por genes que apresentaram um pico de expressão bem distinto no momento em que o micélio produziu primórdios (grupo II). No grupo III houve um aumento de expressão dos genes no micélio que foi coletado após o estresse. Outros genes (grupo IV) diminuíram a sua expressão na fase de micélio rosa avermelhado e de micélio após o estresse hídrico, voltaram a aumentar quando os primórdios apareceram e diminuíram novamente em basidiomatas. No grupo V, os genes apresentaram uma diminuição da expressão na fase do micélio rosa avermelhado, uma elevação de transcritos no micélio após o estresse e em micélio com primórdios, mas a expressão foi bem diminuída em basidiomatas. Um último grupo de genes (grupo VI) apresentou dois picos de expressão, um na fase de micélio rosa avermelhado e o outro na presença de primórdios. Os valores de expressão para cada um dos genes desse macroarranjo encontram-se na Tabela 4 (anexo III).

Um gene para aegerolisina (137) apresentou um pico de expressão na fase de primórdios, porém outros ESTs que provavelmente codificam para proteínas dessa família (38, 62, 109) continuaram presentes em basidiomatas depois que foram induzidas em primórdios (grupo I). Dois representantes desses genes (118 e 178) foram presentes no agrupamento V e um foi detectado unicamente na fase rosa (177). Isso indica que essas seqüências correspondem a genes distintos. Uma análise detalhada desses ESTs por alinhamento de seqüências e PCR quantitativo indica que há pelo menos 2 genes distintos (Pires et al., 2009). O gene que codifica para a provável pleurotolisina B é ativo unicamente em primórdios, mas seu precursor está ativo na fase branca, após o estresse e em primórdios.

Alguns genes foram altamente expressos em basidiomatas maduros, entre eles uma sequência conservada, mas de função desconhecida, um gene para transportador de fosfato inorgânico, um para proteína da família SCAMP, um para PIM1, outro para proteína DIF1 e outro para feruloil esterase B. Os mais altamente expressos na fase branca foram: um gene para septina, outro para citocromo p450 e outro para uma proteína conservada de função desconhecida. Um gene codificando para a provável proteína TAR foi altamente expresso após o estresse.

A expressão de hidrofobinas foi variável, indicando que as seqüências depositadas na membrana representam genes diferentes. Foram detectados 6 perfis distintos, um expresso unicamente em primórdios (55), outro com aumento de transcritos nas fases branca, após o estresse e primórdios (58), outro na fase branca, primórdios e basidiomas (149), outro na fase branca e de primórdios (139), outro na fase branca e após o estresse (154) e ainda outro na fase rosa e de primórdios (111).

A inibição da transcrição na fase de primórdios ocorreu para alguns genes que provavelmente codificam para ornitina aminotransferase, transportador de glicose, inibidor da dissociação de GTP da Rho1, transportador de gluconato, uma endonuclease reparadora de dano por UV, glutathione dissulfite redutase, proteína de divisão celular 123, um fator de transcrição NFX. Esses também podem ser candidatos para análises posteriores, pois indicam que devem ser inibidos para que a frutificação ocorra.

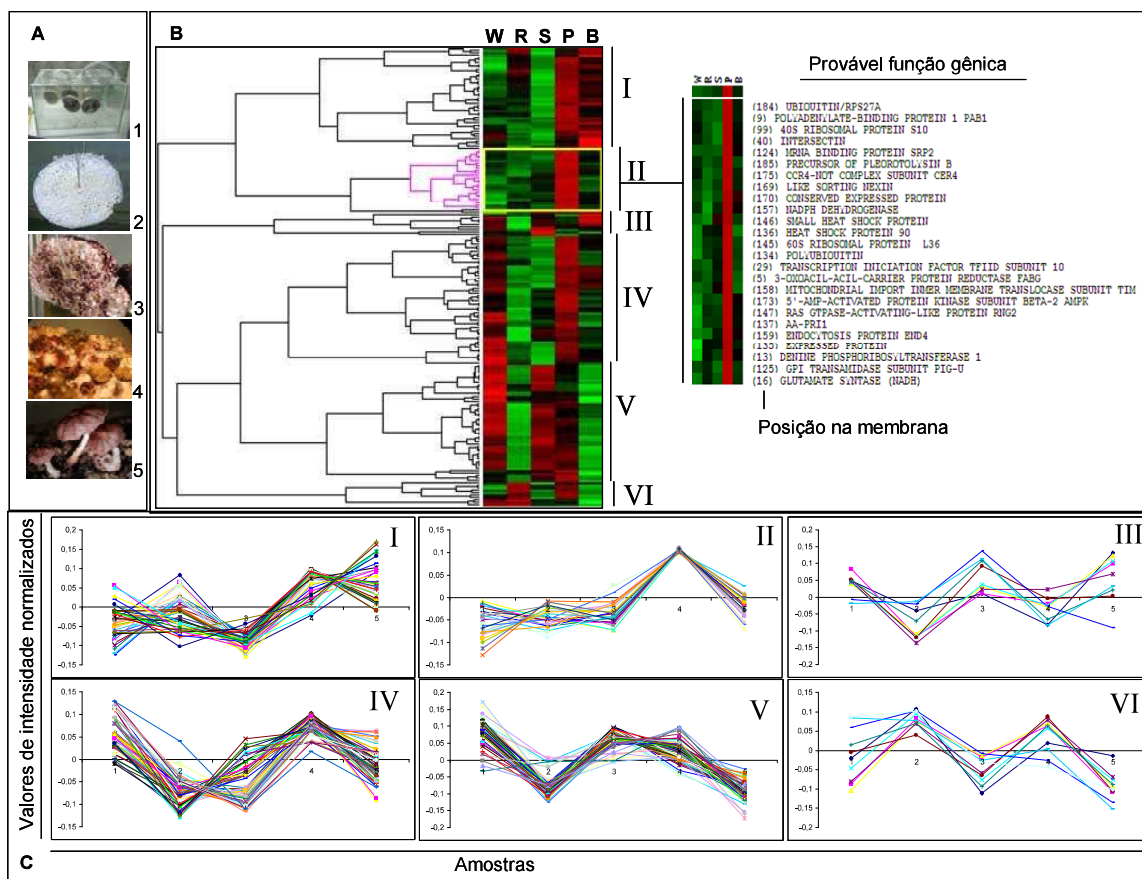


Figura 4. Expressão diferencial de genes em *M. perniciosus* sob condições de frutificação *in vitro*. A. Amostras coletadas durante o cultivo do isolado CP02. A1- Cabine de crescimento contendo micélio fúngico no estágio ‘branco’; A2- Micélio do fungo em estágio amarelo; A3- micélio do fungo em estágio rosa avermelhado; A4- micélio contendo primórdios; A5- basidiomatas. B. Ilustração do agrupamento hierárquico de 192 genes expressos diferencialmente no fungo quando amostrado nos estágios 1 a 5. Em destaque, o agrupamento de genes que apresentou um pico de expressão na presença de primórdios. C. Perfis de expressão dos grupos produzidos pela hierarquização feita em B.

Discussão

Diversos estudos têm sido feitos para entender como o basidiomiceto *M. perniciosa* causa a doença Vassoura-de-Bruixa em cacaueiros. Embora a fase biotrófica do fungo seja a principal responsável pela doença, o patógeno é disseminado unicamente por esporos. Portanto, entender os mecanismos envolvidos em sua reprodução sexual é fundamental para o entendimento completo da doença.

A identificação de genes específicos da fase de frutificação já foi feita para alguns basidiomicetos, principalmente os comestíveis como *Agaricus bisporus*, *Agrocybe cylindracea*, *Pleurotus ostreatus*, *Schizophyllum commune* e *Lentinula edodes* (De Groot et al., 1997; Shim et al., 2006; Kues and Liu, 2000). No caso do *M. perniciosa*, houve interesses múltiplos para a identificação de genes dessa fase. Obter transcritos específicos do fungo seria fundamental para a identificação de genes do seu genoma. Além disso, alguns desses genes podem ser alvos para futuros bloqueios, de forma a impedir o progresso da doença ou para estimulação, para melhorar a produção de esporos em laboratório. A descrição de novos genes ou a utilização de genes com interesses biotecnológicos também foram motivadores para esse levantamento.

Análise ‘*in silico*’ dos transcritos obtidos

A caracterização de ESTs (expressed sequence tags) originadas de bibliotecas de cDNA tem sido útil para a identificação de novos genes em diversos fungos (Ebbole et al., 2004; Maeda et al., 2004; Peter et al., 2003). A biblioteca de ESTs gerada neste trabalho foi muito importante para a validação do genoma de *Moniliophthora perniciosa* (Mondego et al., 2008) porque ela é formada por transcritos gerados durante a fase de frutificação do fungo. Esta fase

é bem distinta, pois ele já está no estágio saprofítico e não apresenta a expressão de genes responsáveis pela interação com a planta. Esta biblioteca também foi utilizada para treinar os programas de análise utilizados para identificar, por exclusão, genes da interação *M. pernicioso* x cacau (Gesteira et al., 2007).

A qualidade da biblioteca foi observada pelo tamanho dos fragmentos, pela redundância, pela qualidade das seqüências e pela presença de genes relacionados ao processo biológico analisado. O tamanho dos fragmentos clonados e a estratégia utilizada (Kit Clontech) permitiram a identificação de vários genes inteiros. A redundância de 44% foi razoável e próxima a valores encontrados em outras bibliotecas (Roche et al., 2005; Rincones et al., 2008). As seqüências geradas após limpeza e agrupamento permitiram a formação de 1.533 unigenes, sendo 276 representados por seqüências agrupadas e 1.257 por seqüências simples. A comparação com outros bancos de dados de seqüências permitiu a identificação por BLAST de 981 prováveis genes também presentes em outros organismos. Dos 552 não identificados, 453 estão presentes em *M. pernicioso* e podem significar novos genes. O alto número de agrupamentos formados com baixo número de seqüências indica que essa biblioteca não foi extensivamente seqüenciada, porém esse número de seqüências geradas foi suficiente para as análises desejadas.

A classificação funcional dos genes identificados por similaridade de seqüências mostrou uma freqüência de 53,1% de não identificados e 12,4% de genes com função desconhecida. Esses valores são próximos dos encontrados para outros fungos (Soanes e Talbot, 2006). Esses autores observaram que fungos patogênicos como *Blumeria graminis* apresentaram um maior número de genes desconhecidos em comparação a outros não patogênicos e atribuem esse fato à maior divergência genética entre eles. O pequeno número de seqüências depositadas e anotadas nos bancos de dados de ESTs obtidas de basidiomicetos

patogênicos também pode ser a causa dessa frequência elevada de transcritos desconhecidos. O número de genes não identificados deve aumentar com o aumento de genomas de basidiomicetos seqüenciados. De qualquer forma, nesse banco de ESTs com funções desconhecidos podem estar genes não descritos ainda relacionados a essa fase de vida do fungo e, portanto, são importantes para futuras análises.

Com relação às funções conhecidas, metabolismo primário e síntese protéica foram as mais abundantes como o esperado, porém citoesqueleto, defesa celular, cromatina e metabolismo do DNA e estímulo abiótico foram menores do que os valores observados para as ESTs obtidas da interação Cacau x *M. pernicioso* (Gesteira et al., 2007). Uma das explicações possíveis é que o número de funções desconhecidas foi muito maior que nas referidas bibliotecas e parte dos genes específicos do fungo são, ainda, realmente desconhecidos. Entretanto, na função de estímulo abiótico, apesar do menor valor proporcional, foram identificados genes referidos na fase de frutificação de outros basidiomicetos.

A abundância relativa de transcritos obtida por comparação de seqüências com um banco de dados de ESTs do próprio *M. pernicioso* serviu para identificar e confirmar genes presentes unicamente nesta fase, como os que codificam para as aegerolisinas. Essas proteínas foram identificadas especificamente na fase inicial da formação de primórdios de *Agrocybe aegerita*, *Pleurotus ostreatus* e *Aspergillus fumigatus*, dentre outros, e está sendo estudada quanto à sua função específica no desenvolvimento de estruturas de frutificação (Berne et al., 2002; Vidic et al., 2005; Berne et al., 2007). Ostreolisinas, que fazem parte dessa família de proteínas, são hemolisinas que possuem a capacidade de formar poros na membrana associando-se preferencialmente a lisofosfolipídios. Os poros são formados pela complexação de pleurotolisina B (Ply B) com a pleurotolisina A (PlyA) (Sepčić et al., 2003).

Um provável gene *pim1* também foi único dessa biblioteca na comparação *in silico*. Ele codifica para uma proteína quinase serino/treonina PIM1 que tem a função de promover o ciclo celular e induzir a formação de tumores (Morishita et al., 2008). Outros genes interessantes são os que codificam para os prováveis reguladores da Rho1, o Rho1-GEF e o inibidor da dissociação Rho1-GTP relacionados à transdução de sinais mediados pelas proteínas da família AMPK.

A comparação *in silico* com os genomas de *L. bicolor* e *U. maydis* também forneceram informações importantes. Os 289 unigenes que apresentaram similaridades somente a *L. bicolor* e não a *U. maydis* podem indicar vias de formação de basidiomatas, uma vez que *L. bicolor* produz basidiomatas enquanto *U. maydis* não. Entre esses, além das hidrofobinas, discutidas anteriormente, destaca-se um gene que codifica para um provável homeodomínio que pode estar relacionado à diferenciação de basidiomatas. Como foi mostrado anteriormente (Pires et al., 2009), *M. perniciosa* apresenta etapas de diferenciação celular para a formação de basidiomatas que implicam em comportamento espacial diferente entre hifas. Em *Drosophila melanogaster*, fatores de transcrição que ativam homeodomínios são responsáveis pelo desenvolvimento embrionário (Li et al., 2008). Além desse, destacam-se também três prováveis genes reguladores: um provável oncogene que codifica para uma proteína de transformação viral Jun, um provável fator de transcrição, que codifica para uma proteína que possui domínio BTB e ‘dedos de zinco’ e outro provável fator de transcrição que codifica para outra proteína também com domínio ‘dedos de zinco’. A formação de basidiomatas deve ocorrer por uma regulação de genes em cascata. Desta forma, esses unigenes identificados podem ser reguladores do desenvolvimento dessas estruturas.

Além desses prováveis genes reguladores, foram similares especificamente a *L. bicolor* outros genes relacionados ao metabolismo primário e secundário, metabolismo de DNA,

transporte através da membrana e síntese protéica. Entre eles são listados alguns que codificam para uma provável enzima da rota do ácido jasmônico, outro para uma enzima da rota de aflatoxinas e outro para uma provável enzima da via de produção de ácido linoléico. Também está presente um gene que codifica para uma lisofosfolipase e outro para o fator de transcrição MBF. A degradação de lisofosfolipídios pode estar relacionada à regulação do desenvolvimento de basidiomatas, porque pode participar da regulação dos prováveis genes que codificam para aegerolisinas encontradas nessa biblioteca. Proteínas dessa família são moduladas por lisofosfolipídios (Sepčić et al., 2003). Por outro lado, fatores do tipo MBF parecem estar envolvidos com pigmentação ou lignificação (Walton et al., 2005).

Análise da abundância de transcritos para 192 genes identificados

Os fragmentos de genes impressos na membrana foram selecionados em função da análise *in silico* da biblioteca de cDNA representativa da fase de frutificação do fungo. Alguns genes de função desconhecida foram incluídos, além de representantes de unigenes mais abundantes. A maioria, porém, foi incluída por apresentar uma função correlata ao desenvolvimento de basidiomatas.

O grupo que apresentou aumento de transcritos especificamente na fase de primórdios e que não permaneceram ativos nos basidiomatas parecem ser interessantes para a identificação de genes que estejam atuando na fase de desenvolvimento inicial de basidiomatas. Nele encontram-se prováveis hemolisinas, além de um gene para uma provável proteína pequena de choque térmico, outro para uma intersectina, além de um para a provável proteína de endocitose END4 e outros para prováveis funções regulatórias como AMPK, RNG2, CER4 e SRP2. Dois genes com funções desconhecidas foram ativos nessa fase e

podem ser estudados com mais detalhes posteriormente. Entretanto, alguns genes específicos que foram incluídos nos outros grupos também merecem destaque.

Houve uma confirmação de que aegerolisinas participam do processo de formação de basidiomatas, como ocorre em outros fungos, porque todos os ESTs colocados na membrana indicaram aumento de transcritos na fase de primórdios, embora alguns tenham sido ativados antes, ou se tratam de precursores. Quanto às hidrofobinas, a abundância de transcritos variou em cada ESTs depositado na membrana, indicando a existência de mais de um gene codificando para essas proteínas. Esses resultados concordam com os obtidos por Santos com análises de RT-PCR (Santos, 2005).

A participação de aegerolisinas e hidrofobinas na formação de primórdios já é conhecida (Berne et al., 2007; Wösten, 2000). As funções exatas das aegerolisinas ainda estão sendo desvendadas (Berne et al., 2009). Já as hidrofobinas formam filmes anfipáticos na superfície das hifas, protegendo-as do contato com a água. No caso específico de formação de primórdios de *M. perniciosa* cultivado ‘*in vitro*’, uma camada superficial hidrofóbica é formada no manto micelial e os primórdios em estágio indiferenciado são formados abaixo dessa camada, emergindo posteriormente (Pires et al., 2009). Essa camada parece ser formada por fusão de hifas e a hipótese é que essas hifas estejam recobertas por hidrofobinas.

A formação do primórdio indiferenciado abaixo da camada hidrofóbica pode ocorrer com a participação das aegerolisinas. Em fungos está sendo hipotetizada a sua participação na agregação de hifas que ocorre na fase inicial dos primórdios (Berne et al., 2002; Berne et al., 2009). A função hemolítica dessas proteínas ocorre na presença de lipídios específicos de células de mamíferos e não de fungos (Sepčić et al., 2003; Sepčić et al., 2005).

O aumento de transcritos detectados para os genes que codificam prováveis intersectina e END4 indica uma regulação de processos endocíticos comuns em estruturas em

desenvolvimento (Fuchs e Steinberg, 2005; Ayscough, 2005). Em primórdios iniciais ocorre remobilização, acúmulo de nutrientes e crescimento hifal (Kues, 2000). O regulador da expressão gênica AMPK é uma proteína kinase ativada por AMPc que promove produção de ATP e inibe o seu consumo. Essas proteínas são ativadas em respostas a vários sinais nutricionais e de estresses e em consequência regulam o estoque de glicogênio e vários outros metabólitos. São mediadoras de sinais que passam para outras proteínas como Rheb, para regular a síntese protéica e o crescimento celular, ou Rab, para regular o aporte de glicose (Beale, 2008). Ela serve como um sensor do estado nutricional da célula (Hue e Rider, 2007). A formação de primórdios deve ocorrer em resposta a estresses nutricionais (Kues, 2000), então esse resultado indica que isto, de fato, está ocorrendo em *M. pernicioso*. Uma análise mais detalhada por qPCR de genes que são regulados pela AMPK indica que isso deve estar realmente ocorrendo durante a indução e formação de primórdios em *M. pernicioso* (Pires et al., 2009). No macroarranjo aqui analisado, o aumento de expressão de um gene que provavelmente codifica para RNG2 na fase de primórdios colabora com essa hipótese. Proteínas Ras do tipo Rng2 são efetores das proteínas GTPases-Rho e estão relacionados à citocinese. A formação de lamelas em primórdios iniciais deve ocorrer por citocinese.

Um gene detectado que apresenta possivelmente a função de fator de transcrição CCR4 também está ativo na fase de primórdios. CCR4 são fatores transcricionais gerais de regulação da transcrição, mas também atuam na degradação do RNA (Collart, 2003). O aumento de expressão desse gene indica que uma regulação pós transcricional está ocorrendo nessa fase do desenvolvimento do fungo. É interessante notar que outro provável regulador pós transcricional também esteja ativo nessa fase, codificando para uma proteína da família SR (proteínas ricas em serina e argininas), que tem função no processamento de íntrons (Graveley, 2000).

Conclusões

Esse é um levantamento pioneiro de genes expressos na fase de frutificação. Foram identificados 1.533 genes por análise *in silico* e 192 foram avaliados quanto ao seu nível de transcritos nas diferentes etapas da fase de frutificação do fungo. Os perfis de expressão produzidos por análise de correlação de Pearson a partir das intensidades geradas no macroarranjo permitiram a identificação de alguns transcritos mais abundantes somente na fase de primórdios, entre eles alguns codificando para proteínas reguladoras como AMPK e RNG2, indicadores de resposta a déficits nutricionais. Genes para hidrofobinas e aegerolisinas foram confirmados como participantes do processo de formação de primórdios.

Os genes identificados por similaridade de seqüências abrem caminho para novos estudos incluindo análises de expressão e funcionais. Uma análise global dos genes dessa biblioteca indica que realmente deve ocorrer ativação da produção de primórdios em resposta a estresses nutricionais e ambientais, porém, somente com análises funcionais posteriores, será possível elucidar a participação dos genes aqui levantados na formação dos basidiomas.

Referências

- Aime MC, Phillips-Mora W: The causal agent of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of *Marasmiaceae*. *Mycologia* 2005, **97**: 1012-1022.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 1997, **25**:3389-3402.
- Ayscough KR: Coupling actin dynamics to the endocytic process in *Saccharomyces cerevisiae*. *Protoplasma* 2005, **226**(1-2):81-8
- Alves SAM, Pomella AWV, Aitken WM, Bergamin Filho, A: Curvas de progresso e gradientes da vassoura-de-bruxa (*Crinipellis pernicioso*) em cacaueiros enxertados em Uruçuca, Bahia. *Fitopatologia Brasileira* 2006, 31:483-491. 2006.
- Beale EG: 5'-AMP-activated protein kinase signaling in *Caenorhabditis elegans*. *Exp Biol Med (Maywood)* 2008, **233** (1):12-20.
- Berne S, Križaj I, Pohleven F, Turk T, Maček P, Sepčić K: *Pleurotus* and *Agrocybe* hemolysins, new proteins hypothetically involved in fungal fruiting. *Biochim Biophys Acta* 2002, **1570** (3): 153-159.
- Berne S, Pohleven J, Vidic I, Rebolj K, Pohleven F, Turk T, Maček P, Sonnenberg A, Sepčić K: Ostreolysin enhances fruiting initiation in the oyster fungus (*Pleurotus ostreatus*). *Mycol Res* 2007, **111** (12): 1431-1436.
- Berne S, Lah L, Sepčić K: Aegerolysins: Structure, function, and putative biological role. *Protein Science* 2009, **18**(4): 694-706.
- Collart MA: Global control of gene expression in yeast by the Ccr4-Not complex. *Gene* 2003 **313**: 1-16

- De Groot PWJ, Peter I, Schaap J, Van Griensven LJLD, Visser J: Isolation of developmentally regulated genes from the edible mushroom *Agaricus bisporus*. *Microbiology* 1997, **143**: 1993-2001.
- Ebbole DJ, Jin Y, Thon M, Pan H, Bhattarai E, Thomas T, Dean R: Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: analysis of Expressed Sequence Tags. *MPMI* 2004, **17** (12): 1337-1347.
- Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998, **95**: 14863-14868.
- Evans HC: Witches' broom disease - A case study. *Cocoa Growers Bulletin* 1981, **32**: 5-19.
- Ewing B, Hillier L, Wendl MC, Green P: Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* 1998, **8**: 175-185.
- Feinberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem* **132**: 6-13. 1983
- Fuchs U, Steinberg G: Endocytosis in the plant-pathogenic fungus *Ustilago maydis*. *Protoplasma* 2005, **226**: 75-80
- Gesteira AS, Micheli F, Carels N, Silva AC da, Gramacho KP, Shuster I, Macedo JN, Pereira GAG, Cascardo JM: Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa*. *Ann Bot* 2007, **100**: 129-140.
- Graveley R: Sorting out the complexity of SR protein functions. *RNA* 2000, **6** (9): 1197-1211.
- Griffith GW, Hedger JN: A novel method for producing basidiocarps of the cocoa pathogen *Crinipellis perniciosa* using a bran-vermiculite medium. *Europ J Plant Pathol* 1993, **99**: 227-230.
- Hibbett DS, Binder M: Evolution of complex fruiting-body morphologies in homobasidiomycete. *Proc. R. Soc. Lond. B* 2002, **269**: 1963-1969.

- Huang X, Madan A: CAP3: A DNA sequence assembly program. *Genome Res* 1999, **9**:868-877
- Hue L, Rider MH: The AMP-activated protein kinase: more than an energy sensor. *Enssays Biochem* 2007, **43**: 121-37
- Journet E-P, van Tuinen D, Gouzy J, Crespeau H, Carreau V, Farmer M-J, et al: Exploring root symbiotic programs in the model legume *Medicago truncatula* using EST analysis. *Nucleic Acids Res* 2002, **30**: 5579–5592.
- Kues U, Liu Y: Fruiting body production in basidiomycetes. *Appl Microbiol Biotechnol* 2000, **54**: 141-152.
- Kues U: Life history and developmental processes in the basidiomycete *Coprinus cinereus*. *Microbiol Mol Biol Rev* 2000, **64**: 316-353.
- Li XY, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VY, Hechmer A, Simirenko L, Stapleton M, Hendriks CLL, Chu HC, Ogawa N, Inwood W, Sementchenko V, Beaton A, Weizmann R, Celniker SE, Knowles DW, Gingeras T, Speed TP, Eisen MB, Biggin MD: Transcription Factors Bind Thousands of Active and Inactive Regions in the Drosophila Blastoderm. *PLSO Biology* 2008, **6** (2): 365-388.
- Lorence A, Chevone BI, Mendes P, Nessler CL: Myo-inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis. *Plant Physiol* 2004, **134**: 1200–1205.
- Macagnan D, Romeiro RS, Souza J, Pomella AWW: Isolation of actinomycetes and endospore-forming bacteria from the cacao pod surface and their activity against the witches' broom and black pod pathogens. *Phytoparasitica* 2006, **34**: 122-132.
- Maeda H , Sano M , Maruyama Y, Tanno T, Akao T, Totsuka Y, Endo M, Sakurada R, Yamagata Y, Machida M, Akita O, Hasegawa F, Abe K, Gomi K, Nakajima T, Iguchi T: Transcriptional analysis of genes for energy catabolism and hydrolytic enzymes in the

filamentous fungus *Aspergillus oryzae* using cDNA microarrays and expressed sequence tags. *Appl Microbiol Biotechnol* 2004, **65**: 74–83

Meinhardt LW, Rincones J, Bailey B, Aime MC, Griffith GW, Zhang D, Pereira G: *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe? *Mol Plant Pathol* 2008, **9** (5), 577–588.

Mondego JMC, Carazzolle MF, Costa GGL, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, García O, Thomazzela DPT, Oliveira BV, Pires ABL, Rio MCS, Araújo MRR, Castro LAB, Gramacho KP, Gonçalves MS, Góes-Neto A, Barbosa LV, Guiltinan MJ, Bailey B, Meinhardt L, Cascardo JCM, Pereira GAG: A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. *BCM Genomics* 2008, **9**:548.

Morishita D, Katayama R, Sekimizu K, Tsuruo T, Fujita N: Pim kinases promote cell cycle progression by phosphorylating and down-regulating p27Kip1 at the transcriptional and posttranscriptional levels. *Cancer Research* 2008, **68**: 5076.

Niella G, Resende ML, Castro HA, Carvalho GA de, Silva LHCP: Aperfeiçoamento da metodologia de produção artificial de basidiocarpos de *Crinipellis perniciosa*. *Fitop Brasileira* 1999, **24**: 523–527.

Peter M, Courty P-E, Kohler A, Delaruelle C, Martin D, Tagu D, Frey-Klett P, Duplessis S, Chalot M, Podila G, Martin F: Analysis of expressed sequence tags from the ectomycorrhizal basidiomycetes *Laccaria bicolor* and *Pisolithus microcarpus*. *New Phytologist* 2003, **159**: 117–129

Pires ABL, Gramacho KP, Silva DC, Goes-Neto A, Silva MM, Muniz-Sobrinho JS, Porto RF, Villela-Dias C, Brendel M, Cascardo JCM, Pereira, GAG: Early development of

- Moniliophthora perniciosa basidiomata and developmentally regulated genes. *BMC Microbiology* 2009, **9**:158
- Purdy LH, Dickstein ER: Basidiocarp development on micelial mats of *Crinipellis perniciosa*. *Plant Dis* 1990, **74**: 493-496.
- Purdy LH, Schmidt RA: Status of cacao witches' broom: biology, epidemiology, and management. *Annu Rev Phytopath* 1996, **34**: 573-594.
- Rincones J, Scarpari LM, Carazzolle MF, Mondego JMC, Formighieri EF, Barau JG, Costa GGL, Carraro DM, Brentani HP, Vilas-Boas LA, Oliveira BV, Sabha M, Dias R, Azevedo RA, Cascardo JM, Meinhardt LW, Pereira GAG: Differential gene expression between the biotrophic-like and saprotrophic mycelia of the Witches' broom pathogen *Moniliophthora perniciosa*. *Mol Plant Microbe Interact* 2008, **21** (7):891-908.
- Rocha HM, Wheeler BEJ: Factors influencing the production of basidiocarps and the deposition and germination of basidiospores of *Crinipellis perniciosa*, the causal agent of witches' broom disease on cocoa (*Theobroma cacao*). *Plant Pathol* 1985, **34**: 319-328.
- Rocha HM, Wheeler BEJ: The water balance as an important factor in basidiocarp production by *Crinipellis perniciosa*, the causal fungus of cocoa witches' broom. *Proc 8th Internat. Cocoa Res. Conf. 1981. Cartagena, Columbia: Cocoa Producers Alliance*, 1982, 381-386.
- Roche JP, Ashley Wackym P, Cioffi JA, Kwitek AE, Erbe CB, Popper P: In silico Analysis of 2085 Clones from a normalized rat vestibular periphery 3' cDNA Library. *Audiol Neurotol*. 2005; **10** (6): 310-322.
- Sambrook J, Russell DW: Molecular Cloning. A Laboratory Manual. Third edition, New York: Cold Spring Harbor Laboratory, 2001.

- Santos Filho LP, Freire ES, Carzola IM: Estimativas de perda de produção de cacau causadas por vassoura-de-bruxa (*Crinipellis pernicioso* (Stahel) Singer) na Bahia. *Agrotropica* 1998, **10** (3): 127-130.
- Santos SC: Caracterização de hydrophobins do fungo *Crinipellis pernicioso* (Stahel) Singer, causador da doença Vassoura-de-Bruxa no cacaueiro. M. S. Thesis, Universidade Estadual de Santa Cruz, Ilhéus – Bahia, Brazil, 2005.
- Sepčić K, Berne S, Potrich C, Turk T, Macek P, Menestrina G: Interaction of ostreolysin, a cytolytic protein from the edible mushroom *Pleurotus ostreatus*, with lipid membranes and modulation by lysophospholipids. *Eur J Biochem* 2003, **270**(6):1199-1210.
- Sepčić K, Berne S, Rebolj K, Batista U, Plemenitaš A, Šentjerc M, Maček P: Ostreolysin, a pore-forming protein from the oyster fungus, interacts specifically with membrane cholesterol-rich lipid domains. *Mycol Res* 2005, **109**: 374-376.
- Shim SM, Kim SB, Kim HY, Rho H-S, Lee HS, Lee MW, Lee UY, Im KH, Lee TS: Genes expressed during fruiting body formation of *Agrocybe cylindracea*. *Mycrobiology* 2006, **34** (4): 209-213.
- Silva SDVM, Luz EDMN, Almeida, OC, Gramacho, KP, Bezerra, JL: Redescritção da sintomatologia causada por *Crinipellis pernicioso* no cacaueiro. *Agrotrópica* 2002, **14** (1).
- Soanes DM, Talbot NJ, Comparative genomic analysis of phytopathogenic fungi using expressed sequence tag (EST) collections. *Mol Plant Pathol* 2006, **7** (1): 61–70
- Telles GP, Braga MDV, Dias Z, Lin TL, Quitzau JAA, Silva FR, Meidanis J: Bioinformatics of the sugarcane EST project. *Genet Mol Biol* 2001: **24**:9-15.
- Telles GP, Silva FR: Trimming and clustering sugarcane ESTs. *Genet Mol Biol* 2001, **24**:17-23.

- Trevizan SDP; Marques M: Impactos socioeconomicos da crise do cacau: um estudo de comunidade-caso. *Agrotropica* 2002, **14**: 127-136.
- Vidic I, Berne S, Drobne D, Maček P, Frangež R, Turk T, Štrus J, Sepči K Temporal and spatial expression of ostreolysin during development of the oyster fungus (*Pleurotus ostreatus*) *Mycol Res* 2005, **109**: 377-382.
- Walton FJ, Idnurm A, Heitman J: Novel gene functions required for melanization of human pathogen *Cryptococcus neoformans*. *Mol Microbiol* 2005, **57**: 1381-1396.
- Wösten HAB, De Vocht ML: Hydrophobins, the fungal coat unraveled. *Biochim Biophys Acta* 2000, **1469**: 79-86.

A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao

Jorge MC Mondego, Marcelo F Carazzolle, Gustavo GL Costa, Eduardo F Formighieri, Lucas P Parizzi, Johana Rincones, Carolina Cotomacci, Dirce M Carraro, Anderson F Cunha, Helaine Carrer, Ramon O Vidal, Raíssa C Estrela, Odalys García, Daniela PT Thomazella, Bruno V de Oliveira, Acássia BL Pires, Maria Carolina S Rio, Marcos Renato R Araújo, Marcos H de Moraes, Luis AB Castro, Karina P Gramacho, Marilda S Gonçalves, José P Moura Neto, Aristóteles Góes Neto, Luciana V Barbosa, Mark J Guiltinan, Bryan A Bailey, Lyndel W Meinhardt, Julio CM Cascardo and Gonçalo AG Pereira

BMC Genomics 2008, **9**:548 doi:10.1186/1471-2164-9-548

Research article

Open Access

A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao

Jorge MC Mondego^{†1}, Marcelo F Carazzolle^{†1}, Gustavo GL Costa¹,
 Eduardo F Formighieri¹, Lucas P Parizzi¹, Johana Rincones¹,
 Carolina Cotomacci¹, Dirce M Carraro², Anderson F Cunha³,
 Helaine Carrer⁴, Ramon O Vidal¹, Raíssa C Estrela¹, Odalys García¹,
 Daniela PT Thomazella¹, Bruno V de Oliveira¹, Acássia BL Pires⁵,
 Maria Carolina S Rio¹, Marcos Renato R Araújo¹, Marcos H de Moraes¹,
 Luis AB Castro⁶, Karina P Gramacho⁷, Marilda S Gonçalves⁸, José P
 Moura Neto⁸, Aristóteles Góes Neto⁹, Luciana V Barbosa¹⁰,
 Mark J Gultinan¹¹, Bryan A Bailey¹², Lyndel W Meinhardt^{†12},
 Julio CM Cascardo^{†5} and Gonçalo AG Pereira^{*1}

Address: ¹Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, 13083-970, Campinas – SP, Brazil, ²Laboratório de Genômica e Biologia Molecular, Hospital A.C. Camargo, 01509-010, São Paulo – SP, Brazil, ³HEMOCENTRO, Laboratório de Genoma e Hemoglobina, Universidade Estadual de Campinas, 13084-878, Campinas – SP, Brazil, ⁴Departamento de Ciências Biológicas, ESALQ, USP, 13418-900, Piracicaba – SP, Brazil, ⁵Laboratório de Genômica e Expressão Gênica, Departamento de Genética e Biologia Molecular, Universidade Estadual de Santa Cruz, 45650-000, Ilhéus – BA, Brazil, ⁶Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica – PqEB – Av. W5 Norte, 70770-900, Brasília – DF, Brazil, ⁷CEPLAC/CEPEC/SEFT, 45600-970, Itabuna – BA, Brazil, ⁸Laboratório de Biologia Molecular – Faculdade de Farmácia, Universidade Federal da Bahia, 40170-290, Salvador – BA, Brazil, ⁹Laboratório de Pesquisa em Microbiologia (LAPEM), Departamento de Ciências Biológicas, Universidade Estadual de Feira de Santana (UEFS), 44031-460, Feira de Santana – BA, Brazil, ¹⁰Laboratório de Biologia Molecular – Departamento de Biologia Geral, Instituto de Biologia, Universidade Federal da Bahia, 40170-290, Salvador – BA, Brazil, ¹¹Department of Horticulture, Pennsylvania State University, University Park, Chester, PA 16802, USA and ¹²Sustainable Perennial Crops Laboratory, USDA-ARS, 10300 Baltimore Av, Bldg. 001, 18 Beltsville MD 20705-2350, USA

Email: Jorge MC Mondego - jmcmondego@lge.ibi.unicamp.br; Marcelo F Carazzolle - mcarazzo@lge.ibi.unicamp.br; Gustavo GL Costa - glacerda@lge.ibi.unicamp.br; Eduardo F Formighieri - eduformi@gmail.com; Lucas P Parizzi - lucas@lge.ibi.unicamp.br; Johana Rincones - johanarp@unicamp.br; Carolina Cotomacci - caca@unicamp.br; Dirce M Carraro - dirce.carraro@hcancer.org.br; Anderson F Cunha - andarf@unicamp.br; Helaine Carrer - hecarrer@esalq.usp.br; Ramon O Vidal - vidal@lge.ibi.unicamp.br; Raíssa C Estrela - raissa@lge.ibi.unicamp.br; Odalys García - odalys@lge.ibi.unicamp.br; Daniela PT Thomazella - danitt@lge.ibi.unicamp.br; Bruno V de Oliveira - bvo@lge.ibi.unicamp.br; Acássia BL Pires - acassia@lge.ibi.unicamp.br; Maria Carolina S Rio - scatolin@lge.ibi.unicamp.br; Marcos Renato R Araújo - marcos.araujo@grialebiometrics.com; Marcos H de Moraes - mhmoraes@lge.ibi.unicamp.br; Luis AB Castro - lbarreto@mct.gov.br; Karina P Gramacho - karina@cepec.gov.br; Marilda S Gonçalves - mari@bahia.fiocruz.br; José P Moura Neto - jpneto@cpqgm.fiocruz.br; Aristóteles Góes Neto - agoesnt@uefs.br; Luciana V Barbosa - veiga@ufba.br; Mark J Gultinan - mjg9@psu.edu; Bryan A Bailey - bryan.bailey@ars.usda.gov; Lyndel W Meinhardt - lyndel.meinhardt@ars.usda.gov; Julio CM Cascardo - cascardo@uesc.br; Gonçalo AG Pereira^{*} - goncalo@unicamp.br

^{*} Corresponding author [†]Equal contributors

Published: 18 November 2008

Received: 17 June 2008

BMC Genomics 2008, 9:548 doi:10.1186/1471-2164-9-548

Accepted: 18 November 2008

This article is available from: <http://www.biomedcentral.com/1471-2164/9/548>

© 2008 Mondego et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The basidiomycete fungus *Moniliophthora perniciosa* is the causal agent of Witches' Broom Disease (WBD) in cacao (*Theobroma cacao*). It is a hemibiotrophic pathogen that colonizes

the apoplast of cacao's meristematic tissues as a biotrophic pathogen, switching to a saprotrophic lifestyle during later stages of infection. *M. perniciosa*, together with the related species *M. roreri*, are pathogens of aerial parts of the plant, an uncommon characteristic in the order Agaricales. A genome survey (1.9× coverage) of *M. perniciosa* was analyzed to evaluate the overall gene content of this phytopathogen.

Results: Genes encoding proteins involved in retrotransposition, reactive oxygen species (ROS) resistance, drug efflux transport and cell wall degradation were identified. The great number of genes encoding cytochrome P450 monooxygenases (1.15% of gene models) indicates that *M. perniciosa* has a great potential for detoxification, production of toxins and hormones; which may confer a high adaptive ability to the fungus. We have also discovered new genes encoding putative secreted polypeptides rich in cysteine, as well as genes related to methylotrophy and plant hormone biosynthesis (gibberellin and auxin). Analysis of gene families indicated that *M. perniciosa* have similar amounts of carboxylesterases and repertoires of plant cell wall degrading enzymes as other hemibiotrophic fungi. In addition, an approach for normalization of gene family data using incomplete genome data was developed and applied in *M. perniciosa* genome survey.

Conclusion: This genome survey gives an overview of the *M. perniciosa* genome, and reveals that a significant portion is involved in stress adaptation and plant necrosis, two necessary characteristics for a hemibiotrophic fungus to fulfill its infection cycle. Our analysis provides new evidence revealing potential adaptive traits that may play major roles in the mechanisms of pathogenicity in the *M. perniciosa*/cacao pathosystem.

Background

Moniliophthora perniciosa, previously known as *Crinipellis perniciosa* (Singer) Stahel, is a hemibiotrophic basidiomycete (Tricholomataceae, Agaricales, Marasmiaceae) fungus that causes Witches' broom disease (WBD) in cacao (*Theobroma cacao* L.) [1-3]. WBD and frosty pod rot (FPR), caused by *Moniliophthora roreri*, are the most devastating diseases of cacao in the Americas [4]. Cacao production in southeastern Bahia, the main production area in Brazil, was severely affected by the introduction of WBD at the end of 1980's [5]. This disease damaged Bahian agribusiness, caused major social problems and has contributed to the degradation of the Atlantic Rainforest ("Mata Atlântica"). This is because cacao producing areas, typically, maintained old-growth native tree species as shade for the cacao plantations, which were converted to pasture [6].

The symptoms displayed by cacao plants during WBD parallel the hemibiotrophic development of *M. perniciosa* [7]. Briefly, the disease begins when fungal spores germinate and infect meristematic tissues, developing into monokaryotic biotrophic hyphae without clamp connections that slowly occupy the intercellular space. This stage of WBD is characterized by the emergence of hypertrophic and hyperplastic anomalous branches, and the formation of parthenocarpic fruits. Infected branches, known as "green brooms", grow without apical dominance, with a phototropic orientation and displaying epinastic leaves [7]. After two to three months the infected tissue turns necrotic (dry brooms) and the hyphae become saprotrophic with two nuclei per cell and clamp connections,

invading the inter and intracellular spaces of the infected tissue [6,8]. This fungal species exhibits primary homothallism as its reproductive strategy [9]; thus, the change from the monokaryotic to the dikaryotic mycelium occurs without the prerequisite of mating between compatible individuals. After alternate wet and dry periods, basidiospores produced by the saprotrophic hyphae release basidiospores that are spread by wind or rain, thus completing the *M. perniciosa* life cycle [7].

The Witches' broom Genome Project <http://www.lge.ibi.unicamp.br/vassoura> involving several Brazilian laboratories was initiated to increase the knowledge of this disease. The genome size, chromosomal polymorphism, genetic variability and the *M. perniciosa* mitochondrial genome have already been described [10-12]. Additionally, a biochemical study revealed the metabolic modifications that occur in cacao plantlets during WBD development [13]. Technical improvements have been achieved in the manipulation of cacao [14,15] and *M. perniciosa* [16,17]. One of the main bottlenecks in *M. perniciosa* research was solved with the development of the *in vitro* production of biotrophic-like cultures [18]. Necrotic inducing proteins expressed by *M. perniciosa* have been characterized [19], and the analysis of EST libraries and DNA microarrays have identified differentially expressed genes during its development [20; A.B.L. Pires *et al.*, unpublished data] and for the interaction of the fungus with cacao [21,22]. Despite the substantial progression in understanding WBD, many questions remain unsolved, mainly those concerning the mechanisms controlling processes such as: (i) the fungal switch from biotrophism

to saprotrophism; (ii) the drastic phenotypic alteration of cacao during disease development and (iii) the death of infected tissues. *In vitro* cultures in our laboratory demonstrate that this fungus has a great capacity to adapt to media containing different sources of carbon and nitrogen and it is able to grow in extremely nutrient-poor media. These results together with the fact that hemibiotrophic fungi, such as *M. perniciosa*, display complex lifestyles [23] suggest that this fungus has a significant genomic and transcriptomic plasticity that contributes to the successful pathogenic mechanisms expressed during its life cycle. In a recent review, Meinhardt *et al.*, give a historical account and summarize the current state of knowledge about WBD [24].

Around two thirds of the known Basidiomycota species are included in the order Agaricales [25]. In addition to being an interesting group for carrying out developmental genetic studies concerning fungal development and reproduction, the Agaricales contain many important industrial species (i.e., edible mushrooms, fiber bleaching fungus), and species with unique lifestyles: saprophytes (i.e., wood-decaying fungus *Phanerochaete chrysosporium*), symbionts (i.e., ectomycorrhizal fungus *Laccaria bicolor*), leaf-litters decomposers (i.e., *Agaricus bisporus*) and root pathogens (i.e., *Armillaria mellea*). Interestingly, *M. perniciosa* and *M. roreri* are members of a group of Agaricales species that are able to infect aerial parts of plants, an uncommon characteristic among this Basidiomycota order.

Genome sequencing and analysis is an important strategy to obtain comprehensive information concerning the metabolism and development of organisms. The initial objective of the *M. perniciosa* Genome Project was to obtain a genome survey sequence using a whole shotgun strategy to provide genomic information for the WBD research community. Recently, the bioinformatics and genomic communities have been debating about the benefits and costs of finishing a complete genome as compared to applying a genome survey strategy [26-32]. It has been proposed that a two-fold genome sequence coverage is sufficient to support a high percentage of EST alignments and exon similarity matches [26,31]. Additionally, gene models resulting from a genome survey can be predicted accurately by the comparison with complete genomes of phylogenetically related organisms [26,31]. This strategy has been used for genomic surveys of dog (1.5× coverage) [31] and of the wine spoilage yeast *Dekkera bruxellensis* (0.4× coverage) [33]. Recently, the genomes of the basidiomycete species *Laccaria bicolor*, *Cryptococcus neoformans*, *Ustilago maydis*, *Coprinopsis cinerea* (*Coprinus cinereus*) and *Phanerochaete chrysosporium* have been determined and many additional Basidiomycota genome projects are ongoing or about to be released <http://www.broad.mit.edu/annotation/cgi>; <http://www.jgi.doe.gov>, which could provide a rich database for additional Basidiomycota genomic survey evaluations.

www.jgi.doe.gov, which could provide a rich database for additional Basidiomycota genomic survey evaluations.

With the existence of several complete Basidiomycota genomes, and both *M. perniciosa* ESTs libraries, and a 1.9 × genome coverage, we decided to conduct a genome survey of *M. perniciosa* to obtain further information about this important phytopathogen. This report describes a survey of the genome sequences of *M. perniciosa*, with specific emphasis on the genes potentially involved in disease development such as a cytochrome P450 monooxygenases, transposable elements, putative plant defense elicitors, pathogenicity effectors, cell wall degrading enzymes, proteins related to methylotrophy and the biosynthesis of plant hormones by the fungus. The identification of such genes in the *M. perniciosa* genomic data lead us to hypothesize a connection between molecular processes involved in the growth phases of the fungus and the progression of WBD.

Results and discussion

Genome assembly and estimation of genome size

A diagram flow describing the bioinformatic procedures applied in the *M. perniciosa* genome survey are depicted in Figure 1. After sequencing, a total of 124,565 reads were obtained and assembled using the whole genome shotgun strategy. During the initial assembly process, a large contig including 6,920 reads was found. This contig was identified as the mitochondrial genome of the fungus (RefSeq NC_005927), which comprised approximately 6% of *M. perniciosa* sequences [12]. The remaining genome sequences were assembled resulting in 17,991 contigs and 7,065 singlets with average lengths of 1,300 bp and 455 bp, respectively. The largest contig consisted of 25,364 bp, and was formed by 513 reads. The sum of all reads was close to 75 Mbp and the total assembly consisted of 26.7 Mbp (Fig. 1).

Previous Feulgen-image analysis experiments estimated *M. perniciosa* genome size to be 32.98 ± 7.95 Mbp [10]. Due to the large error in this estimation we decided to assess the genome size using the genome estimation protocol established in the dog genome survey [31] (more details in Additional File 1). This analysis resulted in a genome size ranging from 38.7 to 39.0 Mbp, a value similar to the genome length of another fungus belonging the order Agaricales, *C. cinerea* (37.5 Mbp), and that agrees with the previous size estimated by Feulgen-image analysis [10].

The Lander Waterman theory (LW) [34] and its applications [35,36] were used to confirm the estimate of the *M. perniciosa* genome size. The theoretical values for the expected number of clusters (contigs + singlets), contigs, gaps, average cluster length and average gap length can be calculated using the LW theory from the effective average

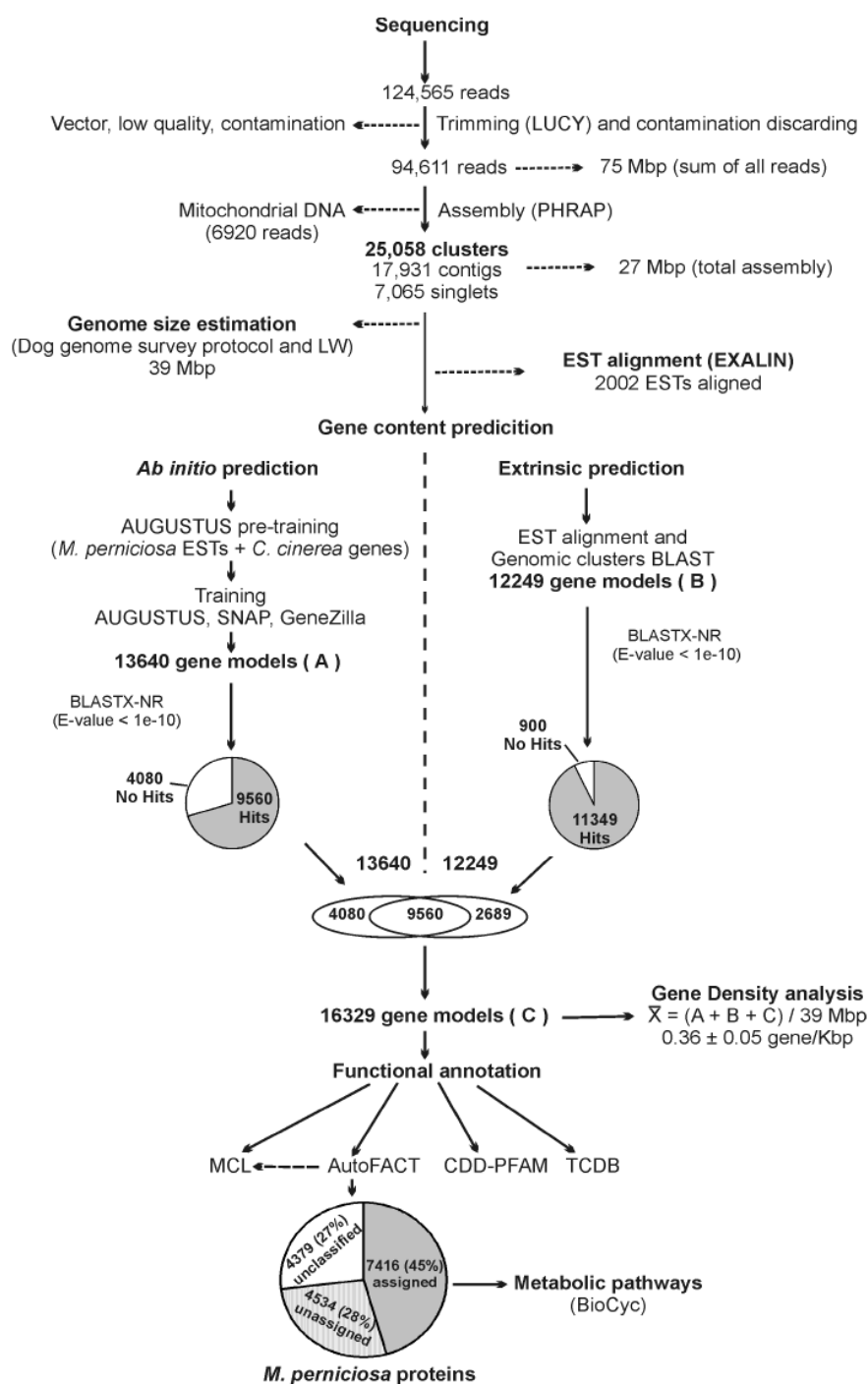


Figure 1
Flow diagram of bioinformatics procedures applied in *M. perniciosus* genome survey.

read length, the number of reads and the genome size as parameters. If the calculated values derived from the genome assembly were close to the theoretical values derived from LW calculations, then it is possible to confirm the estimated genome size. A summary of the results obtained from the assembly data and from the estimation using the LW theory is shown in Table 1. The experimental values agreed with the theoretical calculations, thus supporting our genome size estimation. According to LW analysis the average gap size was 413 bp. To have more information about the distribution of the gap size, we performed a comparison between a set of eukaryotic core proteins (generated by CEGMA pipeline [37]) and *M. perniciosus* contigs (See Additional file 1). Using this methodology, we detected that the average gap size was around 500 ± 300 bp, corroborating with LW average gap size estimate (413 bp). Considering that most of the gaps are around 500 bp and that *C. cinerea* average gene size is 1,678 bp, the majority of *M. perniciosus* genes or partial gene regions are likely to be included in the contigs. Corroborating this notion, we have found all genes encoding proteins of essential metabolic pathways such as Glycolysis, Gluconeogenesis, Pentose Phosphate pathway, and several others, which are available in the website <http://www.lge.ibi.unicamp.br/biocyc/MP> (see below).

In order to estimate the number of sequences misassembled due to repeat regions, we applied the integrated pipeline for assembly validation, called amosvalidate [38] (Further information in Additional file 1). This analysis resulted in 664 contigs with overrepresented regions totaling 1.1 Mbp. Multiplying the number of bases in overrepresented regions by over-coverage estimate resulted in 7.4 Mbp of repeat regions in the genome. Since reads from repetitive regions were eliminated from the estimation of the genome length using dog genome survey protocol (Additional file 1), this misassembling does not invalidates the genome size estimate.

Gene content

The initial step to uncover the gene content of *M. perniciosus* was performed using the genomic sequences together with a library of 3,145 ESTs, previously annotated and partially published [20]. Genes were identified by comparing these two libraries using the program Exalin [39]. This analysis allowed us to identify expressed genes and intron structure. The result was 2,002 ESTs aligned to the

genome contigs. Based on this EST-genome sequence alignment, the average intron length was estimated as 52 bp. This information was then used in the next steps of the process.

A more detailed investigation of the gene content was carried out using a combination of *ab initio* (gene predictor programs) and comparative gene prediction (BLAST-EST sequence alignment). For *ab initio* gene prediction, we first applied the methodology described for the AUGUSTUS gene predictor [40] <http://augustus.gobics.de/binaries/retraining.html>. This methodology essentially trains the AUGUSTUS program using a combination of sequences from the species of interest together with sequences of a phylogenetically related species, specifically with similarities in intron and exon length distributions. Using this approach, the coding content sensors (codon usage, GC content) are trained with sequences of the species of interest, and signal sensors (splice sites, TATA-box, polyadenylation sites, etc) are trained with the sequences of the related species. Sequences from *C. cinerea* were used in the training process; together with selected *M. perniciosus* ESTs with E-value in BLASTX-NR $\leq 1E-10$. The sequences of the ESTs that aligned with proteins in the databank were concatenated, giving rise to a 240 Kbp sequence. Ten copies of this *M. perniciosus* EST concatamers (total of 2.4 Mbp) together with a dataset of genes from *C. cinerea* <http://augustus.gobics.de/datasets>, comprised of 1.2 Mbp, were used to "pre-train" AUGUSTUS. The *M. perniciosus* predictions that came out of this pre-training were compared with the protein databank NR using BLASTp. The resulting predictions with similarities in the NR databank, and with a coverage $\geq 90\%$, were selected. After eliminating redundancies, 134 complete (containing the start codon and stop codon) and 1,136 partial (without the start codon and/or stop codon) *M. perniciosus* gene models were used to train AUGUSTUS [40] and two other gene predictors (SNAP [41] and Genezilla [42]). Predictions with less than 30 amino acids were eliminated. The remaining predictions were grouped into 19,932 overlapping clusters; that is, genomic regions covered by at least one prediction. The predictions in each overlapping cluster were ranked according to the criteria used by the Fungal Genome Initiative at the Broad Institute <http://www.broad.mit.edu>, with some adaptations. The *ab initio* gene finding pipeline generated 13,640 gene models, 9,560 of which contained

Table 1: Comparison between assembly values and values calculated using Lander Waterman theory

	Calculated from assembling	Calculated from LW theory
# Clusters (contigs+singlets)	25,056	24,950
# Contigs	17,991	18,370
# Gaps	-	24,951
Average cluster size (bp)	1,065	1,152
Average gap size (bp)	-	413

significant similarity to GenBank sequences (Fig. 1 and Fig. 2).

The extrinsic prediction methodology consisted of a combination of genomic similarity searches (BLASTX) using contigs and singlets, with the alignment of *M. pernicioso* ESTs in the genomic clusters. 17,991 contigs and 7,065 singlets were submitted to similarity analysis in a databank containing BLASTX-NR plus *P. chrysosporium* proteins. The genomic regions containing homologues in this databank were selected and exon-intron boundaries were determined. We also selected the alignments between ESTs and genomic clusters. Then, we evaluated if there was a superposition of BLAST alignments and EST alignments. After this analysis, the extrinsic prediction methodology revealed 12,249 gene models. Most of these gene models (9,560) were also predicted by the *ab initio* gene predictor programs. Of the 2,689 remaining gene models not predicted by the gene predictor programs, 1,789 presented significant similarity to sequences deposited in the GenBank (E-value $\leq 1E-10$) and 900 did not have any significant similarity in the GenBank (E-value $> 1E-10$) (Fig. 1 and Fig. 2). One of the reasons why these genes were not detected by the gene prediction programs could be the presence of low quality sequences that may have lead to frameshifts, thus making them impossible to detect by these programs. The total number of gene models obtained using both approaches, *ab initio* and extrinsic predictions, is 16,329. Assuming that the average contig length is 1.3 Kbp and the average gap length is 413 bp, it is possible that this total number of gene models (16,329) represents an overestimation. For instance, in our assem-

bly a gene with > 2 Kbp (protein ~ 700 aa) could be represented by two gene models, with each one in different contigs, which suggests a possible redundancy in this gene model prediction. Thus, it is very likely that the real number of gene models will be less and closer to the *ab initio* or extrinsic predictions separately. However, to obtain the maximum amount of information, we decided to evaluate the results of similarity searches in GenBank (BLASTX-NR) for all 16,329 gene models. As a result, 69.5% (11,349) of the gene models showed significant sequence similarity (E-value $\leq 1e-10$) to GenBank proteins. The remaining 30.5% (4,980), designated as "no-hits", were sequences with lower E-value scores (E-value $> 1e-10$). Of these gene models 4,080 were predicted with the *ab initio* gene predictors and 900 with the extrinsic predictions (Fig. 1 and Fig. 2).

As expected, BLASTX analysis against available complete fungal genomes showed that *M. pernicioso* gene models present higher similarity to genes from Basidiomycota (*L. bicolor*, *C. cinerea* and *P. chrysosporium*) than those from Ascomycota (Table 2).

In order to find groups of similar proteins in our dataset, we applied a Markov Clustering (MCL) algorithm [43] to the *M. pernicioso* gene models. Although the output by this method is not highly reliable, they correlate well with "real" gene families and can be applied efficiently to cluster large quantities of genes in a high throughput fashion [44,45]. MCL also helped to the assemble genes without similarity in the GenBank into gene families, which are described in the following sections.

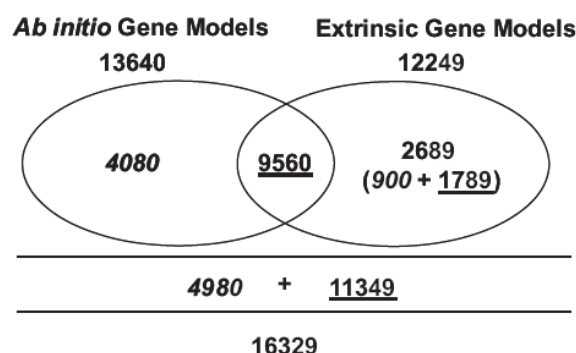


Figure 2
Number of *M. pernicioso* gene models predicted by *ab initio* and/or extrinsic prediction methods. Left ellipse: gene models predicted by *ab initio* methods. Right ellipse: gene models predicted by extrinsic methods. The intersection contains gene models detected by both methods. Underlined: number of gene models with BLASTX-NR E-value similarity $\leq 1e-10$. In italics: number of gene models with BLASTX-NR E-value similarity $> 1e-10$.

Those gene models that were not grouped into gene families by the MCL algorithm have been compared with the genome of *C. cinerea* in order to discard those genes with parts of their sequences present in more than one contig. In order to perform this comparison (TBLASTN) we assume that the length distribution of *M. pernicioso* proteins is similar to those from *C. cinerea*. In this comparison we evaluated similar genes according to the length of the protein (Fig. 3). This comparison showed that there is a clear correlation between complete genes in *C. cinerea* and *M. pernicioso* gene models coding for proteins smaller than 300 aa. However, for larger proteins more than one *M. pernicioso* gene model showed similarity to a single *C. cinerea* protein, thus confirming the overestimation of the number of gene models. Given the fact that the genome sequence is incomplete, there is high probability that we are predicting more than one gene model per gene.

In order to have a measurement of the overestimation and use it to correct our database, the corresponding relationship between the protein size and the number of gene models was calibrated with the genome of *C. cinerea* using

Table 2: Moniliophthora perniciosa predicted gene comparisons

Organism Comparison	Genes with* BLASTX hits	Genes with Top BLASTX Hits
<i>Laccaria bicolor</i>	9910 (60.6%)	4811 (29.27%)
<i>Coprinopsis cinerea</i>	10649 (65.1%)	3281 (19.97%)
<i>Phanerochaete chrysosporium</i>	9311 (57.0%)	2021 (12.30%)
<i>Fusarium graminearum</i>	5904 (36.1%)	57 (0.35%)
<i>Magnaporthe grisea</i>	5645 (34.5%)	39 (0.2%)
<i>Ustilago maydis</i>	5488 (33.6%)	42 (0.26%)
<i>Cryptococcus neoformans</i>	5454 (33.4%)	78 (0.47%)
<i>Neurospora crassa</i>	5127 (31.4%)	16 (0.10%)
<i>Saccharomyces cerevisiae</i>	3296 (20.2%)	4 (0.02%)
All NR database	11349 (69.5%)	—

* – Each *M. perniciosa* gene model was compared to all the proteins from the organisms listed in the table. BLASTX score was defined as 1e-10

the gene models unlinked to gene families identified by the MCL algorithm (Fig. 3). A linear regression formula was used to correct for the overestimation according to the protein size. For example, proteins smaller than 300 amino acids will have only one model representing the gene, while larger proteins will have one model and a fraction of a second model representing its gene. This fraction represents the overestimation of the number of gene models.

In order to estimate the gene density in *M. perniciosa*, and deal with the overestimation evident by the previous analysis, we averaged the total number of gene models obtained by the different predictions: *ab initio* predictions (13,640), extrinsic predictions (12,249) and total number of gene predictions (16,329). This resulted in 14,072 gene predictions that were divided by the 39 Mbp estimated genome size, which gave a gene density of approximately 2.77 ± 0.37 Kbp/gene or 0.36 ± 0.05 gene/Kbp for the genome of *M. perniciosa*. Figure 4 depicts the gene density

comparison between *M. perniciosa*, two eubacteria (*Xyllela fastidiosa* and *Escherichia coli*), one archaea (*Haloquadratum walsbyi*), an apicomplexan (*Theileria parva*), a primitive chordate (*Ciona intestinalis*) and a series of fungal genomes, and this showed that our gene density evaluation is in agreement with gene density data of other filamentous fungi [46].

The overall *M. perniciosa* genomic features are summarized in Table 3. The genomic clusters of *M. perniciosa* were submitted to NCBI (GenomeProject ID 28951, Locus-tag prefix MPER) and the Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the project accession ABRE00000000. The version described in this paper is the first version, ABRE01000000.

Overall Functional Annotation and Metabolic maps

The number of gene models used for the functional annotation and building of the metabolic maps was 16,329, which represents the total number of gene models

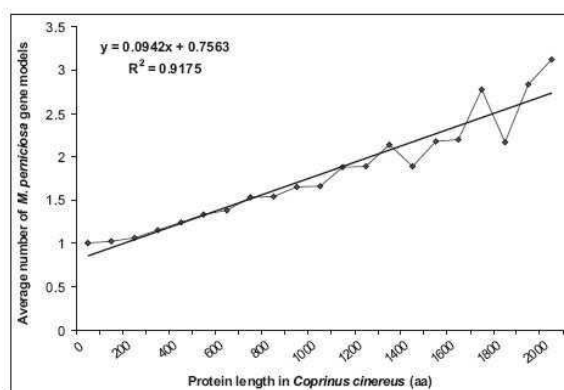


Figure 3
Correlation between the average number of *M. perniciosa* gene models and the length of *C. cinereus* proteins.

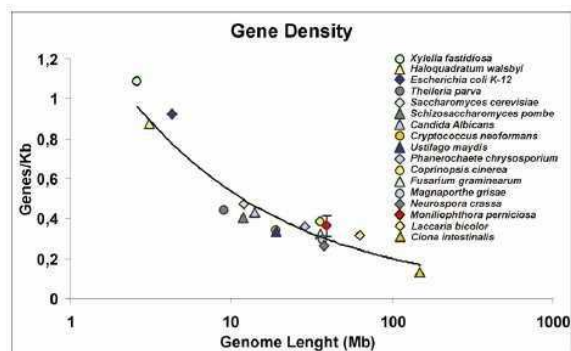


Figure 4
Correlation of various organisms' genome size and number of genes (Gene Density). Error bar in *M. perniciosa* data point depicts the Standard Deviation (SD ± 0.05) of Gene/Kbp ratio using *ab initio* predictions, extrinsic predictions and the sum of both predictions (see text).

Table 3: Moniliophthora perniciosa genome survey features

Genome survey features	Values
Contigs	17,991
Contigs size (average)	1.3 Kbp
Singlets	7,065
Singlets size (average)	455 bp*
Sum of all reads	75 Mbp
Assembly size	26.7 Mbp
Estimated genome size	39 Mbp
Coverage	1.92x
Gene models	
<i>Ab initio</i> gene models	13,640
Extrinsic gene models	12,249
Total gene models	16,329
Gene density	2.77 Kbp/gene (SD $\pm$ 0.39)
	0.36 gene/Kbp (SD $\pm$ 0.05)
Average gene size	651.4 bp
GC content overall	47.7
GC content coding	49.7
GC content non coding	46.8
Intron size (mode)	52 bp
Intron size (average)	63 bp
Exon size (mode)	60 bp
Exon size (average)	166 bp
AutoFACT annotation	
Classified (Hits)	11,950 (73%)
Assigned (Conserved expressed proteins)	7,416 (45%)
Unassigned (Conserved hypothetical proteins)	4,534 (28%)
Unclassified (No Hits)	4,379 (27%)

* – Singlet size after LUCY quality trimming (phred $\geq$ 16).

obtained using *ab initio* and extrinsic predictions. Although this number represents an overestimation of the real number of genes (see above), it was used to maximize the information acquired from these sequences. All data obtained in the gene families and protein domain analyses were corrected according to the normalization procedure developed from the gene models not included in any of the MCL families based on comparisons with the *C. cinerea* genome (see above).

The program AutoFACT [47], an automated annotation tool, was used to evaluate the putative functions of *M. perniciosa* gene models. This program determines the most informative functional description by combining multiple BLAST reports from a number of user selected databases, and provides a consensus result [47]. AutoFACT classifies proteins as "classified proteins", which can be "assigned" (containing similarity to proteins with annotated function) or "unassigned" (containing similarity to proteins with unknown function), and "unclassified proteins" (without similarity to any other protein in data-banks). Using this software, 73% of gene models (11,950) were annotated as previously classified proteins. From these 7,416 (45%) were assigned proteins, and 4,534 (28%) were classified as unassigned proteins, which can

be interpreted as a class of conserved hypothetical proteins (Table 3). The remaining 4,379 gene models (27%) were considered as unclassified proteins. The number of gene models with E-value scores $> 1e-10$ obtained by BLASTX-NR (no-hits, see previous section) was 4,980 and represented 30.5% of the total number of models. The result obtained with AutoFACT increased the number of classified proteins by 3.5% (Fig. 1; Table 3). This indicates that the use of domain classification databases is helpful in the identification and annotation of gene models.

Additionally, we correlated the AutoFACT annotation with the gene families assembled by the MCL algorithm. As depicted in Table 4 and additional file 2, MCL data indicated that the cytochrome P450 monooxygenase family had the largest number of *M. perniciosa* members, followed by a gypsy-like retrotransposon gene family, and a gene family that has similarity to a *C. cinerea* hypothetical protein (EAU86912.1) and to a shitake mushroom *Lenzula edodes* EST (EB016963). This new gene family appears to be related to specific developmental traits of the Agaricales. Also found were new gene families with unknown functions based on their lack of homology with GenBank sequences (Additional file 2; Additional File 3 – Worksheet Unknown Gene Families). These families

Table 4: MCL analysis of *M. perniciosus* gene models

MCL family ID	#members ^a	Norm Factor ^b	Norm #members ^c	Manual annotation ^d
1	175	1.24987	140	Cytochrome P450 monooxygenase
2	143	1.82632	78	Gypsy like-retrotranspon
4	84	1.35205	62	Conserved hypothetical protein
3	84	1.46889	57	Alcohol oxidases and GMC oxidoreductase
5	82	1.26479	65	Gypsy like-retrotranspon
6	76	1.57535	48	Serine threonine kinase
7	64	1.54312	41	Conserved hypothetical protein
8	59	1.54180	38	Conserved hypothetical protein
9	57	1.46370	39	Proteins containing WD-40 and NACHT domain
10	55	-	-	No hits protein family
11	52	1.34190	39	Cytochrome P450
12	50	1.52169	33	Serine threonine kinase
13	46	1.18098	39	Conserved hypothetical protein
14	43	1.21555	35	Cytochrome P450
15	41	1.45275	28	Carboxylesterase/lipase
16	39	1.67524	23	Gypsy like-retrotranspon
17	38	1.26598	30	Hexose transporter
19	35	1.11159	31	MFS transporter
18	35	1.43881	24	Splicing factor
20	32	1.17902	27	Laccase – Multicopper oxidase

a – Number of gene models present in each family

b – Normalization Factor used to normalize the number of gene models present in each family (see methods)

c – Normalization number of gene models present in each family

d – Manual annotation of MCL families according to BLAST similarity analysis

could possibly be linked to *M. perniciosus* physiological characteristics.

A comparison of the gene models with the CDD-PFAM databank [48] was performed to obtain information about protein domains present in *M. perniciosus* proteins. The data obtained were normalized using the procedure described above. Cytochrome P450 monooxygenase was the most prevalent protein domain assigned in *M. perniciosus* (gnl|CDD|40168), followed by protein kinases (gnl|CDD|40170), sugar transporters (gnl|CDD|40184), short chain dehydrogenases (gnl|CDD|40206) and carboxylesterases/lipases (gnl|CDD|40235) (Table 5; Additional File 4). The prevalence of Cytochrome P450 monooxygenase domains in the genome agrees with the results of MCL analysis. CDD-PFAM analysis was also used to evaluate the protein domains in other genome fungi (Additional File 4). The comparisons between *M. perniciosus* and the other fungi analyzed are described below.

A hypothetical metabolic map of *M. perniciosus* was built using BioCyc [49]. This analysis allowed us to annotate 235 metabolic pathways. These include 1358 enzymatic reactions incorporating a total of 2139 enzymes <http://www.lge.ibi.unicamp.br/biocyc/MP>. A comparison of this metabolic map with the fungal model *Saccharomyces cere-*

visiae S288C (documented with 132 pathways, 925 enzymatic reactions, and 675 enzymes – <http://www.lge.ibi.unicamp.br/biocyc/comparatives/MP-SCER>) showed that *M. perniciosus* has more metabolic pathways than *S. cerevisiae*; a result that corresponds to the smaller genome size of *S. cerevisiae* and possibly the more complex lifestyle of *M. perniciosus*. Interestingly, *M. perniciosus* has a higher number of reactions involving O₂, CO₂, H₂O₂, and NAD(P)⁺/NAD(P)H than *S. cerevisiae*, suggesting a greater capacity to use and deal with oxidation-reduction reactions (Additional File 5 – worksheet compounds). We also detected reactions with farnesyl pyrophosphate and dimethylallyl-diphosphate, which are involved in the biosynthesis of the secondary metabolites such as isoprenoids and indoles. Moreover, according to BioCyc analysis, *M. perniciosus* has more amino acid catabolic pathways, alternative carbon sources degradation and biosynthesis routes and C1 compounds (i.e., methanol) utilization and assimilation reactions than *S. cerevisiae* (Additional File 5 – worksheet pathways). These pathways are under manual annotation and will be published on BioCyc web page. The existence of these pathways in *M. perniciosus* suggest ecological and physiological adaptations to environmental stresses; to competition present in its native habitat in the Amazon Basin, and to traits that enable it to colonize cacao and trigger WBD.

Table 5: Top 20 CDD-PFAM domains in *M. perniciosus* proteins

CDD - ID ^a	Domain PFAM ^b	#Hits ^c Domains	%Hits ^d PTN	Rank ^e	Norm Factor	#Hits ^c Domains Norm	%Hits ^d PTN Norm	Rank ^f Norm
gnl CDD 40168	pfam00067, Cytochrome P450	256	1.57%	1	1.3565	188.71427	1.15%	1
gnl CDD 40170	pfam00069, Protein kinase	141	0.86%	2	1.6059	87.80184	0.54%	2
gnl CDD 40184	pfam00083, Sugar (and other) transporters	89	0.54%	3	1.6902	52.65651	0.32%	3
gnl CDD 40206	pfam00106, adh_short, short chain dehydrogenase	84	0.51%	4	1.7141	49.00442	0.30%	4
gnl CDD 40235	pfam00135, Coesterase, Carboxylesterase	57	0.35%	5	1.4738	38.67603	0.24%	5
gnl CDD 40207	pfam00107, ADH_zinc_N, Zinc-binding dehydrogenase	43	0.26%	8	1.2398	34.68221	0.21%	6
gnl CDD 40813	pfam00732, GMC_oxred_N, GMC oxidoreductase	52	0.32%	6	1.5402	33.76192	0.21%	7
gnl CDD 40345	pfam00248, Aldo_ket_red, Aldo/keto reductase family	41	0.25%	9	1.2881	31.83088	0.19%	8
gnl CDD 41245	pfam01185, Hydrophobin, Fungal hydrophobin	26	0.16%	21	0.8460	30.73119	0.19%	9
gnl CDD 40253	pfam00153, Mito_carr, Mitochondrial carrier protein	38	0.23%	12	1.3890	27.35860	0.17%	10
gnl CDD 40748	pfam00665, rve, Integrase core domain	50	0.31%	7	1.8943	26.39449	0.16%	11
gnl CDD 40493	pfam00400, WD40, WD domain, G-beta repeat	35	0.21%	13	1.5151	23.10130	0.14%	12
gnl CDD 40179	pfam00078, RVT, Reverse transcriptase	39	0.24%	11	1.7620	22.13448	0.14%	13
gnl CDD 45101	pfam05199, GMC_oxred_C, GMC oxidoreductase	28	0.17%	16	1.2741	21.97654	0.13%	14
gnl CDD 40588	pfam00501, AMP-binding, AMP-binding enzyme	40	0.24%	10	1.8537	21.57819	0.13%	15
gnl CDD 41412	pfam01360, Monooxygenase, Monooxygenase	25	0.15%	22	1.2253	20.40382	0.12%	16
gnl CDD 40107	pfam00005, ABC_tran, ABC transporter	33	0.20%	14	1.7198	19.18851	0.12%	17
gnl CDD 40367	pfam00270, DEAD, DEAD/DEAH box helicase	30	0.18%	15	1.5947	18.81230	0.12%	18
gnl CDD 40419	pfam00324, AA_permease, Amino acid permease	26	0.16%	20	1.4641	17.75798	0.11%	19
gnl CDD 41783	pfam01753, zf-MYND, MYND finger	25	0.15%	23	1.4288	17.49737	0.11%	20

- a - Number of CDD entry
b - Number and description of PFAM domain
c - Number of gene models containing each CDD-PFAM domain
d - Percentage of gene models containing each CDD-PFAM domain in relation to total number of *M. perniciosus* gene models (proteins).
e - Non-normalized CDD-PFAM ranking
f - Factor used to normalize the number of gene models containing each CDD-PFAM domain (see methods)
g - Normalized number of gene models containing a CDD-PFAM domain
h - Percentage of normalized number of gene models containing each CDD-PFAM domain in relation to total number of *M. perniciosus* gene models (proteins)
i - Normalized *M. perniciosus* CDD-PFAM ranking

Detoxification and general resistance mechanisms: cytochrome P450 monooxygenases, efflux transporters and anti-oxidative apparatus

Based on CDD-PFAM and MCL analyses gene members of the cytochrome P450 monooxygenase superfamily are prevalent in the genome of *M. perniciosa* (Tables 4 and 5; Additional File 3 – Worksheet P450; Additional File 4). Cytochrome P450 monooxygenases play a role in hydroxylation and oxidation processes involved in biosynthesis and degradation of different compounds [50]. Therefore, a large number of gene models similar to cytochrome P450 monooxygenases suggest a significant capacity for synthesis of secondary metabolites, such as hormones or toxins, and for detoxification. Among the fungi analyzed, *M. perniciosa* has the highest number of cytochrome P450 monooxygenase genes (188 gene models) representing 1.15% of the gene models (Fig. 5A). Saprotrophic basidiomycetes *P. chrysosporium* and *C. cinerea* and the hemibiotrophic ascomycetes *Magnaporthe grisea* and *Fusarium graminearum* also have more than one hundred cytochrome P450 monooxygenase genes representing between 1.25% and 0.9% of their gene models (Figure 5A; Additional File 4). The basidiomycete *L. bicolor* (an ectomycorrhizal fungus) and the ascomycete *Neurospora crassa* (a fire-scoured landscape colonizer) follow with fewer genes. *U. maydis* (a biotrophic pathogen), *C. neoformans* (an animal pathogen) and *S. cerevisiae* (a fermentative fungus) have the fewest cytochrome P450 monooxygenase genes of the fungi compared in this study (Fig. 5A; Additional File 4). This analysis clearly demonstrates the prevalence of cytochrome P450 monooxygenases in saprotrophic and hemibiotrophic fungi, which have to hydrolyze complex wood polymers and deal with a highly oxidative environment. As discussed by Gonzalez and Nebert [50], cytochrome P450 monooxygenase polymorphisms may be the product of the "molecular warfare" that occurs during the co-evolution of preys and predators, which produce toxins and detoxifying genes, respectively. This logic can be extrapolated to the plant-fungus interaction, and in that sense, we believe that the plethora of cytochrome P450 monooxygenases in *M. perniciosa* may be critical to detoxification and environmental adaptation as well as for disease development.

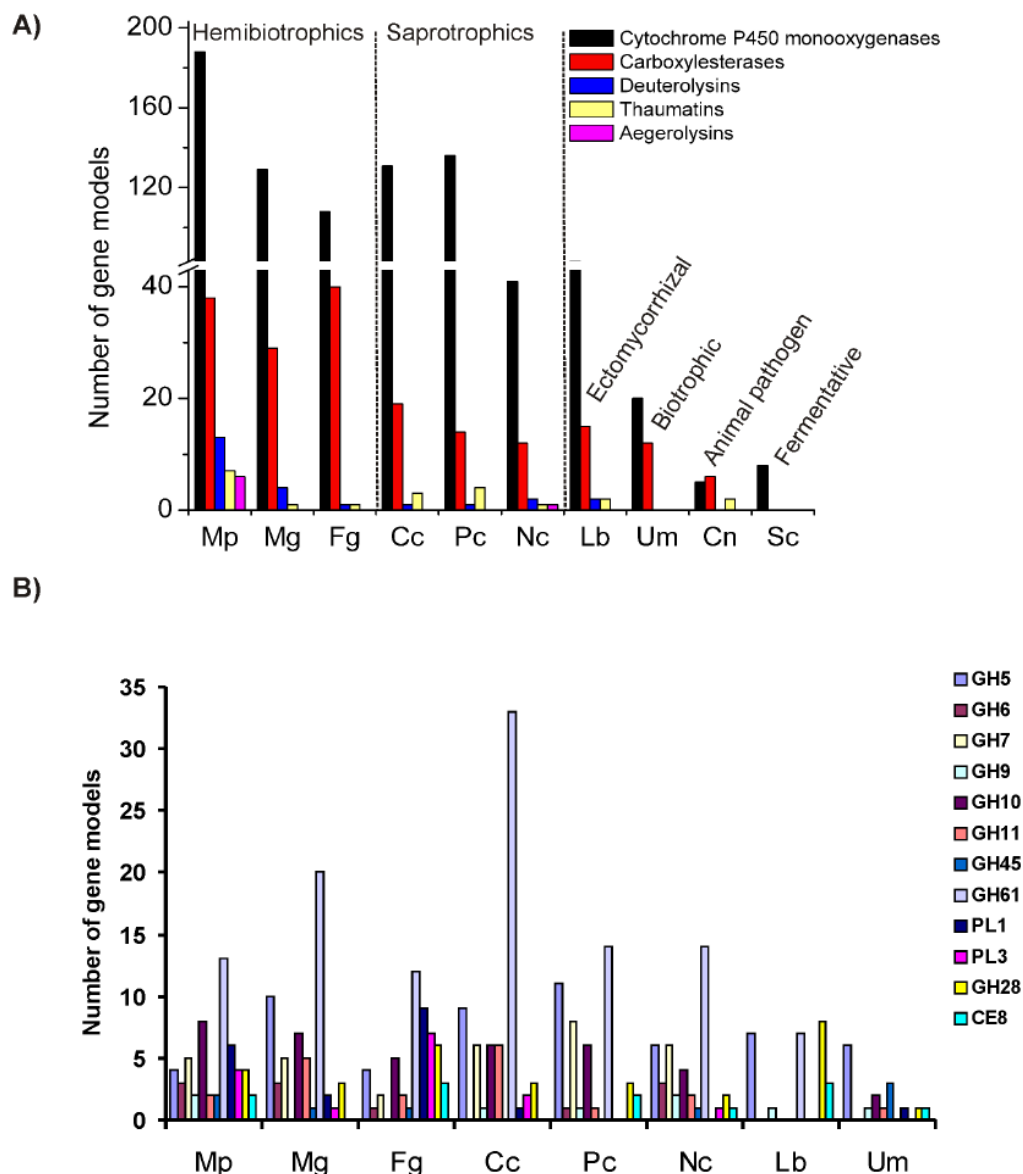
Another set of proteins related to detoxification processes are the efflux transporters. Similarity searches in the Transport Classification Database (TCDB – <http://www.tcdb.org>) [51] which compare the genome of *M. perniciosa* with other fungal genomes (Additional File 3 – Worksheet Functional annotation and Worksheet Transporters) verified an extended set of efflux transporters from classes 3.A.1 (ABC superfamily) and 3.A.3 (P-type ATPase Superfamily). The majority of ABC transporters identified in *M. perniciosa* belong to the Pleiotropic Drug Resistance (PDR) family and the ABC Conjugate (ABCC)

Transporter family. These proteins have been associated with fungal pathogenesis [52] and heavy metal resistance [53]. The members of P-type ATPase superfamily in *M. perniciosa* include phospholipid translocating ATPases family and fungal ENA-ATPases transporters, which are involved with the efflux of excessive Na⁺, and especially K⁺, encountered by fungi during colonization of plants [54].

During the plant defense, reactive oxygen species (ROS) are produced to limit the pathogen invasion [55]. However, pathogenic fungi produce antioxidant enzymes that enable them to neutralize host ROS. *M. perniciosa* contains a myriad of O₂⁻ and H₂O₂ decomposing enzymes such as superoxide dismutases, catalases, peroxiredoxins, glutathione-system, thioredoxin-system enzymes and manganese dependent peroxidases (Additional File 3 – Worksheet Functional Annotation). Therefore, this fungal genome harbors a complete ROS detoxification system. Reports indicate that H₂O₂ favors necrotrophic pathogens infection [56,57]. Furthermore, *M. perniciosa* produces calcium oxalate crystals (COC) [58], and a cacao susceptible genotype accumulates COC during *M. perniciosa* infection, followed by a programmed cell death (PCD) [59]. The degradation of COC produces carbon dioxide and H₂O₂, suggesting that COCs can be important to necrotrophic mycelia development. In addition, oxalate chelates Ca²⁺, an important secondary plant defense messenger and a key cross-linker of pectin in the middle lamella pectin [60], and was found to be a trigger of PCD in plants [61]. These findings suggest that oxalate favors *M. perniciosa* infection by disorganizing plant defense and plant cell wall structure, by facilitating the action of fungal pectinases and possibly triggering PCD in the later stages of WBD [59].

Genome variability: Mating-type genes and transposable elements

Seven *M. perniciosa* gene models were found to be similar to pheromone receptors (Additional File 3 – Worksheet Functional Annotation). As a primary homothallic fungus, *M. perniciosa* does not use its mating type system to outcross, but probably to promote the formation of clamp connections, hyphae dikaryotization and for the expression of pathogenicity genes as in *U. maydis* [62]. Previous reports have indicated that *M. perniciosa* exhibits high genetic variability at the molecular level [11,63,64]. This level of variability may be the reason *M. perniciosa* overcomes resistant genotypes of cacao, such as Scavina 6 [65]. Furthermore, it has been postulated that the genome variability found in homothallic *M. perniciosa* may be due to transposable elements (TEs) and ectopic recombination guided by the numerous copies of these elements found in the genome [10]. The fact that retrotransposons were identified in EST libraries and differentially expressed dur-

**Figure 5**

Comparison of *M. perniciosa* protein families with other fungi. (A) comparison between cytochrome P450 monooxygenases, carboxylesterases, deuterolysins, thaumatin and aegerolysins; (B) comparison of plant cell wall degrading enzymes from fungi that interact with plants. Mp = *Moniliophthora perniciosa*, Lb = *Laccaria bicolor*, Cc = *Coprinopsis cinerea*, Pc = *Phanerochaete chrysosporium*, Um = *Ustilago maydis*, Mg = *Magnaporthe grisea*, Cn = *Cryptococcus neoformans*, Fg = *Fusarium graminearum*, Nc = *Neurospora crassa*, Sc = *Saccharomyces cerevisiae*. The legend at the right refers to the nomenclature of plant cell wall degrading enzymes according to CAZy <http://www.cazy.org>.

ing *M. perniciosus* development [20], indicates that they are active elements, which could contribute to genetic variability. Among the *M. perniciosus* TE families, Gypsy-like retrotransposons were the most abundant, followed by Copia-like retrotransposons (Tables 4 and 5; Additional File 2; Additional File 3 – Worksheet Transposons). TEs similar to *P. chrysosporium* Copia-like elements were found in *M. perniciosus* genome inserted within putative cytochrome P450 monooxygenase genes [66]. Curiously, they were also inserted in a *P. chrysosporium* cytochrome P450 monooxygenase subfamily (Additional file 2 – family 255; Additional File 3 – Worksheet P450), suggesting that these TEs may have a common ancestral origin in Basidiomycota. Retroelements of the tyrosine recombinase (YR) order [67] and DNA transposons (class II transposons) from CACTA [68], hATC [69,70] and Tc/Mariner [71] superfamilies were also found. Finally, a previously described *Boto* DNA transposon [M.V. Queiroz *et al.*, unpublished data] from the PIF/IS5 superfamily [72] was also identified in a MCL family (Additional File 2 – family 251; Additional File 3 – Worksheet Transposons).

Plant Hormonal Disarrangement: Fungal genes related to plant hormones biosynthesis

There is growing evidence of phytohormones being produced by pathogens during some infective processes [73]. For instance, the production of gibberellins (GA); hormones involved in the regulation of stem elongation, seed germination, flowering and fruit maturation; have been identified in phytopathogenic bacteria and fungi that cause overgrowth symptoms, such as *Giberella fujikuroi* and *Sphaceloma manihoticola* [74,75]. A search for homologues of the fungal specific bi-functional ent-kaurene synthase (CPS/KS) responsible for the two-step cyclization of GGDP in fungi [76] identified gene models similar to the N-terminal domain of *G. fujikuroi* CPS/KS but did not detect any sequence similar to the C-terminal domain of this protein in *M. perniciosus*. Another gene model similar to CPS/KS that lacks the C-terminal domain was found in the *Aspergillus niger* genome (AM270241.1). Genes similar to GA-4 desaturase and GA oxidases (cytochrome P450), part of a GA biosynthesis gene cluster present in *G. fujikuroi* and *Phaeosphaeria* sp. were detected (Additional File 3 – Worksheet Functional Annotation). Reinforcing our data mining discover, is the fact that a gibberellin-like compound was detected in basidiospores of *M. perniciosus* [77]. We can theorize that the production of GA by *M. perniciosus* may confer the hyperplastic phenotype of the green broom that resembles stem hyper-elongation caused by GA-producing phytopathogens.

Another interesting discovery in the *M. perniciosus* genome is the presence of genes encoding enzymes of two biosynthetic pathways of indole-3-acetic acid (IAA), the most abundant natural plant auxin. We found a gene similar to

plant nitrilases (E.C 3.5.5.1) which catalyzes the direct conversion of indole-3-acetonitrile into ammonia and IAA. Additionally, potential genes for the IAA-producing tryptamine pathway (one tryptophan decarboxylase, copper amine oxidases and a hypothetical indole-acetaldehyde oxidase) were found (Fig. 6; Additional File 3 – Worksheet Functional Annotation). Furthermore, a manual annotation of EST libraries reveal the presence of an aromatic amino acid aminotransferase, which could make part of Indole-3-pyruvate IAA biosynthetic pathway (Fig. 6). IAA regulates many plant biological processes including cell elongation and fruit ripening. Recently, the presence of IAA in *M. perniciosus* basidiocarps was reported [78]. IAA is produced by other fungal phytopathogens such as biotrophics *U. maydis* and *Taphrina deformans* and hemibiotrophic *Colletotrichum gloeosporioides* sp. [79-81], and induces filamentation and invasive growth in *S. cerevisiae* [82]. Curiously, both Gibberellin and IAA induce fruit parthenocarp and act synergistically in plant organ expansion [83], both of which are traits of WBD.

Cacao plantlets inoculated with *M. perniciosus* spores emit higher levels of ethylene during the late stages of infection than non-inoculated plants [13]. The plant hormone ethylene stimulates elongation at low concentrations, and senescence, fruit ripening, and epinasty at high doses [84]. Auxins stimulate the synthesis of ethylene [85], which together with gibberellins have integrated actions in plant cell death [86] and stem elongation during phytochrome-mediated shade avoidance, a phenomenon that occurs in response to the low red to far-red light ratios (R:FR) under dense canopies [87]. We hypothesize that in a dense and shaded environment, such as a cacao plantation, the low R:FR ratio effects can be increased by the action of the aforementioned phytohormones, explaining the shade

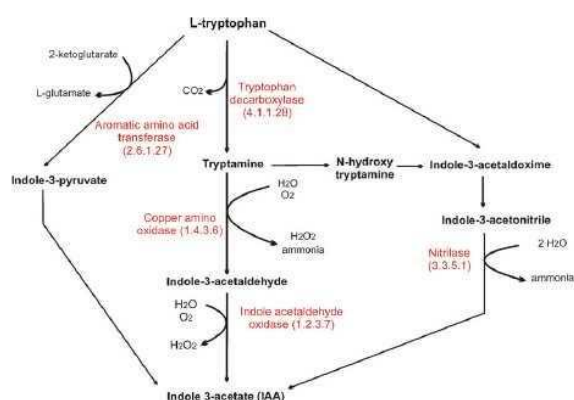


Figure 6
Indole-3-acetate (IAA) biosynthesis pathways. *M. perniciosus* gene models are annotated in red.

avoidance, hypertrophy, and elongation of green brooms. The presence of genes related to plant hormones production in *M. perniciosa* supports previous data [77,78] and suggests that this fungus can influence the plant metabolism and defense, by altering hormonal balance during infection [73,88].

Pathogenicity: Fungal effectors and pathogenicity associated proteins

As a pathogen that colonizes the plant apoplast during its biotrophic stage, *M. perniciosa* may release elicitor or effector proteins into the extracellular medium, which in turn could evade or suppress the plant defense response. Throughout WBD, *M. perniciosa* produces proteins with the potential to kill plant cells, thereby releasing their contents, which are absorbed by the fungus during its saprotrophic stage. Thus, *M. perniciosa* uses a varied arsenal of effector proteins in order to complete its infection cycle. Although some phytopathogenic fungi deliver effector proteins into the cytoplasm by means of haustoria [89,90] this type of structure is absent in *M. perniciosa*. Therefore it is possible that this fungus secretes these proteins into the apoplast, as has been described for other fungi [91]. Effector proteins that are recognized by plant resistance (R) proteins are known as *avr* proteins. No orthologues to the known Ascomycota *avr* genes, not even the *U. maydis* genes contained in "biotrophic clusters" [92], were found in the genome of *M. perniciosa*. However, *ab initio* gene prediction with peptide signal analysis revealed 70 "no hits" small proteins containing secretion signals and at least two cysteines (see methods, Additional File 6), which is a common trait of many proteins that are delivered into the host apoplast by phytopathogens [93]. In order to validate 22 selected *ab initio* predictions, we conducted RT-PCR using RNA from *M. perniciosa* saprotrophic mycelia. We validated the expression of 13 gene models out of 22 tested, all of which contained secretion signals (Additional File 7 – Primers No Hits Cys protein). Possibly, the gene models that were not confirmed by RT-PCR are expressed in other developmental stages of the fungus (basidiome, spore, etc.). Additional File 8 depicts the amplification of three of these genes. Whether these proteins play a role in the pathogenicity of *M. perniciosa* or in elicitation of cacao defense remains to be elucidated.

M. perniciosa contains Necrosis and Ethylene inducing proteins (NEPs) and cerato platanins [19,94], which can act in conjunction with a series of proteinases, hemolysin-like proteins and carboxylesterases/lipases found in the genome (Additional File 3 – Worksheet Functional Annotation). These proteins appear to be part of the destructive arsenal of *M. perniciosa*. The most abundant proteinases in the genome of *M. perniciosa* are deuterolysins, a type of fungal metalloproteinases that are similar to bacterial thermolysin [95]. Compared to other fungi indicates that *M. perniciosa* has a deuterolysin expansion (13 gene

models; Fig. 5A; Additional File 4), suggesting an important role for these proteinases during this fungus development.

We also identified a gene family similar to agaricales *Pleurotus ostreatus* and *Agrocybe aegerita* hemolysin-like aegerolysins. These proteins have cytotoxic properties [96] and seems to play an important role at the initial phase of fungal fruiting by making the fungal membranes permeable during cell signaling [96]. *M. perniciosa* genome contain a family of aegerolysins (6 aegerolysins; Fig. 5A; Additional File 4). None of the other agaricales fungi analyzed (*L. bicolor*, *C. cinerea* and *P. chrysosporium*) contain these proteins. Even though we could not assess the genomic data of the basidiomycete containing aegerolysins, we suggest that the diversification of these proteins in WBD causative agent indicates their importance in *M. perniciosa* development or even in fungi defense and infective process.

Carboxylesterases and lipases are overrepresented in *M. perniciosa* (Tables 4 and 5; Additional Files 2 and 4). According to CDD-PFAM, *M. perniciosa* have 38 gene models annotated as carboxylesterases, approximately twice the number of such proteins in other basidiomycete (ectomycorrhizal *L. bicolor* (15); saprobes *C. cinerea* (19) and *P. chrysosporium* (14); and biotrophic *U. maydis* (12); Figure 5A; Additional File 3); while the hemibiotrophic ascomycete *F. graminearum* and *M. grisea* have 40 and 29 carboxylesterases, respectively. We postulate that the great number of carboxylesterases and lipases in *M. perniciosa*, *F. graminearum* and *M. grisea* is related to their hemibiotrophic lifestyle. In fact, these enzymes are induced during carbon and nitrogen starvation [97], and cell wall degradation [98], two events that occur during hemibiotrophism.

Other genes associated with the plant-pathogen interaction found in our analysis were similar to the SCP-like superfamily proteins, which comprise pathogenesis related (PR) proteins of family 1 (PR-1). Additionally, gene models similar to PR-5/thaumatin superfamily were also detected in the *M. perniciosa* genome (Additional File 3 – Worksheet Functional Annotation). PR proteins are well described as associated with defense reactions in plants against various pathogens [99]. For instance, transgenic plants overexpressing PR-1 proteins were more resistant to oomycete infection [100] and some Thaumatin-like proteins (TLPs) have β -glucanase activity, inhibit xylanase and have antifungal properties [101-103]. Recently, proteins similar to PR-1 and thaumatin have been characterized in animals and fungi [104,105], indicating a conserved and important role in diverse organisms. Based on CDD-PFAM analysis, *M. perniciosa* contains the largest number of thaumatins of any fungus sequenced, so far (7 thaumatins, Fig. 5A; Additional File 4).

Both PR-1 and PR-5 are induced by salicylic acid (SA) in plants [99]. Curiously, the *M. perniciosa* necrotrophic (saprotrophic) mycelia were found to produce and have tolerance to SA in axenic cultures [70]. Plants with WBD have a higher content of SA when compared with healthy plants [106]. *M. perniciosa* tolerance to SA could be explained by the expression of genes encoding salicylate hydroxylases, which were also detected in the genome (Additional File 3 – Worksheet Functional Annotation). In this scenario, high levels of SA could block the synthesis of jasmonic acid (JA), a defense compound against necrotrophic pathogens that acts as a necrosis inducer, thus rendering the plant susceptible to the spread of *M. perniciosa* [78,106]. In addition, it is possible that SA and SA-induced proteins (i.e., PR-1, PR-5) may act to limit competition from other microbial competitors during WBD progression, which would be an important component of the *M. perniciosa* pathogenicity strategy.

Colonization: Plant cell wall degrading enzymes (PCWDE)

Degradation of hemicellulose, cellulose, pectin and depolymerization of lignin are some of the mechanisms that necrotrophic fungi use to colonize plant tissues [107]. We identified genes encoding enzymes involved in degradation of hemicellulose and cellulose, including β -1,4 cellulases, exocellobiohydrolases, endo-beta-1,4-xylanases and endoglucanases; genes encoding lignolytic enzymes including manganese dependent peroxidase and multi-copper polyphenoloxidases (laccases); and genes encoding enzymes involved in pectin degradation, such as pectate lyases, polygalacturonases (pectinases) and pectin methylesterases (pectinesterase) (Additional File 3 – Worksheet Functional Annotation). *M. perniciosa* have an arsenal of plant cell wall degrading enzymes that is similar to that found in the hemibiotrophic pathogens *F. graminearum* and *M. grisea* (Fig. 5B; Additional File 9). Biotrophic *U. maydis* and symbiont *L. bicolor* have a minimal set of PCWDEs what is in accordance with their lifestyle (Fig. 5B; Additional File 9).

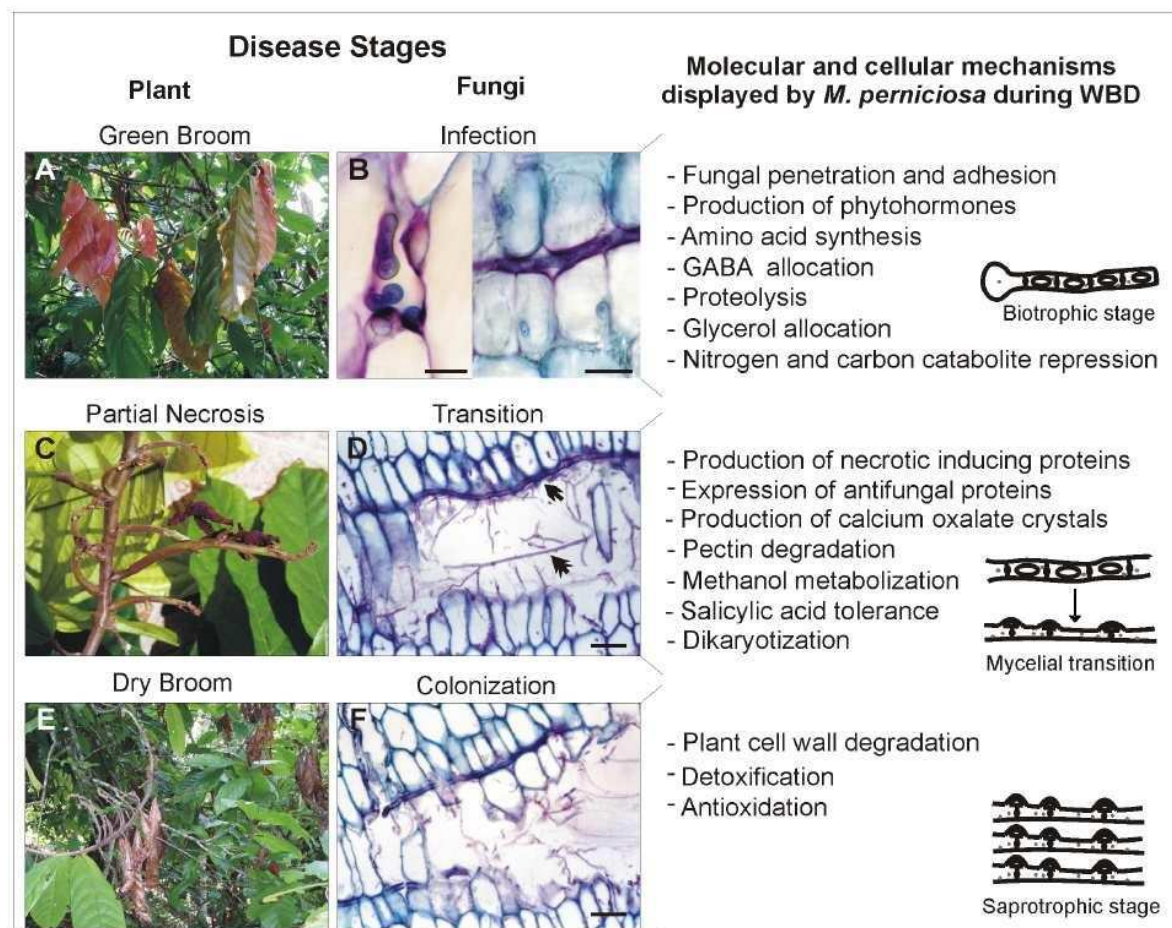
Pectate lyases (PL1 and PL3 according to CAZy nomenclature – <http://www.cazy.org>) cleave pectin, an essential component of plant cell walls. Among the fungi analyzed, *F. graminearum* and *M. perniciosa* contain the largest number of PLs (Fig. 5B; Additional File 9). Unlike *M. grisea*, the other hemibiotrophic analyzed, *F. graminearum* and *M. perniciosa* does not have specialized structures (appressoria) for non-enzymatic penetration of plants, and colonizing the apoplast by breaching the middle lamella barrier (Fig. 7). In addition, both are able to infect dicotyledons that contain cell walls with larger amounts of pectin than monocots [108]. This analysis suggests that PLs have an important role for pathogens that colonize the apoplast of dicotyledons.

Insights into *M. perniciosa* intermediary metabolism and WBD development

Pathogen energy status and the acquisition of host metabolic compounds by the pathogen are factors that determine the outcome of disease onset. Therefore, enzymes of *M. perniciosa* involved in intermediary metabolism (i.e., reactions concerned with storing and generating metabolic energy) may have an important role during WBD (Additional File 3 – Worksheet Functional Annotation). During the first stages of WBD, *M. perniciosa* may encounter a nutrient-poor and extreme oxidative environment containing host defense molecules, such as nitric oxide (NO), an inhibitor of the cytochrome respiratory pathway [109]. Under these conditions, the activity of a NO tolerant alternative oxidase (AOX) may constitute a critical bypass mechanism for the cytochrome pathway [110-112]. A single copy of AOX was detected in *M. perniciosa*, and preliminary experiments indicate that this gene has a higher expression in fungi grown in nutrient-poor media (data not shown).

Following this scenario, in the low-energy status represented by the biotrophic phase, AOX may provide NAD⁺ for the turnover of the mitochondrial tricarboxylic acid cycle (TCA) and the peroxisomal glyoxylate cycle (GLOX). Many reports have documented the importance of GLOX for phytopathogens during host infection [113-115]. All genes coding for proteins of this enzymatic pathway are present in the genome of *M. perniciosa*, including the key enzymes isocitrate lyase and malate synthase (Additional File 3 – Worksheet Functional Annotation).

Curiously, *M. perniciosa* is able to use methanol as the only carbon source (data not shown), indicating that this fungus may have a methylotrophic metabolism. Methylo-trophism have been extensively studied in methylotrophic yeasts such as *Pichia angusta* and recently methanol oxidases (MOX) have been described in filamentous fungi such as *C. fulvum* [116] and wood-degrading basidiomycete *Gloeophyllum trabeum* [117]. A gene encoding a MOX was identified in *M. perniciosa* genome (Additional File 3 – Worksheet Functional Annotation). Interestingly, this gene was previously detected as being overexpressed in biotrophic-mycelia [20]. The methanol catabolism enzymes formaldehyde dehydrogenase and formate dehydrogenase (Additional File 3 – Worksheet Functional Annotation) were also found in the genome, which provides evidence that *M. perniciosa* indeed hydrolyzes methanol. Methanol is, possibly, derived from the first step of pectin hydrolyzation performed by the cell wall degrading enzyme pectin methylesterase [118], or from demethylation of lignin that occurs after manganese peroxidase and/or laccase action [117]. We can not discount the possibility that the biotrophic fungi can use the methanol produced by pectin metabolism during normal

**Figure 7**

An illustrated hypothetical model for WBD. Model correlating classical symptoms of green and dry broom in the field (A, C and E), *M. pernicioso* development inside cacao (B: biotrophic stage, D: transition from biotrophic to saprotrophic stage, F: saprotrophic stage) and molecular and cellular events displayed by the fungus in each developmental stages, based on genes annotated in genome survey (right side of the panel). Notice in B (right side) the presence of biotrophic mycelia in the apoplast surrounded by intact living cells. Also notice in C the presence of biotrophic mycelia (arrowhead) and saprotrophic mycelia (double arrowhead) inside a necrotic region. Micrographs scales: B left side: 15 μ M; B right side: 25 μ M; D: 50 μ M; F: 50 μ M.

cell wall synthesis in early stages of plant expansion [119]. The possibility that *M. pernicioso* is methylotrophic resembles the methyllobacteria nutritional strategy, which provides an ecological advantage over non-methylotrophic microorganisms present in the phyllosphere [120].

Previous data indicated that the utilization of glycerol, instead of fermentable sugars (i.e., glucose), is an important environmental clue for the maintenance of the biotrophic stage [18,121]. Accordingly, our group detected higher amounts of glycerol during the biotrophic fungal phase of the green brooms development than in dry brooms [13]. Additionally, *in vitro* assays showed that the

shift from glucose to glycerol media increased anti-oxidative defenses of *M. pernicioso* mycelia [122]. This result correlates well with green broom environment (high content of glycerol and ROS) raising the possibility that glycerol is a critical metabolite during the initial stages of the disease cycle. Genes involved in glycerol metabolism and uptake are present in *M. pernicioso*, including a biotrophic induced aquaglyceroporin transporter [20] (Additional File 3 – Worksheet Functional Annotation), suggesting that this fungus is able to acquire extracellular glycerol.

After 35 days of WBD, glucose levels increase again, concomitant with a reduction in starch levels [13]. We found

a gene model similar to a secreted glucoamylase in the *M. perniciosa* genome (Additional File 3 – Worksheet Functional Annotation). Gibberellin is an inducer of α -amylase production [123]. We can envisage a disease scenario in which the fungus produces an extracellular amylase and hormones (i.e., gibberellin) that triggers plant amyolytic activity, which then decrease starch content. The resulting glucose can be utilized by the plant since at this stage the green brooms/infected tissues are rapidly growing or even be acquired by the pathogen at the transition phase between biotrophic and necrotrophic stages.

Nitrogen starvation also appears to be a factor that influences the biotrophic lifestyle [124] and is associated with the expression of pathogenicity genes and PCWDE, mainly in hemibiotrophic and biotrophic fungi that deal with nutrient deprivation during early infection [125]. Our data from microarray and EST analysis suggests that nitrogen catabolite repression (NCR) occurs in *M. perniciosa* by the induction of GABA permease, tRNA synthetases and AROM protein [20]. As reported previously [20] *M. perniciosa* contains a gene homologous to *CLNR1* from the hemibiotrophic fungus *C. lindemuthianum* (Additional File 3 – Worksheet Functional Annotation), a global nitrogen regulator that belongs to the AREA/NIT2 family. *CLNR1* activates enzymes and transporters that enable uptake and catabolism of secondary nitrogen sources [126]. The depletion of *CLNR1* impaired the fungal switch to necrotrophy [124], emphasizing the importance of nitrogen catabolism in hemibiotrophic development. Our genomic data mining identified a gene similar to *NPR2*, which encodes a regulatory protein that may act upstream of the AREA/NIT2 protein (Additional File 3 – Worksheet Functional Annotation). *NPR2* is required for the expression of the *M. grisea* pathogenicity gene *MPG1* [127]. In addition, genes that encode enzymes involved in alternative nitrogen sources uptake (i.e., GABA transporter, urea permease and nitrate transporter) and metabolism (i.e., urease, nitrate reductase, nitrite reductase, arginase and uricase) were identified (Additional File 3 – Worksheet Functional Annotation). The presence of such genes in *M. perniciosa* genome indicates that this fungus could allocate and utilize alternative nitrogen sources in the absence of preferential nitrogen sources (glutamine and ammonia) reinforcing our hypothesis that *M. perniciosa* suffers NCR during early stages of WBD.

Conclusion

Our analysis of the *M. perniciosa* genome survey yielded interesting insights and clues into the molecular mechanisms underlying WBD. As far as we know, this is the first phytopathogen included in the order Agaricales sequenced. Therefore, our results support the investigation of pathogenicity mechanisms among Agaricales and Basidiomycete. In addition, we provide an approach for normalization of gene family data in a genome survey that

can aid the genomics community interested in functional analysis in incomplete genome data.

Based on annotated fungal genes from this report and from previous other studies, we designed a hypothetical model for WBD that correlates plant phenotype changes that happen during the disease with the developmental progression of *M. perniciosa* (Fig. 7).

After penetration and adhesion to the plant, the biotrophic fungus slowly grows inside the apoplast causing a series of phenotype changes in cacao, such as hypertrophy and hyperplasia, phototropism and epinasty, by secreting phytohormones that unbalance cacao metabolism (Fig 7A and 7B). In addition, *M. perniciosa* has to deal with nitrogen deprivation in the apoplast, which signals the production of proteins related to the acquisition of alternative nutrient sources, proteolysis and amino acid synthesis. Furthermore, the fungus takes advantage of the increasing content of glycerol in the green broom, an important cue for biotrophic stage maintenance, by expressing aquaglyceroporins. As previously discussed [13], the starch accumulated in early stages of green brooms seems to be metabolized to glucose, which suggests an amyolytic activity exerted by fungi and/or by plant amylases. This increase of glucose is not accompanied by an increase of fructose, but by a sucrose augmentation. Since photosynthesis is not increased during WBD, we suggest that sucrose is translocated from other tissues to the green brooms. Therefore, we hypothesize that these mechanisms cause a source-to-sink transition in stem, turning green brooms into a drain of nutrients.

After numerous physiological and biochemical changes in the plant, which may be caused by the fungal infection, there occurs a transition from the biotrophic to the saprotrophic lifestyle (Fig. 7C and 7D). This change could result from the increase of nutrients in the fungal environment and may be controlled by an AREA/NIT2-like regulator. During this transition phase, the plant displays the beginning of necrosis at the distal portion of the leaves that could be due to the action of NEPs and cerato-platanins that are expressed in the biotrophic hyphae. In addition, *M. perniciosa* produces PCWDEs, such as pectinases, whose action aids the fungus in breaching the middle lamella barrier. Pectin degradation releases methanol, which in turn could be used by *M. perniciosa* as a carbon source, through the action of a MOX and other methanol metabolizing enzymes. Moreover, the calcium released from pectin disruption could be scavenged by the oxalate synthesized by the fungus, a compound that triggers cell death. The release of cell content during necrosis, and the probable aforementioned source-to-sink transition, may influence, or even be indispensable to the *M. perniciosa* switch from biotrophism to saprotrophism. Thus, the postulated carbon and nitrogen catabolite repression dis-

played by the biotrophic mycelia would be switched off, thus causing the mycelial change to its invasive dikaryotic/saprotrophic stage.

During disease progression, and mainly during colonization of saprotrophic hyphae, *M. perniciosa* must deal with an intense oxidative environment. Based on our analysis this stress can be overcome by the action of several antioxidative and detoxifying enzymes. Furthermore, we believe that *M. perniciosa* exerts a negative control on plant defense against necrotrophic/saprotrophic fungi by producing salicylic acid, which would limit competition by other fungi by the action of antifungal proteins. Finally, after alternating wet and dry periods, the formation of the basidiomes produced by saprotrophic hyphae occurs.

Latin American cacao crops suffer tremendous damages caused by WBD, which mainly affects small acreage farmers. Such impact in the relatively primitive cacao cultivation system not only affects the socio-economic status of farmers but also the preservation of the rainforest. The demand for strategies that limit cacao diseases requires an intense effort in understanding the pathogenicity and plant resistance mechanisms. Further sequencing projects of cacao and its pathogens will serve as a background for the integration of transcriptomics, proteomics and metabolomics of these species in a systems biology approach. Such initiatives will provide tools for biological control, crop management and cacao biotechnology to combat cacao diseases. We believe that our report is the first step towards such an integrative initiative and provides insights into the molecular mechanisms of WBD which can aid the cacao's WBD-concerned community to develop control strategies for this plant-fungus interaction.

Methods

Biological material, libraries construction and sequencing

Total DNA was extracted from saprotrophic hyphae of *M. perniciosa* strain FA553 (CP02) maintained in Malt Yeast Extract Agar (Difco) at 27°C. DNA was extracted from grounded mycelia by incubation in CTAB buffer (CTAB 3%, NaCl 1.4 M, EDTA 20 mM pH 8.0, Tris-HCl 10 mM pH 8.0, PVP 1.0%, β -mercaptoethanol 0.2%) at 65°C during 30 min; followed by one phenol:chloroform:isoamyl alcohol (25/24/1) wash, precipitation with sodium acetate pH 5.2 (0.1 Vol) and cold 100% ethanol (2 Vol). DNA was eluted in deionized water and sheared by nebulization and sonication into fragments of approximately 2 Kbp, which were size selected on agarose gels and purified with S.N.A.P. Gel Purification Kit (Invitrogen – Life technologies, USA). DNA fragments were blunt-end ligated into the pCR4Blunt plasmid (Invitrogen – Life technologies, USA). Approximately 50 genomic libraries were constructed, each one corresponding to individually

growing cultures. Sequencing was done in an ABI Prism 3700 sequencer (Applied Biosystems, USA).

Clustering

The resulting chromatograms were submitted to the *M. perniciosa* database and subjected to automatic base calling using the software PHRED [128]. The contaminating vector sequences and low quality shotgun reads, without at least 100 bp with phred score ≥ 16 were trimmed by using the program LUCY [129]. Shotgun reads showing significant sequence similarity (BLASTn, E-value $\geq 1E-30$) with *M. perniciosa* mitochondrial sequences (see above) were removed from subsequent assemblies. The clustering and assembly were performed using the software PHRAP <http://www.phrap.org>. Afterwards, low quality regions of singlets previously evaluated by LUCY (the last window of 10 bp that has an average probability of error given by phred ≤ 10) were trimmed. The remaining clusters were subjected to similarity searches against the NCBI non-redundant protein and nucleotide database using the BLASTx and tBLASTx, respectively, with an E-value cutoff of 1E-5.

Genome Length Statistical Validations

Statistical analyses of the genome length were performed using two approaches. The first one was based on the Dog genome survey using counting of start positions offsets for overlapping reads [31] (for further details see Additional File 1). The second was based on Lander Waterman Theory, [34] and their applications [35,36], which estimate the theoretical values of expected number of clusters (contigs + singlets), contigs, gaps, average cluster size and average gap size, using the effective average read length (L), the total number of reads in the assembly (N) and the estimated genome size (G). L is the average number of base pairs of a read that contributes to the contig through parsing of ace file <http://www.phrap.org>. In our analysis L was equal to 550 bp. In order to estimate the gap size distribution in the *M. perniciosa* genome survey, we performed a comparison between a set of eukaryotic core protein (generated by CEGMA pipeline [37]) and *M. perniciosa* contigs using TBLASTN with threshold of 1e-10 for the E-value (Further details in Additional File 1). The estimation of misassembled sequences due to repetitive regions in the genome was performed using the integrated pipeline amosvalidate [38] (Further details in Additional File 1).

Gene Finding

EST against genome alignments

The alignment of ESTs with genomic sequences was performed using the package GeneSeqer [130] with the pre-built *Aspergillus* intron model. 300 highly confident introns were selected and used as an input for Exalin program [39] that is able to build a splice site model for an organism. The positions of the splice sites as assigned by

Exalin were used to rank overlapping gene predictions (see below).

Ab initio gene models prediction

The *ab initio* gene models prediction was performed with the trainable, open source gene predictors AUGUSTUS [40], SNAP [41] and GENEZILLA [42]. Ten copies of an artificial sequence of 240 Kbp (total of 2.4 Mbp) formed by the concatenation of the *M. perniciosus* ESTs coding regions, together with a *C. cinerea* gene dataset containing 1.2 Mbp were submitted to "pre-training" in AUGUSTUS gene predictor. *M. perniciosus* resulting predictions were compared with the protein databank NR using BLASTp. The predictions with similarities in NR, and with coverage $\geq 90\%$, were selected. After redundancy elimination, *M. perniciosus* gene models were used to train the three gene finders aforementioned. Predictions with less than 30 amino acids were eliminated, and the remaining predictions were grouped in overlapping clusters.

Ab initio gene models ranking

The predictions in each overlapping cluster were ranked according to the criteria used by the Fungal Genome Initiative at Broad institute <http://www.broad.mit.edu>. In each cluster, the "best" *ab initio* gene model according to the stipulated criteria was selected for functional annotation. The criteria for the ranking of the gene models were the following:

1. Manual annotation had priority over all other evidences;
2. Predictions with EST evidences had priority over the predictions without EST evidences;
3. If two predictions had EST evidences, the one with more splice sites in exact agreement with ESTs had priority;
4. Prediction with similarities with known proteins had priority. A prediction was considered to be similar to some known protein if it had an E-value of at most $1e-10$ (BLASTP against NR+*Phanerochaete chrysosporium* protein set);
5. If two predictions had similarity with known proteins, the one with better coverage score had priority. The coverage score was defined as $2 \times CP \times CH / (CP + CH)$, where CP is the coverage of the prediction and CH is the coverage of the similar protein;
6. In clusters without similarity with known proteins and without EST evidence, the priority was for AUGUSTUS, SNAP and GENEZILLA, in this order. This criterion was chosen according to the performance of the three pro-

grams in a dataset of 60 genes structures visually inspected.

A final filter discarded gene predictions reported by only one program, without similarity to known proteins and without EST evidence.

Extrinsic gene models prediction

The extrinsic gene model predictions were performed by two methodologies. First, 17,991 contigs and 7,065 singletons were submitted to similarity analysis in a databank containing BLASTX-NR plus *Phanerochaete chrysosporium* proteins. The genomic regions containing homologues in this databank were selected and assigned as putative gene models. GenomeThreader [131] program was used to make protein-DNA spliced alignments between the BLAST first hit against and the genomic sequence, serving as a guide to delimit the start and stop codons and exon-intron boundaries of the regions of the contigs containing similarity with GenBank.

Concurrently, *M. perniciosus* ESTs aligned with genomic clusters (see above) were inspected to verify if the region in which they aligned contained a BLAST extrinsic prediction gene model. These extrinsic gene models (EST and BLAST) were compared with each other to evaluate the amount of gene models predicted by these methods. After these comparisons, the extrinsic gene models were divided into 4 datasets:

- (i) ESTMODELS: retrieved from the spliced alignments of the ESTs against the genomic clusters not covered to a BLAST extrinsic gene models prediction. Low score spliced alignments and ESTs that seen clearly to be UTR of a neighboring prediction were not included.
- (ii) BLASTMODELS: derived from BLAST extrinsic gene models predictions analysis covering genomic regions without EST evidence.
- (iii) COMBINEDMODELS: gene models derived from genomic sequence regions with BLAST hit and EST evidence.
- (iv) CURATEDMODELS: extrinsic predictions manually annotated for manual correction of merged or split predictions. The genomic survey and gene models nomenclature are depicted in additional file 10.

tRNA prediction

For tRNA prediction, the tRNAscan-SE program [132] was taken into account with the default parameters, which searched for conserved sequences and the characteristic secondary structure of tRNAs.

MCL-families clustering

MCL graph clustering algorithm was applied to generate *M. perniciosus* gene families using WU-TBLASTx "all against all" as the tool used for aligning the gene models [43].

Normalization of gene family data

The normalization of gene family data was performed by comparison between *C. cinerea* proteins and *M. perniciosus* gene models, using TBLASTN with $1E-10$ of E-value threshold. The number of gene models similar to a *C. cinerea* protein was plotted according to protein length. The equation that estimates the number of gene models representing the same protein was generated using linear regression fitting. This equation was used to estimate a normalization factor to each MCL family and CDD-PFAM domain according to the average of protein length of their members.

Automatic Annotation and Metabolic Maps

The automatic annotation program AutoFACT [47] was used for functional annotation of gene models. The set of coding sequences from gene models were submitted to similarity searches against the UNIREF100, UNIREF90, NR, and KEGG databases using BLASTx (E-value $\leq 1E-5$) and against CDD-PFAM using RPS-BLAST (E-value $\leq 1E-5$) [48]. These results were submitted to AutoFACT, which searches for a consensus in the results and output descriptions and statistics about protein domains and families.

For an inference of *M. perniciosus* metabolic maps, we used Pathway Tools (version 11.0), a software of BioCyc databases [49], which generates a metabolic map from a previously annotated genome. The pathways that are probably present in the genome are imported from a reference database, following the Pathway Tools parameters [133]. The annotated genome input was obtained from EST manual annotation and from *M. perniciosus* gene models AutoFACT annotation, using as main information the product name and, if available, E.C. numbers. Metabolic pathways of interest were manually annotated for the elimination of false positives.

The analysis of transporters was made based on a BLASTX search of *M. perniciosus* gene models and other fungal genes against TCDB (Transport Classification Database – <http://www.tcdb.org>), using a threshold of E-value $1E-05$. All classes from third level that contained at least one species with 2% or more of representations were separately represented in the results.

Selection and expression confirmation of no hits *ab initio* gene models

The selection of *ab initio* gene models was performed using a SQL query wizard. As input, we ask for gene models without similarity in GenBank NR, which encoded

proteins that contained at least 2 cysteines and a signal peptide, previously identified by Signal-P 3.0 program [134]. 74 gene models were then selected and their nucleotide sequences were used as template for the design of primers nested in: (i) the sequence encoding the putative signal peptide (SPE); (ii) the sequence encoding the putative first amino acid of mature protein (MAT) and; (iii) the sequence containing the putative stop codon (END). The latter was designed in reverse complement ("reverse") to allow gene amplification using the other two primers ("forward"). RT-PCR analysis was performed to validate the expression of *ab initio* predicted gene models. RNA from saprotrophic mycelia was extracted using hot-phenol method with modifications [135]. Equal amounts of total RNA from CP02 saprotrophic mycelia cultures (24 h, 48 h, 4 days and 7 days) were mixed. After DNase (Invitrogen, USA) treatment, 2 µg of total RNA was reverse transcribed using Superscript RTIII (Invitrogen, USA) in a total volume of 20 µL, following the manufacturer's instructions. PCR reactions were conducted according to primers (MWG, Imprint Genetics Corp) temperature of melting (TMs).

Abbreviations

ABC: ATP Binding Cassette; AOX: Alternative oxidase; bp: base pairs; COC: calcium oxalate crystals; CTAB: Cetyl trimethylammonium bromide; EDTA: Ethylenediaminetetraacetic acid; EST: Expressed Sequence Tag; FPR: Frosty Pod Rot; GA: Gibberellin; GABA: Gamma-aminobutyric acid; GGPP: Geranylgeranyl Diphosphate; IAA: indole-3-acetic acid; JA: Jasmonic acid; Kbp: One thousand base pairs; LW: Lander Waterman; Mbp: One million base pairs; MOX: Methanol oxidase; NAD(P): Nicotinamide adenine dinucleotide phosphate; NCR: nitrogen catabolite repression; NEP: Necrosis and ethylene-inducing proteins; NO: Nitric oxide; PCD: Programmed cell death; PCWDE: Plant cell wall degrading enzymes; PVP: Poly(vinylpyrrolidone); ROS: Reactive oxygen species; RT-PCR: Reverse Transcription-Polymerase Chain Reaction; SA: Salicylic acid; PCR: Polymerase Chain Reaction; PR: Pathogenesis related protein; TCA: Tricarboxylic acid cycle; TE: Transposable Elements; TM: temperature of melting; tRNA: transfer RNA; UTR: untranslated region; WBD: Witches' Broom Disease.

Authors' contributions

JMCM conceived and wrote the article, was responsible for gene models annotation, gene density analysis, primer design and data interpretation. MFC conceived bioinformatics analysis and was responsible for genome assembly, BLASTs, statistical analysis and data interpretation. GGLC was responsible for gene models predictions, Genome Threader analysis, MCL clustering analysis and gene browser design and implementation. EFF participated in genome assembly, was responsible for transporters analy-

sis and annotation, and made figures edition. LPP was responsible for AutoFACT and BioCyc implementation and analysis. JR was responsible for EST library construction and sequencing and participated in genome annotation. CC, DMC, AFC and HC were responsible for *M. perniciosus* DNA extraction, genome shotgun, cloning and maintenance of genomic libraries. ROV was responsible for gene browser design and implementation, tRNA analysis and *M. perniciosus* sketches. RCE performed RT-PCR and aided in primer design. OG, DPTT BVO and MHM participated in genome annotation and were responsible for the characterization of genes encoding NEPs, AOX, MOX and auxin biosynthetic proteins, respectively. ABLP was responsible for EST library construction. MCSR was responsible for microscopic analysis. MRAR participated in genome assembly and in the development of bioinformatics tools. LABC, KPG, MSG, AGN and LBB were coordinators of sequencing groups. JPMN was responsible for sequencing and libraries maintenance. MJG and BAB participated in final manuscript elaboration and provided additional sequencing in order to close gene gaps. LWM participated in pathogenicity gene data interpretation and final manuscript elaboration. JCC was the co-coordinator of *M. perniciosus* genome project and coordinated UESC sequencing group. GAGP conceived and was the coordinator of *M. perniciosus* genome project, participated in data interpretation and in final manuscript elaboration.

Additional material

Additional file 1

Genome statistical validations. A) Estimation of genome length using dog genome survey protocol, B) Estimate of distribution of gap sizes in *M. perniciosus* genome assembly, C) Estimate of misassembly sequences due to repetitive regions.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S1.pdf>]

Additional file 2

MCL analysis of *M. perniciosus* gene models. All predicted proteins were compared all-against-all using WU-TBLASTX. A score ($-\log(E\text{-value})$) for each pair of proteins (u, v) with significant BLAST hits (Evalue $1e-5$) was assigned. The MCL algorithm (inflation parameter 2.0) was applied to find clusters in this graph. This method is fully automatic and protein clusters reported were not subjected to manual curation. ID: number of the MCL family; #members: number of gene models present in each family; Norm factor: factor used to normalize the number of gene models present in each family (see methods); Norm#members: normalized number of gene models present in each family. Annotation: words associated to each family after correlation of gene models with AutoFACT annotation. In parenthesis are the occurrence numbers of each word. Each worksheet shows the ranking of families using normalization factor (Rank_Norm) or not using this factor (Rank_Non_Norm).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S2.xls>]

Additional file 3

Functional annotation of *M. perniciosus* gene models discussed in this paper. ID: gene model; First Hit (BLASTX-NR): Most similar sequence in GenBank; E-value: E-value of most similar sequence; AutoFACT annotation: automatic annotation by AutoFACT; AutoFACT E-value: E-value of AutoFACT annotation; EST: presence (Y) or absence of an EST aligned in this gene model; MCL family: family annotated by MCL analysis.

Worksheet P450: annotation of gene models similar to cytochrome P450 monooxygenases; Worksheet transposons: classification and annotation of gene models similar to transposable elements; Worksheet unknown gene families: annotation of top 20 MCL unknown gene families; Worksheet functional annotation: classification and annotation of gene models similar to efflux transporters, anti-oxidative enzymes, phytohormones biosynthesis related proteins, pheromone receptors, salicylate hydroxylases, effectors/elicitors/pathogenicity associated proteins, cell wall degrading enzymes and intermediary metabolism enzymes (cytochrome pathway bypass, Glyoxylate pathway and oxalate formation, glycerol uptake and metabolism, extracellular sugar degrading enzymes and nitrogen regulation, uptake and metabolism enzymes), EC = enzyme classification (<http://expasy.org/enzyme/>); Worksheet transporters: Relative percentage of transporters distribution in fungi genomes (see methods).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S3.xls>]

Additional file 4

Ranking of CDD-PFAM families annotated in *M. perniciosus*. Gene models were annotated based on CDD-PFAM-ID and ranked. This analysis was performed with other fungi genomes, which CDD-PFAM entries were classified according to *M. perniciosus* ranking. CDD-ID: CDD entry; PFAM Domain: PFAM entry; #Hits Domains: number of gene models containing each CDD-PFAM domain; %Hits Domains: percentage of gene models containing each CDD-PFAM domain in relation to total number of gene models containing a CDD-PFAM domain; %Hits PTN: Percentage of gene models containing each CDD-PFAM domain in relation to total number of *M. perniciosus* gene models; Rank: non-normalized *M. perniciosus* CDD-PFAM ranking; PTNS: proteins in each organism; Norm Factor: factor used to normalize the number of gene models containing each CDD-PFAM domain; # Hits Domains Norm: normalized number of gene models containing a CDD-PFAM domain; %Hits Domains Norm: percentage of gene models containing each CDD-PFAM domain in relation to total number of gene models containing each CDD-PFAM domain; %Hits PTN Norm: Normalized percentage of gene models containing each CDD-PFAM domain in relation to total number of *M. perniciosus* gene models; Rank Norm: normalized *M. perniciosus* CDD-PFAM ranking. Worksheets show the ranking of CDD-PFAM domains using normalization (Rank_Norm) or not using normalization (Rank_Non_Norm).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S4.xls>]

Additional file 5

BioCyc comparison between *S. cerevisiae* and *M. perniciosus* metabolic pathways. Worksheet Compounds: Comparison of number of reactions in each organism containing the compounds described in the table; Worksheet pathways: Comparison of number of pathways in each organism present in each pathway class. The two largest top-level classes, Biosynthesis and Degradation/Utilization/Assimilation, are broken down further to show the distribution of pathways among their next-level subclasses.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S5.xls>]

Additional file 6

Annotation of gene models with no similarity in BLASTX-NR encoding hypothetical small secreted proteins containing at least 2 cysteines. ID: gene model; # residues: number of amino acids of predicted protein encoded by the gene model; # cysteines: number of cysteines in predicted protein; Binomial RT; statistical analysis of cysteines presence in gene models (see methodological details in the file).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S6.xls>]

Additional file 7

Primers used in the amplification of no hits gene models encoding hypothetical small secreted proteins containing at least 2 cysteines. ID: gene model; Set of primers: Group of primers used for the amplification of a gene model. Primer sequence: sequence of primers (SPE – nested in sequence encoding the putative signal peptide; MAT – nested in sequence encoding the putative first amino acid of mature protein; END – nested in sequence containing the putative stop codon). Amplification: positive (Y) or negative (N); EST: presence (Y) or absence (N) of an EST aligned in this gene model; Length: length of amplicon in genomic and cDNA using two combinations of primers (SPE-END; MAT-END).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S7.xls>]

Additional file 8

Examples of amplifications of no hits gene models. PCR amplicons were run on 1% agarose gels. SPE: amplicons resulted from amplification with SPE and END primers; MAT: amplicons resulted from amplification with MAT and END primers; Ctl: water as template (control); Gen: genomic DNA as template; Glu: cDNA from saprotrophic mycelia grown in glucose as template; Cac: cDNA from saprotrophic mycelia grown in cacao extract as template; M: DNA molecular marker.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S8.pdf>]

Additional file 9

Comparison of plant cell wall degrading enzymes in fungi that interact with plants. PFAM entries were correlated with the CAZy nomenclature <http://www.cazy.org> of plant cell wall degrading enzymes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S9.pdf>]

Additional file 10

Genomic survey sequences and gene models nomenclature.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-548-S10.pdf>]

Acknowledgements

We thank CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho de São Paulo) for data processing support, all sequencing technician for excellent technical assistance, especially Robson Dias (UESC – Brazil), Sulamita Freitas Franco (UNICAMP – Brazil) and Vitor Hugo de Almeida e Silva (UNICAMP – Brazil), and Dr. Francisco Javier Medrano for review and commentaries about the article. We are indebted to CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico,

Projects 68.0032/2001-0 and 42.2279/2006-8), SEAGRI (Secretaria de Agricultura, Irrigação e Reforma Agrária do Estado da Bahia), CARGILL/FUN-CAMP/UNICAMP convenio (number C. 3007), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior) for financial support.

References

1. Purdy LH, Schmidt RA: **Status of cacao witches' broom: Biology, epidemiology, and management.** *Annu Rev Phytopathol* 1996, **34**:573-594.
2. Aime MC, Phillips-Mora W: **The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae.** *Mycologia* 2005, **97**(5):1012-1022.
3. Griffith GW, Nicholson J, Nenninger A, Birch RN, Hedger JN: **Witches' brooms and frosty pods: two major pathogens of cacao.** *New Zeal J Bot* 2003, **41**(3):423-435.
4. Evans HC: **Cacao diseases – The trilogy revisited.** *Phytopathology* 2007, **97**(12):1640-1643.
5. Pereira JL, Ram A, Figueiredo JM, de Almeida LC: **La primera aparición de la "escoba de bruja" en la principal área productora de cacao del Brasil.** *Turrialba* 1989, **39**(7):459-461.
6. Pereira JL, deAlmeida LCC, Santos SM: **Witches' broom disease of cocoa in Bahia: Attempts at eradication and containment.** *Crop Protection* 1996, **15**(8):743-752.
7. Evans HC: **Pleomorphism in *Crinipellis perniciosa*, Causal Agent of Witches Broom Disease of Cocoa.** *Trans Br Mycol Soc* 1980, **74**(Jun):515-523.
8. Griffith GW, Hedger JN: **Dual culture of *Crinipellis perniciosa* and potato callus.** *Eur J Plant Pathol* 1994, **100**(6):371-379.
9. Delgado JC, Cook AA: **Nuclear condition of basidia, basidiospores, and mycelium of *Marasmius pernicius*.** *Can J Bot* 1976, **54**(1-2):66-72.
10. Rincones J, Meinhardt LW, Vidal BC, Pereira GA: **Electrophoretic karyotype analysis of *Crinipellis perniciosa*, the causal agent of witches' broom disease of *Theobroma cacao*.** *Mycol Res* 2003, **107**(Pt 4):452-458.
11. Rincones J, Mazotti GD, Griffith GW, Pomela A, Figueira A, Leal GA Jr, Queiroz MV, Pereira JF, Azevedo RA, Pereira GA, Meinhardt LW: **Genetic variability and chromosome-length polymorphisms of the witches' broom pathogen *Crinipellis perniciosa* from various plant hosts in South America.** *Mycol Res* 2006, **110**(Pt 7):821-832.
12. Formighieri EF, Tiburcio RA, Armas ED, Medrano FJ, Shimo H, Carels N, Góes-Neto A, Cotomacci C, Carazzolle MF, Sardinha-Pinto N, 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 JCM, Pereira GAG: **The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stable integrated plasmid.** *Mycol Res* 2008, **112**:1136-1152.
13. Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GA: **Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa*.** *J Exp Bot* 2005, **56**(413):865-877.
14. Gesteira Ada S, Micheli F, Ferreira CF, Cascardo JC: **Isolation and purification of functional total RNA from different organs of cacao tree during its interaction with the pathogen *Crinipellis perniciosa*.** *Biotechniques* 2003, **35**(3):494-496, 498-500.
15. Pirovani CP, Carvalho HA, Machado RC, Gomes DS, Alvim FC, Pomella AW, Gramacho KP, Cascardo JC, Pereira GA, Micheli F: **Protein extraction for proteome analysis from cacao leaves and meristems, organs infected by *Moniliophthora perniciosa*, the causal agent of the witches' broom disease.** *Electrophoresis* 2008, **29**(11):2391-2401.
16. Lima JO, dos Santos JK, Pereira JF, de Resende ML, de Araujo EF, de Queiroz MV: **Development of a transformation system for *Crinipellis perniciosa*, the causal agent of witches' broom in cocoa plants.** *Curr Genet* 2003, **42**(4):236-240.
17. Filho DF, Pungartnik C, Cascardo JC, Brendel M: **Broken hyphae of the basidiomycete *Crinipellis perniciosa* allow quantitative assay of toxicity.** *Curr Microbiol* 2006, **52**(5):407-412.
18. Meinhardt LW, Bellato CM, Rincones J, Azevedo RA, Cascardo JC, Pereira GA: **In vitro production of biotrophic-like cultures of**

- Crinipellis perniciosa*, the causal agent of witches' broom disease of *Theobroma cacao*. *Curr Microbiol* 2006, **52**(3):191-196.
19. Garcia O, Macedo JA, Tiburcio R, Zapparoli G, Rincones J, Bittencourt LM, Ceita GO, Micheli F, Gesteira A, Mariano AC, Schiavinato MA, Medrano FJ, Meinhardt LW, Pereira GA, Cascardo JC: **Characterization of necrosis and ethylene-inducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom in *Theobroma cacao***. *Mycol Res* 2007, **111**(Pt 4):443-455.
 20. Rincones J, Scarpari LM, Carazzolle MF, Mondego JMC, Formighieri EF, Barau JG, Costa GGL, Carraro DM, Brentani HP, Vilas-Boas LA, 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***. *Mol Plant Microbe Interact* 2008, **21**(7):891-908.
 21. Gesteira AS, Micheli F, Carols N, Da Silva AC, Gramacho KP, Schuster I, Macedo JN, Pereira GA, Cascardo JC: **Comparative Analysis of Expressed Genes from Cacao Meristems Infected by *Moniliophthora perniciosa***. *Ann Bot (Lond)* 2007, **100**(1):129-140.
 22. Leal GA, Albuquerque PSB, Figueira A: **Genes differentially expressed in *Theobroma cacao* associated with resistance to witches' broom disease caused by *Crinipellis perniciosa***. *Mol Plant Pathol* 2007, **8**(3):279-292.
 23. Munch S, Lingner U, Floss DS, Ludwig N, Sauer N, Deising HB: **The hemibiotrophic lifestyle of *Colletotrichum* species**. *J Plant Physiol* 2008, **165**(1):41-51.
 24. Meinhardt LW, Rincones J, Bailey BA, Aime MC, Griffith GW, Zhang D, Pereira GA: ***Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe?** *Mol Plant Pathol* 2008, **9**(5):577-588.
 25. James TY, Kauff J, 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 AV, 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, Lücking 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: **Reconstructing the early evolution of Fungi using a six-gene phylogeny**. *Nature* 2006, **443**(7113):818-822.
 26. Bouck J, Miller W, Gorrell JH, Muzny D, Gibbs RA: **Analysis of the quality and utility of random shotgun sequencing at low redundancies**. *Genome Res* 1998, **8**(10):1074-1084.
 27. Branscomb E, Predki P: **On the high value of low standards**. *J Bacteriol* 2002, **184**(23):6406-6409. discussion 6409
 28. Fraser CM, Eisen JA, Nelson KE, Paulsen IT, Salzberg SL: **The value of complete microbial genome sequencing (you get what you pay for)**. *J Bacteriol* 2002, **184**(23):6403-6405. discussion 6405
 29. Gardner SN, Lam MW, Smith JR, Torres CL, Slezak TR: **Draft versus finished sequence data for DNA and protein diagnostic signature development**. *Nucleic Acids Res* 2005, **33**(18):5838-5850.
 30. Taudien S, Ebersberger I, Glockner G, Platzer M: **Should the draft chimpanzee sequence be finished?** *Trends Genet* 2006, **22**(3):122-125.
 31. Kirkness EF, Bafna V, Halpern AL, Levy S, Remington K, Rusch DB, Delcher AL, Pop M, Wang W, Fraser CM, Venter JC: **The dog genome: survey sequencing and comparative analysis**. *Science* 2003, **301**(5641):1898-1903.
 32. Green P: **2x genomes-does depth matter?** *Genome Res* 2007, **17**(11):1547-1549.
 33. Woolfit M, Rozpedowska E, Piskur J, Wolfe KH: **Genome survey sequencing of the wine spoilage yeast *Dekkera* (*Brettanomyces*) *bruxellensis***. *Eukaryot Cell* 2007, **6**(4):721-733.
 34. Lander ES, Waterman MS: **Genomic mapping by fingerprinting random clones: a mathematical analysis**. *Genomics* 1988, **2**(3):231-239.
 35. Wendl MC, Waterston RH: **Generalized gap model for bacterial artificial chromosome clone fingerprint mapping and shotgun sequencing**. *Genome Res* 2002, **12**(12):1943-1949.
 36. Wendl MC, Yang SP: **Gap statistics for whole genome shotgun DNA sequencing projects**. *Bioinformatics* 2004, **20**(10):1527-1534.
 37. Parra G, Bradnam K, Korf I: **CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes**. *Bioinformatics* 2007, **23**(9):1061-1067.
 38. Phillippy AM, Schatz MC, Pop M: **Genome assembly forensics: finding the elusive mis-assembly**. *Genome Biol* 2008, **9**(3):R55.
 39. Zhang M, Gish W: **Improved spliced alignment from an information theoretic approach**. *Bioinformatics* 2006, **22**(1):13-20.
 40. Stanke M, Steinkamp R, Waack S, Morgenstern B: **AUGUSTUS: a web server for gene finding in eukaryotes**. *Nucleic Acids Res* 2004, **32**:W309-312.
 41. Korf I: **Gene finding in novel genomes**. *BMC Bioinformatics* 2004, **5**:59.
 42. Allen JE, Majoros WH, Pertea M, Salzberg SL: **JIGSAW, GeneZilla, and GlimmerHMM: puzzling out the features of human genes in the ENCODE regions**. *Genome Biol* 2006, **7**(Suppl 1):S91-13.
 43. Enright AJ, van Dongen S, Ouzounis CA: **An efficient algorithm for large-scale detection of protein families**. *Nucleic Acids Res* 2002, **30**(7):1575-1584.
 44. Enright AJ, Kunin V, Ouzounis CA: **Protein families and TRIBES in genome sequence space**. *Nucleic Acids Res* 2003, **31**(15):4632-4638.
 45. Tekala F, Latge JP: ***Aspergillus fumigatus*: saprophyte or pathogen?** *Curr Opin Microbiol* 2005, **8**(4):385-392.
 46. Kupfer DM, Reece CA, Clifton SV, Roe BA, Prade RA: **Multicellular ascomycetous fungal genomes contain more than 8000 genes**. *Fungal Genet Biol* 1997, **21**(3):364-372.
 47. Koski LB, Gray MW, Lang BF, Burger G: **AutoFACT: an automatic functional annotation and classification tool**. *BMC Bioinformatics* 2005, **6**:151.
 48. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao LN, He SQ, Hurwitz DI, Jackson JD, Ke ZX, Krylov D, Lanczycki CJ, Liebert CA, Liu CL, Lu F, Lu SN, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang DC, Bryant SH: **CDD: a conserved domain database for interactive domain family analysis**. *Nucleic Acids Res* 2007, **35**:D237-D240.
 49. Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahren D, Tsoka S, Darzentas N, Kunin V, Lopez-Bigas N: **Expansion of the BioCyc collection of pathway/genome databases to 160 genomes**. *Nucleic Acids Res* 2005, **33**(19):6083-6089.
 50. Gonzalez FJ, Nebert DW: **Evolution of the P450 gene superfamily: animal-plant 'warfare', molecular drive and human genetic differences in drug oxidation**. *Trends Genet* 1990, **6**(6):182-186.
 51. Saier MH Jr, Tran CV, Barabote RD: **TCDB: the Transporter Classification Database for membrane transport protein analyses and information**. *Nucleic Acids Res* 2006, **34**:D181-186.
 52. Stergiopoulos I, Zwiers LH, De Waard MA: **The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of subterranean cavities in wheat leaves**. *Mol Plant Microbe Interact* 2003, **16**(8):689-698.
 53. Rogers B, Decottignies A, Kolaczowski M, Carvajal E, Balzi E, Goffeau A: **The pleiotropic drug ABC transporters from *Saccharomyces cerevisiae***. *J Mol Microbiol Biotechnol* 2001, **3**(2):207-214.
 54. Benito B, Garciadeblas B, Rodriguez-Navarro A: **Potassium- or sodium-efflux ATPase, a key enzyme in the evolution of fungi**. *Microbiology* 2002, **148**(Pt 4):933-941.
 55. Apel K, Hirt H: **Reactive oxygen species: metabolism, oxidative stress, and signal transduction**. *Annu Rev Plant Biol* 2004, **55**:373-399.
 56. Able AJ: **Role of reactive oxygen species in the response of barley to necrotrophic pathogens**. *Protoplasma* 2003, **221**(1-2):137-143.
 57. Govrin EM, Levine A: **The hypersensitive response facilitates plant infection by the necrotrophic pathogen *Botrytis cinerea***. *Curr Biol* 2000, **10**(13):751-757.
 58. Rio MCS, Oliveira BV, Thomazella DP, Fracassi da Silva JA, Pereira GAG: **Production of calcium oxalate crystals by the basidiomycete *Moniliophthora perniciosa*, the causal agent of the Witches' Broom Disease of cacao**. *Curr Microbiol* 2008, **56**(4):363-370.

59. Ceita GO, Macêdo JNA, Santos TB, Alemanno L, Gesteira AS, Micheli F, Mariano AC, Gramacho KP, Silva DC, Meinhardt LW, Mazzafera P, Pereira GAG, Cascardo JM: Involvement of calcium oxalate degradation during programmed cell death in *Theobroma cacao* tissues triggered by the hemibiotrophic fungus *Moniliophthora perniciosa*. *Plant Sci* 2007, **173**(2):106-117.
60. Garcia-Brugger A, Lamotte O, Vandelle E, Bourque S, Lecourieux D, Poinssot B, Wendehenne D, Pugin A: Early signaling events induced by elicitors of plant defenses. *Mol Plant Microbe Interact* 2006, **19**(7):711-724.
61. Kim KS, Min JY, Dickman MB: Oxalic Acid Is an Elicitor of Plant Programmed Cell Death during *Sclerotinia sclerotiorum* Disease Development. *Mol Plant Microbe Interact* 2008, **21**(5):605-612.
62. Bolker M: *Ustilago maydis*-a valuable model system for the study of fungal dimorphism and virulence. *Microbiology* 2001, **147**(Pt 6):1395-1401.
63. de Arruda MCCMR, Ferreira MASV, Felipe MSS: Comparison of *Crinipellis perniciosa* isolates from Brazil by ERIC repetitive element sequence-based PCR genomic fingerprinting. *Plant Pathology* 2003, **52**(2):236-244.
64. Andebrhan T, Figueira A, Yamada MM, Cascardo J, Furtak DB: Molecular fingerprinting suggests two primary outbreaks of witches' broom disease (*Crinipellis perniciosa*) of *Theobroma cacao* in Bahia, Brazil. *Eur J Plant Pathol* 1999, **105**(2):167-175.
65. Bartley BGD: Cacao (*Theobroma cacao* L.). Breeding for Durable Resistance in Perennial Crops. *FAO Plant Production and Protection Paper* 1986, **70**:25-42.
66. Larrondo LF, Canessa P, Vicuna R, Stewart P, Wymelenberg A, Vanden, Cullen D: Structure and transcriptional impact of divergent repetitive elements inserted within *Phanerochaete chrysosporium* strain RP-78 genes. *Mol Genet Genomics* 2007, **277**(1):43-55.
67. Poulter RT, Goodwin TJ: DIRS-I and the other tyrosine recombinase retrotransposons. *Cytogenet Genome Res* 2005, **110**(1-4):575-588.
68. DeMarco R, Venancio TM, Verjovski-Almeida S: SmTRCI, a novel *Schistosoma mansoni* DNA transposon, discloses new families of animal and fungi transposons belonging to the CACTA superfamily. *BMC Evol Biol* 2006, **6**:89.
69. Rep M, Does HC van der, Cornelissen BJ: Drifter, a novel, low copy hATlike transposon in *Fusarium oxysporum* is activated during starvation. *Fungal Genet Biol* 2005, **42**(6):546-553.
70. Kempken F, Kuck U: restless, an active Ac-like transposon from the fungus *Tolypocladium inflatum*: structure, expression, and alternative RNA splicing. *Mol Cell Biol* 1996, **16**(11):6563-6572.
71. Monroy F, Sheppard DC: Tafi: a class II transposon of *Aspergillus fumigatus*. *Fungal Genet Biol* 2005, **42**(7):638-645.
72. Zhang X, Feschotte C, Zhang Q, Jiang N, Eggleston WB, Wessler SR: P instability factor: an active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases. *Proc Natl Acad Sci USA* 2001, **98**(22):12572-12577.
73. Robert-Seilaniantz A, Navarro L, Bari R, Jones JD: Pathological hormone imbalances. *Curr Opin Plant Biol* 2007, **10**(4):372-379.
74. Sun S-K, Snyder WC: The bakanae disease of the rice plant. In *Fusarium: diseases, biology and taxonomy* Edited by: Nelson TAT PE, Cook RJ. The Pennsylvania State University Press; 1981:104-113.
75. Rademacher W, Graebe JE: Gibberellin A4 produced by *Sphaceloma manihoticola*, the cause of the superelongation disease of cassava (*Manihot esculenta*). *Biochem Biophys Res Commun* 1979, **91**(1):35-40.
76. Tudzynski B, Kawaide H, Kamiya Y: Gibberellin biosynthesis in *Gibberella fujikuroi*: cloning and characterization of the copalyl diphosphate synthase gene. *Curr Genet* 1998, **34**(3):234-240.
77. Bastos CN, Andebrhan T: Presença de giberelina em basidiosporos de *Crinipellis perniciosa* (Stahel) Singer. *Fitopatologia Brasileira* 1981, **6**:417-423.
78. Kilaru A, Bailey BA, Hasenstein KH: *Moniliophthora perniciosa* produces hormones and alters endogenous auxin and salicylic acid in infected cocoa leaves. *FEMS Microbiol Lett* 2007, **274**(2):238-244.
79. Wolf FT: The Production of Indole Acetic Acid by *Ustilago Zeae*, and Its Possible Significance in Tumor Formation. *Proc Natl Acad Sci USA* 1952, **38**(2):106-111.
80. Maor R, Haskin S, Levi-Kedmi H, Sharon A: In planta production of indole-3-acetic acid by *Colletotrichum gloeosporioides* f. sp. *aeschynomene*. *Appl Environ Microbiol* 2004, **70**(3):1852-1854.
81. Perley JE, Stowe BB: On the ability of *Taphrina deformans* to produce indoleacetic acid from tryptophan by way of tryptamine. *Plant Physiol* 1966, **41**(2):234-237.
82. Prusty R, Grisafi P, Fink GR: The plant hormone indoleacetic acid induces invasive growth in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 2004, **101**(12):4153-4157.
83. Ross JJ, O'Neill DP, Smith JJ, Kerckhoffs LH, Elliott RC: Evidence that auxin promotes gibberellin A1 biosynthesis in pea. *Plant J* 2000, **21**(6):547-552.
84. Pierik R, Tholen D, Poorter H, Visser EJ, Voesenek LA: The Janus face of ethylene: growth inhibition and stimulation. *Trends Plant Sci* 2006, **11**(4):176-183.
85. Swarup R, Parry G, Graham N, Allen T, Bennett M: Auxin cross-talk: integration of signalling pathways to control plant development. *Plant Mol Biol* 2002, **49**(3-4):411-426.
86. Steffens B, Sauter M: Epidermal cell death in rice is regulated by ethylene, gibberellin, and abscisic acid. *Plant Physiol* 2005, **139**(2):713-721.
87. Pierik R, Cuppens ML, Voesenek LA, Visser EJ: Interactions between ethylene and gibberellins in phytochrome-mediated shade avoidance responses in tobacco. *Plant Physiol* 2004, **136**(2):2928-2936.
88. Maor R, Shirasu K: The arms race continues: battle strategies between plants and fungal pathogens. *Curr Opin Microbiol* 2005, **8**(4):399-404.
89. Dodds PN, Lawrence GJ, Catanzariti AM, Ayliffe MA, Ellis JG: The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* 2004, **16**(3):755-768.
90. Kemen E, Kemen AC, Rafiqi M, Hempel U, Mendgen K, Hahn M, Voegelé RT: Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol Plant Microbe Interact* 2005, **18**(11):1130-1139.
91. Rooney HC, Van't Klooster JW, Hoorn RA van der, Joosten MH, Jones JD, de Wit PJ: *Cladosporium Avr2* inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* 2005, **308**(5729):1783-1786.
92. 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 JL, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, Gonzalez-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Grellinger D, Munch K, Rossel N, Scherer M, Vranes M, Ladendorff 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 EV, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BV: Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 2006, **444**(7115):97-101.
93. Rep M: Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiol Lett* 2005, **253**(1):19-27.
94. Zapparoli G, Garcia O, Medrano FJ, Tiburcio R, Costa GL, Pereira GAG: Identification of a second family of genes in *Moniliophthora perniciosa*, the causal agent of Witches' Broom disease in cacao, encoding necrosis inducing proteins similar to cerato platanins. *Mycol Res* 2008.
95. Nonaka T, Dohmae N, Hashimoto Y, Takio K: Amino acid sequences of metalloendopeptidases specific for acyl-lysine bonds from *Grifola frondosa* and *Pleurotus ostreatus* fruiting bodies. *J Biol Chem* 1997, **272**(48):30032-30039.
96. Sepic K, Berne S, Potrich C, Turk T, Macek P, Menestrina G: Interaction of ostreolysin, a cytolytic protein from the edible mushroom *Pleurotus ostreatus*, with lipid membranes and modulation by lysophospholipids. *Eur J Biochem* 2003, **270**(6):1199-1210.

97. Wymelenberg A, Vanden, Minges P, Sabat G, Martinez D, Rajangam AS, Teeri TT, Gaskell J, Kersten PJ, Cullen D: **The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium.** *J Biotechnol* 2005, **118**(1):17-34.
98. Garcia-Conesa MT, Kroon PA, Ralph J, Mellon FA, Colquhoun IJ, Saulnier L, Thibault JF, Williamson G: **A cinnamoyl esterase from *Aspergillus niger* can break plant cell wall cross-links without release of free diferulic acids.** *Eur J Biochem* 1999, **266**(2):644-652.
99. van Loon LC, Rep M, Pieterse CM: **Significance of inducible defense-related proteins in infected plants.** *Annu Rev Phytopathol* 2006, **44**:135-162.
100. Broekaert WF, Terras FRG, BPA C: **Induced and preformed antimicrobial proteins.** In *Mechanisms of Resistance to Plant Diseases* Edited by: Slusarenko AJFR, Van Loon LC. Dordrecht: Kluwer; 2000:371-477.
101. Hu X, Reddy AS: **Nucleotide sequence of a cDNA clone encoding a thaumatin-like protein from *Arabidopsis*.** *Plant Physiol* 1995, **107**(1):305-306.
102. Punja Z: **Transgenic carrots expressing a thaumatin-like protein display enhanced resistance to several fungal pathogens.** *Can J Plant Pathol* 2005, **27**:291-296.
103. Fierens E, Rombouts S, Gebruers K, Goesaert H, Brijs K, Beaupré J, Volckaert G, van Campenhout S, Proost P, Courtin CM, Delcour JA: **TLX1, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family.** *Biochem J* 2007, **403**(3):583-591.
104. Miosga T, Schaaff-Gerstenschlager I, Chwalatzki N, Baur A, Boles E, Fournier C, Schmitt S, Veltin C, Wilhelm N, Zimmermann FK: **Sequence analysis of a 33.1 kb fragment from the left arm of *Saccharomyces cerevisiae* chromosome X, including putative proteins with leucine zippers, a fungal Zn(II)-Cys6 binuclear cluster domain and a putative alpha 2-SCB-alpha 2 binding site.** *Yeast* 1995, **11**(7):681-689.
105. Sakamoto Y, Watanabe H, Nagai M, Nakade K, Takahashi M, Sato T: ***Lentinula edodes* tgl1 encodes a thaumatin-like protein that is involved in lentinan degradation and fruiting body senescence.** *Plant Physiol* 2006, **141**(2):793-801.
106. Chaves FC, Gianfagna TJ: **Necrotrophic phase of *Monilophthora perniciosa* causes salicylic acid accumulation in infected stems of cacao.** *Physiol Mol Plant Pathol* 2006, **69**(1-3):104-108.
107. Aro N, Pakula T, Penttilä M: **Transcriptional regulation of plant cell wall degradation by filamentous fungi.** *FEMS Microbiol Rev* 2005, **29**(4):719-739.
108. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Munsterkott M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC: **The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization.** *Science* 2007, **317**(5843):1400-1402.
109. Millar AH, Day DA: **Nitric oxide inhibits the cytochrome oxidase but not the alternative oxidase of plant mitochondria.** *FEBS Lett* 1996, **398**(2-3):155-158.
110. Juarez O, Guerra G, Martinez F, Pardo JP: **The mitochondrial respiratory chain of *Ustilago maydis*.** *Biochim Biophys Acta* 2004, **1658**(3):244-251.
111. Juarez O, Guerra G, Velazquez I, Flores-Herrera O, Rivera-Perez RE, Pardo JP: **The physiologic role of alternative oxidase in *Ustilago maydis*.** *Febs J* 2006, **273**(20):4603-4615.
112. Joseph-Horne T, Hollomon DW, Wood PM: **Fungal respiration: a fusion of standard and alternative components.** *Biochim Biophys Acta* 2001, **1504**(2-3):179-195.
113. Idnurm A, Howlett BJ: **Isocitrate lyase is essential for pathogenicity of the fungus *Leptosphaeria maculans* to canola (*Brassica napus*).** *Eukaryot Cell* 2002, **1**(5):719-724.
114. Sakai S, Nishide T, Munir E, Baba K, Inui H, Nakano Y, Hattori T, Shimada M: **Subcellular localization of glyoxylate cycle key enzymes involved in oxalate biosynthesis of wood-destroying basidiomycete *Fomitopsis palustris* grown on glucose.** *Microbiology* 2006, **152**(Pt 6):1857-1866.
115. Wang ZY, Thornton CR, Kershaw MJ, Debaio L, Talbot NJ: **The glyoxylate cycle is required for temporal regulation of virulence by the plant pathogenic fungus *Magnaporthe grisea*.** *Mol Microbiol* 2003, **47**(6):1601-1612.
116. Segers G, Bradshaw N, Archer D, Blissett K, Oliver RP: **Alcohol oxidase is a novel pathogenicity factor for *Cladosporium fulvum*, but aldehyde dehydrogenase is dispensable.** *Mol Plant Microbe Interact* 2001, **14**(3):367-377.
117. Daniel G, Volc J, Filonova L, Plihal O, Kubatova E, Halada P: **Characteristics of *Gloeophyllum trabeum* Alcohol Oxidase, an Extracellular Source of H₂O₂ in Brown Rot Decay of Wood.** *Appl Environ Microbiol* 2007, **73**(19):6241-6253.
118. Nakagawa T, Miyaji T, Yurimoto H, Sakai Y, Kato N, Tomizuka N: **A methylotrophic pathway participates in pectin utilization by *Candida boidinii*.** *Appl Environ Microbiol* 2000, **66**(10):4253-4257.
119. Fall R, Benson AA: **Leaf methanol - The simplest natural product from plants.** *Trends Plant Sci* 1996, **1**(9):296-301.
120. Abanda-Nkwatt D, Musch M, Tschiersch J, Boettner M, Schwab W: **Molecular interaction between *Methylobacterium extorquens* and seedlings: growth promotion, methanol consumption, and localization of the methanol emission site.** *J Exp Bot* 2006, **57**(15):4025-4032.
121. Wei Y, Shen W, Dauk M, Wang F, Selvaraj G, Zou J: **Targeted gene disruption of glycerol-3-phosphate dehydrogenase in *Colletotrichum gloeosporioides* reveals evidence that glycerol is a significant transferred nutrient from host plant to fungal pathogen.** *J Biol Chem* 2004, **279**(1):429-435.
122. Santos RX, Melo SC, Cascardo JC, Brendel M, Pungartnik C: **Carbon source-dependent variation of acquired mutagen resistance of *Monilophthora perniciosa*: Similarities in natural and artificial systems.** *Fungal Genet Biol* 2008.
123. Nakayama A, Park S, Zheng-Jun X, Nakajima M, Yamaguchi I: **Immunohistochemistry of active gibberellins and gibberellin-inducible alpha-amylase in developing seeds of morning glory.** *Plant Physiol* 2002, **129**(3):1045-1053.
124. Pellier AL, Lauge R, Veneault-Fourrey C, Langin T: **CLNRI, the AREA/NIT2-like global nitrogen regulator of the plant fungal pathogen *Colletotrichum lindemuthianum* is required for the infection cycle.** *Mol Microbiol* 2003, **48**(3):639-655.
125. Divon HH, Fluhr R: **Nutrition acquisition strategies during fungal infection of plants.** *FEMS Microbiol Lett* 2007, **266**(1):65-74.
126. Marzluft GA: **Genetic regulation of nitrogen metabolism in the fungi.** *Microbiol Mol Biol Rev* 1997, **61**(1):17-32.
127. Lau G, Hamer JE: **Regulatory genes controlling MPG1 expression and pathogenicity in the rice blast fungus *Magnaporthe grisea*.** *Plant Cell* 1996, **8**(5):771-781.
128. Ewing B, Hillier L, Wendt MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment.** *Genome Res* 1998, **8**(3):175-185.
129. Chou HH, Holmes MH: **DNA sequence quality trimming and vector removal.** *Bioinformatics* 2001, **17**(12):1093-1104.
130. Usuka J, Zhu W, Brendel V: **Optimal spliced alignment of homologous cDNA to a genomic DNA template.** *Bioinformatics* 2000, **16**(3):203-211.
131. Gremme G, Brendel V, Sparks ME, Kurtz S: **Engineering a software tool for gene structure prediction in higher organisms.** *Inform Software Technol* 2005, **47**(15):965-978.
132. Lowe TM, Eddy SR: **tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**(5):955-964.
133. Karp PD, Paley S, Romero P: **The Pathway Tools software.** *Bioinformatics* 2002, **18**(Suppl 1):S225-232.
134. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Eng* 1997, **10**(1):1-6.
135. Collart MA, Oliviero S: **Preparation of yeast RNA by extraction with hot acidic phenol.** In *Current Protocols in Molecular Biology* Edited by: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Hoboken KS. New Jersey: John Wiley & Sons, Inc; 2006.

CONCLUSÕES FINAIS

O homobasidiomiceto *M. pernicioso* possui um ciclo de vida característico de fitopatógenos hemibiotróficos. Ele penetra na planta, induz sintomas, torna-se saprofítico e se reproduz sexuadamente para a produção de seus propágulos infectivos, os basidiósporos que re-infectam outras plantas ou outras partes da planta afetada. O conhecimento de todos os detalhes de seu ciclo de vida é importante para o cultivo de espécies susceptíveis e também para o conhecimento de processos que podem ser similares em outras espécies.

Foram apresentadas aqui três partes de um trabalho onde este fungo foi analisado em mais detalhes. No primeiro deles, foi feito um estudo pioneiro sobre a morfogênese de basidiomatas e prováveis genes envolvidos na ativação inicial dessas estruturas e em seu desenvolvimento. O principal objetivo foi aumentar a compreensão do final do ciclo de vida desse patógeno. Por meio de análises feitas com auxílio de estereoscopia, microscopia ótica e de varredura foi possível identificar algumas fases do desenvolvimento dessas estruturas e perceber modificações no padrão geral das hifas no micélio que antecede a frutificação. Foi observada junção de hifas e formação de hifas em anel e ornamentadas como também ocorre em *Pleurotus* sp. Uma camada superficial hidrofóbica foi formada em micélio cultivado e os primeiros sinais da formação dos basidiomatas foram observados internamente a essa camada, na forma de aglomerados com hifas com intensa atividade celular. Foram observados também primórdios imaturos e primórdios diferenciados. A diferenciação foi observada em mais de uma etapa e o padrão apresentado comparado ao de outros basidiomicetos. A seguir, foi feita uma análise de expressão em duas fases contrastantes, uma no início do cultivo e outra quando já havia primórdios presentes de 192 ESTs. Houve

uma distinção clara de dois grupos de genes, ativados ou reprimidos na presença de primórdios. Entre os ativados, destacam-se alguns que codificam para hidrofobinas, citocromo monooxigenases e proteínas da família Ras e entre os reprimidos, os que codificam para calmodulina e PIM1. Doze deles foram re-analisados por qRT-PCR. Essas análises mostraram perfis de expressão que indicam que a ativação do desenvolvimento dos basidiomatas ocorre ainda na fase rosa. Também mostram o envolvimento da rota do AMPc ativado por níveis de glicose e a participação de aegerolisinas e proteínas da família Ras, além do fator MBF.

O segundo trabalho mostra a construção e análise de uma biblioteca de seqüências curtas expressas (ESTs) produzida a partir de uma mistura de RNAs obtidos de amostras coletadas durante o cultivo do fungo em condições de frutificação e que produziu basidiomatas viáveis. Clones dessa biblioteca foram seqüenciados aleatoriamente e as seqüências foram analisadas por programas que possibilitaram a comparação com outras seqüências de bancos de dados disponíveis. O objetivo principal foi identificar genes expressos na fase de frutificação. Para complementar, foi feita novamente uma análise de expressão dos 192 unigenes, semelhantes à já descrita, porém com o acréscimo de três amostras duas obtidas nas fases de micélio rosa e após o estresse e uma de basidiomatas para identificar candidatos mais ou menos expressos nessas fases de desenvolvimento do micélio que antecede os basidiomatas, comparando-os com suas expressões em basidiomatas maduros. A análise dessa biblioteca possibilitou a identificação de 1.533 genes expressos em *M. perniciosus* nessas condições. As comparações ‘*in silico*’ mostraram que 1.209 desses são diferentes dos obtidos em outro banco de ESTs desse fungo, cultivado em outras condições. Também mostram 289 similares a *L. bicolor* e não a *U. maydis*. A análise de abundância de transcritos revelou pelo menos seis perfis de expressão entre os 192 genes analisados. Um

grupo de 25 genes apresentou um pico único na fase de primórdios, destacando-se entre eles os genes que codificam para intersectina, aegerolisina, pleurotolisina B, RNG2 ativada por Ras e um AMPK.

No último trabalho foi mostrado um esforço nacional de um grupo de pesquisadores para seqüenciar o genoma desse patógeno. O principal objetivo foi obter informações que auxiliassem na diminuição da doença Vassoura-de-bruxa que ele causa em *Theobroma cacao*. Esse trabalho resultou, de fato, em uma série de conclusões possíveis sobre a doença e a atuação desse fungo. Foram preditos 14.077 genes modelos a partir dos quais foram identificadas várias rotas metabólicas prováveis. Foi observado que esse fungo codifica para diversas citocromo p450 mono-oxygenases, diversas carboxilesterase, várias deuterolisinas, algumas taumatinas e aegerolisinas. Também foi discutida a presença de genes relacionados à biossíntese de fito-hormônios, de manutenção da variabilidade genética (genes de cruzamento sexual e de elementos de transposição), de mecanismos gerais de detoxificação e resistência, de patogenicidade e de genes relacionados à colonização e frutificação. A identificação desses genes permitiu a elaboração de um modelo para explicar a doença que decorre de sua presença em cacaueiros. O impacto deste trabalho foi bastante alto, principalmente porque foi o primeiro patógeno homobasidiomiceto hemibiotrófico a ser seqüenciado. Além disso, as seqüências geradas e os genes identificados foram disponibilizados no banco de dados internacional, o NCBI, o que contribui para a produção de novos trabalhos científicos sobre o patógeno e sua interação com a planta.

Esses trabalhos foram correlacionados. Os ESTs produzidos foram utilizados para a identificação de genes do genoma de *M. perniciosa* e para correlacionar as alterações morfológicas às mudanças de expressão de alguns candidatos selecionados por indicações de sua participação na frutificação em outros basidiomicetos. Também são complementares. As

informações sobre a morfogênese dos basidiomatas complementam o conhecimento existente sobre o ciclo de vida do fungo e servem para explicar a existência de alguns genes previstos em seu genoma. As informações sobre genes expressos na fase de frutificação do fungo e em momentos específicos dessa fase auxiliam na compreensão do funcionamento do fungo e das vias metabólicas que foram identificadas como resultado do projeto genoma.

Este é um trabalho que tem caráter pioneiro e serve para abrir caminho para novas pesquisas relacionadas a esse patógeno, a outros fungos e a outros sistemas biológicos. Os genes identificados aqui serão de grande utilidade para a comunidade científica.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aime MC, Phillips-Mora W: **The causal agent of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of *Marasmiaceae*.** *Mycologia* 2005, **97**: 1012-1022.
- Albuquerque P, Kyaw CM, Saldanha RR, Brigido MM, Felipe MSS, Silva-Pereira I: **Pbhyd1 and Pbhyd2: two mycelium-specific hydrophobin genes from the dimorphic fungus *Paracoccidioides brasiliensis*.** *Fungal Gen Biol* 2004, **(4)**1: 510–520.
- Alexopoulos CJ, Mims CW, Blackwell M: *Introductory Mycology* (4th Ed.). *John Wiley and Sons*, New York, USA 1996.
- Bahn YS, Xue C, Idnurm A, Rutherford JC, Heitman J, Cardenas ME: **Sensing the environment: lessons from fungi.** *Nat Rev Microbiol* 2007, **5**: 57-69.
- Baker RED, Crowdy SH. **Studies in the Witches' broom disease of cocoa caused by *Marasmius perniciosus* Sthael: Introduction, symptoms and etiology.** *Port-of-Spain: ICTA*, 1943: 28p (Memoir 7).

- Berne S, Križaj I, Pohleven F, Turk T, Maček P, Sepčić K: ***Pleurotus* and *Agrocybe* hemolysins, new proteins hypothetically involved in fungal fruiting.** *Biochim Biophys Acta* 2002, **1570** (3): 153-159.
- Berne S, Lah L, Sepčić K: **Aegerolysins: Structure, function, and putative biological role.** *Protein Sci* 2009, **18**(4): 694-706.
- Berne S, Pohleven J, Vidic I, Rebolj K, Pohleven F, Turk T, Maček P, Sonnenberg A, Sepčić K: **Ostreolysin enhances fruiting initiation in the oyster fungus (*Pleurotus ostreatus*).** *Mycol Res* 2007, **111** (12): 1431-1436.
- Boulianne RP, Liu Y, Aebi M, Lu BC, Kues U: **Fruiting body development in *Coprinus cinereus*: regulated expression of two galectins secreted by a non-classical pathway.** *Microbiology* 2000, **146**: 1841–1853
- Castro, VLSS, Jonsson CM, Melo IS de, Nunes, FV: **Avaliação de risco ecotoxicológico de *Trichoderma stromaticum* usado como biopesticida.** *Ecotoxicology and Environmental Restoration* 2001, **4**(1): 18-24.
- Chowdhury HH, Rebolj K, Kreft M, Zorec R, Maček P, Sepčić K: **Lysophospholipids prevent binding of a cytolytic protein ostreolysin to cholesterol-enriched membrane domains.** *Toxicon* 2008, **51** (8): 1345-1356.
- De Souza JT, Bailey BA, Pomella AWV, Erbe EF, Murphy CA, Bae H, Hebbar PK: **Colonization of cacao seedlings by *Trichoderma stromaticum*, a mycoparasite of the Witches' Broom pathogen, and its influence on plant growth and resistance.** *Biological Control* 2008, **46**: 36-45
- Delgado JC, Cook AA: **Nuclear condition of basidia, basidiospores, and mycelium of *Marasmius perniciosus*.** *Canad J Botany* 1976, **54**: 66-72.
- Evans HC: **Witches' broom disease - A case study.** *Cocoa Growers Bulletin* 1981, **32**: 5-19.
- Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF et al.: **A higher-level phylogenetic classification of the *Fungi*.** *Micol Res* 2007, **111** (5): 509-547.
- Kondoh O, Muto A, Kajiwara S, Takagi J, Saito Y, Shishido K: **A fruiting body-specific cDNA, *mfbAc*, from the mushroom *Lentinus edodes* encodes a high-molecular-weight cell-adhesion protein containing an Arg-Gly-Asp motif.** *Gene* 1995, **154**: 31-37.

- Liu Y, Srivilai P, Loos S, Aebi M, K  s U: **An essential gene for fruiting body initiation in the basidiomycete *Coprinopsis cinerea* is homologous to bacterial cyclopropane fatty acid synthase genes.** *Genetics* 2006, **172**: 873-884.
- Ma A, Shan L, Wang N, Zheng L, Chen L, Xie B: **Characterization of a *Pleurotus ostreatus* fruiting body-specific hydrophobin gene, *Po.hyd*.** *J Bas Microbiol* 2007, **47** (4): 317-324.
- Meihardt LW, Rincones J, Bailey B, Aime MC, Griffith GW, Zhang D, Pereira G: ***Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe?.** *Mol Plant Pathol* 2008, **9** (5), 577-588.
- Miyazaki Y, Kaneko S, Sunagawa M, Shishido K, Yamazaki T, Nakamura M, Babasaki K: **The fruiting-specific *Le.flp1* gene, encoding a novel fungal fasciclin-like protein, of the basidiomycetous mushroom *Lentinula edodes*.** *Cur Genetics* 2007, **51** (6): 367-375.
- Nakazawa T, Miyazaki Y, Kaneko S, Shishido K: **Developmental regulator *Le.CDC5* of the mushroom *Lentinula edodes*: analyses of its amount in each of the stages of fruiting-body formation and its distribution in parts of the fruiting bodies.** *FEMS Microbiol Let* 2006, **261** (1): 60-63.
- Purdy LH, Schmidt RA: **Status of cacao witches' broom: biology, epidemiology, and management.** *Annu Rev Phytopath* 1996, **34**: 573-594.
- Rocha HM, Wheeler BEJ: **Factors influencing the production of basidiocarps and the deposition and germination of basidiospores of *Crinipellis perniciosa*, the causal agent of witches' broom disease on cocoa (*Theobroma cacao*).** *Plant Pathol* 1985, **34**: 319-328.
- Rocha HM, Wheeler BEJ: **The water balance as an important factor in basidiocarp production by *Crinipellis perniciosa*, the causal fungus of cocoa witches' broom.** *Proc 8th Internat. Cocoa Res. Conf. 1981. Cartagena, Columbia: Cocoa Producers Alliance*, 1982, 381-386.
- Rocha HM: **The ecology of *Crinipellis perniciosa* (Stahel) Singer in Witches' broom on cocoa (*Theobroma cacao* L.).** PhD Thesis, University of London, London, England, 1983.

- Samuels GJ, Pardo-Schultheiss R, Hebbar KP, Lumsden RD, Bastos CN, Costa JC, Bezerra JL: **Trichoderma stromaticum sp. nov., a parasite of the cacao witches broom pathogen.** *Micol Res* 2000, 194 (6): 760-764
- Sano H, Kaneko S, Sakamoto Y, Sato T, Shishido K: **The basidiomycetous mushroom *Lentinula edodes* white collar-2 homolog PHRB, a partner of putative blue-light photoreceptor PHRA, binds to a specific site in the promoter region of the *L. edodes* tyrosinase gene.** *Fungal Gen Biol* 2009, 46 (4): 333-341.
- Santos SC: **Caracterização de hidrofobinas do fungo *Crinipellis perniciosa* (Stahel) Singer, causador da doença Vassoura-de-Bruxa no cacauero.** M. S. Thesis, Universidade Estadual de Santa Cruz, Ilhéus – Bahia, Brazil, 2005.
- Silva SDVM, Matsuoka K: **Histologia da interação *Crinipellis perniciosa* em cacaueros suscetível e resistente à vassoura-de-bruxa.** *Fitop Brasileira* 1999, **24**: 54-59.
- Singer R: **A monographic study of the genera “*Crinipellis*” and “*Chaetocalathus*.”** *Lilloa* 1942, **8**:441–534.
- Stahel G: **Contribution to the knowledge of Witchbroom disease.** Surinam Department of Agriculture. Bull. 39. *In: Tropical Agriculture* 1919, **IX** : 167-176.
- Stahel G: ***Marasmius perniciosus* nov. spec.** Dept Landbouw in Suriname Bull 1915, **33**:1–27.
- Takagi Y, Katayose Y, Shishido K: **Intracellular levels of cyclic AMP and adenylate cyclase activity during mycelial development in fruiting body formation in *Lentinus edodes*.** *FEMS Microbiol Let* 1988, **55** (3): 275-278.
- Walther V, Rexer K-H, Kost G: **Ontogeny of *Mycena stylobates* fruit bodies.** *Mycol Res* 2001, **105** (6) : 723±733
- Wösten HAB, De Vocht ML: **Hydrophobins, the fungal coat unraveled.** *Biochim Biophys Acta* 2000, **1469**: 79-86.
- Žužek MC, Maček P, Sepčić K, Cestnik V, Frangež R: **Toxic and lethal effects of ostreolysin, a cytolytic protein from edible oyster mushroom (*Pleurotus ostreatus*), in rodents.** *Toxicon* 2006, **48**: 264-271.

ANEXO I

Tabela 1 – Resultado da anotação da biblioteca e classificação por funções.

EST	Provável Função	Classes de Funções
CP03-EB-001-004-A02-UE.F	12 kDa choque térmico	estímulo abiótico e desenvolvimento
Contig607	18,2 kDa proteína de choque térmico classe I	estímulo abiótico e desenvolvimento
CP03-EB-001-016-G01-UE.F	28 kDa fosfoproteína estável em ácido e calor	estímulo abiótico e desenvolvimento
CP02-EB-001-003-A01-FS.F	Transportador ABC/proteína associada a resistência a muitidrogas	estímulo abiótico e desenvolvimento
Contig610	Aegerolisina	estímulo abiótico e desenvolvimento
Contig458	Aegerolisina Aa-Pri1	estímulo abiótico e desenvolvimento
Contig459	Aegerolisina Aa-Pri1	estímulo abiótico e desenvolvimento
CP03-EB-001-020-F05-UE.F	Quitooligossacarídeo desacetilase	estímulo abiótico e desenvolvimento
CP02-EB-001-002-B09-FS.F	Citocromo P450	estímulo abiótico e desenvolvimento
CP03-EB-001-005-A03-UE.F	Citocromo P450	estímulo abiótico e desenvolvimento
CP03-EB-001-015-G01-UE.F	Citocromo P450	estímulo abiótico e desenvolvimento
CP03-EB-001-025-D09-UE.F	Citocromo P450 hidroxilase/ B-caroteno oxigenase	estímulo abiótico e desenvolvimento
CP03-EB-001-001-D07-UE.F	citocromo P450 monooxigenase pc-3	estímulo abiótico e desenvolvimento
CP03-EB-001-005-C06-UE.F	Epóxido hidrolase	estímulo abiótico e desenvolvimento
CP03-EB-001-025-F05-UE.F	Proteína de superfície ancorada a glicolipídios	estímulo abiótico e desenvolvimento
CP03-EB-001-004-H09-UE.F	Proteína de choque térmico	estímulo abiótico e desenvolvimento
CP03-EB-001-007-F09-UE.F	Proteína de choque térmico	estímulo abiótico e desenvolvimento
Contig647	Heat shock factor-binding protein	estímulo abiótico e desenvolvimento
Contig621	Proteína de choque térmico 16	estímulo abiótico e desenvolvimento
CP03-EB-001-021-C09-UE.G	Proteína de choque térmico 90	estímulo abiótico e desenvolvimento
CP03-EB-001-022-F12-UE.F	Proteína de choque térmico 90	estímulo abiótico e desenvolvimento
CP03-EB-001-010-E02-UE.F	Hsp90 co-chaperone Cdc37	estímulo abiótico e desenvolvimento
CP02-EB-001-002-F02-FS.F	Glicoproteína rica em hidroxiprolina	estímulo abiótico e desenvolvimento
Contig405	MBF 1 (Multiprotein-bridging factor 1)	estímulo abiótico e desenvolvimento
Contig643	MBF 1 (Multiprotein-bridging factor 1)	estímulo abiótico e desenvolvimento
CP03-EB-001-001-B10-UE.F	NADPH–citocromo P450 redutase	estímulo abiótico e desenvolvimento
CP03-EB-001-026-G03-UE.F	NADPH–citocromo P450 redutase	estímulo abiótico e desenvolvimento
CP02-EB-001-004-H04-FS.G	Proteína relatada à patogenicidade	estímulo abiótico e desenvolvimento
Contig460	Pleurotolisin B	estímulo abiótico e desenvolvimento
Contig636	Chaperone HSP31	estímulo abiótico e desenvolvimento
Contig591	Proteína RIC1	estímulo abiótico e desenvolvimento
Contig595	Citocromo P450	estímulo abiótico e desenvolvimento
CP03-EB-001-026-H02-UE.F	Px19-like	estímulo abiótico e desenvolvimento
CP02-EB-001-006-B03-UE.F	Septina	estímulo abiótico e desenvolvimento
Contig618	Pequena proteína de choque térmico	estímulo abiótico e desenvolvimento
Contig402	Pequena proteína de choque térmico	estímulo abiótico e desenvolvimento
Contig423	Pequena proteína de choque térmico	estímulo abiótico e desenvolvimento
CP03-EB-001-019-D06-UE.F	Antígeno 1 associada ao esperma	estímulo abiótico e desenvolvimento
Contig442	Proteína de controle do ciclo celular cwf16	divisão celular
CP03-EB-001-012-C12-UE.F	Proteína do ciclo celular homóloga a 123	divisão celular
CP03-EB-001-024-D11-UE.F	Centrina-3	divisão celular
CP03-EB-001-010-A06-UE.F	Intersectina	divisão celular
Contig320	Fissão mitocondrial 1	divisão celular
CP03-EB-001-014-G05-UE.F	Proteína de replicação do DNA mitocondrial YHM2	divisão celular
CP03-EB-001-004-A03-UE.F	Componente de checagem do fuso mitótico MAD2	divisão celular
CP03-EB-001-024-H06-UE.F	Componente de checagem do fuso mitótico MAD3	divisão celular
CP03-EB-001-001-B02-UE.F	Pim1	divisão celular
CP03-EB-001-014-C08-UE.F	Pim1	divisão celular
CP03-EB-001-024-C10-UE.F	Proteína dif-1	divisão celular
CP03-EB-001-012-E07-UE.F	Proteína 1 supressora de cinetocoro	divisão celular
CP02-EB-001-003-A08-FS.F	Extensina precursor	metabolismo e estrutura da parede celular
CP03-EB-001-023-B02-UE.F	Extensina precursor	metabolismo e estrutura da parede celular
CP03-EB-001-010-C03-UE.F	Extensina precursor (cell wall hydroxyproline-rich glycoprotein)	metabolismo e estrutura da parede celular
Contig415	Hidrofobina SC3	metabolismo e estrutura da parede celular
CP03-EB-001-015-B12-UE.F	Hidrofobina SC3	metabolismo e estrutura da parede celular
CP03-EB-001-025-B04-UE.F	Hidrofobina SC3	metabolismo e estrutura da parede celular
CP03-EB-001-016-D05-UE.F	Hidrofobina SC3	metabolismo e estrutura da parede celular
Contig204	Hidrofobina	metabolismo e estrutura da parede celular
Contig218	Hidrofobina	metabolismo e estrutura da parede celular
Contig422	Hidrofobina	metabolismo e estrutura da parede celular
Contig434	Hidrofobina	metabolismo e estrutura da parede celular
Contig439	Hidrofobina	metabolismo e estrutura da parede celular
Contig581	Hidrofobina	metabolismo e estrutura da parede celular
Contig648	Hidrofobina	metabolismo e estrutura da parede celular
CP03-EB-001-007-C05-UE.F	Hidrofobina	metabolismo e estrutura da parede celular
CP03-EB-001-013-F04-UE.F	Hidrofobina	metabolismo e estrutura da parede celular
Contig630	Hidrofobina	metabolismo e estrutura da parede celular
Contig637	Hidrofobina	metabolismo e estrutura da parede celular
Contig646	Hidrofobina SC3	metabolismo e estrutura da parede celular
CP03-EB-001-016-F01-UE.F	Hidrofobina-2 [Precursor]	metabolismo e estrutura da parede celular
Contig155	Hidrofobina-3	metabolismo e estrutura da parede celular
Contig180	hidrofobina	metabolismo e estrutura da parede celular

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP03-EB-001-013-H09-UE.F	Quitina desacetilase	metabolismo e estrutura da parede celular
CP03-EB-001-022-E03-UE.F	Extensina [precursor]	metabolismo e estrutura da parede celular
Contig156	DNA helicase	cromatina e metabolismo do DNA
CP03-EB-001-025-D08-UE.F	DNA-(sítio apurínico or apirimidínico) liase	cromatina e metabolismo do DNA
CP03-EB-001-015-H12-UE.F	Histona acetiltransferase tipo B subunidade catalítica	cromatina e metabolismo do DNA
Contig158	Histona H4	cromatina e metabolismo do DNA
CP02-EB-001-003-A05-FS.F	Histona H4	cromatina e metabolismo do DNA
CP03-EB-001-016-F10-UE.F	Hipotética proteína C25H2.10c no cromossomo II	cromatina e metabolismo do DNA
CP03-EB-001-005-G11-UE.F	Helicase MusN da família RecQ	cromatina e metabolismo do DNA
CP03-EB-001-010-G02-UE.F	Proteína relatada a actina complexo 2/3 subunidade 5	citoesqueleto
CP03-EB-001-020-G03-UE.F	Tubulina alfa1	citoesqueleto
CP03-EB-001-024-F06-UE.F	Coactosina	citoesqueleto
Contig143	Cofilina	citoesqueleto
CP03-EB-001-016-A03-UE.F	Dineína 8 kDa cadeia leve	citoesqueleto
CP03-EB-001-002-G09-UE.F	Proteína fimbrial	citoesqueleto
Contig473	Profilina	citoesqueleto
CP02-EB-001-002-H12-FS.F	TRIO e proteína que se liga a F-actina	citoesqueleto
Contig624	Proteína 48 responsiva a danos no DNA	defensa e recuperação celular
Contig136	DnaJ	defensa e recuperação celular
CP03-EB-001-010-E03-UE.F	Proteína de reconhecimento mau pareamentos e DNA cruciforme	defensa e recuperação celular
CP03-EB-001-017-C09-UE.F	NUDX fosfolipidase	defensa e recuperação celular
CP03-EB-001-001-G08-UE.F	Peptidil prolil cis-trans isomerase/ Cyclophilin	defensa e recuperação celular
Contig332	Peptidil prolil cis-trans isomerase/ Cyclophilin	defensa e recuperação celular
Contig476	Peptidil prolil cis-trans isomerase/ Cyclophilin	defensa e recuperação celular
Contig195	Peroxiredoxina	defensa e recuperação celular
CP03-EB-001-020-B03-UE.F	Peroxiredoxina	defensa e recuperação celular
CP03-EB-001-007-E08-UE.F	RuvB-like helicase 1	defensa e recuperação celular
Contig469	Snodprot	defensa e recuperação celular
Contig471	Superóxido dismutase	defensa e recuperação celular
CP03-EB-001-024-B01-UE.F	Endonuclease de danos por UV	defensa e recuperação celular
CP03-EB-001-001-H04-UE.F	C6 fator de transcrição	expressão genica e metabolismo de RNA
CP03-EB-001-007-H04-UE.F	CCR4-NOT complexo de transcrição subunidade 6	expressão genica e metabolismo de RNA
CP03-EB-001-003-F09-UE.F	DNA-directed RNA polymerase I polipeptídeo 2	expressão genica e metabolismo de RNA
CP03-EB-001-020-F03-UE.F	RNA polymerase II subunidade 9	expressão genica e metabolismo de RNA
CP03-EB-001-010-H06-UE.F	RNA polymerase III polipeptídeo de 34 kDa	expressão genica e metabolismo de RNA
CP03-EB-001-014-D01-UE.F	F-box/ proteína 7 de repetições WD	expressão genica e metabolismo de RNA
Contig368	Fator de ligação ao GATA	expressão genica e metabolismo de RNA
Contig297	Proteína do grupo de alta mobilidade	expressão genica e metabolismo de RNA
CP03-EB-001-019-F04-UE.F	Proteína Homeodominio	expressão genica e metabolismo de RNA
CP03-EB-001-005-G09-UE.F	Proteína KTI12	expressão genica e metabolismo de RNA
CP03-EB-001-020-A06-UE.F	Sequência 1 amplificada de células T malignas	expressão genica e metabolismo de RNA
Contig374	Mitoferrina-1	expressão genica e metabolismo de RNA
CP03-EB-001-015-E04-UE.F	Mitoferrina-1	expressão genica e metabolismo de RNA
CP03-EB-001-005-H04-UE.F	Proteína de ligação ao mRNA binding	expressão genica e metabolismo de RNA
CP03-EB-001-019-H04-UE.F	Histona H4	expressão genica e metabolismo de RNA
CP02-EB-001-006-C10-UE.F	Proteína 1 contendo 'dedos de zinco' tipo NFX1	expressão genica e metabolismo de RNA
Contig437	Proteína de ligação à sequência de localização nuclear	expressão genica e metabolismo de RNA
CP03-EB-001-025-D04-UE.F	Nucleobindin-2 precursor	expressão genica e metabolismo de RNA
Contig426	Oligoribonuclease	expressão genica e metabolismo de RNA
CP03-EB-001-002-A03-UE.F	Proteína La	expressão genica e metabolismo de RNA
CP03-EB-001-015-H07-UE.F	Proteína de helice anfipática pareada Sin3a	expressão genica e metabolismo de RNA
CP03-EB-001-001-F05-UE.F	Proteína 1 de ligação a poliadenilato	expressão genica e metabolismo de RNA
CP02-EB-001-004-D02-FS.G	Fator 18 de remoção de introns	expressão genica e metabolismo de RNA
CP03-EB-001-024-F08-UE.F	RNA helicase dependente de ATP - fator de remoção de introns	expressão genica e metabolismo de RNA
CP03-EB-001-001-B05-UE.F	Proteína SOF1	expressão genica e metabolismo de RNA
Contig348	Proteína TAR	expressão genica e metabolismo de RNA
CP03-EB-001-014-C10-UE.F	Ribonuclease III	expressão genica e metabolismo de RNA
CP03-EB-001-007-D12-UE.F	Ribonuclease Le2	expressão genica e metabolismo de RNA
CP03-EB-001-001-H03-UE.F	Fator 1 de remoção de introns	expressão genica e metabolismo de RNA
Contig510	Fator de transcrição BTF3	expressão genica e metabolismo de RNA
CP03-EB-001-007-D11-UE.F	Fator II de iniciação da transcrição - cadeia gama	expressão genica e metabolismo de RNA
CP03-EB-001-001-H11-UE.F	Fator de iniciação da transcrição TFIID subunidade 10	expressão genica e metabolismo de RNA
CP03-EB-001-019-B08-UE.F	Regulador transcricional PRZ1	expressão genica e metabolismo de RNA
CP03-EB-001-010-C11-UE.F	Pequena ribonucleoproteína U3 IMP3	expressão genica e metabolismo de RNA
Contig589	Proteína tipo Sm associada ao snRNA U6 - LSm7	expressão genica e metabolismo de RNA
CP03-EB-001-021-G11-UE.F	Proteína tipo Sm associada ao snRNA U6 - LSm7	expressão genica e metabolismo de RNA
Contig629	Proteína tipo Sm associada ao snRNA U6 - LSm8	expressão genica e metabolismo de RNA
CP03-EB-001-002-D04-UE.F	Proteína de transformação viral Jun	expressão genica e metabolismo de RNA
CP03-EB-001-015-B06-UE.F	Proteína POP3 com repetições WD	expressão genica e metabolismo de RNA
CP03-EB-001-017-A06-UE.F	Proteína 38 contendo domínios BTB e 'dedos de zinco'	expressão genica e metabolismo de RNA
CP03-EB-001-002-H09-UE.F	Proteína 13 contendo domínio 'dedos de zinco' CCCH	expressão genica e metabolismo de RNA
CP02-EB-001-006-F05-UE.F	Proteína 'dedos de zinco'	expressão genica e metabolismo de RNA

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP03-EB-001-010-F07-UE.F	Acil-CoA desaturase	Transporte de membrana
CP03-EB-001-017-G01-UE.F	ATP sintase cadeia D, mitocondrial	Transporte de membrana
CP03-EB-001-015-E03-UE.F	ATP sintase subunit beta	Transporte de membrana
CP03-EB-001-005-H06-UE.F	ATPase transportadora de cálcio	Transporte de membrana
Contig444	Cloreto peroxidase	Transporte de membrana
Contig187	Citocromo b-245 cadeia pesada	Transporte de membrana
Contig202	Citocromo b5	Transporte de membrana
Contig324	Citocromo c oxidase polypeptide IV	Transporte de membrana
Contig522	Citocromo c oxidase polypeptide Vib	Transporte de membrana
CP03-EB-001-015-F08-UE.F	Citocromo c oxidase polypeptide VII A	Transporte de membrana
CP03-EB-001-010-B11-UE.F	Citocromo c peroxidase	Transporte de membrana
CP03-EB-001-001-H06-UE.F	Diidrolipoil desidrogenase, precursor mitocondrial	Transporte de membrana
CP03-EB-001-017-E02-UE.F	Monooxygenase contendo Flavina	Transporte de membrana
CP03-EB-001-003-D02-UE.F	Indutor do transporte de gluconato	Transporte de membrana
CP03-EB-001-015-G10-UE.F	Transportador de glicose	Transporte de membrana
CP03-EB-001-019-D04-UE.F	ATPase duplo setor de transporte de H(+)	Transporte de membrana
Contig600	ATPase duplo setor de transporte de H(+), subunidade 5	Transporte de membrana
CP03-EB-001-016-C12-UE.F	Sintase de ATP transportador de hidrogênio	Transporte de membrana
CP03-EB-001-004-A11-UE.F	Transportador de fosfato inorgânico	Transporte de membrana
CP03-EB-001-025-A06-UE.F	Transporte de íons/transporte de fosfato/co-transporte de Sodium-decarboxilato	Transporte de membrana
CP03-EB-001-025-D10-UE.F	Isovaleryl-CoA desidrogenase, precursor mitocondrial	Transporte de membrana
CP03-EB-001-025-H04-UE.F	l-fucose permease	Transporte de membrana
CP03-EB-001-001-E08-UE.F	Translocase mitocondrial membrana interna - subunit tim23	Transporte de membrana
CP03-EB-001-024-C02-UE.F	Translocase mitocondrial membrana interna - subunit tim17	Transporte de membrana
Contig461	Translocase mitocondrial membrana interna - subunit tim9	Transporte de membrana
Contig584	Translocase mitocondrial membrana interna - subunit tim9	Transporte de membrana
CP03-EB-001-003-A06-UE.F	O-metilsterigmatocistina oxidoredutase	Transporte de membrana
CP03-EB-001-020-B02-UE.F	O-metilsterigmatocistina oxidoredutase (Citocromo P450 64)	Transporte de membrana
Contig578	Citocromo C oxidase	Transporte de membrana
Contig611	Peptídeo transportador PTR2A	Transporte de membrana
CP03-EB-001-020-B11-UE.F	Carreador de fosfato	Transporte de membrana
Contig451	Proteolípido 3 de membrana plasmática	Transporte de membrana
CP03-EB-001-016-A08-UE.F	Transportador YER036C dependente de ATP	Transporte de membrana
CP03-EB-001-014-F11-UE.F	monooxygenase inespecífica	Transporte de membrana
CP03-EB-001-012-E02-UE.F	UPF0057 proteína de membrana ZK632.10	Transporte de membrana
CP02-EB-001-003-H03-FS.F	Uracil permease	Transporte de membrana
CP03-EB-001-024-A08-UE.F	Uracil permease	Transporte de membrana
CP03-EB-001-022-D09-UE.F	Anil-alcool desidrogenase [NADP+]	metabolismo primário
Contig151	2-metilcitrate desidratase	metabolismo primário
CP03-EB-001-009-E09-UE.F	2-nitropropane dioxigenase	metabolismo primário
CP03-EB-001-009-A05-UE.F	3-cetoacil-CoA tiolase, peroxissomal [Precursor]	metabolismo primário
Contig644	3-cetoacil-CoA tiolase, peroxissomal [Precursor]	metabolismo primário
CP03-EB-001-007-C09-UE.F	3-cetoacil-CoA tiolase, peroxissomal [Precursor]	metabolismo primário
Contig428	3-ketoreductase	metabolismo primário
CP03-EB-001-001-D02-UE.F	3-oxoacyl-[acyl-carrier-protein] redutase	metabolismo primário
Contig620	3-oxoacyl-[acyl-carrier-protein] redutase 2	metabolismo primário
CP03-EB-001-026-F02-UE.F	Acetil-CoA C-aciltransferase	metabolismo primário
Contig420	Acil-CoA-binding protein	metabolismo primário
CP03-EB-001-007-G08-UE.F	Acil-protein tioesterase 1	metabolismo primário
CP03-EB-001-008-E02-UE.F	Adenine fosforibosiltransferase 1	metabolismo primário
Contig258	Alcool desidrogenase	metabolismo primário
Contig599	Alcool desidrogenase	metabolismo primário
CP03-EB-001-005-D09-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-018-E07-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-018-G06-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-025-D11-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-025-G06-UE.F	Alcool desidrogenase	metabolismo primário
CP02-EB-001-006-G09-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-012-D03-UE.F	Alcool desidrogenase	metabolismo primário
CP03-EB-001-015-D10-UE.F	Aldo/ceto redutase	metabolismo primário
CP03-EB-001-007-B05-UE.F	Aldo-ceto redutase	metabolismo primário
CP03-EB-001-024-E03-UE.F	Alpha/beta hidrolase superfamily protein	metabolismo primário
CP03-EB-001-026-D07-UE.F	Alpha-1,2-mannosiltransferase	metabolismo primário
CP03-EB-001-001-G10-UE.F	Amino-acid N-acetiltransferase	metabolismo primário
CP03-EB-001-001-A06-UE.F	Aminopeptidase	metabolismo primário
CP02-EB-001-004-D03-FS.G	Antranilato sintase componente 2	metabolismo primário
CP03-EB-001-001-B01-UE.F	Anil-alcool desidrogenase	metabolismo primário
CP03-EB-001-004-H04-UE.F	Anil-alcool oxidase	metabolismo primário
Contig455	Aspartato aminotransferase	metabolismo primário
CP03-EB-001-014-B11-UE.F	Aspartato quinase	metabolismo primário
CP03-EB-001-020-E03-UE.F	Proteína de ligação Beta-1,3-glucano	metabolismo primário
CP03-EB-001-016-H11-UE.F	Beta-glucosidase	metabolismo primário

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
Contig384	Beta-manosidase	metabolismo primário
CP03-EB-001-013-F05-UE.F	Carbamoil-fosphate sintase (hidrolizadora de glutamina)	metabolismo primário
CP03-EB-001-019-C04-UE.F	Carbamoil-fosphate sintase arginina específica cadeia maior	metabolismo primário
Contig163	Carboxil esterase	metabolismo primário
Contig67	Coline desidrogenase	metabolismo primário
CP03-EB-001-017-B04-UE.F	Ciclopentanona 1,2-mono-oxigenase	metabolismo primário
Contig628	Cistationina beta-liase	metabolismo primário
CP03-EB-001-003-C06-UE.F	Proteína 1 rica em sisteína	metabolismo primário
CP03-EB-001-016-F05-UE.F	Citidina desaminase	metabolismo primário
CP03-EB-001-025-A01-UE.F	Delta 9- desaturase de ácidos graxos	metabolismo primário
Contig191	Delta(12)- desidrogenase de ácidos graxos	metabolismo primário
CP03-EB-001-002-E08-UE.F	Desidratase do ácido delta-aminolevulinico	metabolismo primário
Contig633	Glicosiltransferase de Dolcil-difosfo-oligosacaridio subunidade DAD1	metabolismo primário
Contig594	DSBA oxidoreductase	metabolismo primário
CP03-EB-001-016-G06-UE.F	Fosfatase duplo específico	metabolismo primário
CP03-EB-001-004-D10-UE.F	Endoglucanase II	metabolismo primário
CP03-EB-001-001-E07-UE.F	Enoil-CoA hidratase	metabolismo primário
CP03-EB-001-004-G07-UE.F	Enoil-CoA hidratase	metabolismo primário
CP03-EB-001-014-G04-UE.F	Oxygenase de ácidos graxos	metabolismo primário
CP03-EB-001-004-F02-UE.F	Sintase de ácidos graxos subunidade alfa	metabolismo primário
Contig225	Formato desidrogenase	metabolismo primário
CP02-EB-001-002-A09-FS.F	Fumarato redutase (NADH)	metabolismo primário
CP03-EB-001-016-D02-UE.F	Glucose-6-fosfato 1-epimerase (hexose-6-fosfato mutarotase)	metabolismo primário
CP02-EB-001-003-A07-FS.F	Glutamate decarboxilase/ GABA sintase	metabolismo primário
CP03-EB-001-008-G12-UE.F	Glutamato sintase (NADH)	metabolismo primário
Contig482	Glutamina sintetase	metabolismo primário
CP03-EB-001-017-D07-UE.F	Glutaredoxina	metabolismo primário
CP02-EB-001-004-D01-FS.G	Glutathiona S-transferase	metabolismo primário
CP03-EB-001-005-E06-UE.F	Glutathiona-disulfito redutase	metabolismo primário
Contig229	Gliceraldeido-3-phosphate desidrogenase	metabolismo primário
CP03-EB-001-013-G07-UE.F	Gliceraldeido-3-fosfato desidrogenase	metabolismo primário
CP02-EB-001-002-A12-FS.F	Glicosidase CRH2 precursor	metabolismo primário
CP03-EB-001-026-E05-UE.F	Imidazoleglycerol-fosfato desidratase	metabolismo primário
CP03-EB-001-010-G06-UE.F	Isopentenil-difosfato Delta-isomerase	metabolismo primário
CP03-EB-001-016-D11-UE.F	Redutoisomerase do ácido cetol/ isomero-redutase do acetohidroxido	metabolismo primário
CP03-EB-001-024-E06-UE.F	L-lactato desidrogenase (FMN-dependent)	metabolismo primário
CP03-EB-001-012-D07-UE.F	Lisofosfolipase/ alfa beta hidrolase	metabolismo primário
CP03-EB-001-004-D04-UE.F	Lisi-tRNA sintetase	metabolismo primário
CP03-EB-001-015-E02-UE.F	Malato desidrogenase	metabolismo primário
CP03-EB-001-003-B11-UE.F	Manan endo-1,4-beta-manosidase	metabolismo primário
Contig23	Metaloproteinase	metabolismo primário
CP03-EB-001-008-G02-UE.F	Metionina aminopeptidase	metabolismo primário
CP03-EB-001-021-B05-UE.G	Metilcrotonoil-CoA carboxilase, cadeia beta	metabolismo primário
CP03-EB-001-021-B05-UE.F	Metilcrotonoil-CoA carboxilase, cadeia beta	metabolismo primário
CP03-EB-001-014-G11-UE.F	Protease de membrana mitocondrial interna subunidade 1	metabolismo primário
CP03-EB-001-015-C07-UE.F	Peroxiredoxina mitocondrial 1	metabolismo primário
Contig583	Proteína de biossíntese do cofator molibidênio 1 A	metabolismo primário
CP03-EB-001-021-E12-UE.F	Proteína de biossíntese do cofator molibidênio 1 A	metabolismo primário
Contig604	Mio-inositol oxigenase	metabolismo primário
CP03-EB-001-009-D05-UE.F	Mio-inositol oxigenase	metabolismo primário
CP03-EB-001-012-E06-UE.F	Mio-inositol oxigenase	metabolismo primário
CP03-EB-001-012-C11-UE.F	NAD quinase 2	metabolismo primário
Contig58	Formato desidrogenase dependente de NAD	metabolismo primário
CP03-EB-001-017-H08-UE.F	NADH Citocrome b5 redutase	metabolismo primário
CP02-EB-001-006-D07-UE.F	NADH desidrogenase	metabolismo primário
CP03-EB-001-014-C02-UE.F	NADH desidrogenase [ubiquinona] 1 subcomplexo alfa subunidade 13	metabolismo primário
Contig445	NADH-quinona oxidoreductase	metabolismo primário
CP03-EB-001-013-C12-UE.F	NADH-ubiquinona oxidoreductase subunidade 21 kDa	metabolismo primário
CP03-EB-001-010-C05-UE.F	NADH-ubiquinona oxidoreductase subunidade 21 kDa	metabolismo primário
CP03-EB-001-024-B12-UE.F	NADPH desidrogenase	metabolismo primário
CP03-EB-001-025-G10-UE.F	NADPH desidrogenase 1	metabolismo primário
CP03-EB-001-002-B04-UE.F	NADPH desidrogenase 2	metabolismo primário
CP03-EB-001-019-E01-UE.F	NADPH oxidase	metabolismo primário
CP03-EB-001-014-F08-UE.F	Nucleoside diphosphate kinase B	metabolismo primário
CP02-EB-001-001-G10-FS.F	Oligo-1,6-glucosidase	metabolismo primário
CP02-EB-001-006-F06-UE.F	Ornithine aminotransferase	metabolismo primário
Contig478	Oxidoreductase, 2-nitropropano dioxigenase	metabolismo primário
CP03-EB-001-014-D05-UE.F	Oxidoreductase, 2-nitropropano dioxigenase	metabolismo primário
CP03-EB-001-004-C07-UE.F	Pentafuntional AROM polipeptideo	metabolismo primário
CP03-EB-001-014-E07-UE.F	Tryptofano periótico proteína 1	metabolismo primário
CP03-EB-001-018-H04-UE.F	fosfo-2-desidro-3-deoxiheptonato aldolase	metabolismo primário
CP03-EB-001-017-H04-UE.F	Phosphoenolpiruvato carboxiquinase [ATP]	metabolismo primário

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP02-EB-001-006-A04-UE.F	Fosfoquetolase	metabolismo primário
CP03-EB-001-026-D03-UE.F	Fosfoproteína fosfatase	metabolismo primário
CP02-EB-001-003-D04-FS.F	Fosfopiruvato hidratase	metabolismo primário
CP03-EB-001-026-B01-UE.F	Anthrnilate sintase componente 1	metabolismo primário
CP03-EB-001-026-B04-UE.F	Oxidoreductase	metabolismo primário
CP03-EB-001-015-F04-UE.F	Sintase ATP vacuolar ATP subunidade G	metabolismo primário
CP03-EB-001-021-D02-UE.G	Prolil aminopeptidase	metabolismo primário
CP03-EB-001-009-C01-UE.F	Fosfatase Slingshot	metabolismo primário
CP03-EB-001-018-C04-UE.F	Trosina-fosfatase	metabolismo primário
CP02-EB-001-002-E10-FS.F	Proteoglicano-4 precursor	metabolismo primário
CP03-EB-001-024-E04-UE.F	Desidrogenase D-3-phosphoglicerato	metabolismo primário
CP03-EB-001-012-C03-UE.F	Dehydrogenase	metabolismo primário
CP03-EB-001-018-C05-UE.F	Piroglutamyl-peptidase 1	metabolismo primário
CP03-EB-001-003-A10-UE.F	Piruvato desidrogenase	metabolismo primário
CP03-EB-001-005-F03-UE.F	Riboflavina quinase	metabolismo primário
Contig301	Ribophorina II	metabolismo primário
CP03-EB-001-007-D05-UE.F	Serine carboxipeptidase	metabolismo primário
CP03-EB-001-014-D06-UE.F	Serine carboxipeptidase	metabolismo primário
CP03-EB-001-025-A09-UE.F	Inibidor IA-1 da serine proteinase	metabolismo primário
CP03-EB-001-018-B06-UE.F	Sulfito oxidase	metabolismo primário
CP03-EB-001-004-F06-UE.F	Tiamine pirofosfato	metabolismo primário
CP03-EB-001-021-E09-UE.G	Tiamina transportador	metabolismo primário
CP03-EB-001-020-A05-UE.F	Tioeterase	metabolismo primário
CP03-EB-001-014-G10-UE.F	Tioredoxina	metabolismo primário
CP03-EB-001-013-E06-UE.F	Tioredoxina 4A	metabolismo primário
CP03-EB-001-020-E05-UE.F	Complexo THO subunidade 4	metabolismo primário
Contig625	Transaldolase	metabolismo primário
CP03-EB-001-004-B03-UE.F	Transaldolase	metabolismo primário
CP03-EB-001-017-A08-UE.F	Transaldolase	metabolismo primário
CP03-EB-001-013-F10-UE.F	Trealose-fosfatase/ trealose fosfate sintase	metabolismo primário
Contig414	Glutamine amidotransferase (GATase1)	metabolismo primário
CP03-EB-001-017-C07-UE.F	Tirosinase	metabolismo primário
CP03-EB-001-014-G02-UE.F	Descarboxilase 1 do ácido glucuronico UDP	metabolismo primário
CP03-EB-001-018-H02-UE.F	Proteína acessória urease UreD	metabolismo primário
CP03-EB-001-025-G01-UE.F	Sintase ATP vacuolar proteolipídio de 16 kDa	metabolismo primário
CP03-EB-001-026-G04-UE.F	Sintase ATP vacuolar ATP subunidade A, isoforma osteoclastica	metabolismo primário
CP03-EB-001-004-F04-UE.F	Sintase ATP vacuolar ATP subunidade C	metabolismo primário
CP03-EB-001-013-E11-UE.F	Ycac/ Isocorismatase hidrolase	metabolismo primário
Contig593	Alcool desidrogenase tipo Znco	metabolismo primário
Contig496	26S proteasomo não-ATPase subunidade regulatória	síntese protéica e processamento
Contig369	Subunidade maior RNA ribossomal 28S	síntese protéica e processamento
CP03-EB-001-024-G10-UE.F	40S ribossomal proteína S0	síntese protéica e processamento
CP03-EB-001-026-D02-UE.F	40S ribossomal proteína S0	síntese protéica e processamento
Contig446	40S ribossomal proteína S10	síntese protéica e processamento
Contig632	40S ribossomal proteína S11	síntese protéica e processamento
CP03-EB-001-026-G02-UE.F	40S ribossomal proteína S11	síntese protéica e processamento
Contig465	40S ribossomal proteína S12	síntese protéica e processamento
Contig601	40S ribossomal proteína S12	síntese protéica e processamento
Contig206	40S ribossomal proteína S13	síntese protéica e processamento
Contig502	40S ribossomal proteína S15	síntese protéica e processamento
CP03-EB-001-017-G11-UE.F	40S ribossomal proteína S19a	síntese protéica e processamento
CP03-EB-001-015-C03-UE.F	40S ribossomal proteína S2	síntese protéica e processamento
CP03-EB-001-021-G04-UE.G	40S ribossomal proteína S20	síntese protéica e processamento
Contig272	40S ribossomal proteína S23	síntese protéica e processamento
Contig356	40S ribossomal proteína S24	síntese protéica e processamento
Contig443	40S ribossomal proteína S26	síntese protéica e processamento
Contig645	40S ribossomal proteína S28	síntese protéica e processamento
Contig65	40S ribossomal proteína S29	síntese protéica e processamento
Contig484	40S ribossomal proteína S3	síntese protéica e processamento
CP03-EB-001-002-G03-UE.F	40S ribossomal proteína S5	síntese protéica e processamento
CP03-EB-001-025-H03-UE.F	40S ribossomal proteína S5	síntese protéica e processamento
CP03-EB-001-020-D03-UE.F	40S ribossomal proteína S9-B	síntese protéica e processamento
Contig477	60S proteína ácida ribossomal P0	síntese protéica e processamento
CP03-EB-001-021-D08-UE.G	60S proteína ácida ribossomal P0	síntese protéica e processamento
CP03-EB-001-004-F12-UE.F	60S proteína ácida ribossomal P1	síntese protéica e processamento
Contig454	60S proteína ácida ribossomal P2	síntese protéica e processamento
CP03-EB-001-013-D08-UE.F	60S proteína ribossomal L10	síntese protéica e processamento
CP03-EB-001-015-D07-UE.F	60S proteína ribossomal L13	síntese protéica e processamento
Contig614	60S proteína ribossomal L16	síntese protéica e processamento
CP03-EB-001-012-G06-UE.F	60S proteína ribossomal L17-B	síntese protéica e processamento
Contig179	60S proteína ribossomal L18	síntese protéica e processamento
Contig617	60S proteína ribossomal L19-3	síntese protéica e processamento

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
Contig336	60S ribossomal proteína I1-a	síntese protéica e processamento
CP03-EB-001-016-E05-UE.F	60S ribossomal proteína L1-B	síntese protéica e processamento
CP03-EB-001-013-C05-UE.F	60S ribossomal proteína L2	síntese protéica e processamento
CP03-EB-001-017-G02-UE.F	60S ribossomal proteína L2	síntese protéica e processamento
CP03-EB-001-001-C07-UE.F	60S ribossomal proteína L24, mitochondrial precursor	síntese protéica e processamento
Contig198	60S ribossomal proteína L26-1	síntese protéica e processamento
Contig475	60S ribossomal proteína L28	síntese protéica e processamento
CP03-EB-001-018-D05-UE.F	60S ribossomal proteína L28	síntese protéica e processamento
CP03-EB-001-015-E10-UE.F	60S ribossomal proteína L3	síntese protéica e processamento
CP03-EB-001-022-H02-UE.F	60S ribossomal proteína L3	síntese protéica e processamento
Contig254	60S ribossomal proteína L30	síntese protéica e processamento
CP03-EB-001-025-E02-UE.F	60S ribossomal proteína L31	síntese protéica e processamento
CP03-EB-001-026-A10-UE.F	60S ribossomal proteína L31	síntese protéica e processamento
Contig421	60S ribossomal proteína L32	síntese protéica e processamento
Contig627	60S ribossomal proteína L32	síntese protéica e processamento
Contig404	60S ribossomal proteína L36	síntese protéica e processamento
Contig441	60S ribossomal proteína L37	síntese protéica e processamento
Contig597	60S ribossomal proteína L37a	síntese protéica e processamento
CP03-EB-001-002-C03-UE.F	60S ribossomal proteína L39	síntese protéica e processamento
CP03-EB-001-014-D10-UE.F	60S ribossomal proteína L39	síntese protéica e processamento
Contig411	60S ribossomal proteína L43	síntese protéica e processamento
CP03-EB-001-001-A07-UE.F	60S ribossomal proteína L5-b	síntese protéica e processamento
CP03-EB-001-024-D12-UE.F	60S ribossomal proteína L6	síntese protéica e processamento
Contig495	60S ribossomal proteína L9	síntese protéica e processamento
Contig457	Asparaginil-tRNA sintetase	síntese protéica e processamento
CP03-EB-001-007-H09-UE.F	Aspartico proteinase	síntese protéica e processamento
CP03-EB-001-013-B09-UE.F	Aspartil protease	síntese protéica e processamento
CP03-EB-001-019-G07-UE.F	RNA helicase dependente de ATP dbp10	síntese protéica e processamento
CP03-EB-001-005-F04-UE.F	RNA helicase dependente de ATP eIF4A	síntese protéica e processamento
CP03-EB-001-009-F11-UE.F	RNA helicase dependente de ATP eIF4A	síntese protéica e processamento
CP03-EB-001-020-E08-UE.F	RNA helicase dependente de ATP HAS1	síntese protéica e processamento
CP03-EB-001-012-D05-UE.F	Chaperone BCS1	síntese protéica e processamento
CP03-EB-001-003-D11-UE.F	Chaperone dnaJ	síntese protéica e processamento
CP03-EB-001-003-B03-UE.F	E2 ubiquitina-conjugada	síntese protéica e processamento
CP03-EB-001-005-F09-UE.F	Fator de alongamento 1-beta	síntese protéica e processamento
CP03-EB-001-017-G10-UE.F	Fator de alongamento 1-beta	síntese protéica e processamento
Contig468	Fator de alongamento 3	síntese protéica e processamento
Contig605	Fator de alongamento 3	síntese protéica e processamento
Contig612	Fator de alongamento 3	síntese protéica e processamento
CP03-EB-001-010-C04-UE.F	Fator de alongamento 3	síntese protéica e processamento
CP03-EB-001-002-F08-UE.F	Fator de iniciação da tradução eucariótica 1A	síntese protéica e processamento
CP03-EB-001-014-E08-UE.F	Fator de iniciação da tradução eucariótica 3	síntese protéica e processamento
CP03-EB-001-024-D10-UE.F	Fator de iniciação da tradução eucariótica 3 subunidade 8	síntese protéica e processamento
CP03-EB-001-009-H06-UE.F	Fator de iniciação da tradução eucariótica 5	síntese protéica e processamento
CP03-EB-001-024-G12-UE.F	Fator de iniciação da tradução eucariótica 5	síntese protéica e processamento
CP03-EB-001-017-F04-UE.F	Metalloproteinase elastinolítica extracelular	síntese protéica e processamento
CP03-EB-001-026-C09-UE.F	Metalloproteinase elastinolítica extracelular	síntese protéica e processamento
CP03-EB-001-007-E12-UE.F	Proteína 1 de ligação a FK506	síntese protéica e processamento
CP03-EB-001-016-F09-UE.F	Proteína 1 de ligação a FK506	síntese protéica e processamento
CP03-EB-001-014-E10-UE.F	Proteína 1 de ligação a FK506	síntese protéica e processamento
CP03-EB-001-010-G10-UE.F	Proteína do sistema H de clivagem de glicina	síntese protéica e processamento
CP03-EB-001-012-D04-UE.F	Proteína 2 rica em glicina que se liga a RNA	síntese protéica e processamento
CP03-EB-001-008-A05-UE.F	subunidade maior RNA ribossomal 28 S	síntese protéica e processamento
CP03-EB-001-015-E11-UE.F	39-S proteína ribossomal L47	síntese protéica e processamento
Contig85	Subunidade maior RNA ribossomal mitocondrial	síntese protéica e processamento
CP03-EB-001-015-B05-UE.F	Peptídeo processamento mitocondrial subunidade beta	síntese protéica e processamento
Contig215	Polyubiquitina	síntese protéica e processamento
Contig93	Polyubiquitina	síntese protéica e processamento
CP03-EB-001-020-H08-UE.F	Polyubiquitina	síntese protéica e processamento
CP03-EB-001-025-D01-UE.F	Polyubiquitina	síntese protéica e processamento
CP03-EB-001-009-H12-UE.F	Predobrimento subunidade 6	síntese protéica e processamento
CP03-EB-001-019-G11-UE.F	Peptidil-tRNA hidrolase 2	síntese protéica e processamento
CP02-EB-001-006-E12-UE.F	Proteassomo subunidade beta tipo 2	síntese protéica e processamento
CP03-EB-001-013-C01-UE.F	Proteassomo subunidade beta tipo 2	síntese protéica e processamento
CP03-EB-001-019-B01-UE.F	Piridoxina biosíntese PDX1	síntese protéica e processamento
CP03-EB-001-015-H01-UE.F	Biogênese de ribossomo RLP24	síntese protéica e processamento
Contig535	Fator de maturação do proteassomo ump1	síntese protéica e processamento
CP03-EB-001-016-H08-UE.F	Fator de tradução proteica SU1	síntese protéica e processamento
Contig270	Proteína bifuncional (metalopectidase e domínio de ligação a quitina)	síntese protéica e processamento
Contig565	Ribossomal proteína L10	síntese protéica e processamento
CP02-EB-001-006-B04-UE.F	Ribossomal proteína 22 subunidade pequena	síntese protéica e processamento
Contig603	Ribossomal proteína HS6-tipo (S12/L30/L7a)	síntese protéica e processamento

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP03-EB-001-019-H08-UE.F	Ribossomal proteína L1p/L10e	síntese protéica e processamento
CP03-EB-001-012-B09-UE.F	Ribossomal proteína L2	síntese protéica e processamento
CP03-EB-001-016-B02-UE.F	Ribossomal proteína L44	síntese protéica e processamento
CP02-EB-001-001-F02-FS.F	Ribossomal proteína L7a	síntese protéica e processamento
CP03-EB-001-020-C09-UE.F	Ribossomal proteína S9	síntese protéica e processamento
CP03-EB-001-009-C10-UE.F	Ribossomo biogênese proteína RLP24	síntese protéica e processamento
CP03-EB-001-016-E02-UE.F	Ribossomo biogênese proteína regulatória	síntese protéica e processamento
CP02-EB-001-006-A12-UE.F	RNA polymerase II	síntese protéica e processamento
CP03-EB-001-002-E05-UE.F	Proteína contendo domínio de reconhecimento de RNA	síntese protéica e processamento
CP03-EB-001-025-H10-UE.F	Proteína de ligação ao RNA	síntese protéica e processamento
CP03-EB-001-019-A08-UE.F	Proteína 8A de ligação ao RNA	síntese protéica e processamento
Contig631	rRNA	síntese protéica e processamento
CP03-EB-001-001-F07-UE.F	Biossíntese de Siroheme	síntese protéica e processamento
Contig566	RNA da pequena subunidade ribossomal	síntese protéica e processamento
Contig111	RNA da pequena subunidade ribossomal	síntese protéica e processamento
CP03-EB-001-007-B07-UE.F	RNA da pequena subunidade ribossomal mitocondrial	síntese protéica e processamento
CP03-EB-001-019-H09-UE.F	Fator de iniciação da transcrição eIF-2B subunidade alfa	síntese protéica e processamento
Contig208	Fator de tradução proteína SU1	síntese protéica e processamento
CP03-EB-001-019-F03-UE.F	tRNA nucleotidiltransferase	síntese protéica e processamento
CP03-EB-001-007-C11-UE.F	Triptofanil-tRNA sintetase	síntese protéica e processamento
CP03-EB-001-013-F09-UE.F	Ubiquinol-citocromo c redutase subunidade ferro-enxofre	síntese protéica e processamento
Contig240	Ubiquinona biosíntese COQ4	síntese protéica e processamento
CP03-EB-001-018-F01-UE.F	Ubiquitina	síntese protéica e processamento
CP03-EB-001-010-F11-UE.F	Ubiquitina carboxil-terminal hidrolase	síntese protéica e processamento
CP03-EB-001-021-E05-UE.G	Ubiquitina enzima de conjugação	síntese protéica e processamento
Contig448	Ubiquitina/RPS27a	síntese protéica e processamento
Contig563	Ubiquitina-enzima de conjugação E2	síntese protéica e processamento
Contig252	Ubiquitina-enzima de conjugação E3	síntese protéica e processamento
CP03-EB-001-016-E10-UE.F	Zutina	síntese protéica e processamento
CP03-EB-001-010-B07-UE.F	12-oxofitodienoato redutase	metabolismo secundário e síntese de hormônio
CP03-EB-001-017-B03-UE.F	4-coumarato-CoA ligase 1	metabolismo secundário e síntese de hormônio
CP03-EB-001-007-G02-UE.F	Quinase ativada por 5'-AMP - subunidade beta-2	metabolismo secundário e síntese de hormônio
CP03-EB-001-015-D06-UE.F	Beta eterase	metabolismo secundário e síntese de hormônio
Contig416	Carotenoide ester lipase	metabolismo secundário e síntese de hormônio
CP03-EB-001-012-F08-UE.F	Cephalosporina esterase/ carboxilesterase	metabolismo secundário e síntese de hormônio
Contig453	Ceratoplatina	metabolismo secundário e síntese de hormônio
CP03-EB-001-020-G02-UE.F	D-arabinitol 2-desidrogenase	metabolismo secundário e síntese de hormônio
CP03-EB-001-008-G04-UE.F	Endo-1,4-beta-xilanase A	metabolismo secundário e síntese de hormônio
CP03-EB-001-017-H07-UE.F	Endopolifosfatase	metabolismo secundário e síntese de hormônio
CP03-EB-001-002-G11-UE.F	Feruloil esterase B	metabolismo secundário e síntese de hormônio
CP03-EB-001-005-H05-UE.F	GPI transamidase subunidade PIG-U	metabolismo secundário e síntese de hormônio
CP03-EB-001-009-F05-UE.F	Lanosterol 14-alpha-demetilase	metabolismo secundário e síntese de hormônio
CP03-EB-001-003-E10-UE.F	Levodione redutase	metabolismo secundário e síntese de hormônio
CP03-EB-001-005-F10-UE.F	Linoleato diol sintase	metabolismo secundário e síntese de hormônio
CP03-EB-001-005-F11-UE.F	Linoleato diol sintase	metabolismo secundário e síntese de hormônio
CP03-EB-001-001-E02-UE.F	Lipoiltransferase	metabolismo secundário e síntese de hormônio
CP02-EB-001-003-A09-FS.F	Mucina-2	metabolismo secundário e síntese de hormônio
Contig487	Nitrilase	metabolismo secundário e síntese de hormônio
CP03-EB-001-020-B08-UE.F	Proteína peroxomal induzida por Oleato POX18	metabolismo secundário e síntese de hormônio
CP03-EB-001-020-F08-UE.F	Proteína de ligação a oxisterol	metabolismo secundário e síntese de hormônio
Contig569	palmitoiltransferase	metabolismo secundário e síntese de hormônio
CP03-EB-001-010-D08-UE.F	Pectinesterase	metabolismo secundário e síntese de hormônio
CP03-EB-001-013-D04-UE.F	Papalisina-1 precursor	metabolismo secundário e síntese de hormônio
CP03-EB-001-025-D07-UE.F	Estearoil-CoA 9-desaturase	metabolismo secundário e síntese de hormônio
CP03-EB-001-004-B09-UE.F	Metabolismo de esteroide	metabolismo secundário e síntese de hormônio
CP03-EB-001-001-F12-UE.F	ATPase da família AAA	sinal de transdução e regulação pós-tradução
CP03-EB-001-025-C05-UE.F	Adenilate ciclase	sinal de transdução e regulação pós-tradução
Contig449	ADP, ATP proteína carreadora	sinal de transdução e regulação pós-tradução
CP03-EB-001-017-E03-UE.F	ADP, ATP proteína carreadora	sinal de transdução e regulação pós-tradução
CP03-EB-001-016-C07-UE.F	ADP, ATP proteína carreadora	sinal de transdução e regulação pós-tradução
Contig162	ARF ativador GTPase	sinal de transdução e regulação pós-tradução
CP03-EB-001-015-F10-UE.F	Proteína de ligação a calcineurina	sinal de transdução e regulação pós-tradução
CP03-EB-001-009-D02-UE.F	Calmodulina	sinal de transdução e regulação pós-tradução
CP03-EB-001-005-E11-UE.F	Rhb1 proteína de ligação a GTP	sinal de transdução e regulação pós-tradução
CP03-EB-001-012-C01-UE.F	Proteína de ligação a GTP Rheb homolog precursor	sinal de transdução e regulação pós-tradução
CP03-EB-001-017-E05-UE.F	Proteína de ligação a GTP SAR1	sinal de transdução e regulação pós-tradução
CP03-EB-001-007-B06-UE.F	Proteína de ligação a GTP ypt1	sinal de transdução e regulação pós-tradução
CP03-EB-001-026-G11-UE.F	Proteína de ligação a GTP ypt2	sinal de transdução e regulação pós-tradução
CP03-EB-001-014-A05-UE.F	Proteína de ligação a GTP subunidade alfa 4	sinal de transdução e regulação pós-tradução
CP03-EB-001-017-A10-UE.F	Proteína de ligação a GTP subunidade gama	sinal de transdução e regulação pós-tradução
CP03-EB-001-025-F02-UE.F	Proteína 2 nucleolar de ligação a GTP	sinal de transdução e regulação pós-tradução
Contig481	Aceptor 1 Rab prenilatado	sinal de transdução e regulação pós-tradução

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP03-EB-001-003-G03-UE.F	Proteína quinase gsk3	sinal de transdução e regulação pós-tradução
CP03-EB-001-015-E06-UE.F	Proteína quinase wis1	sinal de transdução e regulação pós-tradução
Contig640	Proteína RhoA	sinal de transdução e regulação pós-tradução
CP02-EB-001-006-G11-UE.F	RNG2 proteína de ativação Ras GTPase	sinal de transdução e regulação pós-tradução
CP03-EB-001-024-B09-UE.F	Rab-4B proteína relatada a Ras	sinal de transdução e regulação pós-tradução
CP03-EB-001-020-F11-UE.F	Rab-7 proteína relatada a Ras	sinal de transdução e regulação pós-tradução
CP03-EB-001-012-H03-UE.F	Inibidor da dissociação Rho GDP	sinal de transdução e regulação pós-tradução
Contig483	Rho1 fator 1 de mudança do nucleotídeo guanina	sinal de transdução e regulação pós-tradução
CP03-EB-001-015-E01-UE.F	Rho1 fator 1 de mudança do nucleotídeo guanina	sinal de transdução e regulação pós-tradução
Contig639	Quinase Serine/threonine	sinal de transdução e regulação pós-tradução
CP03-EB-001-018-F04-UE.F	Quinase Serine/threonine PRP4	sinal de transdução e regulação pós-tradução
CP03-EB-001-013-C02-UE.F	Quinase Serine/threonine - proteína associada	sinal de transdução e regulação pós-tradução
CP03-EB-001-015-G07-UE.F	Sinal de reconhecimento 9 kDa	sinal de transdução e regulação pós-tradução
Contig410	Pequena COPII capa GTPase sar1	sinal de transdução e regulação pós-tradução
Contig472	Sorting nexin-3	sinal de transdução e regulação pós-tradução
CP03-EB-001-019-C03-UE.F	Sinal de translocação arginina geminada	sinal de transdução e regulação pós-tradução
CP03-EB-001-024-F03-UE.F	1,2-dihidroxi-3-ceto-5-metilipenteno dioxigenase 1	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-015-F09-UE.F	1-fosfatidilinositol-3-fosfato 5-quinase	tráfego de vesículas, secreção e endereçamento de proteínas
Contig485	Fator 3 de ribosilação ADP 3	tráfego de vesículas, secreção e endereçamento de proteínas
Contig344	Proteína 8 relatada a autofagia	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-017-F07-UE.F	Proteína 8 relatada a autofagia	tráfego de vesículas, secreção e endereçamento de proteínas
Contig389	Proteína de seleção do sítio de brotação em levedura	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-020-G05-UE.F	Coatmero subunidade beta	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-008-H04-UE.F	Coatmero subunidade beta/ COPI proteína superfície vesículas	tráfego de vesículas, secreção e endereçamento de proteínas
CP02-EB-001-006-H04-UE.F	Coatmero subunidade zeta	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-024-C03-UE.F	Endocitose proteína end4	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-024-D04-UE.F	ERV14 proteína de vesículas derivadas do retículo endoplasmático	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-020-B04-UE.F	Receptor de benzodiazepina tipo periférico	tráfego de vesículas, secreção e endereçamento de proteínas
Contig335	Proteína de transferência fosfatidilglicerol/fosfatidilinositol	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-003-G05-UE.F	ATPase transportadora de fosfolípidios	tráfego de vesículas, secreção e endereçamento de proteínas
Contig430	Proteína MSF1	tráfego de vesículas, secreção e endereçamento de proteínas
Contig466	Proteína YOS1	tráfego de vesículas, secreção e endereçamento de proteínas
Contig608	Protein 2 secretada protoplastos	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-021-F09-UE.G	Proteína 1 de retenção do RE	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-005-G12-UE.F	Proteína da família SCAMP	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-026-B06-UE.F	Proteína tipo SNF7ESCRT-III complexo	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-018-G01-UE.F	Proteína 4 Iosu Surfeit	tráfego de vesículas, secreção e endereçamento de proteínas
Contig285	Sinaptobrevina (proteína tipo VAMP)	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-003-C05-UE.F	Proteína transportadora	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-021-F08-UE.F	Transportina-1	tráfego de vesículas, secreção e endereçamento de proteínas
Contig479	t-SNARE	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-012-B08-UE.F	Proteína vacuolar de endereçamento protéico	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-017-B08-UE.F	Proteína vacuolar de endereçamento protéico 26	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-003-A03-UE.F	Proteína vacuolar	tráfego de vesículas, secreção e endereçamento de proteínas
CP03-EB-001-026-D09-UE.F	Proteína 94 contendo domínio 'Coiled-coil'	desconhecida
CP03-EB-001-005-E12-UE.F	Proteína expressa conservada	desconhecida
Contig125	Proteína expressa conservada	desconhecida
Contig144	Proteína expressa conservada	desconhecida
Contig145	Proteína expressa conservada	desconhecida
Contig160	Proteína expressa conservada	desconhecida
Contig161	Proteína expressa conservada	desconhecida
Contig189	Proteína expressa conservada	desconhecida
Contig194	Proteína expressa conservada	desconhecida
Contig227	Proteína expressa conservada	desconhecida
Contig232	Proteína expressa conservada	desconhecida
Contig236	Proteína expressa conservada	desconhecida
Contig243	Proteína expressa conservada	desconhecida
Contig257	Proteína expressa conservada	desconhecida
Contig279	Proteína expressa conservada	desconhecida
Contig284	Proteína expressa conservada	desconhecida
Contig287	Proteína expressa conservada	desconhecida
Contig292	Proteína expressa conservada	desconhecida
Contig323	Proteína expressa conservada	desconhecida
Contig325	Proteína expressa conservada	desconhecida
Contig326	Proteína expressa conservada	desconhecida
Contig330	Proteína expressa conservada	desconhecida
Contig331	Proteína expressa conservada	desconhecida
Contig333	Proteína expressa conservada	desconhecida
Contig349	Proteína expressa conservada	desconhecida
Contig378	Proteína expressa conservada	desconhecida
Contig383	Proteína expressa conservada	desconhecida
Contig387	Proteína expressa conservada	desconhecida

Tabela 1 – Continuação

[illegible]

Tabela 1 – Continuação

EST	Provável Função	Classes de Funções
CP03-EB-001-014-B04-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-014-B06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-014-G09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-015-E09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-015-F11-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-A05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-C09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-D09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-D10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-F03-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-G09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-G10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-H10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-017-A11-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-017-B12-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-017-C01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-017-D01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-017-D06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-018-A04-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-018-D11-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-018-E05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-A10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-F10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-G06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-G09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-H02-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-H10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-020-B01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-020-B12-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-020-G11-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-021-B06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-021-C04-UE.G	Proteína expressa conservada	desconhecida
CP03-EB-001-021-C06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-021-C06-UE.G	Proteína expressa conservada	desconhecida
CP03-EB-001-021-F05-UE.G	Proteína expressa conservada	desconhecida
CP03-EB-001-021-H03-UE.G	Proteína expressa conservada	desconhecida
CP03-EB-001-022-B07-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-023-B06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-C04-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-C12-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-D05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-E08-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-E09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-F01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-024-G05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-A02-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-C01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-D06-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-E03-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-E09-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-F01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-026-B03-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-026-B10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-026-D05-UE.F	Proteína expressa conservada	desconhecida
Contig149	Proteína expressa conservada	desconhecida
Contig268	Proteína expressa conservada	desconhecida
Contig28	Proteína expressa conservada	desconhecida
Contig424	Proteína expressa conservada	desconhecida
Contig619	Proteína expressa conservada	desconhecida
CP02-EB-001-002-G03-FS.F	Proteína expressa conservada	desconhecida
CP03-EB-001-002-G05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-003-B10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-005-E01-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-010-B03-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-013-E12-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-B04-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-C05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-016-H05-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-025-H08-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-026-F10-UE.F	Proteína expressa conservada	desconhecida
CP03-EB-001-019-F05-UE.F	Proteína expressa conservada C19C2.15c in chromosome II	desconhecida
CP03-EB-001-020-A04-UE.F	Proteína expressa conservada/ Beta flanking protein	desconhecida
Contig217	Proteína expressa conservada/ NhaC (Na ⁺ /H ⁺ antiporter)	desconhecida
CP03-EB-001-014-D02-UE.F	Proteína expressa conservada/ NhaC (Na ⁺ /H ⁺ antiporter)	desconhecida

ANEXO II

Tabela 3 – Expressão diferencial e agrupamento hierárquico dos 192 ESTs selecionados. A coluna 1 mostra as prováveis funções dos genes impressos na membrana e a sua localização na mesma (números). Os valores de expressão são expressos em LOG correspondentes às intensidades dos RNAs coletados em cada fase em relação ao coletado na fase branca

Genes representados no macroarranjo (prováveis funções)	Rosa x branca	Pós-estresse x branca	Primórdios x branca	Basidiomatas x branca	Grupo
(91) CEFALOSPORINA ESTERASE/CARBOXILESTERASE	0,0834	-0,0473	0,0087	0,0905	I
(126) CALCIUM TRANSPORTE ATPASE	0,1022	-0,0143	0,0744	0,1210	I
(1) ARIL-ÁLCOOL DESIDROGENASE AAD	0,0730	-0,0550	0,0558	0,0712	I
(7) PROTEÍNA EXPRESSA CONSERVADA	0,0641	-0,0097	0,0621	0,1582	I
(50) SPLICE ISOFORMA 1	0,0714	0,0178	0,1339	0,0775	I
(39) PROTEÍNA EXPRESSA CONSERVADA	0,0610	0,0147	0,1317	0,0783	I
(111) HIDROFOBINA SC3	0,0929	0,0235	0,1354	0,0808	I
(110) GLICOSE-6-FOSFATO 1-EPIRAMERASE	0,0981	0,0327	0,1334	0,1291	I
(97) VACUOLAR ATP SINTASE SUBUNIDADE C	0,0948	0,0292	0,1345	0,1177	I
(150) 60S PROTEÍNA ACÍDICA RIBOSSOMAL P2	0,1068	-0,0041	0,1074	0,0705	I
(51) O-METILSTERIGMATOCYSTINA OXIREDUCTASE	0,0844	-0,0216	0,1007	0,0437	I
(74) ATPASE TRANSPORTADORA DE FOSFOLÍPIDIOS	0,0505	-0,0449	0,0851	0,0631	I
(73) PROTEÍNA QUINASE GSK3	0,0363	-0,0567	0,0760	0,0539	I
(61) RECONHECIMENTO DE MAL PAREAMENTO E DNA CRUCIFORME	0,0265	-0,0485	0,0830	0,0639	I
(161) PROTEÍNA ERV14 VESÍCULAS DERMADAS DO RETÍCULO	0,0488	-0,0315	0,1004	0,0492	I
(86) ALDEÍDE DESIDROGENASE	0,0526	-0,0461	0,0864	0,0385	I
(109) AEGEROLISINA	0,0389	-0,0104	0,1144	0,0725	I
(49) DEHYDRATASE DO ÁCIDO DELTA-AMINOLEVULÍNICO	0,0315	-0,0202	0,1077	0,0634	I
(62) PRIA	0,0235	-0,0329	0,0982	0,0558	I
(122) PROTEÍNA TIPO KTI12	0,0267	-0,0261	0,1043	0,0455	I
(121) LINOLEATO DIOL SINTASE LIDS	0,0267	-0,0363	0,0971	0,0364	I
(75) 12-KDA CHOQUE TÉRMICO HSP12	-0,0084	-0,0598	0,0690	0,0557	I
(25) PEPTÍDIL PROLIL CIS-TRANS ISOMERASE / CYCLOPHILIN PP1L1	-0,0004	-0,0431	0,0857	0,0537	I
(26) AMINOÁCIDO N-ACETILTRANSFERASE	0,0180	-0,0549	0,0635	0,0927	I
(63) PROTEÍNA EXPRESSA CONSERVADA	-0,0150	-0,0559	0,0764	0,0122	I
(12) ENOIL-COA HIDRATASE	-0,0330	-0,0619	0,0668	0,0067	I
(174) ACIL-PROTEÍNA THIOESTERASE-1	-0,0055	-0,0342	0,0927	0,0291	I
(112) MFS PEPTÍDIO TRANSPORTADOR	-0,0013	-0,0610	0,0749	-0,0007	I
(52) PROTEÍNA EXPRESSA CONSERVADA	0,0042	-0,0596	0,0767	0,0016	I
(172) CHOQUE TÉRMICO HSP 70	0,0038	-0,0010	0,1068	0,0855	I
(38) AA-PR11	0,0321	-0,0035	0,1153	0,0866	I
(148) COATOMERO SUBUNIDADE ZETA	0,0038	0,0089	0,0990	0,1201	I
(183) COACTOSINA	-0,0325	-0,0234	0,0852	0,0424	I
(162) PROTEÍNA EXPRESSA CONSERVADA	-0,0394	-0,0281	0,0785	0,0442	I
(133) TRANSPORTINA-1	-0,0330	-0,0464	0,0703	0,0525	I
(186) PROTEÍNA DE CHEGAGEM DO FUSO MITÓTICO MAD2	-0,0631	-0,0732	0,0202	0,0676	I
(98) TIAMINA PIROFOSFATO	-0,0507	-0,0954	0,0233	0,0374	I
(37) LANOSTEROL 14-ALFA-DEMETILASE	-0,0819	-0,1095	-0,0102	0,0300	I
(77) TRANSPORTADOR DE FOSFATO INORGÂNICO	-0,0639	-0,0984	-0,0469	0,0708	I
(123) PROTEÍNA DA FAMÍLIA SCAMP	-0,0236	-0,0274	0,0291	0,1362	I
(2) PIM1	-0,0285	-0,0519	0,0296	0,1128	I
(160) PROTEÍNA DIF-1 (ORNITINE TRANSPORTADORA 1 MITOCONDRIAL)	-0,0736	-0,0352	0,0163	0,0900	I
(36) FERULOIL ESTERASE B (PRECURSOR) FAE-1	0,0020	0,0236	0,0489	0,1691	I
(184) UBIQUITINA/RPS27A	-0,0088	-0,0177	0,0992	0,0111	II
(9) PROTEÍNA 1 DE LIGAÇÃO AO POLIADENILATE PAB1	-0,0227	-0,0251	0,0899	0,0006	II
(99) 40S PROTEÍNA RIBOSSOMAL S10	-0,0260	-0,0424	0,0807	0,0002	II
(40) INTERSECTINA	-0,0167	-0,0362	0,0877	0,0091	II
(124) PROTEÍNA DE LIGAÇÃO AO mRNA SRP2	0,0009	-0,0030	0,1101	0,0377	II
(185) PLEOROTOLISINA B	-0,0191	-0,0092	0,0947	-0,0124	II
(175) COMPLEXO CCR4-NOT SUBUNIDADE CER4	-0,0342	-0,0214	0,0855	-0,0050	II
(169) SORTING NEXIN	-0,0195	0,0007	0,1006	0,0097	II
(170) PROTEÍNA EXPRESSA CONSERVADA	-0,0228	-0,0010	0,0990	0,0447	II
(157) NADPH DESIDROGENASE	-0,0417	0,0037	0,0887	0,0249	II
(146) PEQUENA PROTEÍNA DE CHOQUE TÉRMICO	-0,0130	0,0584	0,1093	0,0167	II
(136) PROTEÍNA DE CHOQUE TÉRMICO 90	0,0410	0,0451	0,1283	0,0047	II
(145) 60S PROTEÍNA RIBOSSOMAL L36	0,0311	0,0362	0,1298	0,0230	II
(134) POLIUBIQUITINA	0,0334	0,0446	0,1338	0,0313	II
(29) FATOR DE INICIAÇÃO DA TRANSCRIÇÃO TFIID SUBUNIDADE 10	-0,0086	0,0240	0,1091	0,0035	II
(5) PROTEÍNA REDUTASE 3-OXOACIL-ACIL-CARRIER FABG	-0,0130	0,0124	0,1014	-0,0142	II
(158) TRANSLOCASE MITOCONDRIAL MEMBRANA INTERNA SUBUNIDADE TIM 17	0,0091	0,0495	0,1179	0,0011	II
(173) PROTEÍNA QUINASE ATIVADA POR 5-AMP SUBUNIDADE BETA-2 AMPK	0,0288	0,0236	0,1299	0,0580	II
(147) PROTEIN RING2 ATIVADORA RAS GTPASE	0,0501	0,0440	0,1421	0,0621	II
(137) AA-PR11	0,0402	0,0413	0,1384	0,0562	II
(159) ENDOCITOSE PROTEÍNA END4	0,0086	0,0417	0,1258	0,0577	II
(135) PROTEÍNA EXPRESSA	0,0791	0,0779	0,1557	0,0735	II
(13) ADENINA FOSFORIBOSILTRANSFERASE 1	0,0688	0,0512	0,1483	0,0737	II
(125) GPI TRANSAMIDASE SUBUNIDADE PIG-U	0,0166	-0,0132	0,1075	0,0076	II
(16) GLUTAMATO SINTASE (NADH)	0,0353	0,0124	0,1247	0,0275	II

Tabela 3 – Continuação

Genes representados no macroarranjo (prováveis funções)	Rosa x branca	Pós-estresse x branca	Primórdios x branca	Basidiomatas x branca	Grupo
(182) 1,2-DI-HIDROXI-3-CETO-5-METILTIO-PENTENO DIOXIGENASE 1 (ARD 2)	0,0043	0,0742	0,1017	0,1407	-
(35) PROTEÍNA EXPRESSA CONSERVADA	0,0690	0,1347	0,1392	0,1524	-
(129) PROTEÍNA 1 CONTENDO 'DEDOS DE ZINCO' TIPO NFX1	-0,0620	-0,0263	-0,0898	0,0613	III
(82) PROTEÍNA EXPRESSA CONSERVADA	-0,1366	-0,0513	-0,0731	0,0128	III
(130) NADH DESIDROGENASE	-0,1004	-0,0063	-0,0476	0,0613	III
(22) PROTEÍNA EXPRESSA CONSERVADA	-0,1085	-0,0032	-0,0461	0,0470	III
(171) PEPTIDIL-PROLIL CIS-TRANS ISOMERASE - FKBP	-0,1230	-0,0168	-0,0166	0,0159	III
(154) HIDROFOBINA SC3	-0,1171	0,0292	-0,0405	-0,0354	III
(83) ENDOGLUCANASE II	-0,0824	0,0430	-0,0787	-0,0199	III
(165) PROTEÍNA TAR	-0,0098	0,1048	-0,0135	-0,0566	III
(85) PROTEÍNA DE DIVISÃO CELULAR 123	0,0030	0,0947	-0,0439	0,0349	III
(113) PROTEÍNA 8 RELATADA A AUTOFAGIA PRECURSOR	-0,0792	-0,0971	0,0231	-0,0494	IV
(65) COMPLEXO ARP2/3 SUBUNIDADE 16 KDA (P16-ARC)	-0,0717	-0,0964	0,0276	-0,0276	IV
(53) PIRUVATO DESIDROGENASE (LIPOAMIDA) QUINASE	-0,0684	-0,0799	0,0403	-0,0232	IV
(138) EXTENSINA PRECURSOR (GLICOPROTEÍNA RICA EM PROLINA)	-0,0401	-0,0926	0,0417	-0,0243	IV
(76) PROTEÍNA DE CHEGADA DO FUSO MITÓTICO MAD2	-0,0512	-0,0579	0,0614	0,0015	IV
(27) C6 FATOR DE TRANSCRIÇÃO GAL4	-0,0505	-0,0522	0,0645	-0,0143	IV
(188) MBF 1 (MULTIPROTEIN-BRIDGING FACTOR 1)	-0,0597	-0,0711	0,0420	-0,0700	IV
(100) ENOIL-COA HIDRATASE	-0,0838	-0,0906	0,0220	-0,0701	IV
(88) SUPRESSOR DE CINETOCORO PROTEÍNA 1	-0,0949	-0,1041	0,0053	-0,0879	IV
(114) ATP SINTASE CADEIA D	-0,0641	-0,1093	0,0166	-0,0728	IV
(163) 60S RIBOSSOMAL L43	-0,0136	-0,0676	0,0625	-0,0414	IV
(87) PROTEÍNA 2 RICA EM GLICINA	-0,0186	-0,0683	0,0598	-0,0481	IV
(181) PROTEÍNA EXPRESSA CONSERVADA	-0,0968	-0,0770	0,0200	-0,0023	IV
(45) FATOR DE ALONGAMENTO 3	-0,0952	-0,0689	0,0202	0,0171	IV
(28) DI-HIDROLIPOIL DESIDROGENASE LPD1	-0,0779	-0,0769	0,0297	0,0205	IV
(18) PROTEÍNA EXPRESSA	-0,0724	-0,0653	0,0432	0,0118	IV
(66) GLICINA DESIDROGENASE	-0,0589	-0,0820	0,0387	0,0192	IV
(67) PROTEÍNA EXPRESSA CONSERVADA	-0,0723	-0,0811	0,0354	0,0005	IV
(24) CALMODULINA	-0,0689	-0,0807	0,0388	-0,0090	IV
(151) ASPARAGINIL TRNA SINTETASE	-0,0983	-0,0870	0,0168	-0,0236	IV
(149) HIDROFOBINA	-0,0785	-0,1159	0,0037	-0,0072	IV
(10) BIOSÍNTESE DE SIROHEME MET8	-0,0839	-0,1068	0,0060	0,0042	IV
(152) PROTEÍNA DE LIGAÇÃO AO GTP - YPT1	-0,1233	-0,0810	-0,0018	-0,0706	IV
(48) PECTINESTERASE	-0,1081	-0,0407	0,0204	-0,0642	IV
(46) NADH-UBQUINONA OXIDOREDUCTASE SUBUNIDADE 21 KDA	-0,0904	-0,0440	0,0404	-0,0226	IV
(47) PROTEÍNA EXPRESSA CONSERVADA	-0,0558	-0,0157	0,0715	-0,0269	IV
(43) 12-OXOPHYTODIENATO REDUTASE	-0,0542	-0,0041	0,0718	-0,0366	IV
(41) PROTEÍNA EXPRESSA	-0,0488	-0,0119	0,0796	-0,0061	IV
(8) CITOCROMO P450 MONOOXIGENASE PE-3	-0,0722	-0,0438	0,0520	-0,0436	IV
(6) LIPOILTRANSFERASE LIPB	-0,0700	-0,0415	0,0522	-0,0518	IV
(55) HIDROFOBINA	-0,0753	0,0057	0,0597	0,0089	IV
(44) EXTENSINA PRECURSOR	-0,0617	0,0041	0,0728	0,0082	IV
(79) TRANSALDOSE TAL1	-0,1411	-0,0777	-0,0380	-0,1340	IV
(164) CENTRINA-3	-0,1455	-0,0900	-0,0355	-0,1157	IV
(80) METABOLISMO DO ESTEROL	-0,1335	-0,0663	-0,0138	-0,0938	IV
(115) PEQUENA PROTEÍNA DE CHOQUE TÉRMICO	-0,1389	-0,1141	-0,0390	-0,1260	IV
(54) MANAN ENDO-1,4-BETA-MANOSIDASE	-0,0731	-0,0612	0,0292	-0,0994	IV
(21) MIO-INOSITOL-OXIGENASE	-0,0765	-0,0478	0,0345	-0,0908	IV
(56) PROTEÍNA TRANSPORTADORA	-0,1100	-0,0138	0,0205	-0,0183	IV
(60) LEVODIONA REDUCTASE LVR	-0,1050	0,0080	0,0197	-0,0441	IV
(72) PROTEÍNA DE DIVISÃO CELULAR 123	-0,1044	-0,0012	0,0246	-0,0389	IV
(20) PROTEÍNA EXPRESSA CONSERVADA	-0,1029	-0,0074	0,0276	-0,0417	IV
(192) ANTRANILATO SINTASE COMPONENTE 1	-0,1502	-0,0459	-0,0386	-0,0784	IV
(156) 3-CETOACIL-COA TIOLASE, PRECURSOR PEROXISSOMAL	-0,1618	-0,0721	-0,0551	-0,0970	IV
(139) HIDROFOBINA	-0,1054	-0,1256	-0,0136	-0,0854	IV
(78) PROTEÍNA EXPRESSA CONSERVADA	-0,0979	-0,1219	-0,0103	-0,0999	IV
(90) FATOR DE MUDANÇA DO NUCLEOTÍDEO GUANINA EM RHO1 (RHO-GEF)	-0,1099	-0,1492	-0,0387	-0,1027	IV
(64) ACIL-COA DESATURASE	-0,0916	-0,1270	-0,0063	-0,0581	IV
(11) ATPase FAMÍLIA AAA	-0,0980	-0,1194	-0,0031	-0,0574	IV
(14) METIONINA AMINOPEPTIDASE	-0,1160	-0,1508	-0,0426	-0,0724	IV
(3) PROTEÍNA SOF1	-0,1180	-0,1535	-0,0458	-0,0856	IV
(101) ARIL-ALCOOL OXIDASE	-0,1256	-0,1656	-0,0669	-0,0815	IV
(99) PROTEÍNA EXPRESSA CONSERVADA	-0,1096	-0,1641	-0,0570	-0,0759	IV
(102) PROTEÍNA DE CHOQUE TÉRMICO	-0,0651	-0,1670	-0,0819	-0,1348	IV
(142) PROTEÍNA EXPRESSA CONSERVADA	-0,1116	0,0096	-0,0617	-0,1264	V
(96) TRANSPORTADOR DE GLICOSE	-0,1179	0,0065	-0,0632	-0,1222	V
(191) PROTEÍNA EXPRESSA	-0,1260	-0,0034	-0,0700	-0,1258	V

Tabela 3 – Continuação

Genes representados no macroarranjo (prováveis funções)	Rosa x branca	Pós-estresse x branca	Primórdios x branca	Basidiomatas x branca	Grupo
(131) PROTEÍNA 'DEDOS DE ZINCO' PRDM	-0,1236	-0,0036	-0,0703	-0,1301	V
(119) METILCROTONIL-COA CARBOXILASE, SUBUNIDADE BETA	-0,1334	-0,0237	-0,0937	-0,1439	V
(93) INIBIDOR DA DISSOCIAÇÃO DE RHO-GTP	-0,1372	-0,0263	-0,0925	-0,1428	V
(106) GLUTATIONE-DISULFITO REDUTASE	-0,1310	-0,0303	-0,0899	-0,1564	V
(132) ORNITINE AMINOTRANSFERASE	-0,1209	-0,0051	-0,0967	-0,1263	V
(141) ENDONUCLEASE REPARO DE DANOS UV	-0,0841	-0,0241	-0,1043	-0,1660	V
(59) INDUTOR TRANSPORTE DE GLUTONATO GTI 1	-0,0932	-0,0234	-0,1059	-0,1635	V
(117) TUBULINA ALFA 1	-0,1068	-0,0434	-0,1233	-0,1688	V
(128) SEPTINA	-0,1710	-0,1145	-0,1057	-0,1734	V
(103) CITOCROMO P450	-0,1771	-0,1311	-0,1187	-0,1742	V
(116) PROTEÍNA RAB-7 RELATADA A RAS	-0,0614	-0,0530	-0,0418	-0,1782	V
(104) PROTEÍNA EXPRESSA CONSERVADA	-0,0975	-0,0848	-0,0982	-0,2006	V
(153) PROTEÍNA EXPRESSA CONSERVADA	-0,1221	-0,0725	-0,0445	-0,1652	V
(95) ATP SINTASE SUBUNIDADE BETA	-0,1394	-0,0152	-0,0356	-0,0749	V
(94) PROTEÍNA DE LIGAÇÃO A CALCINERINA	-0,1402	-0,0236	-0,0354	-0,0924	V
(166) ALPHA/BETA HIDROLASE	-0,1350	-0,0033	-0,0522	-0,0949	V
(84) SINTASE DE ÁCIDOS GRAXOS SUBUNIDADE ALFA FAS2	-0,1266	0,0133	-0,0521	-0,0795	V
(167) D-3-PHOSPHOGLICERATO DESIDROGENASE 1	-0,1472	-0,0234	-0,0722	-0,1119	V
(120) PROTEÍNA DE BIOSÍNTESE DO COFATOR MOLIBDÊNIO	-0,1501	-0,0240	-0,0749	-0,1009	V
(69) PROTEÍNA RHEB DE LIGAÇÃO AO GTP	-0,1326	-0,0301	-0,0307	-0,1168	V
(68) PROTEÍNA VACUOLAR DE ENDEREÇAMENTO DE PROTEÍNAS	-0,1346	-0,0277	-0,0321	-0,1117	V
(105) PROTEÍNA EXPRESSA CONSERVADA	-0,1348	-0,0329	-0,0454	-0,1332	V
(81) POLÍPEPTÍDEO PENTAFUNCIONAL AROM - ARO1	-0,1432	-0,0359	-0,0642	-0,1373	V
(180) 40S RIBOSSOMAL PROTEÍNA S13	-0,1514	-0,0400	-0,0796	-0,1347	V
(155) FOSFOLIPASE HIDROXILIZAÇÃO N-ACIL-FOFATIDILETANOLAMINA	-0,1401	-0,0242	-0,0706	-0,1322	V
(143) UBIQUITINA/RPS27A	-0,1322	-0,0074	-0,0434	-0,1050	V
(108) RIBOFLAVINA QUINASE (PRECURSOR) FMN1	-0,1314	-0,0117	-0,0458	-0,1165	V
(190) ATP SINTASE VACUOLAR 16 KDA	-0,0880	0,0316	-0,0236	-0,1138	V
(71) NAD QUINASE 2, PRECURSOR CLOROPLASTO	-0,0885	0,0367	-0,0122	-0,1000	V
(70) PIRANOSE DESIDROGENASE	-0,1025	0,0339	-0,0359	-0,0944	V
(32) PROTEÍNA DE LIGAÇÃO A SEQUÊNCIA DE LOCALIZAÇÃO NUCLEAR	-0,0865	0,0469	0,0134	-0,0609	V
(107) RHB1 - PROTEÍNA LIGAÇÃO A GTP	-0,1133	-0,0173	-0,0397	-0,1444	V
(57) PROTEÍNA 1 RICA EM CISTEÍNA	-0,1169	-0,0299	-0,0280	-0,1390	V
(31) NADPH DESIDROGENASE 2	-0,1074	-0,0111	-0,0146	-0,1245	V
(144) PROTEÍNA RELATADA A RAS - RAB-4B	-0,0912	-0,0021	0,0116	-0,1071	V
(168) L-LACTATO DESIDROGENASE (FMN-DEPENDENTE)	-0,1043	-0,0044	0,0030	-0,1013	V
(19) PROTEÍNA FOSFATASE SLINGSHOT	-0,0886	0,0162	0,0164	-0,0912	V
(118) AGEROLISINA AA-PRI 1	-0,1134	-0,0064	-0,0230	-0,1213	V
(179) PLEOROTOLISINA B	-0,1157	-0,0017	-0,0314	-0,1210	V
(58) HIDROFOBINA	-0,1092	0,0042	-0,0230	-0,1172	V
(178) PRIA	-0,1104	0,0098	-0,0104	-0,0946	V
(42) PROTEÍNA EXPRESSA CONSERVADA	0,0198	0,0212	0,0927	-0,0568	V
(30) PROTEÍNA LA (PRECURSOR DE RNAS POLIMERASE 3)	-0,0246	0,0374	0,0262	-0,1151	V
(33) DESIDRATASE DO ÁCIDO DELTA AMINOLEVULÍNICO HEM 2	-0,0495	0,0291	0,0652	-0,0566	V
(17) COATOMERO SUBUNIDADE BETA/PROTEÍNA SUPERFÍCIE VESÍCULAS	-0,0396	0,0276	0,0805	-0,0369	V
(176) PROTEÍNASE ASPARTATO CARP	-0,0614	0,0215	0,0427	-0,0851	V
(177) AGEROLISINA AA-PRI1	0,0921	-0,0594	0,0278	0,0043	VI
(34) FATOR DE INICIAÇÃO DE TRADUÇÃO EUCARIÓTICA 1A2	0,1174	0,0464	0,1027	-0,0133	VI
(4) NADPH CITOCROMO-P450 REDUTASE	0,1217	0,0603	0,1199	0,0076	VI
(189) PROTEÍNA EXPRESSA	0,0989	-0,0204	0,0716	-0,0200	VI
(15) ENDO-1,4-BETA-XILANASE-A	0,1037	0,0148	0,1105	0,0068	VI
(187) ADENILATE CICLASE	0,0307	-0,0409	0,0657	-0,0672	VI
(23) 2-NITROPROPANO DIOXIGENASE	0,0394	-0,0722	0,0356	-0,0693	VI
(140) PROTEÍNA EXPRESSA CONSERVADA	0,0320	-0,0465	-0,0591	-0,1301	VI
(92) 2-NITROPROPANO DIOXIGENASE-LIKE PROTEIN	-0,0044	-0,0768	-0,0576	-0,1600	VI
(127) FOSFOCETOLASE	0,0107	-0,0834	-0,0906	-0,1300	VI

ANEXO III

Tabela 4 – Lista dos ESTs que apresentaram similaridade a *L. bicolor* e *U. maydis*

EXPRESSED SEQUENCE TAGS (ESTs) COM SIMILARIDADE A <i>LACCARIA BICOLOR</i> E NÃO A <i>USTILAGO MAYDIS</i>				
Contig244	CP02-EB-001-004-G02-FS.G	CP03-EB-001-012-D07-UE.F	CP03-EB-001-017-G11-UE.F	CP03-EB-001-025-A02-UE.F
Contig217	CP02-EB-001-004-C03-FS.G	CP03-EB-001-010-C04-UE.F	CP03-EB-001-017-D12-UE.F	CP03-EB-001-024-F08-UE.F
Contig125	CP02-EB-001-001-C05-FS.F	CP03-EB-001-008-C11-UE.F	CP03-EB-001-017-B03-UE.F	CP03-EB-001-024-D05-UE.F
Contig324	CP02-EB-001-006-E07-UE.F	CP03-EB-001-012-H12-UE.F	CP03-EB-001-018-A03-UE.F	CP03-EB-001-025-B12-UE.F
Contig331	CP02-EB-001-006-F05-UE.F	CP03-EB-001-013-A03-UE.F	CP03-EB-001-018-B05-UE.F	CP03-EB-001-025-D02-UE.F
Contig290	CP02-EB-001-006-C08-UE.F	CP03-EB-001-012-F07-UE.F	CP03-EB-001-017-H09-UE.F	CP03-EB-001-025-B01-UE.F
Contig163	CP02-EB-001-003-A07-FS.F	CP03-EB-001-009-F10-UE.F	CP03-EB-001-017-D01-UE.F	CP03-EB-001-024-E09-UE.F
Contig151	CP02-EB-001-001-F06-FS.F	CP03-EB-001-009-A06-UE.F	CP03-EB-001-017-B10-UE.F	CP03-EB-001-024-E05-UE.F
Contig227	CP02-EB-001-004-D02-FS.G	CP03-EB-001-010-F11-UE.F	CP03-EB-001-017-F10-UE.F	CP03-EB-001-024-G05-UE.F
Contig180	CP02-EB-001-003-B02-FS.F	CP03-EB-001-010-A06-UE.F	CP03-EB-001-017-D02-UE.F	CP03-EB-001-024-F03-UE.F
Contig325	CP02-EB-001-006-E09-UE.F	CP03-EB-001-013-A02-UE.F	CP03-EB-001-018-A04-UE.F	CP03-EB-001-025-C01-UE.F
Contig187	CP02-EB-001-003-C02-FS.F	CP03-EB-001-010-B04-UE.F	CP03-EB-001-017-D07-UE.F	CP03-EB-001-024-F06-UE.F
Contig348	CP03-EB-001-001-B01-UE.F	CP03-EB-001-013-A06-UE.F	CP03-EB-001-018-B06-UE.F	CP03-EB-001-025-D06-UE.F
Contig144	CP02-EB-001-001-F05-FS.F	CP03-EB-001-008-F12-UE.F	CP03-EB-001-017-B09-UE.F	CP03-EB-001-024-D12-UE.F
Contig155	CP02-EB-001-001-H06-FS.F	CP03-EB-001-009-C01-UE.F	CP03-EB-001-017-C07-UE.F	CP03-EB-001-024-E06-UE.F
Contig191	CP02-EB-001-003-G03-FS.F	CP03-EB-001-010-B07-UE.F	CP03-EB-001-017-D10-UE.F	CP03-EB-001-024-F07-UE.F
Contig218	CP02-EB-001-004-D01-FS.G	CP03-EB-001-010-D08-UE.F	CP03-EB-001-017-E10-UE.F	CP03-EB-001-024-F09-UE.F
Contig232	CP02-EB-001-004-G01-FS.G	CP03-EB-001-010-H06-UE.F	CP03-EB-001-017-G07-UE.F	CP03-EB-001-024-H11-UE.F
Contig270	CP02-EB-001-006-B03-UE.F	CP03-EB-001-012-E02-UE.F	CP03-EB-001-017-H07-UE.F	CP03-EB-001-025-A09-UE.F
Contig156	CP02-EB-001-002-A11-FS.F	CP03-EB-001-009-F08-UE.F	CP03-EB-001-017-C10-UE.F	CP03-EB-001-024-E08-UE.F
Contig301	CP02-EB-001-006-C10-UE.F	CP03-EB-001-012-G08-UE.F	CP03-EB-001-017-H12-UE.F	CP03-EB-001-025-B04-UE.F
Contig369	CP03-EB-001-001-C05-UE.F	CP03-EB-001-013-B12-UE.F	CP03-EB-001-018-C04-UE.F	CP03-EB-001-025-D08-UE.F
Contig384	CP03-EB-001-001-D07-UE.F	CP03-EB-001-013-C05-UE.F	CP03-EB-001-018-D11-UE.F	CP03-EB-001-025-E03-UE.F
Contig387	CP03-EB-001-001-H04-UE.F	CP03-EB-001-013-C10-UE.F	CP03-EB-001-018-H02-UE.F	CP03-EB-001-025-F01-UE.F
Contig391	CP03-EB-001-002-B04-UE.F	CP03-EB-001-013-D02-UE.F	CP03-EB-001-019-C03-UE.F	CP03-EB-001-025-F02-UE.F
Contig402	CP03-EB-001-002-C03-UE.F	CP03-EB-001-013-E03-UE.F	CP03-EB-001-019-C04-UE.F	CP03-EB-001-025-G10-UE.F
Contig403	CP03-EB-001-002-C07-UE.F	CP03-EB-001-013-F11-UE.F	CP03-EB-001-019-C06-UE.F	CP03-EB-001-026-A10-UE.F
Contig404	CP03-EB-001-002-D04-UE.F	CP03-EB-001-013-G08-UE.F	CP03-EB-001-019-C12-UE.F	CP03-EB-001-026-B02-UE.F
Contig405	CP03-EB-001-003-A03-UE.F	CP03-EB-001-013-H03-UE.F	CP03-EB-001-019-E01-UE.F	CP03-EB-001-026-B03-UE.F
Contig408	CP03-EB-001-003-A06-UE.F	CP03-EB-001-013-H05-UE.F	CP03-EB-001-019-E12-UE.F	CP03-EB-001-026-C04-UE.F
Contig412	CP03-EB-001-003-B11-UE.F	CP03-EB-001-014-A05-UE.F	CP03-EB-001-019-F02-UE.F	CP03-EB-001-026-E09-UE.F
Contig414	CP03-EB-001-003-E02-UE.F	CP03-EB-001-014-C09-UE.F	CP03-EB-001-019-F03-UE.F	CP03-EB-001-026-G06-UE.F
Contig415	CP03-EB-001-003-E05-UE.F	CP03-EB-001-014-C10-UE.F	CP03-EB-001-019-F04-UE.F	CP03-EB-001-026-H11-UE.F
Contig419	CP03-EB-001-003-F01-UE.F	CP03-EB-001-014-D01-UE.F	CP03-EB-001-019-G05-UE.F	CP03-EB-001-021-H06-UE.G
Contig422	CP03-EB-001-003-F11-UE.F	CP03-EB-001-014-D02-UE.F	CP03-EB-001-019-G06-UE.F	CP03-EB-001-022-F12-UE.F
Contig423	CP03-EB-001-004-B05-UE.F	CP03-EB-001-014-D10-UE.F	CP03-EB-001-019-G12-UE.F	CP03-EB-001-023-B06-UE.F
Contig434	CP03-EB-001-004-D05-UE.F	CP03-EB-001-014-F03-UE.F	CP03-EB-001-019-H08-UE.F	CP03-EB-001-024-A05-UE.F
Contig438	CP03-EB-001-004-F07-UE.F	CP03-EB-001-014-H07-UE.F	CP03-EB-001-019-H10-UE.F	CP03-EB-001-024-B12-UE.F
Contig439	CP03-EB-001-004-F11-UE.F	CP03-EB-001-015-A04-UE.F	CP03-EB-001-020-A04-UE.F	CP03-EB-001-024-C04-UE.F
Contig444	CP03-EB-001-004-H10-UE.F	CP03-EB-001-015-A09-UE.F	CP03-EB-001-020-A05-UE.F	CP03-EB-001-024-C08-UE.F
Contig447	CP03-EB-001-005-A11-UE.F	CP03-EB-001-015-B12-UE.F	CP03-EB-001-020-B02-UE.F	CP03-EB-001-024-C12-UE.F
Contig452	CP03-EB-001-005-B04-UE.F	CP03-EB-001-015-C02-UE.F	CP03-EB-001-020-B12-UE.F	CP03-EB-001-016-D08-UE.F
Contig453	CP03-EB-001-005-B05-UE.F	CP03-EB-001-015-D10-UE.F	CP03-EB-001-020-E06-UE.F	CP03-EB-001-016-D10-UE.F
Contig461	CP03-EB-001-005-B11-UE.F	CP03-EB-001-015-F02-UE.F	CP03-EB-001-020-G06-UE.F	CP03-EB-001-016-G01-UE.F
Contig468	CP03-EB-001-005-C12-UE.F	CP03-EB-001-015-G01-UE.F	CP03-EB-001-020-G10-UE.F	CP03-EB-001-016-G03-UE.F
Contig479	CP03-EB-001-005-E01-UE.F	CP03-EB-001-015-G05-UE.F	CP03-EB-001-021-B06-UE.F	CP03-EB-001-016-G06-UE.F
Contig486	CP03-EB-001-005-E04-UE.F	CP03-EB-001-015-G09-UE.F	CP03-EB-001-021-B08-UE.G	CP03-EB-001-016-H04-UE.F
Contig541	CP03-EB-001-005-G06-UE.F	CP03-EB-001-016-A05-UE.F	CP03-EB-001-021-C04-UE.G	CP03-EB-001-016-H10-UE.F
Contig575	CP03-EB-001-005-G11-UE.F	CP03-EB-001-016-B03-UE.F	CP03-EB-001-021-C06-UE.G	CP03-EB-001-017-A06-UE.F
Contig581	CP03-EB-001-005-H03-UE.F	CP03-EB-001-016-B12-UE.F	CP03-EB-001-021-C07-UE.F	CP03-EB-001-007-D09-UE.F
Contig584	CP03-EB-001-005-H04-UE.F	CP03-EB-001-016-C04-UE.F	CP03-EB-001-021-C09-UE.G	CP03-EB-001-007-E09-UE.F
Contig590	CP03-EB-001-005-H05-UE.F	CP03-EB-001-016-C05-UE.F	CP03-EB-001-021-E09-UE.G	CP03-EB-001-007-E11-UE.F
Contig594	CP03-EB-001-007-A02-UE.F	CP03-EB-001-016-C09-UE.F	CP03-EB-001-021-F02-UE.G	CP03-EB-001-007-F09-UE.F
Contig607	CP03-EB-001-007-C05-UE.F	CP03-EB-001-016-C12-UE.F	CP03-EB-001-021-F08-UE.F	CP03-EB-001-007-F10-UE.F
Contig613	CP03-EB-001-007-D05-UE.F	CP03-EB-001-016-D02-UE.F	CP03-EB-001-021-F08-UE.G	CP03-EB-001-007-G02-UE.F
Contig618	CP03-EB-001-007-D08-UE.F	CP03-EB-001-016-D05-UE.F	CP03-EB-001-021-H03-UE.G	CP03-EB-001-007-H08-UE.F
Contig621	Contig643	Contig647	Contig652	CP03-EB-001-007-H09-UE.F
Contig622	Contig644	Contig648	Contig81	
EXPRESSED SEQUENCE TAGS (ESTs) COM SIMILARIDADE A <i>USTILAGO MAYDIS</i> E NÃO A <i>LACCARIA BICOLOR</i>				
Contig160	Contig58	CP03-EB-001-010-B05-UE.F	CP02-EB-001-003-D01-FS.F	CP03-EB-001-004-A02-UE.F
Contig225	Contig636	CP03-EB-001-010-F08-UE.F	CP02-EB-001-006-A04-UE.F	CP03-EB-001-005-C06-UE.F
Contig368	CP03-EB-001-002-G11-UE.F	CP03-EB-001-015-E11-UE.F	CP02-EB-001-006-B04-UE.F	CP03-EB-001-017-B04-UE.F
Contig401	CP03-EB-001-021-G04-UE.G	CP03-EB-001-016-G09-UE.F	CP03-EB-001-001-C04-UE.F	CP03-EB-001-018-C05-UE.F
Contig466				

ANEXO IV



Campus Soane Nazaré de Andrade, 03 de dezembro de 2007

OFÍCIO Nº 57

Ilma Sra.
MSc. Acássia B.L. Leal
Nesta instituição

Venho por meio desta informar que o requerimento para Autorização para Trabalho em Contenção com OGM projeto

Isolamento e caracterização de genes relacionados à produção de basidiomas de *Crinipellis pernicios*

recebeu o seguinte registro na CIBIO/UESC e foi aprovado.

015/2005

Atenciosamente,

Dr. Martin Brendel

Dr. Martin Brendel
Comissão Interna de Biossegurança
Presidente - Cíbio / Uesc