

JORGE FERNANDO PEREIRA

**CARACTERIZAÇÃO, DISTRIBUIÇÃO E ESTUDO DA ATIVIDADE DE
ELEMENTOS TRANSPONÍVEIS EM *Crinipellis pernicioso*, AGENTE
CAUSAL DA VASSOURA-DE-BRUXA NO CACAUEIRO (*Theobroma cacao*)**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para a obtenção do título de "Doctor Scientiae"

VIÇOSA
MINAS GERAIS - BRASIL
2005

RESUMO

PEREIRA, Jorge Fernando, D.S., Universidade Federal de Viçosa, dezembro de 2005. **Caracterização, distribuição e estudo da atividade de elementos transponíveis em *Crinipellis pernicioso*, agente causal da vassoura-de-bruxa no cacauero (*Theobroma cacao*)**. Orientadora: Marisa Vieira de Queiroz. Conselheiros: Elza Fernandes de Araújo e Sérgio Hermínio Brommonschenkel.

Neste trabalho, elementos transponíveis da *Classe I* e da *Classe II* foram caracterizados no genoma do fungo *Crinipellis pernicioso*, agente causal da enfermidade conhecida como vassoura-de-bruxa no cacauero. *C. pernicioso* apresenta uma alta variabilidade genética e elementos transponíveis têm sido relatados como um dos principais agentes responsáveis pela alta plasticidade e adaptabilidade apresentada por fungos fitopatogênicos. Buscas preliminares feitas no banco de dados do Projeto Genoma da Vassoura-de-Bruxa revelaram a presença de uma seqüência de DNA com similaridade a transcriptase reversa de elementos da *Classe I* pertencente ao grupo *gypsy/Ty3*. Esta seqüência foi amplificada em diferentes isolados de *C. pernicioso* que pertencem aos biótipos C, S e L e distintas regiões geográficas. Seqüenciamento de fragmentos amplificados revelaram a ocorrência de vários eventos de transição G:C para A:T, indicando a existência de um possível mecanismo de silenciamento tipo RIP. Análises de hibridização mostraram a presença de um alto número de cópias da região que codifica a transcriptase reversa e diferentes perfis de hibridização nos isolados dos biótipos C, S e L, indicando que os elementos que contém estas seqüências podem ter um importante papel na variabilidade de *C. pernicioso*. O fragmento de DNA contendo a seqüência de transcriptase reversa foi utilizado para o isolamento de fagos recombinantes em um banco genômico de *C. pernicioso*. Retrotransposons do grupo *gypsy/Ty3*, denominados *CpSaci1* e *CpSaci2*, foram caracterizados a partir de uma nova seqüência presente no banco de dados do Projeto Genoma da Vassoura-de-Bruxa e de um dos fagos recombinantes obtidos. *CpSaci1* possui 6.856 pares de bases (pb) e contém longas repetições terminais (LTRs) diretas de 423 pb que não apresentam regiões normalmente encontradas em LTRs de retrotransposons. Duas seqüências de leitura aberta (ORFs) foram

identificadas codificando peptídeos similares à GAG e à poliproteína POL contendo os domínios protease, transcriptase reversa, RNase H e integrase. A seqüência de DNA de *CpSaci2*, que foi parcialmente caracterizada, possui cerca de 82% de identidade com *CpSaci1*. Estas duas cópias apresentam inserções de bases e códons de parada na região estrutural indicando que se tratam de cópias inativas. Entretanto, análises por RT-PCR revelaram a presença de transcritos *CpSaci* normalmente expressos em *C. pernicioso* cultivado em meio mínimo, e que, sob condições de estresse nutricional, apresentaram uma maior expressão. *CpSaci* está presente em alto número de cópias no genoma dos biótipos C, S e L de *C. pernicioso*, sendo que isolados do biótipo C originados do estado da Bahia apresentam dois perfis de hibridização, e alguns fragmentos de DNA em comum com isolados da região amazônica. Por possuir alto número de cópias e ser ativo, o retrotransposon *CpSaci* provavelmente está relacionado com a geração de variabilidade genética em *C. pernicioso*. Além do retrotransposon *CpSaci*, elementos transponíveis da *Classe II* também foram caracterizados. As cópias denominadas *Boto1* e *Boto2* foram obtidas a partir do banco de dados do Projeto Genoma da Vassoura-de-Bruxa e a partir de um fago recombinante isolado de um banco genômico de *C. pernicioso*, respectivamente. *Boto1* é flanqueado por uma duplicação de 3 pb (TAA) e possui duas ORFs, uma que codifica uma transposase e outra que codifica uma proteína com baixa similaridade a um regulador da transcrição de plantas. As regiões estruturais que codificam as transposases de *Boto1* e *Boto2* são interrompidas por um intron de 53 pb com alto conteúdo A+T. As transposases de *Boto1* e *Boto2* são altamente similares à transposases de elementos *PIF-like*. Estas transposases são relacionadas à transposases de seqüências de inserção do grupo IS5 de bactérias e fazem parte de uma nova família de transposons da *Classe II* denominada *PIFIS5*. A superfamília *PIFIS5* é raramente encontrada em fungos, e como análises filogenéticas indicam que as transposases *Boto-like* formam um ramo evolutivo independente, estas seqüências parecem ter sido transferidas horizontalmente para o genoma de *C. pernicioso*. Transcritos *Boto-like* foram amplificados por RT-PCR em *C. pernicioso* normalmente cultivado em meio de cultura, mas não foram ativados por estresse nutricional. Este é o primeiro relato de elementos da superfamília *PIFIS5* em um fungo fitopatogênico.

ABSTRACT

PEREIRA, Jorge Fernando, D.S., Universidade Federal de Viçosa, December 2005. **Characterization, distribution and activity of transposable elements in *Crinipellis pernicios*, casual agent of the witches' broom disease in cocoa (*Theobroma cacao*).** Adviser: Marisa Vieira de Queiroz. Committee members: Elza Fernandes de Araújo and Sérgio Hermínio Brommonschenkel.

In this work, *Class 1* (RNA) and *Class 2* (DNA) transposable elements were characterized in the genome of *Crinipellis pernicios*, the causal agent of the witches' broom disease of cocoa. *C. pernicios* presents a high genetic variability and transposable elements have been reported as the main agents responsible for the high plasticity and adaptability presented by phytopathogenic fungi. Once the genome of this fungus has been sequenced, preliminary searches in the database of the witches' broom genome project revealed the presence of a DNA sequence with similarity to the reverse transcriptase of *gypsy/Ty3*-like retrotransposons. This sequence was amplified in different *C. pernicios* isolates that belong to the C-, S-, and L-biotypes and to different geographical areas. The sequencing of the amplified fragments revealed the occurrence of several G:C to A:T transitions, indicating the existence of a possible RIP-like silencing mechanism. Southern analyses showed a high copy number of the reverse transcriptase gene with different profiles among the C-, S- and L- biotypes, indicating that the elements that contain these sequences may have an important role in *C. pernicios* genetic variability. The DNA fragment containing the reverse transcriptase gene sequence was used to isolate different recombinant phages from a *C. pernicios* genomic library. *Gypsy/Ty3*-like retrotransposons, named *CpSaci1* and *CpSaci2*, were characterized from a new sequence present in the sequencing project database and from one recombinant phage. *CpSaci1* contains 6.856 bp and long direct terminal repeats (LTRs) of 423 bp that do not present regions usually found in LTRs. Two open reading frame (ORFs) were identified and they code proteins with similarity to GAG and to a POL polyprotein showing protease, reverse transcriptase, RNase H, and integrase domains. Partially characterized DNA sequence of *CpSaci2* possesses about

82% of identity to *CpSaci1*. The copies *CpSaci1* and *CpSaci2* present punctual base insertions and stop codons in the structural region indicating that they are inactive copies. However, RT-PCR analyses revealed the presence of *CpSaci*-like transcripts expressed in *C. pernicioso*, presenting a higher expression under nutritional stress conditions. *CpSaci* is present in high copy number in the genome of the C-, S- and L-biotypes. The C-biotype isolates from Bahia state, Brazil, present two hybridization profiles and some DNA fragments similar to those of the isolates from the Amazon region. Once *CpSaci* presents a high copy number and activity, it is probably related to the generation of genetic variability in *C. pernicioso*. Besides the *CpSaci* retrotransposon, a *Class II* transposable element was also characterized. Two copies, named *Boto1* and *Boto2*, were obtained from the genome sequencing database and from a recombinant phage obtained from a *C. pernicioso* genomic library, respectively. *Boto1* is flanked by a 3 bp duplication (TAA) and possesses two ORFs, one that codes a transposase and another that codes a protein with low similarity to a plant transcription regulator. The structural regions that code the transposases of *Boto1* and *Boto2* are interrupted by a 53 bp intron with high A+T content. The aminoacid sequence of *Boto1* and *Boto2* transposases are highly similar to plant *PIF*-like transposases. These transposases are related to those of bacterial insertion sequences of the group IS5 and they are members of a new superfamily of *Class II* transposons named *PIFIS5*. Once this superfamily is rarely found in fungi and phylogenetic analyses indicate that *Boto1* and *Boto2* transposases present an independent evolutionary branch, it seems that these sequences were horizontally transferred to *C. pernicioso* genome. Although *Boto*-like transcripts were amplified in *C. pernicioso* normally cultivated in culture medium, they were not activated by nutritional stress. This is the first report of elements of the *PIFIS5* superfamily in a phytopathogenic fungus.