Universidade de São Paulo
Escola Superior de Agricultura "Luiz de Queiroz"

Estudo de elementos transponíveis em *Puccinia psidii* Winter, agente causal de ferrugem em *Eucalyptus* spp.

Sarina Tsui

Dissertação apresentada para obtenção do título de Mestra em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba
2015

Sarina Tsui
Bacharela e Licenciada em Ciências Biológicas

Estudo de elementos transponíveis em *Puccinia psidii* Winter, agente causal de ferrugem em *Eucalyptus* spp.

Orientadora:
Profa. Dra. **MARIA CAROLINA QUECINE-VERDI**

Dissertação apresentada para obtenção do título de Mestra em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba
2015

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Tsui, Sarina
Estudo de elementos transponíveis em *Puccinia psidii* Winter, agente causal de ferrugem em *Eucalyptus* spp. / Sarina Tsui. - - Piracicaba, 2015.
133 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura “Luiz de Queiroz”.

1. Biotrófico 2. Genoma 3. Fitopatógeno 4. Eucalipto 5. LTR-Retrotransposon I. Título

CDD 634.9734
T882e

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

Aos meus pais **Man-chak** e **Die-xin**,
pelo incentivo, lições e amor incondicional.

Ao meu amado **Yeu**, pelo afeto, compreensão, paciência
e pela maravilhosa companhia.

Dedico

A toda a minha família pelo apoio
e carinho de sempre.

Especialmente aos meus queridos irmãos
Wen, Jian, Vivian e Carol
por todos os conselhos
e pela confiança depositada em mim ...

Aos meus sobrinh@s
Kevin, Dalwin, Winnie, Yance, Jasmine,
Alick, Rebeca e Justin
por TU-DO!

Ofereço

AGRADECIMENTOS

Agradecimento especial à Profa. Dra. **Aline Aparecida Pizzirani-Kleiner** (*in memorian*) pela oportunidade de primeiro estágio, confiança e apoio em meus primeiros passos na vida científica.

À amiga e orientadora (sempre!) Profa. Dra. **Maria Carolina Quecine-Verdi** pelos ensinamentos, motivação, cumplicidade, animação, amor pela ciência e por sempre confiar em meu trabalho. Sou uma pessoa realmente sortuda por ter sua adorável companhia e amizade durante todos estes anos, estão aqui, os meus agradecimentos mais sinceros.

Ao Prof. Dr. **João Lúcio de Azevedo**, pelo admirável exemplo profissional, ensinamentos, amizade e presença motivadora.

Aos amigos do Laboratório de Genética de Micro-organismos "Prof. João Lúcio de Azevedo", especialmente àqueles que já passaram pelo laboratório pelo auxílio e colaboração durante os primeiros anos de vida científica. Aos **novos companheiros** de bancada, RUCAS e "coffee – coffees": Andréa, Bia, Bruna, Bruno, Carol Silva, Daniel, Fábio, Isaneli, Jaqueline, Jéssica, Joelma, Mariana, Patrícia, Paula, Rafael, Renata, Renatinha, Tiago e Zezo. Sem vocês os meus dias não teriam sido os mesmos (especialmente aos que se identificam com "**corações peludos**").

À equipe do Laboratório de Genética Molecular e de Micro-organismos, BIOAGRO, UFV-MG, especialmente à Profa. Dra. **Marisa Vieira de Queiroz**, Dr. **Mateus F. Santana** e Msc. **Cleydson Silva**, pela receptividade e boa vontade em colaborar no auxílio em técnicas relacionadas à análise de elementos transponíveis.

A **todos** do Laboratório de Genética Molecular e de Micro-organismos, BIOAGRO, UFV-MG, pelo coleguismo e auxílios durante a minha estadia em Viçosa, especialmente: Casley, Jonathan, Kaliane, Lucas, Tiago e Vanessa.

Ao **pessoal** do Laboratório Max Feffer pela amizade e ensinamentos, Ana Paula P., Ana Paula S., Andressa, Fabrício, Fernando, Felipe, Flávia, Hana, Ilara, Ivan, Janaína, Juliana, Lívia, Mariana, Maria Letícia, Marisângela, Nathália B., Simone e Thaís. Agradeço especialmente ao Prof. Dr. Carlos Alberto Labate, Dra. Mônica Labate e Lívia Maria Franceschini pelas dicas profissionais, colaboração e disponibilização dos aparelhos.

Um "Urráh!" para os amigos que trabalham/ram no "grupo Puccinia", **Andressa**, **Felipe**, **Jaqueline**, **Mariana** e **Nathália**. O trabalho só foi possível graças à ajuda de cada um de vocês. Obrigada pela companhia nas belas manhãs de poda e pelas discussões infindáveis, isso acrescentou muito ao meu conhecimento sobre o assunto.

Ao **pessoal** do Laboratório de Genética de Micro-organismos - grupo “Genomics” pela amizade e ensinamentos, especialmente à Profa. Dra. Cláudia Vitorello pelo exemplo profissional, presença motivadora e disponibilização de equipamentos.

Às amigas **Hana** e **Marisângela,** pelo convívio diário durante o estudo infindável para o processo de seleção, bem como a agradável companhia e parceria ao longo das disciplinas. Sem vocês o meu mestrado teria sido em vários tons de cinza. Claro, como poderia esquecer do **Felipe** (o da Mari) pelas boas risadas e amizade.

Às minhas amigas mais “gatíneas” deste mundo: **Maíra**, **Bianca**, **Fernanda**, **Elaine** e **Roberta** pelo companheirismo, pelas dicas, pelas boas risadas, pela alegria, pelo esforço em sempre manter o contato e especialmente por serem o que são desde sempre.

Às minhas amigas e *roommates* mais queridas e engraçadas: **Bianca**, **Bruna** e **Mariane** por toda companhia, amizade, momentos de descontração e conselhos. A nossa combinação deu “*100% match*”, foi um prazer enorme dividir o Lar, doce Lar com vocês.

À minha **família** por me apoiar em decisões importantes, por compreender os momentos de ausência, pelo esforço em dar todo o suporte que precisei ao longo destes anos e claro pelo amor e carinho em todos os momentos.

Ao meu amado **Yeu**, por ser o que você é desde sempre, me proporcionando momentos únicos de felicidade e segurança com sua agradável companhia. Obrigada por acreditar em mim!

Aos **funcionários** do Departamento de Genética pela disposição em dar suporte, amizade, convivência e cumplicidade.

A todos os **professores** do programa de pós-graduação “Genética e Melhoramento de Plantas” que contribuíram positivamente para o meu aprendizado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (processo 2014/04803-3) e à CAPES pelo apoio financeiro.

“O-BRI-GA-DA”
É uma palavra curta demais
para expressar os meus sentimentos de gratidão
por cada um de vocês nesta jornada!

“A essência do conhecimento consiste em
aplicá-lo, uma vez possuído.”
Confúcio

SUMÁRIO

RESUMO....13
ABSTRACT....15
LISTA DE FIGURAS....17
LISTA DE TABELAS....23
1 INTRODUÇÃO....25
1.1 Revisão bibliográfica....27
1.1.1 O gênero *Eucalyptus* e o setor florestal....27
1.1.2 A ferrugem em *Eucalyptus* spp. 30
1.1.3 *Puccinia psidii* Winter 33
1.1.4 Elementos transponíveis....38
1.1.5 Importância de elementos transponíveis na composição do genoma 40
Referências....43
2 ANÁLISE *IN SILICO* DE ELEMENTOS TRANSPONÍVEIS EM *Puccinia psidii*....49
Resumo....49
Abstract....50
2.1 Introdução....51
2.2 Materiais e Métodos....52
2.2.1 Material Biológico e dados preexistentes....52
2.2.2 Mineração de sequências relacionadas a elementos transponíveis....53
2.2.3 Anotação direta dos *contigs* de *P. psidii* MF-1....53
2.2.4 Análise da integridade de elementos transponíveis de *P. psidii*....54
2.2.5 Anotação manual de *contigs* classificados como elementos da Classe I – LTR Retrotransposons de *P. psidii* – MF1....55
2.2.6 Análise comparativa da distribuição de TEs em *P. psidii* MF-1 e de outras espécies de *Puccinia*....55
2.2.7 Análise da integridade de elementos transponíveis de *Puccinia* spp. 57
2.2.8 Análise comparativa da proteína transcriptase reversa 59
2.3 Resultados 60
2.3.1 Mineração de sequências relacionadas a elementos transponíveis....60
2.3.2 Anotação direta dos *contigs* de *P. psidii* MF-1....64
2.3.3 Análise da integridade de elementos transponíveis de *P. psidii*....66

2.3.4 Anotação manual de *contigs* classificados como elementos da Classe I – LTR Retrotransposons de *P. psidii* – MF1 71
2.3.5 Análise comparativa da distribuição de TEs em *P. psidii* MF-1 e de outras espécies de *Puccinia* 74
2.3.6 Análise da integridade de elementos transponíveis de *Puccinia* spp. 77
2.3.7 Análise comparativa da proteína transcriptase reversa 87
2.4 Discussão 88
Referências 99
3 UTILIZAÇÃO DE FERRAMENTAS MOLECULARES NO ESTUDO DE ELEMENTOS TRANSPONÍVEIS EM *Puccinia psidii* 107
Resumo 107
Abstract 108
3.1 Introdução 109
3.2 Materiais e Métodos 110
3.2.1 IRAP na análise de variabilidade genética em populações de *P. psidii* 110
3.2.2 Análise molecular de possíveis “TEs complexos” no genoma de *P. psidii* MF-1 111
3.2.2.1 Desenho de *primers* específicos 111
3.2.2.2 Validação de *primers* específicos 113
3.2.2.3 Clonagem dos *amplicons* de Retrotransposons 114
3.2.2.3.1 Purificação e adição de cauda poli-adenina 114
3.2.2.3.2 Clonagem em vetor pGEM®T Easy Vector 114
3.2.2.3.3 Extração de plasmídeo pGEM®T Easy clonado 115
3.3 Resultados 115
3.3.1 IRAP na análise de variabilidade genética em populações de *P. psidii* 115
3.3.2 Análise molecular de possíveis “TEs complexos” no genoma de *P. psidii* MF-1 116
3.3.2.1 Validação de *primers* específicos 116
3.3.2.2 Obtenção de *amplicons* de Retrotransposons 118
3.3.2.3 Clonagem dos *amplicons* de Retrotransposons 121
3.3.2.4 Extração de plasmídeo pGEM®T Easy clonado 122
3.4 Discussão 122

Referências 127
4 CONSIDERAÇÕES FINAIS 131

RESUMO

Estudo de elementos transponíveis em *Puccinia psidii* Winter, agente causal de ferrugem em *Eucalyptus* spp.

A cultura do eucalipto apresenta grande importância no setor florestal no mundo. No Brasil, 70% da área florestal plantada é destinada ao eucalipto. Entretanto, a ferrugem das mirtáceas, também conhecida como ferrugem do eucalipto, causada pelo fungo *Puccinia psidii* Winter, afeta o enorme potencial produtivo das plantações de eucalipto. A biologia, mecanismos de patogenicidade e genética desse patógeno são pouco conhecidos, apesar de sua importância para o setor florestal. Os elementos transponíveis (TEs) são sequências de DNA com a capacidade de migrar e influenciar a organização, integridade e evolução do genoma hospedeiro. O presente trabalho teve como principal objetivo estudar os TEs presentes no genoma de *P. psidii*, combinando ferramentas *in silico* e moleculares. A classificação dos elementos transponíveis no genoma de *P. psidii* MF-1 foi realizada utilizando *contigs* previamente minerados e remontados, bem como sem seleção prévia dos *contigs*, por meio do programa *RepeatMasker*. Ambas estratégias apontaram o predomínio de elementos da Classe I - LTR Retrotransposons no genoma de *P. psidii* MF-1. O resultado condiz com a composição de TEs em fungos fitopatogênicos descrita na literatura. Algumas análises *in silico*, como verificação de integridade e anotação manual de sequências proteicas foram também realizadas para alguns *contigs* classificados como TEs. Assim, foi possível observar a presença de sequências conservadas pertencentes à região *pol* em LTR Retrotransposons. Além disso, as análises permitiram inferir sobre a existência de TEs híbridos no genoma parcialmente sequenciado de *P. psidii* MF-1. Paralelamente foi também realizada uma análise comparativa entre os TEs presentes nos genomas de *P. graminis*, *P. striiformis, P. triticina* e *P. psidii*. Observou-se que *P. graminis, P. striiformis* e *P. triticina* apresentam maior frequência de elementos da Classe II, do tipo DNA Transposons ao contrário de *P. psidii*, com maior frequência de elementos da Classe I. Interessantemente, a quantidade de elementos desconhecidos foi similarmente alta para todos os quatro genomas avaliados. Este tipo de análise é muito importante, pois evidencia a grande quantidade de famílias de TEs novas a serem descobertas. Elas podem estar potencialmente relacionadas ao silenciamento de genes importantes à virulência destes patógenos. A utilização de TEs no estudo de diversidade genética entre populações é bastante comum. A técnica molecular IRAP foi utilizada para acessar a diversidade entre populações de *P. psidii* originárias de três híbridos de *Eucalyptus* spp., goiabeira, jambeiro e jabuticabeira. No entanto, esta técnica não se mostrou eficiente para detectar polimorfismos existentes entre estas populações. A anotação de TEs foi difícil devido à observação de sequências de elementos sobrepostas, o que podem representar híbridos de TEs, entretanto, visando a confirmação desta hipótese por meio da PCR, alguns *contigs* serão sequenciados e mais estudos devem ser realizados para a continuação desta confirmação. Os resultados apresentados neste trabalho são inéditos e representam uma etapa crucial no entendimento de TEs em fungos do gênero *Puccinia*, em especial do patógeno *P. psidii* para o desenvolvimento de melhores mecanismos de controle de ferrugem.

Palavras-chave: Biotrófico; Genoma; Fitopatógeno; Eucalipto; LTR-Retrotransposon

ABSTRACT

Deciphering the transposable elements in *Puccinia psidii* Winter, causal agent of rust on *Eucalyptus* spp.

The culture of eucalyptus has great importance worldwide in forestry sector. In Brazil, 70% of cultivated forest area is intended for *Eucalyptus*. However, the eucalyptus potential productive has been affected by rust disease, caused by the fungus *Puccinia psidii* Winter. Despite its importance to brazilian and world forest sector, the knowledge of biology, genetic and pathogenic mechanisms of this pathogen is scarce. Transposable elements (TEs) are mobile DNA fragments that influence the organization and development of the host genome. These elements have the ability to move within host genome, and their insertion can cause a wide spectrum of mutations in their hosts. This study aims to decipher the TEs in *P. psidii* genome by combining *in silico* and molecular tools. *P. psidii* MF-1 TEs classification was performed automatically, through *RepeatMasker* software, being observed a predominance of Class I – LTR Retrotransposons in *P. psidii* MF-1 genome. This result is consistent with the TEs composition described in phytopathogenic fungi. Some *in silico* analysis, as integrity and manual annotation of conserved protein sequences from TEs were carried out with *P. psidii* MF-1 contigs classified as transposable elements. The presence of conserved sequences belonging to *pol* region in LTR Retrotransposons was observed. Furthermore, these analysis allowed the inference of hybrid TEs in *P. psidii* MF-1. At the same time, a comparative analysis of TEs present in other *Puccinia* genomes and *P. psidii* MF-1 was also performed. The *P. graminis*, *P. striiformis* and *P. triticina* genomes have higher frequency of Class II - DNA Transposons unlike the results found for *P. psidii*. Interestingly, the number of unknown elements was similarly high for all genomes. This type of analysis is very importante because it shows a great number of potential new TEs families to be discovered. They may be potentially related to the virulence gene silencing of these pathogens. Using TEs for study the fungal genetic diversity is quite common. The IRAP technique was used to access the diversity among *P. psidii* populations originated from three *Eucalyptus* spp. hybrids, guava, syzigium and jabuticaba. However, this technique was not efficient to detect existing polymorphisms between these populations. TEs annotation was labored due to the existence of overlapping elements, which may represent hybrids TEs. PCR tool was used to confirm some sequences annotated as hybrids and more studies are needed to confirm this hyphotesis. The results presented in this study are novel and is a crucial step in understanding the genetic of *P. psidii* pathogen for further improvements of rust control mechanisms.

Keywords: Biotrophic; Genome; Phytopathogen; Eucalyptus; LTR-Retrotransposon

LISTA DE FIGURAS

Figura 1.1 - (**a**) Distribuição e áreas de plantação florestal de *Eucalyptus*, *Pinus* e outras espécies florestais por regiões no Brasil. Os plantios florestais se concentram principalmente nos estados de Minas Gerais, São Paulo, Mato Grosso, Paraná e Santa Catarina; (**b**) Distribuição geográfica das empresas produtoras de papel e celulose. Concentração na região sul e sudeste Fonte: Relatório IBÁ (2015) 29

Figura 1.2 - Levantamento das áreas de risco de ataque do *Eucalyptus* spp. por *P. psidii.* Regiões marcadas com cor preta representam áreas com alto risco de incidência de ferrugem quando comparado com as áreas em cinza. Fonte: BOOTH; OLD; JOVANOVIC, 2000 31

Figura 1.3 - Presença da ferrugem em planta susceptível de *Eucalyptus grandis.* O patógeno infecta tecidos mais tenros da planta com idade até dois anos, como primórdios foliares, inflorescências, gemas, terminais de galhos e haste principal. Imagem fotografada na Estação Experimental de Ciências Florestais de Itatinga, SP 32

Figura 1.4 - O possível ciclo de vida de *Puccinia psidii*, onde os seguintes estágios podem ser encontrados: aécio (estágio I), urédia (estágio II), télia (estágio III) e basídia (estágio IV). Fonte: Adaptado de GLEN *et al*., 2007 35

Figura 1.5 - Processo de transposição de retrotransposons (Classe I) e de DNA transposons (Classe II). Fonte: Adaptado de LISCH, 2013 39

Figura 1.6 - Modificações estruturais e funcionais que podem ser causadas por elementos transponíveis (TEs). Os TEs podem estar envolvidos em mudanças que incluem a “quebra” (*knockout*) da função, introdução de novas funções,

modificações na estrutura dos genes, mobilização e rearranjo de fragmentos genéticos e até mesmo no silenciamento epigenético de genes. Éxons estão representados por retângulos na cor cinza, *enhancers* são círculos de cor verde, repressores são círculos de cor vermelha, TEs são representados como triângulos de cor roxa, término do transposon são triângulos amarelos e proteínas exaptadas são círculos azuis. C^{meth} é 5-metilcitosina. Fonte: Adaptado de LISCH, 2013 42

Figura 2.1 - Anotação funcional do GO termos relacionados às funções moleculares anotadas pelo programa Blast2GO. Os *contigs* são provenientes da montagem pelo programa *Newbler Assembler* 2.6. O gráfico representa os três principais GO termos relacionados aos processos de transposição no genoma, GO:0003682, GO:0000166 e GO:0003677. Os valores foram gerados pelo Blast2GO com corte de GO termos com menos de 5 sequências, independentemente do nível do termo anotado 60

Figura 2.2 - Estrutura básica de um retrotransposon da família Copia. Regiões TSR duplicadas; 5' LTR e 3' LTR; a região *pol* contém os domínios PR (protease), RT (transcriptase reversa), RH (RNase H) e integrase (IN) e *gag* responsável pela proteína do capsídeo viral. Fonte: Adaptado de WICKER *et al*., 2007 85

Figura 2.3 - Representação de LTR - Retrotransposon por meio de sequência do *Supercontig* 2.111 da espécie *P. triticina,* o sítio alvo (TSR – *Target Site Repeat*) está em itálico, as longas terminações repetitivas (LTRs) estão representadas em cinza, a região *gag* e domínios PR (protease), RT (transcriptase reversa), RH (RNase H) e IN (integrasse) da região *pol* estão sublinhados. A fita 3' → 5' é aquela que contém os

nucleotídeos para tradução das proteínas de transposição, neste elemento 86

Figura 2.4 - Agrupamento da região correspondente a transcriptase reversa presentes nos 12 *contigs* de *P. psidii* MF-1 (classificados como Gypsy e Copia) com outras regiões desta proteína de outros elementos de transposição. A análise foi realizada pelo método UPGMA, o alinhamento das proteínas foi realizado por Clustal W com os parâmetros de penalidade padrão. A construção filogenética por UPGMA se deu com o método *Bootstrap* com 1.000 réplicas e no modelo de Poisson. Os números abaixo e acima dos nós indicam a porcentagem em que cada ramo apareceu em análise.....88

Figura 3.1 - Amplificação de DNA de *P. psidii* MF-1 em duas concentrações, 20 ng/µL e 40 ng/µL utilizando as duas combinações de *primers* para a análise de polimorfismo por IRAP. (**M**) Marcador 100 pb (Fermentas); (**1** e **2**) DNA de *P. psidii* MF-1 nas concentrações 20 ng/µL e 40 ng/µL, respectivamente, utilizando a combinação de *primers* CIIRAP 1 e CIIRAP 4; (**3**) DNA de fungo controle utilizando a combinação de *primers* CIIRAP 1 e CIIRAP 4; (**4** e **5**) DNA de *P. psidii* MF-1 nas concentrações 20 ng/µL e 40 ng/µL, respectivamente, utilizando a combinação de *primers* CIIRAP 2 e CIIRAP 4 e (**6**) DNA de fungo controle utilizando a combinação de *primers* CIIRAP 2 e CIIRAP 4. A seta branca indica o mesmo perfil de bandas entre diferentes concentrações do DNA de *P. psidii* MF-1. 115

Figura 3.2 - Amplificação de DNA de *P. psidii* MF-1 e demais *P. psidii* de outros hospedeiros Myrtaceae utilizando as duas combinações de *primers* para a análise de polimorfismo por IRAP. (**M**) Marcador 1Kb (Fermentas); (**1** a **7**) amplificação utilizando a combinação CIIRAP 1 e CIIRAP 4 DNA de (respectivamente) populações de *P. psidii* provenienetes de

três híbridos de *Eucalyptus, E. grandis* x *urophylla* (PEGU); *E. grandis* x *dunni* (clone 1366 [PEGD]) e *E. grandis* x *brassiana* (clone MA2000 [PEGB]); DNA de goiabeira – *Psidium guajava* (PGuava), *jabuticabeira* – *Myrciaria cauliflora* (Pjab) e jambeiro – *Syzygium jambos (*PSyz); *P. psidii* MF-1 na concentração de 40 ng/µL. (**1'** a **7'**) amplificação utilizando a combinação CIIRAP 2 e CIIRAP 4 para o mesmo material na mesma sequência. A seta branca indica a presença de bandas nos diferentes materiais de *P. psidii* analisados 116

Figura 3.3 - Sequências de *contigs* 00075, 00088 e 00095 que originaram os três *amplicons* Pp_1.1, Pp_4.1 e Pp_5.1, respectivamente. Eles foram utilizados para o desenho de *primers* específicos de retrotransposons do tipo Copia. (**a**) Sequência de *contig*00075, com tamanho total de 4.682 pb, a região cinza da sequência representa os 2.083 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_75 6F e Pp_Copia_75 10R e o *amplicon* estimado apresenta 4.424 pb. (**b**) Sequência de *contig*00088, com tamanho total de 4.528 pb, a região cinza da sequência representa os 2.409 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_88 1F e Pp_Copia_88 1R e o *amplicon* estimado apresenta 2.273 pb. (**c**) Sequência de *contig*00095, com tamanho total de 4.474 pb, a região cinza da sequência representa os 3.391 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_95 2F e Pp_Copia_95 2R e o *amplicon* estimado apresenta 3.028 pb 119

Figura 3.4 - Amplificação de DNA genômico de *P. psidii* MF-1 (20 ng/µL) utilizando as combinações de *primers* específicos para retrotransposons: (**M**) Marcador 1Kb (Fermentas); (**1**) Pp_Copia75 6F e Pp_Copia75 10R (combinação 1.1 referente ao *contig*00075); (**2**) Pp_Copia88 1F e Pp_Copia88 1R (combinação 4.1 referente ao *contig*00088) e (**3**) Pp_Copia95 2F e Pp_Copia95 2R (combinação 5.1 referente ao *contig*00095). Os tamanhos aproximados das bandas obtidas para os *amplicons* de Pp_1.1, Pp_4.1 e Pp_5.1 foram estimados em aproximadamente 6.000 pb, 2.500 pb e 3.000 pb, respectivamente de acordo com o marcador molecular 1 Kb 120

Figura 3.5 - (**a**) Placa de meio LB suplementado com antibiótico ampicilina (100 µg.mL^{-1}) e 2 µL/mL de meio de X-Gal 2%, as colônias brancas são os recombinantes positivos (inserção de *amplicon* de retrotransposons em célula de *E. coli* DH5-α). Colônias de coloração azul referem-se à células eletro-competentes de *E. coli* DH5-α que foram transformadas com o vetor pGEM®T Easy sem o inserto de interesse. (**b**) Recombinante pGEM::Pp_5.1, selecionado e repicado em meio LB suplementado com antibiótico ampicilina (100 µg.mL^{-1}) para análises posteriores 121

Figura 3.6 - Extração de plasmídeo realizada com *QIAprep® Spin Miniprep Kit* (QIAgen), para o vetor pGEM®T Easy clonado com três insertos diferentes de *P. psidii* MF-1: (**1**) Pp_1.1 (*contig*0075), (**2**) Pp_4.1 (*contig*00088) e (**3**) Pp_5.1 (*contig*00095). (M) Marcador 1Kb (Fermentas); (M1 a M3) Régua molecular λ (lambda) nas concentrações de 25 ng/µL, 50 ng/µL e 100 ng/µL, respectivamente. Os tamanhos dos vetores clonados são desconhecidos 122

Figura 3.7 - Estratégia de amplificação de *primers* específicos para a técnica IRAP. Os *primers* IRAP alinham-se às regiões LTR dos elementos do tipo LTR Retrotransposons, tanto no sentido 5' →3' como 3' →5'. O produto esperado corresponde ao intervalo entre um elemento LTR Retrotransposon e o próximo 123

LISTA DE TABELAS

Tabela 2.1 - Informações sobre os patógenos utilizados como referência neste trabalho 56

Tabela 2.2 - Listagem de *supercontigs* selecionados para análise de integridade por meio do programa *LTR-Finder* de cada uma das três espécies de *Puccinia* utilizadas para comparar com *P. psdii* MF-1 58

Tabela 2.3 - Mineração de *contigs* possivelmente relacionados a elementos transponíveis..... 61

Tabela 2.4 - Resultados obtidos de acordo com parâmetros utilizados no programa *CLC Genomic Workbench* 6.0 para as três tentativas de *de novo assembly* 63

Tabela 2.5 - Resultados obtidos pela classificação no software do *RepeatMasker* dos *contigs* gerados pelo *De Novo Assembly* com os parâmetros da Tentativa II. As sequências foram comparadas com o banco de dados *TE library* (GIRI). 64

Tabela 2.6 - Resultados obtidos pela classificação no *software* do *RepeatMasker* dos 167.538 *contigs* gerados pelo *De Novo Assembly* 2.6 do genoma parcialmente sequenciado de P*. psidii* MF-1 65

Tabela 2.7 - Classificação de 20 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*..... 67

Tabela 2.8 - Classificação de 10 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*..... 69

Tabela 2.9 - Anotação manual de 25 *contigs* de *P. psidii* MF-1 realizada pelo programa BLASTX pelo NCBI e Uniprot 72

Tabela 2.10 - Informações gerais dos genomas de *P. graminis; P. triticina* e *P. striiformis* retiradas do banco de dados do *Puccinia - Group* do *Broad Institute* e *P. psidii*..... 75

Tabela 2.11 - Distribuição e anotação de principais classes de elementos transponíveis em *P. graminis; P. triticina* e *P. striiformis* retiradas do banco de dados do *Puccinia - Group do Broad Institute* e *P. psidii* MF-1 76

Tabela 2.12 - Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT..... 78

Tabela 2.13 - Domínios proteicos encontrados no elemento Copia presente no Supercontig 2.111 de *P. triticina*, localização dos domínios juntamente a seus respectivos valores de *E-value* dos domínios “blastados” e seus códigos de acesso 87

Tabela 3.1 - Lista de *contigs* de *P. psidii* MF-1 utilizados para o desenho de *primers* específicos para elementos Copia e Gypsy e suas respectivas sequências oligonucleotídicas 112

Tabela 3.2 - (**a**) Resultado da validação preliminar de primers específicos em gradiente de temperatura (45ºC a 60ºC) utilizando Taq DNA Polimerase. (**b**) Resultado da validação de primers específicos utilizando Taq DNA Polimerase do tipo “elongase” em duas temperaturas diferentes de anelamento 117

1 INTRODUÇÃO

"*Feliz aquele que transfere o que sabe*
e aprende o que ensina."
Cora Coralina

No setor florestal brasileiro mais de 70% da área plantada corresponde à cultura de eucalipto (IBÁ, 2015). No entanto, o eucalipto está sujeito ao ataque de patógenos ao longo de seu ciclo de desenvolvimento, uma atenção especial é dada à ferrugem, doença causada pelo fungo basidiomiceto biotrófico *Puccinia psidii* Winter.

O fungo *P. psidii* pertence à ordem Pucciniales (Uredinales), é originário da América do Sul e ataca diversas espécies de mirtáceas das Américas, bem como o eucalipto, o que contribuiu para que o mesmo se espalhasse pelo mundo todo (GLEN *et al*., 2007). Casos de ferrugem relacionados à *P. psidii* vem sendo registrados em diversos locais no mundo como nas Américas, nas Ilhas do Caribe e do Pacífico, Ásia (GLEN *et al*., 2007), África do Sul (ROUX *et al*., 2013) e recentemente na Austrália (CARNEGIE *et al*., 2010).

Apesar de ter sido descrito como causador da ferrugem em eucalipto já na primeira metade do século XX, pouco se sabe sobre a sua interação molecular com *Eucalyptus* spp. (LEITE, 2012) e outros hospedeiros, bem como sobre os mecanismos de resistência do hospedeiro (QUECINE *et al*., 2014).

Os elementos transponíveis (TEs) são sequências de DNA repetidas, dispersas no genoma do hospedeiro com a capacidade de migrar. Estão presentes em muitos genomas eucariotos e podem ocupar até 50% de todo o genoma (MARTIN *et al*., 2008; SANTANA, 2009). Os TEs são capazes de influenciar a organização, a integridade e a evolução do genoma hospedeiro, além de regular a expressão de genes em diferentes espécies (SANTANA, 2009). Assim como mecanismos de silenciamento gênico, uma vez que a movimentação deles no genoma hospedeiro pode gerar rearranjos, novos alelos e influenciar a expressão gênica. A detecção e caracterização de elementos transponíveis em patógenos possibilitam maior conhecimento sobre a organização e a evolução do genoma desses organismos (MARTIN *et al*., 2008; SANTANA, 2009).

Considerando a importância do fungo *P. psidii* e o conhecimento escasso sobre a sua genética e dos mecanismos moleculares envolvidos durante a interação com o hospedeiro. O presente trabalho teve como objetivo geral a caracterização e melhor

entendimento do papel dos elementos transponíveis presentes no genoma do fitopatógeno *P. psidii*.

Assim, inicialmente foi realizada a classificação automática de TEs presentes no genoma parcialmente sequenciado de *P. psidii*, linhagem monopustular MF-1 (LEITE, 2012; QUECINE *et al*., 2012). Também foi realizada a verificação da integridade dos elementos anotados, sendo que algumas sequências foram manualmente anotadas. Paralelamente foi realizada a análise comparativa da presença de TEs com espécies do gênero *Puccinia,* com destaque para a verificação de integridade de elementos Copia e a comparação de sequências de transcriptase reversa entre *P. psidii* MF-1, que também foi realizada utilizando sequências de diversos fungos fitopatogênicos. Os Retrotransposons foram avaliados quanto à sua possível aplicação como marcador molecular para estudos de diversidade de populações de *P. psidii.*

A abordagem deste trabalho é inédita e de suma importância no fornececimentode informações sobre a possível fração majoritária do genoma de *P. psidii,* constituída por elementos repetitivos, visando um melhor entendimento da biologia e evolução do fitopatógeno. Para que futuramente, a complexa interação planta-patógeno seja melhor entendida visando o desenvolvimento de novas estratégias de controle. Além disso, essas sequências podem ser utilizadas como marcador de polimorfismo molecular em estudos de diversidade genética populacional.

1.1 Revisão bibliográfica

1.1.1 O gênero *Eucalyptus* e o setor florestal

Estima-se que a área mundial total de florestas é superior a 4 bilhões de hectares, sendo que o Brasil está entre os 5 países, que juntos, apresentam mais da metade dessa área. A área de florestas plantadas abrange 264 milhões de hectares, ou seja, 7% da floresta total mundial, dentre os quais, aproximadamente 75% são cultivados devido à produção de madeira, fibra, combustível e produtos não-madeireiros (FAO, 2010).

A maioria das florestas tropicais é composta por Angiospermas, que originam madeiras “duras” (*hardwoods*) (GUTMANIS, 2008). Na indústria madeireira, destaca-se o gênero *Eucalyptus*, da família Myrtaceae, composta por mais de 700 espécies de arbustos e árvores de grande porte, que se dividem em subgêneros e híbridos (HARDER *et al*., 1965 apud GUTMANIS, 2008).

As espécies conhecidas do gênero *Eucalyptus* possuem como seu principal centro de origem a Austrália e regiões próximas. Atualmente, o eucalipto pode ser encontrado no mundo inteiro, em regiões de clima tropical e subtropical (entre as latitudes de 40°N e 45°S). As principais espécies de interesse econômico utilizadas em escala comercial são *E. grandis, E. camaldulensis, E. urophylla, E. saligna, E. tereticornis, E. globulus e E. viminalis* (GUTMANIS, 2008).

A expansão da cultura do eucalipto no Brasil se iniciou em 1903 em Rio Claro (São Paulo) com o objetivo de produzir dormentes para ferrovias, lenha para locomotivas e postes para eletrificação. Navarro de Andrade desenvolveu trabalhos experimentais que demonstraram as vantagens do plantio do gênero em relação às espécies nativas (MOURA; GARCIA, 2000), mas o interesse pelo eucalipto como matéria-prima para a produção de celulose surgiu somente na década de 50 (GUTMANIS, 2008).

Destaca-se que atualmente, as plantações de eucalipto no Brasil são uma das mais produtivas do mundo, com média de aumento anual de 40 m^3 ha^{-1} ano^{-1} (GONÇALVES *et al*., 2013; SILVA *et al*., 2013), esta produtividade está relacionada às condições ambientais que refletem a alta adaptação fisiológica do material genético em vários ambientes (STAPE *et al*., 2008; SILVA *et al*., 2013). A produtividade brasileira é resultado de duas políticas públicas adotadas pelo governo federal ao incentivar a produção florestal. A primeira ocorreu de 1965-1987 quando as

plantações de eucalipto podiam utilizar o programa federal de incentivos fiscais. Atualmente, o setor conta com o PNF (Programa Nacional de Florestas) do Ministério do Meio Ambiente (Decreto nº 3420/2000, Brasil, 2000). Evidenciando assim, a importância incontestável do setor florestal no Brasil (IBÁ, 2015).

As estimativas do ano de 2014 mostram que a área brasileira de florestas plantadas atingiram 7,74 milhões de hectares (0,9% do território), houve um aumento de 1,8% em relação a 2013 (IBÁ, 2015). Neste cenário o plantio de *Eucalyptus* representa 71,9% (5,56 milhões de hectares) de toda área florestal plantada no país. Além disso, o setor de florestas plantadas arrecadou em 2014 o valor de R$ 10,23 bilhões em tributos (0,8% da arrecadação nacional) e foram registrados 610 mil empregos totais (IBÁ, 2015).

Em 2014 foi registrado um aumento da produção do setor, destacando-se os principais produtos florestais destinados à exportação: papel (10,40 milhões de toneladas), painéis de madeira reconstituída (7,98 milhões de m^3) e a celulose (16,46 milhões de toneladas). A Europa e a China são os principais importadores de papel e celulose, enquanto os Estados Unidos lideram a importação de celulose, papel, painéis e madeira serrada (IBÁ, 2015).

Os plantios florestais de *Eucalyptus* e *Pinus* se concentram principalmente nos estados de Minas Gerais, São Paulo, Mato Grosso do Sul, Paraná e Santa Catarina (Figura 1.1a). Isso se deve à localização das indústrias de celulose e papel nas regiões Sul e Sudeste (Figura 1.1b) (IBÁ, 2015).

O uso de eucalipto para a produção de madeira tem algumas vantagens como a alta diversidade das espécies e híbridos utilizados. Fato que permite sua utilização e fornecimento de matéria-prima para a produção de celulose, papel, chapas, lâminas, compensados, aglomerados, dormentes, moirões, postes, lenha, carvão vegetal, madeira serrada e processada, caixotaria, estruturas para construção civil, produção do óleo essencial para indústria química e farmacêutica (cosméticos e medicamentos naturais), assim como na apicultura (criação de abelhas com ferrão) e ornamentação (GUTMANIS, 2008).

(**a**) (**b**)

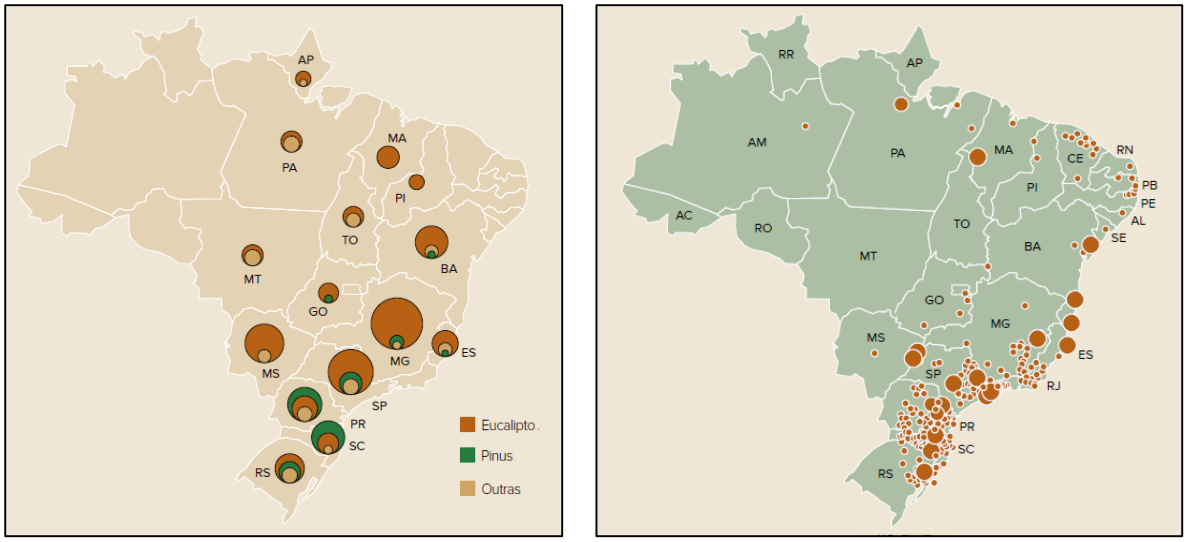


Figura 1.1 – (**a**) Distribuição e áreas de plantação florestal de *Eucalyptus*, *Pinus* e outras espécies florestais por regiões no Brasil. Os plantios florestais se concentram principalmente nos estados de Minas Gerais, São Paulo, Mato Grosso, Paraná e Santa Catarina; (**b**) Distribuição geográfica das empresas produtoras de papel e celulose. Concentração na região sul e sudeste Fonte: Relatório IBÁ (2015)

Do total de 7,74 milhões de hectares de árvores plantadas no Brasil, 34% pertencem a empresas do segmento de celulose e papel. Seguido por 26,8% ocupados por plantios florestais para comercialização da madeira *in natura.* A produção nacional de celulose em 2014 (fibra curta – originária de eucalipto) foi de 16,46 milhões de toneladas, isto faz com que o Brasil ocupe o quarto lugar no *ranking* dos países produtores de celulose de todos os tipos e é o primeiro produtor mundial de celulose de eucalipto (IBÁ, 2015).

A liderança do Brasil na produtividade de celulose se deve as suas condições climáticas e de solo que favorecem o desenvolvimento deste gênero arbóreo, resultando em uma taxa de crescimento maior, e, portanto um índice de produtividade superior. A produção de 1,5 milhão de toneladas de celulose por ano se dá em 140 mil hectares, um quinto da área necessária na Escandinávia para a mesma produtividade (IBÁ, 2015).

O produto interno bruto (PIB) do setor brasileiro de árvores plantadas cresceu 1,7% em 2014, a expansão do volume de exportações de celulose (12,6%) exerceu importante papel nesse desempenho. O setor de árvores plantadas apresenta crescimento 17 vezes maior do que o PIB brasileiro (0,1%), comprovando a importância do setor para a economia nacional. Assim cada hectare de árvore

plantada adicionou R$ 7,8 mil ao PIB nacional em 2014, em comparação ao complexo soja (referência nacional), que adicionou R$ 4,9 mil/ano por hectare plantado e a pecuária com R$ 2,7 mil/ano (IBÁ, 2015).

A espécie *E. grandis* W. Hill ex Maiden é uma das mais importantes nas plantações em todo o mundo (HARWOOD, 2011; SILVA *et al*., 2013) e apresenta uma excelente base genética. Apesar de ser importante para as indústrias florestais pelo ótimo crescimento e formato de caule, muitos dos genótipos dessa espécie e seus híbridos apresentam alta suscetibilidade à ferrugem, principalmente no estado de São Paulo, onde o clima favorece o desenvolvimento do fungo *Puccinia psidii* (FURTADO *et al*., 2008; SILVA *et al*., 2013).

1.1.2 A ferrugem em *Eucalyptus* spp.

O nome ferrugem é dado às doenças causadas por basidiomicetos pertencentes à ordem Pucciniales, esses patógenos são considerados parasitos obrigatórios. Visto que poucas espécies, como o fungo *Melampsora epitea*, podem ser mantidas em condições laboratoriais por culturas axênicas (CARVALHO Jr. *et al*., 1998; APARECIDO, 2001). Em plantações de eucaliptos, a ferrugem é causada pelo fungo *P. psidii* Winter, doença extremamente devastadora e impactante especialmente em regiões tropicais e subtropicais (GLEN *et al*., 2007).

A preocupação com os danos causados pela ferrugem no Brasil e no mundo tem crescido consideravelmente, devido ao aumento do cultivo de espécies que tem aumentado sua importância econômica, tais como, goiaba, uvaia, pitanga e jabuticaba, destinadas à produção de sucos, doces e geleias. Entretanto, no Brasil o maior problema tem sido em relação a sua capacidade de infectar Myrtaceae não nativas como o *Eucalyptus* spp. (COUTINHO *et al*., 1998; FIGUEIREDO, 2001).

O fungo é atualmente considerado um dos patógenos de maior importância na cultura de *Eucalyptus* spp.. Além disso representa uma ameaça em emergência para espécies de mirtáceas nativas e não nativas por todo mundo. Pode acarretar prejuízos econômicos severos em diversas culturas e trazer riscos à conservação da biodiversidade vegetal, especialmente em locais cuja vegetação nativa é majoritariamente formada por Myrtaceae (GLEN *et al*., 2007; CARNEGIE; LIDBETTER, 2012).

O primeiro caso conhecido de ferrugem em eucalipto (*E. citriodora*) ocorreu em 1912 e foi registrado em 1944 (GONÇALVES, 1929 apud GRAÇA *et al*., 2013; JOFFILY, 1944; COUTINHO *et al.*, 1998). O primeiro grande surto da ferrugem ocorreu no Espírito Santo em 1973, causando grandes perdas em viveiros e plantios de *E. grandis* com menos de 18 meses de idade (FERREIRA, 1981;1983). Posteriormente, ataques esporádicos e severos atingiram plantações e viveiros em diversos estados brasileiros (FERREIRA, 1983). A doença ganhou destaque em São Paulo a partir de 1990 (CAMARGO *et al*., 1997 apud TAKAHASHI, 2002).

Booth, Old e Jovanovic (2000) realizaram um levantamento das maiores áreas de risco de ferrugem *P. psidii* na região Neotropical (Figura 1.2), pode-se observar que o Brasil apresenta a maior região potencialmente atingida pelo fungo *P. psidii*, principalmente no Sudeste e Sul do país. Silva *et al*. (2013) analisaram a severidade de ferrugem em todo o estado de São Paulo e demonstram que áreas mais ao sul do estado são as que apresentam maior grau de severidade da doença.

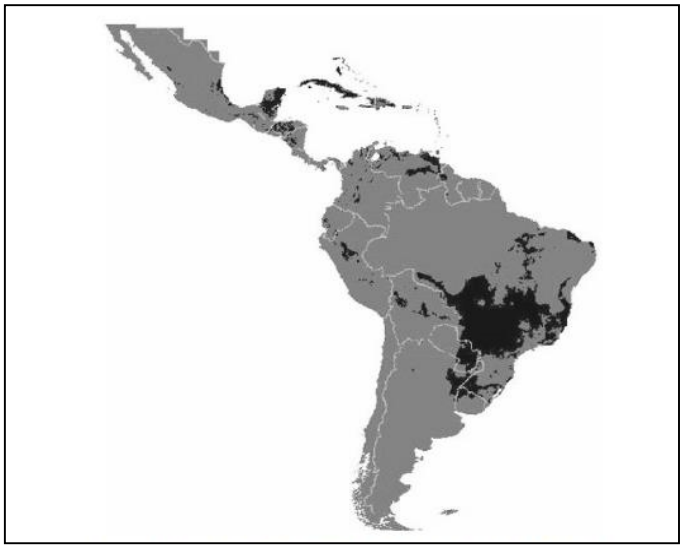


Figura 1.2 – Levantamento das áreas de risco de ataque do *Eucalyptus* spp. por *P. psidii.* Regiões marcadas com cor preta representam áreas com alto risco de incidência de ferrugem quando comparado com as áreas em cinza. Fonte: BOOTH; OLD; JOVANOVIC, 2000

Os danos que a ferrugem causa nas plantações dependem do ambiente, da época do ano, da idade do plantio e dos genótipos utilizados. Takahashi (2002) estimou danos entre 19,79% a 41,09% em termos de produção de volume de madeira, enquanto Santos (2006) verificou um dano médio de 48,3% e Masson (2009) um dano médio de 27,08% em plantações com 19 meses de idade.

A infecção do fitopatógeno se restringe a locais mais tenros da planta com idade até dois anos, como primórdios foliares, inflorescências, gemas, terminais de galhos e haste principal (Figura 1.3). Em culturas jovens, principalmente em viveiros, o patógeno pode atacar brotações recém-emitidas dos troncos, tornando o vegetal impróprio para o plantio (JOFFILY, 1944; APARECIDO, 2001). Em condições de campo, o fungo prejudica o desenvolvimento das plantas, resultando na redução da altura e em casos mais graves a morte dos hospedeiros mais suscetíveis (MORAES *et al*., 1982; APARECIDO, 2001).

Figura 1.3 – Presença da ferrugem em planta susceptível de *Eucalyptus grandis.* O patógeno infecta tecidos mais tenros da planta com idade até dois anos, como primórdios foliares, inflorescências, gemas, terminais de galhos e haste principal. Imagem fotografada na Estação Experimental de Ciências Florestais de Itatinga, SP

Diferentes espécies comerciais de *Eucalyptus* sofrem severas ameaças com *P. psidii*, principalmente as de grande importância na indústria de papel, como *E. cloeziana, E. grandis, E. phaeotricha* e *E. pellita* que são altamente suscetíveis. A espécie *E. citriodora* utilizada na indústria de essências também apresenta susceptibilidade ao patógeno (FIGUEIREDO, 2001).

A primeira manifestação da ferrugem é a presença de pequenas pontuações verde-claras ou amarelo alaranjadas em hospedeiros suscetíveis, após um ou dois dias essas pontuações tornam-se pústulas, as quais ficam carregadas de uma massa de esporos que adquirem coloração amarelo gema de ovo, sendo essa a fase de diagnóstico da doença. Dentro de alguns dias, seguem-se as infecções secundárias devido à disseminação dos esporos pelo orvalho ou chuva (FERREIRA, 1983).

Quanto maior o tempo de infecção na planta, os esporos tornam-se marrons e pode até formar-se uma mistura de mais de um tipo de esporo. Dependendo do caso, é possível a formação de um arco roxo ao redor do sítio de infecção em folhas e

brotos. Após uma ou duas semanas os órgãos atacados apresentam áreas hipertrofiadas e coloração ferrugínea, podendo a doença ser confundida com outros patógenos. Em hospedeiros resistentes, ocorre a indução de uma reação de hipersensibilidade, com a formação de lesões necróticas sem esporulação, porém, dependendo do nível de resistência, pústulas podem ser formadas ao longo dessas lesões (FERREIRA, 1983; GLEN *et al.*, 2007).

Com relação ao controle da doença em *Eucalyptus* spp., o controle químico é utilizado majoritariamente em viveiros de mudas. Em campo o plantio de variedades resistentes têm sido a principal forma de controle, isso se deve, principalmente pela inviabilidade econômica do tratamento químico em campo (APARECIDO, 2001).

Aparentemente uma das formas de controle eficaz é a utilização de variedades resistentes (GLEN *et al*., 2007). No entanto, frequentemente a resistência é quebrada e a preocupação com o ambiente favorece o desenvolvimento de técnicas alternativas como o controle biológico por ação hiper-parasitária sobre as estruturas esporíferas de *P. psidii* (SANTOS *et al*., 1998 apud APARECIDO, 2001).

1.1.3 *Puccinia psidii* Winter

O fungo *P. psidii* Winter (Basidiomycota, Pucciniales) foi descrito pela primeira vez no Brasil em 1884 por George Winter no sudeste brasileiro como agente causador de ferrugem em goiabeira (*Psidium guajava*), o termo “psidii” refere-se ao nome genérico da goiaba (GRAÇA *et al*., 2013). Assim o agente causal de ferrugem em goiabeira e eucalipto foram considerados coespecíficos por serem morfologicamente idênticos (JOFFILY, 1944; GRAÇA *et al*., 2013).

A hipótese amplamente difundida de que o patógeno de goiabeira passou a infectar o gênero *Eucalyptus* por um processo conhecido como *host jump*, ou seja, o patógeno se move entre dois hospedeiros separados por uma grande distância taxonômica (STUKENBROCK; MCDONALD, 2008 apud GRAÇA *et al*., 2013) nunca foi criticamente testada para confirmar a sua veracidade (GRAÇA *et al*., 2013).

A história evolutiva de *P. psidii* é pouco conhecida, assim como sua origem e muitas são as perguntas a respeito de seu ciclo de vida e sistema de reprodução. Destaca-se ainda que, sua abrangência de hospedeiros é considerada incomum em espécies causadoras de ferrugem (GRAÇA *et al*., 2013; CRISTANCHO *et al*., 2014). Aparentemente, o patógeno se originou em diferentes espécies da família Myrtaceae,

especialmente da subfamília Myrtoidea na América do Sul (GLEN *et al*., 2007; MORIN *et al*., 2012).

Com o passar do tempo o patógeno tornou-se bastante agressivo em algumas espécies de interesse econômico, como *Eucalyptus* spp. (FERREIRA, 1983; COUTINHO *et al*., 1998; TESSMANN *et al*., 2001; ALFENAS *et al*., 2004). A principal limitação nos estudos deste patógeno deve-se ao seu comportamento obrigatoriamente biotófico, ou seja, ele cresce somente em células vivas da planta hospedeira, fato que dificulta a criação de hipóteses e então respostas sobre o mecanismo de infecção e outras condições biológicas associadas a este fungo (FIGUEIREDO; PASSADOR, 2008).

Até o presente momento, informações sobre o ciclo de vida das ferrugens tropicais e subtropicais ainda são escassas, visto que a ênfase maior é dada a espécies de clima temperado (FIGUEIREDO; PASSADOR, 2008). O ciclo de vida de *P. psidii* é alvo de controvérsias entre diferentes autores, devido ao pleomorfismo (capacidade de adquirir diversas formas de acordo com o período do ciclo de vida, reprodutivo ou ainda de acordo com as condições ambientais) e à complexidade dos ciclos de vida das ferrugens (MORIN; TALBOT; GLEN, 2014).

O ciclo atualmente aceito é apresentado por Glen *et al.* (2007), o qual considera *P. psidii* como uma ferrugem autoécia, ou seja, que completa seu ciclo de vida em um único hospedeiro e macrocíclica com ciclo incompleto, devido à ausência do estágio de espermagônio. Assim, os seguintes estágios podem ser encontrados: aécio (estágio 1), urédia (estágio 2), télia (estágio 3) e basídia (estágio 4). Apesar de admitir o estágio de aécio, nenhum hospedeiro aecial foi identificado (Figura 1.4) (COUTINHO *et al.*, 1998; GLEN *et al.*, 2007).

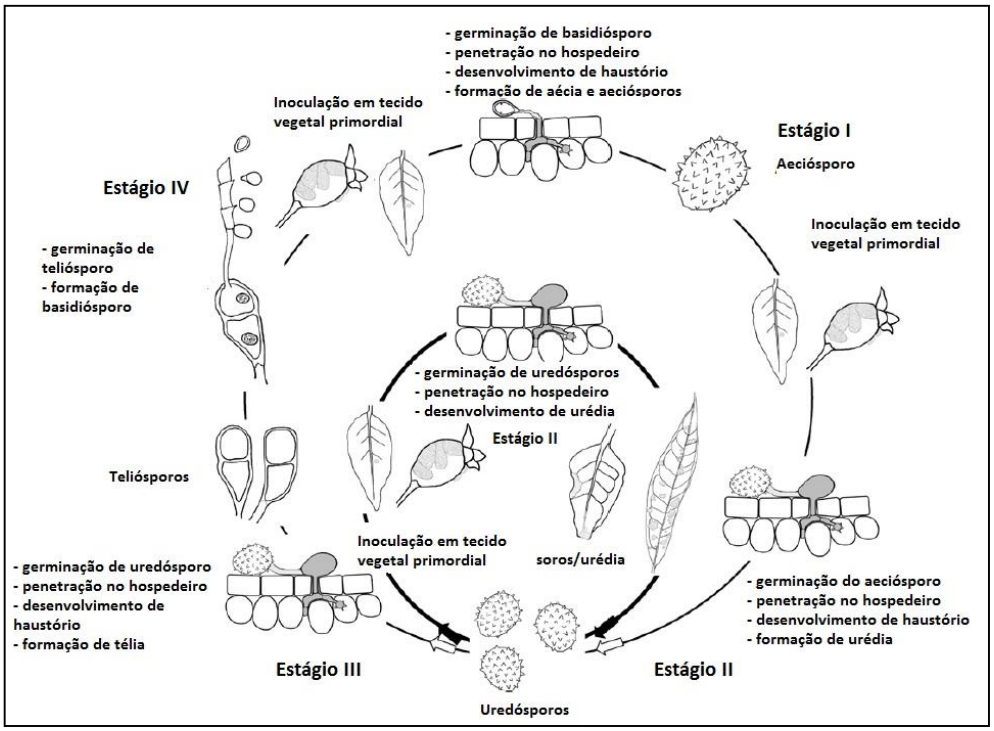


Figura 1.4 - O possível ciclo de vida de *Puccinia psidii*, onde os seguintes estágios podem ser encontrados: aécio (estágio I), urédia (estágio II), télia (estágio III) e basídia (estágio IV). Fonte: Adaptado de GLEN *et al*., 2007

A respeito dos aspectos gerais do ciclo de vida de ferrugens, existem ferrugens macrocíclicas completas, constituídas por cinco estágios e também são conhecidas ferrugens microcíclicas. Além disso, o ciclo pode depender de um único hospedeiro (autoécio) ou somente se completar através de dois hospedeiros distintos (heteroécio) (FIGUEIREDO; PASSADOR, 2008).

Espécies causadoras de ferrugem apresentam diversas estruturas esporogênicas e esporos com diferentes morfologias e funções, o entendimento destas só foi possível, estudando o macrociclo, sendo distinguidos em estrutura esporogênica e esporo, respectivamente: espermagônio/ espermácias, aécio/ aeciósporo, urédia/ uredósporos, télio/ teliósporos e basídia/ basidiósporos (FIGUEIREDO; PASSADOR, 2008).

A especificidade de *P. psidii* quanto a seus hospedeiros foi estudada por inoculações artificiais com amostras desse patógeno provenientes dos mais diferentes espécies da família Myrtaceae, sendo assim, evidenciada a variação fisiológica e/ou patogênica nas populações do patógeno (FERREIRA 1981; CASTRO *et al*., 1983; COUTINHO; FIGUEIREDO 1984; COELHO; ALFENAS; FERREIRA, 2001; APARECIDO; FIGUEIREDO; FURTADO, 2003; GRAÇA *et al*., 2013).

Graça *et al*. (2013) analisaram a diversidade genética de *P. psidii* proveniente de uma série de hospedeiros como jabuticabeira – *M. cauliflora*; goiabeira – *P. guajava*; pitanga – *E. uniflora*; goiabeira-brasileira – *P. guineense*; jambeiro – *S. jambos* e jambolão – *S. cumini* no Brasil e Uruguai, os autores rejeitaram a hipótese de origem de *P. psidii* no gênero *Eucalyptus* no Brasil via patógeno de goiabeira.

Existem pelo menos dois biotipos pelas análises realizadas, sendo um associado à *Eucalyptus* e outro a goiabeira, mais estudos são necessários para determinar se estes biotipos são *forma speciales* ou espécies emergentes (GRAÇA *et al*., 2013). Apesar da relação hospedeiro-específica apresentada pelas populações estudadas de *P. psidii*, Graça *et al*. (2013) ressaltam o cuidado com a ferrugem causada por este patógeno, sendo a mesma uma grande ameaça às espécies da família Myrtaceae ainda não detectadas como hospedeiros (COUTINHO *et al*., 1998).

Durante muitos anos os fitopatologistas australianos temiam a possível invasão do fungo *P. psidii* sensu lato (s.l.) na Austrália, devido à grande gama de hospedeiros suscetíveis, cerca de 50% de todos os táxons da família Myrtaceae se encontram na Austrália (GLEN *et al*., 2007). Sabe-se atualmente que 129 espécies de 32 gêneros em 9 tribos são considerados susceptíveis ao *P. psidii*, destas, 86 são espécies endêmicas australianas (GLEN *et al*., 2007; MORIN *et al*., 2012).

O fungo *P. psidii* foi detectado na Austrália pela primeira vez no estado de *New South Wales* em abril de 2010 (CARNEGIE *et al*., 2010), a origem e a maneira como ele foi introduzido são desconhecidos até o presente momento (CARNEGIE *et al*., 2010; MORIN *et al*., 2012). Inicialmente o fitopatógeno foi detectado em espécies nativas da Austrália, como a *Agonis flexuosa,* sendo classificado como *Uredo rangelli*, no entanto, foi observado o desenvolvimento de teliósporos completamente compatíveis aos de *P. psidii* sensu stricto (s.s.) e os autores julgaram ideal considerar o fitopatógeno como *P. psidii* sensu lato (s.l.). Entretanto, a relação taxonômica de *P. psidii* e *U. rangelli* está sendo ainda estudada (CARNEGIE *et al*., 2010; MORIN *et al*., 2012).

Após a invasão de *P. psidii* na Austrália em três hospedeiros nativos (CARNEGIE *et al*., 2010), espécies importantes para a indústria florestal foram testadas quanto à susceptibilidade a este “novo” patógeno. O patógeno *P. psidii* infectou e completou o seu ciclo de vida nas seguintes espécies vegetais: *Eucalyptus agglomerata, E. cloeziana, E. grandis, E. pilularis* e *Melaleuca quinquenervia*. Assim, as observações permitiram inferir que a ferrugem recém-chegada na Austrália é

semelhante ao *P. psidii* s.s. e possui uma gama abrangente de hospedeiros da família Myrtaceae, o que representa um grande risco aos ecossistemas naturais e à indústria florestal (CARNEGIE; LIDBETTER, 2012; MORIN *et al*., 2012).

Em maio de 2013, uma planta ornamental *Myrtus communis* na província de *KwaZulu-Natal*, África do Sul foi detectada apresentando sintomas típicos de infecção por *P. psidii*, além disso, a infecção provocou a morte do vegetal (ROUX *et al*., 2013). A análise filogenética deste patógeno com outras 17 sequências de *P. psidii* obtidas de outros hospedeiros por todo o mundo mostrou que todos os isolados de *P. psidii* agruparam em um único táxon diferente do grupo externo utilizado (*Puccinia cygnorum, P. hordei* e *P. recondita*) (ROUX *et al*., 2013).

Aparentemente não existe uma associação entre a presença ou ausência de sintomas de ferrugem e a relação filogenética dos táxons vegetais da família Myrtaceae, ou seja, o desenvolvimento deste fitopatógeno em uma espécie de um grupo não significa necessariamente que uma espécie ou gênero do mesmo grupo irá apresentar susceptibilidade (MORIN *et al*., 2012). Acredita-se na existência de especializações fisiológicas dentro da espécie *P. psidii*, pois o patógeno começou a parasitar o gênero *Eucalyptus* (proveniente da Austrália) desde o início de seu cultivo no Brasil (JOFFILY, 1944), sendo que, até recentemente, não se tinha conhecimento de ferrugem causada por *P. psidii* na Austrália (APARECIDO, 2001).

Ressalta-se que *P. psidii* afeta muitas espécies de *Eucalyptus* em todo o mundo, variando no seu grau de severidade, sendo *E. grandis* e seus híbridos os mais susceptíveis (ALFENAS *et al*., 2004; QUECINE *et al*., 2014). Em um dos poucos estudos sobre a variabilidade genética desse patógeno, Quecine *et al*. (2014) analisaram a diversidade de populações de *P. psidii* utilizando duas técnicas independentes de cultura, PCR-DGGE e T-RFLP. Ambas as técnicas revelaram que as populações de *P. psidii* provenientes de outras espécies de Myrtaceae (goiabeira, jambeiro e jabuticabeira) agruparam em um clado, enquanto que as populações originadas de *E. grandis* (clone D901 e clone 1600) foram agrupadas em outro clado. É provável que goiabeira, jambeiro e jabuticabeira sejam importantes na manutenção da diversidade do fitopatógeno enquanto este se adapta ao processo de infecção no gênero *Eucalyptus*. Assim, é evidente que o estudo deste patógeno é ainda incipiente, o que justifica a necessidade de estudos relacionados à genômica do mesmo.

1.1.4 Elementos transponíveis

A maioria dos organismos eucariotos apresenta grande porção do genoma composto por DNA repetitivo (RUIZ-SILVA, 2012). Enquanto algumas sequências repetidas possuem sua função esclarecida (DNA ribossomais, centroméricos e teloméricos), o papel da grande maioria do DNA repetitivo é desconhecido, e, até pouco tempo era considerado “DNA lixo” (SCHMIDT; HESLOP-HARRISON, 1998).

Alguns estudos mostram a importância destes para a manutenção e estrutura do genoma, como no processo de replicação do DNA (LI *et al*., 2002), recombinação (BIET *et al*., 1999), expressão gênica (LIU *et al.,* 2001) e diferenciação de cromossomos sexuais (POKORNA *et al*., 2011), ou seja, na organização estrutural e funcional do genoma em geral (RUIZ-SILVA, 2012).

Sequências genômicas repetidas com a capacidade de migrar dentro do genoma hospedeiro, afetá-lo e gerar plasticidade, são denominadas elementos transponíveis (TEs – *Transposable Elements*). Essas sequências mostram alta diversidade e em alguns genomas constituem a maior fração deles, em fungos, chegam a representar 3 a 21% do genoma (MARTIN *et al*., 2008; SANTANA, 2009).

Até o momento, com poucas exceções, a grande maioria dos genomas investigados apresentam elementos móveis (WALLAU, 2010). A grande distribuição e variabilidade dos TEs presentes nos seres eucarióticos sugere que sua origem se deu em eventos antigos, com íntimas ligações com os primeiros genomas de fungos analisados, principalmente em grupos de Ascomycota, Zigomycota e Basidiomycota (DABOUSSI, 1997).

Há uma classificação dos elementos transponíveis para fins didáticos e científicos em ordem hierárquica: classe, subclasse, ordem, superfamília, família e subfamília (DABOUSSI, 1997; WICKER *et al*., 2007; LISCH, 2013). Basicamente, a classificação é determinada pela presença ou ausência de um RNA intermediário durante o processo de transposição (Figura 1.5).

Na Classe I (*copy and paste*), a partir de uma cópia de RNA do transposon é sintetizado um cDNA, via transcriptase reversa (codificada pelo próprio elemento) capaz de se inserir em outro sítio alvo (Figura 1.5). Dessa forma, cada ciclo completo de transposição uma nova cópia é produzida, consequentemente, os retrotransposons (elementos da Classe I) são os maiores colaboradores na fração repetitiva do genoma (DABOUSSI, 1997; WICKER *et al*., 2007).

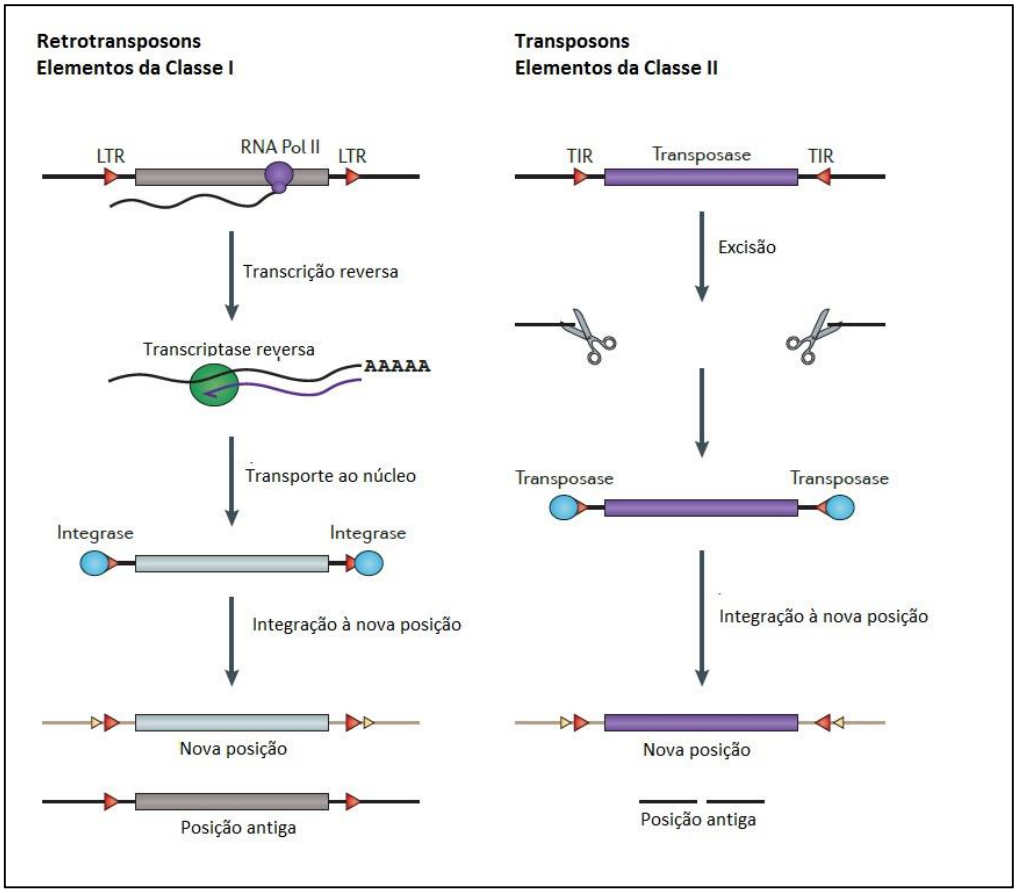


Figura 1.5 - Processo de transposição de retrotransposons (Classe I) e de DNA transposons (Classe II). Fonte: Adaptado de LISCH, 2013

Os elementos pertencentes à Classe II (*cut and paste*) sofrem uma excisão por meio da enzima transposase seguida de integração sem a intermediação de RNA (Figura 1.5) (DABOUSSI, 1997; WICKER *et al*., 2007). Apesar das diferenças nos mecanismos de transposição, a integrase de alguns elementos de RNA e a transposase de alguns elementos de DNA provavelmente tem uma origem em comum (WALLAU, 2010).

Os retrotransposons são divididos em cinco ordens, de acordo com o mecanismo de transposição e organização em LTR Retrotransposon, DIRS-*like*, Penelope-*like*, LINEs e SINEs (DABOUSSI, 1997; WICKER *et al*., 2007; LISCH, 2013).

Os LTR Retrotransposons (*Long Terminal Repeat*) apresentam longas terminações repetitivas (LTRs) diretas que flanqueiam uma região codificadora e são predominantemente encontradas em eucariotos. Esta ordem dispõe-se de duas ORFs (*Open Reading Frames*), *gag* e *pol*, a *gag* é responsável pela codificação de proteínas estruturais que formam uma partícula semelhante a vírus (proteínas do capsídeo)

enquanto a *pol* codifica uma protease, uma transcriptase reversa, uma RNase e uma integrase (HAVECKER *et al*., 2004; SANTANA, 2009).

Dentre os LTR Retrotransposons, destacam-se duas superfamílias, a *Gypsy* e a *Copia* que diferem na disposição de regiões que codificam a transcriptase reversa e a integrase na região *pol*. Em fungos fitopatogênicos, o grupo *Gypsy/Ty3* é o mais frequente entre os elementos móveis (DABOUSSI; CAPY, 2003).

Retrotransposons que não possuem longas terminações repetitivas são classificadas como não LTR e são principalmente representados por SINEs (*Short Interspersed Elements*) que não codificam as proteínas necessárias para a transcrição reversa, precisando de outros elementos móveis para sua transposição e LINEs (*Long Interspersed Elements*) capazes de codificar as proteínas necessárias para a transcrição reversa (RUIZ-SILVA, 2012).

O efeito da transposição de TEs depende muito do local em que ela ocorre no genoma, assim, danos severos devido à presença de elementos transponíveis no genoma podem ser amenizados ou evitados por mecanismos de silenciamento de TEs (*TE-silencing mechanism*) presentes em certos fungos, sendo RIP (*Repeat Induced Point Mutation* - mutação de ponto induzida por repetição) o mais comumente estudado (SANTANA, 2009).

Nos últimos anos, o mecanismo de silenciamento de TEs por RIP têm sido descrito em diversas classes de fungos, especialmente no subfilo Pucciniomycotina (Basidiomycota), em que o processo e o local de “hipermutação” são bem conservados (SANTANA *et al*., 2012; HORNS *et al*., 2012).

1.1.5 Importância de elementos transponíveis na composição do genoma

A presença de elementos transponíveis pode apresentar um importante papel na formação e mudança no genoma do hospedeiro ao longo de milhões de anos de evolução. Os TEs podem se integrar ao genoma por dois processos, a transferência vertical (TV) que se caracteriza pela herança do material genético que ocorre entre espécies ancestrais e seus descendentes e transferência horizontal (TH) em que espécies isoladas reprodutivamente são capazes de transferir material genético entre si (WALLAU, 2010).

Em fungos fitopatogênicos a detecção e a caracterização de elementos transponíveis permitem a obtenção de maior conhecimento sobre aspectos

fisiológicos, evolutivos e mecanismos de silenciamento gênico. Possibilitam ainda, a aplicação em sistemas de inativação gênica e podem ser utilizados como marcadores moleculares para traçar o perfil populacional de fitopatógenos (SANTANA, 2009; SANTANA *et al*., 2012).

Destaca-se que alguns fatores (estresses abióticos e bióticos) estão envolvidos com a taxa de atividade de transposição. Além disso, existem diversas estratégias de silenciamento, responsáveis pelo controle desses elementos, sendo as mais conhecidas a metilação, iRNA e RIP (SANTANA, 2009; SANTANA *et al*., 2012).

Há décadas sabe-se que os TEs são capazes de produzir uma gama de variações na expressão gênica e função de organismos eucarióticos, esta informação faz com que os cientistas acreditem que a atividade destes elementos exercem uma função importante na adaptação evolutiva destes organismos em seu ambiente. Os mecanismos pelos quais os TEs podem influenciar a expressão gênica e consequentemente a evolução do genoma são os mais diversos (Figura 1.6) (LISCH, 2013).

Assim, fica claro que o sequenciamento do genoma de organismos eucariotos apresenta fundamental importância para o conhecimento da diversidade, estrutura, evolução e possíveis efeitos de TEs para o hospedeiro. Diversos genomas de fungos fitopatogênicos já foram sequenciados e investigados quanto a presença de TEs, como exemplos: *Moniliophthora perniciosa* e *Cochiliobolus heterostrophus* (SANTANA, 2009), *Puccinia triticina* (FELLERS *et al*., 2013), *Microbotryum lychnidis-dioicae*, *Puccinia graminis*, *Melampsora laricis-populina* e *Rhodotorula graminis* (HORNS *et al*., 2012), *P. psidii* (TAN *et al*., 2014), *Puccinia striiformis* f. sp. *tritici* (CANTU *et al*., 2011), 30 espécies dos gêneros *Puccinia, Uromyces* e *Melampsora* (CRISTANCHO *et al*., 2014) entre outros. Assim, fica evidenciado que novas estratégias como a análise *in silico* podem e devem ser aplicadas em conjunto com as ferramentas da biologia molecular para o melhor entendimento dos TEs (PEREIRA, 2005; SANTANA, 2009; SANTANA *et al*., 2012).

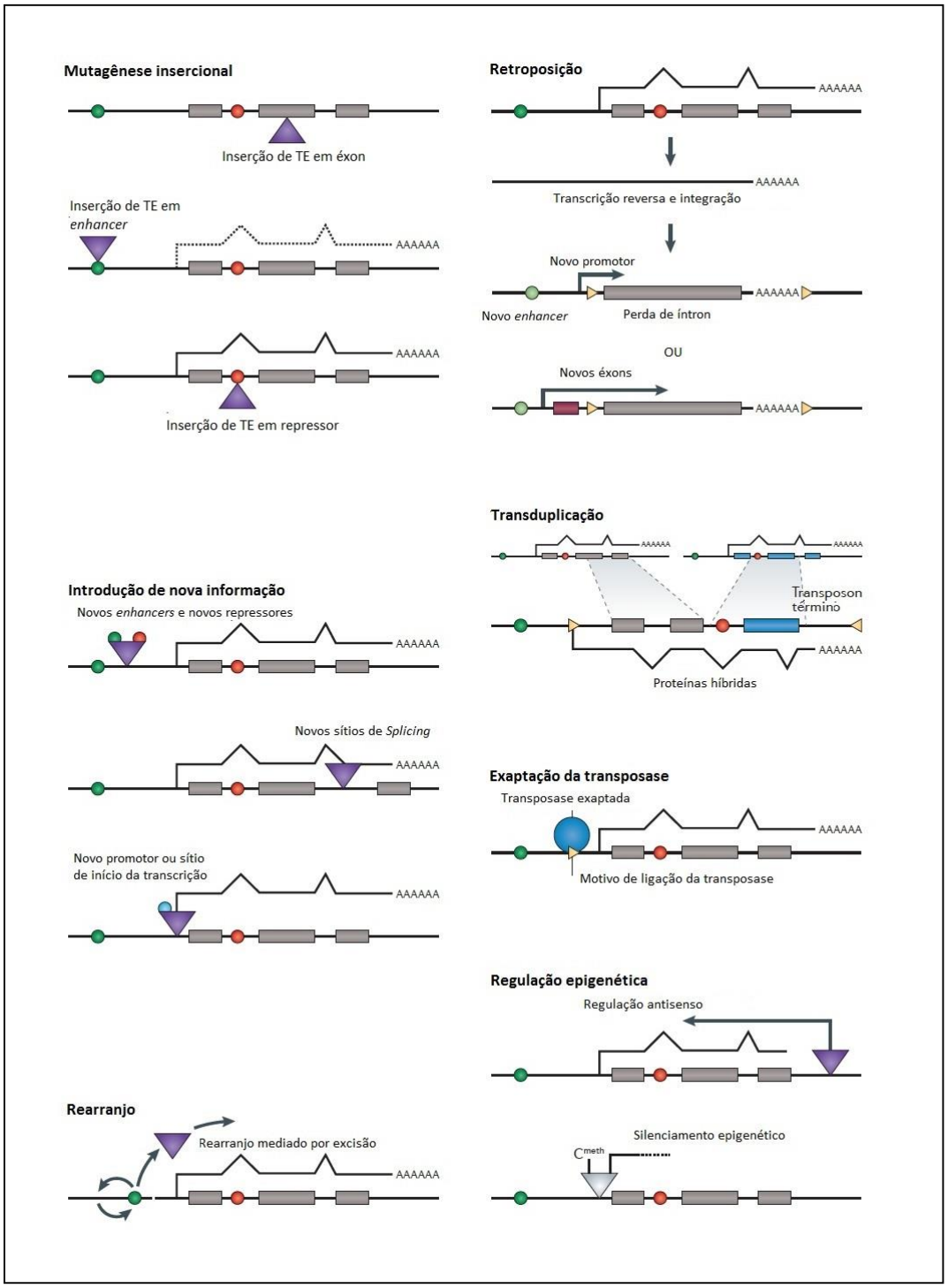


Figura 1.6 - Modificações estruturais e funcionais que podem ser causadas por elementos transponíveis (TEs). Os TEs podem estar envolvidos em mudanças que incluem a “quebra” (*knockout*) da função, introdução de novas funções, modificações na estrutura dos genes, mobilização e rearranjo de fragmentos genéticos e até mesmo no silenciamento epigenético de genes. Éxons estão representados por retângulos na cor cinza, *enhancers* são círculos de cor verde, repressores são círculos de cor vermelha, TEs são representados como triângulos de cor roxa, término do transposon são triângulos amarelos e proteínas exaptadas são círculos azuis. C^{meth} é 5-metilcitosina. Fonte: Adaptado de LISCH, 2013

Referências

ALFENAS, A.C.; ZAUZA, E.A.V.; MAFIA, R.G.; ASSIS, T.F. (Ed.). **Clonagem e doenças do eucalipto**. 2. ed. Viçosa: Universidade Federal de Viçosa, 2004. 500 p.

APARECIDO, C. C.; FIGUEIREDO, M. B.; FURTADO, E.L. Groups of physiological variability in *Puccinia psidii* populations. **Summa Phytopathologica**, Jaguariúna, SP, v. 29, p. 234–238, 2003.

APARECIDO, C.C. **Estudos ecológicos sobre *Puccinia psidii* Winter – ferrugem das mirtáceas**. 2001, 69 p. Dissertação (Mestrado em Agronomia) – Universidade Estadual de São Paulo, Botucatu, 2001.

BIET, E.; SUN, J.; DUTREIX, M. Conserved sequence preference in DNA binding among recombinant proteins: abnormal effect of ssDNA secondary structure. **Nucleic Acids Research**, London, GB, v. 27, n.2, p. 596-600, 1999.

BOOTH, T.H.; OLD, K.M.; JOVANOVIC, T. A preliminary assessment of high risk areas for *Puccinia psidii* (*Eucalyptus* rust) in the Neotropics and Australia. **Agriculture, Ecosystems and Environment**, Amsterdam, NL, v. 82, p. 295–301, Dec. 2000.

BRASIL. Decreto n. 3.420, de 20 de abril de 2000. Dispõe sobre a criação do Programa Nacional de Florestas – PNF, e dá outras providências. Publicado originalmente no **Diário Oficial** [da] República Federativa do Brasil na Edição Extra, 22 Abr. 2000. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto/D3420.htm>. Acesso em: 09 dez.2014.

CANTU, D.; GOVINDARAJULU, M.; KOZIK, A.; WANG, M.; CHEN, X.; KOJIMA, K.K.; JURKA, J.; MICHELMORE, R.W.; DUBCOVSKY, J. Next Generation Sequencing Provides Rapid Access to the Genome of *Puccinia striiformis* f.sp. *tritici*, the Causal Agent of Wheat Stripe Rust. **PLos ONE**, San Francisco, US, v.6, n.8, Aug. 2011. Disponível em: < DOI: 10.1371/journal.pone.0024230.t001.>. Acesso em: 05 jul. 2015.

CARNEGIE, A.J.; LIDBETTER, J.R. Rapidly expanding host range of *Puccinia psidii* sensu lato in Australia. **Australasian Plant Pathology**, Clayton, AU, v. 41, p. 13–29, Jan. 2012.

CARNEGIE, A.J.; LIDBETTER, J.R.; WALKER, J.; HORWOOD, M.A.; TESORIERO, L.; GLEN, M.; PRIEST, M. J. *Uredo rangelii*, a taxon in the guava rust complex, newly recorded on Myrtaceae in Australia. **Australasian Plant Pathology**, Clayton, AU, v. 39, p. 436-466, Sep. 2010.

CARVALHO Jr., A. A.; MARTINS, E. M. F.; FIGUEIREDO, M. B. Cultura axênica de *Melampsora epitea*, ferrugem do chorão (*Salix babylonica*), a partir de urediniosporos. **Fitopatologia Brasileira**, Brasília, DF, v. 23, p. 379-385, 1998.

CASTRO, H. A.; KRÜGNER, T.L.; IDERIHA, C. H. F; CAPPELLO, M. S. C.; MARCHI, A. B. Cross-inoculation of *Eucalyptus,* guava (*Psidum guajava*), and rose apple (*Syzygium jambos*) with *Puccinia psidii*. **Fitopatologia Brasileira**, Brasília, DF, v. 8, p. 491–497, 1983.

COELHO, L.; ALFENAS, A. C.; FERREIRA, F. A. Physiologic variability of *Puccinia psidii* – the rust of *Eucalyptus*. **Summa Phytopathologica**, Jaguariúna, SP, v. 27, p. 295–300, 2001.

COUTINHO, L.N.; FIGUEIREDO, M.B. Estudos sobre especializações fisiológicas em *Puccinia psidii* Winter. **Summa Phytopathologica***,* Jaguariúna, SP, v. 10, p. 55-56, 1984.

COUTINHO, T.; WINGFIELD, M.J.; ALFENAS, A.C.; CROUS, P.W. Eucalyptus rust: a disease with the potential for serious international implications. **Plant disease**, Saint Paul, US, v. 82, n. 7, p. 819-825, Jul. 1998.

CRISTANCHO, M.A.; BOTERO-ROZO, D.O.; GIRALDO, W. TABIMA, J.; RIAÑO-PACHÓN, D.M.; ESCOBAR, C.; ROZO, Y.; RIVERA, L.F.; DURÁN, A.; RESTREPO, S.; EILAM, T.; ANIKSTER, Y.; GAITÁN, A.L. Annotation of a hybrid partial genome of the coffee rust (*Hemileia vastatrix*) contributes to the gene repertoire catalog of the Pucciniales. **Frontiers in Plant Science**, Lausanne, v. 31, n. 5, Oct. 2014. Disponível em: <DOI: 10.3389/fpls.2014.00594.>. Acesso em: 05 jul. 2015.

DABOUSSI, M.J. Fungal transposable elements and genome evolution. **Genetica**, The Hague, NL, v. 100, p. 253 – 260, 1997.

DABOUSSI, M.J.; CAPY, P. Transposable elements in filamentous fungi. **Annual Review of Microbiology**, Palo Alto, US, v. 57, p. 275-299, 2003.

FAO. **Global forest in sustainable forest management - A statement of principles.** Roma, 2010. 16 p.

FELLERS, J.P.; SOLTANI, B.M.; BRUCE, M.; LINNING,R.; CUOMO, C.A.; SZABO, L.J.; BAKKEREN, G. Conserved loci of leaf and stem rust fungi of wheat share synteny interrupted by lineage-specific influx of repeat elements. **BMC Genomics**, London, GB, v. 14, n. 60, Jan. 2013. Disponível em: <DOI: 10.1186/1471-2164-14-60.>. Acesso em: 05 jul. 2015.

FERREIRA, F.A. Ferrugem do eucalipto. **Revista árvore**, Viçosa, MG, v. 7, n. 2, p. 91-109,1983.

_____. Ferrugem do eucalipto-ocorrência, temperatura para germinação de uredosporos, produção de teliosporos, hospedeiros alternativos e resistência. **Fitopatologia Brasileira**, Brasília, DF, n. 6, p. 603-604, 1981.

FIGUEIREDO, M. B. Doenças fúngicas emergentes em grandes culturas. **O Biológico**, São Paulo, BR, v. 63, p. 29-32, Jan. 2001.

FIGUEIREDO, M.B.; PASSADOR M.M. Morfologia, funções dos soros e variações dos ciclos vitais das ferrugens. **Arquivos do Instituto Biológico**, São Paulo, SP, v. 75, n. 1, p.117-134, Jan. 2008.

FURTADO, E.L.; SANTOS, C.A.G.; MASSON, M.V. Impacto potencial das mudanças climáticas sobre a ferrugem do eucalipto no Estado de São Paulo**.** In: GHINI R., HAMADA E. (Ed.), **Mudanças Climáticas. Impactos Sobre Doenças de Plantas no Brasil**, Brasília, Embrapa Informação Tecnológica, 2008. p. 273–286.

GLEN, M.; ALFENAS, A.C.; ZAUZA, A.V.; WINGFIELD, M.J.; MOHAMMED, C. *Puccinia psidii*: a threat to the Australian environment and economy a review. **Australasian Plant Pathology**, Clayton, AU, v. 36, n. 1, p. 1-16, Jan. 2007.

GONÇALVES, J.L.M.; ALVARES, C.A.; HIGA, A.R.; SILVA, L.D.; ALFENAS, A.C.; STAHL, J.; FERRAZ, S.F.B.; LIMA, W.P.; BRANCALION, P.H.S.; HUBNER, A.; BOUILLET, J.P.D.; LACLAU, J.P.; NOUVELLON, Y.; EPRON, D. Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in Brazilian eucalypt plantations. **Forest Ecology and Management**, Amsterdam, NL, v. 301, p. 6–27, Aug. 2013.

GRAÇA, R.N.; ROSS-DAVIS, A.L.; KLOPFENSTEIN, N.B.; KIM, M.S.; PEEVER, T.P.; CANNON, P.G.; AUN, C.P.; MIZUBUTI, E.S.G.; ALFENAS, A.C. Rust disease of eucalypts, caused by *Puccinia psidii*, did not originate via host jump from guava in Brazil. **Molecular Ecology**, Oxford, GB, v. 22, p. 6033–6047, Nov. 2013.

GUTMANIS, G. **Modulação de gene *ugp* e análise das alterações na composição dos carboidratos da parede celular primária e secundária de *Nicotiana tabacum* e *Eucalyptus grandis***. 2008, 142 p. Tese (Doutorado em Ciências) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2008.

HARWOOD, C. New Introductions: doing it right. In: WALKER, J. (Ed.), **Developing a Eucalypt Resource**: Learning from Australia and Elsewhere. Christchurc: Wood Technology Research Centre, University of Canterbury, 2011. p. 125–136.

HAVECKER, E.R.; GAO, X.; VOYTAS, D.F. The diversity of LTR retrotransposons. **Genome Biology**, London, GB, v. 5, May. 2004. Disponível em < DOI: 10.1186/gb-2004-5-6-225>. Acesso em: 30 jun.2015.

HORNS, F.; PETIT, E.; HOOD, M.E. Patterns of repeat-induced point mutation in transposable elements of basidiomycetes fungi. **Genome Biology and Evolution**, Oxford, GB, v. 4, n. 3, p. 240-247, Jan. 2012

INDÚSTRIA BRASILEIRA DE ÁRVORES. São Paulo, 2015. 64 p.

JOFFILY, J. Ferrugem do eucalipto. **Bragantia**, Campinas, SP, v. 4, n. 8, p.475 - 487, 1944.

LEITE, T.F. **Análise histológica da interação planta patógeno entre *Eucalyptus grandis* e *Puccinia psiddi* Winter e identificação de genes diferencialmente expressos em plantas resistentes e susceptíveis inoculadas com o patógeno por meio da técnica de RNA-Seq.** 2012, 177 p. Tese (Doutorado em Ciências) Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.

LI, Y.C.; KORD, A.B.; FAHIMA, T.; BERLES A.; NERO E. Microsatellites: genomic distribution, putative functions and mutation mechanisms: a review. **Molecular Ecology**, Oxford, GB, v. 11, n. 12 p. 2453-2465, Dec. 2002.

LISCH, D. How important are transposons for plant evolution? **Nature Reviews Genetics**, London, GB, v. 14, n. 1, p. 49–61, Dec. 2013.

LIU, Z.; LI, P.; KOCABAS, A.; KARSI, A.; JU, Z. Microsatellite containing genes from the channel catfish brain: evidence of trinucleotide repeat expansion in the coding region of nucleotide excision repair gene RAD23B. **Biochemical and Biophysical Research Communications**, New York, US, v. 30, n. 289, p. 317-324, Nov. 2001.

MARTIN, F.; AERTS, A.; AHRÉN, D.; BRUN, A.; DANCHIN, E.G.; DUCHAUSSOY, F.; GIBON, J.; KOHLER, A.; LINDQUIST, E.; PEREDA, V.; SALAMOV, A.; SHAPIRO, H.J.; WUYTS, J.; BLAUDEZ, D.; BUÉE, M.; BROKSTEIN, P.; CANBÄCK, B.; COHEN, D.; COURTY, P.E.; COUTINHO, P. M.; DELARUELLE, C.; DETTER, J.C.; DEVEAU, A.; DIFAZIO, S.; DUPLESSIS, S.; FRAISSINET-TACHET, L.; LUCIC, E.; FREY-KLETT, P.; FOURREY, C.; FEUSSNER, I.; GAY, G.; GRIMWOOD, J.; HOEGGER, P.J.; JAIN, P.; KILARU, S.; LABBÉ, J.; LIN, Y.C.; LEGUÉ, V.; LE TACON, F.; MARMEISSE, R.; MELAYAH, D.; MONTANINI, B.; MURATET, M.; NEHLS, U.; NICULITA-HIRZEL, H.; OUDOT-LE SECQ, M.P.; PETER, M.; QUESNEVILLE, H.; RAJASHEKAR, B.; REICH, M.; ROUHIER, N.; SCHMUTZ, J.; YIN, T.; CHALOT, M.; HENRISSAT, B.; KÜES, U.; LUCAS, S.; VAN DE PEER, Y.; PODILA, G.K.; POLLE, A.; PUKKILA, P.J.; RICHARDSON, P.M.; ROUZÉ, P.; SANDERS, I.R.; STAJICH, J.E.; TUNLID, A.; TUSKAN, G.; GRIGORIEV, I.V. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. **Nature**, London, GB, v. 452, n. 7183, p. 88-92, Mar. 2008.

MASSON, M. V. **Ferrugem do eucalipto:** planejamento evasivo, estimativa de dano e análise da viabilidade do controle químico. 2009 174 p. Dissertação (Mestrado em Proteção de Plantas), Faculdade de Ciências Agronômicas da UNESP, Botucatu, 2009.

MORAES, T.S. de A.; GONÇALVES, E.L.; RESENDE, G.C. de; MENDES, J.; SUITER, F.W. Evolução da ferrugem causada pela *Puccinia psidii* Winter em *Eucalyptus* spp. Dados preliminares. **IPEF Circular Técnica**, Piracicaba, SP, v. 144, p. 1-12, 1982.

MORIN, L.; AVEYARD, R.; LIDBETTER, J.R.; WILSON, P.G. Investigating the Host-Range of the Rust Fungus *Puccinia psidii* sensu lato across Tribes of the Family Myrtaceae Present in Australia. **PLOS One**, San Francisco, US, v. 7, n. 4, Apr. 2012. Disponível em: <DOI: 10.1371/journal.pone.0035434.>. Acesso em: 30 jun. 2015.

MORIN, L.; TALBOT, M.J.; GLEN, M. Quest to elucidate the life cycle of *Puccinia psidii* sensu lato. **Fungal biology**, Amsterdam, GB, v. 118, n. 2, p. 253-263, Feb. 2014.

MOURA, A.; GARCIA, C. H. (Ed.) **A cultura do eucalipto no Brasil**. São Paulo: Sociedade Brasileira de Silvicultura, 2000. 112 p.

PEREIRA, J.F. **Caracterização, distribuição e estudo da atividade de elementos transponíveis em Crinipellis perniciosa, agente causal da vassoura-de-bruxa no cacaueiro (*Theobroma cacao*)**. 2005. 182 p. Tese (Doutorado em Microbiologia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2005.

POKORNA, M.; KRATOCHVIL, L.; KEJNOVSKY, E. Microsatellite distribution on sex chromosomes at different stages of heteromorphism and heterochromatinization in two lizard species (Squamata: Eublepharidae: *Coleonyx elegans* and Lacertidae: *Eremias velo*x). **BMC Genetics**, London, GB, v. 12, Oct. 2011. Disponível em < DOI: 10.1186/1471-2156-12-90.>. Acesso em: 30 jun.2015.

QUECINE, M.C.; BINI, A.P.; ROMAGNOLI, E.R.; ANDREOTE, F.D.; MOON, D.H.; LABATE, C.A. Genetic Variability in *Puccinia psidii* Populations as Revealed by PCR-DGGE and T-RFLP Markers. **Plant Disease**, Saint Paul, US, v. 98, n.1, p.16-23, Jan. 2014.

QUECINE, M.C.; MOON, D.; FRANCHESCHINI, L.M.; BINI, A.P.; LEITE, T.F.; LABATE, C.A. De novo partial genome assembly of the biotrophic eucalyptus rust pathogen *Puccinia psidii*. In: INTERNATIONAL CONGRESS ON MOLECULAR PLANT-MICROBE INTERACTION, 15., 2012, Kioto. **Abstracts ...** Kioto: IS-MPMI, 2012. 1 CD-ROM.

ROUX, J.; GREYLING, I.; COUTINHO, T.A.; VERLEUR, M.; MICHAEL J.; WINGFIELD, M.J. The Myrtle rust pathogen *Puccinia psidii* discovered in Africa. **IMA Fungus**, Ashtead, UK, v. 1, n. 4, p. 155-159, Jun. 2013.

RUIZ-SILVA, J. **Isolamento e caracterização de elementos transponíveis em espécies do gênero *Brycon* (Characidae, Bryconinae)**. 2012, 81 p. Dissertação (Mestrado em Ciências Biológicas) – Universidade Estadual de São Paulo, Botucatu, 2012.

SANTANA, M.F. **Retrotransposons LTR no genoma de *Moniliophthora perniciosa* e *Cochliobolus heterostrophus*.** 2009, 81 p. Dissertação (Mestrado em Microbiologia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2009.

SANTANA, M.F.; SILVA, J.C.F.; BATISTA, A.D.; RIBEIRO, L.E.; SILVA, G.F.; ARAÚJO, E.F.; QUEIROZ, M.V. Abundance, distribution and potential impact of transposable elements in the genome of *Mycosphaerella fijiensis.* **BMC Genomics**, London, GB, v. 13, n. 720, Dec. 2012. Disponível em <DOI: 10.1186/1471-2164-13-720.>. Acesso em: 30 jun.2015.

SANTOS, C. **Estudos epidemiológicos da ferrugem do eucalipto causado por Puccinia psidii em plantios irrigados.** 2006. 75 p. Dissertação (Mestrado em Proteção de Plantas), Faculdade de Ciências Agronômicas, Universidade Estadual de São Paulo, Botucatu, 2006.

SCHMIDT, T.; HESLOP-HARRISON, J.S. Genomes, genes, and junk: the large-scale organization of plant chromosomes. **Trends in Plant Science**, Oxford, GB, v. 3, p. 195-199, May. 1998.

SILVA, P.H.M.; MIRANDA, A.C.; MORAES, M.L.T.; FURTADO, E.L.; STAPE, J.L.; ALVARES, C.A.; SENTELHAS, P.C.; MORI, E.S.; SEBBENN, A.M. Selecting for rust (*Puccinia psidii*) resistance in *Eucalyptus grandis* in São Paulo State, Brazil. **Forest Ecology and Management**, Amsterdam, NL, v. 303, p. 91–97, Sep. 2013.

STAPE, J.L.; BINKLEY, D.; RYAN, M. G. Production and carbon allocation in a clonal *Eucalyptus* plantation with water and nutrient manipulations. **Forest Ecology and Management**, Amsterdam, NL, v. 255, p. 920–930, Abr. 2008.

TAKAHASHI, S. S. **Ferrugem do eucalipto:** índice de infecção, análise temporal e estimativas de danos relacionados a intensidade da doença no campo. 2002, 113 p. Dissertação (Mestrado em Proteção de Plantas) – Faculdade de Ciências Agronômicas, Universidade Estadual de São Paulo, Botucatu, 2002.

TAN, M.K.; COLLINSA, D.; CHENB, Z.; ENGLEZOUA, A.; WILKINS, M.R. A brief overview of the size and composition of the myrtle rust genome and its taxonomic stat. **Mycology**, Abingdon, GB, v. 5, n. 2, p. 52–63, Jun. 2014.

TESSMANN, D.J.; DIANESE, J.C.; MIRANDA, A.C.; CASTRO, L.H.R. Epidemiology of a Neotropical rust (*Puccinia psidii*): periodical analysis of the temporal progress in a perennial host (*Syzygium jambos*). **Plant Pathology**, Oxford, GB, v. 50, 725–731, Dec. 2001.

WALLAU, G.L. **História evolutiva de elementos transponíveis da superfamília *Tc1- Mariner* em drosofilídeos**. 2010. 85 p. Dissertação (Mestrado em Ciências Biológicas) – Universidade Federal de Santa Maria, Santa Maria, 2010.

WICKER, T; SABOT, F; HUAN-VAN, A.; BENNETZEN, J. L.; CAPY, P.; CHALHOUB, B.; FLAVELL, A.; LEROY, P.; MORGANTE, M.; PANAUD, O.; PAUX, E.; SANMIGUEL, P.; SCHULMAN, A. H. A unified classification system for eukaryotic transposable elements. **Nature**, London, GB, v. 8, n. 12 p. 973–982, Dec. 2007.

2 ANÁLISE *IN SILICO* DE ELEMENTOS TRANSPONÍVEIS EM *Puccinia psidii*

Resumo

A alta velocidade com que os dados de sequenciamento por meio de plataformas de nova geração (NGS – *Next Generation Sequencing*) estão sendo gerados e disponibilizados certamente contribui com estudos mais detalhados de elementos transponíveis (TEs). O estudo desses elementos é importante para o entendimento da biologia de fitopatógenos, pois uma fração significante nos genomas de organismos eucariotos é composta por TEs. Assim, apesar de promissor, o estudo de TEs representa também um grande desafio. No presente estudo, foi realizada a anotação automática por meio do programa *RepeatMasker* dos potenciais TEs presentes no genoma parcialmente sequenciado de *Puccinia psidii* (linhagem monopustular MF-1). Duas estratégias foram adotadas: a primeira por meio da classificação de 586 "*supercontigs*" remontados e originários da mineração de *contigs* cujos GO termos (*Blast2GO*) eram relacionados a TEs; a segunda estratégia foi a anotação dos 167.538 *contigs* que formam o genoma parcial de *P. psidii* MF-1 pelo *RepeatMasker.* Ambas estratégias apontaram a predominância de elementos da Classe I – LTR Retrotransposons no genoma de *P. psidii* MF-1. A anotação com 586 *"supercontigs"* apresentou 99% das sequências analisadas compostas por elementos repetitivos, como esperado, o que demonstra a eficiência da mineração de sequências. Já a anotação de 167.538 *contigs* apresentou 10% das sequências compostas por elementos repetitivos. Foram selecionados 20 e 10 *contigs* classificados como LTR Retrotransposons e DNA Transposons, respectivamente, para a verificação da integridade destes elementos em *P. psidii* MF-1. Nenhum dos *contigs* analisados apresentou regiões de interesse, como LTRs ou TIRs, nem mesmo TSRs. Foram selecionados 25 *contigs* classificados como elementos Copia e Gypsy, eles foram avaliados quanto à integridade e anotados manualmente por meio da ferramenta BLASTX. Não foram encontradas regiões LTRs e TSRs. Dentre os 25 *contigs,* 44% apresentaram todos os domínios proteicos da região *pol*. Além disso, as regiões UBN2_2 e CH, relacionadas à transposição foram encontradas nos *contigs* 00084 e 00126. A análise comparativa entre os conteúdos de TEs presente no genoma de *P. psidii* MF-1*, P. graminis, P. striiformis* e *P. triticina* foi realizada. A frequência de TEs pertencentes à Classe II em *Puccinia* spp. é maior, compondo em média 2% dos genomas, comparado ao predomínio de TEs da Classe I em *P. psidii* MF-1 (0,02% de LTR Retrotransposons e fração insignificante de DNA transposons). A análise da integridade de 15 *supercontigs* contendo elementos Copia de *Puccinia* spp. revelou identidade de LTRs superior a 97%, consenso 5'-TG...GA-3' e presença de TSRs duplicados. Por fim, fez-se uma análise comparativa entre sequências de transcriptase reversa de 12 *contigs* de *P. psidii* MF-1, 2 de *Puccinia* spp. e 5 de diversos fungos fitopatogênicos. As 12 proteínas transcriptase reversa de *P. psidii* MF-1 agruparam com as 2 proteínas provenientes de elementos Copia de *Puccinia* spp. em dois clados diferentes, o que mostra a diversidade destas proteínas no genoma de *P. psidii*. Considerando os resultados preliminares, análises futuras serão necessárias para obter uma abordagem mais ampla de TEs presentes no genoma do fitopatógeno *P. psidii*, tornando possível o melhor entendimento da biologia do fungo.

Palavras-chave: Anotação; *RepeatMasker*; LTR Retrotransposon; Integridade; Comparação de TEs; *Puccinia* spp.; Transcriptase reversa

Abstract

Next Generation Sequencing (NGS) platforms generate data at very high speedy, that increase the TEs (transposable elements) studies aiming to understand phytopathogens biology, due to the fact that a great fraction of eukaryotic organisms is composed by TEs. The TEs may represents a great potential activity in virulence mechanisms of pathogenic fungi. Thus, although promising, the TEs study also represents a huge challenge. Thereby, in this study, an annotation by *RepeatMasker* of potential TEs present in *Puccinia psidii* (mono-pustular strain MF-1) partial genome was conducted. Two strategies were adopted, the first by classifying 586 “supercontigs” re-assembled and originated by mining GO terms related to TEs contigs (Blast2GO); second classifying the 167.538 contigs from the *P. psidii* MF-1 partial genome. Both strategies showed Class I – LTR Retrotransposons predominance in the *P. psidii* MF-1 genome. The 586 “supercontigs” annotation showed about 99% repetitive elements, as expected. The 167.538 contigs annotation showed 10% of the sequences composed of repetitive elements. Aiming to proceed a TE integrity analysis in *P. psidii* MF-1, 20 and 10 contigs classified as LTR Retrotransposons and DNA Transposons, respectively, were selected and evaluated. None of analyzed contigs showed regions of interest such as LTRs, TIRs even TSRs. There is 25 contigs classified as Copia and Gypsy elements that were selected to an integrity analysis and then a manually annotation by BLASTX tool. No LTRs or TSRs regions were found. Among the 25 contigs, 44% presents all protein sequences of the pol region. Furthermore, UBN2_2 and CH regions, related to transposition events, belonging to contigs 00084 and 00126 were also found. A comparative analysis between TEs contents present in *P. psidii* MF-1, *P. graminis, P. striiformis* and *P. triticina* was carried out. The Class II TEs frequency of *Puccinia* spp. was greater, composing about 2% of the genome, when compared to the predominance of Class I TEs in *P. psidii* MF-1 genome (0.02% of LTR retrotransposon and insignificant fraction of DNA Transposon). An integrity analysis of 15 supercontigs containing Copia elements in *Puccinia* spp. was also conducted. An identity greater than 97% of LTR regions, 5’-TG...GA-3’ consensus and presence of duplicated TSRs were observed in this analysis. Finally, a comparative analysis of the reverse transcriptase protein was conducted with 12 contigs sequences from *P. psidii* MF-1, 2 from *Puccinia* spp. and 5 from various phytopathogenic fungi. All *P. psidii* MF-1 reverse transcriptase protein sequences grouped with both *Puccinia* spp. in two different clades, suggesting the variability of this protein. Considering this preliminary results, further analysis will be needed for a broader approach for TEs in the pathogen *P. psidii* genome, enabling a better understanding of the fungal biology.

Keywords: Annotation; *RepeatMasker*; LTR Retrotransposon; Integrity; TEs Comparison; *Puccinia* spp.; Reverse transcriptase

2.1 Introdução

A importância do setor florestal brasileiro se deve a vasta área territorial ocupada para a produção de espécies madeireiras, como o *Eucalyptus*; a geração de tributos na escala de bilhões de reais; além de milhões de empregos relacionados ao setor (IBÁ, 2015).

No entanto, a produtividade é afetada pelo ataque de patógenos ao longo do desenvolvimento da cultura, sendo a ferrugem, doença causada pelo fungo basidiomiceto biotrófico *Puccinia psidii* Winter, a principal responsável por severos danos no setor (GLEN *et al*., 2007). O fungo *P. psidii* afeta diversas espécies da família Myrtaceae, bem como o eucalipto (COUTINHO *et al*., 1998; ALFENAS *et al*., 2004; GLEN *et al*., 2007; QUECINE *et al.*, 2014). O primeiro relato de ferrugem em eucalipto ocorreu no século XX, mesmo assim o conhecimento sobre a genética e a biologia do patógeno são ainda incipientes (LEITE, 2012; QUECINE *et al*., 2014).

Entretanto, com a enorme quantidade de dados gerados pelas plataformas de sequenciamento de nova geração, o conhecimento sobre estrutura e composição dos diversos genomas de interesse vem sendo atingidos com maior rapidez. A grande maioria dos dados apresentados mostra que existe uma fração significante nos genomas de organismos eucariotos composta por elementos transponíveis (TEs) e em fungos chegam a representar 3 a 21% do genoma (MARTIN *et al*., 2008; SANTANA, 2009). Os TEs são sequências de DNA repetitivas com capacidade de se movimentarem dentro dos genomas e com isso influenciam na sua reestruturação, pois a movimentação de TEs pode gerar novos rearranjos, novos alelos e até modificar a expressão gênica (SANTANA, 2009; SANTANA *et al*., 2012). Há uma grande diversidade e variação entre o número de cópias, distribuição e tipos de TEs dentre as espécies, resultado das características intrínsecas dos elementos transponíveis e as forças evolutivas que atuam sobre esses (DABOUSSI, 1997).

Dada a abundância, complexidade e diversidade de TEs e a velocidade com que os dados de sequenciamento estão sendo gerados, a identificação e anotação de TEs presentes nos organismos de estudos tem sido um grande desafio (WICKER *et al*., 2007). A detecção e caracterização de TEs em fitopatógenos podem possibilitar um maior conhecimento sobre a organização e a evolução do genoma desses fungos, bem como os mecanismos de ativação e silenciamento gênico, resultado da movimentação dos mesmos (MARTIN *et al*., 2008; SANTANA, 2009).

Assim, o presente estudo teve como principais objetivos a análise da frequência e classificação preliminar de TEs no genoma parcialmente sequenciado de *P. psidii*; anotação de regiões conservadas; comparação da composição de elementos repetitivos do genoma com outras espécies de *Puccinia* causadoras de ferrugens e análise da enzima transcriptase reversa de diversos fungos fitopatogênicos. Dessa forma, visando, um melhor entendimento da evolução do genoma deste fitopatógeno para que então, futuramente, a complexa interação planta-patógeno seja melhor entendida para o desenvolvimento de novas estratégias de controle.

2.2 Materiais e Métodos

2.2.1 Material biológico e dados preexistentes

O material biológico utilizado foi a linhagem monopostular MF-1 do fungo *P. psidii*. Esta linhagem foi obtida de uma monoprogênie susceptível de *E. grandis*, D901, na área experimental da Suzano Papel e Celulose, Itapetininga-SP e purificada em monopústulas por meio de repiques sucessivos (LEITE, 2012). O DNA dessa linhagem foi utilizada para a extração do material genético por meio do *DNAeasy Plant Mini Kit* (Qiagen) e posterior sequenciamento parcial do genoma.

O genoma parcialmente sequenciado de *P. psidii* MF-1 foi montado a partir de sequências (*reads*) geradas pelo sequenciamento *454 Roche GS FLX - Titanium* por meio do programa *Newbler Assembler* 2.6 (MARGULIES *et al*., 2005). O tamanho mínimo dos *contigs* gerados foi de 250 pb. Foram obtidos no total 5.955.419 *reads*, que perfazem 2.486.033.413 pb, o que corresponde à aproximadamente uma cobertura de 20 X o genoma estipulado para diversas espécies de *Puccinia,* cujo genoma varia entre 67 e 164 Mb (LEONARD; SZABO, 2005).

Por meio do *de novo assembly* realizado pelo programa *Newbler Assembler* 2.6 foram utilizadas aproximadamente 50% do total de sequências obtidas, perfazendo 473.271.316 pb alinhados, resultando em 167.538 *contigs* de tamanho médio de 1.412 pb. O maior *contig* possui tamanho aproximado de 30.053 pb (dado não publicado).

A anotação funcional do genoma parcial de *P. psidii* MF-1 realizada por meio do programa Blast2GO (CONESA; GÖTZ *et al*., 2005) resultou na categorização de sequências em termos do Consórcio *Gene Ontonlogy* (GO) (ASHBURNER *et al*., 2000) (http://www.geneontology.org) correspondentes a três categorias funcionais:

processo biológico, função molecular e componente celular, utilizando os parâmetros-padrão indicados pelo programa.

2.2.2 Mineração de sequências relacionadas a elementos transponíveis

Inicialmente, as 167.538 sequências (*contigs*) geradas pela montagem no programa *Newbler Assembler* 2.6 e anotadas automaticamente pelo programa Blast2GO foram mineradas de acordo com os GO termos de função molecular. Os *contigs* com GO termos relacionados à ligação a cromatina, nucleotídeo e DNA, GO:0003682, GO:0000166 e GO:0003677 respectivamente, que estão diretamente associados a grande presença de transposons no genoma desse fungo foram selecionados.

Foi realizada uma busca no programa Blas2GO por *contigs* contendo as palavras-chave relacionadas a processos de transposição ou elementos de transposição (pol, gag, rve, integrase, poliproteína, retrotransponível, retrotransposon, *tick*, transcriptase reversa, RNAse, poliproteína-copia, retrotransposon gypsy, DNA transposon, hAT, mutator, Tc1 entre outros).

2.2.3 Anotação direta dos *contigs* de *P. psidii* MF-1

Para anotação automática dos elementos transponíveis presentes no genoma de parcialmente sequenciado de *P. psidii* MF-1 foram utilizadas duas estratégias, ou seja, anotados dois grupos de sequências obtidas diferencialmente.

Primeiramente, a fim de se obter sequências maiores que permitissem a classificação mais fiel de elementos transponíveis minerados a partir dos *contigs* foi utilizado o programa *CLC Genomic Workbench* 6.0 para a montagem das sequências obtidas no item 2.2.2. Para tanto, foram testados diferentes parâmetros (não-pareamento; custo de inserção; custo de deleção; tamanho; similaridade) em três tentativas de montagem denominadas Tentativa I, II e III (Tabela 2.4).

O melhor resultado, levando em consideração o N50 obtido foi utilizado para classificação pelo programa *RepeatMasker* (CHEN, 2004; KOHANY *et al*., 2006) (http://www.repeatmasker.org), utilizando os parâmetros padrão do programa com a alteração da fonte da sequência (*sequence source*) para Fungi.

A outra estratégia foi utilizar diretamente os 167.538 *contigs* de *P. psidii* MF-1 por meio do programa *RepeatMasker*, utilizando os parâmetros: RMBlast; banco de

dados FNGREP.ref; baixa sensibilidade (para não pegar qualquer sequência e considerar como elemento transponível).

2.2.4 Análise da integridade de elementos transponíveis de *P. psidii*

Após a identificação de sequências contidas no genoma parcialmente sequenciado de *P. psidii* pelo programa *RepeatMasker* (item 2.2.3) utilizando a segunda estratégia, os 20 maiores *contigs* que apresentaram regiões com identidade a algum elemento transponível da Classe I – Retrotransposon LTR tiveram seus 500 pb a jusante ou a montante submetidos a verificação da integridade do elemento com o programa *on-line LTR-Finder* (XU; WANG, 2007) (http://tlife.fudan.edu.cn/ltr_finder/), sendo o banco de dados de tRNA de Eukarya para predição o único parâmetro ajustado, os demais parâmetros utilizados foram a opção padrão indicada pelo programa.

Foi realizada também uma busca de regiões de terminações de TEs por meio do alinhamento de regiões similares pelo BLASTN – NCBI, selecionando-se a opção "*align two or more sequences*" (http://www.ncbi.nlm.nih.gov), somente para confirmação da presença/ausência de LTRs.

Elementos com regiões cujos *hits* apresentaram identidade com elementos da Classe II também foram submetidos à uma análise preliminar de integridade do elemento. Assim, os maiores 10 *contigs* classificados como DNA transposons foram selecionados e também tiveram cerca de 500 pb de suas sequências a jusante ou a montante analisados pelo programa *on-line TIR Finder* (GAMBIN *et al.,* 2013) (http://bioputer.mimuw.edu.pl/tirfindertool/) para encontrar as regiões TIRs (*Terminal Inverted Repeats*) na opção padrão indicada pelo programa.

Somente para confirmação da presença/ausência de TIRs foi realizada também uma busca de regiões das terminações pelo BLASTN – NCBI, selecionando-se a opção "*align two or more sequences*" (http://www.ncbi.nlm.nih.gov).

Com o auxílio da ferramenta *on-line ReverseComplement* (http://www.bioinformatics.org/sms/rev_comp.html) a região à montante da possível sequência de transposição é "transcrita" em seu reverso, pois os elementos da Classe II apresentam terminações invertidas.

2.2.5 Anotação manual de *contigs* classificados como elementos da Classe I – LTR Retrotransposons de *P. psidii* -MF1

Foram selecionados 25 *contigs* classificados pelo *RepeatMasker* pela segunda estratégia, como elementos da família Copia e/ou Gypsy e estes foram analisados pelo *LTR-Finder* (pela mesma metodologia do item 2.2.4) além disso, essas sequências também foram submetidas à anotação manual utilizando-se a ferramenta BLASTX (ALTSCHUL *et al*., 1990) pelo NCBI (http://www.ncbi.nlm.nih.gov) e Uniprot (http://www.uniprot.org/).

Os *contigs* que apresentaram o domínio relacionado à transposição (transcriptase reversa) foram ainda analisados pelo programa *ORF-Finder* (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). A anotação consistiu na determinação das regiões de sequências de leitura abertas ou ORFs (*Open Reading Frames*) exclusivas de LTR-Retrotransposons, como *gag* e *pol,* a fim de se identificar possíveis sequências relacionadas à transposição (protease, transcriptase reversa, RNase H, integrase) e talvez regiões importante para auxiliar a transcrição de genes de transposição como PPT – *Polipurina Treac* e CH – *Chromatin Organization Modifier*.

2.2.6 Análise comparativa da distribuição de TEs em *P. psidii* MF-1 e de outras espécies de *Puccinia*

Os genomas dos fungos *P. graminis, P. striiformis* e *P. triticina* foram obtidos na base de dados do *Puccinia Group Genome Databases, Broad Institute* (http://www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.html), cujas informações estão listadas na Tabela 2.1. As sequências do genoma dos três fitopatógenos foram também analisadas quanto a presença de elementos repetitivos por meio do programa *RepeatMasker* utilizando os parâmetros: RMBlast; banco de dados FNGREP.ref e baixa sensibilidade.

Tabela 2.1 – Informações sobre os patógenos utilizados como referência neste trabalho

Organismos de Referência	**Banco de Dados**	**Número de Acesso**	**Quantidade de *contigs***	**Características***
Puccinia graminis f. sp. *tritici*	*Puccinia - Group* Database do Broad Institute	AAWC01	4.557	Causador da ferrugem do colmo. Ocorre nas culturas do trigo, cevada, centeio, triticale e em algumas outras espécies de *Poacea* (Gramíneas).
Puccinia triticina	*Puccinia - Group* Database do Broad Institute	ADAS01	24.838	Causador da ferrugem da folha do trigo.
Puccinia striiformis	*Puccinia - Group* Database do Broad Institute	AEEW01	17.295	Causador da Ferrugem da tarja, que ocorre nas lavouras de trigo e cevada.

* Informações retiradas de *Puccinia Group Genome Databases, Broad Institute*

2.2.7 Análise da integridade de elementos transponíveis de *Puccinia* spp.

Após a classificação preliminar de elementos transponíveis nas três espécies de *Puccinia* pelo programa *RepeatMasker* (item 2.2.6), 5 *supercontigs* de cada espécie foram selecionados para a realização da análise de integridade dos elementos contidos em cada uma das sequências de *supercontig*. A busca foi realizada por meio da palavra “Copia” no arquivo de alinhamento e classificação gerado pelo *RepeatMasker,* assim, a ordenação dos resultados correspondem a *supercontigs* de maior tamanho com *hits* relacionados à família Copia.

Os *supercontigs* de cada uma das três espécies que apresentaram *hits* no alinhamento do *RepeatMasker* relacionados a LTR-Retrotranposons do tipo Copia foram selecionados e por meio da base de dados do *Puccinia Group – Broad Institute* (http://www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.htmll), cada um dos *supercontigs* selecionados de cada uma das três *Puccinia* foi individualmente salvo em formato FASTA para posterior análise por meio do programa *LTR-Finder* (XU; WANG, 2007) (http://tlife.fudan.edu.cn/ltr_finder/), assim como foi feito com os *contigs* de *P. psidii* MF-1 (item 2.2.4), com base no banco de dados de tRNA de Eukarya para predição e demais parâmetros na opção padrão indicada pelo programa. Neste caso, o arquivo inteiro em formato FASTA de cada um dos *supercontigs* selecionados (Tabela 2.2) para este tipo de análise foi submetido ao programa para análise de longas terminações repetitivas (LTR).

Tabela 2.2 – Listagem de *supercontigs* selecionados para análise de integridade por meio do programa *LTR-Finder* de cada uma das três espécies de *Puccinia* utilizadas para comparar com *P. psdii* MF-1

Espécie	***Supercontigs*** **selecionados**[1]	**Tamanho (pb)**	**Classificação (*RepeatMasker*)**
***Puccinia graminis f.* sp. tritici**	PGR_Supercontig2.13	1.556.540 pb	Copia
	PGR_Supercontig2.15	1.374.611 pb	Copia
	PGR_Supercontig2.16	1.217.956 pb	Copia
	PGR_Supercontig2.17	1.242.959 pb	Copia
	PGR_Supercontig2.148	64.884 pb	Copia
Puccinia triticina	PT_Supercontig2.1	3.059.345 pb	Copia
	PT_Supercontig2.10	1.388.629 pb	Copia
	PT_Supercontig2.14	1.210.875 pb	Copia
	PT_Supercontig2.111	309.745 pb	Copia
	PT_Supercontig2.120	278.607 pb	Copia
Puccinia striiformis	PST_Supercontig1.1	1.913.627 pb	Copia
	PST_Supercontig1.10	1.236.861 pb	Copia
	PST_Supercontig1.11	1.262.058 pb	Copia
	PST_Supercontig1.101	330.717 pb	Copia
	PST_Supercontig1.1212	1.461 pb	Copia

Nota: [1] Nomenclatura utilizada para se referir às sequências de *supercontigs* dos respectivos fungos causadores de ferrugem, sendo PGR – *P. graminis*; PT – *P. triticina* e PST – *P. striiformis*. A numeração ao lado da palavra "*Supercontig*" foi a mesma encontrada no arquivo destas sequências no banco *Puccinia Group*

2.2.8 Análise comparativa da proteína transcriptase reversa

Os 11 *contigs* do fungo *P. psidii* MF-1 que apresentaram a região completa do domínio proteico para transcriptase reversa na anotação manual por BLASTX (item 2.2.5) foram submetidos à comparação com sequências desta proteína depositadas no GenBank por meio da ferramenta BLAST pelo NCBI (http://www.ncbi.nlm.nih.gov).

Os *contigs* de *P. psidii* MF-1 00084; 00088; 00090; 00092; 00095; 00096; 00104; 00107; 00110; 00138 e 00238 junto com os *supercontigs* 2.111 de *P. triticina* e 2.13 (elemento anotado como de número 6) de *P. graminis* foram analisados na comparação de transcriptase reversa.

As sequências de outros fungos fitopatogênicos utilizadas na comparação encontram-se disponíveis no GenBank pelos respectivos números de acesso e identificadores *GenInfo Identifier* (gi): *Magnaporthe grisea* (T18350 e gi|7493955), *Tricholoma matsutake* (BAA78624.1 e gi|5002509), *Colletotrichum gloesporioides* (AAG24792.1 e gi|10946132), *Melampsora lini* (AAS66950.1 e gi|45504831) e o patégeno *Fusarium oxysporum* (AAB33090 e gi|10764493) contendo uma sequência conservada de transposase.

Realizou-se um alinhamento de sequências por meio do programa MEGA v. 5.05, para a construção da árvore filogenética utilizou-se o teste UPGMA e as sequências proteicas foram agrupadas por teste de *Bootstrap* com 1.000 replicações utilizando o Modelo de Poisson. O grupo externo utilizado foi a sequência de uma proteína transposase (presente em elementos da Classe II – DNA Transposons) do fungo *F. oxysporum*, pertencente ao filo Ascomycota

2.3 Resultados

2.3.1 Mineração de sequências relacionadas a elementos transponíveis

A categorização dos *contigs* de acordo com a função molecular demonstrou a predominância de GO termos relacionados a ligação a cromatina - GO:0003682 (20% dos *contigs*), nucleotídeo - GO:0000166 (28% dos *contigs*) e DNA - GO:0003677 (23% dos *contigs*), que estão diretamente associados a grande presença de transposons no genoma desse fungo (Figura 2.1). Esse fato foi observado pelo grande número de sequências "blastadas" com elementos transponíveis durante o processo de anotação.

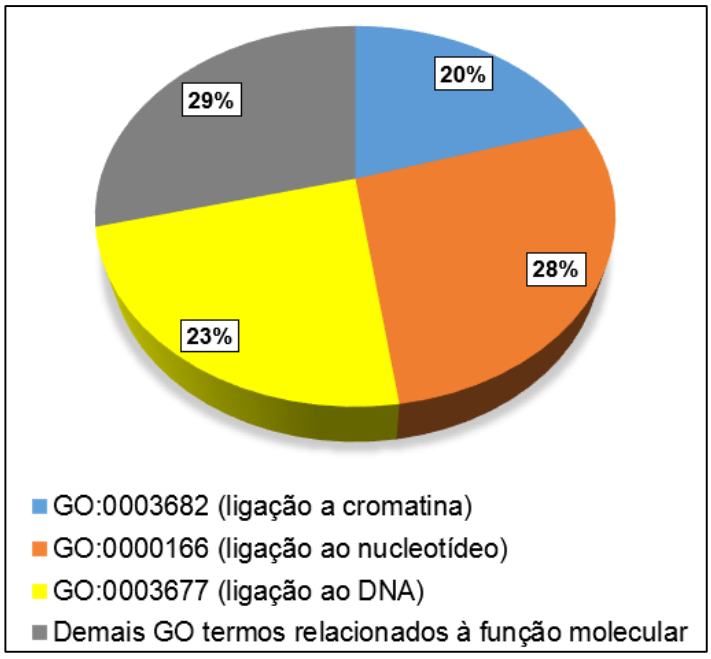


Figura 2.1 – Anotação funcional do GO termos relacionados às funções moleculares anotadas pelo programa Blast2GO. Os *contigs* são provenientes da montagem pelo programa *Newbler Assembler* 2.6. O gráfico representa os três principais GO termos relacionados aos processos de transposição no genoma, GO:0003682, GO:0000166 e GO:0003677. Os valores foram gerados pelo Blast2GO com corte de GO termos com menos de 5 sequências, independentemente do nível do termo anotado

Nesta estratégia adotada, os *contigs* já anotados pelo Blast2GO com termos conhecidamente relacionados a elementos de transposição foram minerados, totalizando 50.476 *contigs* que foram agrupados em 30 "grupos" de acordo com a descrição da anotação pelo Blast2GO.

O grupo 18 caracterizado pela presença do termo *Polyprotein*, presentes em retrotransposons, na anotação automática pelo Blast2GO continha a maior quantidade

de *contigs* minerados, 10.935 *contigs* sendo 16,4% dos *contigs* apresentavam mais de 1.000 pb.

Os grupos 25 e 26 contendo a anotação do Blast2GO como enzima transcriptase reversa apresentaram respectivamente 5.279 e 5.301 *contigs*, neste caso cerca de 10% dos *contigs* eram maiores do que 1.000 pb. Os grupos 12 e 13 agrupando anotações de enzima retropepsina e integrase apresentaram 3.158 e 3.738 *contigs*, respectivamente, sendo que cerca de 15% dos *contigs* eram maiores do que 1.000 pb. Finalmente, os grupos 1, 27 e 28 continham a menor quantidade de *contigs* minerados, 13 e 11 *contigs* e a maior parte deles contendo entre 250 a 1.000 pb (Tabela 2.3).

Tabela 2.3 – Mineração de *contigs* possivelmente relacionados a elementos transponíveis

(continua)

Grupo	Termo anotado no Blast2GO	Quantidade de *contigs*	Tamanho	
			250 até < 1000	> 1000
1	*ac transposase*	13	7	6
2	*ac-like transposase*	17	12	5
3	*copia ltr rider*	83	38	45
4	*copia polyprotein*	487	313	174
5	*copia protein-like*	16	10	6
6	*copia-like*	175	106	69
7	*copia-type proteins*	221	121	100
8	*gag polymerase*	288	270	18
9	*gag -int-pol protein*	134	80	54
10	*gag-pol*	2.541	2.081	460
11	*Gypsy*	1.496	1.221	275
12	*hiv-1 retropepsin*	3.158	2.667	491
13	*Integrasse*	3.738	3.252	486
14	*ltr retrotransposon*	40	18	22
15	*mutator-like pao retrotransposon*	349	276	73
16	*peptidase*	34	34	
17	*pol protein*	2.109	1.756	353
18	*Polyprotein*	10.935	9.136	1.799

Tabela 2.3 – Mineração de *contigs* possivelmente relacionados a elementos transponíveis
(conclusão)

Grupo	Termo anotado no Blast2GO	Quantidade de *contigs*	Tamanho	
			250 até < 1000	> 1000
19	*polyprotein of retroviral*	1.272	1.231	41
20	*related to ty3b-ty3b protein*	119	106	13
21	*retroelement polyproteins*	15	9	6
22	*retrotransposable* elemento	4.801	4.004	797
23	*Retrotransposons retrovirus-related pol*	5.965	4.785	1.180
24	*polyprotein from transposon*	530	355	175
25	*rnase h* e *transcriptase reverse transcriptase*	5.279	4.557	722
26	*proteins reverse transcriptase* e	5.301	4.567	734
27	*retroelement pol polyprotein*	11	8	3
28	*tc1-like transporase*	11	6	5
29	*Transposase*	1.204	988	216
30	*Transposon*	134	72	62

Os grupos tiveram suas respectivas sequências exportadas em formato FASTA, foram remontados e então classificados pelo programa *RepeatMasker* (item 2.2.2).

A primeira montagem realizada com os *contigs* minerados, denominada Tentativa I resultou em 2.228 "*supercontigs"* (assim denominados pelo grupo de pesquisa para fins didáticos, uma vez que contém maior número de pares de bases do que os *contigs* resultantes da montagem dos *reads* do sequenciamento). A Tentativa II gerou 586 *supercontigs* e finalmente a Tentativa III gerou 879 *supercontigs* (Tabela 2.4). A Tentativa II apresentou valor N50 de 2.779 pb, considerado superior comparado as montagens Tentativa I e III, cujos valores de N50 foram equivalentes a 1.601 pb e 1.889 pb.

Por meio das três montagens de *contigs* utilizando diferentes parâmetros, a montagem gerado pela Tentativa II foi considerada a mais “adequada”, pois apresentou o maior valor de N50 equivalente a 2.779 pb.

Tabela 2.4 – Resultados obtidos de acordo com parâmetros utilizados no programa *CLC Genomic Workbench* 6.0 para as três tentativas de *de novo assembly*

Montagem	**N50 (pb)**	***Supercontigs* gerados**	**Bases montadas em “*supercontigs*” (pb)[1]**
Tentativa I	1601	2.228	30.580.549
Tentativa II	2779	586	7.427.173
Tentativa III	1889	879	3.819.102

Nota: ([1]) A denominação “*supercontig*” para estas sequências foi assim feita pelo grupo de pesquisa para fins didáticos, uma vez que contém maior número de pares de bases do que os *contigs* resultantes da montagem dos *reads* do sequenciamento

Por meio da comparação de sequências depositadas no banco de elementos transponíveis do programa *RepeatMasker*, os 586 *supercontigs* gerados pela montagem Tentativa II (item 2.2.3) resultaram em 1.376 *hits* relacionados a elementos de transposição, sendo que 574.865 nucleotídeos foram identificados como sendo fragmentos de alguma família de elemento de transposição (Tabela 2.5).

Dentre os elementos repetitivos classificados pelo programa, os mais abundantes foram os LTR Retrotransposons, representando cerca de 99% dessas sequências no genoma, principalmente do tipo Copia (1.119 *hits* com 481.982 nucleotídeos), seguidos por DIRS com 11 *hits* e Gypsy com 17 fragmentos. Dentre os transposons de DNA classificados, representando apenas 0,60% do genoma, os mais abundantes foram os da família Mariner/Tc1 com 13 *hits* correspondentes a 2.215 nucleotídeos (Tabela 2.5).

Tabela 2.5 – Resultados obtidos pela classificação no software do *RepeatMasker* dos *contigs* gerados pelo *De Novo Assembly* com os parâmetros da Tentativa II. As sequências foram comparadas com o banco de dados *TE library (GIRI)*

Classe de Elemento Transponível	**Número de *hits***	**Tamanho (pb)**	**Frequência (%)[1]**
DNA transposon	29	3.395	0,60
Crypton	1	25	-
EnSpm/CACTA	3	495	-
Harbinger	3	176	-
Mariner/Tc1	13	2.215	-
MuDR	3	130	-
P	1	40	-
hAT	3	189	-
LTR Retrotransposon	1.330	568.398	98,87
Copia	1.119	481.928	-
DIRS	11	1.249	-
Gypsy	192	84.133	-
Retrotransposons Não-LTR	17	2.895	0,50
I	1	80	-
Tad1	13	2.171	-
Repetições interspersas	3	177	0,030
Elementos transponíveis	1376	574.688	99,96

Nota: (1) Tamanho dos fragmentos que apresentaram *hits* com TEs no *RepeatMasker* e o total de nucleotídeos que foram analisados (574.865 nucleotídeos)

2.3.2 Anotação direta dos *contigs* de *P. psidii* MF-1

A segunda estratégia de anotação de elementos transponíveis de *P. psidii* MF-1 se deu pela utilização do programa *RepeatMasker* para efetuar a classificação de forma direta de 167.538 *contigs* (resultantes do *Newbler Assembler* 2.6).

Foi observado que 10% do genoma parcialmente sequenciado do fungo foi anotado como elemento transponível, destes, 3% correspondem a elementos considerados desconhecidos (Tabela 2.6).

A quantidade de LTR Retrotransposons foi de 66 *hits* representando 0,02% do genoma, enquanto a proporção composta por DNA transposon, 4 *hits* apresentaram frequência praticamente insignificante no genoma (Tabela 2.6).

A primeira estratégia gerou resultados (item 2.3.1) mais refinados com relação à anotação de elementos transponíveis, comparando as duas estratégias observa-se que a composição de diferentes classes de TEs se mantém. Em ambas, ocorre predominância de LTR Retrotransposons. Apesar disso, a anotação gerada pela segunda estratégia (este item) é que foi utilizada na análise comparativa de composição de TEs com demais *Puccinia* spp. As sequências de *supercontigs* destas espécies não foram remontadas.

Tabela 2.6 **–** Resultados obtidos pela classificação no *software* do *RepeatMasker* dos 167.538 *contigs* gerados pelo *De Novo Assembly* do genoma parcialmente sequenciado de *P. psidii* MF-1

Classe de Elemento Transponível	**Número de *hits***	**Tamanho (pb)**	**Frequência (%)**
DNA transposon	4	238	0,00
TcMariner-Tigger	0	0	-
hAT-Charlie	0	0	-
LTR Retrotransposon	66	6585	0,02
ERVL	0	0	-
ERVL-MaLRs	0	0	-
ERVL-classe I	0	0	-
ERVL-clase II	0	0	-
Retrotransposons Não-LTR	0	0	0,00
SINEs	0	0	-
LINEs	0	0	-
L3/CR1	0	0	-
Penelope	0	0	-
Elementos Desconhecidos	4.808	989.760	3,01
Total de Elementos anotados[1]	21.279	32.882.244	10,05%

Nota: ([1]) Demais elementos não listados nesta tabela foram anotados como “DNA repetitivo” (elementos inter-dispersos; RNA pequeno; DNA satélite; repetições simples; repetições de baixa complexidade) e portanto não necessariamente elementos de transposição

Sinais utilizados:

(-) não foi encontrado nenhum elemento

2.3.3 Análise da integridade de elementos transponíveis de *P. psidii* MF-1

A busca por regiões terminais da família LTR - Retrotransposons pelo *LTR-Finder* e NCBI não permitiu a detecção de alguma sequência relacionada a longas terminações repetitivas (LTRs) nos 20 *contigs* analisados (Tabela 2.7) que foram selecionados a partir da montagem do genoma parcial de *P. psidii* MF-1 pelo programa *Newbler Assembler* 2.6 (item 2.3.2).

Similarmente, a utilização dos programas *TIR Finder* e NCBI para a busca por de regiões terminais da família Tc1-Mariner também não se mostrou eficiente, pois não foram detectadas sequências relacionadas a TIRs nos 10 *contigs* analisados (Tabela 2.8).

Tabela 2.7 – Classificação de 20 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*

(continua)

Contig	**Tamanho (pb)**	**Posição na Sequência**		**Fita**	**Classificação (*RepeatMasker*)**	**Identificação de LTRs**	
		Início	**Término**			**LTR-Finder**	**NCBI**
00052	3.773	3	3.763	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00075	3.535	20	3.411	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00083	3.415	1	3.415	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00111	3.242	12	3.214	+	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00116	3.221	6	3.212	+	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00146	3.069	728	3.069	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00151	3.043	424	3.041	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00152	3.058	61	3.027	C	LTR/Gypsy	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00153	3.062	11	3.061	C	LTR/Gypsy	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00157	3.018	2	2.325	+	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00167	2.973	6	2.902	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00229	2.824	23	2.823	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência

Tabela 2.7 – Classificação de 20 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*

(conclusão)

Contig	**Tamanho (pb)**	**Posição na Sequência**		**Fita**	**Classificação (*RepeatMasker*)**	**Identificação de LTRs**	
		Início	**Término**			**LTR-Finder**	**NCBI**
00251	2.743	1	2.741	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00263	2.754	57	2.752	C	LTR/Gypsy	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00306	2.656	1	2.567	C	LTR/Gypsy	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00378	2.542	16	2.542	+	LTR/Gypsy	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00395	2.478	112	2.473	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00401	2.454	2	2.395	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00443	2.444	3	2.404	C	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência
00450	2.431	41	2.431	+	LTR/Copia	Nenhuma sequência LTR identificada	Não há LTR flanqueando a sequência

Observações:
* As sequências dos *contigs* foram comparadas com o banco de dados FNGREP.ref que inclui 1.726 elementos transponíveis identificados em Fungi, inclusive elementos repetitivos de *P. triticina, P. graminis* e *P. striiformis*;
** As repetições estão apresentadas pela posição que ocupam em cada *contig*, respectivamente, fita de DNA, classificação do elemento pelo *RepeatMasker* e identificação de sequências LTR pelo LTR-Finder e NCBI por pareamento

Nota: Sinais utilizados:
(+) fita de DNA no sentido 5' → 3'
(C) fita de DNA no sentido 3'→ 5'

Tabela 2.8 – Classificação de 10 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*

(continua)

Contig	**Tamanho (pb)**	**Posição na Sequência**		**Fita**	**Classificação (*RepeatMasker*)**	**Identificação de TIRs**	
		Início	**Término**			**Repeat Finder**	**NCBI**
00032	5.507	1.308	1.355	+	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00047	5.129	1.134	1.209	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00175	3.942	3.279	3.409	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00446	3.065	100	569	+	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00463	3.034	1.183	1.269	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00501	2.977	1.166	1.226	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00630	2.785	1.053	2.785	+	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
00778	2.656	1.126	1.825	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência

Tabela 2.8 – Classificação de 10 *contigs* de *P. psidii* MF-1 em elementos repetitivos pelo *RepeatMasker*

(conclusão)

Contig	**Tamanho (pb)**	**Posição na Sequência**		**Fita**	**Classificação (*RepeatMasker*)**	**Identificação de TIRs**	
		Início	**Término**			**Repeat Finder**	**NCBI**
00886	2.569	685	911	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência
01210	2.378	1.878	1.956	C	Mariner/Tc1	Nenhuma sequência TIR identificada	Não há TIR flanqueando a sequência

Observações:
* As sequências dos *contigs* foram comparadas com o banco de dados FNGREP.ref que inclui 1.726 elementos transponíveis identificados em Fungi, inclusive elementos repetitivos de *P. triticina, P. graminis* e *P. striiformis*;
** As repetições estão apresentadas pela posição que ocupam em cada *contig*, respectivamente, fita de DNA, classificação do elemento pelo *RepeatMasker* e identificação de sequências TIRs pelo TIR Finder e NCBI por pareamento

Nota: Sinais utilizados:
(+) fita de DNA no sentido 5' → 3'
(C) fita de DNA no sentido 3'→ 5'

2.3.4 Anotação manual de *contigs* classificados como elementos da Classe I – LTR Retrotransposons de *P. psidii* MF – 1

A utilização do programa *LTR-Finder* não se mostrou eficiente para a determinação de regiões importantes dos *contigs* classificados como elementos da Classe I – LTR Retrotransposons, pois nenhum dos 25 *contigs* apresentaram longas terminações repetitivas (LTRs). Os 25 *contigs* selecionados foram individualmente anotados por meio da ferramenta BLASTX pelo NCBI e Uniprot. A anotação consistiu na determinação de regiões relacionadas à transposição (Tabela 2.9). Dentre os 25 *contigs* anotados, observou-se que 9 *contigs* não apresentaram nenhum domínio proteico relacionado a transposição na região analisada.

Os demais 16 *contigs* foram analisados pelo NCBI e também pelo *ORF-Finder*, observou-se que 11 destes *contigs* (68%) apresentavam a região relacionada à proteina transcriptase reversa (região *pol)* inteiros (presença de códon de iniciação “ATG” e de terminação, sendo “TGA” na maioria dos *contigs*). Além disso, observou-se que 44% destas sequências analisadas (*contigs* 00081, 00083, 00084, 00088, 00090, 00092, 00094, 00095, 00096, 00107 e 00110) apresentaram todos os domínios proteicos da região *pol* (integrase, RNase H e transcriptase reversa) (Tabela 2.9).

Duas regiões importantes para a transposição encontradas nos *contigs* 00084 e 00126, respectivamente foram a UBN2_2 e CH (*Chromatin Organization Modifier*). A região UBN2_2 é uma família proteica encontrada exclusivamente em LTR Retrotransposons, especialmente do tipo Copia, sendo estudada em plantas e fungos (CDD: NCBI database). A região CH é conservada e apresenta cerca de 50 aminoácidos presente em uma variedade de proteínas cromossômicas, aparentemente tem um importante papel na organização do úcleo em organismos eucarióticos (CDD: NCBI database), também existem estudos indicando que esta região exerce um papel importante na transcrição de genes de transposição (SANTANA, 2009).

Tabela 2.9 – Anotação manual de 25 *contigs* de *P. psidii* MF-1 realizada pelo programa BLASTX pelo NCBI e Uniprot

(continua)

Contig	**Tamanho (pb)**	**Classificação (*RepeatMasker*)**	**Fita**	**Localização do elemento**	**Domínios encontrados por BLASTX**
00001	32.576 pb	Copia	+	20.282 – 20.385	-
					transcriptase reversa
00010	8.067 pb	Gypsy	C	1.452 – 7.921	Rnase H
00019	6.641 pb	Copia	+	3.724 – 3.773	-
00023	6.067 pb	Copia	C	889 – 2.027	RNase H
					transcriptase
00028	5.815 pb	Gypsy	+	2.030 – 4.790	reversa
00036	5.284 pb	Gypsy	C	2573 – 4.963	-
00048	5.073 pb	Copia	C	52 – 306	-
00051	5.032 pb	Copia	C	968 – 1.017	-
00055	4.990 pb	Copia	+	1.712 – 4.198	Integrase
					RNase H
00064	4.834 pb	Copia	C	239 – 3.184	Integrase
00066	4.727 pb	Gypsy	C	1.656 – 4.727	-
					RNase H
					Integrase
00075	4.682 pb	Copia	C	322 – 2.403	*gag*
					Integrase
00078	4.605 pb	Copia	+	433 – 4.091	RNase H
					transcriptase reversa;
					Rnase e
00081	4.597 pb	Gypsy	+	103 – 4.559	integrase
00082	4.573 pb	Gypsy	C	63 - 682	-
					Integrase
					Transcriptase reversa
00083	4.581 pb	Copia	+	970 – 3.784	RNase H
					Integrase
					RNase H
					Transcriptase reversa
00084	4.560 pb	Copia	C	342 – 4.533	UBN2_2*

Tabela 2.9 – Anotação manual de 25 *contigs* de *P. psidii* MF-1 realizada pelo programa BLASTX pelo NCBI e Uniprot

(continuação)

Contig	**Tamanho (pb)**	**Classificação (*RepeatMasker*)**	**Fita**	**Localização do elemento**	**Domínios encontrados por BLASTX**
00088	4.528 pb	Copia	+	1.212 – 3.620	Integrase Transcriptase reversa e RNase H
00090	4.499 pb	Copia	+	1.676 – 4.331	Transcriptase reversa RNase H Integrase
00092	4.481 pb	Copia	+	1.949 – 4.300	Transcriptase reversa RNase H Integrase
00094	4.474 pb	Copia	C	79 – 3.166	RNase H Transcriptase reversa Integrase
00095	4.489 pb	Copia	+	321 – 3.711	Integrase Transcriptase reversa RNase H
00096	4.455 pb	Copia	+	1.303 – 3.597	Integrase Transcriptase reversa RNase
00104	4.446 pb	Copia	C	1 – 4.026	RNase H Transcriptase reversa
00107	4.411 pb	Copia	C	204 – 2.755	RNase; transcriptase reversa integrase
00110	4.834 pb	Copia	C	150 – 3.114	RNase transcriptase reversa integrase
00113	4.363 pb	Gypsy	+	1.629 – 4.363	-

Tabela 2.9 – Anotação manual de 25 *contigs* de *P. psidii* MF-1 realizada pelo programa BLASTX pelo NCBI e Uniprot

(conclusão)

Contig	**Tamanho (pb)**	**Classificação (*RepeatMasker*)**	**Fita**	**Localização do elemento**	**Domínios encontrados por BLASTX**
000126	4.270 pb	Gypsy	+	207 – 4.230	*Chromatin organization modifier* (chromo)* transcriptase
000138	4.187 pb	Gypsy	C	180 – 2.188	reversa like
000170	3.982 pb	Gypsy	+	971 – 2.204	- transcriptase reversa e
000238	3.630 pb	Gypsy	C	731 – 3.186	Rnase

Notas: Sinais utilizados:
(+) fita de DNA no sentido 5' → 3'
(C) fita de DNA no sentido 3'→ 5'
(-) não foram encontrados domínios proteicos

(*) domínios **Chromo** e **UBN2_2** não são diretamente relacionados a transposição, mas auxiliam na regulação de demais genes diretamente relacionados a transposição (região *pol*)

2.3.5 Análise comparativa da distribuição de TEs em *P. psidii* MF-1 e de outras espécies de *Puccinia*

A Tabela 2.10 apresenta os dados disponíveis de outros genomas de *Puccinia* obtidos por meio do acesso ao banco de dados do *Puccinia – Group Genome Database* do *Broad Institute*, sendo que a coluna "*RepeatMasker*" apresenta a porcentagem de elementos mascarados identificados pelo programa.

Assim o valor observado em *P. psidii* MF-1 está mais próximo da quantidade de elementos mascarados apresentados por *P. striiformis* pela mesma metodologia de análise, com 11,53% e 10,05%, respectivamente. Observa-se também que a concentração de GC (%) presente no genoma de *P. psidii* é de 33,7%, valor considerado mais baixo com relação as demais *Puccinia* spp., sendo o menor encontrado em *P. graminis*, cuja concentração de GC (%) é de 43,3% (*Broad Institute*), quando comparado aos outros dois genomas.

Tabela 2.10 - Informações gerais dos genomas de *P. graminis; P. triticina* e *P. striiformis* retiradas do banco de dados do *Puccinia - Group do Broad Institute* e *P. psidii*

Organismo	Cobertura do Genoma	Tamanho do Genoma (Mb)	GC (%)	Quantidade de *Contigs*	RepeatMasker (%)
P. graminis [a]	12 X	81,52	43,35	4.557	18,06
P. triticina [a]	31 X	106,57	46,72	24.838	16,31
P. striiformis [a]	81 X	117,31	44,43	17.295	11,53
P. psidii [b]	10 X[c]	ND	33,7	167.538	10,05

Notas: Letras utilizadas na tabela representam:
[a] banco de dados *Puccinia - Group do Broad Institute;*
[b] As informações apresentadas de *P. psidii* são estimativas obtidas até o presente momento pelo grupo, como a cobertura do genoma;
[c] Cobertura referente aos dados obtidos até o presente momento pelo sequenciamento por meio da plataforma 454 Roche;
ND - não determinado

Observou-se que os 167.538 *contigs* de *P. psidii* MF-1 classificados pelo programa *RepeatMasker*, apresentam um predomínio da Classe I – LTR Retrotransposons com 0,02% de frequência no genoma, comparado a uma fração bem menor de elementos da Classe II – DNA Transposons representados por apenas 238 pb distribuídos em 4 *hits* de elementos, que conferem 0,00% à classe, ou seja, insignificante (Tabela 2.11). Diferentemente, nos genomas de *P. graminis; P. triticina* e *P. striiformis* retiradas do banco de dados do *Puccinia - Group do Broad Institute,* em que se observa o predomínio de elementos da Classe II – DNA Transposons, compondo em média 2% de todo o genoma das três *Puccinia* avaliadas. Finalmente a fração de elementos desconhecidos representa a maior parte, sendo de 8,6% para *P. graminis;* 8,15% para *P. triticina;* 5,5% para *P. striiformis* e 3% para *P. psidii* (Tabela 2.11).

Tabela 2.11 – Distribuição e anotação de principais classes de elementos transponíveis em *P. graminis; P. triticina* e *P. striiformis* retiradas do banco de dados do *Puccinia - Group do Broad Institute* e *P. psidii* MF-1

Espécie	LTR Retrotransposons			Retrotransposons não-LTR	DNA Transposons			Elementos Desconhecidos		
	Número de elementos	Tamanho (pb)	Frequência no genoma (%)	Frequência no genoma (%)	Número de elementos	Tamanho (pb)	Frequência no genoma (%)	Número de elementos	Tamanho (pb)	Frequência no genoma (%)
P. graminis*	843	136.346	0,15	0,00	11.256	1.781.712	2,01	22.960	7.613.551	8,59
P. triticina*	1.760	281.842	0,21	0,00	20.841	2.750.176	2,03	30.946	11.036.850	8,15
P. striiformis*	2.002	560.656	0,48	0,00	22.440	3.071.228	2,62	23.235	6.418.652	5,47
***P. psidii* MF-1**	66	6.585	0,02	0,00	4	238	0,00	4.808	989.760	3,01

* banco de dados *Puccinia - Group do Broad Institute;*

2.3.6 Análise da integridade de elementos transponíveis de *Puccinia* spp.

As sequências classificadas como elementos Copia da Classe I das três espécies de *Puccinia* foram analisadas pelo programa *LTR-Finder* e apresentavam pelo menos 1.072 pb e no máximo 18.688 pb (Tabela 2.12). Os elementos da família Copia foram escolhidos para as análises, pois na anotação manual de elementos da Classe I de *P. psidii* MF-1, a maioria dos elementos analisadas com domínios proteicos pertencentes a região *pol* eram pertencentes à família Copia, como o objetivo é a realização de uma comparação entre estas sequências então utilizou-se somente sequências Copia das três *Puccinia*.

Os elementos desta família apresentam longas terminações repetitivas (LTRs) diretas conservadas com tamanhos variados dentre as sequências analisadas e duas ORFs (*Open Reading Frames*, ou seja, sequências de leitura aberta) a *gag* e *pol*, sendo que esta última é composta por sequencias de proteínas: protease, transcriptase reversa, RNase H e integrase.

A maioria das LTR 5’ e 3’ apresentaram identidade superior a 97% (Tabela 2.12), observou-se que as LTRs apresentam um consenso 5’ – TG...CA – 3’. Foram observados também sinais de inserção do elemento transponível, neste caso, representado pelas TSR (*Target Site Repeat*), que são duplicados no momento da inserção, estes sítios apresentam 4 até 6 pb e flanqueiam a LTR 5’ e a LTR 3’. Nas 15 sequências de *Supercontigs* analisadas das três *Puccinia*, diferentes elementos Copia foram analisados pelo *LTR-Finder*, os sítios TSR encontrados em *Puccinia* foram bastante diferentes entre eles.

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continua)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
Pst_Supercontig_1.1						
[1]	CCCCC	TGTAAGGGTC	2.117	AGGGCTTACA	348	97,1
[2]	GGCTG	TGTAAACCTT	2.447	AACCCTTACA	379	99,5
Pst_Supercontig_1.10	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	AGAAC	TGAAGAGAT	1.058	AGTTTCAACA	302	99,7
Pst_Supercontig_1.11	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	CAGTA	TGCTAGAA	19.752	TACTCCTACA	104	98,1
[2]	TTTT	TGGTCATCTC	16.850	GATTCATCA	287/292	92,1
[3]	GAAGT	TGTAAGGACT	2.105	AGGGGTTACA	327	100
[4]	GCAG	TGTGGAGCCC	4.092	TGCCGGATCA	421/431	83,4
[5]	CTTGC	TGTAAGAGCT	2.915	GGCTTACA	108	93,7
[6]	ACTTT	TGTGGTGACA	2.346	GGTCACAACA	124	98,4
[7]	CCTG	TGGTGACAAG	1.497	GGTCACAACA	122	98,4
[8]	AAGGC	TGTAAGGATT	2.263	AGGGCTTACA	329	98,8

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continuação)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
Pst_Supercontig_1.101						
[1]	GGAAT	TGTCATACGA	**11.715**	CCGTGTGACA	**867**	**96,3**
Pst_Supercontig_1.1212	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	CCCCC	TGTAAGGGTC	2.117	AGGGCTTACA	348	97,1
[2]	GGCTG	TGTAAACCTT	2.447	AACCCTTACA	379	99,5
[3]	CTGAG	TGAGGGAAGA	1.046	ATTTCAACAC	293/287	94,9
PG_Supercontig_2.13	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	GTTCC	TGTCAGCCGG	1.770	ACCGCTGACA	498	100
[2]	TGCAT	TGTTAATTGT	4.048	GATCTCAACA	622/627	98,6
[3]	GAATC	TGTTGAGGGG	1.239	CATTCCAACA	210	100
[4]	CGGAC	TGTTGAGTAT	4.874	GATCTCAACA	685/686	99,4
[5]	ACGT	TGTAAAGCTC	9.607	GTATACCACA	1089/1090	99,9
[6]	TGTTG	TGTCAGCCTA	6.933	CAAGCTGACA	1579/1580	98,9
[7]	GATAG	TGTAAAACCC	1.956	AGTCCTTACA	341	99,7

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continuação)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
PG_Supercontig_2.13						
[8]	GCGC	TGTGGAACTA	8.292	GCTCACCACA	754	100
[9]	ACATA	TGTAAGAAAG	4.426	AACTCTAACA	320	98,1
[10]	TAGCG	TGAGTGCTT	7.545	CCAAAAAGCA	159/166	87,3
PG_Supercontig_2.15	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	CCAC	TGAGCTGGTT	7.328	GCCTGGACCA	639/651	94,3
[2]	AACAA	TGATGGCAAT	8.031	AAAACAATCA	1541/1538	99
[3]	CCCCT	TGTAGTGACT	1.788	AGTCACAACA	105	99
[4]	CACCGG	TGTAAGACTC	10.603	TTGACTACCA	908	100
[5]	TGCCG	TGTAAGAACC	4.833	AGTCATTACA	436	99.8
[6]	GAAAC	TGTTGAGGTG	12.333	AAATCCAACA	282	98,9
[7]	GATTG	TGTTGAGTTT	4.658	AATCTTTTCA	726	100
[8]	AGCTC	TGAAAGATTT	4.167	AATTCCAACA	818	100
[9]	CTTC	TGTCAGCCTC	17.236	GTGGCTGACA	1085	98,4
[10]	CCGT	TGTGGTAGTC	6.377	GAGTCTTACA	930	99,2

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continuação)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
PG_Supercontig_2.16						
[1]	CAAGC	TGTTGAGGTG	1.269	CATTCCAACA	182	98,9
[2]	ATATT	TGATGGAATT	1.045	AATTCCAACA	365	100
[3]	AAATG	TGTTGGAATT	18.960	AATTCTTTCA	858	100
[4]	CCTT	TGTATCATAT	10.349	ACCATCAACA	617	98,5
[5]	AATTC	TGTTGGAATT	3.982	TTTTTCATCA	879	99,7
[6]	CCGG	TGTGGTAAAT	10.055	GAGTCTTACA	970	99,9
PG_Supercontig_2.17	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	CTCTG	TGTTGTAACC	5 682	GGTGCGAACA	166	100
[2]	CTTAC	TGTGAGCCTG	6.402	AAGGCTGACA	975	100
[3]	ATTTT	TGTTGAAAAA	1.037	AAATTCAACA	322	99,7
[4]	AGCTC	TGTAAGGACT	2.101	AGGGCTTACA	328	99,1
[5]	CCGT	TGTGGTCATT	9.380	GAGTCTTACA	1010	96,3
[6]	CTTAT	TGTCAGCTGG	1.730	ACCGCTGACA	499	100
[7]	ATTTT	TGTAAACCCC	2.035	AGTCCTTACA	318	100

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continuação)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
PG_Supercontig_2.17						
[8]	CCTTT	TGTCAGCCAC	4.887	GAGGCTGACA	1324	99,6
[9]	GCGG	TGTAAGACTC	9.764	GACTACCACA	910	96,6
[10]	CCAAG	TGTTGGAGTG	10.686	TGTCACAACA	123	99,2
[11]	GTGCG	TGTAAAGAAC	6.526	GCTTTTGACA	399	99,5
[12]	CTGAG	TGAAGAAACT	1.091	AATTCCAACA	390	100
[13]	CCGT	TGTGGTAATT	11.327	GAGTCTTACA	1066	99,8
[14]	CCGT	TGTGGAACTA	4.122	GCTTACCACA	953	99,1
PT_Supercontig_2.1	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	CACAA	TGTTGAATTA	1.323	ACCCTCAACA	192	99
[2]	GTGCA	TGTCAGCCGT	8.811	CGCCCTGACA	661/682	91,5
[3]	GCGG	TGTGGTGGAC	12.173	GAGCTCCTCA	771/784	96,4
[4]	AAATA	TGTTGAGGGA	1.191	CATCTCAACA	201	97,5
[5]	AGAGC	TGTTGAGGTT	1.072	GATTCTGACA	437/454	94,9
[6]	TGTT	TGGTATCAAT	12.447	GATTCTTACA	124	96
[7]	CTTGT	TGTGAGGGCC	4.418	AGTCGTTACA	438	98,6

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(continuação)

Sequência analisada	TSR	Início LTR 5'	Tamanho da sequência (pb)	Término LTR 3'	Tamanho LTR (pb)	% Similaridade
PT_Supercontig_2.1						
[8]	CTTGC	TGATGAGCTA	4.740	TATCTTAACA	348	100
[9]	AGAAC	TGAAAAAACA	5.772	AAATCTCACA	667	99,3
[10]	CTAAT	TGTGAAGACC	1.359	AACCCTTACA	329	98,2
[11]	CACGG	TGTTGGATGG	5.572	TGTCACAACA	123	100
[12]	GGCTG	TGTAACCCCC	2.535	GGCTCTTACA	327/329	97,6
[13]	AAGGG	TGTAGAACTG	6.724	TGATATAACA	225	98,7
PT_Supercontig_2.10	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	AAAAG	TGTTGGATTT	1.181	AATCCCTTCA	504	98,2
[2]	CGGAG	TGTAAGGGTT	2.475	GGTCTTCACA	329	99,1
[3]	GACC	TGTTGTAACC	6.230	AATCACTACA	112	97,3
PT_Supercontig_2.14	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	GTTCC	TGTAAGCGGT	2.582	GGGGGGTTAC	337/343	97,1
[2]	GTCAC	TGTTAGGTCC	6.752	GCTCATAACA	195	97,9

Tabela 2.12 – Descrição da análise de integridade dos elementos pertencentes a família Copia da Classe I – LTR Retrotransposons de diversos *supercontigs* de *P. graminis*, *P. striiformis* e *P. triticina*, representados respectivamente por PG, Pst e PT

(conclusão)

Sequência analisada	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
PT_Supercontig_2.14						
[3]	GGGTC	TGTTGAGATG	1.193	TCCCTCAACA	201	99,5
[4]	GACTC	TGTTGTGGCA	7.348	GCGCTCTACA	240	97,5
[5]	GGAGA	TGAAAGATTT	4.183	TAATCTCACA	814	99
PT_Supercontig_2.111	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	GAAAG	TGTTGAGAAT	4.020	ACTCTCAACA	663	98,6
PT_Supercontig_2.120	**TSR**	**Início LTR 5'**	**Tamanho da sequência (pb)**	**Término LTR 3'**	**Tamanho LTR (pb)**	**% Similaridade**
[1]	GTAAG	TGTGAGGATC	18.688	ATTTGTTACA	448	98
[2]	ACTAC	TGTTGAGGTT	4.626	CACCGCAACA	299/297	99,3

Observações:

* A região TSR se refere à sequência olinucleotídica específica de cada elemento de transposição, composta por no máximo 6 nucleotídeos e que se apresenta duplicada quando ocorre a inserção do elemento;

** Apenas os 10 primeiros (LTR 5') e 10 últimos (LTR 3') nucleotídeos que compõem a região das LTRs (longas terminações repetitivas) estão sendo apresentadas na tabela, observa-se um consenso 5' – TG...CA – 3' para estas LTRs do tipo Copia;

*** O tamanho da sequência se refere à toda região flanqueada pelos LTR 5' e LTR 3', sendo que os respectivos tamanhos delas estão representados na coluna "Tamanho LTR";

**** A similaridade entre as repetições LTR 5' e LTR 3' estão apresentadas na última coluna, sendo superior a 97% na maioria dos elementos analisados

A análise da integridade de 79 elementos classificados como da família Copia pelo *RepeatMasker* pertencentes à *P. graminis*, *P. striiformis* e *P. triticina* revelaram que todos os elementos analisados pertencem pelo menos à Classe I com certeza, pois eles apresentam as principais regiões de um elemento deste tipo, TSR (duplicada) e consenso das regiões LTRs em 5' – TG...CA – 3', de acordo com Wicker *et al*. (2007) é o padrão comum encontrado em elementos LTR - Retrotransposons da Classe I, como Copia, Gypsy, Bel-Pao, Retrovirus e ERV.

A análise de sequências conservadas de proteínas dos 79 elementos de *Puccinia* spp. por meio de BLASTX via NCBI e Uniprot não foi realizada devido à este não ser um dos principais interesses do presente trabalho. Alguns destes 79 elementos foram analisados quanto à presença de sequências proteicas características de Copia. Aqueles elementos cuja região correspondente à transcriptase reversa estava completa foram selecionados para a etapa de análise comparativa entre enzimas de transcriptase reversa de *Puccinia* spp. e diversos fitopatógenos (item 2.2.8).

Sabe-se que os elementos da família Copia são formados por duas ORFs: *gag* e *pol,* sendo que esta última é composta por domínios de PR (protease), RT (transcriptase reversa), RH (RNase H) e integrasse (IN) no sentido 5' → 3' da fita. Wicker *et al*. (2007) discutem a presença de uma região à jusante do domínio da protease, denominada proteinase aspártica (AP) que não foi encontrada para os elementos transponíveis analisados de *Puccinia* spp. pela metodologia apresentada (Figura 2.2).

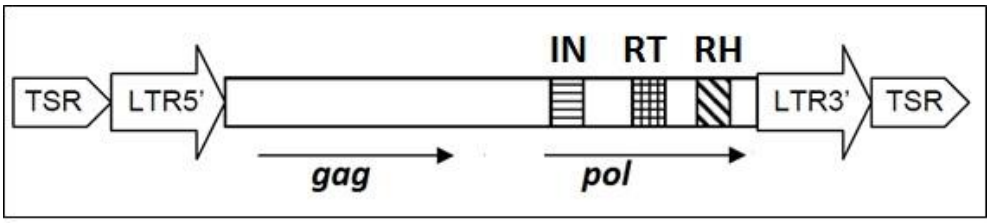


Figura 2.2 – Estrutura básica de um retrotransposon da família Copia. Regiões TSR duplicadas; 5' LTR e 3' LTR; a região *pol* contém os domínios PR (protease), RT (transcriptase reversa), RH (RNase H) e integrase (IN) e *gag* responsável pela proteína do capsídeo viral. Fonte: Adaptado de WICKER *et al*., 2007

Assim, pode-se afirmar que a anotação de 79 elementos transponíveis de *Puccinia* spp. com relação às regiões TSR e LTR foi efetiva e os elementos desta família podem ser representados pela anotação manual do Supercontig 2.111

pertencente à espécie *P. triticina* como sendo “padrão” para esta família de elementos (Figura 2.3).

A anotação de domínios proteicos do PT_Supercontig2.111 se deu na leitura da fita 3’ → 5’ em sua primeira *frame* de leitura (-1), assim sendo observa-se que o padrão esperado para elementos do tipo Copia (Figura 2.2) foi encontrado no elemento PT_Supercontig2.111 de *Puccinia* spp. anotado (representado por *P. triticina*) (Figura 2.3).

Todos os domínios proteicos relacionados à transposição e classificação do elemento foram encontrados na mesma *frame* de leitura (-1), a ordem dos domínios região pol são: IN (integrase), RT (transcriptase reversa) e RH (RNase H), esperada para elementos Copia, no sentido 5’ → 3’.

> PT_Supercontig_2.111

*GAAAG*TGTTGAGAATCAAACTGAAGTAATTGCAAAAGTTAGAAAAAGAAAAACAAAGAAGAATTTTGAGAGAAAGGAAAAATACCTGAAGGCAAACACAACACA
GAAGAGCACGACAAAACAAAGAGATGACACAGAGAAGAAAGAAAGAGCAGACTAAAGTAGAACAATGTTAATAGGGTCAGATGATGAGAGAAAGAGAAAACA
CGACATACGGCCAGGAGCAGACTTGGGCGGAGGAAGTGAGTGAGCTTGAGGTTTGAATTTGAAGACCGATTTGAAAGGTGTCTTCGGGAGGAACAGAGCAG
TCAGAGAGTTTAAGAAAGATTTGAGACTGCCAATGTAATGCGTCGAGTGATGATTCGAGTGAGTTCCCCAAGGGAGGAGGAAAGAGGCGAGATTCTGCGAGT
TTAAATATGTTGAGCGGTCAAGGATAGCTGATGAGAAGTTGTGGTTGAGAACAGTTCAGGTTTGACGGAAAGGGAGATGCGAATTGAGTTGTCCAAGAGCGT
GATGAGTGAGAATTGAGTTGTCCGAAATAGTGAAGATGAGTGAATAGAGGAAAAATAGAGAGAGAGCCGGAGAGGAATATGTTGTGTGACAGGTAACATGAC
AGGTCCATTACAGGCAACAAGACCTCAAGTTGAGGCGAGGGCTTGACTCTCAACACCCCCCCTTTCCTCTAGTCGCAAAAGTCCTATGGCGGATAGGGATTT
TACAAGAGCTGGGTGTGGGACCGATTTTGTTAGGAAGTCGGCTAGCATGTCAAAGGTGGGAGTGTACTGGAGAACAATCTTGGAAGATTTGATTACGTCGCG
AATGAAGTGGAGTTGAATTTCCACATGCTTCATTTTCTTTGAGTTGGTGTTTGCTTCAAAGTTTGCAACCGCGATGCACCCTTGATTGTCTTTGTGAACAACAAT
GGGCTCTTCAAGTTTGAAAGTTGAAGTTTGATGCAGAGTTGACGAAGCCAGAGTAGCTCGCAGGAGAAGTTGGTCAAGGCACGGTATTCCGCTTCGCAAGA
GGATAAGGAGACAACGGGTTGCTTCTTGGTTTGCCAACATACGAGGTGATTTTTGACCAATGAAAGGAAGCCGGTGACGGATCTCCGAGTGACCGGACAGTT
GCCCCAATCTGCGTCAGTGAATCCACAAAGAGGAGACGAAAGCCCGCGAGAGTACACCAAGCCGTAGTCTGGGGTACCGCTGAGATATCGTAGGACGTGTT
TGAAGGCGTCCCAGTGTTGCATGCCGGGATTCTCCAAAAACTGCGAAAGATTGCTTACAGCGTAAGCAATGTTGGGACGTGTTGCGGTGCTCAAGTAGCTTA
AAGCCCCAATGGCGCTCCGGTAATTTTTGCCAAGGGAGAGGAATTTCTTGATGTCTTCTTGTGTCGCCTTGCCTAGATGAGAGTTTGGGACGAGTGGAGTTAC
CGCGGGTCTGCACAAACTCATTCTGTACAGCTCCAAAACCGAGTTGACGTAGTGTTGTTGAGTGAGAATGATGGCGCCGGGTTGATGAAGGATTTTTATTCCG
AGTAAGATGTCTGCTTTACCAAGGTCTTTCATGCTGAATTCCTTTTTGATGGTGATTTTGAAAGGCTCTACTTGCGGTCCGAAGACAGCAATGTCATCGACATG
AATGAACAGATAAATCGGATTAGGTGTGAGGCAATAGAATACACATGGGTCGGAAACAGAACATGTGAAATCAACGCTGATTAGCCAAGCCGAGAGACGATT
GTACCACGCAAGGGGAGCCTGTTTGAGTCCATAGATTGCTTTGTGTAGTTTCAAACAGAAAACCTTTTTGTCGGCGGAGACTCCTTGAGGAATTGACAGATAG
ACCGACTCTTCAAGATCGGCGTTGAGGAAGGCGGTTTTTACATCTAATTGTTGAATATCCCAGTTCATGGCGGTAGCGAACAAAATGAGTACCAGAGTGAAT
TAAGTCATCCAGTGGGCGCAAATGTTTTAGAGAAGTCAACACCGTAGGTCTGCGAAAACCCCTGCGCGCATAGACGAGCTTTGAATTCAATGGAGCAAGAGC
CGTTTTTCTTCTTCAGGAACACCCAGGTTGTCCCGACTAGTTTCACGCCAGGAATGATTTTGATGATTGACCAAACTTTGAGAGCTTCCATGGCGTCCAATTCC
TTTTTGATAGCGGTGATCCATTGTTGAGATTCACGACACTTGATTGCTTGGTTGAAGTGTTTAGGAATATTGTCGGTGGCCTGCATAGCAAAGGCCCGAGGCC
TTCGAGAATGGTGAATGATGTTGTCGGTAGTGATGTCTCCTTGAATGATGTTGGAGGTGCTGACTTGAGGGATTTCCTGAGGGGAAGTATCAACGATCTGAGA
GGCGGGATCTTCAGGAATGGATTCCGATTTGGAACTATCTGAAGTGTCAGACTCCGAGGAAGAGTCGGAGTGTTCAGGACAGTCATGGAAATTGTCCTCATC
AGTGGAGTCGTCCGTTACAAGAAGAGAATTATTGACAACAAGGGGTTCATCACTGGAATCGGAATTGATGTTAGATCCGGATAGAGAAGGAAAAACTGATTCG
TTGAATGAAGCGTGCCTGGTACGAATAAGCTTCCCATTGCTAATTCTTATGACTCTGTAGACTGCCCTATCATTTTCAAATCCAACGAGGATACCCAATTCTCC
CGTGTTGGCCAGTTTCCAAGATTGGCGATGTTTTGGAACCAATATGGCAACTTGACAACCAAAAGCCTTGATGCGTCAAAGAGGAGCAGGCGAGTTGGTCCA
CAATTCAAAAGGGGATTTGTTTGCGCGAGAGGGTGTAGCAAGCAGGTTGGAGAGGAAAGTAGCAGTATTGACCGCCTCACCCCAATAGGTTTTCGGGAGCTT
GCTGGTGGAAAGTAGGCATTGAGCTTTGTCAAGAATTGTTCTGTTGGCCCGTTTGGCAAATCCGTTGAGTTCCGGGGTGTAGGGAGGAGAAAACATGTGATT
GATCCCCTTTTCTGCTACAAAGCTTTTGAAGTCATGATTTACAAATTCGCCGCCATTGTTGGAAACTATCTCTTTCACTCGTCGTTCCTGGATGTTTTCTACCAG
GGCAATAAGTAGCTTGAGGGACCCCAACGCTTCCGATTTTGCCTTTAGAAATCAAACAAATTCAAAAGATGAGAATTGATCAACAACAGTCATGAAAAATTTGA
ATCCGGAAAGAGTGGGAGGAGTGATTGGTCCAACTAAGTCTATGTGGAGGCGATGAAGTGGGAATGAAGCAGGAGAAAAATGGTTGGAGAAGGGAAGACGA
GTCAT**TTTGGAGCAAAGACACACTTCACAAGAGTCGTTCGGAGTAGAGGAGGACAATCTCATGAGTTTGAGCGTTTTTTCGCTTGGATGGCCGAGTTGGAG**
GTGCCAAATGTTTGGAGTCGTTGCGGATTTCGAGAGAAGAGCGACAGGATTAGAATAAGATGAGTGGAGAAGTCCGTTAACAATGTGACCGGTAAACATGA
CACTTTTTCCGTCGGAAACGGAGAAAGAGTCGGCCTGTTTTTGAATAATGAGTGTCCCAGGAAAAAGTAGACCAAGGGAAATTAATTGTTGCGATATGGAGGG
TACATATAGGCAATCCAAGAGCTCAATGAGTTTTCCATTCGAGAATATTTTTGCTGTGCCATGACCTTCAGCAAATAGAGAGTCGTTCGGATTCCCAGTGGTGA
TTGAAACTTCTTCATTCACTAGCTTGGTGAATAAAGAACAATCGCTGAGCATGTGATGTGTCGCCGCGGTGTCCATAACAAGTGTTTGAGTGGAAAGTCTTCC
GGATGAAACAATGAAGGCCGAGGCTTGGGCGAGAGAGGCACTTGCATTTTTTGGCTTCTGAGTTCAAAGATTGGGACACTTATGATAGCATCGGCTTTCAGG
GTGAGAGGTGGACAGCGGATTATGTTCGGTTCCTGCGCATAGATGAACAATCTTTTGAGGAAATTTGGAGGGAACAGCGGTGACAAGAGCGGAAGCAGATTG
TTTTCCATTATCTTGATGCATGGCTTTTGACAACGTGTGGGTTCTGAAGAGTCGGAGAGCATCGAGACAGCTATATGGTGTGGCAACGGTTTCTTCAGATAAG
GCAATCTTGTCCACGATCATGTCAACTTCCTTGTTCCCCATAAGTTTTCCGAGGATTTCGTAGGAGAGAACATCGCCAGGCATCTTCTTTTCAACAAGTTCGAT
TTCCCGAAGTTGTTTTCGGATAGTTTTGATGAAGAGTTGAAGATTGCCTGTGTACTTGATGGCTGACCATTTCATGAAGACGTGCCCTCAATTGATGATAGATT
GAGAGGCTTAACAGTCGGCAATTCTTTACCAAAGAAGGACAGCGCTGTCATGGGTTTCTGAGTTAACAACTTCATTAAAGCATTTCTCTTCCAGTCTGGCGGT
AATGAGAGAGATTGCGAGATAGTTTTCTTCATGAAAGGTAGTTTTGATGTCGGCATCTGCATCCTTATTGACAGTTTTTTGACAAATGTCGAAAAGCTTTCTTCC
TTGTAGGTAGATCATCATACGGTCGCTCCAATGCGCGTAGTTATCGCCATTGAGCAGTGGGATGTTTTGTCGGCTTAAGTTGTCGGCTCCATCGGAGTTGGAC
ATGTTGTGTGTTTGCCTTGGAGTTTAGATTCAAGGAATAGCGGTTCCGAGAACAAGACTGTAAATC*TGTTGAGAATCAAACTGAAGTAATTGCAAAAGTTAGAA*
AAAGAAAAACAAAGAAGAATTTTGAGAGAAAGGAAAAATACCTGAAGGCAAACACAACACAGAAGAGCACGACAAAACAAAGAGATGACACAGAGAAGAAAGA
AAGAGCAGACTAAAGTAGAACAATGTTAATAGGGTCAGATGATGAGAGAAAGAGAAAACACGACATACGGCCAGGAGCAGACTTGGGCGGAGGAAGTGAGT
GAGCTTGAGGTTTGAATTTGAAGACCGATTTGAAAGGTGTCTTTGGGAGGAACAGAGCAGTCAGAGAGTTTAAGAAAGATTCGAGACTGCCAATGTAATGTGT
CGAGTGATGATTCGAGTGAGTTCCCCAAGGGAGGAGGAAAGAGGCGAGATTCTGCGAGTTTAAATATGTTGAGCGGTCAAGGATAGCTGATGAGAAGTTGTG
GTTGAGAACAGTTCAGGTTTGACGGAAAGGGAGATGCGAATTGAGTTGTCCAAGAGCGTGATGAGCGAGAATCGAGTTGTCTGAAATAGTGAAGATGAGCAA
ATAGAGGAAAAATAGAGAGAGAGCCGGAGAGGAATATGTTGTGTGACAGGTAACATGACAGGTCCATTACAGGCAACAAGACCTCAAGTTGAGGCAAGGGCT
TGACTCTCAACA*GAAAG*

LTR 3’

RH

RT

IN

Gag

LTR 5’

Figura 2.3 – Representação de LTR - Retrotransposon por meio de sequência do Supercontig 2.111 da espécie *P. triticina,* o sítio alvo (TSR – *Target Site Repeats*) está em itálico, as longas terminações repetitivas (LTRs) estão representadas em cinza, a região *gag* e domínios PR (protease), RT (transcriptase reversa), RH (RNase H) e IN (integrasse) da região *pol* estão sublinhados. A fita 3’ → 5’ é aquela que contém os nucleotídeos para tradução das proteínas de transposição, neste elemento

A localização dos domínios juntamente aos seus respectivos valores de *E-value* dos domínios “blastados” e seus códigos de acesso estão representadas na Tabela 2.13. Dessa forma é possível inferir que a classificação dos elementos Copia realizada por meio do programa *RepeatMasker* se mostrou efetiva.

Tabela 2.13 – Domínios proteicos encontrados no elemento Copia presente no Supercontig 2.111 de *P. triticina*

Domínio proteico	**Código de acesso**	**Localização (pb)[1]**	**Tamanho (pb)**	***E-value***
RNase H (RH)	cd09272	725 – 1.150	426	$1.51\ e^{-28}$
Transcriptase reversa (RT)	pfam07727	1.415 – 2.128	713	$5.41\ e^{-53}$
Integrase (IN)	pfam00665	2.897 – 3.253	356	$6.70\ e^{-12}$
região *gag*	pfam13976	3.290 – 3.448	158	$1.67\ e^{-03}$

Nota: ([1]) Localização dos domínios conservados

2.3.7 Análise comparativa da proteína transcriptase reversa

A análise comparativa da proteína transcriptase reversa por meio da construção de uma árvore filogenética permitiu observar que as sequências analisadas formaram dois grandes grupos: A e B (Figura 2.4). Sendo que o grupo A pode ser sub-dividido em dois principais clados, A1 e A2. O grupo A1 é composto por 5 sequências de transcriptase reversa originárias dos seguintes *contigs* de *P. psidii* MF-1: Contig00084; Contig00088; Contig00095; Contig00104 e Contig00110 que agruparam à sequência desta proteína presente no Supercontig 2.111 da espécie *P. triticina* (Puccinia Group), sendo que as transcriptases dos contigs 00088 e 00110 de *P. psidii* MF-1, foram agrupadas em um mesmo clado com a sequência pertencente a *P. triticina* (Supercontig 2.111). As demais sequências presentes no grupo A1 provenientes de *P. psidii* MF-1 apresentaram-se agrupadas entre si.

O grupo A2 é formado por dois clados, um agrupando o Contig00096 e o Contig00107 e o segundo agrupando os *contigs* 00090 e 00092, todos pertencentes à *P. psidii* MF-1. Por fim, observa-se que o grupo A reúne todas as sequências acima descritas com a transcriptase reversa do fungo *T. matsutake*, pertencente ao filo Basidiomycota, ao qual pertencem a *Puccinia* spp..

O grupo B também pode ser analisado separadamente em grupo B1, em que as sequências proteicas de transcriptase de outros fungos fitopatogênicos se agruparam em um clado, sendo eles: *C. gloesporioides* e *M. grisea* ambos fungos do filo Ascomycota que agrupou com o fungo *M. lini*, pertencente ao filo Basidiomycota.

O grupo B2 agrupou as demais sequências proteicas de *P. psidii* MF-1, bem como a a sequência de transcriptase reversa do fungo *P. graminis* (elemento 6 do Supercontig 2.13 – *Puccinia Group*), demonstrando que deve haver três grupos distintos de transcriptase reversa no genoma de *P. psidii,* um mais similar à *P. triticina*, outro mais similar à *P. graminis* e outro que apresenta baixa similaridade com as transcriptases reversas dos demais fungos analisados.

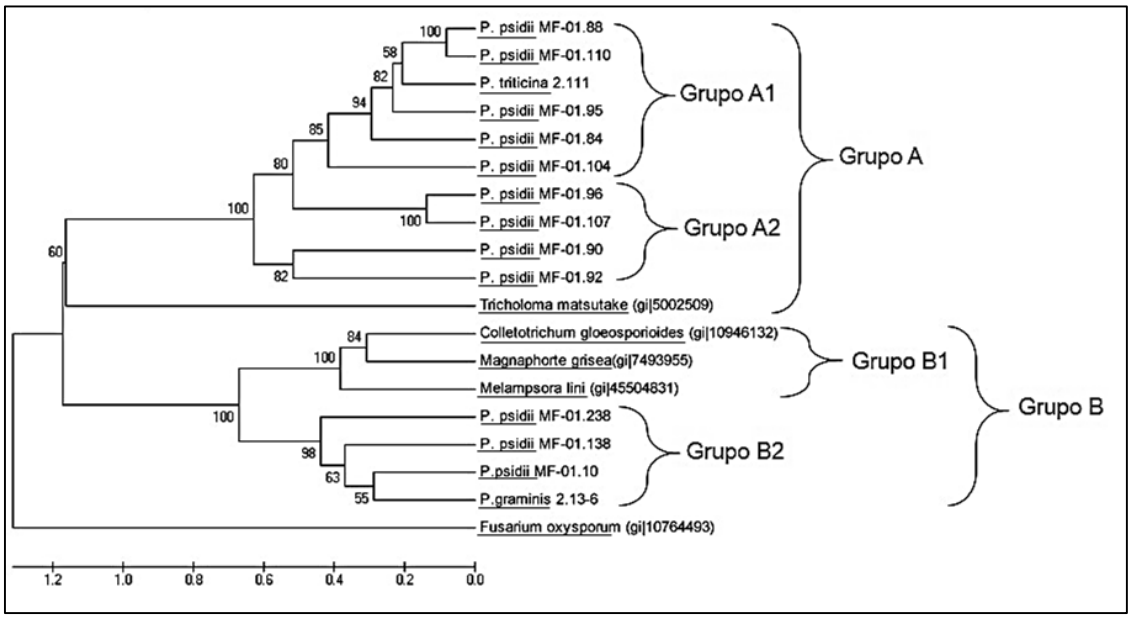


Figura 2.4 – Agrupamento da região correspondente a transcriptase reversa presentes nos 12 *contigs* de *P. psidii* MF-1 (classificados como Gypsy e Copia) com outras regiões desta proteína de outros elementos de transposição. A análise foi realizada pelo método UPGMA, o alinhamento das proteínas foi realizado por Clustal W com os parâmetros de penalidade padrão. A construção filogenética por UPGMA se deu com o método *Bootstrap* com 1.000 réplicas e no modelo de Poisson. Os números abaixo e acima dos nós indicam a porcentagem em que cada ramo apareceu em análise

2.4 Discussão

O sequenciamento obtido até o presente momento de diversos fungos causadores de ferrugem (Basidiomycota, Pucciniales) em diversas culturas têm revelado genomas de tamanho superior à média encontrada para genomas do reino Fungi que é de 44,2 Mb (TAVARES *et al*., 2014). O tamanho médio de genoma de

fungos Basidiomycota é de 49,9 Mb, enquanto a média do tamanho dos fungos causadores de ferrugem são em torno de 225,3 Mb. Dentro desse grupo, encontra-se um dos organismos com o maior genoma já descrito para fungos, *Puccinia chrysanthemi* cujo genoma possui 806,5 Mpb (TAVARES *et al*., 2014).

Destaca-se que fungos causadores de ferrugem têm o genoma cerca de cinco vezes maior do que a média dos fungos. Uma das hipóteses para esse fenômeno é que a ordem Pucciniales representa um grupo em que a expansão do tamanho do genoma é um evento comum e frequente, sendo que uma das razões para a expansão do tamanho destes genomas é devido à alta atividade de elementos transponíveis (TAVARES *et al*., 2014).

A atividade mutacional de TEs pode resultar em interrupção de genes e rearranjos cromossômicos, podendo contribuir com a diversidade genética dos respectivos hospedeiros. As inserções são potenciais alvos de processos epigenéticos e também podem resultar em novos padrões de expressão gênica (BLEYKASTEN-GROSSHANS; FRIEDRICH; SCHACHERER, 2013). Assim, o estudo de elementos de transposição em espécies de fungos fitopatogênicos, como *P. psidii* causador de ferrugem de mirtáceas, justifica-se, pois pode representar uma potencial ferramenta no desenvolvimento de mecanismos de controle das doenças causadas por estes organismos (RUIZ-SILVA, 2012).

No entanto, existe uma dificuldade considerável na classificação de TEs em *P. psidii* MF-1, isso se deve a diversos fatores como (i) o fato de o genoma estar parcialmente sequenciado; (ii) o tamanho dos *contigs* utilizados serem muito pequenos; (iii) a base de dados de elementos repetitivos (*TE library – RepBase* do GIRI) é bastante restrita.

Duplessis *et al.* (2014) mostraram que fungos causadores de ferrugem apresentam cerca de 50% de seu genoma composto por TEs, fato que torna o sequenciamento e montagem destes genomas mais complexo, devido ao componente repetitivo e grande tamanho do genoma (DUPLESSIS *et al*., 2011; ZHENG *et al*., 2013).

Trata-se de um genoma cujo sequenciamento ainda está em andamento e portanto passível de melhoria, o tamanho dos *contigs* utilizados nas análises de anotação de elementos são pequenos e fragmentados quando comparados com as sequências de outras *Puccinia* para o mesmo fim. No caso de *P. graminis* o maior *supercontig* possui 86 *contigs* e apresenta tamanho de 3.081.398 pb enquanto o

menor *supercontig* possui 1 *contig* de 2.878 pb comparado a *P. psidii* MF-1 do presente estudo, cujo maior "*supercontig*" montado possui apenas 8.362 pb.

A restrição do banco de elementos é um fator de se considerar, uma vez que existem algumas diferenças marcantes entre o conteúdo de TEs entre as diversas espécies, no que diz respeito ao número de cópias, repertório de TEs e suas respectivas proporções de tamanho total, além da presença de cópias-fóssil que sofreram mutações (BLEYKASTEN-GROSSHANS; FRIEDRICH; SCHACHERER, 2013), fato que dificulta o sistema de classificação baseado no alinhamento de sequências similares (WICKER *et al*., 2007).

Entretanto, apesar das limitações discutidas, obteve-se sucesso na análise inédita de TEs presentes em *P. psidii*. A mineração dos 167.538 *contigs* que representam o genoma partial de *P. psidii* MF-1 resultou em 50.476 *contigs* contendo termos relacionados a transposição, ou seja, aproximadamente 30% dos *contigs* que compõem o genoma. Destaca-se o grupo 18 (*Polyprotein*) com 16% dos *contigs* com mais de 1.000 pb. Antes de realizar a anotação automática no *RepeatMasker,* já foi visto um maior predomínio de sequências relacionadas a elementos transponíveis da Classe I, detentores da região *pol* (também denominada *polyprotein*) responsável por codificar proteínas necessárias à transposição intermediada por um RNA (protease, transcriptase reversa, RNase H e integrase) (SANTANA, 2009).

A remontagem dos 50.476 *contigs* não foi capaz de resultar em TEs completos, talvez seja resultado do tamanho pequeno ou apenas reflexo de uma história evolutiva que deve ser melhor investigada. Acredita-se também que isso seja resultado da presença de híbridos de TEs no genoma de *P. psidii* MF-1*.* Uma vez que a classificação dos 586 "*supercontigs"* no programa *RepeatMasker* revelou uma sobreposição e interrupção de fragmentos pertencentes a elementos transponíveis dentre os *contigs* que compõem a sequência utilizada para realizar a comparação com o banco de TEs do GIRI.

Destaca-se que a presença de TEs híbridos é um assunto bastante discutido na literatura para demais espécies de fungos, como em *S. cerevisiae*, em que Bleykasten-Grosshans; Friedrich e Schacherer (2013) analisaram o conteúdo de elementos retrotrasponíveis LTR, da família Ty (1-5) e verificaram a presenta de híbridos Ty1/2, compostos por terminações TYB maiores que 300 pb e também por

híbridos curtos de Ty2 contendo 60 pb, ambos faz com que estes elementos não sejam anotados como da família Ty2.

Fellers *et al*. (2013) ao estudar o genoma de *P. triticina* (Pt) observaram a classificação de elementos repetitivos pelo RepBase contra a base de dados de demais *Puccinias* (*P. graminis* – Pgr; *P. triticina* – Pt e *P. striiformis* – Pst) em elementos completos e incompletos de TIRs, LTRs (Copia e Gypsy), Mariner, Mutator, Harbinger, Helitron, hAT e DNA transposons. Alguns elementos da família Copia foram encontrados inseridos dentro de elementos da família Gypgy (Pt1F16 e PtHSP02).

Quando comparam-se as estratégias de classificação utilizadas para entender a composição de elementos de transposição no genoma parcialmente sequenciado de *P. psidii* MF-1, seja por análise de *contigs* minerados pela anotação por Blast2GO e remontados em "*supercontigs*" (item 2.3.1) ou de todos os 167.538 *contigs* no programa *RepeatMasker* (item 2.3.2) observa-se o predomínio de elementos da Classe I, principalmente LTR - Retrotransposons.

O programa *RepeatMasker* utilizado na classificação de elementos transponíveis em *P. psidii* realiza a identificação de TEs por meio da comparação das sequências genômicas obtidas com as sequências descritas e depositadas em um banco de TEs, que pode ser do próprio usuário ou público disponibilizado pelo GIRI (*Genetic Information Research Institute*) (*TE library* – RepBase: http://www.girinst.org/Rpbase-Update.html) (JURKA *et al*., 2005).

Pode-se dizer que a primeira estratégia é um método mais refinado, pois as sequências montadas em 586 "*supercontigs*" foram selecionadas como "relacionadas à transposição". Observa-se que 99% de 574.688 pb minerados e anotados correspondem à sequências relacionadas a eventos de transposição classificados como LTR Retrotransposons. Entretanto essa estratégia pode gerar erros como a montagem equivocada das sequências, formando quimeras/ híbridos de TEs.

A segunda estratégia é mais ampla, pois pode incluir sequências relacionadas a TEs, sem ter sido anotadas em GO termos. Observou-se que de 32.882.244 pb anotados, 10% apresentam alguma relação com elementos transponíveis e 989.760 pb representam sequências repetitivas desconhecidas. Foram anotados 21.279 *contigs* como elementos de transposição, destes 0,02% foram classificados como LTR Retrotransposons.

A anotação automática de TEs de *P. psidii* MF-1 baseia-se no alinhamento de sequências similares aquelas depositadas no banco *TE library*. Dessa forma, elementos exclusivos da espécie de estudo, elementos não conservados e até mesmo elementos desconhecidos até o presente não são “blastados” e portanto eliminados da classificação, subestimando a composição real de TEs no genoma de *P. psidii* MF-1.

Assim, nota-se a importância da adoção de mais de uma estratégia de análise. A segunda estratégia foi o método utilizado no presente trabalho para prosseguir às análises, tanto de integridade dos elementos, comparação com outras *Puccinia* bem como da enzima transcriptase reversa.

Cantu *et al*. (2011) também acessaram o genoma do fungo *Puccinia striiformis* f. sp. *tritici* (PTS), causador de ferrugem amarela (*wheat stripe rust*) por sequenciamento de nova geração Illumina e verificaram que elementos repetitivos do tipo LTR Retrotransposons foram os mais abundantes dentro do genoma, correspondendo a 7,2% do genoma estimado em 64,8 Mb, dentro desta superfamília o tipo Gypsy mostrou-se mais frequente, com 52% de frequência dentro desta categoria.

Resultados semelhantes foram observados para *P. psidii* MF-1 no presente estudo onde os LTR Retrotransposons se mostraram mais abundantes. Observou-se uma leve diferença entre os transposons de DNA e LTR retrotransposons, os primeiros correspondem a 8,2% do genoma estimado de *Puccinia striiformis* f. sp. *tritici*, sendo os mais abundantes, do tipo hAT com 13,6% dentro desta classe. Totalizando 17,8% de elementos transponíveis classificados no genoma do fungo (CANTU *et al*., 2011). No caso de *P. psidii*, os elementos hAT foram um dos tipos mais abundantes de DNA transposon, apresentando 3 *hits* contendo 189 nucleotídeos.

Similarmente, Santana *et al*. (2012) estudaram o genoma do fungo ascomiceto *M. fijiensis*, causador de Sigatoka-negra em banana e verificaram que o genoma deste fungo é composto em sua maioria por retrotransposons, sendo 11,5% correspondente a LTR Retrotransposons e somente 0,17% de transposons de DNA foram observados, dos quais os mais abundantes foram do tipo Tc1-Mariner, com 59 cópias degeneradas correspondendo a 0,13% do genoma. Ao todo foram encontrados 11,7% de elementos transponíveis em todo o genoma de *M. fijiensis*.

A anotação de TEs de *P. psidii* s.l. (*myrtle rust*), estimado em 103 – 145 Mpb pelos autores Tan *et al.* (2014) resultou em 27% do genoma composto por elementos

de transposição, sendo que 21,2% correspondem a elementos da classe I – LTR Retrotransposons, corroborando com os dados obtidos até o momento no presente estudo.

Em *P. striiformis* f. sp. *tritici* (PTS) foi observado que próximo de 50% do genoma é composto por elementos transponíveis (CANTU *et al*., 2011), acredita-se que em *P. psidii* MF-1 os elementos transponíveis possam compor 50% ou até mais do seu genoma, exercendo um importante papel na evolução deste organismo e até na quebra de barreiras de resistência em potenciais hospedeiros.

O fungo *Hemileia vastatrix* (Hva) causador de ferrugem em cafeeiro (*Coffea arabica*) estudado por Cristancho *et al.* (2014) apresentou 70% do genoma composto por elementos repetitivos, sendo que a maior fração (38%) corresponde a LTR retrotransposons, 7,2% são DNA transposons e 27% são elementos desconhecidos. A base de dados utilizadas para a anotação foi o FNGREP.ref que inclui 1.726 elementos transponíveis identificados em Fungi, a mesma base de dados utilizada na anotação automática do genoma parcial de *P. psidii* MF-1 no presente trabalho.

Sun *et al.* (2003) discutem que genomas cuja incidência de frações repetitivas é alta, como o fungo *H. vastatrix,* são complexos de serem montados e analisados, o que dificulta a obtenção de um genoma completamente sequenciado. A expansão do genoma resultante da replicação de TEs têm sido descritas em fungos filamentosos e pode ser devido a replicação de genes relacionados à virulência do patógeno (KEMEN; JONES, 2012).

Huang *et al*. (2014) verificaram que a ocorrência de alta diversidade de genes *Avr* (virulência) em *M. grisea*, fungo causador de ferrugem em arroz (*Oriza sativa*), está diretamente relacionada a presença e atividade de sequências repetitivas.

Além disso, fungos basidiomicetos não patogênicos como *Laccaria bicolor* (MARTIN *et al*., 2008) e *Coprinopsis cinerea* (STAJICH *et al*., 2010) apresentam uma proporção reduzida de elementos repetitivos. Mais uma vez ressalta-se que a expansão de genomas de fungos causadores de ferrugem pode ser consequência de eventos de transposição, estes aumentam a diversidade de genes *Avr* o que influencia diretamente na susceptibilidade e gama de hospedeiros afetados em escala mundial.

Até o presente momento, observa-se que a distribuição de elementos repetitivos apresentada por *P. psidii* MF-1 está bastante coerente com a literatura analisada de TEs em basidiomicetos. Existe uma constante presença de LTR Retrotransposons e que as famílias mais frequentes de transposons de DNA são hAT

e Mariner. Os elementos da Classe I – LTR Retrotransposons são os mais comuns em Fungi (MUSZEWSKA *et al.*, 2011), fato que é encontrado em diversos fungos causadores de ferrugem (CANTU *et al*., 2011; DUPLESSIS *et al*., 2011; TAN *et al.*, 2014).

Um elemento de transposição é dito completo quando apresenta regiões codificadoras de proteínas relacionadas a maquinaria da transposição, terminais repetitivos conservados e TSD (*Target Site Duplication*) ou TSR (*Target Site Repeat*) (SANTANA, 2009).

O conteúdo de sequências repetitivas de genomas que são sequenciados por plataformas de NGS, como Illumina, tendem a subestimar a abundância de elementos repetitivos devido à dificuldade de montagem dos *reads* que são pequenos e repetitivos com regiões longas por meio desta tecnologia (SANTANA *et al.*, 2014). Acredita-se que este tipo de tecnologia tende a eliminar grande parte dos TEs durante a montagem dos *contigs* (SANTANA *et al.*, 2014), fato semelhante pode ter ocorrido com o genoma de *P. psidii* MF-1 em estudo.

Essa hipótese foi investigada pela busca de regiões terminais de duas famílias de elementos de transposição, LTR Retrotransposons e Tc1-Mariner, ambos amplamente estudados para utilização como marcador molecular em fungos, devido a sua abundância nos genomas eucariotos.

Nenhum dos 20 *contigs* avaliados apresentaram as terminações esperadas, talvez isso seja fruto de seus tamanhos, em média 2.937 pb, Santana *et al*. (2012) avaliaram 5.000 pb a jusante e a montante da região do genoma de interesse (*M. fijiensis*) encontrando as regiões terminais dos LTR retrotransposons, sugerindo assim, que a presente a análise feita com *contigs* de tamanho pequeno não é viável para a avaliação da integridade dos elementos de transposição da Classe I – LTR Retrotransposons.

Devido à importância de elementos Tc1-Mariner e seu potencial de utilização para estudos de diversidade populacional em Fungi (BRAGA *et al*., 2014). Foi verificada a integridade de 10 *contigs* que apresentaram *hits* relacionados a superfamília Tc1-Mariner pelo *RepeatMasker* em *P. psdiii*. Os 10 *contigs* avaliados apresentavam em média 3.404 pb e o menor *contig* possui 2.378 pb, a análise foi realizada somente *in silico.*

Os elementos desta superfamília apresentam geralmente um tamanho em torno de 1.300 a 2.400 pb e são constituídos por estruturas simples que apresentam

uma ORF (*Open Reading Frame*) responsável por codificar transposase e são flanqueadas por repetições invertidas terminais (TIRs – *Terminal Inverted Repeats*) (PLASTERK *et al*., 1999), apesar dessas características, nenhum se apresentou íntegro no genoma de *P. psidii* MF-1.

Cristancho *et al.* (2014) anotaram o genoma híbrido de *H. vastratrix,* constituído pela montagem de oito isolados da espécie, com relação à presença de LTRs e TIRs bem como a similaridade com outras sequências de Pucciniales. Foram descobertas quatro novas famílias de TEs no genoma e a anotação destas novas famílias foi realizada manualmente.

A realização da etapa de anotação manual de *contigs* classificados como elementos da Classe I – LTR Retrotransposons apresentou dois objetivos: *i)* verificar se a organização de ORFs corresponde à classificação atribuída aos elementos e presença de sequências conservadas; *ii)* comparação entre sequências de transcriptase reversa com demais fungos fitopatogênicos depositados no GenBank.

A análise foi realizada com 25 *contigs* da família Copia e Gypsy a fim de buscar uma similaridade entre as proteínas para confirmar a presença de transcriptase reversa no genoma parcialmente sequenciado de *P. psidii* MF-1*,* corroborando com aumento da confiança dos dados gerados pela anotação automática no *RepeatMasker.*

Foi observado que entre os 25 *contigs* anotados manualmente quanto às sequências proteicas presentes, 44% (*contigs* 00081, 00083, 00084, 00088, 00090, 00092, 00094, 00095, 00096, 00107 e 00110) apresentaram todos as regiões da ORF *pol* (integrase, RNase H e transcriptase reversa) (Tabela 2.9), sendo esta análise de importância imprescindível para o esclarecimento e melhor classificação de elementos de transposição em *P. psidii* MF-1.

Como observado pelos resultados obtidos, a classificação de TEs em *P. psidii* MF-1 apresenta uma dificuldade considerável. Tendo em vista a disponibilidade de informações no Banco de Dados da *Broad Institute – Puccinia Group* a respeito do genoma de outras *Puccinia* spp. causadoras de ferrugem em outras espécies vegetais. Foi realizada uma busca pelo genoma sequenciado e depositado de três espécie *P. graminis, P. striiformis* e *P. triticina,* para uma análise comparativa entre os conteúdos de TEs.

A média do tamanho dos genomas de *P. graminis*, *P. triticina* e *P. striiformis* depositados no *Puccinia Group* é de 100 Mb. Enquanto o sequenciamento e

montagem de *contigs* que formam o genoma parcial de *P. psidii* MF-1 apresenta em torno de 473 Mb (dado não publicado).

Tavares *et al.* (2014) apresentam a espécie *P. chrysanthemi* com 806,5 Mpb como o segundo maior genoma entre 32 organismos de Pucciniales avaliados. A variação entre o menor (*P. graminis* e *P. triticina* com média de 75 Mb por citometria de fluxo) e maior genoma estudados entre onze espécies do gênero Puccinia é de mais de dez vezes.

Os fungos causadores de ferrugem são conhecidos pela sua alta especialização em hospedeiros, consequência de uma história evolutiva comum com suas respectivas plantas hospedeiras (DUPLESSIS *et al*., 2011). No entanto, não é o que acontece com *P. psidii*, capaz de infectar uma série de hospedeiros da família Myrtaceae, apresentando uma gama de hospedeiros nada restrita.

Tavares *et al.* (2014) não encontraram uma característica comum que relacionasse os genomas de maior tamanho influenciada pelo tipo de ciclo sexual, dessa forma o papel dele continua incerto. Uma vez que os três maiores genomas de fungos descritos até o momento *H. vastatrix, Phakospora pachyrhizi* e *P. chrysanthemi* dependem unicamente de ciclo assexuado. Assim, a atividade de TEs representa uma fonte potencialmente importante para a geração de diversidade, mesmo para espécies que possuam ciclo sexuado (TAVARES *et al*., 2014).

A comparação entre o conteúdo de TEs entre *P. psidii* e outras três *Puccinia* spp. mostra-se importante para melhorar a compreensão da possível influência desta classe de DNA para a alta capacidade de *P. psidii* causar ferrugem em diferentes hospedeiros, visto a relação entre presença de TEs, expansão de genoma e genes relacionados a virulência (HUANG *et al.*, 2014).

Observou-se que a concentração de GC (%) presente no genoma de *P. psidii* é de 33,7% (dado não publicado), valor considerado muito inferior próximo ao encontrado em *P. graminis* (43,3%), *P. striiformis* (44,4%) e *P. triticina* (46%) (*Broad Institute*). Cristancho *et al.* (2014) ao estudarem o genoma de *H. vastatrix* observaram que a concentração de GC foi igualmente 33%, valor inferior aos demais genomas analisados, como *M. larici-populina* com 41% e *P. graminis* f. sp. *tritici* com 43,3%. Os autores discutiram esse fato como consequência do colapso de sequências repetitivas formadas por “GC” reduzindo então a representação deste dinucleotídeo no genoma. Pode-se inferir que a baixa quantidade de GC em *P. psidii* MF-1 frente a um possível

genoma de grande tamanho se deve a erro de montagem, o que não reflete a situação real do genoma.

A quantidade de elementos mascarados apresentados por *P. psidii* está mais próximo daquele apresentado por *P. striiformis*,10% e 11%, respectivamente. O número de elementos estimado para *P. psidii* MF-1 (66 elementos) é 33 vezes inferior ao observado em *P. striiformis* (2.002 elementos), acredita-se na possível eliminação de sequências durante a montagem e também na existência de famílias novas no genoma de *P. psidii* MF-1.

A frequência de LTR Retrotransposons em *P. psidii* MF-1, *P. graminis, P. triticina* e *P. striiformis* foi de 0,02%, 0,15%, 021% e 0,48%, respectivamente e nenhum retrotransposon não LTR foi detectado para os quatro genomas. Aparentemente os TEs de *P. psidii* diferem das demais espécies analisadas pois, observa-se o predomínio de elementos da Classe II – DNA Transposons, que em outras espécies de *Puccinia,* compõe em média 2% de todo o genoma. Finalmente a fração de elementos desconhecidos representa a maior parte, sendo de 8,6% para *P. graminis;* 8,15% para *P. triticina;* 5,5% para *P. striiformis* e 3% para *P. psidii*.

Acredita-se na existência de especificidade e características intrínsecas de TEs em diversos organismos eucariotos, sendo necessários estudos que demonstrem e identifiquem a presença de novas famílias de elementos de transposição, algumas potencialmente relacionadas ao controle de patogenicidade de fungos causadores de sérias doenças, como é o caso da *P. psidii*, que está sendo investigada ineditamente em detalhes no presente estudo.

Entretanto destaca-se que a classificação de algumas famílias pode ser mais precisa do que outras, pois existe uma diferença considerável entre o número de bases mascaradas e a soma de bases anotadas no arquivo, por exemplo, fragmentos menores do que 10 pb não são anotados, mas são mascarados (JURKA *et al*., 2005). Dessa forma, é comum a ocorrência de uma leve superestimação do número de cópias que contêm TEs, pois o arquivo gerado (.tbl) resume o número de repetições identificadas e agrupadas em superfamílias e a soma dos nucleotídeos que pertencem a cada superfamília (JURKA *et al*., 2005).

A análise da integridade de 79 elementos Copia de *Puccinia* spp. permite inferir que os TEs da Classe I foram recentemente inseridos ou reativados em seu respectivo genoma hospedeiro, isso se deve ao fato da alta similaridade (acima de 97%) entre as LTRs 5’ e 3’ de cada elemento ser um indicativo de inserção recente, visto que no

momento da inserção as regiões LTRs são idênticas e com o acúmulo de mutações elas vão se diferenciando. Outro fator importante que indica a atividade recente deste elementos, é de que os sítios de inserção (TSR) encontrados foram bastante variáveis (Tabela 2.12), sugerindo atividade recente.

A maioria dos elementos de transposição duplicam seu sítio alvo no momento da inserção (CHALVET *et al.,* 2003; RAY *et al*., 2007), e de acordo com Neafsey *et al*. (2004) a presença de vários sítios de inserção indica fortemente a atividade recente de transposição, pois espera-se que por seleção purificadora e deriva genética que um ou poucos sítios de inserção sejam fixados para diminuir os efeitos das inserções causadas por eventos de transposição.

Os elementos da Classe I – LTR Retrotranposons podem ser constituídos de poucas centenas de pares de bases até milhares de pares de bases, mantendo sempre o padrão de início 5’ – GT – 3’ e término 5’ – CA – 3’ (DU *et al*., 2010). Esses elementos presentes nos genomas de *Puccinia* spp. analisadas possuem de 1.072 pb e no máximo 18.688 pb, tamanho coerente ao descrito na literatura.

Todas as regiões LTRs estudadas apresentaram o padrão característico de início 5’ – GT – 3’ e término 5’ – CA – 3’, sugerindo assim, que os elementos de *Puccinia* spp. classificados como Copia pelo *RepeatMasker* são dados de alta confiabilidade e pelo menos 20% destas sequências foram recentemente inseridas devido à presença de 5’ LTR e 3’ LTR idênticas (similaridade de 100%) e TSR duplicados com alta variabilidade.

A transcriptase reversa foi primeiramente descrita em 1970 por Baltimore como uma enzima que catalisa a replicação de DNA em retrovírus (XIONG; EICKBUSH, 1990), desde então, muitos elementos genéticos de diversos organismos foram descritos apresentando sequências cujas ORFs codificavam proteínas extremamente similares à transcriptase reversa de retrovírus (ROGERS, 1985 apud XIONG; EICKBUSH, 1990).

Xiong e Eickbush, (1990) ao analisarem os aminoácidos referentes à transcriptase reversa de diversos organismos como animais, plantas, protozoários e bactérias sugeriram uma provável origem comum para as diversas formas dessa proteína nas mais diferentes formas de vida.

Na análise realizada no presente trabalho, a região da transcriptase reversa é praticamente a única comum a todos os TEs da Classe I e portanto pode ser utilizada para a análise genética e então compreensão destes (XIONG; EICKBUSH, 1990).

Apesar de outras sequências proteicas como RNase H, protease e integrase também serem comuns a todos os TEs da Classe I, a similaridade destas sequências não é tão consistente para a determinação de relações filogenéticas entre a diversidade de elementos existentes (DOOLITTLE *et al*., 1989 apud XIONG; EICKBUSH, 1990).

Destaca-se ainda que os trabalhos de análise de TEs em diversos genomas de Fungi revelam a abundância de elementos da Classe I – LTR Retrotransposons (MUSZEWSKA *et al.*, 2011; CANTU *et al*., 2011; DUPLESSIS *et al*., 2011; TAN *et al.*, 2014). Dessa forma uma comparação entre os organismos utilizando a enzima transcriptase reversa foi realizada para melhor classificação do elemento.

Observou-se um alinhamento significativo entre nove proteínas de *P. psidii* MF-1 com a transcriptase reversa de *T. matsutake,* um fungo basidiomiceto no clado denominado Grupo A que inclui a proteína de *P. triticina*.

Já as demais três sequências de *P. psidii* MF-1 agrupam com a transcriptase reversa de *P. graminis* em Grupo B2 e todas as demais proteínas de fungos formam um clado nitidamente distinto, o Grupo B1. Assim, pode-se inferir que as doze proteínas transcriptase reversa de *P. psidii* MF-1 analisadas pertencem a um elemento de transposição da família Copia, pois as sequências se agruparam a sequências de transcriptase reversa provenientes de elementos Copia de ambas *Puccinia* spp.

Apesar de ser conservada, a transcriptase reversa deve apresentar padrões de distinções dentro dos seus sete domínios conservados (XIONG; EICKBUSH, 1990), pois não foi observada a formação de um único clado monofilético composto por proteínas da *P. psidii* MF-1, mas elas se agrupavam relativamente entre elas, mostrando assim a diversidade existente entre os elementos presentes no genoma de um único organismo.

Referências

ALFENAS, A.C.; ZAUZA, E.A.V.; MAFIA, R.G.; ASSIS, T.F. (Ed.). **Clonagem e doenças do eucalipto**. 2. ed. Viçosa: Universidade Federal de Viçosa, 2004. 500 p.

ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, London, GB, v. 215, n. 3, p. 403-410, Oct. 1990.

ASHBURNER, M.; BALL, C.A.; BLAKE, J.A.; BOTSTEIN, D.; BUTLER, H.; CHERRY, J.M.; DAVIS, A.P.; DOLINSKI, K.; DWIGHT, S.S; EPPIG, J.T.; HARRIS, M.A.; HILL, D.P.; ISSEL-TARVER, L.; KASARSKIS, A.; LEWIS, S.; MATESE, J.C.; RICHARDSON, J.E.; RINGWALD, M.; RUBIN, G.M.; SHERLOCK, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. **Nature Genetics**, New York, US, v. 25, n. 1, p. 25-29, May. 2000.

BLEYKASTEN-GROSSHANS, C.; FRIEDRICH, A.; SCHACHERER, J. Genome-wide analysis of intraspecific transposon diversity in yeast. **BMC Genomics**, London, GB, v. 14, n. 399, Jun. 2013. Disponível em: <DOI:10.1186/1471-2164-14-399.>. Acesso em: 05 jul. 2015.

BRAGA, R.M.; SANTANA, M.F.; COSTA, R.V.; BROMMONSCHENKEL, S.H.; ARAÚJO, E.F.; QUEIROZ, M.V. Transposable elements belonging to the Tc1-Mariner superfamily are heavily mutated in *Colletotrichum graminicola*. **Mycologia**, New York, US, v. 106, n. 4, Jun. 2014. Disponível em: <DOI: 10.3852/13–262.>. Acesso em: 05 jul. 2015.

CANTU, D.; GOVINDARAJULU, M.; KOZIK, A.; WANG, M.; CHEN, X.; KOJIMA, K.K.; JURKA, J.; MICHELMORE, R.W.; DUBCOVSKY, J. Next Generation Sequencing Provides Rapid Access to the Genome of *Puccinia striiformis* f.sp. *tritici*, the Causal Agent of Wheat Stripe Rust. **PLos ONE**, San Francisco, US, v.6, n.8, Aug. 2011. Disponível em: <DOI: 10.1371/journal.pone.0024230.t001.> Acesso em: 05 jul. 2015.

CONSERVED DOMAIN DATABASE. Disponível em: < http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>. Acesso em: 05 jul. 2015.

CHALVET, F.; GRIMALDI, C.; KAPER, F.; LANGIN, T.; DABOUSSI, M.J. Hop, an active Mutator-like element in the genome of the fungus *Fusarium oxysporum*. **Molecular Biology and Evolution**, Chicago, US, v. 20, n. 8, p. 1362-1375, May. 2003.

CHEN, N. Using RepeatMasker to identify repetitive elements in genomic sequences. **Current Protocols in Bioinformatics**, Cambridge, UK, chap. 4, unit. 4.10, May, 2004. Disponível em: <DOI:10.1002/0471250953.bi0410s05.>. Acesso em: 05 jul. 2015.

CONESA, A.; GÖTZ, S.; GARCIA-GOMEZ, J.M.; TEROL, J.; TÁLON, M.; ROBLES, M. Blast2GO: A Universal Tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, Oxford, GB, v. 21, n. 18, p. 3674-3676, Sep. 2005.

COUTINHO, T.; WINGFIELD, M.J.; ALFENAS, A.C.; CROUS, P.W. Eucalyptus rust: a disease with the potential for serious international implications. **Plant disease**, Saint Paul, US, v. 82, n. 7, p. 819-825, Jul. 1998.

CRISTANCHO, M.A.; BOTERO-ROZO, D.O.; GIRALDO, W. TABIMA, J.; RIAÑO-PACHÓN, D.M.; ESCOBAR, C.; ROZO, Y.; RIVERA, L.F.; DURÁN, A.; RESTREPO, S.; EILAM, T.; ANIKSTER, Y.; GAITÁN, A.L. Annotation of a hybrid partial genome of the coffee rust (*Hemileia vastatrix*) contributes to the gene repertoire catalog of the Pucciniales. **Frontiers in Plant Science**, Lausanne, v. 31, n. 5, Oct. 2014. Disponível em: <DOI: 10.3389/fpls.2014.00594.>. Acesso em: 05 jul. 2015.

DABOUSSI, M.J. Fungal transposable elements and genome evolution. **Genetica**, The Hague, NL, v. 100, p. 253 – 260, 1997.

DU, J.; TIAN, Z.; HANS, C.S.; LATEN, H.M.; CANNON, S.B.; JACKSON, S.A.; SHOEMAKER, R.C.; MA, J. Evolutionary conservation, diversity and specificity of LTR-retrotransposons in flowering plants: insights from genome-wide analysis and multi-specific comparison. **The Plant Journal**, Oxford, GB, v. 63, n. 4, p. 584–598, Aug. 2010.

DUPLESSIS, S.; BAKKEREN, G.; HAMELIN, R. Advancing knowledge on biology of rust fungi through genomics. **Advances in Botanical Research**, London, GB, v. 70, p. 173 – 209, 2014. Disponível em: <DOI:10.1016/B978-0-12-397940-7.00006-9>. Acesso em 05 jul. 2015.

DUPLESSIS, S.; CUOMO, C.A.; LIN, Y.C.; AERTS, A.; TISSERANT, E.; VENEAULT-FOURREY, C.; JOLY, D.L.; HACQUARD, S.; AMSELEM, J.; CANTAREL, B.L.; CHIU, R.; COUTINHO, P.M.; FEAU, N.; FIELD, M.; FREY, P.; GELHAYE, E.; GOLDBERG, J.; GRABHERR, M.G.; KODIRA, C.D.; KOHLER, A.; KÜES, U.; LINDQUIST, E.A.; LUCAS, S.M.; MAGO, R.; MAUCELI, E.; MORIN, E. MURAT, C.; PANGILINAN, J.L.; PARK, R.; PEARSON, M.; QUESNEVILLE, H.; ROUHIER, N.; SAKTHIKUMAR, S.; SALAMOV, A.A.; SCHMUTZ, J.; SELLES, B.; SHAPIRO, H.; TANGUAY, P.; TUSKAN, G.A.; HENRISSAT, B.; VAN DE PEER, Y.; ROUZÉ, P.; ELLIS, J.G.; DODDS, P.N.; SCHEIN, J.E.; ZHONG, S.; HAMELIN, R.C.; GRIGORIEV, I.V.; SZABO, L.J.; MARTIN, F. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. **Proceedings of the National Academy of Sciences USA**, Washington, US, v.108, n. 22, p. 9166-9171, May. 2011.

FELLERS, J.P.; SOLTANI, B.M.; BRUCE, M.; LINNING,R.; CUOMO, C.A.; SZABO, L.J.; BAKKEREN, G. Conserved loci of leaf and stem rust fungi of wheat share synteny interrupted by lineage-specific influx of repeat elements. **BMC Genomics**, London, GB, v. 14, n. 60, Jan. 2013. Disponível em: <DOI: 10.1186/1471-2164-14-60.>. Acesso em: 05 jul. 2015.

GAMBIN, T.; STARTEK, M.; WALCZAK, K.; PASZEK, J.; GRZEBELUS, D.; GAMBIN, A. TIRFinder: A Web Tool For Mining Class II Transposons Carrying Terminal Inverted Repeats. **Evolutionary Bioinformatics**, Boston, US, v. 9, p. 17-27, Sep. 2013. Disponível em: < DOI: 10.4137/EBO.S10619>. Acesso em: 05 jul. 2015.

GLEN, M.; ALFENAS, A.C.; ZAUZA, A.V.; WINGFIELD, M.J.; MOHAMMED, C. *Puccinia psidii*: a threat to the Australian environment and economy a review. **Australasian Plant Pathology**, Clayton, AU, v. 36, n. 1, p. 1-16, Jan. 2007.

HUANG, J.; SI, W.; DENG, Q.; LI, P.; YANG, S. Rapid evolution of avirulence genes in rice blast fungus *Magnaporthe oryzae*. **BMC Genetics**, London, GB, v. 15, n. 45, p. 2156-15-45, Apr. 2014.

INDÚSTRIA BRASILEIRA DE ÁRVORES. São Paulo, 2015. 64 p.

JURKA, J.; KAPITONOV, V. V.; PAVLICEK, A.; KLONOWSKI, P.; KOHANY, O.; WALICHIEWICZ, J. Repbase Update, a database of eukaryotic repetitive elements. **Cytogenetic and Genome Research**, New York, US, v. 110, p. 462–467, 2005.

KEMEN, E.; JONES, J.D. Obligate biotroph parasitism: can we link genomes to lifestyles? **Trends in Plant Science**, Oxford, GB, v. 17, n. 8, p. 448–457, May. 2012.

KOHANY, O.; GENTLE, A.J.; HANKUS, L; JURKA, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmiter and Censor. **BMC Bioinformatics**, London, GB, v. 7, n. 474, Oct, 2006. Disponível em: <DOI: 10.1186/1471-2105-7-474.>. Acesso em: 05 jul. 2015.

LEITE, T.F. **Análise histológica da interação planta patógeno entre *Eucalyptus grandis* e *Puccinia psiddi* Winter e identificação de genes diferencialmente expressos em plantas resistentes e susceptíveis inoculadas com o patógeno por meio da técnica de RNA-Seq.** 2012, 177 p. Tese (Doutorado em Ciências) Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.

LEONARD, K.J.; SZABO, L.J. Stem rust of small grains and grasses caused by *Puccinia graminis*. **Molecular Plant Pathology**, Oxford, GB, v. 6, n. 2, p. 99–111, Mar. 2005.

MARGULIES, M.; EGHOLM, M.; ALTMAN, W.E.; ATTIYA, S.; BADER, J.S.; BEMBEN, L.A.; BERKA, J.; BRAVERMAN, M.S.; CHEN, Y.J.; CHEN, Z.; DEWELL, S.B.; DU, L.; FIERRO, J.M.; GOMES, X.V.; GODWIN, B.C.; HE, W.; HELGESEN, S.; HO, C.H.; IRZYK, G.P.; JANDO, S.C.; ALENQUER, M.L.; JARVIE, T.P.; JIRAGE, K.B.; KIM, J.B.; KNIGHT, J.R.; LANZA, J.R.; LEAMON, J.H.; LEFKOWITZ, S.M.; LEI, M.; LI, J; LOHMAN, K.L.; LU, H.; MAKHIJANI, V.B.; McDADE, K.E.; McKENNA, M.P.; MYER.S, E.W.; NICKERSON, E.; NOBILE, J.R.; PLANT, R.; PUC, B.P.; RONAN, M.T.; ROTH, G.T.; SARKIS, G.J.; SIMONS, J.F.; SIMPSON, J.W.; SRINIVASAN, M.; TARTARO, K.R.; TOMASZ, A.; VOGT, K.A.; VOLKMER, G.A.; WANG, S.H.; WANG, Y.; WEINER, M.P.; YU, P.; BEGLEY, R.F.; ROTHBERG, J.M. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, London, GB, v. 437, n. 7057, p. 376-380, Sep. 2005.

MARTIN, F.; AERTS, A.; AHRÉN, D.; BRUN, A.; DANCHIN, E.G.; DUCHAUSSOY, F.; GIBON, J.; KOHLER, A.; LINDQUIST, E.; PEREDA, V.; SALAMOV, A.; SHAPIRO, H.J.; WUYTS, J.; BLAUDEZ, D.; BUÉE, M.; BROKSTEIN, P.; CANBÄCK, B.; COHEN, D.; COURTY, P.E.; COUTINHO, P. M.; DELARUELLE, C.; DETTER, J.C.; DEVEAU, A.; DIFAZIO, S.; DUPLESSIS, S.; FRAISSINET-TACHET, L.; LUCIC, E.; FREY-KLETT, P.; FOURREY, C.; FEUSSNER, I.; GAY, G.; GRIMWOOD, J.; HOEGGER, P.J.; JAIN, P.; KILARU, S.; LABBÉ, J.; LIN, Y.C.; LEGUÉ, V.; LE TACON, F.; MARMEISSE, R.; MELAYAH, D.; MONTANINI, B.; MURATET, M.; NEHLS, U.; NICULITA-HIRZEL, H.; OUDOT-LE SECQ, M.P.; PETER, M.; QUESNEVILLE, H.; RAJASHEKAR, B.; REICH, M.; ROUHIER, N.; SCHMUTZ, J.; YIN, T.; CHALOT, M.; HENRISSAT, B.; KÜES, U.; LUCAS, S.; VAN DE PEER, Y.; PODILA, G.K.; POLLE, A.; PUKKILA, P.J.; RICHARDSON, P.M.; ROUZÉ, P.; SANDERS, I.R.; STAJICH, J.E.; TUNLID, A.; TUSKAN, G.; GRIGORIEV, I.V. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. **Nature**, London, GB, v. 452, n. 7183, p. 88-92, Mar. 2008.

MUSZEWSKA, A.; HOFFMAN-SOMMER, M.; GRYNBERG, M. LTR Retrotransposons in Fungi. **PLoS ONE**, San Francisco, US, v. 6, n. 12, Dec. 2011. Disponível em: <DOI: 10.1371/journal.pone.0029425.>. Acesso em 05 jul. 2015.

NEAFSEY, D.E.; BLUMENSTIE, J.P.; HARTL, D.L. Different regulatory mechanisms underlie similar transposable element profiles in pufferfish and fruitflies. **Molecular Biology and Evolution**, Chicago, US, v. 21, n. 12, p. 2310-2318, Dec. 2004.

PLASTERK, R.H.A; IZSVAK, Z.; IVICS, Z. Resident aliens - the Tc1/mariner superfamily of transposable elements. **Trends in Genetics**, Cambridge, GB, v. 15, n. 8, p. 326–332, Aug. 1999.

Puccinia Group Genome Databases, **Broad Institute of Harvard and MIT.** Disponível em: <http://www.broadinstitute.org/>. Acesso em 05 jul. 2015.

QUECINE, M.C.; BINI, A.P.; ROMAGNOLI, E.R.; ANDREOTE, F.D.; MOON, D.H.; LABATE, C.A. Genetic Variability in *Puccinia psidii* Populations as Revealed by PCR-DGGE and T-RFLP Markers. **Plant Disease**, Saint Paul, US, v. 98, n.1, p.16-23, Jan. 2014.

RAY, D.A.; PAGAN, H.J.; THOMPSON, M.L.; STEVENS, R.D. Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*. **Molecular Biology and Evolution**, Chicago, US, v. 24, n. 3, p. 632-639, Dec. 2006.

RUIZ-SILVA, J. **Isolamento e caracterização de elementos transponíveis em espécies do gênero *Brycon* (Characidae, Bryconinae)**. 2012, 81 p. Dissertação (Mestrado em Ciências Biológicas) – Universidade Estadual de São Paulo, Botucatu, 2012.

SANTANA, M.F. **Retrotransposons LTR no genoma de *Moniliophthora perniciosa* e *Cochliobolus heterostrophus*.** 2009, 81 p. Dissertação (Mestrado em Microbiologia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2009.

SANTANA, M.F.; SILVA, J.C.F.; BATISTA, A.D.; RIBEIRO, L.E.; SILVA, G.F.; ARAÚJO, E.F.; QUEIROZ, M.V. Abundance, distribution and potential impact of transposable elements in the genome of *Mycosphaerella fijiensis.* **BMC Genomics**, London, GB, v. 13, n. 720, Dec. 2012. Disponível em < DOI: 10.1186/1471-2164-13-720>. Acesso em: 30 jun.2015.

SANTANA, M.F.; SILVA, J.C.F.; MIZUBUTI, E.S.G.; ARAÚJO, E.F.; CONDON, B.J.; TURGEON, B.G.; QUEIROZ, M.V. Characterization and potential evolutionary impact of transposable elements in the genome of *Cochliobolus heterostrophus*. **BMC Genomics**, London, GB, v.15, n. 536, Jun. 2014. Disponível em: <DOI: 10.1186/1471-2164-15-536.>. Acesso em 05 jul. 2015.

STAJICH, J.E.; WILKE, S.; AHRÉN, D.; AU, C.A.; BIRREN, B.W.; BORODOVSKY,M.; BURNS, C.; CANBÄCK, B.; CASSELTON, L.A.; CHENG, C.K.; DENG, J.; DIETRICH, F.S.; FARGO, D.C.; FARMAN, M.L.; GATHMAN, A.C.; GOLDBERG, J.; GUIGÓ, R.; HOEGGER, P.J.; HOOKER, J.B.; HUGGINS, A.; JAMES, T.Y.; KAMADA, T.; KILARU, S.; KODIRA, C.; KÜES, U.; KUPFER, D.; KWAN, H.S.; LOMSADZE, A.; LI, W.; LILLY, W.W.; MA, L.J.; MACKEY, A.J.; MANNING, G.; MARTIN, F.; MURAGUCHI, H.; NATVIG, D.O.; PALMERINI, H.; RAMESH, M.A.; REHMEYER, C.J.; ROE, B.A.; SHENOY, N.; STANKE, M.; TER-HOVHANNISYAN, V.; TUNLID, A.; VELAGAPUDI, R.; VISION, T.J.; ZENG, Q.; ZOLAN, M.E.; PUKKILA, P.J. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (Coprinuscinereus). **Proceedings of the National Academy of Sciences USA**, Washington, US, v. 107, n. 26, p. 11889–11894, Jun. 2010.

SUN, X.; LE, H.D.; WAHLSTROM, J.M.; KARPEN, G.H. Sequence analysis of a functional Drosophila centromere. **Genome Research,** New York, US, v. 13, n. 2 p. 182–194, Feb. 2003.

TAN, M.K.; COLLINSA, D.; CHENB, Z.; ENGLEZOUA, A.; WILKINS, M.R. A brief overview of the size and composition of the myrtle rust genome and its taxonomic stat. **Mycology**, Abingdon, GB, v. 5, n. 2, p. 52–63, Jun. 2014.

TAVARES, S.; RAMOS, A.P.; PIRES, A.S.; AZINHEIRA, H.G.; CALDEIRINHA, P.; LINK, T.; ABRANCHES, R.; SILVA, M. do C.; VOEGELE, R.T.; LOUREIRO, J.; TALHINHAS, P. Genome size analyses of Pucciniales reveal the largest fungal genomes. **Frontiers in Plant Science**, Lausanne, CH, v. 26, n. 5, Aug. 2014. Disponível em: < DOI: 10.3389/fpls.2014.00422.>. Acesso em 05 jul. 2015.

WICKER, T; SABOT, F; HUAN-VAN, A.; BENNETZEN, J. L.; CAPY, P.; CHALHOUB, B.; FLAVELL, A.; LEROY, P.; MORGANTE, M.; PANAUD, O.; PAUX, E.; SANMIGUEL, P.; SCHULMAN, A. H. A unified classification system for eukaryotic transposable elements. **Nature**, London, GB, v. 8, n. 12 p. 973–982, Dec. 2007.

XIONG, Y.; EICKBUSH, T.H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. **The EMBO Journal**, Heidelberg, DE, v.9, n.10, p.3353-3362, Oct. 1990.

XU, Z.; WANG, H. LTR_FINDER: an efficient tool for the prediction of full-lenght LTR retrotransposons. **Nucleic Acids Research**, London, GB, v. 35, Jul. 2007. Disponível em: <DOI: 10.1093/nar/gkm286.>. Acesso em: 05 jul. 2015.

ZHENG, W.; HUANG, L.; HUANG, J.; WANG, X.; CHEN, X.; ZHAO, J.; GUO, J.; ZHUANG, H.; QIU, C.; LIU, J.; LIU, H.; HUANG, X.; PEI, G.; ZHAN, G.; TANG, C.; CHENG, Y.; LIU M.; ZHANG, J.; ZHAO, Z.; ZHANG, S.; HAN, Q.; HAN, D.; ZHANG, H.; ZHAO, J.; GAO, X.; WANG, J.; NI, P.; DONG, W.; YANG, L.; YANG, H.; XU, J.R.; ZHANG, G.; KANG, Z. High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. **Nature Communications**, New York, US, v. 4, n. 2673, Oct. 2013. Disponível em: <DOI: 10.1038/ncomms3673.>. Acesso em: 05 jul. 2015.

3 UTILIZAÇÃO DE FERRAMENTAS MOLECULARES NO ESTUDO DE ELEMENTOS TRANSPONÍVEIS EM *Puccinia psidii*

Resumo

O presente trabalho investigou os elementos transponíveis em *Puccinia psidii,* agente causal da ferrugem em eucalipto, por meio de ferramentas moleculares, visando um melhor conhecimento desses elementos, complementando as análises realizadas *in silico*. Uma maneira interessante de acessar a diversidade genética de organismos cujo genoma apresenta uma fração considerável composta por elementos transponíveis (TE) é utilizar o polimorfismo existente entre os elementos de transposição como marcador molecular. *Primers* específicos para a técnica IRAP (*Inter-Retrotransposon Amplified Polymorphism*) foram utilizados em duas combinações: CIIRAP1 e CIIRAP4; CIIRAP2 e CIIRAP4. A variabilidade genética dentre populações de *P.psidii* foi analisada utilizando populações provenientes de três híbridos de *Eucalyptus* (PEGU, PEGD e PEGB) e de três hospedeiros Myrtaceae (goiabeira, jambeiro e jabuticabeira). O acesso à diversidade genética de diferentes populações de *P. psidii* não se mostrou eficaz por meio de IRAP. Uma das hipóteses propostas é que a especificidade dos *primers* utilizados deve ser tão grande a ponto de não alinhar em regiões conservadas de LTRs (*Long Terminal Repeats*) em outras espécies de fungos, ou os TEs presentes no genoma de *P. psidii* não possuem LTRs conservadas, sendo denominados “TEs complexos” ou “híbridos de TEs”. A presença de “TEs complexos” ou “híbridos de TEs” são fenômenos comuns em sequências genômicas. Os TEs presentes em *P. psidii* MF-1, em análise *in silico,* apresentaram diversas sequências com sobreposição de fragmentos pertencentes a TEs. Acredita-se na presença de híbridos de TEs no genoma de *P. psidii* MF-1*.* A fim de se confirmar esta hipótese e verificar a ordem de sequências proteicas que compõem as regiões *gag* e *pol*, foram desenhados dezoito *primers* com base na sequência de seis *contigs* (00075; 00081; 00084; 00088; 00095 e 00104) de *P. psidii* MF-1, anotados com fragmentação/sobreposição de elementos. Os *primers* foram utilizados em doze combinações para amplificar fragmentos referentes aos seis *contigs* contendo elementos potencialmente híbridos de *P. psidii* MF-1. As reações de PCRs (*Polymerase Chain Reaction*) foram realizadas com a utilização da enzima “elongase” (*Accu LA Taq DNA Polymerase*) que permite a amplificação de fragmentos de até 20 Kb. A amplificação mostrou-se eficiente e reproduzível para três combinações referentes aos *contigs* 00075, 00088, 00095 (*amplicons* Pp_1.1, Pp_4.1 e Pp_5.1). O objetivo desta etapa foi obter o sequenciamento destas possíveis regiões formadas por híbridos de TEs e confirmar que não foi fruto de erros durante a montagem, no entanto, até o presente momento, o sequenciamento ainda não foi alcançado. Uma vez obtidas essas sequências, a presença de similaridade entre os *contigs* utilizados para o desenho de *primers* específicos e sequências dos *amplicons* poderá revelar a presença de híbridos de TEs no genoma de *P. psidii* MF-1. Certamente, a confirmação da hipótese levantada será de suma importância para o melhor entendimento da evolução, bem como a patologia do fungo. É provável ainda, que novos elementos transponíveis exclusivos de *P. psidii* sejam descobertos e estejam relacionados a adaptabilidade do fungo a diversas espécies de Myrtaceae.

Palavras-chave: Ferramentas moleculares; *Primers* CIIRAP; Variabilidade; Populações de *P. psidii*; Hídridos de TEs

Abstract

This study investigated the transposable elements (TEs) in *Puccinia psidii*, causal agent of eucalyptus rust, using molecular tools aiming a better knowledge of its genome and pathogenic biology, complementing *in silico* analysis. An interesting way to assess the organism genetic diversity whose genome has a considerable TEs fraction is the usage of existing polymorphism between TEs as molecular markers. Specific primers for the Inter-retrotransposon Amplified Polymorphism (IRAP) technique were used in two combinations CIIRAP 1 and CIIRAP 4; CIIRAP 2 and CIIRAP 4. Genetic variability among *P. psidii* populations from *Eucalyptus* hybrids (PEGU, PEGD and PEGB) and three Myrtaceae hosts (guava, syzygium and jabuticaba) was analyzed. The genetic diversity access was not effective through IRAP. The primers specificity may be as great as not annealing in the conserved LTRs (Long Terminal Repeats) regions in other fungal species or TEs present in *P. psidii* MF-1 genome have no conserved LTRs, than they are called “complex TEs” or “hybrid TEs”. Complex TEs or hybrid TEs are common in genomic sequences. TEs present in *P. psidii* MF-1 presented several sequences showed overlapping fragments belonging to TEs, it is believed that this occurrence indicate the presence of hybrid TEs in *P. psidii* MF-1 genome. In order to confirm this hypothesis and to verify the order of protein sequences comprising *gag* and *pol* regions, eighteen primers were designed to six TEs in *P. psidii* MF-1 (contigs00075; 00081; 00084; 00088; 00095 and 00104). The primers were used in twelve combinations to amplify fragments for the six contigs. The molecular tool used was Polymerase Chain Reactions (PCR) using an elongase enzime (Accu LA Taq DNA polymerase) that allows the amplification of fragments up to 20 Kb. Amplification was efficient and reproducible for three combinations from contigs 00075, 00088, 00095 (amplicons Pp_1.1, Pp_4.1 and Pp_5.1). The purpose of this step was to obtain the sequencing of these possible regions formed by hybrid TEs, however, to date has not obtained the sequencing. The presence of similarity between the contigs used to design specific primers and sequences of the amplicons will reveal the presence of hybrids TEs. Certainly, the confirmation of our hypothesis is very important to reveal the presence / absence of TEs hybrids, which allows a better understanding of the evolution and the pathology of the fungus. The discovery of new exclusive transposable elements in *P. psidii* could be related to adaptability of the fungus to several Mytaceae species.

Keywords: Molecular Tools; CIIRAP Primers; Variability; Populations of *P. psidii*; TE Hydrids

3.1 Introdução

Dentre as sequências de DNA repetitivo que podem ser encontradas dispersas no genoma hospedeiro, destacam-se os elementos transponíveis (TEs). A importância do estudo deste tipo de DNA repetitivo se deve à sua capacidade de se mover de um local a outro dentro do genoma (HUA-VAN *et al*., 2005), além disso, a inserção destes elementos pode provocar um amplo espectro de mutações no genoma que os comportam (KIDWELL; LISCH, 2000). Os TEs são considerados importantes na evolução e adaptabilidade de genomas de fungos fitopatogênicos, devido à influência deles na redução ou aumento da virulência (KANG *et al*., 2001).

Visto a abundância e diversidade de TEs e a velocidade com que os dados de sequenciamento são gerados, sua identificação e anotação tem sido um grande desafio (WICKER *et al*., 2007). A detecção e caracterização de TEs em fitopatógenos podem possibilitar maior conhecimento sobre a organização e a evolução do genoma desses fungos, bem como os mecanismos de silenciamento gênico, uma vez que a movimentação deles no genoma hospedeiro pode gerar rearranjos, novos alelos e influenciar a expressão gênica (MARTIN *et al*., 2008; SANTANA, 2009).

Além do importante papel no genoma de fungos fitopatogênicos, a atividade de TEs pode ainda, gerar variabilidade em fungos, que por sua vez podem ser os responsáveis pela alta plasticidade e adaptabilidade apresentada por diversos fungos (DABOUSSI; CAPY, 2003). A atividade de transposição destes elementos depende de algumas variáveis ambientais como condições de temperatura, irradiação, estresse oxidativo e ciclo sexual (SANTANA, 2009).

Analisar elementos transponíveis presentes em genomas de fungos fitopatogênicos pode contribuir para estudos evolutivos e de transferência genética, além disso, estas sequências de DNA repetitivas podem ser utilizadas como marcadores para traçar o perfil evolutivo de populações (MURATA *et al*., 2008; SANTANA, 2009).

A técnica IRAP (*Inter-Retrotransposon Amplified Polymorphism*) é amplamente utilizada devido a sua baixa complexidade (KALENDAR *et al*., 2011; SANTANA *et al*., 2013), permite a obtenção de marcas polimórficas em regiões em que os retrotransposons estão inseridos. Assim *primers* para regiões conservadas, como as LTRs (longas terminações repetitivas) são desenhados e utilizados para detectar perfis de variabilidade genética entre populações de fungos (KALENDAR *et al*., 1999).

O fungo basidiomiceto biotrófico *Puccinia psidii* Winter, pertence à ordem Pucciniales, cujas 6.000 espécies descritas são todas parasitas obrigatórias de plantas, conhecidas comumente como “ferrugem” e causam na maioria dos casos grandes prejuízos econômicos (SWERZ, 1994 apud LEITE, 2012). Destaca-se ainda que o setor florestal apresenta grande importância para o Brasil, devido à vasta área territorial ocupada para a produção de espécies madeireiras, como o *Eucalyptus* spp., bem como a geração de tributos e de empregos relacionados ao setor (IBÁ, 2015). Assim, a presença de doenças como a ferrugem causada pelo fungo *P. psidii* certamente acarreta severos danos econômicos para o setor florestal (GLEN *et al*., 2007).

Dessa maneira, os principais objetivos desta etapa foram: *i)* analisar a variabilidade genética entre diferentes populações de *P. psidii*, causadores de ferrugem em cultura de *Eucalyptus* spp. bem como em espécies frutíferas de Myrtaceae por meio da técnica IRAP; *ii)* entender e acessar a possível presença de híbridos de TEs no genoma de *P. psidii* (linhagem monopustular MF-1).

A expectativa é encontrar algum padrão específico entre os diversos hospedeiros e seu possível patógeno exclusivo e alcançar um melhor entendimento da composição de TEs no genoma deste fitopatógeno para que então, futuramente, a complexa interação planta-patógeno seja melhor entendida para o desenvolvimento de novas estratégias de controle.

3.2 Materiais e Métodos

3.2.1 IRAP na análise de variabilidade genética em populações de *P. psidii*

A técnica IRAP foi utilizada para a análise de *P. psidii* MF-1 e seis populações de *P. psidii* isoladas de diversos hospedeiros da família Myrtaceae. Os hospedeiros selecionados para esta análise, foram populações de *P. psidii* MF-1 provenientes de três híbridos de *Eucalyptus, E. grandis* x *urophylla* (PEGU); *E. grandis* x *dunni* (clone 1366 [PEGD]) e *E. grandis* x *brassiana* (clone MA2000 [PEGB]), todos coletados na área experimental da Futuragene, Itapetininga-SP. Além de *Eucalyptus* spp., populações de *P. psidii* provenientes de três espécies de Myrtaceae hospedeiras, também foram analisadas, goiabeira – *Psidium guajava* (PGuava), jabuticabeira – *Myrciaria cauliflora* (Pjab) e jambeiro – *Syzygium jambos (*PSyz) (QUECINE *et al*., 2014).

Foram utilizadas duas combinações de *primers* específicos para IRAP (SANTANA *et al*., 2013), CIIRAP1 (5’ CGTACGGAACACGCTACAGA 3’) com CIIRAP4 (5’ CTTTTGACGAGGCCATGC 3’) e CIIRAP2 (5’ AATAACGTCTCGGCCTTCAG 3’) com CIIRAP4 (5’ CTTTTGACGAGGCCATGC 3’) previamente utilizados por Santos *et al*. (2012).

As reações de PCR para os *primers* de IRAP foram conduzidas em volume final de 25µL contendo tampão 1X Taq DNA Polymerase Buffer 10X (Invitrogen), 1,3 µL de $MgCl_2$ (Invitrogen) [25 mM], 1 µL de dNTP mix (Thermo Scientific) [10 mM], 0,5 µL de cada *primer* [10 µM], 40 ng de DNA genômico e 0,3 µL de Taq DNA Polimerase [5 U/µL] (Invitrogen) e o volume final foi completado com H_2O miliquê. As reações de PCR foram realizadas em termociclador (GeneAmo PCR System 9700, Applied Biossystems) com desnaturação inicial de 2 min a 94°C, seguida de seis ciclos de 30 seg a 94°C, 30 seg a 50°C e 2 min a 72°C. Posteriormente, 24 ciclos com 30 seg adicionais a cada tempo de extensão (a 72°C) a cada seis ciclos. A extensão final é de 10 min a 72°C. O produto das reações foi analisado em gel de agarose (1% p/v) juntamente com marcador de peso molecular 1Kb (Sigma) a 3 volts.cm^{-1}. O corante-intercalante de DNA utilizado para tingir as amostras foi o SYBER Green (0,003 µL.$amostra^{-1}$) (Invitrogen). Após a eletroforese, o gel foi foto-documentado sobre luz UV.

3.2.2 Análise molecular de possíveis “TEs complexos” no genoma de *P. psidii* MF-1

3.2.2.1 Desenho de *primers* específicos

Alguns dos *contigs* que foram manualmente anotados (item 2.3.4) foram selecionados para o desenho de *primers* específicos. A Tabela 3.1 contém os seis *contigs* (00075; 00081; 00084; 00088; 00095 e 00104) que foram utilizados para o desenho de *primers* específicos para elementos Copia e Gypsy (*Contig*00081) em *P. psidii* MF-1.

O programa *OligoPerfect™ Designer* (http://goo.gl/2nrzy2) foi utilizado para realizar o desenho de *primers* e os pares que abrangiam toda a região do *contig* anotado como elemento de transposição foram selecionados. Os parâmetros utilizados foram tamanho entre 17-23 pb, temperatura de anelamento entre 57- 63°C, concentração de CG entre 30-60% e tamanho do produto entre 2.000 - 4.000 pb. Para

verificar a existência de dímeros foi utilizado o *software Oligo Analysis Tool* (http://goo.gl/37JHNg).

Tabela 3.1 – Lista de *contigs* de *P. psidii* MF-1 utilizados para o desenho de *primers* específicos para elementos Copia e Gypsy e suas respectivas sequências oligonucleotídicas

Contig	***Primer***	**Sequência 5' → 3'**
00075	Pp_Copia75 6 F	C T C C C C T C C A T A A C A C A T T A
	Pp_Copia75 6 R	G A T G C C G C A T T A T T A G A C T C
	Pp_Copia75 6 F	C T C C C C T C C A T A A C A C A T T A
	Pp_Copia75 10 R	A T C A A C C A A T T T G C C A A T A C
00081	Pp_Gypsy81 1 F	G A A A A A C G T T C T G T T C A A G G
	Pp_Gypsy81 1 R	A A T T C A C C A C T C A A T T C G A C
	Pp_Gypsy81 1 F	G A A A A A C G T T C T G T T C A A G G
	Pp_Gypsy81 4 R	C T T C A T T C T G T T C C A C T G G T
00084	Pp_Copia 84 1 F	T G G A C T T C A T C T A C C A C C T C
	Pp_Copia 84 1 R	T G T C T G G T C C T A T G G A A G A C
	Pp_Copia 84 1 F	T G G A C T T C A T C T A C C A C C T C
	Pp_Copia 84 8 R	A C A C A T T C A A A A G G G A A G T G
00088	Pp_Copia 88 1 F	T A G T C A C C A C T C A C G A T G A A
	Pp_Copia 88 1 R	G A T C G A T T G T C A G T C C T T G T
	Pp_Copia 88 1 F	T A G T C A C C A C T C A C G A T G A A
	Pp_Copia 88 8 R	G T C T T T G A T G T C C A C G A T T T
00095	Pp_Copia95 2 F	T G G A C T G C A A G A T C A C T A T G
	Pp_Copia95 2 R	C A T C A A C G T G A A T G T A C A G C
	Pp_Copia95 3 F	A A A T C A C C G A C A C A A A A G A T
	Pp_Copia95 2 R	C A T C A A C G T G A A T G T A C A G C
000104	Pp_Copia104 4 F	T T C A C T G G T A G G A G C A A G T T
	Pp_Copia104 1 R	T A G T C A A C C A T C T C C T C C A C
	Pp_Copia104 10 F	C A A A A T T C C T T C T T C A C T G G
	Pp_Copia104 1 R	T A G T C A A C C A T C T C C T C C A C

3.2.2.2 Validação de *primers* específicos

O DNA do fungo utilizado para a validação dos *primers* específicos para Retrotransposons de *P. psidii* MF-1 foi extraído utilizando-se o *DNAeasy Plant Mini Kit* (Qiagen), seguindo-se as recomendações do fabricante. Foram realizadas 12 combinações entre os pares de *primers* desenhados (Tabela 3.1) para as seis sequências de *contigs* anotados.

Primeiramente, uma análise preliminar foi realizada com *primers* específicos para os *contigs* 00075, 00081, 00088 e 00104 e a reação foi realizada nas seguintes condições: volume final de 25 µL, contendo tampão 1X Taq DNA Polymerase Buffer 10X (Invitrogen), 1,5 µL de $MgCl_2$ (Invitrogen) [25 mM], 1 µL de dNTP mix (Thermo Scientific) [10 mM], 0,5 µL de cada *primer* [10 µM], 10 ng de DNA genômico e 0,3 µL de Taq DNA Polimerase [5 U/µL] (Invitrogen) e o volume final foi completado com H_2O miliquê. As reações de cPCR (PCR convencional) foram realizadas em termociclador (Therm–1000–2, Axygen) com desnaturação inicial de 10 min a 94°C, seguida de 35 ciclos 1 min a 94°C, 45 seg (utilizando gradiente de temperaturas 45°C e 60°C) e extensão por 4 min a 72 °C, a extensão final foi de 7 min por 72°C. O produto das reações foi analisado em gel de agarose (1,2 % p/v) juntamente com marcador de peso molecular 1Kb (Sigma) a 3 volts.cm^{-1}. O corante-intercalante de DNA utilizado para tingir as amostras foi o SYBER Green (0,003 µL.$amostra^{-1}$) (Invitrogen). Após a eletroforese, o gel foi foto-documentado sobre luz UV.

Posteriormente, a fim de avaliar a especificidade dos *primers* desenhados, as combinações referentes aos *contigs* 00075; 00081; 00084; 00088; 00095 e 00104 foram testadas. As reações de amplificação foram realizadas com uma enzima Taq DNA “elongase” em um volume final de 25 µL, contendo 1X AccuTaq LA 10X Buffer (Sigma - Aldrich), 1 µL de dNTP mix (Thermo Scientific) [10 mM], 1 µL de DMSO - Dimetil Sulfóxido (Sigma - Aldrich), 1 µL de cada *primer* [10 µM], 20 ng de DNA genômico, 0,5 de AccuTaq LA DNA Polymerase [0,05 U/µL] (Sigma - Aldrich) e o volume final foi completado com H_2O miliquê. As reações de PCR foram realizadas em termociclador (Therm–1000–2, Axygen) com desnaturação inicial de 30 seg a 98°C, seguida de 30 ciclos com desnaturação de 15 seg a 94°C, anelamento dos *primers* por 20 seg (temperaturas 59°C e 65°C) e extensão por 15 min a 68 °C, a extensão final foi de 10 min por 68°C. O produto das reações foi analisado em gel de agarose (1,2 % p/v) juntamente com marcador de peso molecular 1Kb (Sigma) a 3

volts.cm^{-1}. O corante-intercalante de DNA utilizado para tingir as amostras foi o SYBER Green (0,003 µL.amostra^{-1}) (Invitrogen). Após a eletroforese, o gel foi foto-documentado sobre luz UV.

3.2.2.3 Clonagem dos *amplicons* de Retrotransposons

3.2.2.3.1 Purificação e adição de cauda poli-adenina

Os *amplicons* Pp_1.1 (6.000 pb), Pp_4.1 (2.500 pb) e Pp_5.1 (3.000 pb), gerados a partir dos *contigs* 00075, 00088 e 00095, respectivamente, foram purificados seguindo as recomendações do fabricante pelo *Illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare). Os produtos purificados foram concentrados em aparelho Vacufuge® 5301 Concentrator (Eppendorf) para obtenção de 200 ng de DNA. Foi realizada uma etapa de adição de cauda poli-adenina aos produtos de PCR antes de proceder com a clonagem do vetor pGEM®T Easy (Promega). Seguindo o protocolo *Addition of 3' overhangs to PCR products* (Promega), assim sendo, 4 µL do *amplicons* purificados (200 ng) foram adicionados a 1 µL de Taq DNA Polymerase Buffer 10X (Invitrogen), 0,5 µL de dATP [10 mM], 0,6 de $MgCl_2$ (Invitrogen) [25 mM], 1 µL de Taq DNA Polimerase [5 U/µL] (Invitrogen) e o volume final foi completado com H_2O, a mistura foi incubada por 30 min a 72ºC em termociclador (GeneAmo PCR System 9700, Applied Biossystems).

3.2.2.3.2 Clonagem em vetor pGEM®T Easy Vector

Os produtos de PCR purificados e poli-adenilados foram submetidos à ligação com o vetor pGEM®T Easy utilizando o *pGEM®T Easy Vector System I Kit* (Promega) na razão molar de inserto (*amplicons* purificados e poli-adelinados):vetor de 3:1, segundo as recomendações do fabricante. Posteriormente, foi realizada a transformação do vetor resultante da clonagem dos insertos Pp_1.1, Pp_4.1 e Pp_5.1, originários dos *contigs* 00075, 00088 e 00095, respectivamente, em células eletro-competentes de *E. coli* DH5.

O meio Luria Bertani – LB (SAMBROOK; RUSSEL, 2001) suplementado com o antibiótico ampicilina (100 µg.mL^{-1}) e 2 µL/mL de 5-Bromo- 4chloro- 3indolyl β-D-galactopyranoside 2% (X-Gal) (Sigma) foi utilizado para a seleção de recombinantes positivos.

3.2.2.3.3 Extração de plasmídeo pGEM®T Easy clonado

A extração de plasmídeo foi conduzida a partir de culturas frescas dos recombinantes inoculados em 5 mL de meio líquido LB suplementado com antibiótico ampicilina (100 $\mu g.mL^{-1}$) e crescidos em frascos de Erlenmeyer com tamanho de quatro vezes maior ao volume do inóculo (permite maior aeração), sob agitação a 200 rpm a 37ºC por 18 horas. O precipitado obtido pela centrifugação dessas suspensões bacterianas foram utilizados na extração plasmidial por *QIAprep® Spin Miniprep Kit* (QIAgen), seguindo as recomendações do fabricante.

A integridade e concentração dos plasmídeos extraídos nesta etapa do trabalho foram verificadas em gel de agarose 1,2% a 3 $volts.cm^{-1}$ juntamente com o marcador de peso molecular λ (Fermentas) e 1Kb (Sigma) a 3 $volts.cm^{-1}$. O corante-intercalante de DNA utilizado para tingir as amostras foi o SYBER Green (0,003 $\mu L.amostra^{-1}$) (Invitrogen). Após a eletroforese, o gel foi fotodocumentado sobre luz UV.

3.3 Resultados

3.3.1 IRAP e variabilidade genética em populações de *P. psidii*

Os *primers* CIIRAP 1, CIIRAP 2 e CIIRAP 4 que foram testados preliminarmente com o DNA de *P. psidii* MF-1 resultaram na amplificação de 4 bandas similares quando utiliza-se o DNA de *P. psidii* MF-1 na concentração de 20 ng/µL e 40 ng/µL pela combinação de *primers* CIIRAP 2 e CIIRAP 4 (Figura 3.1).

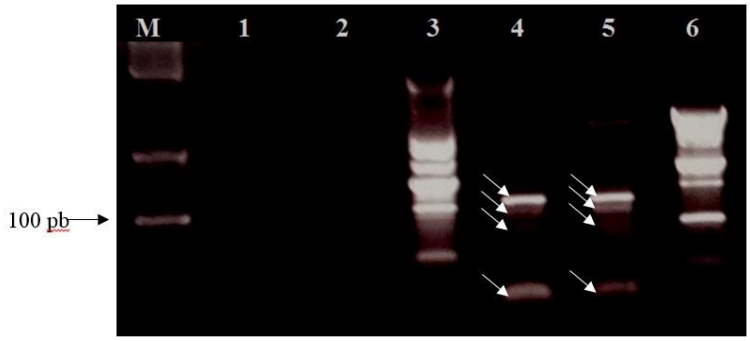


Figura 3.1 – Amplificação de DNA de *P. psidii* MF-1 em duas concentrações, 20 ng/µL e 40 ng/µL utilizando as duas combinações de *primers* para a análise de polimorfismo por IRAP. (**M**) Marcador 100 pb (Fermentas); (**1** e **2**) DNA de *P. psidii* MF-1 nas concentrações 20 ng/µL e 40 ng/µL, respectivamente, utilizando a combinação de *primers* CIIRAP 1 e CIIRAP 4; (**3**) DNA de fungo controle utilizando a combinação de *primers* CIIRAP 1 e CIIRAP 4; (**4** e **5**) DNA de *P. psidii* MF-1 nas concentrações 20 ng/µL e 40 ng/µL, respectivamente, utilizando a combinação de *primers* CIIRAP 2 e CIIRAP 4 e (**6**) DNA de fungo controle utilizando a combinação de *primers* CIIRAP 2 e CIIRAP 4. A seta branca indica o mesmo perfil de bandas entre diferentes concentrações do DNA de *P. psidii* MF-1

Observa-se que quando são utilizados 20 ng/µL e 40 ng/µL de DNA de *P. psidii* MF-1 pela combinação de *primers* CIIRAP 1 e CIIRAP 4, nenhum produto é observado (Figura 3.1). Estabelecida a concentração ótima de DNA, a PCR para amplificação das demais populações de *P. psidii* resultaram na obtenção de apenas duas bandas para a população de *P. psidii* provenienete do material PEGB quando utilizada a combinação de *primers* CIIRAP 1 e CIIRAP 4 (Figura 3.2).

Observa-se ainda que quando a combinação CIIRAP 2 e CIIRAP 4 é utilizada, ocorre a amplificação nas amostras das populações do fitopatógeno provenientes de PEGD (apenas uma única banda bem fraca), PEGB (apresentaram quatro bandas) e PGuava (também uma única banda bem fraca).

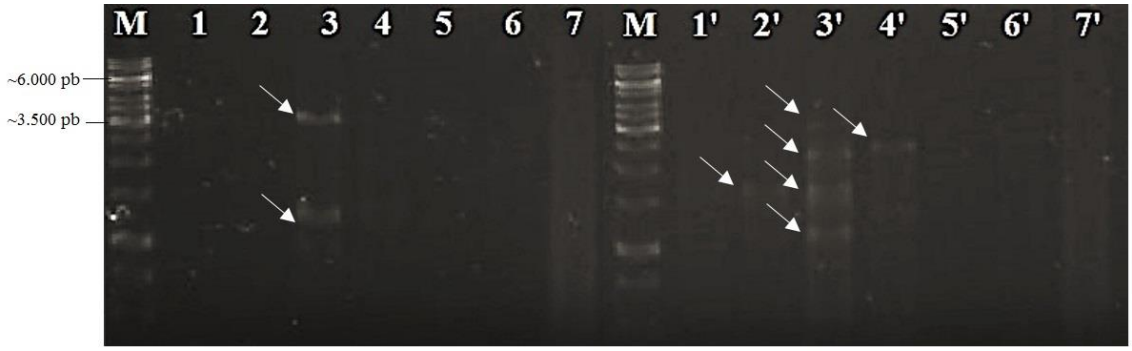


Figura 3.2 – Amplificação de DNA de *P. psidii* MF-1 e demais *P. psidii* de outros hospedeiros Myrtaceae utilizando as duas combinações de *primers* para a análise de polimorfismo por IRAP. (**M**) Marcador 1Kb (Fermentas); (**1** a **7**) amplificação utilizando a combinação CIIRAP 1 e CIIRAP 4 DNA de (respectivamente) populações de *P. psidii* provenienetes de três híbridos de *Eucalyptus, E. grandis* x *urophylla* (PEGU); *E. grandis* x *dunni* (clone 1366 [PEGD]) e *E. grandis* x *brassiana* (clone MA2000 [PEGB]); DNA de goiabeira – *Psidium guajava* (PGuava), *jabuticabeira – Myrciaria cauliflora* (Pjab) e jambeiro – *Syzygium jambos (*PSyz); *P. psidii* MF-1 na concentração de 40 ng/µL. (**1'** a **7'**) amplificação utilizando a combinação CIIRAP 2 e CIIRAP 4 para o mesmo material na mesma sequência. A seta branca indica a presença de bandas nos diferentes materiais de *P. psidii* analisados

3.3.2 Análise molecular de possíveis "TEs complexos" no genoma de *P. psidii* MF-1

3.3.2.1 Validação de *primers* específicos

A validação preliminar dos *primers* específicos utilizando a enzima Taq DNA Polimerase (convencional) não gerou amplificação de produtos (Tabela 3.2a). A segunda etapa corresponde à utilização de Taq DNA Polimerase "elongase" que permitiu a observação de que a temperatura de anelamento a 59ºC foi a mais eficiente (Tabela 3.2b). Apenas os *primers* correspondentes aos *contigs* 00075, 00088 e 00095

resultaram em *amplicons* de tamanho esperado para serem utilizados para clonagem e análises posteriores. Assim, dentre as doze combinações de *primers* avaliadas, sete resultaram em pelo menos uma amplificação e todas apresentaram-se específicas no anelamento a 59ºC. Destas sete combinações, apenas três apresentaram amplificação passível de reprodução sob as mesmas condições e permitiram avançar na etapa seguinte (transformação).

Tabela 3.2 – (**a**) Resultado da validação preliminar de *primers* específicos em gradiente de temperatura (45ºC a 60ºC) utilizando Taq DNA Polimerase. (**b**) Resultado da validação de *primers* específicos utilizando Taq DNA Polimerase do tipo "elongase" em duas temperaturas diferentes de anelamento

(continua)

(a)

Identificação das combinações[1]	**Combinação de *primers* específicos**	**Temperatura (ºC)**			
		45.0	**50.8**	**55.9**	**60.0**
Pp_1.1	Pp_Copia75 6 F e 10	-	-	-	-
Pp_1.2	Pp_Copia75 6 F e 6 R	-	-	-	-
Pp_2.1	Pp_Gypsy81 1 F e 1R	-	-	-	-
Pp_2.2	Pp_Gypsy81 1 F e 4 R	-	-	-	-
Pp_4.1	Pp_Copia 88 1 F e 1 R	-	-	-	-
Pp_4.2	Pp_Copia 88 1 F e 8 R	-	-	-	-
Pp_6.1	Pp_Copia104 4 F e 1 R	-	-	-	-
Pp_6.2	Pp_Copia104 10 F e 1 R	-	-	-	-

(b)

Identificação das combinações[1]	**Combinação de *primers* específicos**	**Temperatura (ºC)**	
		59.0	**65.0**
Pp_1.1	Pp_Copia75 6 F e 10	X	-
Pp_1.2	Pp_Copia75 6 F e 6 R	X	-
Pp_2.1	Pp_Gypsy81 1 F e 1R	-	-
Pp_2.2	Pp_Gypsy81 1 F e 4 R	X	-
Pp_3.1	Pp_Copia 84 1 F e 1 R	X	-
Pp_3.2	Pp_Copia 84 1 F e 8 R	X	-
Pp_4.1	Pp_Copia 88 1 F e 1 R	X	-

Tabela 3.2 – (**a**) Resultado da validação preliminar de *primers* específicos em gradiente de temperatura (45ºC a 60ºC) utilizando Taq DNA Polimerase. (**b**) Resultado da validação de *primers* específicos utilizando Taq DNA Polimerase do tipo "elongase" em duas temperaturas diferentes de anelamento

(conclusão)

(b)

Identificação das combinações[1]	Combinação de *primers* específicos	Temperatura (ºC)	
		59.0	**65.0**
Pp_4.2	Pp_Copia 88 1 F e 8 R	-	-
Pp_5.1	Pp_Copia95 2 F e 2 R	X	-
Pp_5.2	Pp_Copia95 3 F e 2 R	-	-
Pp_6.1	Pp_Copia104 4 F e 1 R	-	-
Pp_6.2	Pp_Copia104 10 F e 1 R	-	-

Nota: Sinais utilizados: (X) representa a amplificação de produto de PCR da respectiva combinação de *primers* na temperatura indicada;

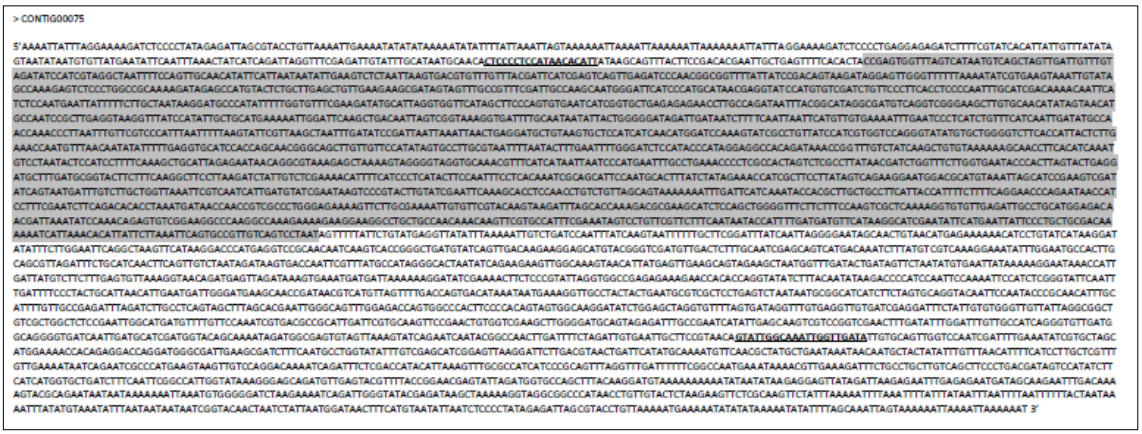
(-) representa que não houve amplificação de produto de PCR da respectiva combinação de *primers* na temperatura indicada;

(1) código atribuído às diferentes combinações referentes a cada um dos *contigs* utilizados para o desenho destes *primers*, o código da combinação foi também utilizado para "nomear" os *amplicons* gerados pela utilização da respectiva combinação de *primers* indicada

3.3.2.2 Obtenção de *amplicons* de Retrotransposons

Foram obtidos três *amplicons* de elementos Copia em *P. psidii* MF-1 (item 3.2.2.2), por meio da combinação dos *primers* Pp_Copia75 6F e Pp_Copia75 10R (combinação Pp_1.1 referente ao *contig*00075); Pp_Copia88 1F e Pp_Copia88 1R (combinação Pp_4.1 referente ao *contig*00088) e Pp_Copia95 2F e Pp_Copia95 2R (combinação Pp_5.1 referente ao *contig*00095) (Figura 3.3). Todos os produtos resultaram do anelamento de *primers* a temperatura de 59ºC (recomendada pelo fabricante dos *primers*). Os *amplicons* foram denominados Pp_1.1, Pp_4.1 e Pp_5.1 cujos tamanhos foram respectivamente 6.000 pb, 2.500 pb e 3.000 pb (Figura 3.4).

(a)



(b)

Figura 3.3 – Sequências de *contigs* 00075, 00088 e 00095 que originaram os três *amplicons* Pp_1.1, Pp_4.1 e Pp_5.1, respectivamente. Eles foram utilizados para o desenho de *primers* específicos de retrotransposons do tipo Copia. (**a**) Sequência de *contig*00075, com tamanho total de 4.682 pb, a região cinza da sequência representa os 2.083 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_75 6F e Pp_Copia_75 10R e o *amplicon* estimado apresenta 4.424 pb. (**b**) Sequência de *contig*00088, com tamanho total de 4.528 pb, a região cinza da sequência representa os 2.409 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_88 1F e Pp_Copia_88 1R e o *amplicon* estimado apresenta 2.273 pb. (**c**) Sequência de *contig*00095, com tamanho total de 4.474 pb, a região cinza da sequência representa os 3.391 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_95 2F e Pp_Copia_95 2R e o *amplicon* estimado apresenta 3.028 pb

(continua)

(c)

> CONTIG00095

5'AAACTTTAACAGATTTACAGTCTCGAAGCTCTTTGAATTTCTCTTTACGAGGTGGACTGCAAGATCACTATGACTGACAAAATCACCGACACAAAAGATCTAACAAACATTCCAATCTTAGATGGCACCAACTTTAGCCAATGGAATATTCGAATGAAGATTCATCTTCGTGCCAAAGA
TCTTTTAGATGTTTGTGAAAAGTCACTTCCTGGAGATGCAACTATCCCTGCTACAAACAAATGGCTCAAAGCTAGTTACGATGCCATAAACATCATAAACTACTCGAATCAGCGAGCGTGTTTTTCTAGAAGTCATCAACTCAGAGACGACCGAAAAAGCAAACTTGCTCTGGGCCAAAAT
AAATAAGCAGTATGCCTCAAAAAGGGCAATCAACAGGGGAAGAATTTGGATGGAGTTTCAAAGATACTTTTACAATGGTGATCTCCAAGTCTACATAGACTCCTGCAGAAAGCTTTTAATGGAACTTGAAACTGTAAGCATCAAGATTCCCAAAGAGTTACTTTCGTACTCTCTGCTTG
GAAAACTGGCTGGAGACTCTAAACTTCATCAGCTTGTTGAAACTCTCACCTTGAATGAAGAGCTTATCGAGAACCCAGAGCTCATACTCACTCGATTACAAGACTATGTTCACTTAATTGGATCACAAGATCCTCCTCCAACTAATCTACCTTCAGCACTAATATCAACCATAGACAATTCT
TTCAAAATAATACATTTTTGTACCAACGGCAAGCATAACCCAAAGAGCACAACTCACAAGAAAGAACAATGTTGGGCAGAAAACCCACAATTAAGGCCAAATCAAAGGACAACAAGCAAAGAAAGTTTCAACCAGCCGCTCATTTCTCGACAGCTACTGCACTTGTAACCTCGTTGGAA
GACGAAAAAACCAGCACCCATCAAGTTATCTTAGACTGTGGTGCTACGCATCACATGTTCAACTCAAAATCCTTCTTTACCTCTCTAAATAGCCTTCCACCGTTTACTGTCACCACAGGGGACTCAAGCAGCTCACTCTTGGCTCTTGGCACTGGTAGCGTCGAATTATTTTGCAACAGTCA
GCCTCTAGTATTGATGGACTGTCTGTACGTTCCTAAACTTAGCTGTAATCTCATAAGCTTACTCGCCCTGTTTGAAAAGAACATCACTGTTCATCGAGTAGAAGATAAATTTACCCTTGAATCCAATAGCAAGACAATTCTAAAAGGGAAAATACTCAACAAGCTAATGCACATCGATTAC
ATGTTGCCGACCACCCTCATTACTACAAAGGATCAAGCTCTGTGGCATAGGAGGCTAGGGCACCCAGGCACACTAGTGCTCAAGCGCATGGGCCTTCCCTATGAAATCACCACCTGTTCAGTTTGCGAAATCAACAAGTCTCACAAACAACCATTCAATCATCAGTTCGACAGTGCAGC
TCAACCTCTAGACTGCATTCACCTGGATTTGATAGGACCAGTAGTTCCTGCTTCAATATCTGGTTGTCAGTACATTCTCACCATTGTAGACCAAGCGACATCCTTCAAAATAGTAAGTTTCTAAAAAAGAAGTCTGAAGCATTCGATCAGTTTTGTATGTCAAGACCTTCATGGAAAATCA
ACAAGATCGAAAAGGTTAAGAGTCTTGTCAGACAGAGGGGGCGAATTCATGAACAAGAACTTCAGGGAACTTGCTGAGAAGTATGGATTCATGCACATCTTCTCTCCTCCCCAAAACACCCCAGCATAATGGATATGCTGAGAGAGCAAATAGGACAATCCTAGAAAAAGCGAGATGC
CTACTTAATTCATCAAACTTGCCTAATCCTTACTGGGCCGAAGCGGTAAATACTGCTACTCTACTCTCAAACCTGGTCCCAACCCCTTCCAGGTTGAATCGATCCCCTTACTCGCTCTGGACCAATCAACCACCTCGAGTTAAACGTCTTAGGACGTTTGGATGTCTTGCATTTATTCCTACT
GCAAGAAATCATTGCTCTTGGAAATTGGGACCAGCAGGTACGGAAGGAATTCTATTGGGGTATGAGAATGAAAATACATCCTATCCATTCTGCAAATCTCAGATGCCAAAATCATTGTCACCAAACATGTGACTTTTGATGAGAATGAGCTCCCAAGACTAAAAGGGTTTGTTGAGAG
CAGTTCACTGAATCTAGATAATTTCACAACGATGGTAGATGAAGCCCATCTTGTTGAACAACAAGGAGACGTAAGCTGTTCAACGCTGGTAGATGAAGTCCGGGTTGACAGTTTAGATGAACGTTCCTCGTCAGACATGGTAGATGAAGTCCATGCTGATGAGGAAGTATCTCAGTCT
CTTGGATCAGTAACTGCTCCAGCAACTGTCCCTCGCATTAAGGTCATAGGACCACGACACCCTACCTTGGTGTCATCAGACATCAACAATCTGAACATACTACCGTACCCAAGGAGAGCAGATGCCCTGCTCTCCTCCGCGGACGTCACACCCCATTCTTTTAAAGGTGCACTACTATCAC
CTGAAAAAGACTTCTGGATGGTTGCCATCGATAAGGAGCTCCTGTCAATGAGCAACCTGAAAGTTTGGGAGGTAACAGAACGACAAGACCACTACAAGCTTGTCGGAACAACGTGGGTATTTAGAATGAAAAGAGATCACACTAATAAAGTGACAGAGTACAAAGCTCGACTTTGTG
CTCAGGGCTTCACTCAAACGCAAGGGATTGATTTTGAGAAGACTTATGCCCCAACCGGGCGCCTCAATTCGTTGCGAATGCTGATTGCATTTGCTTCAGCCAAAAGCCTAAAGTTCCACCAAATTGACATCAAGAGCGCATTCTTGAATGCACCACTGAAGGAGGAGGTATATCTTCAT
GTTCCTCAAGGTTTGTCTTTGGACCAAACAAAGTACTGCCTTCGATTGAAAAAGGCAATCTACGGACTGAAACAGGCACCGCTTGTCTGGTACCAACACTTGAAGGAATATCTAATCAGCATCAATTTCTCTGCATGTGTGCTTGACGCCTGTGTGTTCTACCGATCGGAACCCAGACCT
GTATGGCTGTACATTCACGTTGATGACATTGCTGTCATCGGTGAAGAGGTTGAGGACTTCAAGAAGGAAATTCGCTCTCAAGTTTGAAATTAAGGATCTAGGCGCGGCTGATCTCATGCTGGAGTAAAAGTAACTCATGGCCCTGATTCGATAAGCCTGGACCAGCAACATTTGCTG
AATCGCTACTTGAACAGTACGGTATGGGTGAATGTAGAACAGTCTCCACCCCTCTGGTACCGAACGAACAGTTGGGTACTGCGACGGAGGAAGAAATGATTGCCCTCAGAAACCTCAAGGTTAACTACCGTAGTGCAATAGGTAGTATTAATTATCTAAGCACTGCGACTCGCCCTGA
CCTGTTGTTTGCAGTCAGTTGTCTATCTCAGTACTTAGAGAATCCTGGGCTCAAGCATTGGCAGGCTTTCATGCACGTTCTACGGTATTTAAGAGGATCTCTGGACGTCTGCCTTGAGTATCCGAGGGGAGGAAGCCAAGGCATCACTGCCTGGAGTGATGCAGACTGGGGGAATTGT
CGGTCCACCAGACGGTCCATAACGGGCTATCTAGCGACCTTCCATGGGTGCCTAGTACTATGGAAGACCAGGAAGCAACCCTTGGTATCTATTTCGACCGCAGAAGCCGAGTACAAGGCTCTGTGCGACCTCACTTCCGAACTCCTTTGGCTGAAACAGTGGTCTGAAGAGGCAAGAA
TTCTGAAATTAGAGGCGCCTATCAATGTCTGGGAAGATAATCAGAGTTGCATTAACACCGCCAACGGTGACTGCAACTTCAATGGGCGAAGAATGAAGCATGTAGACATTCAACTACATTTCGTCAAAGAAGTAATCCGGTCGTCAATCATCAACTTTGTTACGCACCAAGCACCAAAA
TGCTGGCAGATTTCCTCACCAAATCCGTCAACAAGGTAAGCCTACTCAAGTCTCTCGAAGCACTGGGCATTTTGAGAATAGGAGTGAGGGGGGGATGTGAAAGAATCTGATCTAATAGCCTCAAAGCGTCTTAGATCATGTGAGCACTCCTATCCCAAAAGTGCACTAGTCTCCTTGTTC
CCAGTAATACCACGGCTTACTCAAGCTACTCTTGACCTAGGAATTACGATGCTCCTGAAGATATAAAGCATCTCCTTGAACGACGAAGGAGGGGACATTTTGGGTTTTCATTCACTTATCATCTCTCTACTCTTTATTATTACTTCCTTCCTTATTTGTCTTTTCTGTTTCTTTCATCACATTCA
TTCCTAAATTATTGTCTAAATATTAGTTGCTCTTACATTTGCCTGTTGTCGTTGCTCTCCATTACACTTGTAGCTCAAACTTGCTTTCCACCTTAGGTACTTTCTTACTCACTCTTTTCTGATTGTTTACTGCATATTAATGAAATATCAACCTTAGCTTCTG 3'

Figura 3.3 – Sequências de *contigs* 00075, 00088 e 00095 que originaram os três *amplicons* Pp_1.1, Pp_4.1 e Pp_5.1, respectivamente. Eles foram utilizados para o desenho de *primers* específicos de retrotransposons do tipo Copia. (**a**) Sequência de *contig*00075, com tamanho total de 4.682 pb, a região cinza da sequência representa os 2.083 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_75 6F e Pp_Copia_75 10R e o *amplicon* estimado apresenta 4.424 pb. (**b**) Sequência de *contig*00088, com tamanho total de 4.528 pb, a região cinza da sequência representa os 2.409 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_88 1F e Pp_Copia_88 1R e o *amplicon* estimado apresenta 2.273 pb. (**c**) Sequência de *contig*00095, com tamanho total de 4.474 pb, a região cinza da sequência representa os 3.391 pb classificados como TE pelo programa *RepeatMasker*, os oligonucleotídeos sublinhados e em negrito acima (*primer forward*) e abaixo (*primer reverse*) correspondem aos *primers* Pp_Copia_95 2F e Pp_Copia_95 2R e o *amplicon* estimado apresenta 3.028 pb

(conclusão)

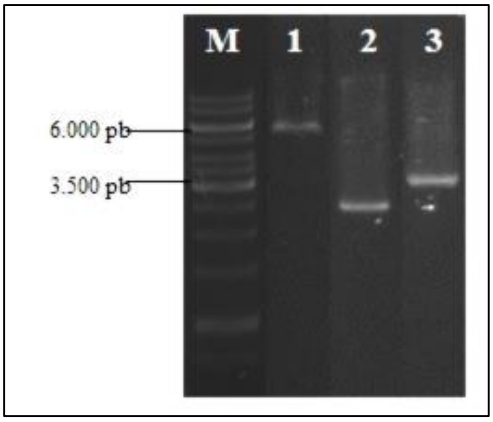


Figura 3.4 - Amplificação de DNA genômico de *P. psidii* MF-1 (20 ng/µL) utilizando as combinações de *primers* específicos para retrotransposons: (**M**) Marcador 1Kb (Fermentas); (**1**) Pp_Copia75 6F e Pp_Copia75 10R (combinação 1.1 referente ao *contig*00075); (**2**) Pp_Copia88 1F e Pp_Copia88 1R (combinação 4.1 referente ao *contig*00088) e (**3**) Pp_Copia95 2F e Pp_Copia95 2R (combinação 5.1 referente ao *contig*00095). Os tamanhos aproximados das bandas obtidas para os *amplicons* de Pp_1.1, Pp_4.1 e Pp_5.1 foram estimados em aproximadamente 6.000 pb, 2.500 pb e 3.000 pb, respectivamente de acordo com o marcador molecular 1 Kb

Os *primers* referentes as combinações Pp_1.2, Pp_2.2; Pp_3.1 e Pp_3.2 (Tabela 3.2) não resultaram em *amplicons* para a etapa seguinte de transformação devido à impossibilidade de obter novos produtos destes a partir destes *primers* em subsequentes repetições, portanto a quantidade de produto de PCR necessária às demais etapas não foi atingida e então elas não foram utilizadas. As únicas combinações que não apresentaram nenhuma amplificação em nenhuma das reações de PCR realizadas são referentes ao *contig*00104 (combinações Pp_6.1 e Pp_6.2).

3.3.2.3 Clonagem dos *amplicons* de Retrotransposons

Todos os três produtos de ligação dos *amplicons* de retrotransposons Pp_1.1, Pp_4.1 e Pp_5.1 clonados no vetor pGEM®T Easy resultaram em pelo menos uma cultura recombinante (Figura 3.5a). As colônias transparentes/brancas referem-se aos recombinantes de elementos em *E.coli*, assim essas colônias foram selecionadas e repicadas em meio sólido LB suplementado com ampicilina (100 µg.mL^{-1}) para posteriores análises (extração de plasmídeo) (Figura 3.5b).

(a) **(b)**

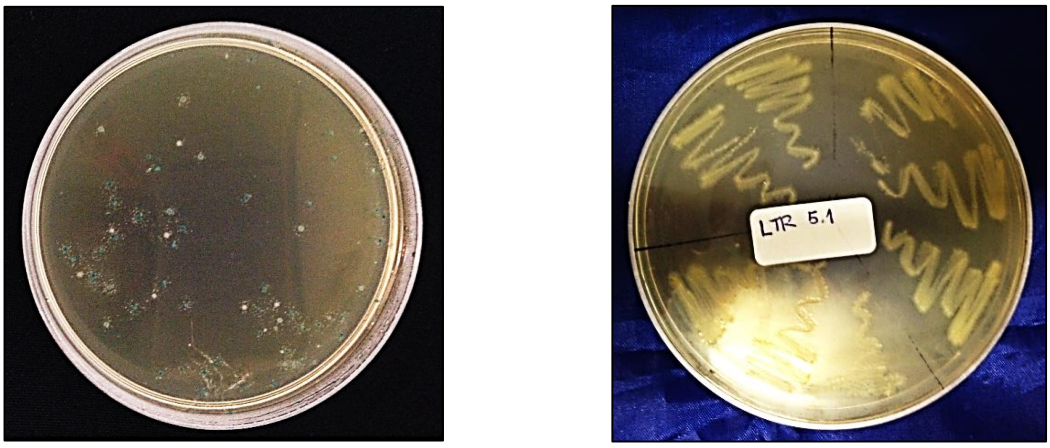


Figura 3.5 - (**a**) Placa de meio LB suplementado com antibiótico ampicilina (100 µg.mL^{-1}) e 2 µL/mL de meio de X-Gal 2%, as colônias brancas são os recombinantes positivos (inserção de *amplicon* de retrotransposon em célula de *E. coli* DH5-α). Colônias de coloração azul referem-se à células eletro-competentes de *E. coli* DH5-α que foram transformadas com o vetor pGEM®T Easy sem o inserto de interesse. (**b**) Recombinante pGEM::Pp_5.1, selecionado e repicado em meio LB suplementado com antibiótico ampicilina (100 µg.mL^{-1}) para análises posteriores

3.3.2.4 Extração de plasmídeo pGEM®T Easy clonado

A realização da extração de plasmídeos pGEM::Pp_1.1, pGEM::Pp_4.1 e pGEM::Pp_5.1, mostrou-se eficiente (Figura 3.6).

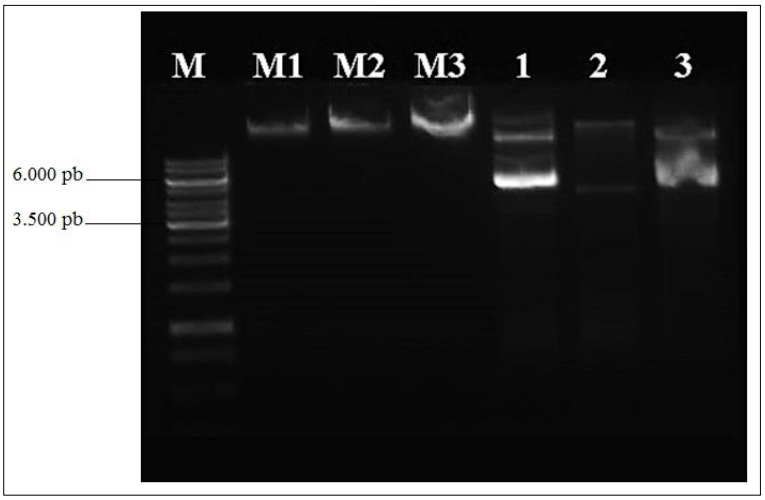


Figura 3.6 – Verificação de extração de plasmídeo realizada com *QIAprep® Spin Miniprep Kit* (QIAgen), para o vetor pGEM®T Easy clonado com três insertos diferentes de *P. psidii* MF-1: (**1**) Pp_1.1 (*contig*0075), (**2**) Pp_4.1 (*contig*00088) e (**3**) Pp_5.1 (*contig*00095). (M) Marcador 1Kb (Fermentas); (M1 a M3) Régua molecular λ (lambda) nas concentrações de 25 ng/µL, 50 ng/µL e 100 ng/µL, respectivamente. Os tamanhos dos vetores clonados são desconhecidos

O DNA plasmidial obtido de todos os três recombinantes positivos (pGEM::Pp_1.1, pGEM::Pp_4.1 e pGEM::Pp_5.1) foram enviados para a realização do sequenciamento no Centro de Estudos do Genoma Humano, São Paulo, Brasil.

A alta similaridade entre as duas sequências (modelo e obtida após amplificação) confirmará a presença de sequências de híbridos de TEs no genoma de *P. psidii* MF-1, permitindo consequentemente, que a hipótese de erro de montagem do genoma parcial seja descartada.

3.4 Discussão

O modo de amplificação e inserção no genoma dos elementos de transposição, especialmente retrotransposons, permite a utilização dos mesmos para discriminação de espécies e/ou genótipos devido a sua abundância nos diversos genomas (SANTANA, 2009). Cada evento de transposição resulta em um polimorfismo de inserção passível de detecção por meio de técnicas moleculares, TEs no genoma podem apresentar-se com inúmeras cópias e, portanto gerar centenas e até milhares de marcas polimórficas. Assim a presença ou ausência de um TE em certo *locus* pode

ser usada como marca molecular de *fingerprinting* de um determinado genótipo, mapa de ligação e investigação em diversidade e filogenia (GRZEBELUS, 2006; SANTANA, 2009).

O polimorfismo existente entre os TEs vem sendo amplamente utilizado como ferramenta para acessar a diversidade genética de organismos que apresentam uma quantidade moderadamente alta de elementos de transposição, algumas das técnicas baseadas na análise de perfis gerados por amplificação via PCR e hibridização conhecidas até o momento são *Transposon Display*, IRAP (*Inter-Retrotransposon Amplified Polymorphism*), REMAP (*Retrotransposon-Microsatellite Polymorphism*), IMP (*Inter-MITE Polymorphism*), S-SAP (*Sequece-Specific Amplification Polymorphism*) e RBIP (*Retrotransposon Based Insertion Polymorphism*) (GRZEBELUS, 2006).

A técnica IRAP é muito utilizada devido a sua baixa complexidade (KALENDAR *et al*., 2011) e seu poder em revelar polimorfismos em regiões em que os retrotransposons estão inseridos, assim *primers* para regiões conservadas, como as LTRs são desevolvidos (Figura 3.7) (KALENDAR *et al*., 1999).

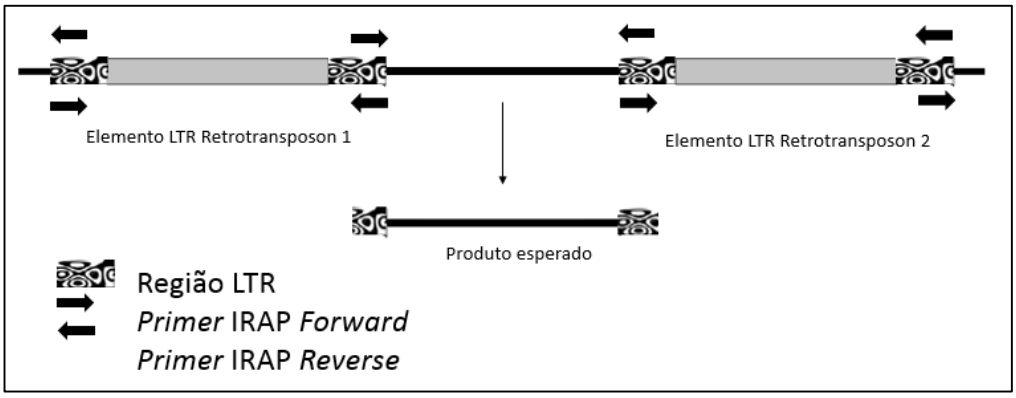


Figura 3.7 – Estratégia de amplificação de *primers* específicos para a técnica IRAP. Os *primers* IRAP alinham-se às regiões LTR dos elementos do tipo LTR Retrotransposons, tanto no sentido 5' →3' como 3' →5'. O produto esperado corresponde ao intervalo entre um elemento LTR Retrotransposon e o próximo

Alguns estudos indicam que os TRIMs (*Terminal-Repeat Retrotransposon in Miniature*), elementos abundantes em diversas espécies, considerados não-autônomos por não possuírem um ou mais genes essenciais para a transposição, podem ser utilizados para estudos de diversidade entre espécies correlacionadas (SANTANA *et al*., 2013). Santos *et al*. (2012) caracterizaram um elemento TRIM-*like*, chamado *RetroCI1*, que pode ser utilizado para análises de variabilidade genética em

diferentes gêneros de fungos pela técnica de IRAP (*Inter-Retrotransposon Amplified Polymorphism*), como *Colletotrichum* (SANTOS *et al*., 2012), *Moniliophthora, Cryptococcus, Sporobolomyces, Rhodotorula, Stemphylium, Microsphaeropsis, Leptosphaerulina, Cochliobolus, Ampelomyces, Epicoccum, Cladosporium, Cercospora, Penicillium, Cladophialophora, Fusarium, Myrothecium, Annulohypoxylon, Xylaria, Anthostomella, Hypoxylon, Diaporthe* e *Phomopsis* (SANTANA *et al*., 2013). Uma das grandes vantagens de se utilizar marcadores com base em IRAP, é a sua alta capacidade de rastrear eventos de inserção e sua subsequente verticalização por meio de filogenia (KALENDAR *et al*., 1999).

Dentre os fungos avaliados por Santana *et al.* (2013), seis isolados eram Basidiomycota (*Moniliophthora* spp., *Cryptococcua* sp. e *Rhodotorula* sp.) o número de *amplicons* gerados variaram entre 250 pb a 4.000 pb e a combinação dos *primers* (CIIRAP 1 e CIIRAP 4 ou CIIRAP2 e CIIRAP 4) determinou a quantidade de polimorfismos detectados em cada espécie. Apesar de *primers* CIIRAPs serem descritos na literatura como ótimos para detecção de polimorfismos inter e intra espécies (SANTANA *et al*., 2013) no presente estudo sobre diversidade populacional de *P. psidii* a técnica não se mostrou eficiente.

Apesar do uso de TEs serem amplamente empregados como marcadores em fungos, para *P. psidii* a técnica IRAP não foi eficiente na detecção de polimorfismo entre as populações estudadas. Esse resultado não era esperado, uma vez que não condiz com os resultados obtidos pelas análises *in silico* (capítulo 2), que demonstraram que mais de 90% dos elementos presentes em *P. psidii* MF-1 são do tipo LTR Retrotransposons e que cerca de 50% de todo o DNA genômico é devido à elementos de transposição (item 2.3.1).

Além disso, diferente do que foi demonstrado por Santana *et al*. (2013), em que a técnica IRAP permitiu a análise de diversidade entre 27 espécies de fungos. Na análise de populações de *P. psidii*, a técnica não foi passível de repetição, ou seja, não houve a reprodutibilidade da amplificação. Destaca-se que Santana (2009) ao estudar 36 isolados do fungo patogênico *Moniliophthora perniciosa* por IRAP ressaltou a importância da reprodutibilidade dos perfis de bandas do DNA genômico em questão, que foram realizadas por repetição da reação de PCR com os *primers* selecionados.

A análise de diversidade de populações de diferentes hospedeiros de *P. psidii* realizada no presente trabalho utilizou parte do material genético de *P. psidii*

proveniente dos hospedeiros: *E. grandis* x *urophylla* (PEGU); *E. grandis* x *dunni* (clone 1366 [PEGD]) e *E. grandis* x *brassiana* (clone MA2000 [PEGB]), goiabeira – *Psidium guajava* (PGuava), jabuticabeira – *Myrciaria cauliflora* (Pjab) e jambeiro – *Syzygium jambos (*PSyz), já utilizados por Quecine *et al*. (2014) por meio de PCR-DGGE e T-RFLP.

Tanto a técnica PCR-DGGE como a T-RFLP revelaram que as populações de *P. psidii* provenientes de outras espécies de Myrtaceae (goiabeira, jambeiro e jabuticabeira) agruparam em um clado, enquanto que as populações originadas de *E. grandis* (clone D901 e clone 1600) foram agrupadas em outro clado. Relatou-se a existência de diferença e alta diversidade entre diversos patógenos *P. psidii* originários de goiabeira, jambeiro e jabuticabeira e *Eucalyptus* spp. Ressalta-se portanto, a existência de polimorfismo entre as diferentes populações de *P. psidii*, no entanto no presente trabalho a técnica de IRAP não permitiu detectar estes polimorfismos.

Os *primers* CIIRAP utilizados neste estudo foram construídos a partir de sequências conservadas de LTRs da espécie *Colletotrichum lindemuthianum* 1 – *RetroCl1* (SANTOS *et al*., 2012). Visto a ineficiência do IRAP para *P. psidii* MF-1, o desenvolvimento de *primers* específicos a partir de regiões de LTR conservadas no genoma de *P. psidii* representa uma possível alternativa para viabilizar a utilização da técnica no estudo de diversidade populacional.

Destaca-se a relevância desta técnica ser otimizada para o melhor entendimento de polimorfismos apresentados pelas diferentes populações de *P. psidii*, pois até o presente momento, não foram encontrados estudos de fungos da ordem Pucciniales utilizando marcadores de polimorfismos construídos a partir de elementos transponíveis. Futuramente, a avaliação do perfil de LTR retrotransposons em diversas populações de *P. psidii* pode fornecer subsídios para que a técnica seja mais explorada permitindo futuros estudos.

Por outro lado, ao longo do desenvolvimento deste trabalho foi observada uma grande dificuldade na classificação *in silico* de TEs presente no genoma parcialmente sequenciado de *P. psidii* MF-1, devido a uma série de fatores já discutidos anteriormente. Os TEs estão incompletos (nenhuma sequência terminal repetitiva foi encontrada dentre os *contigs* do fungo analisados) e aparecem como potenciais híbridos de TEs.

Acredita-se na possibilidade de existência de elementos que são cópias-fóssil (potenciais alvos de mutação) ao longo do genoma, frações de elementos inseridos

de forma aninhada em outros elementos, resultantes da inserção de um elemento dentro de outro, podem causar a inativação do elemento “receptor” da inserção transposicional, como consequência há a dificuldade em verificar as regiões conservadas por meio dos programas convencionais.

Certamente, a análise *in silico* isoladamente não permite responder a este problema na classificação, que apesar de pouco provável, pode ser um erro de montagem. A verificação da integridade de elementos de transposição de duas famílias distintas por programas tradicionalmente utilizados para este fim em diversos fungos fitopatogênicos não se mostrou tão eficiente para a classificação de TEs presentes em *P. psidii*.

Vale ressaltar que os TEs complexos ou híbridos de TEs são fenômenos comumente observados em sequências genômicas, podendo causar confusão na anotação de elementos transponíveis (WICKER *et al*., 2007). Alguns são resultados da integração de um elemento dentro (aninhado) de outro elemento de transposição e também podem resultar de recombinação intra-cromossomal e até mesmo de variações na replicação.

Inserções antigas tendem a sofrer mutações e consequentemente inserções e rearranjos que irão interromper a continuidade e reduzir a similaridade destes a famílias de TEs conhecidas (WICKER *et al*., 2007). Neste caso, Wicker *et al*. (2007) sugerem que a anotação seja feita individualmente com fragmentos de pelo menos 80 nucleotídeos, elementos intensamente interruptos e novos são classificados como “desconhecidos”, isso pode estar ocorrendo com o TEs presentes no genoma de *P. psidii,* sendo assim, essa estratégia foi adotada no presente trabalho.

Dentre os vinte e cinco *contigs* anotados manualmente (item 2.3.4), seis *contigs* (00075, 00081, 00084, 00088, 0095 e 00104) foram selecionados para efetuar o desenho de *primers* específicos e então a amplificação da possível região correspondente a um híbrido de TE. A escolha dos mesmos se baseou no fato de apresentarem a maior quantidade de sequências proteicas e sua classificação pelo programa *RepeatMasker*, que apresentou sobreposição e/ou interrupção de fragmentos pertencentes a elementos transponíveis. Apesar do fato de que até o presente momento, essas sequências ainda não estarem sequenciadas, o tamanho dos *amplicons* obtidos, condiz com o tamanho esperado. Acredita-se que a presença de fragmentos sobrepostos e interrompidos nas sequências referentes a TEs em *P. psidii* MF-1 são realmente resultado da presença de híbridos de TEs em seu genoma.

Epidemiologicamente, os elementos de transposição são importantes para o conhecimento da organização da especificidade entre hospedeiro e patógeno e da rapidez em que o patossistema muda. Neste contexto, o modo como estes elementos dispersam ou se conservam no genoma de fungos representam uma marca valiosa no estudo de populações de patógenos de plantas e animais (GABRIEL *et al*., 2006). Comprovando-se a real existência de híbridos de TEs no genoma do fungo *P. psidii* pode-se então inferir sobre a atividade de replicação de retrotransposons que pode produzir novas inserções no genoma aumentando assim o polimorfismo intra e inter espécies.

Visto toda influência que os elementos de transposição, especialmente os da Classe I – LTR Retrotransposons podem exercer sobre o genoma do hospedeiro. Ressalta-se a importância do uso concomitante de análises *in silico* e confirmações dos dados por técnicas moleculares. Por isso, espera-se que a análise não só *in silico*, mas também com o uso de ferramentas moleculares, permita a verificação de que as sequências disponíveis do genoma de *P. psidii* MF-1 são realmente híbridos de TEs*,* a confirmação desses dados inéditos, é de suma importância para o melhor entendimento da biologia do patógeno, bem como desvendar a confusão verificada no sistema de classificação *in silico* por alinhamento e similaridade com TEs de bancos de dados observada.

Referências

DABOUSSI, M.J.; CAPY, P. Transposable elements in filamentous fungi. **Annual Review of Microbiology**, Palo Alto, US, v. 57, p. 275-299, 2003.

GABRIEL, A.; DAPPRICH, J.; KUNKEL, M.; GRESHAM, D.; PRATT, S.C.; DUNHAM, M.J. Global Mapping of Transposon Location, **PLos ONE**, San Francisco, US, v. 2, n. 12, Dec. 2006. Disponível em: <DOI: 10.1371/journal.pgen.0020212.> Acesso em: 07 jul. 2015.

GLEN, M.; ALFENAS, A.C.; ZAUZA, A.V.; WINGFIELD, M.J.; MOHAMMED, C. *Puccinia psidii*: a threat to the Australian environment and economy a review. **Australasian Plant Pathology**, Clayton, AU, v. 36, n. 1, p. 1-16, Jan. 2007.

GRZEBELUS, D. Transposon insertion polymorphism as a new source of molecular markers. **Journal of Fruit and Ornamental Plant Research**, Skierniewice, PL, v. 14, suppl. 1, p. 21–29, 2006.

HUA-VAN, A.; LE ROUZIC, A.; MAISONHAUTE, C.; CAPY, P. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. **Cytogenetic and Genome Research**, New York, US, v. 110, p. 426-440, 2005.

INDÚSTRIA BRASILEIRA DE ÁRVORES. São Paulo, 2015. 64 p.

KALENDAR, R.; FLAVELL, A.J.; ELLIS, T.H.N.; SJAKSTE, T.; MOISY, C.; SCHULMAN, A.H. Analysis of plant diversity with retrotransposon based molecular markers. **Heredity**, Edinburgh, GB, v. 106, n. 4, p. 520–530, Apr. 2011.

KALENDAR, R.; GROB, T.; REGINA, M.; SUONIEMI, A.; SCHULMAN, A. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. **Theoretical and Applied Genetics**, Berlin, DE, v. 98, p. 704–711, 1999.

KANG, S.; LEBRUN, M.H.; FARRALL, L.; VALENT, B. Gain of virulence caused by insertion of a Pot3 transposon in a Magnaporthe grisea avirulence gene. **Molecular Plant-Microbe Interactions**, Saint Paul, US, v. 14, n. 5, p. 671-674, May. 2001.

KIDWELL, M.G.; LISCH, D.R. Transposable elements and host genome evolution. **Trends in Ecology & Evolution**, Cambridge, GB, v. 15, n. 3, p. 95-99, Mar. 2000.

LEITE, T.F. **Análise histológica da interação planta patógeno entre *Eucalyptus grandis* e *Puccinia psiddi* Winter e identificação de genes diferencialmente expressos em plantas resistentes e susceptíveis inoculadas com o patógeno por meio da técnica de RNA-Seq.** 2012, 177 p. Tese (Doutorado em Ciências) Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.

MARTIN, F.; AERTS, A.; AHRÉN, D.; BRUN, A.; DANCHIN, E.G.; DUCHAUSSOY, F.; GIBON, J.; KOHLER, A.; LINDQUIST, E.; PEREDA, V.; SALAMOV, A.; SHAPIRO, H.J.; WUYTS, J.; BLAUDEZ, D.; BUÉE, M.; BROKSTEIN, P.; CANBÄCK, B.; COHEN, D.; COURTY, P.E.; COUTINHO, P. M.; DELARUELLE, C.; DETTER, J.C.; DEVEAU, A.; DIFAZIO, S.; DUPLESSIS, S.; FRAISSINET-TACHET, L.; LUCIC, E.; FREY-KLETT, P.; FOURREY, C.; FEUSSNER, I.; GAY, G.; GRIMWOOD, J.; HOEGGER, P.J.; JAIN, P.; KILARU, S.; LABBÉ, J.; LIN, Y.C.; LEGUÉ, V.; LE TACON, F.; MARMEISSE, R.; MELAYAH, D.; MONTANINI, B.; MURATET, M.; NEHLS, U.; NICULITA-HIRZEL, H.; OUDOT-LE SECQ, M.P.; PETER, M.; QUESNEVILLE, H.; RAJASHEKAR, B.; REICH, M.; ROUHIER, N.; SCHMUTZ, J.; YIN, T.; CHALOT, M.; HENRISSAT, B.; KÜES, U.; LUCAS, S.; VAN DE PEER, Y.; PODILA, G.K.; POLLE, A.; PUKKILA, P.J.; RICHARDSON, P.M.; ROUZÉ, P.; SANDERS, I.R.; STAJICH, J.E.; TUNLID, A.; TUSKAN, G.; GRIGORIEV, I.V. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. **Nature**, London, GB, v. 452, n. 7183, p. 88-92, Mar. 2008.

MURATA, H.; BABASAKI, K.; SAEGUSA, T.; TAKEMOTO, K.; YAMADA, A.; OHTA, A. Traceability of Asian Matsutake, Specialty Mushrooms Produced by the Ectomycorrhizal Basidiomycete *Tricholoma matsutake*, on the Basis of Retroelement-Based DNA Markers. **Applied and Environmental Microbiology**, Washington, US, v. 73, n. 7, p. 2023-2031, Apr. 2008.

QUECINE, M.C.; BINI, A.P.; ROMAGNOLI, E.R.; ANDREOTE, F.D.; MOON, D.H.; LABATE, C.A. Genetic Variability in *Puccinia psidii* Populations as Revealed by PCR-DGGE and T-RFLP Markers. **Plant Disease**, Saint Paul, US, v. 98, n.1, p.16-23, Jan. 2014.

SAMBROOK, J.; RUSSEL, D.W. (Ed.) **Molecular cloning:** a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press. 2001. 2344p.

SANTANA, M. F.; BATISTA, A. D.; RIBEIRO, L. E.; ARAÚJO, E. F.; QUEIROZ, M. V. Terminal repeat retrotransposons as DNA markers in fungi. **Journal of Basic Microbiology**, Berlin, DE, v. 53, n. 10, p. 823-827, Oct. 2013.

SANTANA, M.F. **Retrotransposons LTR no genoma de *Moniliophthora perniciosa* e *Cochliobolus heterostrophus*.** 2009, 81 p. Dissertação (Mestrado em Microbiologia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2009.

SANTOS, L.V.; QUEIROZ, M.V.; SANTANA, M.F.; SOARES, M.A.; de BARROS, E.G.; de ARAÚJO, E.F.; LANGIN, T. Development of new molecular markers for the *Colletotrichum* genus using *RetroCl1* sequences**. World Journal of Microbiology and Biotechnology**, Oxford, GB, v. 28, n. p. 3, p. 1087–1095, Mar. 2012.

WICKER, T; SABOT, F; HUAN-VAN, A.; BENNETZEN, J. L.; CAPY, P.; CHALHOUB, B.; FLAVELL, A.; LEROY, P.; MORGANTE, M.; PANAUD, O.; PAUX, E.; SANMIGUEL, P.; SCHULMAN, A. H. A unified classification system for eukaryotic transposable elements. **Nature**, London, GB, v. 8, n. 12 p. 973–982, Dec. 2007.

4 CONSIDERAÇÕES FINAIS

A área de florestas plantadas no Brasil apresenta 70% de sua composição correspondente à cultura de espécies de *Eucalyptus.* Ao longo do seu desenvolvimento, o eucalipto está sujeito a diversas doenças, sendo que a ferrugem, causada pelo fungo *Puccinia psidii*, é considerada a grande ameaça à cultura e também a diversas outras espécies da família Myrtaceae. A ferrugem é conhecida como patógeno do *Eucalyptus* spp. desde o século XX, porém informações a respeito da interação molecular planta-patógeno, mecanismos de resistência, assim como a biologia do fungo e sua virulência são incipientes.

Os organismos eucariotos podem apresentar em seu genoma até 50% de elementos transponíveis (TEs), uma classe de DNA repetitivo capaz de influenciar a organização, integridade e evolução do genoma hospedeiro, além de regular a expressão gênica. Especialmente em espécies de fungos patogênicos, o estudo de elementos transponíveis permite obter maior conhecimento sobre diversos mecanismos, como o silenciamento gênico de genes relacionados à virulência, além de servir como ótimo marcador de polimorfismo no estudo de diversidade populacional.

Dessa maneira, o estudo da atividade de elementos transponíveis em *P. psidii* Winter, é de suma importância para o melhor entendimento da biologia do patógeno e indiretamente no controle de ferrugem de eucalipto no Brasil e no mundo.

O capítulo 2 do presente trabalho descreve a classificação e caracterização *in silico* de possíveis elementos transponíveis do genoma parcialmente sequenciado de *P. psidii* (linhagem monopustular MF-1) por duas estratégias, ambas apontaram a alta frequência de elementos da Classe I – LTR Retrotransposons. A verificação da integridade de *contigs* não apresentou resultados positivos, nenhuma das regiões importantes na identificação (LTRs para Classe I e TIRs para Classe II e TSRs) foram encontradas.

Sugere-se como hipótese que há alta taxa de mutação dos elementos que compõem o genoma de *P. psidii* MF-1, alterando as regiões conservadas a ponto de não serem passíveis de detecção por algoritmos convencionais. As mutações podem também ser responsáveis por gerar sequências recombinantes deles, descritos na literatura como híbridos de TEs ou TEs complexos e a classificação destas sequências torna-se bastante confusa.

A anotação manual permitiu observar que existem sequências conservadas relacionadas à transposição. Destaca-se a importância da incorporação de análises manuais às automáticas para refinar os resultados e confirmação de dados gerados.

A distribuição de TEs apresentada por *P. psidii* MF-1 corrobora com os dados presentes na literatura, que relata uma maior incidência de elementos LTR Retrotransposons e elementos mais comuns no reino Fungi, além disso, esta distribuição foi também descrita em outras espécies de fungos, causadores de ferrugem.

A comparação com outras *Puccinia* spp. e *P. psidii* MF-1 com relação ao repertório de elementos transponíveis apresentou divergência. As espécies *P. graminis*, *P. striiformis* e *P. triticina* apresentam maior frequência de elementos da Classe II – DNA Transposons, enquanto *P. psidii* MF-1 tem seu genoma majoritariamente constituído por LTR Retrotransposons.

Visto a importância destes na geração de polimorfismos e variabilidade genética no genoma hospedeiro devido à capacidade de “deixar uma cópia e se transpor”. A fração de LTR Retrotransposons pode estar relacionada à capacidade de *P. psidii* apresentar-se como patógeno de diversas espécies de Myrtaceae, fato não observado às demais *Puccinia* spp. que apresentam uma relação mais estreita espécie-específica de infecção.

Ressalta-se portanto, a importância da continuidade destes estudos para o melhor entendimento em escala evolutiva de *P. psidii* bem como a diversidade populacional desta espécie para o delineamento eficiente de uma ferramenta de controle biológico.

No capítulo 3, visando uma aplicação prática do estudo de TEs, foi avaliada a técnica IRAP para análise de diversidade populacional em *P. psidii*. Este estudo é importante para o entendimento da biologia do fungo e dos mecanismos envolvidos à quebra de resistência em diversas espécies Myrtaceae.

Destaca-se então a importância de otimizar e até desenvolver novas técnicas de análise de diversidade populacional baseada em polimorfismos gerados por LTR Retrotransposons para a compreensão deste fungo biotrófico causador de ferrugem. Uma vez desenvolvidos, os marcadores moleculares podem ser também utilizados em outros fungos da ordem Pucciniales e até do filo Basidiomycota, já que as sequências são conservadas.

Como observado pelos resultados obtidos, existe uma dificuldade considerável na classificação de TEs em *P. psidii.* Tendo em vista a limitação da disponibilidade de informações no Banco de Dados do *Puccinia Group Genome Database - Broad Institute* e também da complexidade do genoma de fungos da ordem Pucciniales, considerados organismos de tamanhos de genoma grandes.

A identificação e caracterização de TEs no genoma de fungos fitopatogênicos podem fornecer informações importantes sobre o mecanismo de transposição, ativação, distribuição horizontal e vertical, mecanismos de silenciamento gênico, organização do genoma hospedeiro e estudo de diversidade genética e populacional. Todos estes aspectos são relevantes e podem contribuir de alguma maneira para a melhor compreensão da relação planta-patógeno, mecanismos de virulência, entre outros.

Visto toda influência que os elementos de transposição, especialmente os da Classe I – LTR Retrotransposons podem exercer sobre o genoma do hospedeiro, pode-se dizer que o presente trabalho, obteve resultados inéditos, permitindo o avanço no entendimento da biologia e melhor compreensão do genoma desse fitopatógeno, que poderá contribuir futuramente para o controle da ferrugem, como o direcionamento de programas de melhoramento de *Eucalyptus* spp, resistentes à *P. psidii*, sendo essa uma valiosa contribuição para o setor florestal.