

UNIVERSIDADE ESTADUAL DE CAMPINAS

Carolina Cotomacci

**Análise genômica do sistema *mating type* de
Crinipellis pernicioso, fungo causador da vassoura-
de-bruxa em *Theobroma cacao***

**Tese apresentada ao Instituto de
Biologia para obtenção do Título de
Mestre em Biologia Funcional e
Molecular, na área de Bioquímica**

**Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira
Co-orientador: Prof. Dr. Lyndel Wayne Meinhardt**

2004

Campinas, 02 de fevereiro de 2004

BANCA EXAMINADORA:

Prof. Dr. Gonçalo Amarante Guimarães Pereira

(orientador)

Prof. Dr. Cláudio Luiz Messias

Prof. Dr. José Camillo Novello

Prof. Dr. Marcelo Brocchi

Dedico este trabalho,

Aos meus pais Antonio Carlos e Zorilda, pelo amor, segurança, estímulo e confiança que me deram desde meus primeiros rabiscos.

Aos meus irmãos Fernando, Gustavo e Giovana, pelo amor, estímulo e amizade.

Ao meu namorado Wagner, pelo seu amor, estímulo e compreensão.

A todos vocês, pela paciência nos meus momentos difíceis.

*A todos,
por existirem e estarem ao meu lado.*

*A todos,
o meu eterno amor.*

Agradecimentos,

A Deus, nosso pai criador, agradeço por tudo que sou, por tudo que tenho.

Ao meu orientador, Gonçalo, pela confiança e oportunidades concedidas.

Ao meu co-orientador, Lyndel, por me socorrer quando precisei.

À banca examinadora, pela grande contribuição a este trabalho.

À CAPES, ao Governo da Bahia e ao CNPQ, pelo apoio financeiro.

À todos os integrantes do Projeto Genoma Crinipellis perniciososa.

À Johana pelas importantes correções neste trabalho e por sua amizade.

À Mariciene pelo microarray, pelas correções e por sua amizade.

Ao Zé Pedro por me ensinar trabalhar com ferramentas de filogenia.

Ao Eduardo, Marcos Renato e Marcelo por me ajudarem com a bioinformática.

À Eliane e Andréa por resolverem todos os problemas burocráticos para a realização deste trabalho.

Aos meus queridos “filhos de pesquisa”, Vitor, Sula, Camila e Gabriel que me ajudaram muito.

À Diana, por ter sido a porta de entrada neste laboratório e por sua amizade.

À Raquel, por ter assumido a coordenação de DNA no meu lugar para eu terminar a tese e por sua amizade.

À Andréa, pelo sequenciamento e amizade.

Ao Anderson, por ter me ensinado muito.

Às amigas Lilica, Odalys, Naiara e Camila, pela amizade.

A todos do LGE que tornaram possível a execução deste trabalho.

À Andréia e Marina, secretárias do departamento de Bioquímica, pela ajuda e carinho.

A todos que contribuíram direta ou indiretamente para a realização deste trabalho.

ÍNDICE

RESUMO	IX
ABSTRACT.....	X
1.0 INTRODUÇÃO.....	11
1.1 A DOENÇA VASSOURA-DE-BRUXA EM <i>THEOBROMA CACAO</i> L (CACAUEIRO).....	11
1.2 CICLO DE VIDA DE <i>CRINIPELLIS PERNICIOSA</i> EM <i>THEOBROMA CACAO</i>	11
1.3 O SISTEMA SEXUAL EM FUNGOS	15
1.3.1 <i>O ciclo sexual dos basidiomicetos</i>	16
1.3.2 <i>Heterotalismo e homotalismo</i>	17
1.4 O SISTEMA MATING TYPE	19
1.4.1 <i>O Sistema Mating type em <i>Saccharomyces cerevisiae</i></i>	21
1.4.2 <i>O Sistema Mating type de <i>Basidiomicetos</i></i>	22
1.5 A GENÔMICA E SUAS CONTRIBUIÇÕES	40
1.5.1 <i>O projeto genoma de <i>Crinipellis pernicioso</i></i>	42
OBJETIVOS	44
2.0 MATERIAIS E MÉTODOS	44
2.1 SEQÜENCIAMENTO DO GENOMA DE <i>CRINIPELLIS PERNICIOSA</i>	45
2.1.1 <i>Construção de bibliotecas genômicas</i>.....	45
2.1.2 <i>Seqüenciamento do genoma de <i>Crinipellis pernicioso</i></i>.....	46
2.1.3 <i>Busca por genes mating type</i>.....	47
2.2 NÍVEL DE SIMILARIDADE E FILOGENIA DOS GENES DE MATING TYPE.....	48
2.3 ANÁLISE DAS REGIÕES CONSERVADAS DAS PROTEÍNAS HD1, HD2 E RC	48
3.0 RESULTADOS	50
3.1 IDENTIFICAÇÃO DOS GENES DO SISTEMA MATING TYPE DE <i>C. PERNICIOSA</i>.....	50

3.1.1	<i>Seqüenciamento do genoma de Crinipellis perniciososa</i>	50
3.1.2	<i>Busca por genes mating type</i>	50
3.2	NÍVEL DE SIMILARIDADE E FILOGENIA DOS GENES DE MATING TYPE	52
	GENE <i>BE U. MAYDIS</i>	53
	GENE <i>BW U. MAYDIS</i>	53
	GENE <i>PRA1 U. MAYDIS</i>	53
	GENE <i>PRA2 U. MAYDIS</i>	53
3.3	ANÁLISE DAS REGIÕES CONSERVADAS DAS PROTEÍNAS HD1, HD2 E RC	58
4.0	DISCUSSÃO	62
5.0	CONCLUSÕES	69
6.0	REFERÊNCIAS BIBLIOGRÁFICAS	70

RESUMO

A doença vassoura-de-bruxa em *Theobroma cacao* (cacaueiro), causada pelo fungo *Crinipellis perniciosa*, é uma das doenças de maior impacto econômico nos países produtores de cacau, sendo o Brasil um destes países. Esse fungo infecta os tecidos meristemáticos do cacaueiro em duas fases: parasítica e saprofítica. Pesquisas com outros fitopatógenos têm demonstrado que a mudança da fase parasítica para a saprofítica é regulada por genes do sistema *mating type* tornando este estudo extremamente importante para inferir estratégias de combate à doença.

O presente trabalho teve por finalidade identificar os genes *mating type* de *C. perniciosa* através da análise dos dados gerados pelo seu projeto genoma. Para a identificação dos genes do sistema *mating type* (*Hd1*, *Hd2*, *Rc* e *Fe*) foram feitas buscas no banco de dados do genoma de *C. perniciosa* comparando seqüências similares àquelas codificadas pelos fungos basidiomicetos *Coprinus cinereus*, *Coprinus bilanatus*, *Schizophyllum commune* e *Ustilago maydis*, identificadas e disponibilizadas em rede.

Foram identificados seis genes do sistema *mating type* de *C. perniciosa*. Um gene que codifica a proteína regulatória *Hd1*, um gene que codifica a proteína regulatória *Hd2* e quatro genes que codificam as proteínas receptoras de ferormônio *Rc1*, *Rc2*, *Rc3* e *Rc4*. Não foram identificados genes que codificam ferormônios. Portanto, concluímos que a organização molecular do sistema *mating type* de *C. perniciosa* é tetrapolar, contendo o *locus HD* bialélico e o *locus FRF* multialélico.

ABSTRACT

The witch's broom disease in *Theobroma cacao*, caused by mushroom *Crinipellis pernicioso*, is one of the diseases with the biggest economic impact in cocoa producing countries, and Brazil is one of them. The mushroom infects the meristematic tissues of the cocoa tree in two phases: parasitic and saprophytic. Research with other phytopathogens have demonstrated that the change from the parasitic to the saprophytic phase is regulated by genes of the mating type system, making this study extremely important to infer fighting strategies to this disease.

This work's proposal is to identify mating type genes of *C. pernicioso* through the analysis of data generated by its genome project. The genes identification of the (Hd1, Hd2, Rc and Fe) mating type system was made by database search in the *C. pernicioso* genome comparing similar sequences with the ones codified by the basidiomycetes mushrooms *Coprinus cinereus*, *Coprinus bilanatus*, *Schizophyllum commune* and *Ustilago maydis*, identified and available in the internet.

Six genes were identified in *C. pernicioso* mating type system. One gene that codifies the regulatory protein Hd1, another that codifies the regulatory protein Hd2 and four genes that codify the pheromone receptor proteins Rc1, Rc2, Rc3 and Rc4. Genes that codify pheromones were not identified. Therefore, we concluded that the molecular organization of *C. pernicioso* mating type system is tetrapolar, containing the bialelic HD locus and the multialelic FRF locus.

1.0 INTRODUÇÃO

1.1 A doença Vassoura-de-Bruxa em *Theobroma cacao* L (Cacaueiro)

A doença vassoura-de-bruxa em *Theobroma cacao* L (cacaueiro), causada pelo fungo basidiomiceto *Crinipellis perniciosus* (Stahel) Singer, é a doença de maior impacto econômico nos países produtores de cacau da América do Sul e das ilhas do Caribe (Silva 1999).

Detectada pela primeira vez em 1989 no sul da Bahia, região responsável por 84,5% da produção nacional de cacau, o equivalente na época a 15% da produção mundial (Pereira *et al* 1989), a doença vassoura-de-bruxa alastrou-se consideravelmente ocasionando queda na produção e ameaçando a competitividade da cacauicultura nacional. A queda do preço do cacau, associada a condições climáticas altamente propícias à disseminação do fungo causador da vassoura-de-bruxa, dificultou a exploração desse produto o qual foi abandonado em inúmeras fazendas da região (Trevisan 1996).

1.2 Ciclo de vida de *Crinipellis perniciosus* em *Theobroma cacao*

Os sintomas da doença vassoura-de-bruxa foram descritos pela primeira vez em 1895 (Stell 1928). Porém, o fungo causador da doença somente foi caracterizado em 1915 por Stahel (Stahel 1915), que lhe deu o nome de *Marasmius perniciosus*. Singer em 1942 (Singer 1942) transferiu esta espécie para o gênero *Crinipellis* e o nome de *Crinipellis perniciosus* (Stahel) Singer tem sido usado até o momento (Purdy 1996).

Crinipellis perniciosus pertence à família Tricholomataceae da ordem Agaricales, divisão Basidiomycota. Foram descritas três variedades infectantes para esta espécie: *perniciosus*,

ecuatoriensis e *citriniceps* (Purdy 1996). Outros autores têm dividido essa espécie em quatro biótipos ou patótipos dependendo da espécie que hospedam: biótipo S, biótipo B, biótipo L e biótipo C (Griffith 1994). O biótipo S (*Solanum*) causa vassoura-de-bruxa em solanaceas (Bastos 1985), o biótipo B (*Bixa*) em *Bixa orellana* (Bastos *et al* 1988) e o biótipo C (cacau) em *Theobroma cacao*, enquanto que o biótipo L (liana) não causa vassoura-de-bruxa em seu hospedeiro (Hedger *et al* 1987).

O ciclo de vida de *C. perniciosa* é comparável em muitos aspectos com de outros basidiomicetos, em que dois núcleos dentro do basídio sofrem cariogamia seguida de meiose e migração dos núcleos aos quatro basidiósporos (Delgado 1976; Alexopoulos 1996).

Os basidiósporos de *C. perniciosa* infectam os tecidos meristemáticos do cacauzeiro, tais como gemas vegetativas, frutos, brotos e almofadas florais (Baker 1957; Thorold 1975; Evans 1981) (Orchard 1988) (Fig.1-2), e germinam formando um micélio primário (monocariótico) entre as células (micélio intercelular) (Fig.1-3). Nessa fase, denominada parasítica ou biotrófica, o fungo se localiza entre as células do tecido infectado, o qual permanece vivo e apresenta nítidos sintomas de hiperplasia e hipertrofia. Nessa fase o tecido infectado é denominado de vassoura verde e permanece assim por alguns dias (Fig.1-4).

A partir de um determinado momento, por razões ainda desconhecidas, ocorre a fusão de hifas, gerando um micélio dicariótico (secundário), reconhecível ao microscópio pela formação dos grampos de conexão (delgado) (Fig. 1-5). Esta fase é a saprofítica, em que o tecido do hospedeiro torna-se necrosado (Evans 1979; Calle *et al* 1982; Evans 1980). A morfologia do ramo infectado pelo fungo guarda semelhança com a forma de uma vassoura, surgindo assim o nome popular de vassoura-de-bruxa ou vassoura seca (Evans 1980) (Fig. 1-6). Essa fase pode permanecer por longos períodos, até mesmo alguns anos.

Periodicamente, sob condições ambientais favoráveis de luminosidade e de períodos intermitentes de umidade e seca (Purdy 1996) (Fig.1-7), o micélio secundário sofre uma

morfogênese complexa para formar os corpos de frutificação (basidiocarpos) (Fig. 1-8). Dos basidiocarpos emergem os basídios, os quais carregam os basidiósporos que quando liberados podem infectar novos ramos (Fig.1-1). As condições climáticas do sul do estado da Bahia favorecem a produção de basidiósporos durante todo o ano (Trevisan 1996).

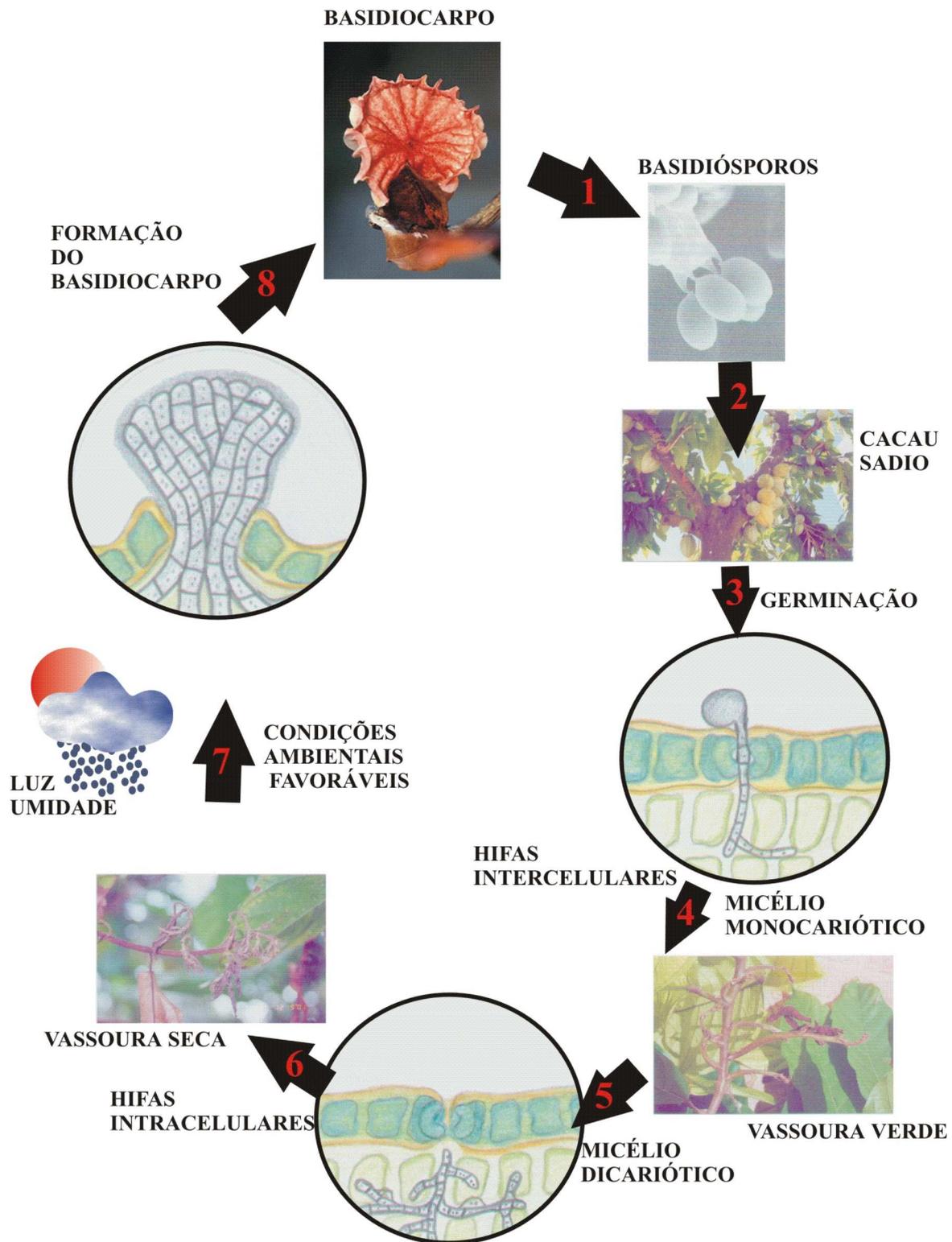


Figura 1. Ciclo de vida de *Crinipellis perniciosus* em *Theobroma cacao*.

As setas indicam os estágios do desenvolvimento: 1) formação dos basidiósporos (Delgado 1976). 2) Dispersão e infecção dos esporos no cacauzeiro. 3) Germinação do esporo penetrando o tecido vegetal através dos estômatos e formando o micélio monocariótico intercelular (ilustração). 4) Hipertrofia e hiperplasia do tecido cacauzeiro, vassoura verde. 5) Formação do micélio dicariótico intracelular através de gramplos de conexão (ilustração). 6) Necrose do tecido cacauzeiro formando a vassoura seca. Seta 7) Período de condições ambientais favoráveis como umidade e luz. 8) Formação do basidiocarpo (ilustração). 9) Basidiocarpo maduro.

1.3 O Sistema Sexual em Fungos

Dentro de fungos quatro distintos filos são reconhecidos: Chytridiomycota, Zygomycota, Ascomycota e Basidiomycota. Outros grupos normalmente classificados como fungos são geralmente separados dos fungos verdadeiros e classificados em Protozoa ou Chromista (Rapper 1966).

A complexidade de mecanismos celulares e moleculares do sistema sexual, os quais promovem o cruzamento, estão relacionados com a complexidade relativa de seus ciclos de vida. A posição da mitose no ciclo sexual determina o tipo de ciclo de vida e a extensão da mitose determina o nível de complexidade obtido pelo organismo, conforme descrito abaixo (Hiscock 1999).

No ciclo de vida haplôntico somente as células haplóides podem-se dividir por mitose e a fase diplóide é representada pelo zigoto. Esse ciclo de vida é característico de protistas, algas unicelulares e certos fungos como *Schizosaccharomyces pombe*.

No ciclo de vida diplôntico a mitose ocorre somente em células diplóides. São encontrados em organismos multicelulares, particularmente animais e certas algas, mas também pode ocorrer em diatomos unicelulares e em certos fungos como *Saccharomyces cerevisiae*.

No ciclo de vida haplo-diplôntico, geralmente referido como alternância de gerações, a mitose está tanto na fase haplóide como na diplóide. Este ciclo é tipicamente encontrado em plantas terrestres e muitas algas.

A maioria dos eucariotos exibem um ciclo de vida que corresponde a uma das três formas básicas descritas acima. Entretanto, no grupo dos fungos em várias espécies de ascomicetos e basidiomicetos o ciclo de vida é caracterizado pela separação temporal entre a fusão de hifas mononucleadas (plasmogamia) e a fusão nuclear (cariogamia). Uma fase nova é introduzida nesse espaço temporal entre as fases haplóide e diplóide: a fase dicariótica. Durante esta nova fase, os núcleos haplóides podem multiplicar-se por mitose sincronizada por um grande período de tempo.

No final, os dois núcleos se fundem e formam um zigoto, que rapidamente entra em meiose. Dois ciclos de vida podem existir (Meinhardt 1990; Hiscock 1999):

- Ciclo de vida haplo dicariótico: nesse ciclo a fase haplóide consiste de crescimento vegetativo (assexual) e de formação do corpo de frutificação (sexual). A fase dicariótica é restrita a uma curta vida dicariótica confinada ao zigoto, o qual sofre meiose para produzir esporos haplóides. Esse ciclo é típico de ascomicetos filamentosos.
- Ciclo de vida dicariótico: nesse ciclo os corpos vegetativos e reprodutivos existem como dicarióticos. A fase diplóide é confinada ao zigoto, o qual sofre meiose para produzir esporos e algumas células geradas por mitoses ambos haplóides. Esse ciclo é típico dos basidiomicetos.

1.3.1 O ciclo sexual dos basidiomicetos

Os basiodiósporos germinam e formam um micélio que pode ser inicialmente multinucleado. Este micélio torna-se monocariótico através da formação de septos dividindo o micélio em células uninucleadas, o qual é capaz de manter-se em crescimento vegetativo indefinido. Quando hifas encontram e se unem há formação do micélio dicariótico através da plasmogamia (fusão de hifas). Se as hifas forem compatíveis, há o desenvolvimento de grampos de conexão assegurando a distribuição de um núcleo de cada tipo às células-filhas, o qual constitui uma característica marcante dos basidiomicetos e que permite o desenvolvimento e a manutenção do estágio dicariótico por tempo indefinido.

Quando as condições ambientais são favoráveis para a reprodução, parte do micélio dicariótico sofre uma morfogênese complexa para formar o basidiocarpo (Kues *et al* 2000). Algumas células do basidiocarpo transformam-se em basídios dicarióticos que possuem dois núcleos

haplóides. Dentro destes basídios ocorre a cariogamia para formar um único núcleo diplóide, e posterior meiose para a formação de quatro basidiósporos (Rapper 1966; Alexopoulos 1996; Purdy 1996; Raven 1996; Casselton 1998; Hiscock 1999; Kues 2000; Casselton 2002) (Fig.2).

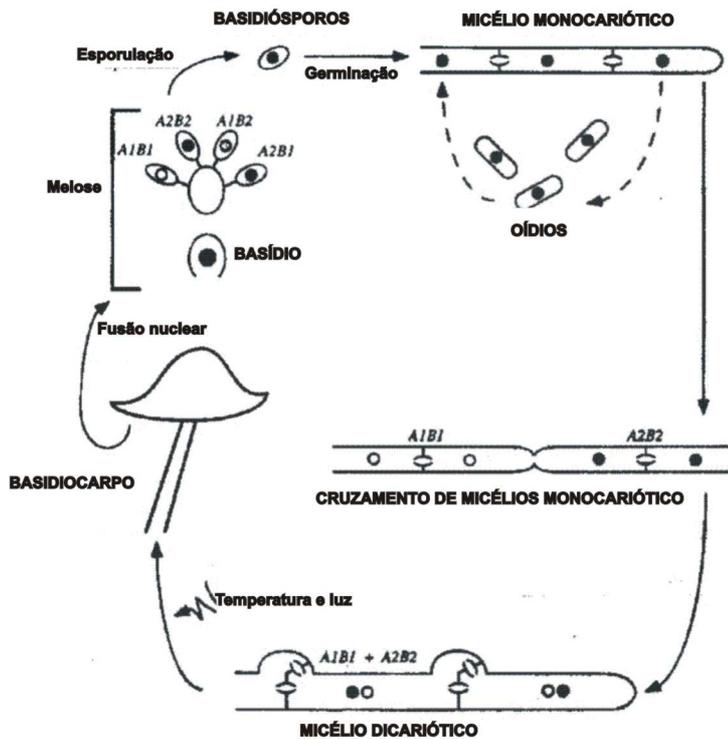


Figura 2. Ciclo de vida dos basidiomicetos (Casselton and Olesnicky 1998).

1.3.2 Heterotalismo e homotalismo

Heterotalismo é o termo designado às espécies que são auto estéreis e o termo homotalismo designa as espécies que são auto férteis (Rapper 1966; Alexopoulos 1996).

Em espécies de basidiomicetos heterotálicos (auto-estéreis) ocorre a plasmogamia ou somatogamia que consiste na união de duas hifas monocarióticas geneticamente diferentes onde o número haplóide de cada uma é igual a n . Esta fusão origina um micélio dicariótico com dois núcleos haplóides ($n+n$) geneticamente diferentes. Quando condições ambientais são favoráveis há a formação do corpo de frutificação seguido de cariogamia e meiose originando quatro basidiósporos

cada um contendo um núcleo geneticamente diferente (Rapper 1966; Alexopoulos 1996; Purdy 1996; Raven 1996; Kues 2000) (Fig. 3).

Em espécies homotálicas (auto-férteis), a formação do micélio dicariótico se produz a partir de um único micélio monocariótico em que ocorre plasmogamia entre suas hifas (Rapper 1966; Alexopoulos 1996; Purdy 1996). Este homotalismo também é chamado de homotalismo primário (Griffith 1994) (Fig. 3).

Segundo Alexopoulos, os fungos homotálicos não possuem sistema *mating type* (Alexopoulos 1996).

O homotalismo secundário, ou melhor, o pseudohomotalismo ocorre em alguns fungos heterotálicos. Estes fungos desenvolveram um interessante mecanismo para tornar-se auto-férteis. Após a meiose, dois núcleos geneticamente diferentes migram para um único basidiósporo. Este basidiósporo ao germinar forma um micélio binucleado auto-fértil não necessitando da união de hifas (Alexopoulos 1996; Purdy 1996; Raven 1996; Kues 2000) (Fig. 3).

Segundo Delgado, *C. pernicioso* é um fungo homotálico em que o micélio desenvolve-se a partir de um basidiósporo uninucleado levando a formação de um micélio primário monocariótico que evolui para um micélio dicariótico com a formação de grampos de conexão, sem a evidência de anastomose entre hifas (Delgado 1976).

Segundo Griffith, a preponderância de heterotalismo entre os basidiomicetos é considerada uma indicação de que o heterotalismo é a condição ancestral ao homotalismo neste grupo. Em seus estudos sobre os diferentes biótipos de *C. pernicioso* foi encontrado um biótipo heterotálico (biótipo L) sugerindo que o fungo causador de vassoura-de-bruxa em cacauero (biótipo C) homotálico primário evoluiu de um biótipo heterotálico (Griffith 1994).

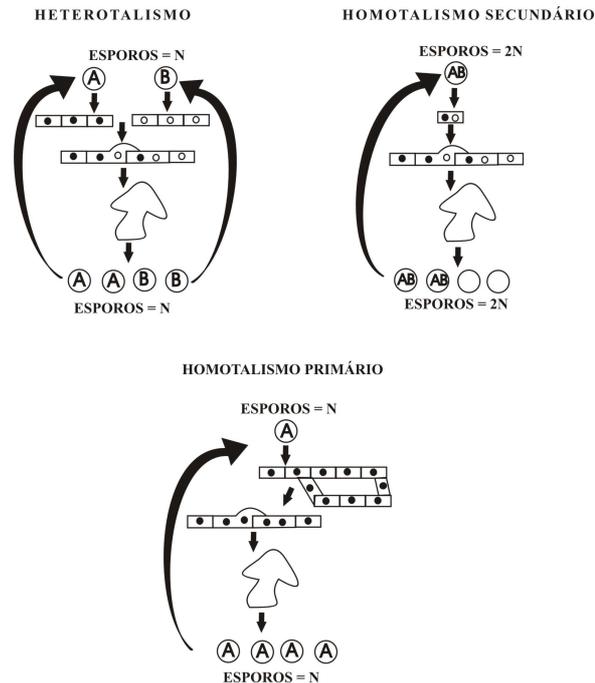


Figura 3. Heterotalismo, homotalismo primário e secundário.

Ciclo sexual de fungos basidiomicetos mostrando a estratégia de reprodução auto estéril (HETEROTALISMO) e as estratégias auto férteis (HOMOTALISMO SECUNDÁRIO E PRIMÁRIO). Em ambos são mostrados os esporos, micélios e corpos de frutificação. As setas indicam o caminho do ciclo. (Alexopoulos 1996; Kues 2000).

1.4 O sistema *mating type*

Os fungos ascomicetos e basidiomicetos têm contribuído muito para o entendimento da biologia celular eucariótica. O estudo do sistema *mating type* tem providenciado uma compreensão detalhada dos caminhos de sinalização celular e transcrição de genes específicos.

Os fungos em sua maioria não são diferencialmente separados em “machos” e “fêmeas” sendo morfologicamente indivíduos idênticos. A distinção como forma de evitar o auto cruzamento e promover o “*outbreeding*” é feita mediante mecanismos genéticos através de um sistema de reconhecimento molecular de incompatibilidade conhecido como *Sistema Mating Type* (Rapper 1966).

Os genes que determinam a incompatibilidade sexual em fungos estão localizados em um ou mais complexos *loci* conhecidos como *locus mating type*.

Há um número de diferenças importantes nos sistemas de incompatibilidade ocorrendo em vários grupos e esses sistemas podem ser distinguidos por dois critérios principais:

1. O número de *loci* no sistema . Bipolar e tetrapolar são termos comumente aceitos para sistemas com um *locus* e com dois *loci*, respectivamente, e denotam o número de progênies dos cruzamentos. Um cruzamento bipolar produz dois tipos de genes de incompatibilidade (*mating types*) e o cruzamento tetrapolar produz quatro.
2. O número de alelos em cada *locus*. Os termos bialélico e multialélico distinguem entre sistemas com dois alelos por *locus* e aqueles com vários alelos por *locus*.

Mating type em ascomicetos, zigomicetos e oomicetos são governados por um único *locus mating type* com duas formas distintas. A maioria dos basidiomicetos possuem múltiplos *mating types* determinados por um, dois ou três diferentes *loci* genéticos cada qual podendo ter duas ou mais formas diferentes (Rapper 1966; Hiscock 1999).

O reconhecimento de parceiros compatíveis envolve dois mecanismos genéticos interligados. O primeiro mecanismo a ser ativado opera através de ferormônios e receptores de ferormônios que desencadeiam um sinal que leva a ativação de segundo mecanismo e de genes regulatórios do sistema sexual. O segundo mecanismo envolve o reconhecimento e a interação molecular de dois homeodomínios de proteínas (HD1 e HD2) para gerar uma proteína heterodimérica com função de regulador da transcrição (Casselton 1998; Hiscock 1999; Kues 2000; Casselton 2002).

Para um melhor entendimento do mecanismo de ação desse sistema mostraremos primeiramente como o *sistema mating type* age na levedura ascomiceto *S. cerevisiae* e posteriormente como age em fungos basidiomicetos.

1.4.1 O Sistema *Mating type* em *Saccharomyces cerevisiae*

O sistema *mating type* mais amplamente estudado é o da levedura *S. cerevisiae* (Herskowitz 1988; Herskowitz 1989). Esse sistema é caracterizado por apenas um *locus* conhecido como *locus MAT* com duas formas diferentes, *MAT a* e *MAT alfa*, cada uma representada por dois genes que codificam proteínas do tipo homeodomínio (*a1* e *alfa2*), as quais regulam a reprodução sexual .

Em *S. cerevisiae* as células haplóides produzem e secretam feromônios, os quais devem ser reconhecidos por receptores específicos de células do tipo sexual oposto, desencadeando uma resposta que leva ao cruzamento entre as células e a geração do diplóide (Bender 1989; Herskowitz 1989; Jackson 1990; Kurjan 1992; Kurjan 1993).

Quando duas células de tipos opostos encontram-se a resposta gerada pela ligação do ferormônio ao respectivo receptor (*fator a* com o receptor STE3 e *fator alfa* com o receptor STE2) é transmitida internamente através de proteínas G heterotriméricas [subunidades: Gpa1 (alfa), STE4 (beta) e STE18 (gama)] acopladas ao receptor.

Ao receber o sinal, a proteína G heterotrimérica promove a dissociação de G-alfa das subunidades beta e gama. O dímero G-beta/gama dissociado gera um sinal transmitido via caminho quinase: STE11, MAPKK (STE7) e MAP quinase (FUS3 ou KSS1), o qual culmina regulando os fatores de transcrição STE12 (regula a expressão dos genes requeridos para *mating*.) e FAR1 (controla a parada do ciclo celular) (Kurjan 1992; Kurjan 1993; Bardwell *et al* 1994; Elion 1995; Herskowitz 1995; Banuett 1998)

O recente diplóide *a / alfa* reprime, através do STE12, a expressão dos genes requeridos para a manutenção da fase haplóide, entre eles os para a produção de ferormônio e sua recepção, inibindo assim nova fusão de células.

A fusão celular traz as proteínas *a1* e *alfa2* juntas e estes dois distintos homeodomínios interagem para dar uma nova função regulatória, a qual tem uma mudança na preferência de ligação ao DNA. Portanto, o heterodímero *a1 / alfa2* junto com outros fatores de transcrição reprimem a expressão de todos os genes haplóides específicos, incluindo o gene *rme1*, que codifica um repressor de transcrição de genes requeridos para meiose e para a esporulação (Herskowitz 1989; Kurjan 1992; Kurjan 1993; Bardwell *et al* 1994).

1.4.2 O Sistema *Mating type* de Basidiomicetos

1.4.2.1. Aspectos comuns

Em basidiomicetos heterotálicos o sistema *mating type* é responsável por promover *outbreeding*, manutenção do estágio dicariótico, a repressão da oidização e o início da formação do corpo de frutificação (Kertesz-Chaloupkova *et al* 1998; Kues 2000).

O sistema de incompatibilidade sexual em fungos basidiomicetos é muito complexo devido este ser regido por dois *loci mating type* (tetrapolar), cada um dos quais possuindo múltiplas versões (multialélico). Além dessa complexidade, a existência de variações na nomenclatura dos *loci* e dos genes dificulta o estudo e a compreensão desse sistema.

Portanto, para melhor entender o *sistema mating type* em basidiomicetos, neste trabalho os termos serão padronizados de acordo com a tabela 1.

Os fungos basidiomicetos mais estudados neste campo são: *Ustilago maydis*, *Schizophyllum commune*, *Coprinus cinereus* e *Coprinus bilanatus*. Em todos eles, o *sistema mating type* envolve dois mecanismos genéticos interligados. Esses mecanismos estão divididos em dois *loci*. Um *locus* é responsável pela codificação de duas proteínas do tipo homeodomínio e é denominado *locus HD*. O outro *locus* é responsável pelo reconhecimento entre receptores de ferormônios e ferormônios e é denominado *locus FRF*. Esses *loci* assemelham-se aos encontrados no sistema *mating type* de *S.*

cerevisiae, porém possuem uma organização molecular mais complexa e as proteínas sintetizadas desempenham um papel um pouco diferente, mesmo sendo homólogas às da levedura.

Um ponto importante é que os genes de ambos *loci* *HD* e *FRF* podem ter múltiplos alelos e a interação entre os seus produtos controla um complexo sistema de compatibilidade/incompatibilidade que promove o estágio dicariótico. (Rapper 1966; Puhalla 1970; Casselton 1998).

ORGANISMOS	MECANISMOS GENÉTICOS DE RECONHECIMENTO DO SISTEMA MATING TYPE					
	ATRAVÉS DE PROTEÍNA REGULATÓRIA			ATRAVÉS DE FERORMÔNIO E RECEPTOR		
	LOCUS	Gene para HD1	Gene para HD2	LOCUS	Gene para FERORMÔNIO	Gene para RECEPTOR
<i>Saccharomyces cerevisiae</i>	<i>MAT</i>	<i>MATα</i>	<i>MATα</i>		<i>Fator α e α</i>	<i>STE3, STE2</i>
<i>Ustilago maydis</i>	Mat b	<i>bE</i>	<i>BW</i>	Mat a	<i>Mfa1, Mfa2</i>	<i>Pra1, Pra2</i>
<i>Coprinus cinereus</i>	A	<i>a1 b1, d1</i>	<i>a2, b2, d2</i>	B	<i>phb1,2,3 (1,2)</i>	<i>rcb1,2,3</i>
<i>Schizophyllum commune</i>	A	<i>Z</i>	<i>Y</i>	B	<i>bap 1,2,3</i> <i>bbp 1,2,3</i>	<i>Bar</i> <i>Bbr</i>
<i>Coprinus bilanatus</i>	A	<i>a1</i>	<i>a2</i>			
Nomenclatura usada neste trabalho	HD	<i>Hd1</i>	<i>Hd2</i>	FRF	<i>Fe</i>	<i>Rc</i>

Tabela.1 – Tabela para referência de terminologia dos *loci mating type* em fungos basidiomicetos.

1.4.2.2 Locus *FRF*

O *locus FRF* é subdividido em dois sub *loci* *alfa* e *beta*. Cada *sub locus* é composto de um gene que codifica um receptor de ferormônio e um gene que codifica um precursor de ferormônio. A interação do ferormônio maduro ao receptor da célula oposta é responsável por desencadear um processo de sinalização celular que também regula genes relevantes para *mating* (Fig.4).

O modelo melhor entendido da cascata de sinalização entre ferormônios e receptores é o de *S. cerevisiae*, embora não sejam codificados por genes *mating type*, ferormônios e receptores similares são encontrados em basidiomicetos (Jackson 1990; Casselton 1998).

Os ferormônios são dodecâmeros carboximetilados e prenilados, uma mistura de lipopeptídeos de 12 aminoácidos que diferem em um único resíduo. Seus precursores consistem de aproximadamente 22 aminoácidos na seqüência N-terminal, de uma seqüência do ferormônio maduro e de um motivo C-terminal - Caax (onde C é uma cisteína, a é alifático e x é um aminoácido qualquer).

O processamento de maturação do ferormônio inicia com a farnesilação de resíduos de cisteína no motivo Caax pelo complexo farnesil transferase, seguido por clivagem (remoção) dos três aminoácidos terminais por proteases de prenilação. Em seguida acontece a carboximetilação pela carboxil metiltransferase, e finalmente a remoção da região precursora aminoterminal através de um processamento N-terminal que envolve dois passos proteolíticos. (Chen *et al* 1997; Casselton 1998; Hiscock 1999). Após estes estágios sucessivos de maturação, o ferormônio pode ser secretado por um caminho não clássico (Berkower 1991; Cadwell *et al* 1995) ou pelo caminho clássico de leveduras (Cadwell *et al* 1995; Casselton 1998).

Os receptores de ferormônios pertencem à família de receptores rodopsinas, os quais são compostos de sete hélices *transmembranas* hidrofóbicas (7-TM), de um domínio curto extracelular N-terminal, de três loops internos e três loops externos e de uma longa cauda carboxi-terminal intracelular (Nakayama *et al* 1985; Hagen *et al* 1986; Dohlman *et al* 1991; Vaillancourt *et al* 1997; Casselton 1998; O'Shea *et al* 1998; Hiscock 1999; Olesnicky *et al* 1999; Kues 2000; Olesnicky *et al* 2000) (Fig.5).

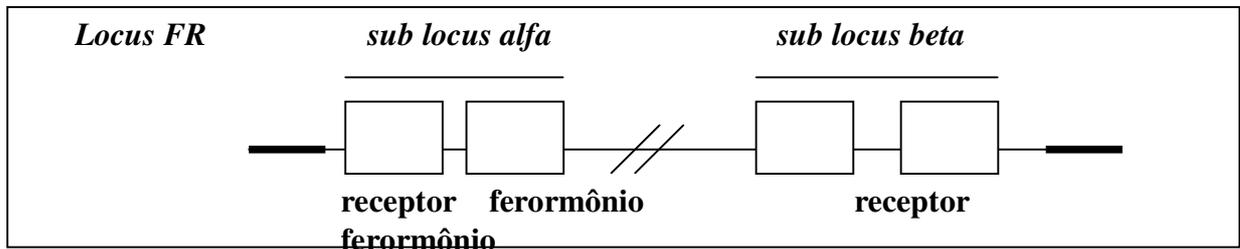


Figura 4. Ilustração do *Locus FRF*.

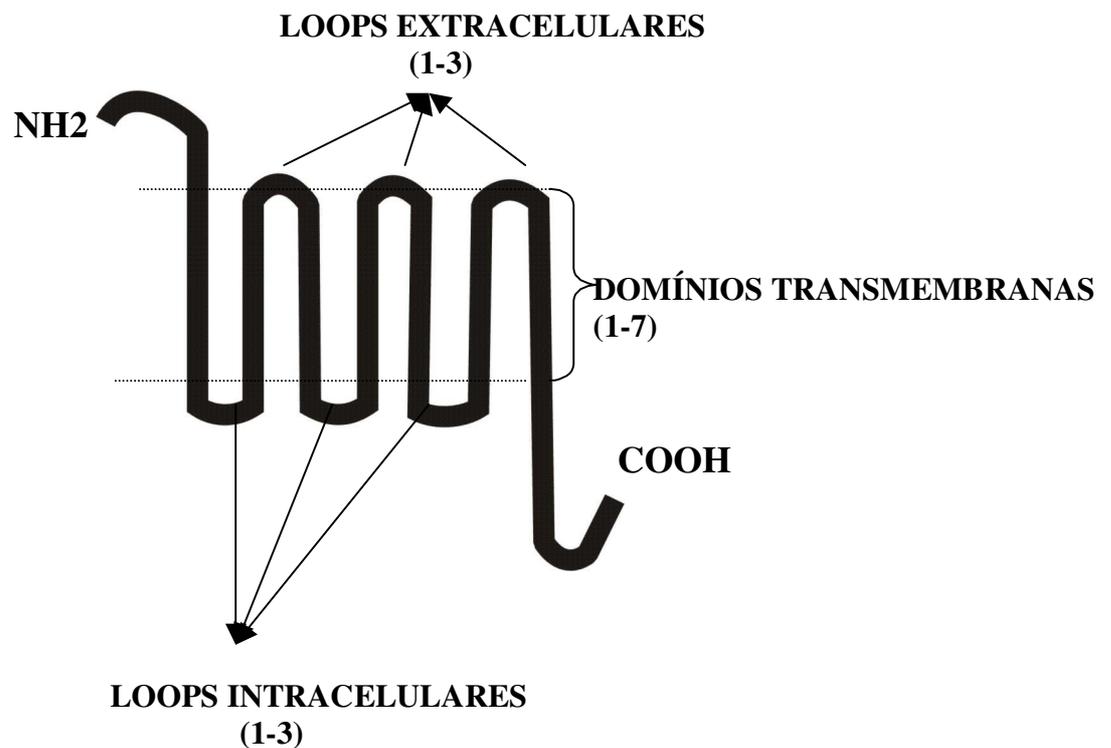


Figura 5. Modelo da estrutura do receptor de ferormônio (7-TM). Estão representados o curto domínio amino terminal (NH₂), as sete alfa hélices - domínios transmembranas, os três loops intracelulares, os três loops extracelulares e a longa cauda carboxi terminal intracelular (Casselton and Olesnicky 1998).

1.4.2.2 Locus HD

O locus HD é também sub-dividido em dois *sub loci* denominados *alfa* e *beta*. Cada *sub locus* contém dois genes que codificam ambas proteínas regulatórias HD1 e HD2 (Fig.6). Essas proteínas têm a função de juntas formarem uma proteína regulatória heterodimérica, conhecida como heterodímero, a qual deve regular a transcrição de genes relevantes para promover a dicarionização.

A proteína HD2 é um típico homeodomínio de 60 aa formado por três hélices contendo uma seqüência característica e altamente conservada, a WF-N-R, dentro da terceira hélice (hélice de reconhecimento) (Gehring *et al* 1994; Burglin 1997). O homeodomínio da classe HD1 é considerado atípico devido a um desvio da seqüência consenso WF-N-R da hélice de reconhecimento e por ter aminoácidos extras entre as hélices 1 e 2 (Burglin 1997) e diferentemente de HD2 a proteína HD1 possui sinais de localização nuclear clássicos conhecidos como NLS1 e NLS2 (Scott *et al* 1989; Kronstad 1990; Tymon *et al* 1992) (Fig.7).

Em basidiomicetos, várias proteínas HD1 e HD2 são produzidas e combinações protéicas compatíveis derivadas de diferentes haplótipos precisam ser discriminadas das interações incompatíveis dentro do mesmo haplótipo. O domínio responsável pela especificidade desta discriminação está localizado em uma região altamente variável localizada na porção N-terminal do homeodomínio (Scott *et al* 1989; Casselton 1998; Hiscock 1999; Kues 2000).

Tem sido mostrado que a substituição de resíduos dentro do domínio de reconhecimento de proteínas HD incompatíveis, afeta especificamente a proteína tornando-a compatível (Yee 1993; Banham *et al* 1995; Kämper *et al* 1995; Yue *et al* 1997).

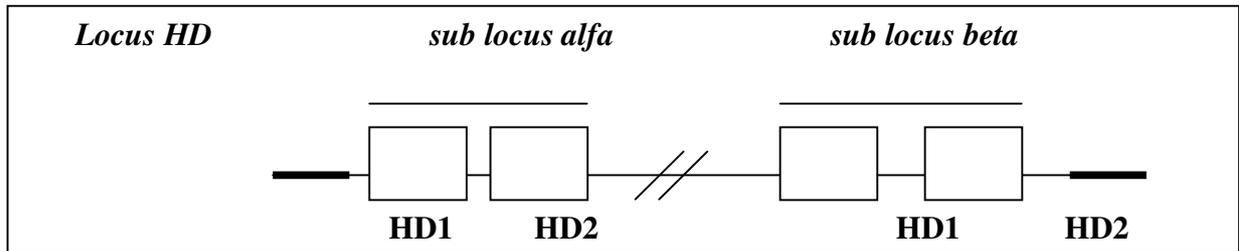


Figura 6. Ilustração do *Locus HD*

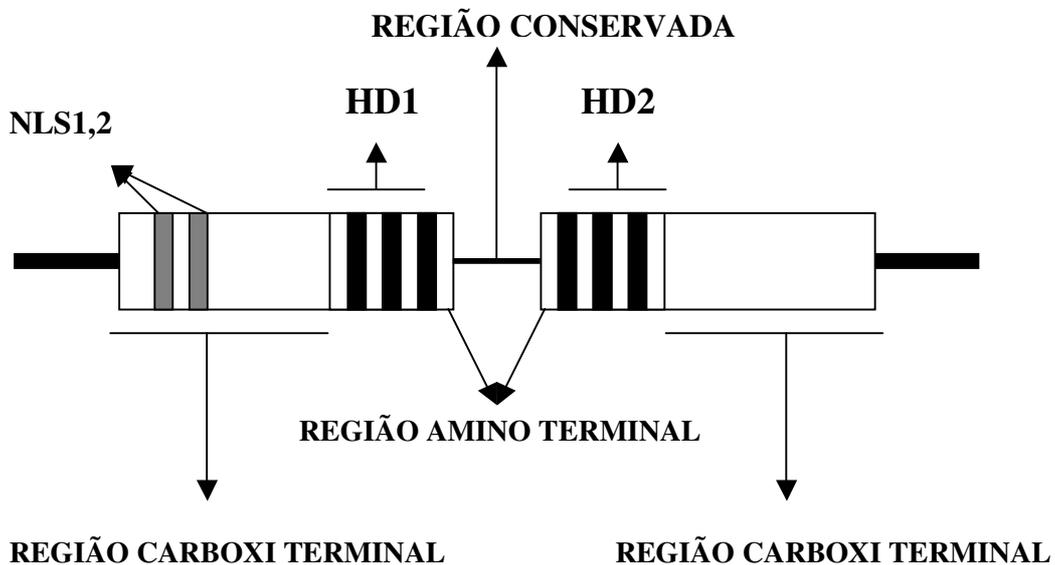


Figura 7. Estrutura dos genes HD1 e HD2.

As regiões dos homeodomínios HD1 e HD2, as regiões amino terminais e as regiões carboxi terminais estão representadas. A região entre as duas proteínas é conservada. Retângulos pretos representam as três hélices do homeodomínio. Retângulo cinza representa as regiões NLS1,2 (Casselton 1998; Hiscock 1999; Kues 2000).

Nenhuma das duas proteínas isoladamente é capaz de desempenhar o papel de fator de transcrição, o qual adquirem após a dimerização. A proteína HD1, diferentemente da proteína HD2, não possui a hélice de reconhecimento ao DNA. Portanto, é a proteína HD2 que desempenha esse papel. A interação do homeodomínio com o DNA tem sido determinada por ressonância magnética nuclear e cristalografia de raioX (Gehring *et al* 1994). Porém, como essas proteínas são dimerizadas

no citoplasma e precisam entrar no núcleo para fazerem seu trabalho, a proteína HD1 acaba tendo um papel essencial no reconhecimento molecular, pois contém um sinal de localização nuclear NLS1 e NLS2 (Scott *et al* 1989; Kronstad 1990; Gillissen *et al* 1992; Tymon *et al* 1992; Asante *et al* 1996; Spit *et al* 1998). Desta forma, a proteína HD2 não pode entrar no núcleo sem primeiro associar-se à proteína HD1, e embora a proteína HD1 possa entrar no núcleo sem associar-se a HD2 não há o reconhecimento ao sítio alvo do DNA (Casselton 1998).

Interessantemente, tem sido encontrado dentro da região carboxi terminal um domínio similar a cauda de 20 aminoácidos de *alfa2* de *S. cerevisiae*, o qual tem a função de promover a interação entre as proteínas HD1 e HD2 (Dranginis 1990; Kues *et al* 1994).

1.4.2.4 Coordenação das atividades reguladas pelos loci *FRF* e *HD*

Na maioria dos basidiomicetos, o micélio dicariótico surge da fusão de dois micélios monocarióticos compatíveis, um evento coordenado pelo *Sistema Mating Type* (Rapper 1966; Kues *et al* 2001; Casselton 2002) (Fig.8). Esse processo tem sido estudado com detalhes em *Coprinus cinereus* e *Ustilago maydis*, embora ainda existam muitas dúvidas sobre o seu funcionamento (Bolker 2001).

Em *C. cinereus*, hifas monocarióticas, encontram e se fundem aparentemente sem a necessidade de reconhecimento extra celular por ferormônios (Kothe 1996; Casselton 1998).

Após a fusão celular, os ferormônios que são constitutivamente produzidos nestas células, encontram os respectivos receptores e, se houver compatibilidade, é desencadeado um processo de cascata de sinalização, similar ao da levedura *S. cerevisiae*, culminando na mitose dos núcleos doadores e na dissolução dos septos. O núcleo doador entra na hifa receptora e migra até alcançar o seu topo (Kues 2000).

Durante esse percurso, a sinalização desencadeada pelo *locus FRF* ativa a produção das proteínas regulatórias do *locus HD* (Bolker 2001) liberando-as no citoplasma. Se a proteína HD1 ou HD2 do núcleo doador for compatível com a proteína HD2 ou HD1 do núcleo receptor, respectivamente, é formado, no citoplasma, uma proteína heterodimérica (HD1HD2 ou HD2HD1) (Magaie 1995) que entra no núcleo receptor e promove a regulação de genes envolvidos no processo de dicariose (Kertesz-Chaloupkova *et al* 1998; Bolker 2001) (Fig.9).

O heterodímero formado promove o pareamento dos núcleos e a divisão mitótica sincronizada, simultaneamente com a formação de especializados grampos de conexão na posição onde um novo septo irá aparecer. Durante a prófase mitótica o núcleo doador entra no grampo de conexão e o núcleo receptor permanece na célula adjacente ao grampo. Em anáfase os núcleos dividem-se sendo que o núcleo doador mantém um curto fuso com seu núcleo filho e o núcleo receptor mantém um longo fuso com seu núcleo filho. Septos são formados dividindo os núcleos em três células: o núcleo doador permanece no grampo de conexão, o núcleo receptor permanece na célula adjacente ao grampo e os dois núcleos filhos ficam juntos no topo hifal.

Com o grampo de conexão formado, o *locus FRF*, promove a fusão do grampo à célula adjacente formando uma só célula que terá um núcleo de cada tipo. Este mecanismo coordenado de formação do dicário assegura que cada célula dicariótica terá um núcleo de cada tipo (Iwasa *et al* 1998; Kues 2000) (Fig.8).

A manutenção do estágio dicariótico pode permanecer por longos períodos. O corpo de frutificação só é produzido quando as condições ambientais são favoráveis, como umidade e luz. O *sistema mating type* aparentemente tem participação apenas na regulação do início da formação do corpo de frutificação, mas pouco se conhece sobre o seu papel nesta etapa (Tymon *et al* 1992; Kues *et al* 1994; Kues *et al* 1998). Quando os corpos de frutificação são formados, rapidamente ocorre a fusão nuclear (cariogamia) seguida da meiose em células especiais chamadas basídios. Da meiose surgem os basidiósporos (esporos) iniciando um novo ciclo de vida.

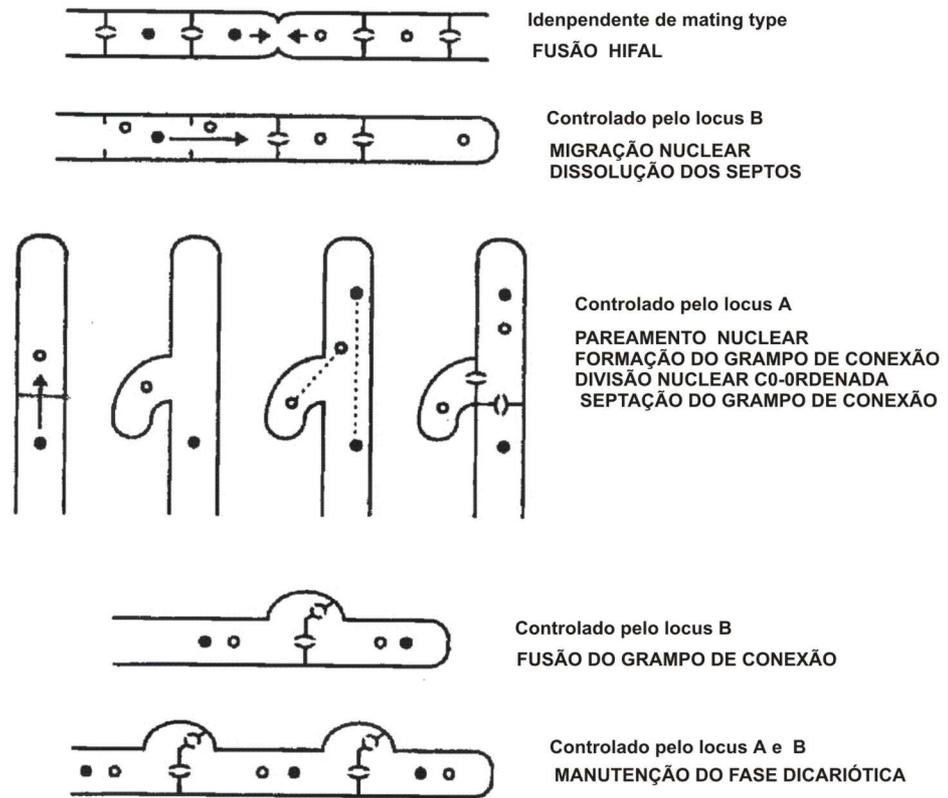


Figura 8. Atividade coordenada dos genes *mating type* na formação e manutenção da fase dicariótica (Kues 2000).

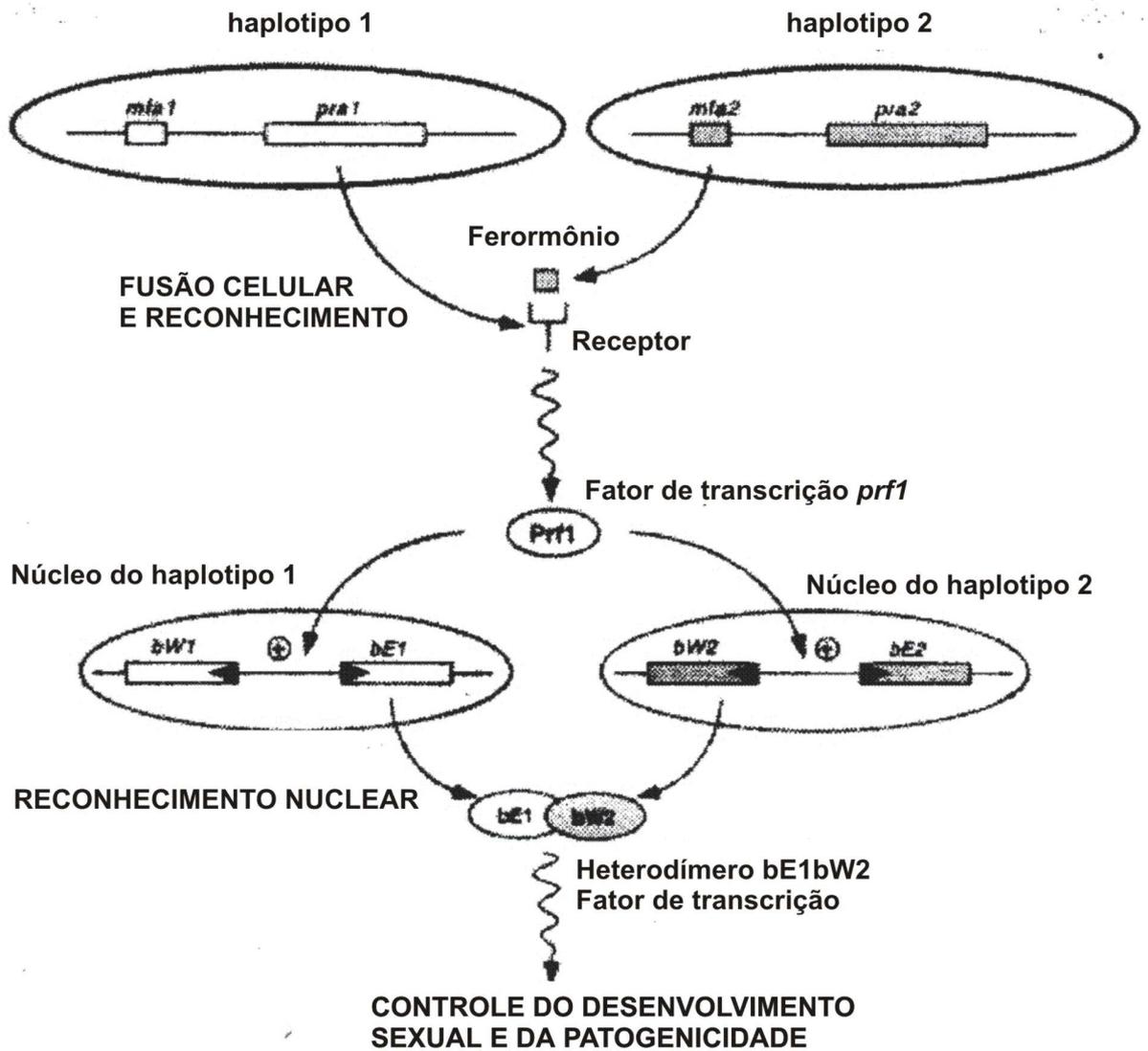


Figura 9. Interação entre os genes *mating type* de *Ustilago maydis* (Bolker 2001).

1.4.2.5 Características do Sistema Mating Type de *Ustilago maydis*

No *U. maydis*, cada locus b (locus HD) possui dois genes denominados de *bE* (HD1) e *bW* (HD2), os quais são transcritos diferentemente e inseparáveis por eventos de recombinação, permanecendo juntos como uma unidade genética. Os genes *bE* codificam a proteína HD1 de 473 aminoácidos, e os genes *bW* codificam a proteína HD2 de 644 aminoácidos (Schulz *et al* 1990; Kues *et al* 1994). *bE* e *bW* têm reduzida similaridade em sequência, sendo essa limitada à região conservada do homeodomínio e nos domínios carboxiterminais. As regiões aminoterminal destas proteínas são altamente variáveis (Schulz *et al* 1990; Gillissen *et al* 1992).

Quando dois haplótipos compatíveis fundem-se, por exemplo, os haplótipos 1 e 2 (Fig.10), quatro proteínas diferentes estão presentes na fusão celular: *bE1*, *bE2*, *bW1*, *bW2*. Apesar disso somente duas combinações protéicas são possíveis: *bE1* X *bW2* e *bE2* X *bW1*, pois as proteínas codificadas por um mesmo locus são sempre incompatíveis.

U. maydis como *S. cerevisiae*, possuem somente duas versões de genes que codificam ferormônios (*mfa1* e *mfa2*) e duas de receptores (*pra1* e *pra2*). Estes genes estão localizados no locus a (locus *FRF*) e distribuídos em dois sub loci denominados de *a1* e *a2* (Fig.11).

No sub locus *a1* estão localizados os genes *mfa1* e *pra1*. No sub locus *a2* estão localizados os genes *mfa2* e *pra2*, sendo que entre eles estão localizados dois genes, *lga2* e *rga2*, que são ativados pelo caminho de resposta ao ferormônio, mas cujas funções são desconhecidas (Bölker *et al* 1992; Urban *et al* 1996).

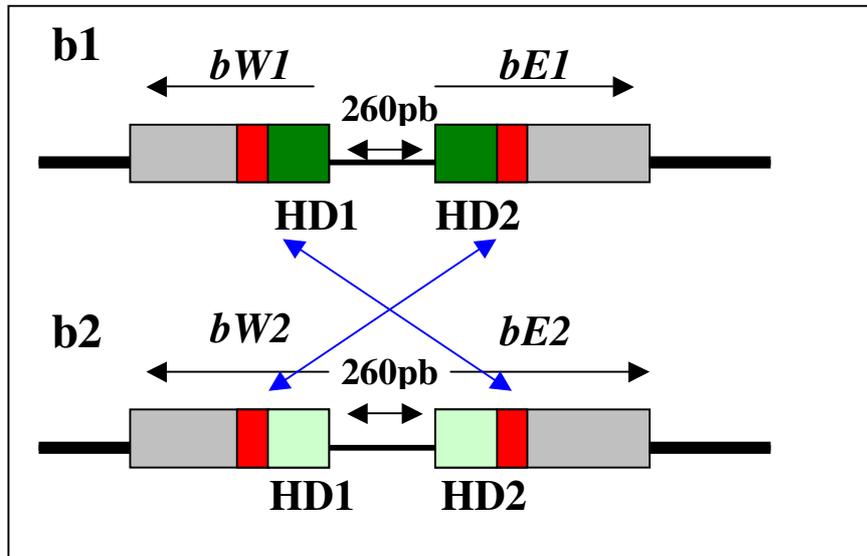


Figura 10. Organização molecular do locus *mating type b* de *U. maydis*. O locus multialélico *b* contém dois genes transcritos diferentemente, *bE* e *bW*, os quais codificam polipeptídeos de 437 e 644 aminoácidos respectivamente. A sequência da região aminoterminal ao homeodomínio das duas proteínas e o intervalo de 260pb da região promotora são diferentes para cada *b*. Os loci dos haplótipos *b1* e *b2* estão ilustrados, e a sequência variável de DNA, a região aminoterminal, específica para cada, está representada por caixas verde claro e verde escuro, respectivamente. Caixas vermelhas indicam as sequências que codificam homeodomínios. Caixas cinzas representam a região carboxiterminal. As setas azuis representam as duas possíveis combinações para formar o heterodímero. (Casselton 1998).

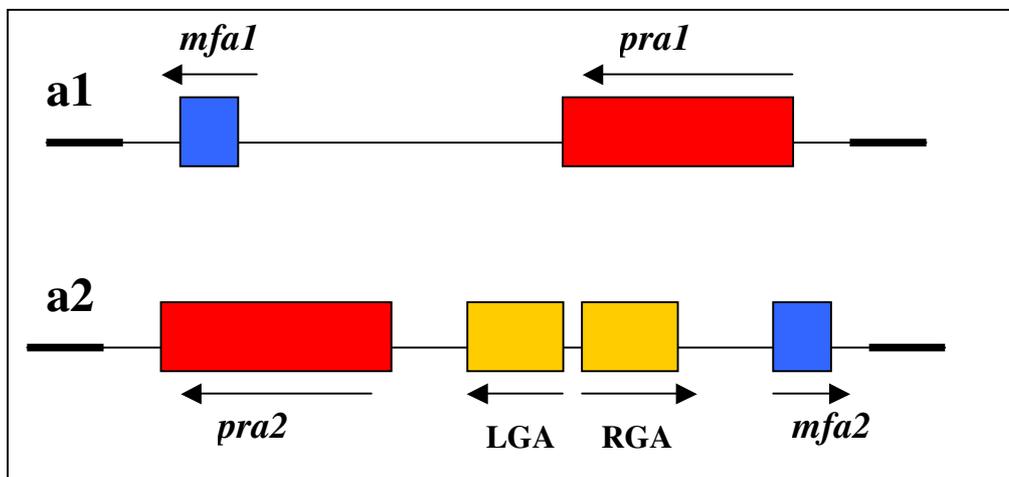


Figura 11. Organização molecular do locus *mating type a* de *U. maydis*. O subloci *a1* tem aproximadamente 4,5kb, contém dois genes *mfa1* (codifica para ferormônio) e *pra1* (codifica para receptor). O locus *a2*, de aproximadamente 8 kb, contém dois genes *mfa2* (codifica para ferormônio), *pra2* (codifica para receptor) e dois genes com funções desconhecidas, *lga2* e *rga2*. Flechas indicam a direção da transcrição dos genes e linhas finas indicam regiões flanqueadas de homologia e bordadas por sequências específicas de *mating type*.(Casselton 1998)

1.4.2.6 Características do Sistema Mating Type de *Schizophyllum commune*

O locus A (*locus HD*) de *Schizophyllum commune* é dividido em dois *sub loci* localizados no mesmo cromossomo denominados de *A alfa* e *A beta* (Papazian 1951; Day 1960; Raper 1960). Cada *sub locus* contém um par de genes funcionalmente redundantes que codificam as proteínas HD1 e HD2. No *sub locus A alfa* a proteína HD1 (890 aa) é codificada pelo gene Z (930 aa) e a proteína HD2 pelo gene Y. O *sub locus beta* está parcialmente caracterizado com a identificação de apenas o gene V, não estando definido se esse codificaria a proteína HD1 ou HD2 (Stankis *et al* 1992; Shen *et al* 1996).

Assim como em *U. maydis*, apenas um par destes dois *sub loci* precisam ser diferentes entre parceiros para ter uma interação compatível e promover a heterodimerização de HD1 e HD2.

Além dos genes envolvidos em *mating*, o no intervalo que compreende o locus A encontram-se outros três genes que aparentemente não participam do processo. No *sub locus A alfa* estão localizados o gene X, de função desconhecida (Marion *et al* 1996), e o gene *mip*, que codifica uma peptidase intermediária mitocondrial (Casselton *et al* 1995; Isaya *et al* 1995). Entre os dois *sub-loci* encontra-se o gene *pab-1*, o qual é necessário para a produção de ácido para-aminobenzóico (paba), e o gene *ade-5* (aminoimidazol ribonucleotídeo), que participa da síntese da adenina (Rapper 1966; Giasson *et al* 1989; Mutasa *et al* 1990; Kues *et al* 2001) (Fig.12).

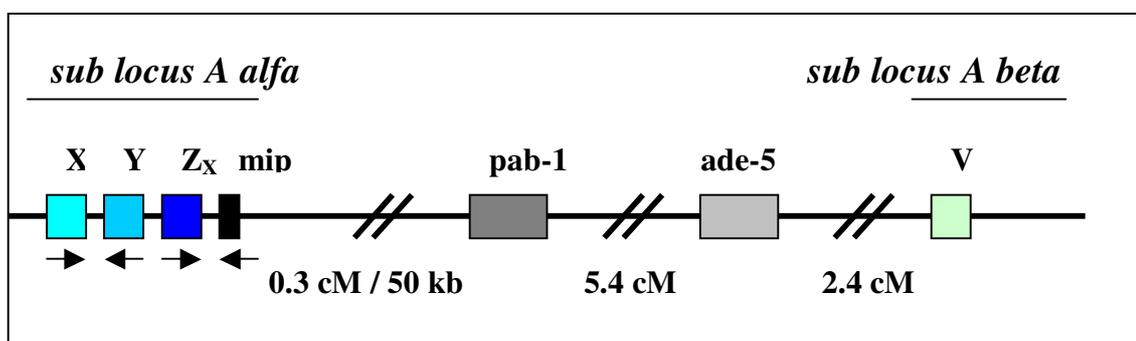


Figura 12. Organização molecular do locus mating type A de *S. commune*. O *sub locus A alfa* de *S.* contém um único par de genes multialélicos diferentemente transcritos HD1 e HD2 designados de Z e Y, respectivamente (Casselton 1998).

Em *S. commune*, o locus *B* (*FRF*) está separado em dois *sub loci*: *sub locus B alfa* e *sub locus B beta* (Parag 1971). Baseado nas seqüências dos *sub loci Balfa1* e *Bbeta1*, o locus *B* parece conter dois genes para receptor de ferormônio, *Rc*, (*bar*, *bbr*) e seis genes para ferormônios, *Fe*, [*bap*(1,2,3), *bbp*(1,2,3)]. Os *sub loci Balfa* e *Bbeta* representam duas subfamílias independentes de genes e os *sets* de genes são cassetes funcionais, os quais permanecem intactos (Fig. 13).

Os ferormônios codificados pelos genes do *sub locus Balfa* - *bap* (1,2,3) - podem apenas interagir com receptores do próprio sub-loci *Balfa* - (*bar*), mas de haplótipos compatíveis. Ele não interage nem com o receptor codificado pelo gene do seu próprio sub-locus nem com nenhum receptor codificado por genes do sub-loci *Bbeta* - (*bbr*). O mesmo ocorre com os ferormônios do *sub locus Bbeta* (Casselton 1998).

Assim sendo, em um cruzamento entre dois haplótipos de *S. commune* são gerados doze possíveis combinações, sendo que dessas apenas uma é necessária para o sucesso do cruzamento (Fig. 14).

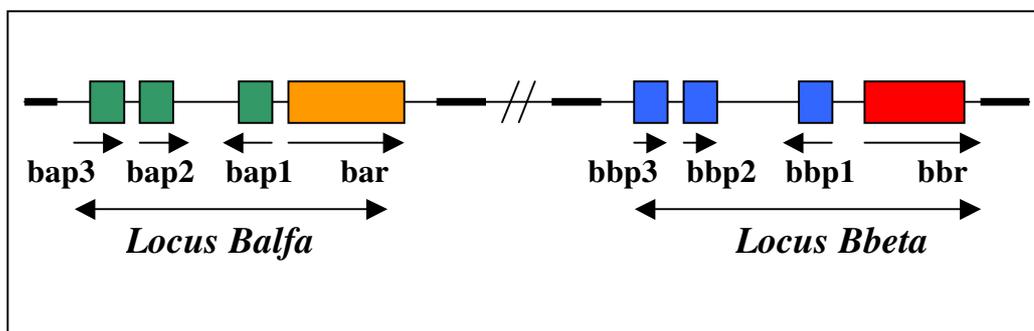


Figura 13 Organização molecular do locus *B* de *S. commune*. Cada *sub locus* tem aproximadamente 8kb e possui quatro genes. O *sub locus Balfa* possui três genes *bap3*, *bap2* e *bap1* (caixas verdes) que codificam ferormônios e um gene *bar* (caixa laranja) que codifica o receptor. O *sub locus Bbeta* também possui três genes *bbp3*, *bbp2* e *bbp1* (caixas azuis) que codificam ferormônios e um gene *bbr* (caixa vermelha) que codifica o receptor. Os *sub loci* estão separados por aproximadamente 3.5 unidades de mapeamento (Casselton 1998).

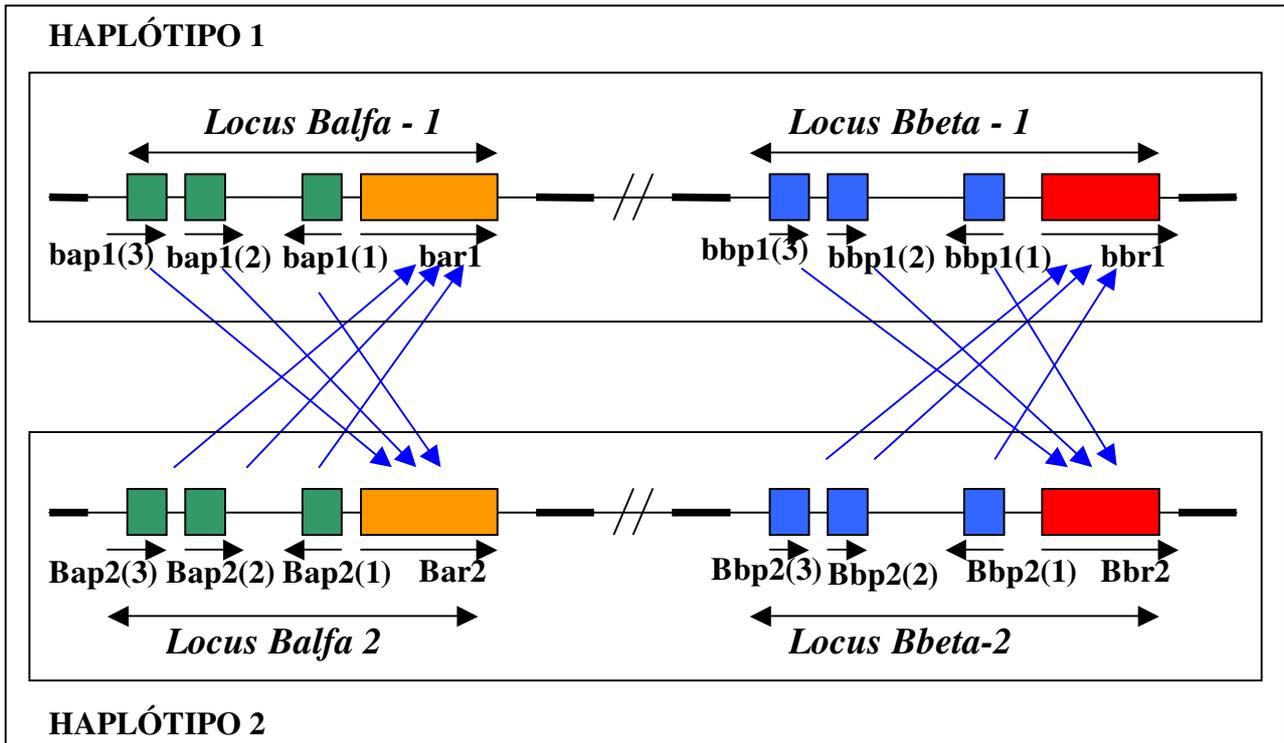


Figura 14 Cruzamento entre dois haplótipo de *S. commune*. Dois *loci B* são mostrados, um do organismo 1 e outro do organismo 2. As 12 setas azuis indicam cada uma das possíveis interações entre ferormônios e os seu receptores compatíveis (Casselton 1998).

1.4.2.7 Características do Sistema Mating Type de *Coprinus cinereus*

Da mesma forma que em *S. commune*, em *Coprinus cinereus* foram identificados no *locus A* (*HD*) dois *sub loci*, o *sub locus A alfa* e o *sub locus A beta*. Sua organização molecular é mais conhecida devido ao fato desse *sub-loci* estarem muito próximos, o que permitiu o seu isolamento molecular como um todo (Mutasa *et al* 1990; May *et al* 1991; Pardo *et al* 1996).

O *sub locus A alfa* contém um par de genes, *a1* e *a2*, que codificam as proteínas HD1 e HD2, respectivamente. O *sub locus A beta* contém dois pares de genes, *b1* e *b2* e *d1* e *d2*, que também codificam pares das proteínas HD1 e HD2, respectivamente (Kues 1992; Kues 1993; Lukens 1996). O *sub locus A alfa* está separado do *sub locus A beta* por aproximadamente 7.0 kb da seqüência de DNA (equivalente a 0.07 unidades do mapeamento) (Fig. 15).

Devido à existência de muitos genes funcionalmente redundantes, é raro encontrar todos os genes presentes em um único *locus*. De todos os haplótipos até então analisados, em apenas um, o haplótipo A44, foram encontrados todos os seis genes (Pardo *et al* 1996). No haplótipo A6 um dos genes *a* (*a1*) e um dos genes *d* (*d2*) foram perdidos, enquanto que no haplótipo A5 somente um dos genes *d* (*d1*) está ausente.

Como em *U. maydis*, apenas uma combinação compatível de HD1 e HD2 é necessária e suficiente para promover o desenvolvimento sexual (Pardo *et al* 1996). Entretanto, estas combinações devem vir de alelos de um mesmo subgrupo, aqui considerados *a*, *b* e *d*, de um haplótipo compatível. Por exemplo, o produto do gene *a1* (HD1) trabalhará apenas com genes *a2* (HD2) do haplótipo compatível. Ele não será capaz de interagir com *b2* ou *d2* dele mesmo ou do haplótipo compatível.

Como em *S. commune*, há também em *C. cinereus* genes não ligados a *mating*, mas que flanqueiam o *locus A*. Um destes é o gene *mip* que codifica uma peptidase intermediária mitocondrial uma metalo-endopeptidase (Casselton *et al* 1995; Isaya *et al* 1995), que flanqueia o *sub locus Aalfa*. Flanqueando o *sub locus Abeta* tem o gene *beta-fg*, um curto gene conservado de função desconhecida (Kues 1992; Kues *et al* 1994), e o gene *ade-8* (Day 1960; Lukens 1996). Outro gene é o *pab-1*, gene necessário para a produção de ácido para-aminobenzóico (*paba*), que flanqueia entre o *sub loci Aalfa* e *Abeta* (Rapper 1966; Giasson *et al* 1989; Mutasa *et al* 1990). A distancia entre *mip* e *pab-1* é de 40kb (Mutasa *et al* 1990; Kues 1992).

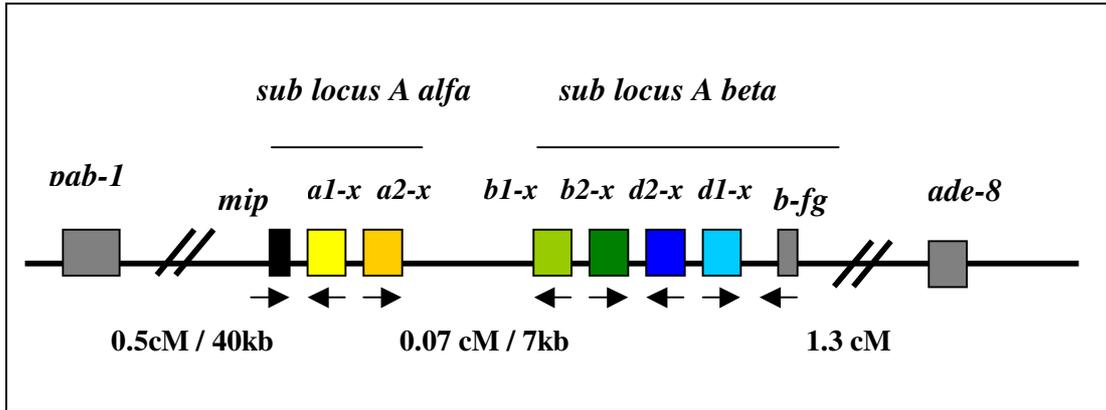


Figura 15: Organização molecular do *locus mating type A* de *C. cinereus*. O *locus A* contém três pares de genes multialélicos transcritos diferentemente (a, b e d). Cada par destes genes codificam duas proteínas diferentes com motivos diferentes de homeodomínios designados HD1 e HD2. Os genes *a* são representados por caixas amarelas, os genes *b* são representados por caixas verdes e os genes *d* são representados por caixas azuis (Casselton 1998).

O *locus B* (*locus FRF*) possui três *sets* de genes funcionalmente independentes representados como subfamílias 1, 2 e 3. Cada uma destas subfamílias contém dois genes para ferormônios (*phb1-1*, *phb1-2*, *phb2-1*, *phb2-2*, *phb3-1* e *phb3-2*) e um gene para receptor de ferormônio (*rcb1*, *rcb2* e *rcb3*) (Casselton 1998) (Fig. 16).

Dois organismos precisam formar um único par de receptor e ferormônio dentro de uma mesma subfamília para ser compatível. Os ferormônios de uma única família somente podem estimular todos os receptores dentro da mesma subfamília, exceto é claro, ele próprio.

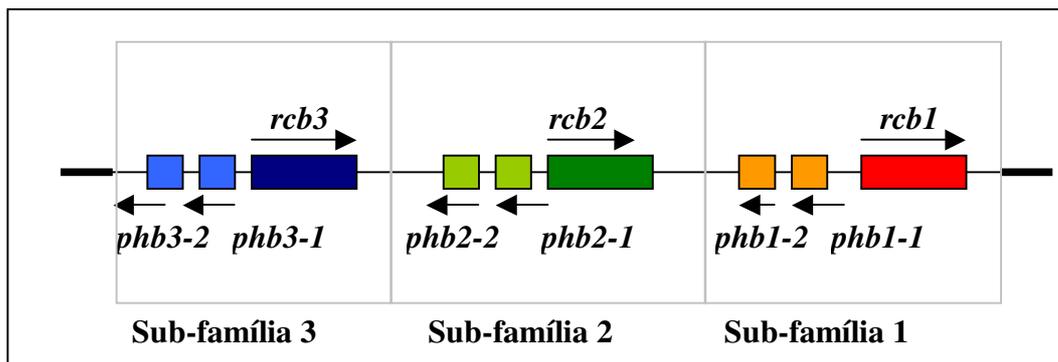


Figura 16: Organização molecular do *locus B* de *C. cinereus*. O *locus B* tem aproximadamente 17 kb e contém três subfamílias de genes funcionalmente independentes delimitados por linhas cinzas imaginárias. Os genes que codificam ferormônios (*phb*) estão representados por caixas pequenas e os genes que codificam receptores (*rcb*) por caixas grandes. As subfamílias estão representadas com cores diferentes: azul, verde e vermelho / subfamílias 3, 2 e 1, respectivamente. (Casselton 1998)

1.4.2.8 Características do Sistema Mating Type de *Coprinus bilnatus*

Coprinus bilnatus é um homobasidiomiceto homotático secundário ao contrário de *C. cinereus* e *S. commune* que são heterotáticos. É o terceiro homobasidiomiceto e o primeiro homotático secundário a ter o locus A clonado (locus HD).

Igualmente à *C. cinereus* e *S. commune*, o locus A é dividido em dois sub loci A alfa e A beta. O sub locus A alfa contém os genes *a1* e *a2* que codificam as proteínas HD1 e HD2, respectivamente. Embora ainda não tenha sido provado, provavelmente o sub locus A beta contém também dois genes que codificam as proteínas HD1 e HD2 (*b1* e *b2*, respectivamente).

Um gene que codifica para uma peptidase intermediária mitocondrial (*mip*) flanqueia diretamente o sub locus A alfa. O gene *pab-1*, requerido para a síntese do ácido para-aminobenzóico está a 39kb upstream de *mip*. O arranjo estrutural desta região cromossomal assemelha-se ao heterotático *C. cinereus* (Kues *et al* 2001) (Fig.17). O sub locus FRF ainda não foi caracterizado.

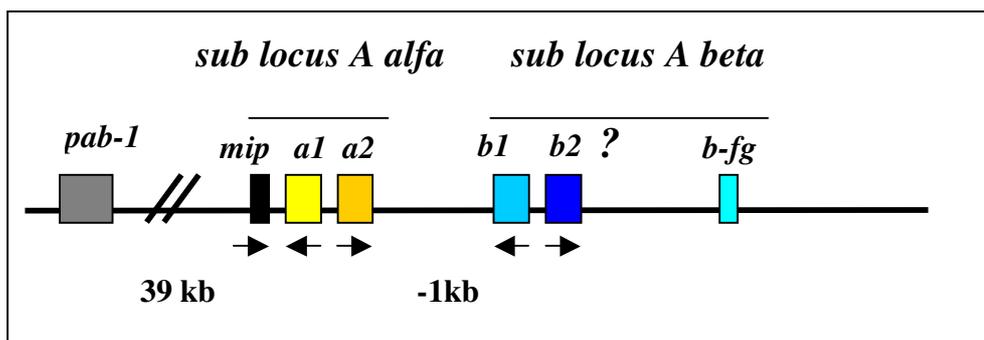


Fig.17: Organização molecular do locus mating type A de *C. bilnatus*. O locus A contém dois pares de genes multialélicos transcritos diferentemente (*a*, *b*). Cada par destes genes codificam duas proteínas dissimilares com diferentes motivos de homeodomínios designados HD1 e HD2. Os genes *a* são representados por caixas amarelas, *b* são representados por caixas azuis (Kues *et al* 2001).

1.5 A Genômica e suas Contribuições

A genômica é o estudo do genoma de um organismo, que é o conjunto de genes que codificam o organismo.

A genômica trouxe uma nova visão para o campo científico. A identidade dos organismos, que antigamente eram baseados em suposições, hoje com o advento da tecnologia é possível estudar seus mais íntimos aspectos.

A bioinformática, que é a aplicação da tecnologia de informação e gerenciamento de dados à área da biologia molecular (Gibas 2001), foi o ponto chave para o início e avanço da genômica. As ferramentas computacionais permitiram o aumento da velocidade e qualidade das informações geradas reduzindo tempo e erros.

Procurar similaridade por comparação de seqüências é uma das principais técnicas usadas pelos biólogos computacionais para inferir seus trabalhos. A mais popular ferramenta para este propósito é o BLAST (*Basic Local Alignment Search Tool*). (Altschul *et al* 1997; Pertsemlidis 2001).

O programa BLAST realiza comparações entre pares de seqüências procurando por regiões de similaridade pelo emprego de algoritmos heurísticos. Essa ferramenta pode produzir muitos alinhamentos. Para saber qual desses alinhamentos é significativo deve ser analisado um valor de expectativa, o famoso *e-value*. *E-value* é o valor que indica a probabilidade de um acaso. Portanto, quanto menor for o valor do *e-value* menor é a chance daquele alinhamento ter ocorrido ao acaso, representando grande significância biológica (Pertsemlidis 2001).

No Brasil a genômica explodiu como ciência de ponta com o Projeto Genoma da *Xylella fastidiosa* (Simpson *et al* 2000) e hoje temos vários projetos genomas em andamento no país. Contudo, embora a genômica tenha sido uma das grandes descobertas do século, apenas o estudo

das seqüências não é suficiente para o entendimento da programação genética e, portanto, para a compreensão do funcionamento do organismo como um todo. A pós-genômica, com o estudo comparativo dos genomas, a análise global da expressão gênica e o estudo das proteínas (proteoma) surgiu para fechar complementar as informações das seqüências. O seqüenciamento de um genoma pode ser feito de duas formas diferentes: através do DNA total, ou genômico, ou pelo mRNA (seqüências codantes). No caso de DNA total a técnica mais usada é conhecida como *shotgun* (Venter *et al* 1998), que consiste na quebra aleatória do genoma em inúmeros fragmentos pequenos possíveis de serem seqüenciados e que são posteriormente montados como um quebra cabeça pela bioinformática através de vários programas computacionais (Phred/Phrap/Consed) (Ewing *et al*¹ 1998; Ewing *et al*² 1998).

A outra forma para fazer genoma é via o mRNA. Esta forma é uma boa alternativa para genomas grandes porque só é seqüenciado a parte codante, mas por outro lado raramente é possível identificar todos os genes do organismo. Neste caso a estratégia desenvolvida que tem obtido grande sucesso é a técnica brasileira chamada ORESTES (Neto *et al* 2000).

Após o seqüenciamento, uma etapa essencial para a finalização do projeto genoma é a anotação das seqüências geradas. Anotar em termos biológicos é dar uma função possível para o gene. Inicialmente as seqüências geradas são comparadas às seqüências depositadas nos bancos de dados públicos, como o *Genbank* (Benson *et al* 2002), que possui grande número de seqüências de genes anotadas. As comparações são feitas através de programas computacionais específicos como o BLAST (Altschul *et al* 1997), e a partir do nível de similaridade encontrada pode-se traçar a relação evolutiva entre genes, ou mesmo organismos, identificando-se assim possíveis homólogos. A partir desse processo é possível se inferir a provável função de genes novos. É o que tem sido denominado de genética “*in silico*”.

Neste momento cabe diferenciar o termo homologia, similaridade e identidade. Identidade de seqüências é o termo empregado para a ocorrência de exatamente o mesmo nucleotídeo ou

aminoácido na mesma posição do alinhamento. Similaridade é uma medida que verifica quanto similar uma seqüência é da outra, levando-se em conta as possibilidades de substituições. Homologia refere-se a seqüências que são similares e que têm a mesma origem ancestral (Pertsemlidis 2001).

A análise filogenética é um processo que desenvolve hipóteses sobre a relação evolutiva. Para tanto, utiliza a comparação das seqüências, primeiro alinhando-as e depois escolhendo um método para inferir a árvore filogenética. Um dos métodos mais usados é o método baseado na distância, na qual é definida uma matriz de distância entre cada par de seqüências e estas são agrupadas de acordo com essas distâncias. Pode-se fazer uma árvore baseada na junção de vizinhos usando-se a matriz de distância. Neste caso as seqüências mais próximas são consideradas vizinhas e o algoritmo pesquisa as distâncias menores (Graur 1999; Brown 2001; Varga 2003).

1.5.1 O projeto genoma de *Crinipellis perniciosa*

O projeto Genoma de *Crinipellis perniciosa* (www.lge.ibi.unicamp.br/vassoura) nasceu na metade do ano de 2000 idealizado pelo Dr. Gonçalo Amarante Guimarães Pereira. Inicialmente foi feito um projeto piloto, no qual foi construída uma biblioteca de *ShotGun* e seqüenciados cerca de 500 fragmentos que foram comparados contra o banco de dados mundial *GenBank*. Nesta fase inicial minha participação no projeto era de técnica executando o seqüenciamento.

O projeto genoma foi aprovado pelo Governo da Bahia que, através da Secretaria da Agricultura, se posicionou favoravelmente ao seu financiamento. Desta forma foi formada a Rede de Genômica do Estado da Bahia composta por quatro grupos: Laboratório de Genômica e Expressão (LGE) da Universidade Estadual de Campinas (SP), o Centro de Pesquisas do Cacau (CEPEC) da CEPLAC (BA), a Universidade Estadual de Santa Cruz (BA) e o Centro Nacional de Recursos Genéticos (CENARGEN) da EMBAPA (DF).

No início do ano de 2001 foi deflagrado o seqüenciamento, nesta época eu assumia a coordenação de DNA, tendo como responsabilidades a manutenção do fungo na fase micelial, a extração do DNA genômico, a construção de bibliotecas genômicas, o envio e distribuição de clones aos laboratórios de seqüenciamento e seu respectivo controle e suporte técnico. Neste mesmo ano foi necessário que eu ministrasse um treinamento de seqüenciamento ao laboratório CEPEC. No ano de 2002 entraram no projeto mais quatro outros laboratórios financiados pelo CNPQ para ampliar a rede genômica do Estado da Bahia, a Universidade Estadual de Feira de Santana – BA, a Universidade Federal da Bahia – BA) e a Universidade Católica de Salvador – BA. Também foi necessário neste ano que eu ministrasse treinamento de seqüenciamento a estes novos laboratórios.

A estratégia de seqüenciamento foi semelhante a normalmente empregada em projetos genoma de ESTs, baseado no fato de que fungos possuem alta densidade de genes, *introns* pequenos e pouca seqüência repetitiva.

Com isso, o seqüenciamento foi feito nos plasmídeos apenas em uma direção, empregando para isso o *primer* universal (denominado F). Com essa estratégia objetivamos cobrir no menor espaço de tempo e com a menor quantidade de seqüenciamento possível, o máximo do genoma, sempre tendo em mente que o resultado inicial é um rascunho das seqüências.

Quando uma determinada seqüência (*read*), entretanto, é identificada como importante para um determinado objetivo, o plasmídeo que contém o fragmento é resgatado e neste procede-se o seqüenciamento nas duas direções. Assim sendo, completar genes no esquema do nosso projeto é uma etapa posterior e é feito por demanda.

Cada *read* obtido é imediatamente comparado com o bando de dados NR do *Genbank* (Benson *et al* 2000) via o programa BLASTx, que traduz dinamicamente cada *read* nos 6 quadros de leitura possíveis e compara o resultado das proteínas formadas com o banco de proteínas. O resultado dessas comparações, em formato de hipertexto, fica armazenado no banco de dados e disponível para a pesquisa por palavra chave.

OBJETIVOS

OBJETIVO GERAL:

- Identificar, a partir dos dados gerados no projeto genoma de *Crinipellis pernicioso*, os prováveis genes de *mating type* desse fungo e a partir daí deduzir sua organização molecular.

OBJETIVOS ESPECÍFICOS:

- Identificar os genes do sistema *mating type* por similaridade às seqüências de fungos depositadas no *Genbank*.
- Analisar o nível de similaridade dos genes homólogos de *mating type* entre basidiomicetos, que apresentam seqüências conhecidas, posicionando filogeneticamente *C. pernicioso* em relação às demais espécies.
- Analisar se os genes de *C. pernicioso* possuem as regiões conservadas características destes genes do sistema *mating type*.
- Deduzir a possível organização molecular do sistema *mating type* de *C. pernicioso*

2.0 MATERIAIS E MÉTODOS

2.1 Seqüenciamento do genoma de *Crinipellis perniciososa*

2.1.1 Construção de bibliotecas genômicas

Para a construção de biblioteca genômica foi usado o DNA total do fungo *C. perniciososa* na fase micelial dicariótica. O fungo foi crescido em meio malte líquido sob agitação e em temperatura de 28 °C por 20 dias. O DNA foi extraído segundo métodos tradicionais de extração de DNA de fungos e plantas (Talbot 2001).

A biblioteca genômica foi baseada na estratégia de *shotgun* (Venter *et al* 1998), sendo usado como método para obtenção dos fragmentos a quebra física através de do uso de sonicador e nebulizador (Surzcki 1990; Surzcki 2000). Os fragmentos foram selecionados através de corrida em gel de eletroforese em tamanhos de 1-2 kb e 2-4kb e purificadas com S.N.A.P. Gel Purification Kit (Invitrogen –Life technologies).

Os vetores usados para a inserção dos fragmentos foram escolhidos conforme sua eficácia na produção de clonagens em larga escala para genomas. Inicialmente foi usado o vetor pUC18, um plasmídio largamente usado para clonagens. Este vetor fazia parte de um pacote (SureClone kit – Amershan-Pharmacia) de “ferramentas” necessárias para o sucesso da clonagem e utilizava a propriedade da enzima ligase para a inserção do fragmento. O outro vetor usado foi o pCR4Blunt que também pertencia a um pacote (TOPO Shotgun Subcloning kit – Invitrogen – Life technologies) de “ferramentas” necessárias para o sucesso da clonagem e utilizava a propriedade da enzima topoisomerase para a inserção do fragmento.

Os plasmídios recombinantes foram transformados em células competentes de *E. coli* DH10b resistentes a ampicilina através da técnica de eletroporação. Os plasmídios possuíam a propriedade de identificação de moléculas recombinantes.

As células clonadas foram selecionadas das células sem inserto pela coloração quando colocadas em meio sólido com X-gal. Os clones eram colocados para crescer em microplacas de 96 poços e estocados em temperatura de – 80°C.

2.1.2 Seqüenciamento do genoma de *Crinipellis pernicios*

A extração do DNA plasmidial foi realizada em grande escala sendo produzida em microplacas de 96 poços seguindo a técnica de lise alcalina (Engebrecht *et al* 1998). Esse método utiliza três fases para a extração e uma para a precipitação do DNA. As fases de extração compreendiam a uma ressuspensão das células em um meio tamponante contendo glicose, EDTA e Tris-HCl, seguida por uma lise alcalina para promover a denaturação das proteínas e do DNA através de SDS e NaOH, seguida por uma solução contendo acetato de potássio com a função de renaturação do DNA plasmidial e precipitação das proteínas e do DNA cromossomal. O DNA plasmidial foi precipitado com isopropanol.

O seqüenciamento foi feito através de seqüenciadores automáticos, 377, 3700, 3100 (Applied Biosystems) e o Mega Bace (Amersham Pharmacia). Foram usados os respectivos pacotes “kits” de “ferramentas” indicadas pelos fabricantes.

De uma maneira geral estes “kits” de seqüenciamento são baseados em dideoxinucleotídeos marcados por fluorescência baseado na técnica de Sanger (Sanger *et al*) que quando visualizados em seqüenciadores, permitem a identificação da seqüência a ser descoberta.

Montagem dos fragmentos

No projeto genoma de *C. pernicioso* os eletroferogramas das seqüências obtidas foram submetidos a um filtro para eliminar as seqüências do vetor e possíveis contaminações de organismos com que trabalhamos no laboratório, como *E. coli*, *Xylella fastidiosa* e *S.cerevisiae*. Em seguida foi verificada a qualidade das seqüências através dos programas Phred (Ewing *et al*¹ 1998; Ewing *et al*² 1998) em sistema operacional Linux. Essas seqüências estão armazenadas em um banco de dados interno de consulta restrita (Projeto genoma *Crinipellis pernicioso* – www.lge.ibi.unicamp.br/vassoura).

2.1.3 Busca por genes *mating type*

Para identificar possíveis genes do sistema *mating type* de *C. pernicioso* foram feitas duas etapas de procura: uma por palavra-chave e outra por comparação de seqüências:

Na primeira etapa foi feita uma busca, no banco de dados do genoma de *C. pernicioso*, por possíveis reads candidatos através da ferramenta de procura por palavra-chave usando termos como “*mating*”, “*pheromone*”, “*receptor*”, entre outros, através do programa GENE PROJECT (GP) (Carazolle *et a*).

Em uma segunda etapa, as seqüências de genes de organismos caracterizados como envolvidos no processo de *mating type* identificados na etapa anterior *Coprinus cinereus*, *Coprinus bilanatus*, *Schizophyllum commune* e *Ustilago maydis* foram isolados e comparados contra o banco de *C. pernicioso* através da ferramenta tBLASTn (Altschul *et al* 1997) incorporada ao GP.

Esses reads selecionados foram alinhados e clusterizados como um “quebra-cabeça”, a fim de se obter a maior seqüência possível “*contig*”, usando o programa PHRAP incorporado ao GP.

Para obter uma extensão adicional desses *contigs*, o consenso de cada um deles foi comparado por BLASTn com os *reads* de *C. pernicioso*. Também como forma de aumentar o *contig* e/ou melhorar sua qualidade foi feito o seqüenciamento no sentido inverso dos clones selecionados.

2.2 Nível de similaridade e filogenia dos genes de *mating type*

Para analisar o nível de similaridade foram feitas comparações entre as seqüências nucleotídicas e de aminoácidos das proteínas HD1, HD2 e RC dos fungos *C. pernicioso*, *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis*, através do programa BLASTtwo-seq.

Para avaliar se os genes encontrados de *C. pernicioso* eram homólogos aos genes dos basidiomicetos e identificar com qual organismo o parentesco é mais próximo foi feita uma análise evolutiva pelo método de distância. Para isso foram usadas as seqüências dos genes que codificam as proteínas HD1, HD2 e RC da tabela 2 e as seqüências dos *contigs* de *Crinipellis pernicioso* no formato FASTA, gerando com isso árvores filogenéticas.

Primeiramente as seqüências foram alinhadas pelo programa CLUSTAL-X (Thompson and Jeanmougin 1997), seguidas por análise do número de modificações necessárias para se converter uma seqüência em outra através do método de distancia Junkes-Cantor DNA dist – Neighbor (Nei 1987) do pacote de programas Philip 3.6 (Felsenstein, 1993) analisados por *bootstrap* (Brown 1994) com *data set* de 100 e usando como *outgroup* as seqüências do fungo ascomiceto *Sacharomyces cerevisiae*. As árvores filogenéticas resultantes foram visualizadas pelo programa Tree View 1.6.1 (Page, 1996).

Nesse método as distâncias são estimadas partindo-se do número de modificações necessárias para se converter uma seqüência em outra. Quanto menor o número de modificações, mais provável a proximidade evolutiva entre as duas seqüências e, portanto, entre os organismos que as hospedam, considerando-se tratar-se de uma seqüência herdada verticalmente.

2.3 Análise das regiões conservadas das proteínas HD1, HD2 e RC

Foram comparadas as três classes de proteínas, HD1, HD2 e RC de *C. perniciososa* com as de *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis*, através do alinhamento das seqüências de aminoácidos pelo programa Clustal_X (Thompson 1997) e alinhadas pelo programa GeneDoc (Nicholas 1997).

HD1

Para a classe de proteínas HD1 foram usadas as seqüências das proteínas:

C. cinereus (b1, d1), *C. bilanatus* (a1) e *S. commune* (Z).

HD2

Para a classe de proteínas HD2 foram usadas as seqüências das proteínas:

C. cinereus (a2, b2), *C. bilanatus* (a2) e *S. commune* (Y).

RC

Para a classe de proteínas RC (receptor de ferormônio) foram usadas as seqüências das proteínas dos organismos:

C. cinereus (rcb 1, rcb2, rcb3) e *S. commune* (bar, bbr).

3.0 RESULTADOS

3.1 Identificação dos genes do sistema *mating type* de *C. pernicioso*

3.1.1 Seqüenciamento do genoma de *Crinipellis pernicioso*

Um total aproximado de 137.000 *reads* foram seqüenciados de bibliotecas de *shotgun* com insertos com tamanho médio de 2,0 a 4,0kb, totalizando cerca de 125.000.000 pares de base até o momento. Isso perfaz cerca de 4 vezes o tamanho do genoma de *C. pernicioso*, definido em aproximadamente 30,3 MB (Rincones *et al* 2003).

As seqüências obtidas geraram um banco de dados disponível para consulta através de comparações DNA:DNA (BlastN), Proteína:DNA traduzido (tBlastN) ou DNA traduzido:DNA traduzido (tBlastX), gerenciadas pelo programa GP (Carazolle *et al*).

3.1.2 Busca por genes *mating type*

Inicialmente, a busca por possíveis candidatos feita por palavra-chave, foi pouco restritiva. Os termos “*mating*”, “*pheromone*” e “*receptor*”, foram os que geraram resultados mais consistentes. Foram identificados *reads* com similaridade aos genes que codificam as proteínas HD1, HD2 e RC (tabela 1) e não foi identificada similaridade aos genes que codificam ferormônios (FE). Esses genes identificados eram de organismos filogeneticamente relacionados a *C. pernicioso* e dentre estes estavam os fungos basidiomicetos *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis*.

Na segunda etapa foram isoladas as seqüências de genes *mating type* que codificam as proteínas HD1, HD2 e RC dos fungos basidiomicetos: *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis* (tabela 2) conforme o resultado por palavra-chave. Foi descartada a possibilidade de se fazer

a busca de genes que codificam ferormônios, pois essas seqüências são muito diferentes e possuem apenas alguns trechos conservados de quatro aminoácidos.

As seqüências codificadas pelos genes *b1*, *d1*, *a1*, *Z* e *bE*, apresentados na tabela 2, usados independentemente, via tBlastN, identificaram 2 possíveis *reads* pertencentes ao provável gene HD1 de *C. pernicioso*. Esses *reads* quando alinhados e clusterizados geraram 1 *contig* de tamanho 561pb. Esse *contig* foi re-seqüenciado nas duas direções e o *read* resultante foi adicionado ao *contig*.

As seqüências codificadas pelos genes *a2*, *b2*, *a2*, *Y* e *bW*, apresentados na tabela 2, usados independentemente, via tBlastN, identificaram 2 possíveis *reads* pertencentes ao provável gene HD2 de *C. pernicioso*. Estes *reads* quando alinhados e clusterizados geraram 1 *contig* de tamanho 625pb. Este *contig* foi re-seqüenciado nas duas direções e o *read* resultante foi adicionado ao *contig*.

As seqüências codificadas pelos genes *rcb1*, *rcb2*, *rcb3*, *bbp*, *bar*, *pra1* e *pra2* apresentados na tabela 2, usados independentemente, via tBlastN, identificaram 7 possíveis *reads* pertencentes ao provável gene HD2 de *C. pernicioso*. Esses *reads* quando alinhados e clusterizados geraram 4 *contigs*, *Rc1*, *Rc2*, *Rc3* e *Rc4* de tamanhos 712, 1114, 886, 112 pb, respectivamente. Esses *contigs* foram re-seqüenciados nas duas direções e o *read* resultante foi adicionado ao *contig*.

ORGANISMOS		GENES			GENES	
		<i>LOCUS HD</i>			<i>LOCUS FRF</i>	
		HD1		HD2		RC
	<u>N^o</u> <u>ACESSO</u> <u>GENBANK</u>		<u>N^o</u> <u>ACESSO</u> <u>GENBANK</u>		<u>N^o</u> <u>ACESSO</u> <u>GENBANK</u>	
<i>C. cinereus</i>	X62336.1	<i>b1</i>	X79686.1	<i>a2</i>	AAF01418	<i>rcb1</i>
	X79688.1	<i>d1</i>	X79687.1	<i>b2</i>	AAF01419	<i>rcb2</i>
					AAF01420	<i>rcb3</i>
<i>C. bilanatus</i>	AAK1706	<i>a1</i>	AAK17070	<i>a2</i>		
<i>S. commune</i>	P37938	<i>Z</i>	A46203	<i>Y</i>	Q92275	<i>bar</i>
					P78741	<i>bbp</i>
<i>U. maydis</i>	P22020	<i>bE</i>	D42094	<i>bW</i>	P31303	<i>pra1</i>
					AAD56044	<i>pra2</i>

Tabela 2. Genes do sistema *mating type* e os respectivos números de acesso ao *Genbank*. (Benson *et al* 2000).

3.2 Nível de similaridade e filogenia dos genes de *mating type*

Os níveis de similaridade analisados foram colocados em tabelas demonstrativas. Observa-se que o organismo com maior nível de similaridade a *C. pernicioso* é o mesmo que está mais próximo a *C. pernicioso* filogeneticamente, segundo os dados abaixo.

Não foram identificadas similaridades entre as seqüências nucleotídicas apenas entre as seqüências de aminoácidos de todas as proteínas.

- Comparação por similaridade através do valor do *e-value* entre as proteínas HD1:

	<i>Gene Hd1 C. Perniciosa</i>
<i>Gene b1 C. cinereus</i>	Z-14
<i>Gene d1 C. cinereus</i>	2e-09
<i>Gene a1 C. bilanatus</i>	4e-12
<i>Gene Z S. commune</i>	2.3
<i>Gene bE U. maydis</i>	X

Tabela 3. Cruzamentos entre as seqüências através do programa *BLASTtwo Seq*. **X** representa nenhuma similaridade encontrada.

- Comparação por similaridade através do valor do *e-value* entre as proteínas HD2:

	<i>Gene HD2 C. perniciosa</i>
<i>Gene a2 C. cinereus</i>	5E-06
<i>Gene b2 C. cinereus</i>	1e-06
<i>Gene a2 C. bilanatus</i>	6e-08
<i>Gene Y S. commune</i>	1e-04
<i>Gene bW U. Maydis</i>	X

Tabela 4. Cruzamentos entre as seqüências através do programa *BLASTtwo Seq*. **X** representa nenhuma similaridade encontrada.

- Comparação por similaridade através do valor do *e-value* entre as proteínas HD2:

	<i>Genes de C. perniciosa</i>			
	<i>Rc1</i>	<i>Rc2</i>	<i>Rc3</i>	<i>Rc4</i>
<i>Gene rcb1 C. cinereus</i>	1e-06	9e-10	6E-21	2E-19
<i>Gene rcb2 C. cinereus</i>	7e-10	1e-09	2e-23	1e-14
<i>Gene rcb3 C. cinereus</i>	1e-07	6e-29	1e-59	4e-40
<i>Gene bbp S. commune</i>	0.021	2e-06	7e-28	1e-13
<i>Gene bar S. commune</i>	6e-12	2e-23	3e-55	1e-38
<i>Gene pra1 U. maydis</i>	0.001	3e-04	2e-14	2e-05
<i>Gene pra2 U. maydis</i>	5.6	X	7e-16	8e-04

Tabela 5. Cruzamentos entre as seqüências através do programa *BLASTtwo Seq*. **X** representa nenhuma similaridade encontrada.

A comparação das três classes de proteínas, HD1, HD2 e RC (genes da tabela 2), pelo método de distância NEIGHBOR–Junkes-Cantor (Nei 1987) permitiu inferir três árvores filogenéticas visualizadas através do programa Tree view 1.6.1 (Page, 1996).

Na análise da seqüência HD1 de *C. pernicioso* foi gerada a árvore filogenética representada na figura 18 mostrando o parentesco evolutivo mais próximo ao fungo basidiomiceto *Coprinus cinereus*.

Na análise da seqüência HD2 de *C. pernicioso* foi gerada a árvore filogenética representada na figura 19 mostrando o parentesco evolutivo mais próximo ao fungo basidiomiceto *Coprinus cinereus*.

Na análise da seqüência RC de *C. pernicioso*, não foi possível alinhar todos os quatro possíveis genes de receptores de ferormônio devido aos diferentes tamanhos, desta forma foi utilizado o *contig Rc3* para gerar a árvore filogenética representada na figura 20. O parentesco evolutivo mais próximo a *C. pernicioso* também foi com *C. cinereus*.

1º Árvore:

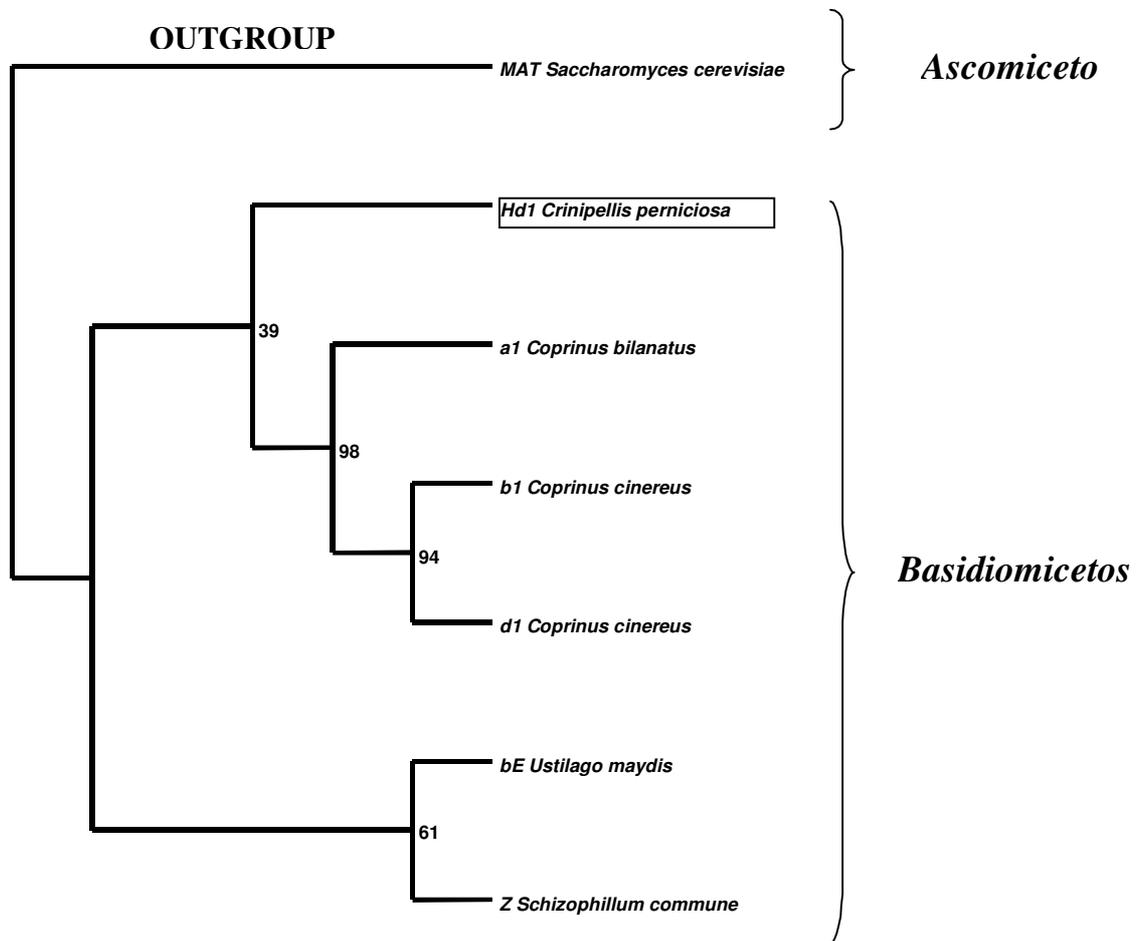


Figura 19 – Árvore filogenética dos genes que codificam a proteína HD1. Valores de *bootstrap* são mostrados. Dados gerados por Neighbor. *Outgroup* usado: Mat de *S. cerevisiae*.

2º Árvore:

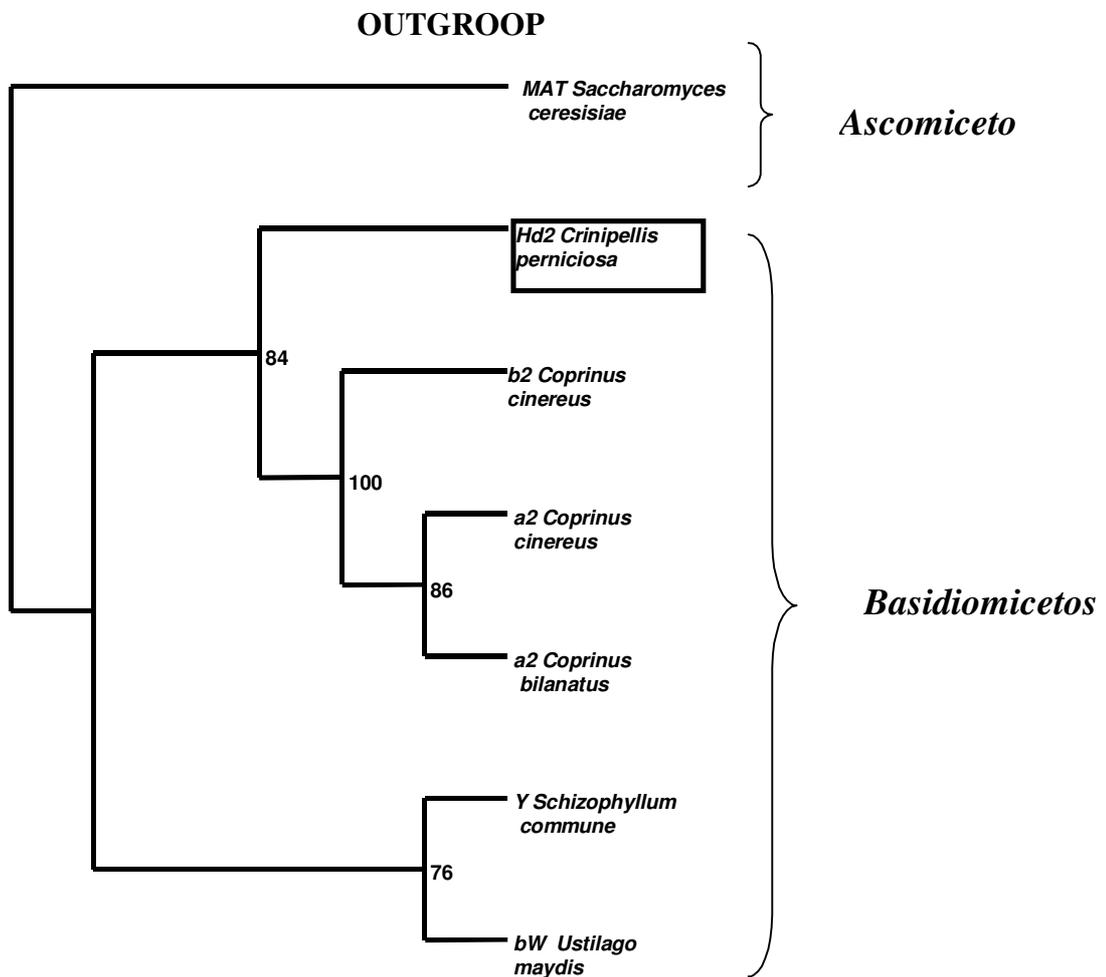


Figura 20 – Árvore filogenética dos genes que codificam a proteína HD2. Valores de *bootstrap* são mostrados. Dados gerados por Neighbor. *Outgroup* usado: Mat de *S. cerevisiae*.

3º Árvore:

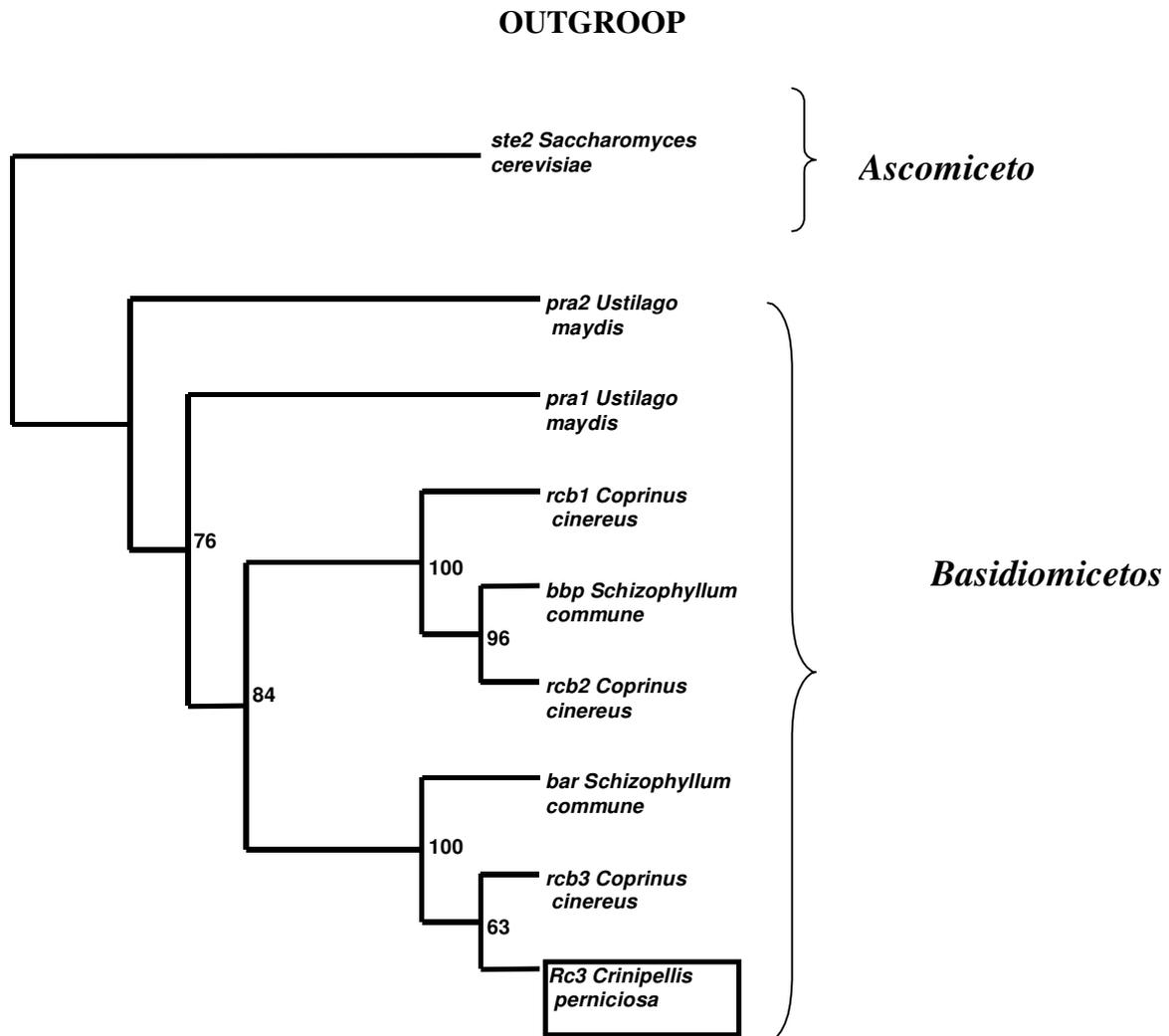


Figura 21 – Árvore filogenética dos genes que codificam a proteína RC. Valores de *bootstrap* são mostrados. Dados gerados por Neighbor. *Outgroup* usado: Mat de *S. cerevisiae*.

3.3 Análise das regiões conservadas das proteínas HD1, HD2 e RC

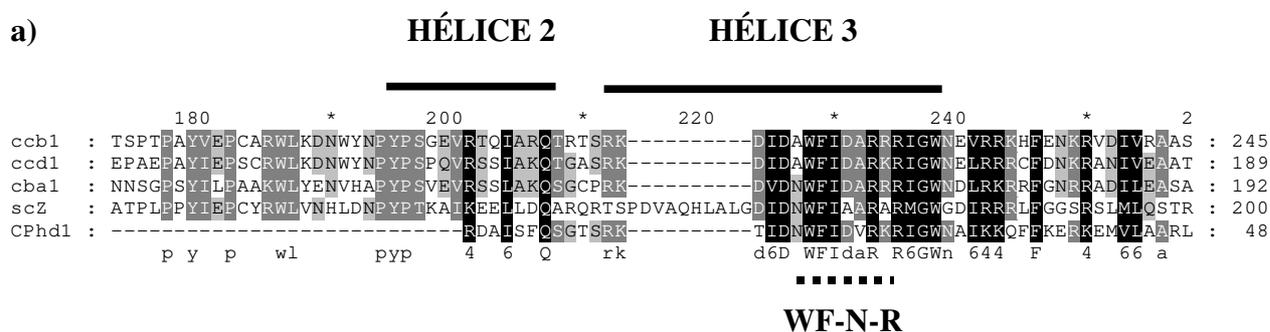
HD1

O alinhamento realizado entre as proteínas HD1 de outros organismos e a aprovável proteína HD1 de *C. pernicioso* identificou três regiões de similaridade:

a) região conservada do homeodomínio de HD1, mais especificamente a segunda e a terceira hélice – hélice de reconhecimento - (Gehring *et al* 1994; Casselton 1998), incluindo o sinal WF-N-R extremamente conservados (Scott *et al* 1989) (Fig.21-a).

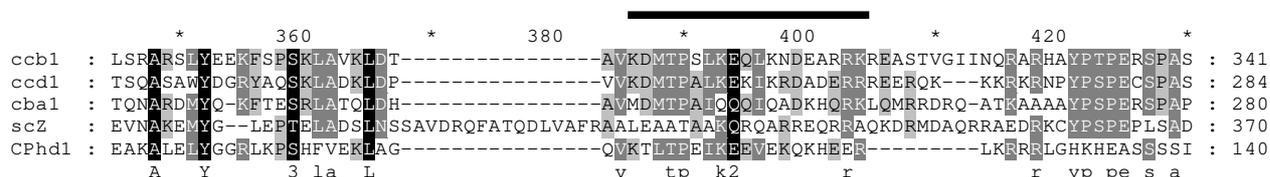
b) curta seqüência na região carboxi-terminal “cauda alfa2” cuja função em *Saccharomyces cerevisiae* (Kues *et al* 1994; Hiscock 1999; Stark *et al* 1999) é a heterodimerização entre as proteínas. Seqüências homólogas também foram encontradas em *C. cinereus* (Kues *et al* 1994; Bradane 1999), *S. commune* (Stankis *et al* 1992; Magae *et al* 1995) e *U. maydis* (Kronstad 1990; Gillissen *et al* 1992) (Fig.21-b).

c) sinal de localização nuclear NLS1 específico a proteína HD1 (Tymon *et al* 1992; Gehring *et al* 1994) (Fig.21-c).



b)

“CAUDA ALFA2”



c)

NLS1

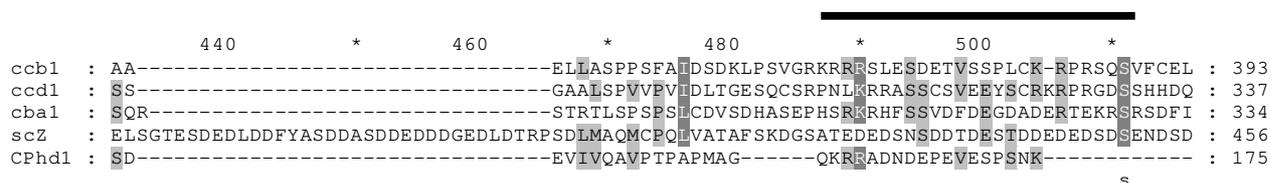


Figura 21. Alinhamento entre proteínas HD1. *C. pernicioso* (Cphd1), *C. cinereus* (ccb1=b1, ccd1=d1), *C. bilanatus* (cba1=a1) e *S. commune* (scz=Z). Resíduos de aminoácidos estão marcados em três níveis de conservação: preto (100%), cinza escuro (80%) e cinza claro (60%). a) Regiões das hélices do homeodomínio e o sinal WF-N-R. b) Região da “cauda alfa2”. c) Sinal de localização nuclear NLS1.

HD2

O alinhamento realizado entre as proteínas HD2 e a aprovável proteína de *C. pernicioso* identificou similaridade com a região conservada do homeodomínio de HD2, as hélices de reconhecimento, incluindo o sinal WF-N-R extremamente conservado na terceira hélice – hélice de reconhecimento - (Scott *et al* 1989; Gehring *et al* 1994; Casselton 1998) (Fig. 22).

HÉLICE 1

```

*      100      *      120      *      140      *      160      *
cca2 : LQTECKNKIKTQALSDLPDTS-----ASRHRLQIINKTFTKLYITRHAPRLREQVIAQLETKLAAPPSKQAFNVHYIPVLEKYYF : 166
ccb2 : QRTLHLTNYQRMCGGYLELSKQQGMETDPISVRNMQTLVERS---FQLISRITQSEIMERYQOTRQQLVKKRTPFNSEYTPVLEKYYF : 168
cba2 : LRRVHLDNVEQTCKRYIKANQATLATEPARHFRLQAVCAKQRTTFQSKVAQIQADCLAKCHQAMPQESKRTPENAEYTPVLEKYYF : 166
scy  : WRYSLDESLSQAARSVSPHNLHLLSTLRFRLYTEQVERWAVQVLQVAEQWRAEMEKQRAHIAASTDKSKKPRPKFHSYTPVLEKYYF : 165
hd2  : -----f eY P LEkYF

```

HÉLICE 2

HÉLICE 3

```

180      *      200      *      220      *      240      *      260
cca2 : EY---NAYPTAQDRALARKSMMSARQIEVWFQNHRRRARKEGKPLQ----- : 210
ccb2 : EY---NAYPSARDREWLARKTMMMSVRQIEVWFQNHRRRARKEGIHFE----- : 212
cba2 : EY---NAYPSVRDRERLARKSMMTPROIEVWFQNHRRRARKEGRPLQ----- : 210
scy  : RF---NAYPTYADRRVLAEKTGMLTRQITVWFQNHRRRARKEGKPLRMTPTAKIPLEEFERERENLARKLLPMLPPHLRPITLGNNKT : 249
hd2  : EL---DAYPSIRDQEVMAOKSGMTRQIEVWFQNSRQPAS----- : 51
      e nAYP Dr lA K M RQIEVWFQNhR rA
          *****

```

WF-N-R

Figura 22. Alinhamento entre proteínas HD2. *C. pernicioso* (hd2), *C. cinereus* (ccb2 = b2 ; cca2 = a2), *C. bilanatus* (cba2 = a2) e *S. commune* (scy=Y). Resíduos de aminoácidos estão marcados em três níveis de conservação: preto (100%), cinza escuro (80%) e cinza claro (60%). As regiões das hélices do homeodomínio e o sinal WF-N-R estão representados.

RC

O alinhamento realizado entre as proteínas RC (receptor de ferormônio) e as aprováveis proteína RC de *C. pernicioso* (Rc1, Rc2, Rc3 e Rc4) identificou uma região de similaridade que compreende os setes domínios transmembranas (TMD), três *loops* intracelulares e três *loops* extracelulares do receptor (Nakayama *et al* 1985; Hagen *et al* 1986; Dohlman *et al* 1991; Wendland *et al* 1995; Vaillancourt *et al* 1997; Casselton 1998; O'Shea *et al* 1998) (Fig. 23).

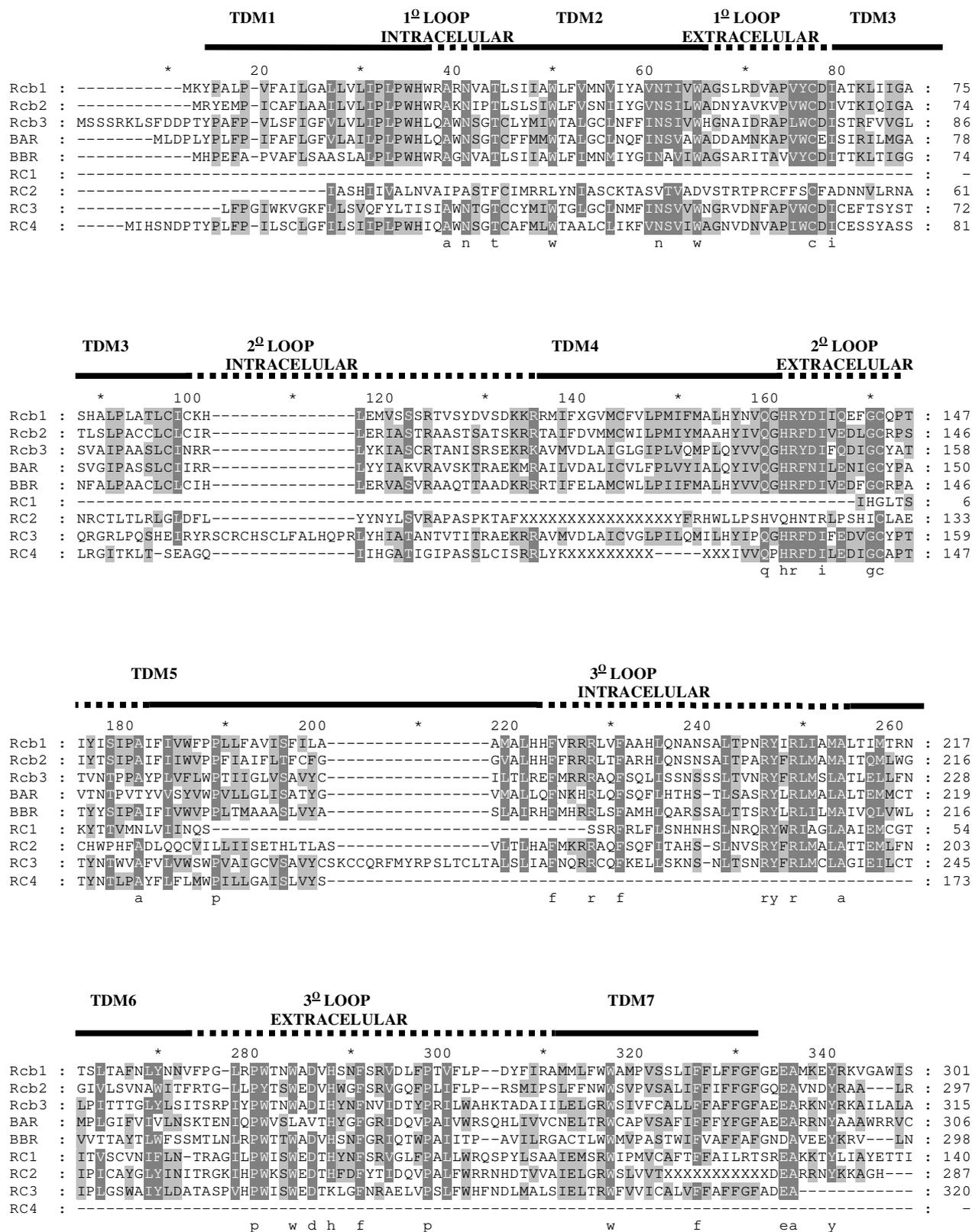


Figura 23. Alinhamento das proteínas RC. *C. pernicioso* (Rc1, Rc2, Rc3 e Rc4), *C. cinereus* (rcb 1, rcb2, rcb3) e *S. commune* (bar, bbr). Resíduos de aminoácidos estão marcados em três níveis de conservação: preto (100%), cinza escuro (80%) e cinza claro (60%). As regiões dos domínios transmembranas (TMD 1-7), os loops intracelulares (1-3) e os loops extracelulares do receptor estão representados.

4.0 DISCUSSÃO

Acredita-se que o surgimento da vida na terra tenha sido um evento único, a partir do qual diversos organismos evoluíram e divergiram, submetidos aos processos de mutações casuais, de portes variados, seguidas por seleção adaptativa. Assim sendo, os organismos, de uma forma geral, guardam entre si mecanismos biológicos comuns e esses estão em grande parte refletidos pela similaridade da informação genética que carregam. Ou seja, na seqüência dos seus genes ou das proteínas codificadas por eles. No final do século passado, o mundo viveu uma verdadeira revolução na biologia.

O desenvolvimento da biologia molecular permitiu a associação entre genes específicos e os fenômenos que a genética tradicionalmente estudava ao nível fenotípico. Posteriormente, a genômica permitiu que novos organismos pudessem ser estudados a partir da comparação da seqüência do seu material genético, para o qual informações experimentais não estavam disponíveis, com a seqüência de genes com funções bem estabelecidas em outros organismos. Era a aplicação da propriedade transitiva da matemática na biologia. Ou seja, se o organismo X possui o gene A que realiza a função K e o organismo Y possui o gene B que é altamente similar ao gene A, a conclusão que se chega é que provavelmente o organismo B possui a capacidade de realizar a função K. É o que tem sido denominado de biologia “*in silico*” e embora esse tipo de procedimento não possa ser considerado conclusivo e necessite de comprovação experimental, ele abre a perspectiva de se fazer uma biologia baseada em hipóteses. Ou seja, ao invés das investigações de processos biológicos em organismos pouco estudados começarem na base da tentativa e erro, elas já se apoiam em possibilidades e hipóteses que podem ser testadas.

Baseado neste método “*in silico*” foi definido neste trabalho a investigação do sistema *mating type* de *C. perniciosus* apoiado em comparação com sistemas semelhantes existentes na

literatura dos organismos *Coprinus cinereus*, *Coprinus bilanatus*, *Schizophyllum commune*, *Ustilago maydis* e *Saccharomyces cerevisiae*.

Neste trabalho foram encontrados no banco de dados de *C. pernicioso* as seqüências parciais de seis possíveis genes do sistema *mating type*.

Na primeira análise verificamos que a comparação entre as seqüências nucleotídicas dos prováveis genes não apresentaram similaridade a nenhum dos genes usados como modelo (obtidos dos fungos basidiomicetos *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis*). Entretanto, quando comparamos as prováveis seqüências das proteínas codificadas por esses genes verificamos altos valores de similaridade, indicando um funcionamento semelhante para o sistema. O fato das seqüências de nucleotídeos não apresentarem grande semelhança não é surpresa, visto que os organismos, durante a evolução tendem a uma utilização preferencial de *códons*, o que acarreta em variabilidade da seqüência sem que necessariamente traga como consequência variação da identidade dos aminoácidos.

As proteínas HD1 de *C. cinereus* e *C. bilanatus* apresentaram valores estatísticos *e-value* em torno de e^{-12} significando que dificilmente essa similaridade ocorreu ao acaso. Da mesma forma as proteínas HD2 de *C. cinereus*, *C. bilanatus* e *S. commune* também apresentaram valores estatísticos de *e-value* de e^{-4} a e^{-8} . Interessante, com as proteínas de receptores RC, similaridades ocorreram com todos os organismos modelos comparados e com valores de *e-value* extremamente baixos (14 à 59 casas), as similaridades mais próximas a zero e portanto mais impossíveis de terem ocorrido ao acaso foram as de *C. cinereus* e *S. commune*.

Essas diferenças de similaridade entre as seqüências das prováveis proteínas podem ser explicadas pelo fato de elas serem proteínas de alta especificidade de ligação, mas com funções semelhantes.

A segunda análise confirmou os resultados de similaridade. A análise filogenética demonstrou que os fragmentos gênicos de *C. perniciosus* são homólogos aos genes dos fungos modelos e o parentesco mais próximo em relação a essas proteínas.

As árvores filogenéticas das proteínas HD1, HD2 e RC mostraram que *C. perniciosus* está filogeneticamente mais próxima a *C. cinereus* mostrando *bootstrap* de 39, 84 e 100, respectivamente. Estes dados conferem com os resultados de níveis de similaridade realizados.

Desta forma seguimos com a terceira análise, a análise de regiões conservadas e características, usando apenas as proteínas mais similares. Novamente, embora tenham sido usadas seqüências parciais, algumas regiões funcionais puderam ser claramente identificadas.

O alinhamento realizado entre as proteínas HD1 de *C. cinereus*, *C. bilanatus* e a provável proteína Hd1 de *C. perniciosus* identificou três regiões de similaridade. A primeira região foi a do homeodomínio, mais especificamente a segunda e a terceira hélice. A proteína Hd1 de *C. perniciosus* possui aminoácidos idênticos aos de *C. cinereus* e *C. bilanatus* na segunda hélice (Arg,xx,Ile,xx,Glu). Também foi verificado que a proteína Hd1 possui aminoácidos idênticos aos de *C. cinereus* e *C. bilanatus* da terceira hélice (Ile, Asp,x,Trp, Phe, Ile,xx,Arg,x Arg, Ile, Gly, Trp) incluindo o sinal modificado e extremamente conservado WF-N-R (WFIDAR, para *C. cinereus* e *C. bilanatus*) e (WFIDVR, para *C. perniciosus*) (Scott *et al* 1989; Gehring *et al* 1994; Casselton 1998).

A segunda região de similaridade foi a “cauda alfa2” de 20 aminoácidos na região carboxi-terminal. Esta cauda tem função em *S. cerevisiae* (Kues *et al* 1994; Hiscock 1999; Stark *et al* 1999) de heterodimerização entre as proteínas HD1 e HD2. Seqüências homólogas também foram encontradas em *C. cinereus* (Kues *et al* 1994; Bradane 1999). Destes 20 aminoácidos, 13 são conservados (a média dos aa conservados de *C. cinereus* e *C. bilanatus* é 14) e um deles com identidade ao único aminoácido idêntico em *C. cinereus* e *C. bilanatus* (Gln) (Kues *et al* 1994; Hiscock and Kues 1999; Stark *et al* 1999).

A terceira região de similaridade também está na região carboxi-terminal, na qual foi identificado o primeiro sinal de localização nuclear NLS1 da proteína HD1 (Tymon *et al* 1992; Gehring *et al* 1994). Aminoácidos conservados de *C. cinereus* e *C. bilanatus* (Arg, Arg,x7,Val, x3, Ser,x, Lys) foram encontrados em *C. pernicioso*.

Podemos concluir que o fragmento gênico *Hd1* de *C. pernicioso*, apesar de parcial codifica a proteína de heterodimerização HD1 presente nos sistemas *mating type* de fungos, pois possui a região do homeodomínio, a região da cauda alfa2 responsável pela heterodimerização e o exclusivo sinal de localização nuclear.

O alinhamento realizado entre as proteínas HD2 de *C. cinereus*, *C. bilanatus* e a provável proteína Hd2 de *C. pernicioso* também identificou a região do homeodomínio, especialmente a região de reconhecimento ao DNA. Todos os aminoácidos idênticos (Tyr, -, Pro, -, Leu, Gln, -, Try, Phe) da primeira hélice, entre as espécies de *C. cinereus* e *C. bilanatus*, também estão presentes na proteína Hd2. Da mesma forma na segunda e na terceira hélice, todos os aminoácidos idênticos de *C. cinereus* e *C. bilanatus* também estão presentes na proteína Hd2. Na segunda hélice os aminoácidos (Asp, Ala, Tyr, Pro, Thr, xx, Asn, 4x, Leu, Ala, x, Lys, Ser) e na terceira hélice os aminoácidos (Arg, Glu, Ile, x, Val, Met, x, Glu, 5x, Ala) incluindo o sinal de localização nuclear WF-N-R (WFQNHR, para *C. cinereus* e *C. bilanatus*) (WFQNSR, para *C. pernicioso*), aminoácidos extremamente conservados na terceira hélice – hélice de reconhecimento - (Scott *et al* 1989; Gehring *et al* 1994; Casselton 1998)

Podemos concluir que o fragmento gênico *Hd2* de *C. pernicioso* codifica a proteína de heterodimerização HD2 presente nos sistemas *mating type* de fungos, pois possui a região do homeodomínio especialmente o sinal WF-N-R da hélice de reconhecimento ao DNA.

No alinhamento das proteínas Rc de *C. pernicioso* aos de *C. cinereus* e *S. commune*, devido os genes *Rc1*, *Rc2*, *Rc3* e *Rc4* serem apenas fragmentos nem todos identificaram as mesmas regiões.

Rc1 apresentou similaridade às regiões do segundo e terceiro *loop* extracelular, terceiro *loop* intracelular e os domínios transmembranas TDM5,6 e7. Rc2 apresentou similaridade as regiões do primeiro, segundo e terceiro *loop* intracelular, primeiro, segundo e terceiro *loop* extracelular e os domínios transmembranas TDM2,3,4,5,6 e7. Rc3 apresentou similaridade as regiões do primeiro, segundo e terceiro *loop* intracelular, primeiro, segundo e terceiro *loop* extracelular e os domínios transmembranas TDM1,2,3,4,5,6 e7. Rc4 apresentou similaridade as regiões do primeiro e segundo *loop* intracelular, primeiro e segundo *loop* extracelular e os domínios transmembranas TDM2,3,4,5.

Podemos concluir que os quatro fragmentos gênicos de *C. pernicioso* codificam proteínas que possuem as características dos receptores rodopsinas presentes nos fungos basidiomicetos, pois apresentaram similaridades às sete hélices transmembranas hidrofóbicas (7-TM), aos três loops internos e aos três loops externos (Nakayama *et al* 1985; Hagen *et al* 1986; Vaillancourt *et al* 1997; Casselton 1998; O'Shea *et al* 1998; Hiscock 1999; Olesnicky *et al* 1999; Kues 2000; Olesnicky *et al* 2000). Outra conclusão importante é que embora regiões desses fragmentos estejam sobrepostas funcionalmente, as suas seqüências de nucleotídeos são diferentes, indicando claramente que se tratam de alelos.

Esses dados associados ao fato do biótipo L de *C. pernicioso* ter um ciclo sexual controlado por um sistema tetrapolar (Griffith 1994) podemos deduzir que o biótipo C de *C. pernicioso* também tem um sistema tetrapolar, com um *locus* codificando as proteínas HD e outro *locus* codificando as proteínas RC e que o locus HD é bialélico (um par de genes *Hd1* e *Hd2*) e o locus *FRF* é multialélico (quatro genes pra receptores *Rc1*, *Rc2*, *Rc3* e *Rc4*). Desta forma podemos deduzir que a organização molecular do sistema *mating type* de *C. pernicioso* é a seguinte:

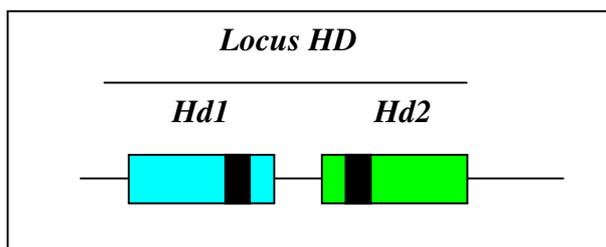


Figura 24. Provável organização molecular do *locus HD* de *C. pernicioso*. Caixas pretas representam as regiões dos homeodomínios. Caixa azul representa o gene *Hd1* e a caixa verde representa o gene *Hd2*.

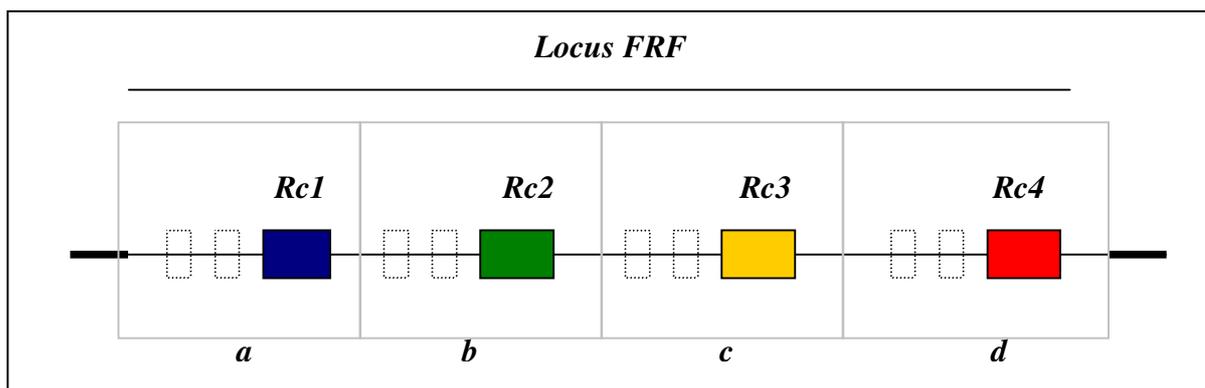


Figura 25. Provável organização molecular do *locus FRF* de *C. pernicioso*. Caixas coloridas representam os genes *Rc1*, *Rc2*, *Rc3* e *Rc4* divididos em 4 prováveis subfamílias (*a*, *b*, *c* e *d*). Caixas pontilhadas representam regiões onde podem estar localizados os genes de ferormônio.

Os genes de *C. pernicioso* identificados neste trabalho representam a primeira identificação de componentes do sistema *mating type* de fungos basidiomicetos homotáticos primários. Este achado contraria o postulado por Alexopoulos (Alexopoulos 1996) que afirma que fungos homotáticos primários não possuem sistema *mating type*. Entretanto, a simples existência das seqüências não é suficiente para provar que esses genes realmente são expressos, embora seja provável uma vez que o fungo apresenta grampos de conexão durante a sua fase dicariótica. Além disso, os primeiros experimentos com *microarray* de *C. pernicioso*, realizados pela aluna de doutoramento Maricene Sabha, mostrou a expressão do gene *Rc3*, aparentemente constitutivo para as condições testadas (comunicação pessoal).

Embora não possamos afirmar se estes genes do sistema *mating type* estão ativos (não foram realizados trabalhos de expressão gênica) o fato de outros pesquisadores terem identificado a formação de grampos de conexão indica muito provavelmente que esses genes são expressos.

Portanto, embora *C. pernicioso*, não use o sistema *mating type* para promover *outbreeding*, pois é um organismo homotático, esse sistema deve ter a função de promover e manter a fase dicariótica, bem como promover o início do corpo de frutificação, além de provavelmente regular os genes responsáveis pela patogenicidade como visto em *U. maydis* (Schulz *et al* 1990; Bolker 2001).

A maneira como as proteínas do sistema *mating type* do *locus HD* funcionam em *C. pernicioso* pode ser similar ao observado em mutantes de *C. cinereus*, onde ocorreu a fusão dos genes que codificam as proteínas HD1 e HD2 fazendo com que seja produzida em uma só unidade genética a proteína que funciona como um heterodímero. Dessa forma o organismo não precisa encontrar um parceiro compatível para haver a formação do heterodímero (Kues *et al* 1994).

Em relação ao *locus FRF*, estudos de mutantes mostraram que apenas algumas mudanças de aminoácidos na região de reconhecimento da ligação receptor/ferormônio tornaram o organismo auto compatível não necessitando de um parceiro para a ativação da cascata de sinalização (Olesnicky *et al* 1999; Olesnicky *et al* 2000)

Embora os resultados deste trabalho sejam significativos para compreendermos que um sistema *mating type* bifatorial deve estar controlando a patogenicidade de *C. pernicioso* em cacau, outros trabalhos devem ser feitos nesse sentido. Podemos sugerir que seja feita a clonagem de todo os *loci mating type* de *C. pernicioso*, não só deste biótipo em estudo, mas do biótipo heterotático, comparando-as entre si e com as espécies de basidiomicetos e que sejam feitos estudos de mutação e indução de reprodução em *C. pernicioso in vitro*.

5.0 CONCLUSÕES

Podemos concluir a partir deste trabalho que *Crinipellis perniciososa* possui genes do sistema *mating type*. Foram identificados seis genes: um gene que codifica a proteína HD1, um gene que codifica a proteína HD2 e quatro genes que codificam proteínas receptores de ferormônio.

Esses genes são homólogos aos genes de *C. cinereus*, *C. bilanatus*, *S. commune* e *U. maydis*.

O organismo mais próximo à *C. perniciososa* para estudos do *locus HD* e *locus FRF* é o fungo *C. cinereus*.

Podemos concluir que *C. perniciososa* é um fungo homotálico com um sistema *mating type* tetrapolar (com um *locus HD* e um *locus FRF*) e multialélico no *locus FRF*.

6.0 REFERÊNCIAS BIBLIOGRÁFICAS

- Alexopoulos, C. J. (1996). Introductory mycology. U. S. A., John Wiley & Sons, Inc.
- Altschul, S. F *et al.* (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**: 3389-3402.
- Asante-Owusu, *et al.* (1996). "Heterodimerization between two classes of homeodomain proteins in the mushroom *Coprinus cinereus* brings together potential DNA-binding and activation domains." Gene **172**(1): 25-31.
- Baker, R. E. D. and Holliday, P. (1957). Witches' broom disease of cocoa (*Marasmius perniciosus* Stahel). Richmond, Commonwealth Mycological Institute.
- Banham, A. H. *et al.* (1995). "An N-terminal dimerization domain permits homeodomain proteins to choose compatible partners and initiate sexual development in the mushroom *Coprinus cinereus*." Plant Cell **7**: 773-783.
- Banuett, F. (1998). "Signalling in the yeast: An informational cascade with links to the filamentous fungi." Microbiol Mol Biol Rev **62**: 249-274.
- Bardwell, L., *et al.* (1994). "Signal propagation and regulation in the mating pheromone response pathway of the yeast *Saccharomyces cerevisiae*." Dev. Biol. **166**: 363-379.
- Bastos, C. N., *et al.* (1988). "Comparação morfológica de isolados de *Crinipellis perniciososa*." Fitopatologia Brasileira. **13**: 202-206.
- Bastos, C. N., Evans, H. C. (1985). "A new pathotype of *Crinipellis perniciososa* (witches' broom disease) on solanaceous hosts." Plant Pathology **34**: 306-312.
- Bender, A., and Sprague, G. F., Jr. (1989). "Pheromones and pheromones receptors are the primary determinants of mating specificity in the yeast *Saccharomyces cerevisiae*." Genetics **121**: 463-476.
- Benson, D. A., *et al.* (2002). "GenBank." Nucleic Acid Research **30**(1): 17-20.
- Berkower, C. and Michaelis, S. (1991). "Mutational analysis of the yeast a-factor transporter STE6, a member of the ATP binding cassette (ABC) protein superfamily." Embo J **10**: 3777-3785.
- Bolker, M. (2001). "*Ustilago maydis* - a valuable model systems for the study of fungal dimorphism and virulence." Microbiology **147**: 1395-1401.
- Bölker, M., *et al.* (1992). "The a mating type locus of *U. maydis* specifies cell signalling components." Cell **68**: 441-450.
- Bradane, H., May, G. (1999). "The divergence-homogenization duality in the evolution of the *bl* mating type of *Coprinus cinereus*." Mol Biol Evol **16**: 975-986.
- Brown, J.K.M.(1994). Bootstrap hypothesis tests for evolutionary trees and other dendrograms. In Proceedings of the National Academy of Sciences of the United States of America. **91**, pages 12293-12297.
- Brown, J. R. (2001). "Genomic and Phylogenetic Perspectives on the Evolution of Prokaryotes." Syst Biol **50**: 497-512.
- Burglin, T. R. (1997). "Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals." Nucleic Acids Res **25**: 4173-4180.
- Cadwell, G. A., *et al.* (1995). "Fungal lipopeptide mating pheromones, a model system for the study of protein preylation." Microbiol. Rev. **59**: 406-422.
- Calle, H., *et al.* (1982). "Histology of witches' broom caused in cocoa by *Crinipellis perniciososa*." Phitopatologia **72**: 1474-1481.
- Carazolle, M. F., *et al.* "Software Gene Projects". LGE UNICAMP .
- Casselton, L. A. (2002). "Mate recognition in fungi." Heredity **88**(2): 142-7.

- Casselton, L. A., *et al.* (1995). "Mating type control of sexual development in *Coprinus cinereus*." Can J. Bot **73**: S266-S272.
- Casselton, L. A. and N. S. Olesnicky (1998). "Molecular genetics of mating recognition in basidiomycete fungi." Microbiol Mol Biol Rev **62**(1): 55-70.
- Chen, P., *et al.* (1997). "Biogenesis of the *Sacharomyces cerevisiae* mating pheromone a-factor." J. Cell Biol. **136**: 251-269.
- Day, P. R. (1960). "The structure of the A mating-type locus in *Coprinus lagopus*." Genetics **45**: 641-650.
- Delgado, J. C., and Cook, A. A. (1976). "Nuclear condition of the basidia, basidiosporos, and mycelium of *Marasmius perniciosus*." Can J. Bot **54**: 66-72.
- Dohlman, H. G., *et al.* (1991). "Model systems for study of 7-transmembrane-segment receptors." Annu Rev Biochem **60**: 653-688.
- Dranginis, A. M. (1990). "Binding of yeast a1 and alfa2 as a heterodimer to the operator DNA of a haploid-specific gene." Nature **347**: 682-685.
- Elion, E. A. (1995). "Ste5: A meeting place for MAP kinases and their associates." Trends Genet. **5**: 322-327.
- Engebrecht, J., *et al.* (1998). Unit 1.6. Minipreps of Plasmid DNA. Current Protocols in Molecular Biology. F. M. Ausubel, Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. & K. Struhl eds. U.S.A., John Wiley and Sons. **1**: 1.6.1 - 1.6.3.
- Evans, H. C. (1980). "Pleomorphism in *Crinipellis pernicioso*, Causal Agent of Witches' Broom Disease of Cocoa." Transactions of the British Mycological Society **74**(3): 515-523.
- Evans, H. C. (1981). "Witches' broom disease: a case study." Cocoa Growers Bulletin **32**: 5-19.
- Evans, H. C. (1980). "Pleomorphism in *Crinipellis pernicioso*, Causal Agent of Witches' Broom Disease of Cocoa." Transactions of the British Mycological Society **74**(3): 515-523.
- Evans, H. C., and Bastos, C. N. (1979). "Uma avaliação do ciclo de vida da vassoura de bruxa (*Crinipellis pernicioso*) do cacauero." Fitopatologia Brasileira **4**(1): 515-523.
- Ewing, B., *et al*¹. (1998). "Base-calling of automated sequencer traces using phred. I. Accuracy assessment." Genome Res. **8**: 175-185.
- Ewing, B., *et al*². (1998). "Base-calling of automated sequencer traces using phred. II. Error probabilities." Genome Res. **8**: 186-194.
- Felsenstein, J. Phylip (phylogeny inference package) version 3.6. (1993). Department of Genetics, University of Washington, Seattle.
- Gehring, W. J., *et al.* (1994). "Homeodomain-DNA recognition." Cell **78**: 211-223.
- Giasson L., S. C. A., *et al.* (1989). "Cloning and comparison of the Aa mating-types alleles of the basidiomycete *Schizophillum commune*." Mol Gen Genet **218**: 72-77.
- Gibas, C., and Jambeck, P. (2001). Desenvolvendo bioinformática. Rio de Janeiro, Ed. Campus.
- Gillissen, B., *et al.* (1992). "A two-component regulatory system for self/non-self recognition in *Ustilago maydis*." Cell **68**: 647-657.
- Graur, D., and Li, W. H. (1999). Fundamentals of Molecular Evolution. U.S.A., Sinauer Associates, Inc.
- Griffith, G. W., and Hedger J. N. (1994). "The breeding biology of biotypes of the witches' broom pathogen of cocoa, *Crinipellis pernicioso*." Heredity **72**: 278-289.
- Hagen, D. C., *et al.* (1986). "Evidence that yeast STE3 genes encode a receptor for the peptide pheromone a-factor: Gene sequence and implication for the structure of the presumed receptor." Proc. Natl. Acad. Sci. USA **83**: 1418-1422.
- Hedger, J. N., *et al.* (1987). "Variability of populations of the witch's broom disease of cocoa (*Crinipellis pernicioso*)." Transactions of the British Mycological Society **88**: 533-546.
- Herskowitz, I. (1988). "Life cycle of the budding yeast *Saccharomyces cerevisiae*." Microbiol Rev **52**: 536-553.

- Herskowitz, I. (1989). "A regulatory hierarchy for cell specialization in yeast." Nature **342**: 749-757.
- Herskowitz, I. (1995). "MAP kinase pathways in yeast, for mating and more." Cell **80**: 187-197.
- Hiscock, S. J. and U. Kues (1999). "Cellular and molecular mechanisms of sexual incompatibility in plants and fungi." Int Rev Cytol **193**: 165-295.
- Isaya, G., *et al.* (1995). "Mammalian mitochondrial intermediate peptidase - structure-function analysis of the new homolog from *Schizophyllum commune* and relationship to thimet oligopeptidases." Genomics **28**: 450-461.
- Iwasa, M., *et al.* (1998). "The two nuclei in the dicaryon of the homodasidiomycete *Coprinus cinereus* change position after each conjugate division." Fungal Genet Biol **23**: 110-116.
- Jackson, C. L., and Hartwell, L.H. (1990). "Courtship in *S. cerevisiae*, both cell types choose mating partners by responding to the strongest pheromone signal." Cell **63**: 1039-1051.
- Kämper, J., *et al.* (1995). "Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomain proteins in *Ustilago maydis*." Cell **81**: 73-83.
- Kertesz-Chaloupkova, K., *et al.* (1998). "Blue light overrides repression of asexual sporulation by mating type genes in the basidiomycete *Coprinus cinereus*." Fungal Genet Biol **23**: 95-109.
- Kothe, E. (1996). "Tetrapolar fungal mating types: sexes by thousands." FEMS Microbiol **18**: 65-87.
- Kronstad, J. W., and S. A. Leong (1990). "The b mating-type locus of *Ustilago maydis* contains variable and constant regions." Genes Dev. **4**: 1384-1395.
- Kues, U. (2000). "Life history and developmental processes in the basidiomycete *Coprinus cinereus*." Microbiol Mol Biol Rev **64**(2): 316-53.
- Kües, U., and L. A. Casselton (1993). "The origin of multiple mating types in mushrooms." J. Cell Sci. **104**: 227-230.
- Kues, U., *et al.* (1994). "Two classes of homeodomain proteins specify the multiple a mating types of the mushroom *Coprinus cinereus*." Plant Cell **6**(10): 1467-75.
- Kues, U. and L. A. Casselton (1992). "Molecular and functional analysis of the A mating type genes of *Coprinus cinereus*." Genet Eng (N Y) **14**: 251-68.
- Kues, U., *et al.* (1994). "A chimeric homeodomain protein causes self-compatibility and constitutive sexual development in the mushroom *Coprinus cinereus*." Embo J **13**(17): 4054-9.
- Kues, U., *et al.* (1998). "The A mating type and blue light regulate all known differentiation processes in the basidiomycete *Coprinus cinereus*." Mol Gen Genet **260**(1): 81-91.
- Kues, U., *et al.* (2001). "The chromosomal region containing *pab-1*, *mip*, and the A mating type locus of the secondarily homothallic homobasidiomycete *Coprinus bilanatus*." Curr Genet **39**: 16-24.
- Kues, U., *et al.* (2001). "The chromosomal region containing *pab-1*, *mip*, and the A mating type locus of the secondarily homothallic homobasidiomycete *Coprinus bilanatus*." Curr Genet **39**(1): 16-24.
- Kues, U., and Liu, Y. (2000). "Fruiting body production in basidiomycetes." Appl Microbiol Biotechnol **54**: 141-152.
- Kues, U., *et al.* (1994). "A mating-type factors of *Coprinus cinereus* have variable numbers of specificity genes encoding two classes of homeodomain proteins." Mol Gen Genet **245**(1): 45-52.
- Kues, U., *et al.* (1992). "The combination of dissimilar alleles of the A α and A β gene complexes, whose proteins contain homeodomain motifs, determines sexual development in the mushroom *Coprinus cinereus*." Genes Dev. **4**: 568-577.
- Kurjan, J. (1992). "Pheromone response in yeast." Annu. Rev. Genet. **61**: 1097-1129.
- Kurjan, J. (1993). "The pheromone response pathway in *Saccharomyces cerevisiae*." Annu. Rev. Genet. **27**: 147-179.
- Lukens, L. H. Y., and G. May (1996). "Correlation of genetic and physical maps at the A mating type locus of *Coprinus cinereus*." Genetics **144**: 1471-1477.

- Magae, Y., *et al.* (1995). "Interaction of the Aalpha Y mating type and Z mating type homeodomains proteins of *Schizophyllum commune* detected by the two-hybrid systems." Biochim Biophys Acta **211**: 1071-1076.
- Marion, A. L., *et al.* (1996). "The Aa mating-type locus of *Schizophyllum commune*: structure and function of gene X." Curr Genet **29**: 143-149.
- May, G., *et al.* (1991). "Molecular analysis of the *Coprinus cinereus* mating type A factor demonstrates an unexpectedly complex structure." Genetics **128**(3): 529-38.
- Meinhardt, F. A., K. (1990). "Sex determination and sexual differentiation in filamentous." Crit Rev Plant Sci **9**: 329-341.
- Mutasa, E. S., *et al.* (1990). "Molecular organisation of an A mating type factor of the basidiomycete fungus *Coprinus cinereus*." Curr Genet **18**: 223-229.
- Nakayama, N., *et al.* (1985). "Nucleotide sequence of STE2 and STE3, cell-type specific sterile genes from *Sacharomyces cerevisiae*." Embo J **4**: 2643-2648.
- Nei, M. and Saitou, N. (1987). The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. Molecular Biology Evolution, **4**(4):406-425.
- Neto, E. D. *et al.*, (2000). "Shotgun sequencing of the human transcriptome with ORF expressed sequence tags." PNAS **97**: 3491-3496.
- Nicholas, K. B., *et al.* (1997). "GeneDoc: Analysis and Visualization of Genetic Variation." Embnet.News, **4**(2)
- Olesnicky, N. S., *et al.* (1999). "A constitutively active G-protein-coupled receptor causes mating self-compatibility in the mushroom *Coprinus*." Embo J **18**(10): 2756-63.
- Olesnicky, N. S., *et al.* (2000). "Self-compatible B mutants in *Coprinus* with altered pheromone-receptor specificities." Genetics **156**(3): 1025-33.
- Orchard, J. E. K. H. (1988). "Photosynthesis, carbohydrate translocation and metabolism and fungal tissues in cacao seedlings infected with *Crinipellis pernicioso*." Actas , Cocoa Producers Alliance: 325-330.
- O'Shea, S. F., *et al.* (1998). "A large pheromone and receptor gene complex determines multiple B mating type specificities in *Coprinus cinereus*." Genetics **148**(3): 1081-90.
- Papazian, H. P. (1951). "Incompatibility factors and a related gene in *Schizophyllum commune*." Genetics **36**: 441-459.
- Page, R. D. M. (1996). "TreeView. Na application to display phylogenetic trees on personal computer. Computer Applications in the Biosciences. **12**: 357-358.
- Parag, Y., and Y. Koltin (1971). "The structure of the incompatibility factors of *Schizophyllum commune*: constitution of three classes of B factors." Mol Gen Genet **112**: 43-48.
- Pardo, E. H., *et al.* (1996). "Multiple versions of the A mating type locus of *Coprinus cinereus* are generated by three paralogous pairs of multiallelic homeobox genes." Genetics **144**(1): 87-94.
- Pereira, J. L., *et al.* (1989). "Primeira ocorrência de vassoura-de-bruxa na principal região produtora de cacau do Brasil." Agrotrópica **1**(1): 79-81.
- Pertsemlidis A., and Fondom, J. W. (2001). "Having a BLAST with bioinformatics (and avoiding BLAST phemy)." Genome Biology **2**: 1-10.
- Puhalla, J. E. (1970). "Genetics studies of the b incompatibility locus of *Ustilago maydis*." Genet. Res. **16**: 229-232.
- Purdy, L. H. and Schmidt, R. A. (1996). "Status of Cacao Witches' Broom: Biology, Epidemiology and Management." Annual Review of Phytopathology(34): 573-594.
- Raper, J. R., *et al.* (1960). "The genetic structure of the incompatibility factors of *Schizophyllum commune*: the A factor." Proc. Natl. Acad. Sci. USA **46**: 833-842.
- Rapper, J. R. (1966). "Genetics of sexuality in higher fungi." Ronald Press.
- Raven, P. H., *et al.* (1996). Biologia Vegetal. R.J., Guanabara Koogan.

- Rincones, J., *et al.* (2003). "Electrophoretic karyotype analysis of *Crinipellis pernicioso*, the causal agent of witches' broom disease of *Theobroma cacao*." *Mycol Res* 107(Pt 4): 452-8.
- Sanger, F., *et al.* (1977). DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
- Schulz, B., *et al.* (1990). "The b alleles of *U. maydis*, whose combinations programme pathogenic development, code for polypeptides containing a homeodomain-related motif." *Cell* 60: 295-306.
- Scott, M. P., *et al.* (1989). "The structure and function of the homeodomain." *Biochim Biophys Acta* 989: 25-48.
- Shen, G. P., *et al.* (1996). "Cloning and characterization of a Schizophyllum gene with Abeta6 mating-type activity." *Curr Genet* 29: 136-142.
- Silva, S. D. V., and Matsuoka, K. (1999). "Histologia da Interação *Crinipellis pernicioso* em Cacaueiros Suscetível e Resistente à Vassoura-de-Bruxa." *Fitopatologia Brasileira* 24(1): 54-59.
- Simpson, A. J., *et al.* (2000). "The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis." *Nature* 406: 151-157.
- Singer, R. (1942). "A Monograph Study of the Genera *Crinipellis* and *Chaetochalatus*." *Lilloa* 8: 503.
- Spit, A., *et al.* (1998). "A role for heterodimerization in nuclear localization of homeodomain protein." *Proc. Natl. Acad. Sci. USA* 95: 6228- 6233.
- Stahel, G. (1915). *Marasmius perniciosus* Nov. Spec., Bulletin Department Van den Landbouw in Suriname.
- Stankis, M. M., *et al.* (1992). "The Aalfa mating locus of *Schizophyllum commune* encodes two dissimilar, multiallelic, homeodomain proteins." *Proc. Natl. Acad. Sci. USA* 89: 7169-7173.
- Stark, M. R., *et al.* (1999). "A trans-acting peptide activates the yeast a1 repressor by raising its DNA-binding affinity." *Embo J* 18: 1621-1629.
- Stell, F. (1928). *Witch-broom Disease of Cacao and its Control*. Trinidad & Tobago.
- Surzcki, S. J. (1990). A fast method to prepare random fragment sequencing libraries using a new procedure of DNA shearing by nebulization and electroporation. *The International conference on the status and future of research on the human genome. Genome II*. San Diego, CA: 51.
- Surzcki, S. J. (2000). *Basic methods in molecular biology*. New York, N.Y., Springer-Verlag.
- Talbot, N. (2001). *Molecular and Cellular Biology of Filamentous Fungi*. New York, Oxford.
- Thompson, J. D., *et al.* (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." *Nucleic Acids Res* 25: 4876-4882.
- Thorold, C. A. (1975). *Diseases of cocoa*. Oxford, Clarendon Press.
- Trevisan, S. D. P. (1996). "Mudanças no Sul da Bahia associadas à vassoura-de-bruxa do cacau." *Actas. Lagos, Cocoa Producer's Alliance*.
- Tymon, A. M., *et al.* (1992). "A fungal mating type protein that regulates sexual and asexual development contains a POU-related domain." *Embo J* 11(5): 1805-13.
- Urban, M., *et al.* (1996). "Identification of the pheromone response element in *Ustilago maydis*." *Mol. Gen. Genet.* 251: 31-37.
- Vaillancourt, L. J., *et al.* (1997). "Multiple genes encoding pheromones and a pheromone receptor define the bbeta1 mating type specificity in *Schizophyllum commune*." *Genetics* 146: 541-551.
- Varga, J. (2003). "Mating type gene homologues in *Aspergillus fumigatus*." *Microbiology* 149: 816-819.

- Venter, J. C., *et al* (1998). "Shotgun sequencing of the human genome." Science **280**: 1540-1542.
- Wendland, J., *et al.* (1995). "The mating type locus B alpha 1 of *Schizophyllum commune* contains a pheromone receptor and putative pheromone genes." Embo J **14**: 5271-5278.
- Yee, A. R., and J. W. Kronstad (1993). "Construction of chimeric alleles with altered specificity at the *b* incompatibility locus of *Ustilago maydis*." Proc. Natl. Acad. Sci. USA **90**: 664-668.
- Yue, C. L., *et al.* (1997). "The specificity determinant of the Y mating-type proteins of *Schizophyllum commune* is also essential for Y-Z protein binding." Genetics **145**: 253-260.