**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Caracterização *in silico* de sequências polimórficas na população de
Sporisorium scitamineum e análise da expressão de genes vizinhos aos
polimorfismos**

Suzane Saito

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

**Piracicaba
2015**

Suzane Saito
Bacharel e Licenciada em Ciências Biológicas

**Caracterização *in silico* de sequências polimórficas na população de *Sporisorium*
scitamineum e análise da expressão de genes vizinhos aos polimorfismos**
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora:
Profa. Dra. **CLAUDIA BARROS MONTEIRO-VITORELLO**

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**

Saito, Suzane

Caracterização *in silico* de sequências polimórficas na população de *Sporisorium scitamineum* e análise da expressão de genes vizinhos aos polimorfismos / Suzane Saito. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2015.
67 p. : il.

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

1. *Sporisorium scitamineum* 2. Carvão da cana-de-açúcar 3. Polimorfismos 4. AFLP
5. Expressão gênica I. Título

CDD 633.61
S158c

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

Aos meus pais
E a minha irmã

Dedico

Ao meu namorado e melhor amigo

Ofereço

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), ao Departamento de Genética e ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas pela formação que me proporcionaram e pela oportunidade de realizar o mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro e bolsa concedida.

À Profa. Marie Anne Van Sluys por ter oferecido auxílio na execução deste trabalho e à Tatiana Silveira Corrêa pelo auxílio no sequenciamento e por toda a paciência.

À Profa. Maria Lúcia Carneiro Vieira e a Alessandra Palhares por auxiliarem no início do trabalho.

À Patricia Daiane Shacker por ter, pronta e atenciosamente, atendido ao meu pedido de ajuda com a análise de dados.

A duas Professoras que desde o início foram meus exemplos de pessoas e de profissionais a serem seguidos: Profa. Claudia Barros Monteiro-Vitorello, que nesses anos de trabalho foi muito mais que uma orientadora; e Profa. Maria Carolina Quecine-Verdi. São pessoas com quem sempre pude contar, tanto profissionalmente quanto pessoalmente. Serei sempre grata pelos ensinamentos, amizade, carinho e preocupação. Muito, muito obrigada a vocês duas!

A TODOS do grande Laboratório de Genética de Microrganismos “João Lúcio de Azevedo” pela amizade construída nesses anos trabalhando juntos e pela mão sempre estendida. Agradeço especialmente ao Genomics Group, que se tornou uma segunda família. Sinto-me imensamente grata e honrada por ter feito parte desse grupo! A convivência com todos os que passaram por aqui desde o início da formação do grupo foi mais do que especial para mim, aprendi muito com cada um de vocês. Muito obrigada! Obrigada pela mão sempre estendida e pelo conforto que vocês sempre me ofereceram. Com certeza vou levar vocês comigo para a vida.

Um agradecimento especial para o **Guilherme Pereira de Oliveira**, com você o caminho é menos difícil, mais divertido e muito mais gostoso. Obrigada por todo o incentivo, apoio, paciência, dedicação, carinho e amor.

Aos meus grandes amigos, com quem sempre poderei contar: **Amanda Morato do Canto** e **Danilo Soares Gissi**. Obrigada por me entenderem e apoiarem!

À minha irmã, **Lia Saito**, minha mãe, **Luzia Mitiko Seino**, e meu pai, **Lauro Katsushiko Saito**, que, com muito amor, me incentivam e apoiam sempre! Muito obrigada!

Aos meus amigos de longa data, que, apesar da distância e do tempo, sempre torceram por mim: **Juliana Raposo Semeghini**, **Luiza Ferreira**, **Lívia Tieri Kuga**, **Carolina de Moura Trindade**, **Carolina Ando Matsuno**, **Natália Grego Carvalho**, **Maria Eduarda Rocha**, **Arthur Vieira Fisch** e **Guilherme Hashioka**. Muito obrigada a todos vocês!

“The important thing is not to stop questioning. Curiosity has its own reason for existence. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery each day. Never lose a holy curiosity.”

Albert Einstein

SUMÁRIO

RESUMO	11
ABSTRACT	13
1 INTRODUÇÃO.....	15
2 REVISÃO BIBLIOGRÁFICA	17
2.1 O carvão da cana-de-açúcar.....	17
2.1.1 O fungo <i>Sporisorium scitamineum</i>	18
2.1.2 Tipo de reação sexual em <i>S. scitamineum</i>	19
2.1.3 Diversidade genética de <i>S. scitamineum</i>	20
2.2 Marcadores moleculares tel-RFLP e AFLP	22
2.3 Genes e seus elementos regulatórios	23
2.3.1 Conceito de gene	23
2.3.2 Elementos regulatórios	24
3 MATERIAL E MÉTODOS.....	27
3.1 AFLP de isolados de <i>S. scitamineum</i>	27
3.1.1 Extração e digestão do DNA genômico	27
3.1.2 Preparo e ligação de adaptadores.....	28
3.1.3 Reações de amplificação	29
3.1.3.1 Pré-amplificação.....	29
3.1.3.2 Amplificação seletiva	30
3.1.4 Eletroforese em gel de poliacrilamida desnaturante.....	30
3.1.4.1 Separação dos <i>loci</i> amplificados.....	30
3.1.4.2 Coloração e visualização dos <i>loci</i>	30
3.2 Reamplificação de fragmentos de DNA recuperados do gel de AFLP	31
3.2.1 Pré-amplificação.....	31
3.2.2 Reamplificação	32
3.3 Sequenciamento e análise <i>in silico</i>	32
3.3.1 Contexto genômico e análise de genes próximos aos polimorfismos	32
3.3.2 Estudo da expressão gênica	33
4 RESULTADOS	35
4.1 Reamplificação de fragmentos de DNA recuperados do gel de AFLP	35
4.2 Contexto genômico dos polimorfismos encontrados por AFLP	36
4.3 Análise da expressão gênica em três condições diferentes.....	46
5 DISCUSSÃO	49
5.1 Considerações Gerais	49
5.2 Genes mais expressos <i>in planta</i>	49
5.3 Genes mais expressos em 5 DAI.....	51

5.4 Co-regulação de genes adjacentes.....	52
5.5 Há a formação de raças em <i>S. scitamineum</i> ?	52
6 CONSIDERAÇÕES FINAIS	53
REFERÊNCIAS	55
ANEXO.....	63

RESUMO

Caracterização *in silico* de sequências polimórficas na população de *Sporisorium scitamineum* e análise da expressão de genes vizinhos aos polimorfismos

O fungo *Sporisorium scitamineum* é o agente causador de uma das doenças mais importantes da cultura da cana-de-açúcar, o carvão da cana, e vem sendo alvo de diversos estudos genéticos para a elucidação de seu mecanismo de patogenicidade e interação com o seu hospedeiro. Dentre esses estudos há os que visam acessar a variabilidade genética entre isolados desse fungo utilizando diversos marcadores moleculares, porém todos os experimentos indicaram baixa variabilidade, com exceção de isolados provenientes da Ásia. O primeiro estudo realizado com esse intuito em isolados brasileiros utilizou o tel-RFLP e AFLP, obtendo uma variabilidade de 34 e 3%, respectivamente. O AFLP é muito utilizado para a detecção de variação genética em indivíduos próximos, porém, nesse caso, apresentou baixa variação, o que nos levou a caracterizar, no presente trabalho, os poucos polimorfismos encontrados através dessa técnica. Os fragmentos polimórficos foram reamplificados, sequenciados e caracterizados quanto ao seu contexto genômico. Como essas sequências estão dispersas ao longo do genoma e geralmente estão em regiões intergênicas, foi feita a análise de expressão dos genes vizinhos a esses polimorfismos em três condições distintas. Desses genes, 70% apresentaram maior número de *reads* mapeadas *in planta* que em meio de cultura, indicando que são mais importantes durante a interação planta-patógeno. Dessa forma, mesmo havendo uma baixa variabilidade detectada por AFLP, essas diferenças podem ser importantes na regulação desses genes, podendo influenciar na agressividade do patógeno e na sua defesa.

Palavras-chave: *Sporisorium scitamineum*; Carvão da cana-de-açúcar; Polimorfismos; AFLP; Expressão gênica

ABSTRACT

***In silico* characterization of polymorphic sequences in *Sporisorium scitamineum* population and expression analysis of polymorphisms neighbour genes**

The fungi *Sporisorium scitamineum* is the causal agent of one of the most important diseases in sugarcane, the sugarcane smut, and it has been the target of several genetic studies aiming the elucidation of pathogenicity and host interaction mechanisms. Among these studies are those that aim to access the genetic variability among isolates of this fungus using various molecular markers techniques, however in all of them the experiments indicated low variability, with exception of the Asian isolates. The first work performed with Brazilian isolates was done using tel-RFLP and AFLP, and detected a variability of 34 and 3%, respectively. The AFLP is widely used to detect variability in close related individuals, but in this case, the method revealed low polymorphism. These results led us to characterize in the present work the few polymorphisms detected. The polymorphic fragments were isolated, reamplified, sequenced and *in silico* characterized regarding the genomic context. Since these sequences are dispersed along the genome and generally they are in intergenic regions, an expression analysis of the polymorphisms neighbor genes was made at three different conditions. Of these genes, 70% presented higher numbers of mapped reads *in planta* than in culture media, which indicates that they are important during the plant-pathogen interaction. Thus, even with low variability detected via AFLP, these differences can be important in the gene regulation and may influence in the pathogen aggressivity and in its defense.

Keywords: *Sporisorium scitamineum*; Sugarcane smut; Polymorphisms; AFLP; Gene expression

1 INTRODUÇÃO

O fungo basidiomiceto *Sporisorium scitamineum* causa uma das doenças mais importantes da cultura da cana-de-açúcar, o carvão da cana, e provavelmente teve a sua origem na Ásia, local onde possui a sua maior variabilidade genética. Vários trabalhos já tentaram acessar a variabilidade genética de isolados desse fungo, porém, mesmo utilizando diversas técnicas de marcadores moleculares, encontrou-se uma baixa variação entre os isolados, com exceção daqueles provenientes da Ásia. E, apesar da uniformidade genética encontrada entre os isolados, foi descrito um comportamento diferencial de cultivares de cana-de-açúcar em resposta a diferentes isolados do patógeno, porém ainda não há consenso sobre a formação de raças nessa espécie.

Outro trabalho foi realizado para acessar a variabilidade genética entre isolados brasileiros desse fungo utilizando o AFLP (*Amplified Fragment Length Polymorphism*) e o tel-RFLP (*Telomere-associated Restriction Fragment Length Polymorphism*). Esse trabalho obteve pouca variação utilizando o AFLP, apenas 3%, porém obteve-se maior variação utilizando o tel-RFLP, 34% de variação. Os dados não apresentaram agrupamento por localidade, mas ficou claro que o fluxo desses isolados é baixo e que os isolados mais distantes do estado de São Paulo foram provavelmente levados a partir de isolados desse estado.

A técnica de AFLP já foi muito utilizada para se distinguir indivíduos geneticamente próximos, o que vai de encontro com os resultados obtidos no trabalho utilizando isolados brasileiros. Além disso, essa técnica permite a detecção de marcadores moleculares dispersos pelo genoma, diferente do tel-RFLP. Por isso, os fragmentos polimórficos obtidos através do AFLP foram utilizados para confecção do presente trabalho, que teve como objetivo: (1) a reamplificação de fragmentos de DNA polimórficos entre os diferentes isolados de *S. scitamineum*; (2) caracterizar esses polimorfismos quanto ao seu contexto genômico e (3) analisar a expressão de genes próximos aos polimorfismos em três diferentes condições, já que os fragmentos estão, em sua maioria, em regiões intergênicas.

cana plantada e da severidade da infestação, podendo chegar a 100% de perda (FIGUEIREDO, 2008).

Figura 2 - Diferentes morfologias do chicote em cana-de-açúcar infectada com *Sporisorium scitamineum*. a - chicote longo; b - chicote fechado; c - chicote torcido; d - chicote curto; e - chicotes múltiplos (SUNDAR et al., 2012)

2.1.1 O fungo *Sporisorium scitamineum*

S. scitamineum é um fungo biotrófico, parasita de tecidos meristemáticos, pertencente ao filo Basidiomycota, classe Ustilaginomycetes, ordem Ustilaginales e família Ustilaginaceae (HAUKSWORTH et al., 1995). Inicialmente, o fungo era conhecido como *Ustilago scitaminea*, mas foi removido do gênero *Ustilago* e realocado para o *Sporisorium*, entre outros motivos, porque foram encontradas diferenças nas estruturas reprodutivas formadas durante o seu ciclo de vida (PIEPENBRIG et al., 2002; STOLL et al., 2003).

Essa espécie possui ciclo de vida dimórfico, com uma fase haploide (leveduriforme) e uma fase de hifa dicariótica. O seu ciclo (Figura 3) inicia-se com a germinação dos teliósporos ($2n$), em seguida, ocorre a meiose, formando os esporídeos (n). Esporídeos sexualmente compatíveis se fundem e formam a hifa dicariótica infectiva. O desenvolvimento da hifa dicariótica no tecido vegetal resulta na formação de um apressório para a invasão dos tecidos meristemáticos da planta. O fungo se desenvolve dentro da planta ocupando tecidos associados aos feixes vasculares periféricos (LLOYD et al., 1980), manipulando a fisiologia do hospedeiro para a realização da esporogênese que associada a formação do chicote resulta na liberação dos teliósporos. O ciclo sexuado só se completa na presença de interação com a planta hospedeira (ALBERT e SHENCK, 1996; KRONSTAD e LEONG, 1990).

Figura 3 - Ciclo de vida do fungo *S. scitamineum*. A cana-de-açúcar emite o chicote, liberando os teliósporos ($2n$), que quando germinam emitem o pró-basídio que sofre meiose (R!), formando os esporídios (n). Quando esporídios de tipos de reação sexual diferentes se encontram e se fundem, forma-se a hifa dicariótica infectiva ($n + n$) (TANIGUTI et al., 2015)

Nos últimos anos, vem se buscando um melhor entendimento desse patossistema, nesse sentido vários estudos vêm sendo conduzidos desde a análise de marcadores genéticos para resistência, sequenciamento do genoma completo, e análises de transcriptoma na ausência e durante a interação entre a planta e o patógeno (HEINZE et al., 2001; THOKOANE e RUTHERFORD, 2001; BUTERFIELD et al., 2004; BÓRRAS-HIDALGO et al., 2005; LAO et al., 2008; QUE et al., 2011; PALHARES, 2014; QUE et al., 2014; TANIGUTI et al., 2015).

2.1.2 Tipo de reação sexual em *S. scitamineum*

O tipo de reação sexual (*mating type*) em fungos causadores de carvão é regulado por dois *loci*, *a* e *b*, que carregam genes conservados. No *locus a* estão os genes que codificam para os ferormônios e para os receptores de ferormônio, enquanto que no *locus b*, os genes codificam as duas subunidades de um fator de transcrição heterodimérico (BAKKEREN et al., 2008). Os dois *loci* que controlam o *mating type* podem estar ligados em um mesmo cromossomo e segregam com uma unidade, dando origem ao sistema bipolar de *Ustilago hordei* e *S. scitamineum* (SCHIRAWSKI et al., 2005), ou estarem separados e segregarem

independentemente, sendo esse o sistema tetrapolar encontrado em *S. reilianum* e *U. maydis* (QUE et al., 2014; LAURIE et al., 2012). Em *S. scitamineum*, os dois *loci* estão localizados no cromossomo 2 a 59 Kpb de distância um do outro (TANIGUTI et al., 2015).

Dentre os genes que compõe o *locus a* (Figura 4A) estão os genes necessários para codificar o sistema ferormônio-receptor, como as proteínas *Mfa* (ferormônio lipopeptídeo) e *pra* (receptor transmembrana). Ferormônios produzidos por esporídeos são reconhecidos por receptores de membrana compatíveis (SNETSELAAR e MIMS, 1992). Além desses genes que codificam para o ferormônio e o receptor, existem outros genes que ainda não possuem função conhecida (BORTFELD et al., 2004).

O *locus b* (Figura 4B) é composto por dois genes, cada um codifica para uma subunidade de um fator de transcrição heterodimérico que controla a expressão de genes responsáveis pela manutenção do crescimento da hifa dicariótica na planta (QUE et al., 2014). Para que o fator de transcrição seja funcional, é necessário que as subunidades que o formam sejam provenientes de alelos distintos do *locus b* (KRONSTAD e STABEN, 1997).

Figura 4 - Localização dos genes que compõe o *locus a* (A) e o *locus b* (B) que compõem o sistema de *mating type* do fungo *Sporisorium scitamineum*

2.1.3 Diversidade genética de *S. scitamineum*

O fungo *S. scitamineum*, provavelmente, teve a sua origem na Ásia, onde ocorre a maior diversidade entre os isolados obtidos no mundo, inclusive no Brasil (RABOIN et al., 2007). Várias técnicas foram utilizadas para acessar a variabilidade genética entre os isolados de *S. scitamineum* como estudos morfológicos e de inoculação em casa de vegetação (GILLASPIE et al., 1983), análise de sequência de nucleotídeos das regiões ITS (SINGH et al., 2005), RAPD (XU et al., 2004), AFLP (BRAITHWAITE et al., 2004) e microssatélites (RABOIN et al., 2007). Todos esses experimentos indicaram baixa variabilidade genética entre os isolados, com exceção dos isolados provenientes da Ásia (BRAITHWAITE et al., 2004).

Apesar da uniformidade genética encontrada entre os isolados, foi descrito um comportamento diferencial de cultivares de cana-de-açúcar em resposta a diferentes linhagens do patógeno, o que levou ao registro de duas raças de *S. scitamineum* no Havaí (COMSTOCK e HEINZ, 1977). Em Taiwan, Paquistão, Filipinas e, inclusive, no Brasil, também foram descritas diferentes raças para esse patógeno, porém, em nenhum dos trabalhos houve uma uniformidade nas técnicas para a caracterização dessas raças (LEU et al., 1976; FERREIRA e COMSTOCK, 1989; COMSTOCK, 2000). No Estado de São Paulo, diferentes populações de *S. scitamineum* foram analisadas com base na sua agressividade, esse trabalho indicou diferenças entre as populações (RAGO et al., 2009).

A variação de agressividade entre as populações como descrita em RAGO et al. (2009) é um forte indicativo de que existe uma variabilidade genética entre isolados, ainda não descrita. REIS (2012) apresentou um trabalho no qual uma coleção de 14 isolados de *S. scitamineum* coletados em diferentes regiões canaveiras do Brasil (Figura 5) foram analisados utilizando duas metodologias: tel-RFLP (*Telomere-associated Restriction Fragment Length Polymorphism*) e AFLP (*Amplified Fragment Length Polymorphism*) de maneira a se comparar as duas técnicas.

Figura 5 - Pontos de coleta dos isolados obtidos por REIS (2012) (MONTEIRO-VITORELLO et al., 2013)

Os resultados de REIS (2012) demonstraram que a variação genética pode ser mais facilmente acessada utilizando tel-RFLP do que através de AFLP. Utilizando a técnica de AFLP foram encontrados 40 *loci* polimórficos (3%) entre 1.311 analisados a partir da combinação de 2 enzimas de corte raro, 1 de corte frequente e 17 combinações de *primers*. Para o tel-RFLP, 3 enzimas de restrição e sondas construídas a partir de sequência telomérica

gerou 102 *loci*, sendo 35 polimórficos (34,3%). O que indica que a variação genética acessada está principalmente ligada a regiões subteloméricas.

Os dados não apresentaram agrupamento por localidade, mas ficou claro que o fluxo desses isolados é baixo e que os isolados mais distantes do estado de São Paulo foram provavelmente levados a partir de isolados desse estado. Os derivados haplóides de tipos de reação sexual opostos de cada linhagem dicariótica (teliósporo) permaneceram dentro do mesmo grupo, com exceção de uma única linhagem haplóide. Esse resultado indica que, geralmente, a fusão de esporódeos acontece entre aqueles originados do mesmo teliósporo, como descrito anteriormente em RABOIN et al. (2007).

2.2 Marcadores moleculares tel-RFLP e AFLP

Marcadores moleculares são *loci* cromossômicos, genes ou sequências de DNA que apresentam segregação mendeliana e podem ter ação dominante ou codominante (VIEIRA et al., 2004). Além disso, os marcadores são *loci* representativos das diferenças que ocorrem em nível genômico, podendo ou não ser correlacionados ao fenótipo apresentado pelo organismo (AGARWAL et al., 2008).

Não há uma única técnica que apresente todos os aspectos importantes para a detecção de marcadores moleculares, por isso, a escolha da técnica deve ser feita de forma a se considerar qual seria o principal aspecto a ser estudado e analisado para cada situação. Cada técnica pode diferir entre si em aspectos como a abundância de marcadores relevantes detectados, repetibilidade, requerimentos técnicos e custo (AGARWAL et al., 2008).

Com a técnica do RFLP, os polimorfismos são detectados através de uma sonda marcada quimicamente. A geração dos padrões de fragmentos é realizada utilizando uma endonuclease e essas diferenças são devido a substituições de nucleotídeo ou rearranjos cromossômicos. Esse é uma técnica que apresenta um número de polimorfismos relativamente alto, são marcadores codominantes e é altamente reproduzível (AGARWAL et al., 2008), fornecendo um grande número de marcadores quando utilizada para a análise de genomas (BECKMANN e SOLLER, 1983).

O tel-RFLP consiste na técnica do RFLP utilizando como sonda uma sequência que corresponde à região telomérica. Dessa forma, o padrão de bandas detectado será associado a essa região, na qual se tem maior interesse.

Já o AFLP consiste na amplificação de fragmentos gerados por duas endonucleases diferentes, uma se corte raro e outra de corte frequente, ligação de adaptadores e amplificação

seletiva de fragmentos para a formação de um padrão de bandas (VOS et al., 1995). Essa técnica gera um *fingerprint* em qualquer DNA, independente de sua fonte e sem que haja um conhecimento prévio sobre a sua sequência. Além disso, pode ser utilizada para distinguir indivíduos proximamente relacionados, por exemplo, em nível de subespécie (ALTHOFF et al., 2007), apesar de REIS (2012) não ter identificado um grande número de polimorfismos.

A técnica do AFLP é considerada a que mais se aproxima do ideal para a análise de marcadores moleculares, já que é relativamente barata, de fácil execução e gera marcadores confiáveis. A maior desvantagem da técnica é ela ser um marcador dominante, o que dificulta um estudo preciso sobre os estados alélicos (MULLER e WOLFENBARGER, 1999).

Uma das principais diferenças entre o tel-RFLP e o AFLP é que os marcadores nesse último estão dispersos pelo genoma, em diversas regiões. O presente trabalho buscou caracterizar alguns dos *loci* que se apresentaram polimórficos entre os isolados brasileiros de *S. scitamineum* detectados utilizando a técnica de AFLP (REIS, 2012; MONTEIRO-VITORELLO et al., 2013).

2.3 Genes e seus elementos regulatórios

2.3.1 Conceito de gene

A visão de como se define um gene foi se modificando ao longo da história, conforme novas descobertas no campo da genética foram feitas. Para os primeiros geneticistas o gene era considerado um *locus* com tamanho determinado por mutações que ativavam ou desativavam uma característica de interesse. Posteriormente, foi estabelecida a regra “um gene, uma enzima” o qual depois se tornou “um gene, um polipeptídio”. Em 1953, com WATSON e CRICK, foi esclarecido como o DNA poderia funcionar como a molécula da hereditariedade. Desde então a biologia molecular avançou rapidamente e descobriu-se que alguns genes podem codificar para moléculas de RNA funcional e não para proteínas (GERSTEIN et al., 2007).

Hoje, “gene” é definido como “região localizável de sequência genômica, a qual corresponde a uma unidade de herança e que está associada a regiões regulatórias, regiões transcritas e/ou outras regiões funcionais” (PEARSON, 2006). Porém, essa definição ainda possui pontos problemáticos em relação a descobertas que estão sendo feitas pelo projeto ENCODE (THE ENCODE PROJECT CONSORTIUM, 2004), como por exemplo, (1) sequências de DNA não anotadas e que são transcritas em RNA; (2) elementos regulatórios

que estão dispersos no genoma, podendo estar dentro da ORF (*Open Reading Frame*) do gene ou mesmo em outro cromossomo; (3) TSSs (*Transcription Start Site*) não anotados (GERSTEIN et al., 2007; ZHANG et al., 2007).

2.3.2 Elementos regulatórios

A complexidade e a quantidade de transcritos que são gerados a partir de um único “gene” depende inteiramente de elementos regulatórios da expressão gênica. Esses elementos regulatórios podem atuar, em eucariotos, no núcleo ou no citoplasma (Figura 6), em diferentes etapas da expressão gênica, sendo em nível de DNA, RNA ou polipeptídeos (SNUSTAD e SIMMONS, 2006).

Figura 6 - Esquema apontando as diversas etapas de regulação da expressão gênica em eucariotos. Modificado de ALBERTS et al. (2008)

A regulação da expressão gênica é imprescindível para que o organismo possa ter uma resposta adaptativa devido à fase do seu ciclo de vida ou ao ambiente em que está inserido. As células de um indivíduo necessitam modificar a sua regulação o tempo todo para que as proteínas necessárias naquele momento e para aquela situação sejam sintetizadas e ativadas. Por isso há uma gama de possibilidades de regulação para a expressão gênica, assim a célula pode obter tanto uma resposta rápida, quanto uma mais lenta para se adequar a determinada situação.

Em situações de estresse, por exemplo, o organismo necessita de uma resposta imediata, além de uma resposta específica para a mudança que está ocorrendo no ambiente. As modificações pós-traducionais ocorrem para que a resposta seja a mais imediata possível,

e a regulação da transcrição será essencial para a resposta mais lenta, de adaptação longa e para fases de recuperação (DE NADAL et al., 2011).

Não necessitando de uma resposta imediata, para a maioria dos genes, o passo mais importante é o controle transcricional, assim se garante que não haverá gasto energético produzindo intermediários supérfluos para a expressão desses genes (ALBERTS et al., 2008). Os mecanismos para a regulação da transcrição em eucariotos tem sido estudado há muitos anos, como resultado, alguns grupos de elementos regulatórios que afetam a transcrição já foram descritos, como os promotores, *enhancers*, silenciadores e *insulators* (revisado por RAZIN et al., 2015), porém, ainda há muito que ser descrito quanto ao mecanismo de ação e a interação entre esses grupos.

Os promotores são sequências que estão imediatamente *upstream* ao gene que regulam, eles são o sítio preferencial de ligação da RNA polimerase II no DNA. Os promotores canônicos possuem o TATA-box, uma sequência rica em A/T, na qual proteínas TBP (*TATA-binding protein*) se ligam para a montagem do complexo de pré-iniciação da polimerase II (SMALE e KADONAGA, 2003). Porém, o recrutamento de proteínas TBP necessita de outros fatores de transcrição. Há promotores que não possuem o TATA-box, esses ocorrem em ilhas CpG e, geralmente, controlam gene *housekeeping* (RAZIN et al., 2015).

Há grupos de elementos regulatórios que estão mais distantes do gene que será regulado, como os *enhancers* que ativam a transcrição e os silenciadores que a suprimem. Os *enhancers* não possuem uma sequência consenso típica, mas é associada a regiões em que se ligam fatores de transcrição e facilita a ligação da RNA polimerase II ao promotor gênico. A sua localização pode ser *upstream*, *downstream* ou dentro de um gene, e a sua posição pode ser mais próxima a um gene que não é regulado por esse mesmo *enhancer*. Mas apesar da regulação ocorrer à distância, a conformação cromossômica os mantém espacialmente próximos dos promotores aos quais regulam (DE LAAT e GROSVELD 2003; RAZIN et al., 2015).

Os silenciadores são o grupo menos definido dos elementos regulatórios. Eles foram inicialmente descritos como tendo o mecanismo oposto aos *enhancers* (BRAND et al., 1985), porém, hoje, sabe-se que o seu mecanismo é diferente, funcionando em nível de reconfiguração da cromatina e não necessitam estar próximo aos promotores aos quais regulam, como os *enhancers*.

Já os *insulators* fazem parte dos elementos de arquitetura, que facilitam a comunicação entre regiões distantes do genoma e facilitam a formação de *loops* estáveis no

genoma. Dependendo das condições, os *insulators* podem desempenhar duas funções, prevenir a comunicação de *enhancers* com os promotores (*enhancer-blocking activity*) e/ou prevenir que a heterocromatina se espalhe (*barrier activity*). Essas duas funções podem estar presente em um mesmo *insulator* ou podem estar presentes de forma separada. O mecanismo de como essas funções são desempenhadas por *insulators* ainda está em aberto e necessita ser descrita (RAZIN et al., 2015).

A Figura 7 ilustra como esses elementos regulatórios podem agir sobre a transcrição de um determinado gene e como pode ocorrer a interação entre eles.

Figura 7 - Regulação gênica típica de um eucarioto: região promotora da transcrição e a sua interação com outros elementos regulatórios. Retirado de ALBERTS et al. (2008)

3 MATERIAL E MÉTODOS

3.1 AFLP de isolados de *S. scitamineum*

Previamente, uma coleção de 25 isolados brasileiros de *S. scitamineum* (Tabela 1), provenientes de 13 chicotes coletados em 4 estados brasileiros, foi utilizada para a detecção de variação molecular por AFLP (REIS, 2012). Essa técnica tem como base a amplificação por PCR de um conjunto de fragmentos gerados a partir da digestão de DNA genômico com combinações de duas enzimas de restrição, uma de corte raro e outra de corte frequente gerando um padrão de bandas em um gel de poliacrilamida (MONTEIRO-VITORELLO et al., 2013).

Tabela 1 - Amostras isoladas por REIS (2012) a partir de chicotes coletados de plantas infectadas de cana-de-açúcar. A identificação para cada isolado foi dada em laboratório.

Identificação	Local	Variedade
Ssc04 A e B	Barreiras - BA	IAC91 5155
Ssc05 A e B	Frutal - MG	SP79 1011
Ssc11 A e B	Jaboticabal - SP	IAC91 1099
Ssc17 A e B	Macatuba - SP	CTC12
Ssc18 A e B	Paranaíba - MS	IACSP94 4004
Ssc31 A e B	Serra Azul - SP	IACSP04 4084
Ssc33 B	Vista Alegre do Alto - SP	RB855453
Ssc35 A e B	Limeira D'Oeste - MG	IACSP93 3046
Ssc36 A e B	Valparaíso - SP	RB72454
Ssc38 A e B	Guaíra - SP	SP90 3414
Ssc39A e B	Ribeirão Preto - SP	IAC98 2053
Ssc40 A e B	Araçatuba - SP	-
Ssc41 A e B	Conchal - SP	-

3.1.1 Extração e digestão do DNA genômico

Foram utilizados inóculos de células leveduriformes haplóides para o crescimento em meio cultura líquido YEDP (3% (massa/volume) extrato de levedura, 1% peptona, 1% glicose), em uma incubadora sob agitação de 200 rpm, a 30 °C, por 24 h.

As células foram transferidas para um tubo de 50 mL e centrifugado por 5 min a 10.000 rpm. O sobrenadante foi descartado e o precipitado, liofilizado por 24 h para posterior

maceração com nitrogênio líquido. Foram acrescentados 6 mL de solução de lise (1% SDS; 0,15 M NaCl, 60 mM EDTA pH 8,0 e 400 mM Tris-HCl) e 600 µL de 2-mercaptoetanol e a amostra foi homogeneizada e, em seguida, transferida para microtubos de 1,5 mL. Adicionaram-se 600 µL de Clorofane (25 mL fenol, 24 mL clorofórmio e 1 mL de álcool isoamílico) às amostras, que foram homogeneizadas por inversão durante 1 min. Em seguida, centrifugou-se a 10.000 rpm por 12 min para posterior transferência do sobrenadante para um novo microtubo. Adicionou-se o mesmo volume de uma solução de Clorofil (24 mL de clorofórmio e 1 mL de álcool isoamílico) homogeneizando por inversão. A mistura foi centrifugada a 10.000 rpm por 12 min e a sua fração superior foi transferida para um novo microtubo. Adicionou-se 1 volume de etanol gelado homogenando por inversão. A amostra foi centrifugada novamente a 10.000 rpm por 1 min, descartou-se o sobrenadante e adicionou-se 500 µL de etanol 70% gelado. Uma nova centrifugação foi realizada a 10.000 rpm por 1 min, descartando-se o sobrenadante. Esperou-se o álcool evaporar para que o precipitado fosse ressuspensionado em 100 µL de TE acrescido de 3 µL de RNase (na concentração 10 mg/mL).

O DNA genômico foi digerido com *MseI* como enzima de restrição de corte frequente e *EcoRI/PstI* como de corte raro. Para a digestão, foram utilizados 400 ng DNA genômico com 6 U de enzima de restrição em tampão 1X da enzima, 1X BSA (New England Bio Labs®) e água milli-Q esterilizada para completar o volume da reação para 30 µL. O material foi incubado no termociclador Gen Amp® PCR System 9700 Thermocycler (Applied Biosystems) por 12 h a 37°C e, em seguida, 20 min a 65°C para a inativação da enzima. Após a digestão, 5 µL da amostra foram submetidos à eletroforese em gel de agarose 1%, TBE 1X, 60 V, por 3 h.

3.1.2 Preparo e ligação de adaptadores

O preparo dos adaptadores complementares aos fragmentos clivados foi manipulado conforme a endonuclease utilizada:

Adaptadores *EcoRI/PstI*: Consiste na preparação de uma solução contendo 5 µM de adaptador *forward*, 5 µM de adaptador *reverse* e água Milli-Q esterilizada para completar volume 200 µL.

Adaptador *MseI*: Uma solução foi preparada contendo 50 µM de adaptador *forward*, 50 µM de adaptador *reverse*, e água milli-Q esterilizada para completar volume 200 µL.

Para o preparo dos adaptadores foi utilizada as seguintes condições: 65°C por 10 min, 37°C por 10 min e, por fim, 25°C por 10 min. A solução foi estocada a -20°C.

A ligação dos adaptadores foi realizada acrescentando-se ao material digerido uma solução de 30 µL contendo 0,25 µM de adaptadores da enzima de corte raro e 2,5 µM de adaptadores da enzima de corte frequente, tampão 1X para a enzima T4 DNA ligase, 1,7 U de enzima T4 DNA ligase (New England Bio Labs®) e água Milli-Q esterilizada para completar volume da reação. O material foi incubado por 14 h a 16°C para a ligação e a 65°C por 20 min para a inativação da enzima. O produto de ligação foi diluído de modo a padronizar sua concentração final a 5 ng/µL. A reação foi estocada a -20°C.

3.1.3 Reações de amplificação

3.1.3.1 Pré-amplificação

Tabela 2 - Sítios de restrição, sequências de adaptadores e de iniciadores usados para as análises de AFLP

Enzima	Sítio de Restrição	Especificação	Sequência
EcoRI	5'-GAATTC-3'	Adaptador <i>forward</i>	5'-CTC GTA GAC TGC GTA CC-3'
	3'-CTTAAG-5'	Adaptador <i>reverse</i>	5' -AAT TGG TAC GCA GTC TAC-3'
		Iniciador E+A	5'-GAC TGC GTA CCA ATT CA-3'
		Iniciador E+ANN	5' - GAC TGC GTA CCA ATT CAN N-3'
PstI	5'-CTGCAG-3'	Adaptador <i>forward</i>	5'-CTC GTA GAC TGC GTA CAT GCA-3'
	3'-GACGTC-5'	Adaptador <i>reverse</i>	5'-TGT ACG CAG TCT AC- 3'
		Iniciador P+A	5'-GAC TGC GTA CAT GCA GA-3'
		Iniciador P+ANN	5'-GAC TGC GTA CAT GCA GAN N-3'
MseI	5'-TTAA-3'	Adaptador <i>forward</i>	5'-GAC GAT GAG TCC TGA G-3'
	3'-AATT -5'	Adaptador <i>reverse</i>	5'-TAC TCA GGA CTC AT-3'
		Iniciador M+C	5'-GAT GAG TCC TGA GTA AC-3'
		Inciador M +CNN	5'-GAT GAG TCC TGA GTA ACN N-3'

A pré-amplificação foi realizada utilizando para cada reação, um volume final de 20 µL, onde foi acrescentado 1,5 mM MgCl₂; 0,5 mM dNTPs; 300 nM de cada iniciador complementar ao adaptador onde foi adicionado um nucleotídeo seletivo a extremidade 3': oligo *EcoRI*+A ou *PstI*+A e oligo *MseI*+C, apresentados na Tabela 2; tampão 1X da *Taq* DNA polimerase; 3 U *Taq* polimerase (Fermentas™); 3 µL (15 ng) do produto de ligação e água Milli-Q esterilizada para completar o volume final. As condições para a amplificação

por PCR foram: desnaturação inicial de 2 min a 94°C; seguido de 26 ciclos de 1 min a 94°C, 1 min a 56°C, 1 min a 72°C; e uma extensão final de 5 min a 72°C. A reação de pré-amplificação foi diluída em 5X em água Milli-Q esterilizada e armazenada a -20°C.

3.1.3.2 Amplificação seletiva

Para cada reação de amplificação seletiva foram adicionados na solução 1,5 mM MgCl₂; 0,2 mM dNTPs; 250 nM do *primer* *EcoRI*+ANN ou *PstI* +ANN (Tabela 2) e 300 nM do iniciador *MseI*+CNN (Tabela 2), tampão 1X da *Taq* DNA polimerase; 3 U *Taq* polimerase (Fermentas TM); 1,5 µL (5 ng) de reação de pré-amplificação diluída e água Milli-Q esterilizada para completar o volume final de 20 µL. As condições para PCR foram: desnaturação inicial 94°C por 2 min; seguida de 12 ciclos a 94°C por 30 s, 65°C por 30s e 72°C por 1 min; 23 ciclos nas mesmas condições, apenas modificando a temperatura de anelamento para 56°C; e extensão final de 3 min a 72°C. Ao final, a reação foi estocada a -20°C.

3.1.4 Eletroforese em gel de poliacrilamida desnaturante

3.1.4.1 Separação dos *loci* amplificados

Aos produtos da amplificação foi adicionado o tampão desnaturante (98% de formamida; 10 mM de EDTA pH 8.0; 0,002 % de azul bromofenol e 0,002% de xileno cianol). Esta mistura foi submetida à desnaturação a 94°C por 5 min e em seguida mantida em gelo. Uma alíquota de 3 µL foi aplicada em gel de poliacrilamida 5% (acrilamida/bisacrilamida [19:1]; 7,5 M de uréia e tampão TBE 1X [89 mM de Tris-base, 89 mM de ácido bórico; 2,5 mM EDTA 2Na]) após 1 h de pré-corrida a 70 W. A eletroforese foi realizada no aparelho Sequi-Gen® (Bio-Rad) por 4 h a 70 W. Como marcador de peso molecular utilizou-se 2 µL de uma mistura dos marcadores (Invitrogen) de peso molecular de 25 e 100 pb (cada um a 40 ng/µL na mistura).

3.1.4.2 Coloração e visualização dos *loci*

A coloração foi realizada em solução de nitrato de prata (AgNO₃), seguindo protocolo adaptado de CRESTE et al. (2001):

Fixação: A fixação do gel foi realizada por imersão em solução de etanol 10% e ácido acético 1%, agitando por 10 min, seguida de uma lavagem com água destilada e deionizada, sob agitação durante 1 min.

Oxidação: O pré-tratamento foi realizado em solução de ácido nítrico 1,5%, sob agitação durante 3 min, seguindo-se de uma nova lavagem com água destilada e deionizada durante 1 min.

Impregnação com prata: A impregnação do gel se deu em solução de nitrato de prata (AgNO_3) 0,2% sob agitação durante 20 min, na ausência de luz. Foi seguida de duas novas lavagens com água destilada e deionizada, com duração de 30 s cada.

Revelação: A revelação dos géis foi em uma solução de 2 litros contendo 60 g de Na_2CO_3 e 2 mL de formaldeído (37%). Nesta etapa, 1 litro da solução foi aplicado sobre o gel até que comecem a se formar as bandas. Após este passo, a solução foi descartada e o restante da solução foi adicionado para finalizar o processo de revelação.

Bloqueio: O bloqueio da revelação foi realizado em solução de ácido acético 5% durante 5 min, seguindo-se uma nova lavagem com água destilada e deionizada durante 1 min. Após estas etapas, a placa contendo o gel foi colocada na posição vertical em um local seco, arejado e à temperatura ambiente para sua secagem. O gel foi digitalizado pelo scanner (Color Page-Genius®).

3.2 Reamplificação de fragmentos de DNA recuperados do gel de AFLP

As bandas polimórficas foram retiradas diretamente do gel de AFLP preparado por REIS (2012) e acondicionadas em TE a 4°C por 16 h para que o DNA presente no gel de poliacrilamida se difundisse na solução. A mistura foi centrifugada a 10.000 rpm por 1 min, recuperando o sobrenadante para um novo tubo e utilizando-o para as próximas etapas de pré-amplificação e reamplificação.

3.2.1 Pré-amplificação

Para a pré-amplificação, foram utilizados os iniciadores que deram origem aos fragmentos polimórficos, *EcoRI*+A ou *PstI*+A e *MseI*+C (Tabela 2). O volume final de cada reação foi de 20 µL, no qual foi adicionado 1,5 mM MgCl_2 ; 2 mM dNTPs; 2,5 µM de cada iniciador; tampão 1X da *Taq* DNA polimerase; 3 U *Taq* polimerase (Fermentas™); 3 µL do DNA obtido do gel de poliacrilamida e água Milli-Q esterilizada para completar o volume

final. As condições para a amplificação por PCR foram: desnaturação inicial de 2 min a 94°C; seguido de 26 ciclos de 1 min a 94°C, 1 min a 56°C, 1 min a 72°C; e uma extensão final de 5 min a 72°C. A reação de pré-amplificação foi diluída em 5X em água Milli-Q esterilizada e armazenada a -20°C.

3.2.2 Reamplificação

A reamplificação foi realizada utilizando os mesmos iniciadores da pré-amplificação. O volume final de cada reação foi de 20 µL, no qual foi adicionado 1,5 mM MgCl₂; 0,2 mM dNTPs; 2,5 µM de cada iniciador; tampão 1X da *Taq* DNA polimerase; 1,5 U *Taq* polimerase (Fermentas™); 1,5 µL da reação de pré-amplificação diluída em 5X e água Milli-Q esterilizada para completar o volume final. As condições para a amplificação por PCR foram: desnaturação inicial de 2 min a 94°C; seguido de 12 ciclos de 30 s a 94°C, 30 s a 65°C, 1 min a 72°C; mais 23 ciclos de 30 s a 94°C, 30 s a 56°C, 1 min a 72°C; e uma extensão final de 2 min a 72°C. As reações de reamplificação foram analisadas em gel de agarose 1,5% e armazenadas a -20°C.

3.3 Sequenciamento e análise *in silico*

Para o sequenciamento, foi feita a purificação dos produtos de PCR utilizando GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) seguindo as recomendações do fabricante.

O sequenciamento de todas as amostras foi realizado utilizando os *primers* da etapa de pré-amplificação (Tabela 2) pelo método de Sanger, no Laboratório de Genômica e Elementos de Transposição (GaTE) sob a coordenação da Profa. Dra. Marie-Anne Van Sluys. E as análises *in silico* foram realizadas em banco de dados local, no software CLC Genomics Workbench (CLCbio) com o genoma completo de *S. scitamineum* Ssc39B e dados de RNAseq (TANIGUTI et al., 2015; PALHARES, 2014).

3.3.1 Contexto genômico e análise de genes próximos aos polimorfismos

Para a análise dos genes localizados próximos aos polimorfismos identificados por AFLP, foram realizados BLASTx (ALTSCHUL et al., 1990) nos bancos de dados públicos

como o NCBI¹ e o UniProt², foi feita a identificação de domínios em proteínas preditas através do InterPro (MITCHELL et al., 2015) e foram analisados a presença de peptídeo sinal com o programa SignalP (PETERSEN et al., 2011) e a presença de hélices transmembrana com o programa TMHMM (KROGH et al., 2001).

3.3.2 Estudo da expressão gênica

Para o estudo da expressão gênica foi feito no CLC Genomics Workbench (CLCbio) o mapeamento das *reads* de fungo obtidas através de RNAseq (TANIGUTI et al., 2015; PALHARES, 2014) utilizando como referência as sequências das CDS dos genes analisados. Foram obtidos *heatmaps* através da ferramenta R (R CORE TEAM, 2014) utilizando os valores normalizados de *reads* mapeadas em cada gene. As condições consideradas no estudo foram em meio de cultura, 5 dias após a inoculação (DAI) em cana-de-açúcar e 200 DAI (TANIGUTI et al., 2015; PALHARES, 2014).

¹ <http://blast.ncbi.nlm.nih.gov/>

² <http://www.uniprot.org/>

4 RESULTADOS

4.1 Reamplificação de fragmentos de DNA recuperados do gel de AFLP

No total, foram 45 bandas polimórficas retiradas do gel de AFLP (ANEXO 1). Sendo que 9 delas foram obtidas utilizando a combinação de iniciadores E+A e M+C e 36 com as combinações P+A e M+C.

A pré-amplificação foi realizada como descrito no item 3.2.1, porém não foi possível visualizar em gel de agarose nenhum fragmento de DNA para nenhuma das amostras, provavelmente pela reduzida quantidade de DNA presente na reação.

Seguiu-se com a reamplificação, como descrito no item 3.2.2, porém, poucas amostras apresentaram fragmentos amplificados em gel de agarose (Figura 8). Em seguida, foi realizada nova amplificação, dessa vez, utilizando como DNA *template* a primeira reação de reamplificação, resultando em um melhor resultado (Figura 9).

Figura 8 - Eletroforese em gel de agarose 1,5% da primeira reação de reamplificação. M - Marcador 100 pb (Invitrogen); 1 a 45 - amostras das 45 bandas polimórficas retiradas do gel de AFLP

Figura 9 - Eletroforese em gel de agarose 1,5% da segunda reação de reamplificação. M - Marcador 100 pb (Invitrogen); 1 a 45 - amostras das 45 bandas polimórficas retiradas do gel de AFLP

Apesar de muitas amostras apresentarem fragmentos de diversos tamanhos, foi possível recuperar, para algumas amostras, o fragmento do tamanho esperado. As amostras 03, 05 e 07 da combinação de iniciadores E+A e M+C e as amostras 14, 15, 25, 37, 40 e 41 da combinação P+A e M+C foram sequenciadas com sucesso.

4.2 Contexto genômico dos polimorfismos encontrados por AFLP

Todas as amostras sequenciadas foram analisadas e localizadas no genoma do fungo *S. scitamineum* (TANIGUTI et al., 2015) em cinco diferentes cromossomos e em uma região do DNA mitocondrial (mtDNA) (Figura 10). Em algumas dessas regiões foi detectado o motivo que é reconhecido pelo Prf1 (*Pherormone response factor 1*), esse motivo, chamado de PRE (*Pheromone Response Element*), possui 9 nucleotídeos, 5'- ACA AAG GGA -3', essa sequência foi identificada e caracterizada em *U. maydis* (URBAN et al., 1996), e identificada em outros fungos causadores de carvão (KELLNER et al., 2011). Segundo HARTMMAN et al. (1996), quando há uma substituição na sequência na posição 3 (A por G) e na posição 6 (G por C ou G por A), não ocorre a ligação de Prf1 com o motivo ou a ligação que ocorre é muito fraca. E, quando ocorre uma substituição em outra posição, a ligação fica mais fraca que com a sequência original.

Figura 10 - Contexto genômico dos polimorfismos encontrados através do AFLP. Setas em vermelho indicam a localização dos polimorfismos e as barras verdes indicam a localização de motivos PRE reconhecido por Prf1 próximos aos polimorfismos

No caso dos PREs que foram identificados na Figura 10, todos eles apresentam apenas uma modificação na sequência original de 9 nucleotídeos, sendo que essa substituição não está nas posições 3 ou 6 dessa sequência, com exceção do PRE que está logo *downstream* ao gene g3509, no cromossomo 08, esse apresenta uma substituição na posição 6, de G para A.

Em *U. maydis*, o Prf1 possui duas grandes funções durante o desenvolvimento do fungo, (1) obter uma alta quantidade de transcritos dos genes *Mfa* (ferormônio) e *pra* (receptor para o ferormônio), pertencentes ao *locus a* e (2) manter a expressão dos genes pertencentes ao *locus b* durante o desenvolvimento infeccioso (URBAN et al., 1996; HARTMMAN et al., 1996; HARTMMAN et al., 1999). Ou seja, Prf1 é um fator de transcrição necessário para o crescimento filamentoso e para a patogenicidade. E, além de os PREs serem encontrados nas vizinhanças dos genes induzidos por ferormônio, ele também pode atuar a distância, porém o mecanismo molecular de como isso ocorre ainda não foi elucidado (URBAN et al., 1996; HARTMMAN et al., 1996).

Algumas das amostras sequenciadas são variações do mesmo *locus*, como as amostras 05 e 07 que se alinham na mesma região do cromossomo 02 (Figura 11) e as amostras 25, 37 e 41, no cromossomo 05 (Figura 12).

Figura 11 - Alinhamento das amostras 05 e 07 sequenciadas com a região do cromossomo 02 do genoma de *S. scitamineum*

Figura 12 - Alinhamento das amostras 25, 37 e 41 sequenciadas com a região do cromossomo 05 do genoma de *S. scitamineum*

As amostras 05 e 07, polimorfismos presentes no cromossomo 02, estão localizadas *downstream* ao gene g981 (Figura 10). Ao realizar o BLASTx em bancos de dados públicos, esse gene não apresentou *hits* com boa cobertura, identidade e *e-value* considerável. Também foram analisados os genes vizinhos a esse, o gene g980 também não apresentou *hits* consideráveis, em ambos não foi possível identificar domínios conservados, peptídeo sinal ou hélices transmembrana; já o gene g982, apresentou alta similaridade (78 a 60% de identidade) com proteínas hipotéticas de organismos próximos como *S. reilianum* (NCBI: CBQ71130.1; UniProt: E6ZV73), *U. maydis* (NCBI: XP_011388918.1; UniProt: A0A0D1CTQ3) e *U. hordei* (NCBI: CCF49939.1 e CAJ41910.1; UniProt: I2FSP6 e Q2A763), e também apresentou alta similaridade (59% de identidade) com fator de transcrição C6 dos organismos

Melanopsichium pennsylvanicum (NCBI: CDI55354.1; UniProt: A0A077RCP8) e *Pseudozyma hubeiensis* (NCBI: XP_012192243.1; UniProt: R9PBA1).

Ao analisar os domínios presentes na proteína predita a partir da sequência de DNA do gene g982, foram detectados dois domínios. O primeiro seria relacionado a um *zinc finger* (IPR001138), um motivo estrutural de proteínas. Em fungos, várias proteínas de regulação transcricional possuem em sua porção N-terminal um motivo rico em cisteína relacionado à interação com DNA que é dependente de íons de zinco. O segundo é um domínio específico presente em fatores de transcrição em fungos, geralmente presente em proteínas que atuam na regulação da transcrição em diversos processos celulares e metabólicos (IPR007219).

As amostras 25, 37 e 41 são polimorfismos que estão localizados *upstream* ao gene g2477, localizado no cromossomo 05 (Figura 10). Esse apresentou *hits* com proteínas hipotéticas em organismos próximos, porém não foram identificados domínios conservados, peptídeo sinal ou hélices transmembrana em sua sequência predita. O gene g2476 obteve *hits* com alta similaridade (96 a 92% de identidade) para uma 3-fosfoglicerato desidrogenase de *S. reilianum* (NCBI: CBQ70639.1; UniProt: E6ZTT2), *Pseudozyma brasiliensis* (NCBI: EST06176.1; UniProt: V5E7C4), *P. antarctica* (NCBI: GAK63958.1 e GAC72541.1; UniProt: M9LM08), *P. hubeiensis* (NCBI: XP_012190606.1; UniProt: R9P6Q2), *U. hordei* (NCBI: CCF54427.1; UniProt: I2G5I4), *U. maydis* (NCBI: XP_011387167.1; UniProt: A0A0D1CYA8), entre outros. Além disso, um dos resultados obtidos foi uma D-3-fosfoglicerato desidrogenase 2 manualmente curada no Swiss-Prot de *Saccharomyces cerevisiae* (P40510) que possui 75,4% de similaridade e *e-value* de 0.0. Ao analisar os domínios presentes na proteína predita, foi detectado o domínio catalítico 2-hidroxiácido desidrogenase (IPR006139), da família de proteínas D-3-fosfoglicerato desidrogenase tipo 2 (IPR029015).

Já o gene g2478, *downstream* ao gene g2477, apresentou *hits* (69 a 56% de identidade) para proteínas hipotéticas de *U. maydis* (NCBI: XP_011387166.1; UniProt: A0A0D1CE33) e *P. hubeiensis* (NCBI: XP_012190612.1; UniProt: R9PFZ8) e para proteínas relacionadas as membros LAP3 do GAL regulon de *S. reilianum* (NCBI: CBQ70638.1; UniProt: E6ZTT1) e *Melanopsichium pennsylvanicum* (NCBI: CDI52170.1; UniProt: A0A077QRQ0). A proteína foi identificada como sendo parte da família protéica peptidase C1B, bleomicina hidrolase (IPR004134).

A amostra 14 alinha em uma região do cromossomo 08 (Figura 13) que fica na região *upstream* de dois genes divergentes g3507 e g3508 (Figura 10). O gene g3507 apresentou *hits* de alta similaridade (83 a 73% de identidade) para proteínas relacionadas à cinesina de *S.*

reilianum (NCBI: CBQ69686.1; UniProt: E6ZR29), *U. maydis* (NCBI: XP_011391611.1; UniProt: A0A0D1DS04), *U. hordei* (NCBI: CCF49194.1; UniProt: I2FQK1), *P. hubeiensis* (NCBI: XP_012191901.1; UniProt: R9PAD9) e *M. pennsylvanicum* (NCBI: CDI55762.1; UniProt: A0A077R958). Os domínios identificados na sequência foram o domínio cinesina motora (IPR001752) e o domínio nucleosídeo trifosfato hidrolase contendo P-loop (IPR027417).

Figura 13 - Alinhamento da amostra 14 sequenciada com a região do cromossomo 08 do genoma de *S. scitamineum*

O gene g3508 é altamente similar (68 a 67% de identidade) a uma provável SAICAR (Fosforibosilamidoimidazole-succinocarboxamida) sintetase de *S. reilianum* (NCBI: CBQ69685.1; UniProt: E6ZR28), *U. maydis* (NCBI: XP_011391610.1; UniProt: A0A0D1DW00), *P. hubeiensis* (NCBI: XP_012191900.1; UniProt: R9PAB8) e *P. brasiliensis* (NCBI: EST08601.1; UniProt: V5EDR7), entre outros. Sendo que há alguns *hits* para proteínas manualmente curadas do Swiss-Prot, como a SAICAR sintetase de *S. cerevisiae* (P27616) que possui evidência experimental em nível de proteína e uma similaridade de 58.8% com a proteína predita de Ssc39 apresentando *e-value* de 1.5e-65. Quanto aos domínios identificados na sequência, esse foi compatível com o que foi encontrado no BLASTx, SAICAR sintase/ADE2 (IPR028923).

Outro gene analisado próximo a esse polimorfismo foi o g3509, gene convergente ao g3508, que apresentou *hits* de alta similaridade (87 a 73% de identidade) com uma proteína

relacionada à cinesina em *S. reilianum* (NCBI: CBQ69684.1; UniProt: E6ZR27), *U. hordei* (NCBI: CCF49192.1; UniProt: I2FQJ9) e *M. pennsylvanicum* (NCBI: CDI55751.1; UniProt: A0A077R930). Os domínios identificados nessa sequência foram os mesmos domínios identificados no gene g3507.

A sequência gerada pelo sequenciamento da amostra 15 alinhou a uma região do cromossomo 20 (Figura 14), na porção final do gene g6116 (Figura 10). Esse gene apresentou *hits* de alta similaridade (93 a 79% de identidade) com uma proteína relacionada ao complexo mediador, subunidade soh1, de *S. reilianum* (NCBI: CBQ69913.1; UniProt: E6ZRQ9), *U. hordei* (NCBI: CCF50305.1; UniProt: I2FTR2), *M. pennsylvanicum* (NCBI: CDI55974.1; UniProt: A0A077RDL3) e *P. antarctica* (NCBI: GAK67102.1; UniProt: M9MES1). Além disso, é altamente similar (87.8%) a subunidade 31 do complexo da RNA polimerase II mediador da transcrição (Q4P308) de *U. maydis* que está depositado no Swiss-Prot. Essa proteína faz parte da família protéica do complexo mediador, subunidade Med31 (IPR008831).

Figura 14 - Alinhamento da amostra 15 sequenciada com a região do cromossomo 20 do genoma de *S. scitamineum*

O gene g6115, que está *upstream* ao g6116, e apresentou *hits* (86 a 74% de identidade) com glutathione sintase de *S. reilianum* (NCBI: CBQ69912.1; UniProt: E6ZRQ8), *P. hubeiensis* (NCBI: XP_012191310.1; UniProt: R9P8L3), *P. antarctica* (NCBI: GAC75544.1; UniProt: M9LY24) e *M. pennsylvanicum* (NCBI: CDI55973.1; UniProt: A0A077R9B2), entre outros organismos. Os domínios identificados na sequência de proteína são todos relacionados à glutathione sintase (IPR014049; IPR016185; IPR004887; IPR014709). E o gene g6117, que está *downstream* ao gene g6116, apresentou *hits* (91 a 80% de identidade) para proteínas

hipotéticas em fungos causadores de carvão e para permease em *P. hubeiensis* (NCBI: XP_012191308.1; UniProt: R9P8N5). O domínio identificado na sequência de proteína foi o transportador de triose fosfato (IPR004853), não foi identificado peptídeo sinal, mas foram identificadas hélices transmembranas (Figura 15).

Figura 15 - Identificação de hélices transmembrana na proteína codificada pelo gene g6117 localizado no cromossomo 20. Figura gerada pelo programa TMHMM (KROGH et al., 2001)

A sequência da amostra 40 alinha com a região do cromossomo 22 (Figura 16) no qual ficam localizados os genes g6514 e g6515 de forma convergente, sendo que o polimorfismo está localizado entre os dois genes (Figura 10). O gene g6514 só possui *hits* com proteínas hipotéticas e não foi possível identificar nenhum domínio nos bancos de dados.

Figura 16 - Alinhamento da amostra 40 sequenciada com a região do cromossomo 22 do genoma de *S. scitamineum*

O gene g6515 apresentou *hits* (87 a 79% de identidade) para uma proteína relacionada ao transporte de quinato em *S. reilianum* (NCBI: CBQ71353.1; UniProt: E6ZVN5), *U. hordei* (NCBI: CCF53555.1; UniProt: I2G312) e *M. pennsylvanicum* (NCBI: CDI56640.1; UniProt: A0A077R292). O domínio identificado na sequência foi um facilitador do transporte de açúcar/inositol (IPR020846; IPR003663). A proteína não apresentou peptídeo sinal, mas possui hélices transmembrana (Figura 17), possuindo uma estrutura compatível com um transportador transmembrana de açúcar.

Figura 17 - Identificação de hélices transmembrana na proteína codificada pelo gene g6515 localizado no cromossomo 22. Figura gerada pelo programa TMHMM (KROGH et al., 2001)

Por fim, a amostra 03 se alinhou ao DNA mitocondrial (Figura 18), *downstream* ao gene *cox3* (Figura 10). Esse gene possui *hits* de alta similaridade (99 a 93% de identidade) para a subunidade 3 da citocromo c oxidase de *S. reilianum* (NCBI: CBQ72551.1; UniProt: E6ZZ94), *U. maydis* (NCBI: YP_762698.1; UniProt: Q0H8X4), *U. porosa* (ADN68156.1; UniProt: G9B945), entre outros organismos. A sequência predita da proteína apresentou o domínio para a citocromo c oxidase III (IPR000298), sem peptídeo sinal e com hélices transmembrana (Figura 19).

Figura 18 - Alinhamento da amostra 03 sequenciada com a região do DNA mitocondrial de *S. scitamineum*

Figura 19 - Identificação de hélices transmembrana na proteína codificada pelo gene *Cox3* localizado no mtDNA. Figura gerada pelo programa TMHMM (KROGH et al., 2001)

O gene *cob*, que está próximo ao *cox3*, apresentou *hits* com alta similaridade (97 e 90% de identidade) com a proteína mitocondrial citocromo b de *S. reilianum* (NCBI: CBQ72547.1; UniProt: E6ZZ90), *U. maydis* (NCBI: YP_762699.1; UniProt: Q0H8X3), entre outros organismos que são filogeneticamente mais distantes. Os domínios identificados na sequência de proteína são todos relacionados ao citocromo b (IPR027387; IPR016174; IPR005798). Naturalmente a proteína não possui peptídeo sinal, pois participa da cadeia respiratória na membrana interna da mitocôndria e apresenta hélices transmembranas (Figura 20).

Figura 20 - Identificação de hélices transmembrana na proteína codificada pelo gene *cob* localizado no mtDNA. Figura gerada pelo programa TMHMM (KROGH et al., 2001)

Esses resultados foram resumidos em forma de tabela no Anexo B.

4.3 Análise da expressão gênica em três condições diferentes

A análise de expressão dos genes próximos aos polimorfismos foi feita utilizando três condições diferentes: *S. scitamineum* em meio de cultura, 5 dias após inoculado (DAI) e 200 DAI (Figura 21). Alguns dos genes analisados não apresentaram *reads* alinhados em nenhuma das condições, o que pode indicar que a anotação pode estar errada, sendo esse um pseudogene, ou que o gene é expresso em uma condição não considerada no estudo. Além disso, geralmente há menos *reads* alinhados na condição de 5 DAI, já que há uma menor concentração de RNA do fungo nas amostras desse tempo em comparação com as amostras das outras duas condições.

Figura 21 - Análise da expressão de genes próximos aos polimorfismos detectados através do AFLP em três condições distintas, meio de cultura, 5 DAI e 200 DAI

Os genes que apresentaram maior número de *reads* mapeados em meio de cultura foram os genes g2476 e g2478, ambos localizados no cromossomo 05. O gene g2476 apresentou 3100 *reads* mapeadas em meio de cultura; em 5 DAI, 970,33 e 200 DAI foi a condição que obteve o menor número (857,17). O gene g2478 também apresentou *reads* mapeadas nas três condições, 416,47 em meio de cultura, em 5 DAI foi a condição que obteve menos *reads* mapeadas (132,07) e em 200 DAI foram mapeadas um número de *reads* próximo ao encontrado em meio de cultura (371,47).

Como dito anteriormente, é mais difícil obter um número alto de *reads* mapeadas em 5 DAI, porém alguns genes apresentaram esse número maior para essa condição que para as outras duas. O gene g3508 do cromossomo 08 apresentou 660,37 *reads* mapeadas em 5 DAI, 179,93 em meio de cultura e em 200 DAI houve poucas *reads* mapeadas (72,13). O gene g6116, localizado no cromossomo 20, apresentou 365,1, 260,43 e 37,67 *reads* mapeadas em 5 DAI, 200 DAI e meio de cultura, respectivamente. E os dois genes analisados do cromossomo 22: o gene g6514 possui 1234,5 *reads* mapeadas em 5 DAI enquanto que em meio de cultura (613) e 200 DAI (781,73) esse gene possui quase metade de *reads* mapeadas; o gene g6515 possui um pouco menos de *reads* que o gene anterior, 566,77 em 5 DAI, 230,7 em 200 DAI e 88,2 em meio de cultura.

A maioria dos genes analisados obteve um número de *reads* mapeadas maior em 200 DAI, sendo que, entre eles, muitos não obtiveram *reads* mapeadas em 5 DAI. O gene g982, do cromossomo 02, apresentou alguns no meio de cultura (53,73) e em 200 DAI (256,37), mas não em 5 DAI. Os genes g3507 e g3509, localizados no cromossomo 08, também não obtiveram *reads* mapeadas em 5 DAI. O g3507 apresentou 102,2 e 585,53 em meio de cultura e 200 DAI, respectivamente, e o g3509, 40,77 e 164,77. Já os genes g6115 e g6117, localizados no cromossomo 20, apresentaram *reads* mapeadas nas três condições. O gene g6115 possui 302,6, 370,43 e 193,03 *reads* mapeadas em 5 DAI, 200 DAI e meio de cultura, respectivamente, e o gene g6117, 302,6, 333,43 e 135,17, respectivamente.

Quanto aos genes do mtDNA, apresentaram maior número de *reads* mapeadas em 200 DAI, sendo que para o gene *cox3* essa foi a única condição na qual foram mapeadas *reads* (30). O gene *cob* apresentou um número próximo (28,67) ao do *cox3* em 200 DAI, 13,67 *reads* mapeadas em meio de cultura e nenhuma *read* em 5 DAI.

Os genes g980 e g981 localizados no cromossomo 02 e o gene g2477 do cromossomo 05 não apresentaram *reads* mapeadas nessa análise.

5 DISCUSSÃO

5.1 Considerações Gerais

Os 45 fragmentos polimórficos gerados por AFLP foram submetidos ao processo de reamplificação, porém apenas 20% dos polimorfismos detectados puderam ser caracterizados nesse trabalho. A principal dificuldade nessa etapa foi a amplificação de fragmentos inespecíficos. Os fragmentos sequenciados com sucesso foram alinhados em diferentes cromossomos, inclusive no mtDNA, o que pode indicar que os polimorfismos detectados através de AFLP estão bem dispersos no genoma de *S. scitamineum*. Além disso, foi possível detectar diferentes alelos do mesmo *locus* através da caracterização desses polimorfismos.

Através das análises de expressão dos genes vizinhos aos polimorfismos, foi possível identificar três genes que não apresentam *reads* mapeadas em nenhuma das condições analisadas e que também não apresentaram *hits* em análises em bancos de dados. Esses genes são os g980 e g981 localizados no cromossomo 02 e o gene g2477, no cromossomo 05. Esses genes podem estar anotados de forma errada, sendo eles pseudogenes, ou então são genes que ainda não foram caracterizados e que são expressos em condições não consideradas nesse trabalho.

Além disso, foi possível identificar genes que são mais expressos *in planta*, os mais expressos em 5 DAI e genes adjacentes que podem apresentar co-regulação da transcrição, como será abordado nos itens seguintes.

5.2 Genes mais expressos *in planta*

A maioria dos genes analisados (~70%) apresentou maior número de *reads* mapeados *in planta*, quando há a interação planta-patógeno, o que indica que os genes analisados são importantes para a interação e desenvolvimento da doença.

O gene g982, localizado no cromossomo 02, foi relacionado a fatores de transcrição em sua caracterização e apresentou mais *reads* em 200 DAI que em meio de cultura, mas em comparação com os demais genes analisados não apresentou, em números, muitas *reads* mapeadas. Ainda assim, esse fator de transcrição pode ser muito importante para a regulação transcricional de genes que são essenciais para a resposta de adaptação do fungo ao ambiente dentro de seu hospedeiro.

Os genes g3507 e g3509, localizados no cromossomo 08, foram identificados como sendo relacionados à proteína motora cinesina. Essa proteína é um microtúbulo que atua no

transporte de organelas dentro da célula e possui inúmeras isoformas que podem ser reguladas de formas diferentes. Isso pode explicar o porquê da quantidade de *reads* mapeadas nesses genes é pequena. Em *S. cerevisiae*, a KAR3 é essencial para a fusão nuclear e as proteínas CIN8 e KIP1 são necessárias para a mitose, interagindo com as fibras do fuso (BRADY, 1995). Estudos em *U. maydis* em relação à cinesinas também foram realizados, demonstrando que elas são importantes para o crescimento filamentosso direcionado (WEDLICH-SÖLDNER et al., 2002; LENZ et al., 2006; STEINBERG, 2015), sendo importante para a penetração do fungo nos tecidos da planta. Isso pode ser um indicativo do porque esses genes são mais expressos *in planta* que em meio de cultura, apesar de não ter sido detectada nenhuma *read* em 5 DAI, mas isso pode ser devido à amostra que possui pouco RNA de fungo nesse tempo.

Os genes localizados no cromossomo 20, também possuem mais *reads* mapeadas *in planta* que em meio de cultura, apesar da diferença entre um ambiente e outro não serem muito diferentes para os genes g6115 e g6117. O primeiro foi relacionado a uma glutatona sintase e o segundo a um transportador de triosefosfato. A glutatona está envolvida no metabolismo de resposta a estresse, especialmente ao estresse oxidativo (GASCH et al., 2000; GRANT, 2001), além disso, está envolvido em funções celulares indispensáveis ao funcionamento celular, como a proteção do DNA mitocondrial e a estabilização de membranas e proteínas (PÓCSI et al., 2004). Isso explica por que ocorre a expressão desse gene em condições ótimas, como o meio de cultura, e a expressão é intensificada durante a interação com a planta. Quanto ao transportador de triosefosfato, sua expressão pode ser levemente intensificada durante a interação com o hospedeiro porque é necessário que haja mais transportadores de membrana para açúcares, já que em meio de cultura há maior disponibilidade de açúcares para o fungo do que *in planta*.

Já o gene g6116, também localizado no cromossomo 20, é o que apresenta maior diferença de expressão entre as condições com e sem interação com o hospedeiro. Esse é o único gene no qual o polimorfismo está localizado dentro da ORF, que codifica para a subunidade 31 com complexo mediador da transcrição (soh1). O complexo mediador facilita a ligação do complexo de pré-iniciação da transcrição e possui diversas subunidades que atuam em conjunto, algumas subunidades não são essenciais para o funcionamento do complexo, como o soh1 (CONAWAY e CONAWAY, 2011). Além da função de atuar na regulação transcricional, o soh1 também interage com fatores envolvidos com o reparo de DNA (FAN et al., 1996). Além de ser curioso o fato de que essa subunidade seja mais expressa *in planta*, também é o fato de que o polimorfismo encontrado através de AFLP foi encontrado na ORF desse gene, já que as subunidades do complexo mediador são extremamente conservadas.

Os genes mitocondriais analisados obtiveram um maior número de *reads* mapeadas em 200 DAI, sendo que o gene *cox3* só obteve nessa condição e o *cob* apresentou poucas *reads* mapeadas em meio de cultura. Esses genes estão relacionados à cadeia respiratória e podem estar mais expressos *in planta* durante essa fase da infecção por ser a fase em que o fungo está em esporogênese, portanto, o seu metabolismo está mais ativo.

5.3 Genes mais expressos em 5 DAI

Os genes que apresentaram maior número de *reads* mapeadas em 5 DAI podem ser especialmente importantes para a interação planta-patógeno, já que nesse período da infecção é mais difícil obter RNA de fungo, havendo maior quantidade de RNA proveniente da planta para o sequenciamento.

O gene 3508, localizado no cromossomo 08, possui 3X mais *reads* mapeadas em 5 DAI do que em meio de cultura e 9X mais do que em 200 DAI. Esse gene codifica para uma SAICAR sintetase, atuando na síntese *de novo* de nucleotídeos purínicos (MANJUNATH et al., 2015).

Os genes g6514 e g6515, localizados no cromossomo 22, também apresentaram um número consideravelmente maior de *reads* mapeadas em 5 DAI. O gene g6514 foi o gene que apresentou o maior número de *reads* mapeadas em 5 DAI, porém esse gene não foi relacionado a nenhuma proteína já descrita, somente com proteínas hipotéticas. Com a sua elevada expressão no tempo inicial de infecção, pensou-se, em um primeiro momento, que se tratasse de um efector, porém não foi identificado em sua sequência predita um peptídeo sinal, necessário para a secreção da proteína. Existem alguns efetores descritos que não possuem o peptídeo sinal e são secretados por uma via não convencional que ainda não foi descrita (LIU et al., 2014). Essa proteína, indiscutivelmente é importante para a interação inicial de *S. scitamineum* e cana-de-açúcar, porém ainda necessita de mais estudos para que seja possível caracterizá-la.

Já o gene g6515 é relacionado a um transportador de quinato, um açúcar associado à via do shikimato. O quinato já foi apontado como um foco de manipulação por fungos fitopatogenicos nos estágios iniciais para que a produção de corismato seja suprimida, dessa forma, evita-se a síntese de ácido salicílico e diminui as defesas da planta (PARKER et al., 2009). O que é condizente com o observado, já que o aumento na expressão ocorre nas etapas iniciais e foram identificadas hélices transmembranas na sequência de proteína.

O gene g2476, localizado no cromossomo 05, apesar de possuir maior número de reads mapeadas em meio de cultura, também apresentou um número expressivo de *reads* em 5 DAI, o que chamou a atenção pela dificuldade de se conseguir *reads* em 5 DAI, como foi dito anteriormente. Esse gene está relacionado com uma 3-fosfoglicerato desidrogenase, enzima que participa da síntese de L-serina.

5.4 Co-regulação de genes adjacentes

A co-regulação de genes adjacentes ocorre entre os diversos grupos de eucariotos de forma frequente, entre pares de genes que podem estar em diferentes orientações (divergente, *tandem* ou convergente), o que sugere que a regulação não é resultado apenas de um promotor bidirecional. É possível que haja uma pressão evolutiva para que se favoreça a co-regulação de genes adjacentes que são altamente expressos ou regulados (ARNONE et al., 2012; ARNONE et al., 2014). Um par de genes analisados que pode ser produto de co-regulação são os dois genes localizados no cromossomo 22.

Esses genes são bastante requeridos durante a fase inicial da infecção, decaindo a transcrição em 200 DAI, dessa forma, pode ser que sejam co-regulados. Porém, necessita-se de maiores estudos para confirmar se essa co-regulação realmente ocorre.

5.5 Há a formação de raças em *S. scitamineum*?

Apesar de poucos polimorfismos detectados via AFLP entre os isolados de *S. scitamineum*, esses polimorfismos estão próximos a genes que são importantes para o desenvolvimento da doença. Essas regiões intergênicas podem ter influência na transcrição desses genes, o que poderia influenciar na agressividade do patógeno e na sua defesa contra o mecanismo de defesa da planta. Porém ainda é necessário que se estabeleça a interação entre os polimorfismos e a expressão gênica dos seus genes vizinhos de forma experimental.

6 CONSIDERAÇÕES FINAIS

Poucos polimorfismos foram detectados através do AFLP entre os isolados brasileiros, porém esses que foram detectados podem estar em regiões importantes para a agressividade do fungo, influenciando em seu mecanismo de patogenicidade.

Dentre os isolados estudados por REIS (2012), somente o isolado Ssc18A não apresentou nenhum dos polimorfismos caracterizados nesse trabalho; o Ssc35A foi o único isolado que apresentou o polimorfismo proveniente da amostra 03, localizado no mtDNA, e os isolados Ssc35A e Ssc36B foram os únicos que apresentaram o polimorfismo caracterizado no cromossomo 02.

Dentre os polimorfismos caracterizados, somente um foi detectado dentro de uma ORF, g6116, gene localizado no cromossomo 20. Esse polimorfismo está mais presente nos indivíduos de *mating type* B.

Os polimorfismos detectados e que foram caracterizados com sucesso nesse trabalho estão próximos de genes importantes para a interação planta-patógeno, sendo que cerca de 70% dos genes analisados são mais expressos *in planta*. Porém, para que se confirme a influência desses polimorfismos na expressão desses genes, são necessários estudos experimentais futuros.

Além disso, foi identificado um gene interessante para a interação inicial de *S. scitamineum* e cana-de-açúcar: o g6514, localizado no cromossomo 22, mas que não está caracterizado. Esse pode ser um gene importante a ser estudado em trabalhos futuros, já que possui uma expressão bem elevada nos momentos iniciais de infecção.

REFERÊNCIAS

- ALBERT, H.H.; SCHENCK, S. PCR amplification from a homolog of the *bE* matingtype genes a sensitive assay for the presence of *Ustilago scitaminea* DNA. **Plant Disease**, Saint Paul, v. 80, n. 10, p. 1189-1192, 1996.
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. Control of gene expression. In: ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular biology of the cell**, New York, Taylor & Francis, 2008. p. 411-499.
- ALTHOFF, D.M.; GITZENDANNER, M.A.; SEGRAVES, K.A. The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. **Systematic Biology**, Oxford, v. 56, p. 477–484, 2007.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, San Diego, v. 215, p. 403-410, 1990.
- AGARWAL, M.; SHRIVASTAVA, N.; PADH, N. Advances in molecular marker techniques and their applications in plant sciences. **Plant Cell Reports**, Rockville, v. 27, p. 617–631, 2008.
- ARNONE, J.T.; ARACE, J.R.; SOORNEEDI, A.R.; CITINO, T.T.; KAMITAKI, T.L.; MCALEAR, M.A. Dissecting the cis and trans elements that regulate adjacent-gene coregulation in *Saccharomyces cerevisiae*. **Eukaryot Cell**, Washington, v. 13,n.6, p. 738-748, 2014.
- ARNONE, J.T.; ROBBINS-PIANKA, A.; ARACE, J.R.; KASS-GERGI, S.; MCALEAR, M.A. The adjacent positioning of co-regulated gene pairs is widely conserved across eukaryotes. **BMC Genomics**, London, v. 10, p. 13-546, 2012.
- BAKKEREN, G.; KÄMPER, J.; SCHIRAWSKI, J. Sex in smut fungi: structure, function and evolution of mating-type complexes. **Fungal Genetics and Biology**, Orlando, v. 45, suppl. 1, p. S15-S21, 2008.
- BECKMANN, J.S.; SOLLER, M. Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. **Tag Theoretical and Applied Genetics**, Berlin, v.67, n. 1, p. 35-43, 1983.
- BORRAS-HIDALGO, O.; THOMMA, B.; CARMONA, E.; BORROTO, C.J.; PUJOL, M.; ARENCIBIA, A.; LOPEZ, J. Identification of sugarcane genes induced in disease-resistant somaclones upon inoculation with *Ustilago scitaminea* or *Bipolaris sacchari*. **Plant Physiology and Biochemistry**, Dorchester, v. 43, n.12, p. 1115-1121, 2005.
- BORTFELD, M.; AUFFARTH, K.; KAHMANN, R.; BASSE, C.W. The *Ustilago maydis* a2 mating-type locus genes *lga2* and *rga2* compromise pathogenicity in the absence of the mitochondrial p32 family protein Mrb1. **Plant Cell**, Rockville, v. 16, p. 2233– 2248, 2004.
- BRADY, S. A kinesin medley: biochemical and functional heterogeneity. **Trends in Cell Biology**, Cambridge, v. 5, p. 159-164, 1995.

- BRAITHWAITE, K.S.; BAKKEREN, G.; CROFT, B.J.; BRUMBLEY, S.M. Genetic variation in a worldwide collection of the sugarcane smut fungus *Ustilago scitaminea*. In: AUSTRALIAN SOCIETY SUGARCANE TECHNOLOGIST CONGRESS, 2004, Mackay, **Proceedings...** Mackay: AuSSCT, 2004. v. 26. p. 224-232.
- BRAND, A.H.; BREEDEN, L.; ABRAHAM, J.; STERNGLANZ, R.; NASMYTH, K. Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. **Cell**, Cambridge, v. 41,n.1, p. 41-48, 1985.
- BUTTERFIELD, M.K.; RUTHERFORD, R.S.; CARSON, D.L.; HUCKETT, B.I. Application of gene discovery to varietal improvement in sugarcane. **South African Journal of Botany**, South Africa, v. 70, n.1, p. 167-172, 2004.
- CASAGRANDE, M.V. **Avaliação da incidência da doença e estimativa de danos ocasionados pelo carvão (*Ustilago scitaminea* Sydow) em variedades de cana-de-açúcar**. 1998. 86p. Dissertação (Mestrado em Fitopatologia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1998.
- COMSTOCK, J.C. Smut. In: ROTT, P.; BAILEY, R.A.; COMSTOCK, J.C.; CROFT, B.J.; SALEM SAUMTALLY, A. (Ed.). **A guide to sugarcane diseases**. Montpellier: CIRAD; ISSCT, 2000. p. 181-185.
- COMSTOCK, J.C.; HEINZ, D.J. A new race of culmicolous smut of sugarcane in Hawaii. **Sugarcane Pathologists Newsletter**, Aiea, v. 19, p. 24-25, 1977.
- COMSTOCK, J. C.; LENTINI, R. S. Sugarcane smut disease. Gainesville: University of Florida, IFAS Extension, 2005. Disponível em: <<http://edis.ifas.ufl.edu/pdf/SC/SC00800.pdf>>. Acesso em: 08 jun. 2015.
- CONAWAY, R.C.; CONAWAY, J.W. Origins and activity of the Mediator complex. **Seminars in Cell and Developmental Biology**. Stover Court, v. 22, n.7, p. 729-734, 2011.
- CRESTE, S.; TULMANN-NETO, A.; FIGUEIRA, A.V.O. Detection of single sequence repeat polymorphism in denaturing polyacrylamide sequencing gels by silver straining. **Plant Molecular Biology Report**, Athens, v. 19, p. 299-306, 2001.
- DE LAAT, W.; GROSVELD, F. Spatial organization of gene expression: the active chromatin hub. **Chromosome Research**, Dordrecht, v. 11, n.5, p. 447-459, 2003.
- DE NADAL, E.; AMMERER, G.; POSAS, F. Controlling gene expression in response to stress. **Nature Reviews Genetics**, New York v. 12, p. 833-845, 2011.
- FAN, H-Y.; CHENG, K.K.; KLEIN, H.L. Mutations in the RNA Polymerase II Transcription Machinery Suppress the Hyperrecombination Mutant Hpr1Δ of *Saccharomyces cerevisiae*. **Genetics**, Bethesda, v. 142, n.3, p. 749-759, 1996.
- FAOSTAT. Food and Agricultural commodities production – Countries by commodity. Disponível em: <http://faostat3.fao.org/browse/rankings/countries_by_commodity/E>. Acesso em: 20 jul. 2015.

FERREIRA, S.A.; COMSTOCK, J.C. Smut. In: RICAU, C.; EGAN, B.T.; GILLASPIE JR, A.G.; HUGHES, C.G. (Ed.). **Diseases of sugarcane**: major diseases. Amsterdam: Elsevier, 1989. p. 211-229.

FIGUEIREDO, P. Breve história da cana-de-açúcar e do papel do instituto agrônomo no seu estabelecimento no Brasil. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C. M. de; ANDRADE LANDELL, M.G. de. **Cana-de-açúcar**. Campinas: Instituto Agrônomo, 2008. p. 31-41.

GASCH, A.P.; SPELLMAN, P.T.; KAO, C.M.; CARMEL-HAREL, O.; EISEN, M.B.; STORZ, G.; BOTSTEIN, D.; BROWN, P.O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. **Molecular Biology of the Cell**, Bethesda, v. 11, p. 4241–4257, 2000.

GRANT, C.M. Role of the glutathione/glutaredoxin and thioredoxin systems in yeast growth and response to stress conditions. **Molecular Microbiology**, Parañaque, v. 39, p. 533–541, 2001.

GERSTEIN, M.B.; BRUCE, C.; ROZOWSKY, J.S.; ZHENG, D.; DU, J.; KORBEL, J.O.; EMANUELSSON, O.; ZHANG, Z.D.; WEISSMAN, S.; SNYDER, M. What is a gene, post-ENCODE? History and updated definition. **Genome Research**, Long Island, v. 17, p. 669-681, 2007.

GILLASPIE, A.G Jr.; MOCK, R.G.; DEAN, J.L. Differentiation of *Ustilago scitaminea* isolates in greenhouse tests. **Plant Diseases**, Saint Paul, v. 67, n. 4, p. 373-375, 1983.

HARTMANN, H.A.; KAHMANN, R.; BOLKER, M. The pheromone response factor coordinates filamentous growth and pathogenicity in *Ustilago maydis*. **The EMBO Journal**, Heidelberg, v. 15,n.7, p. 1632-1641, 1996.

HARTMANN, H.A.; KRÜGER J.; LOTTSPEICH, F.; KAHMANN, R. Environmental Signals Controlling Sexual Development of the Corn Smut Fungus *Ustilago maydis* through the Transcriptional Regulator Prf1. **The Plant Cell**, Rockville, v.11, p. 1293–1305, 1999.

HAUKSWORTH, D.L.; KIRK, P.M.; SUTTON, B.C.; PEGLER, D.N. **Ainsworth & Bisby's dictionary of the fungi**. 8th ed. Cambridge: CAB International, 1995. 616p.

HEINZE, B.S.; THOKOANE, L.N.; WILLIAMS, N.J.; BARNES, J.M.; RUTHERFORD, R.S. The smut-sugarcane interaction as a model system for the integration of marker discovery and gene isolation. **Proceedings of South African Sugar Technologist Association**, South Africa, v. 75, p. 88-93, 2001.

KELLNER, R.; VOLLMEISTER, E.; FELDBRU, M.; BEGEROW, D. Interspecific Sex in Grass Smuts and the Genetic Diversity of Their Pheromone-Receptor System. **PLoS Genetics**, Stanford, v. 7, n.12, e1002436, 2011.

KROGH, A.; LARSSON, B.; VON HEIJNE, G.; SONNHAMMER, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. **Journal of Molecular Biology**, San Diego, v. 305, p. 567-580, 2001.

KRONSTAD, J.W.; LEONG, S.A. The b mating-type locus of *Ustilago maydis* contains variable and constant regions. **Genes**, Amsterdam, 4, p. 1384–1395, 1990.

KRONSTAD, J.W.; STABEN, C. Mating type in filamentous fungi. **Annual Review of Genetics**, Palo Alto, v. 31, p. 245–276, 1997.

LAO, M.; ARENCIBIA, A.D.; CARMONA, E.R.; ACEVEDO, R.; RODRIGUEZ, E.; LEON, O.; SANTANA, I. Differential expression analysis by cDNA-AFLP of *Saccharum* spp. after inoculation with the host pathogen *Sporisorium scitamineum*. **Plant Cell Reports**, Rockville, p. 27,n.6, p. 1103-1111, 2008.

LAURIE, J.D.; ALI, S.; LINNING, R.; MANNHAUPT, G.; WONG, P.; GÜLDENER, U.; MÜNSTERKÖTTER, M.; MOORE, R.; KAHMANN, R.; BAKKEREN, G.; SCHIRAWSKI, J. Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. **The Plant Cell**, Rockville, v. 24, n. 5, p. 1733-1745, 2012.

JH LENZ, J.H.; SCHUCHARDT, I.; STRAUBE, A.; STEINBERG, G. A dynein loading zone for retrograde endosome motility at microtubule plus-ends. **The EMBO Journal**, Heidelberg, v. 25, p. 2275-2286, 2006.

LEU, L.S.; TENG, W.S.; WANG, Z.N. Culmicolous smut of sugarcane in Taiwan, Resistant trial. **Taiwan Sugar Experiment Stantion Research Report**, v. 74, p. 37-45, 1976.

LIU, T.; SONG1, T.; ZHANG, X.; YUAN, H.; SU, L.; LI, W.; XU, J.; LIU, S.; CHEN, L.; CHEN, T.; ZHANG, M.; GU, L.; ZHANG, B.; DOU, D. Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. **Nature Communications**, New York, v. 5, p. 1-10, 2014.

LLOYD, H.L.; PILLAY, M. The development of an improved method for evaluating sugarcane for resistance to smut. **Proceedings of South Africa Sugar Technology Association**, Durban, v. 54, p. 168-172, 1980.

MANJUNATH, K.; JEYAKANTHAN, J.; SEKAR, K. Catalytic pathway, substrate binding and stability in SAICAR synthetase: A structure and molecular dynamics study. **Journal of Structural Biology**, Maryland, v. 191, p. 22-31, 2015.

MITCHELL, A.; CHANG, H-Y.; DAUGHERTY, L.; FRASER, M.; HUNTER, S.; LOPEZ, R.; MCANULLA, C.; MCMENAMIN, C.; NUKA, G.; PESSEAT, S.; SANGRADOR-VEGAS, A.; SCHEREMETJEW, M.; RATO, C.; YONG, S-Y.; BATEMAN, A.; PUNTA, M.; ATTWOOD, T.K.; SIGRIST, C.J.A.; REDASCHI, N.; RIVOIRE, C.; XENARIOS, I.; KAHN, D.; GUYOT, D.; BORK, P.; LETUNIC, I.; GOUGH, J.; OATES, M.; HAFT, D.; HUANG, H.; NATALE, D.A.; WU, C.H.; ORENGO, C.; SILLITOE, I.; MI, H.; THOMAS, P.D.; FINN, R.D. (2015). The InterPro protein families database: the classification resource after 15 years. **Nucleic Acids Research**, Oxford, doi: 10.1093/nar/gku1243, 2014.

MONTEIRO-VITORELLO, C.B.; QUECINE, M.C.; REIS, G.; PALHARES, A.; CASTELO-BRANCO, L.; KMIT, M.C.P.; SOUZA, S.C.; VIEIRA, M.L.C. Establishment of a genetic platform to study *Sporisorium scitamineum* in Brazil. **International Society of Sugar Cane Technologists**, São Paulo, BM24, 2013.

MUELLER, U.G.; WOLFENBARGER, L.L. AFLP genotyping and fingerprinting. **Trends in Ecology and Evolution**, Cambridge, vol. 14, no. 10, p. 389-394, 1999.

PALHARES, A.C. **Análise, via RNAseq, do transcriptoma da cana-de-açúcar e identificação de genes expressos em resposta a *Sporisorium scitamineum*, o agente causal do carvão**. 2014. 146p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2014.

PARKER, D.; BECKMANN, M.; ZUBAIR, H.; ENOT, D.P.; CARACUEL-RIOS, Z.; OVERY, D.P.; SNOWDON, S.; TALBOT, N.J.; DRAPER, J. Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by *Magnaporthe grisea*. **The Plant Journal**, Oxford, v. 59.n.5, p. 723-737, 2009.

PEARSON, H. Genetics: what is a gene? **Nature**, London, v. 441, p. 398-401, 2006.

PETERSEN, T.N.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, London, v. 8, p. 785-786, 2011.

PIEPENBRING, M.; STOLL, M.; OBERWINKLER, F. The generic position of *Ustilago scitaminea* and *Ustilago esculenta* (Ustilaginales). **Mycological Progress**, München, v. 1, n. 1, p. 71-80, 2002.

PLANTWISE. Pest Distribution Map. Disponível em:

<<http://www.plantwise.org/KnowledgeBank/PWMap.aspx>>. Acesso em: 10 jun. 2015.

PÓCSI, I.; PRADE, R.A.; PENNINCKX, M.J. Glutathione, Altruistic Metabolite in Fungi. **Advances in Microbial Physiology**, Oxford, v. 49, p. 1-76, 2004.

QUE, Y.; XU, L.; LIN, J.; RUAN, M.; ZHANG, M.; CHEN R. Differential Protein Expression in Sugarcane during Sugarcane-*Sporisorium scitamineum* Interaction Revealed by 2-DE and MALDI-TOF-TOF/MS. **Comparative and Functional Genomics**, Singapore, doi:10.1155/2011/989016, 2011.

QUE, Y.; XU, L.; WU, Q.; LIU, Y.; LING, H.; LIU, Y.; ZHANG, Y.; GUO, J.; SU, Y.; CHEN, J.; WANG, S.; ZHANG, C. Genome sequencing of *Sporisorium scitamineum* provides insights into the pathogenic mechanisms of sugarcane smut. **BMC Genomics**, London, v. 15, n. 996, p. 1-19, 2014.

R CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Viena, 2014. Disponível em: <<http://www.R-project.org/>>. Acesso em: 31 de jul. 2015.

RABOIN, L.M.; SELVI, A.; OLIVEIRA, K.M.; PAULET, F.; CALATAYUD, C.; ZAPATER, M.F.; BROTTIER, P.; LUZARAN, R.; GARSMEUS, O.; CARLIER, J.; D'HONT, A. Evidence for the dispersal of a unique lineage from Asia to America and Africa in the sugarcane fungal pathogen *Ustilago scitaminea*. **Fungal Genetics and Biology**, Orlando, v. 44, p. 64-76, 2007.

RAGO, A.M.; CASAGRANDE, M.V.; MASSOLA JUNIOR, N.S. Variabilidade patogênica de *Ustilago scitaminea* no Estado de São Paulo. **Summa Phytopathologica**, Jaguariúna, v. 35, n. 2, p. 93-97, 2009.

RAZIN, S.V.; GAVRILOV, A.A.; ULYANOV, S.V. Transcription-controlling regulatory elements of the eukaryotic genome. **Molecular Biology**, Athens, v. 49, p. 185-194, 2015.

REIS, G.V. **Diversidade genética de isolados do fungo *Sporisorium scitamineum* analisada através de *fingerprinting* da região telomérica**. 2012. 84p. Dissertação (Mestrado em Microbiologia Agrícola) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.

SCHIRAWSKI, J.; HEINZE, B.; WAGENKNECHT, M.; KAHMANN, R. Mating type loci of *Sporisorium reilianum*: novel pattern with three a and multiple b specificities. **Eukaryotic Cell**, Washington, v.4, p. 1317–1327, 2005.

SMALE, S.T.; KADONAGA, J.T. The RNA polymerase II core promoter. **Annual Review of Biochemistry**, Palo Alto, v. 72, p. 449-479, 2003.

SINGH, N.; SOMAI, B.M.; PYLLAY, D. Molecular, profiling demonstrates limited diversity amongst geographically separate strains of *Ustilago scitaminea*. **FEMS Microbiology Letters**, Amsterdam, v. 247, p. 7-15, 2005.

SNETSELAAR, K.M.; MIMS, C.W. Sporidial fusion and infection of maize seedlings by the smut fungus *Ustilago maydis*. **Mycologia**, New York, v. 84, p. 193-203, 1992.
 SNUSTAD, D.P.; SIMMONS, M.J. **Principles of Genetics**. 4th. Hoboken: John Wiley, 2006. 866p.

STEINBERG, G. Kinesin-3 in the basidiomycete *Ustilago maydis* transports organelles along the entire microtubule array. **Fungal Genetics and Biology**, Wisconsin, v. 74, p. 59–61, 2015.

STOLL, M.; PIEPENBRING, M.; BEGEROW, D.; OBERWINKLER, F. Molecular phylogeny of *Ustilago* and *Sporisorium* species (Basidiomycota, Ustilaginales) based on internal transcribed spacer (ITS) sequences. **Canadian Journal of Botany**, Ottawa, v. 81, p. 976–984, 2003.

TANIGUTI, L.M.; SCHAKER, P.D.C.; BENEVENUTO, J.; PETERS, L.P.; CARVALHO, G.; PALHARES, A.; QUECINE, M.C.; NUNES, F.R.S.; KMIT, M.C.P.; WAI, A.; HAUSNER, G.; AITKEN, K.S.; BERKMAN, P.J.; FRASER, J.A.; MOOLHUIJZEN, P.M.; COUTINHO, L.L.; CRESTE, S.; VIEIRA, M.L.C.; KITAJIMA, P.J., MONTEIRO-VITORELLO, C.B. Complete genome sequence of *Sporisorium scitamineum* and biotrophic interaction transcriptome with sugarcane. **PLoS One**, São Francisco, v. 10, n. 6, p.1-31, 2015.

THE ENCODE PROJECT CONSORTIUM. The ENCODE (ENCyclopedia Of DNA Elements) Project. **Science**, Washington, v. 306, n. 5696, p. 636-640, 2004.

THOKOANE L, RUTHERFORD R. cDNA-AFLP differential display of sugarcane (*Saccharum* spp. hybrids) genes induced by challenge with the fungal pathogen *Ustilago scitaminea* (sugarcane smut). **Proceedings of South African Sugar Technologist Association**, Durban, v. 75, p. 104-107, 2001.

TOKESHI, H. Doenças de cana-de-açúcar (híbridos de *Saccharum* spp.). In: KIMATI, H.; AMORIM, L.; REZENDE, J.A.M.; BERGAMIN FILHO, A.; CAMARGO, L.E.A. (Ed.). **Manual de fitopatologia**. São Paulo: Agronômica Ceres, 2005. v. 2, p. 207- 225.

URBAN, M.; KAHMANN, R.; BÖLKER, M. Identification of the pheromone response element in *Ustilago maydis*. **Molecular Genetics and Genetics**, Heidelberg, v. 251, p. 31–37, 1996.

VIEIRA, M.L.C.; NATAL, A.; SILVA FILHO, M.C. Genética e melhoramento vegetal. In: MIR, L. (Org). **Genômica**, São Paulo: Atheneu, 2004. chap.33, p. 679-699.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRIJTERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, Oxford, v. 23, p. 4407–4414, 1995.

WATSON, J.D.; CRICK, F.H.C. Molecular structure of nucleic acids. **Nature**, London, v. 171, n. 4356, p. 737-738, 1953.

WEDLICH-SÖLDNER, R.; STRAUBE, A.; FRIEDRICH, W.; STEINBERG, G. A balance of KIF1A-like kinesin and dynein organizes early endosomes in the fungus *Ustilago maydis*. **The EMBO Journal**, Heidelberg, vol. 21, n. 12, p. 2946-2957, 2002.

XU, L.; QUE, Y.; CHEN, R. Genetic Diversity of *Ustilago scitaminea* in Mainland China. **Sugar Tech**, Cham, v. 6, n. 4, p. 267 – 271, 2004.

ZHANG, Z.D.; PACCANARO, A.; FU, Y.; WEISSMAN, S.; WENG, Z.; CHANG, J.; SNYDER, M.; GERSTEIN, M.B. Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. **Genome Research**, Long Island, v. 17, n. 6, p. 787-797, 2007.

ANEXO

Anexo A - Bandas polimórficas que foram retiradas do gel de AFLP para reamplificação e sequenciamento
(continua)

Amostra	Combinação de iniciadores		Tamanho
01	E+AGA	M+CTG	352 pb
02	E+AGA	M+CTG	167 pb
03	E+AGA	M+CTG	152 pb
04	E+AGC	M+CTA	170 pb
05	E+AGA	M+CGC	230 pb
06	E+AGA	M+CGC	122 pb
07	E+AAA	M+CTC	167 pb
08	E+AAA	M+CTC	165 pb
09	E+AAA	M+CTC	154 pb
10	P+AGA	M+CAA	575 pb
11	P+AGA	M+CAA	550 pb
12	P+AGA	M+CAA	525 pb
13	P+AGA	M+CAA	300 pb
14	P+AGA	M+CAA	260 pb
15	P+AGA	M+CAA	223 pb
16	P+AGA	M+CAA	220 pb
17	P+AGA	M+CAA	215 pb
18	P+AGA	M+CAA	210 pb
19	P+AGA	M+CAA	187 pb
20	P+AAC	M+CTA	294 pb
21	P+AAC	M+CTA	265 pb
22	P+AAC	M+CTA	215 pb
23	P+AAC	M+CTA	198 pb
24	P+AAC	M+CTA	190 pb
25	P+AAA	M+CTC	247 pb
26	P+AAA	M+CTC	130 pb
27	P+ACA	M+CTA	178 pb
28	P+AAA	M+CAT	370 pb
29	P+AAA	M+CAT	310 pb
30	P+AAA	M+CAT	230 pb
31	P+AAA	M+CAT	190 pb
32	P+AAA	M+CAT	156 pb
33	P+AGA	M+CAG	545 pb
34	P+AGC	M+CAT	398 pb
35	P+AGC	M+CAT	275 pb
36	P+ACA	M+CAT	575 pb
37	P+AGA	M+CAC	245 pb

Anexo A - Bandas polimórficas que foram retiradas do gel de AFLP para reamplificação e sequenciamento
(conclusão)

Amostra	Combinação de iniciadores		Tamanho
38	P+ACT	M+CTG	460 pb
39	P+ACT	M+CTG	422 pb
40	P+ACT	M+CTG	148 pb
41	P+ACT	M+CTG	145 pb
42	P+AGA	M+CTC	245 pb
43	P+AGA	M+CTC	212 pb
44	P+AAC	M+CGC	620 pb
45	P+AAC	M+CGC	218 pb

Anexo B - Resumos dos resultados da análise do contexto genômico dos polimorfismos encontrados por AFLP

Amostra	Local	Gene	BLASTx	InterPro	SignalP	TMHMM
05 e 07	Chr02	g980	-			
		g981	-			
		g982	Fator de transcrição	<i>zinc finger</i> - IPR001138	N	N
25, 37 e 41	Chr05	g2476	3-fosfoglicerato desidrogenase	D-3-fosfoglicerato desidrogenase tipo 2 - IPR029015	N	N
		g2477	Proteína hipotética	-	N	N
		g2478	Membro LAP3 do GAL regulon	Peptidase C1B – bleomicina hidrolase - IPR004134	N	N
14	Chr08	g3507	cinesina	Cinesina motora - IPR001752 Nucleosídio trifosfato hidrolase - IPR027417	N	N
		g3508	SAICAR sintetase	SAICAR sintase/ADE2 - IPR028923	N	N
		g3509	cinesina	Cinesina motora - IPR001752 Nucleosídio trifosfato hidrolase - IPR027417	N	N
15	Chr20	g6115	glutaciona sintase	Glutaciona sintase - IPR014049, IPR016185, IPR004887, IPR014709	N	N
		g6116	soh1 – subunidade do complexo mediador	Subunidade 31 do complexo mediador - IPR008831	N	N
		g6117	Permease	Transportador triose fosfato - IPR004853	N	S
40	Chr22	g6514	Hipotética	-	N	N
		g6515	Transporte de quinato	Facilitador do transporte de açúcar/inositol - IPR020846, IPR003663	N	S
03	mtDNA	<i>cob</i>	citocromo b	Citocromo b - IPR027387, IPR016174, IPR005798	N	S
		<i>cox3</i>	Subunidade 3 da citocromo c oxidase	Citocromo c oxidase III - IPR000298	N	S