GEOVANI TOLFO RAGAGNIN

Aplicação do RBIP-qPCR em elementos de transposição adjacentes à genes de resistência para, posteriormente, verificar a ocorrência de associação entre genes de resistência e elementos de transposição.

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Biotecnologia da Universidade de São Paulo, Instituto Butantã e Instituto de Pesquisas Tecnológicas para a obtenção do Título de Mestre em Biotecnologia.

Área de Concentração: Biotecnologia

Orientador(a): Prof(a). Dr.(a) Marie-Anne Van Sluys

Versão Corrigida. A versão original eletrônica, encontra-se disponível na Secretaria de Pós-Graduação que aloja o Programa.

São Paulo

2018

RESUMO

RAGAGNIN, GEOVANI. Aplicação do RBIP-qPCR em elementos de transposição adjacentes à genes de resistência para, posteriormente, verificar a ocorrência de associação entre genes de resistência e elementos de transposição. 2018. 111 f. Dissertação (Mestrado em Interunidades em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2018.

Cana-de-açúcar é uma planta agrícola com grande importância econômica para o Brasil. Os cultivares modernos de cana-de-açúcar apresentam uma alta complexidade genômica, devido ao seu genoma ser poliploide e com número cromossômico variando entre 100-120. As plantas de cana-de-açúcar estão sujeitas ao ataque de diversos tipos de patógenos, e sua tolerância depende da ação de genes de resistência. Estes podem desencadear uma cascata de reações ou interagir diretamente com uma molécula efetora, resultando em tolerância por parte da planta à presença do patógeno. O genoma dos cultivares moderno apresentam inúmeros elementos de transposição posicionados em diferentes regiões do genoma. Devido a essa abundância, os elementos de transposição estão sendo usados como marcadores moleculares, por exemplo, RBIP (Retrotransposon-Based Insertion Polymorphism), que identifica polimorfismos baseado na inserção de elementos do tipo retrotransposons. Esta técnica é capaz de diferenciar a presença e a ausência de um elemento de transposição em uma determinada região do genoma. Para a aplicação em organismos poliploides foi desenvolvida a técnica RBIP-qPCR, permitindo dosar a presença e ausência de um elemento de transposição em um organismo poliploide. Sendo assim este trabalho tem como objetivo aplicar a técnica de RBIP-qPCR em elementos de transposição alocados adjacentes a genes de resistência a fim de, posteriormente, verificar a existência de associação entre genes de resistência e elementos de transposição. Para isso foram analisadas três famílias de genes de resistência, *I2C-2*, *Xa21* e *Pti1*. Foram realizadas análises de composição genômica de cada uma das regiões, genômica comparativa entre as variedades de R570 e SP80-3280 para a região genômica que contém o gene *I2C-2*, além de uma análise de expressão de genes e elementos de transposição vizinhos neste loco. A partir destes estudos foi possível selecionar um elemento de transposição, scDEL 5.1, para aplicar a técnica de RBIP-qPCR. Os resultados das análises de reconhecimento do ambiente genômico determinou a escolha de um elemento para aplicação da técnica. Verificamos que na ausência de scDEL 5.1 a metodologia RBIP-qPCR foi eficaz em todos os genomas testados. As análises de presença de scDEL 5.1 no loco revelaram que elementos com elevado número de cópias no genoma não são adequados para a aplicação da metodologia.

Palavras-Chave: Cana-de-açúcar. RBIP-qPCR. Genes de Resistência. Elementos de Transposição. Genômica Comparativa.

ABSTRACT

RAGAGNIN, GEOVANI. Aplicação do RBIP-qPCR em elementos de transposição adjacentes à genes de resistência para, posteriormente, verificar a ocorrência de associação entre genes de resistência e elementos de transposição.. 2018. 111 f. Dissertation (Master thesis in Interunits in Biotechnology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2018.

Sugarcane is an agricultural plant with great economic importance for Brazil. The modern cultivars of sugarcane present a high genomic complexity, due to their genome being polyploid and with chromosome number varying between 100-120. Sugarcane plants are targets of several types of pathogens, and their tolerance depends on the action of resistance genes. These can trigger a cascade of reactions or interact directly with an effector molecule, resulting in tolerance by the plant to the presence of the pathogen. The genome of modern cultivars presents several transposable elements allocated in different regions of the genome. Due to this abundance, transposable elements are being used as molecular markers, such as RBIP (Retrotransposon-Based Insertion Polymorphism), which identifies polymorphisms based on the insertion of retrotransposon. This technique is able to differentiate the presence and absence of a transposable element in a particular region of the genome. For the application in polyploid organisms the RBIP-qPCR technique was developed, allowing measuring the presence and absence of a transposable element in a polyploid organism. Therefore, the objective of this work is to apply the RBIP-qPCR technique to on a transposable element localized adjacent to resistance genes in order to, latterly, verify the existence of association between resistance genes and transposable elements. For this, three families of resistance genes, *I2C-2*, *Xa21* and *Pti1* were analyzed. Genomic composition analyzes of each region, comparative genomics between the R570 and SP80-3280 varieties for the genomic region containing the *I2C-2* gene, as well as an analysis of genes and elements expression neighbors at this locus were performed. From these studies it was possible to select a transposition element, scDEL 5.1, to apply the RBIP-qPCR technique. The results of the genomic environment recognition analysis determined the choice the transposable element for the application of the technique. We verified that in the absence of scDEL 5.1 the RBIP-qPCR methodology was effective in all genomes tested. Analysis of the presence of scDEL 5.1 in the locus revealed that elements with high copy number in the genome are not suitable for the application of the methodology.

Keywords: Sugarcane. RBIP-qPCR. Resistance Genes. Transposable Elements. Comparative Genomics.

1 INTRODUÇÃO

A cana-de-açúcar é uma gramínea pertencente à família das Poaceae, que compreende mais de 600 gêneros e 10.000 espécies (KELLOGG, 2001), classificada dentro do gênero *Saccharum* e, devido aos cruzamentos entre espécies desde o século XIX, é considerada uma planta híbrida.

A cana-de-açúcar é hoje a principal fonte de açúcar do mundo, de onde provém cerca de 75% do açúcar mundial, e atualmente se tornando uma importante fonte de bio combustível (CARRUTHER, 1994; MATSUOKA et al., 2014).

Atualmente existem três principais programas de melhoramento genético de cana-de-açúcar no Brasil, o do Instituto Agrônomo de Campinas (IAC), responsável por liberar variedades de canas com a denominação IAC. O programa de melhoramento do Centro de Tecnologia Canavieira (CTC), responsável por liberar as variedades denominadas SP e CTC, e ainda o programa de melhoramento genético de cana-de-açúcar da RIDESA, responsável por liberar as variedades denominadas RB de cana-de-açúcar (NOVACANA, 2013).

1.1 Histórico dos cultivares de Cana-De-Açúcar

A cana-de-açúcar utilizada hoje no mercado provém de cultivares modernos, sendo usada em campos de produção, além de instituições que a utilizam em programas de melhoramento, visando à melhoria da planta sobre diversas características como concentração de açúcar, resistência a patógenos, concentração de fibras além de outros atributos.

O processo de melhoramento genético de cultivares de cana-de-açúcar resulta de uma sequência de processos realizados com o objetivo de aumentar a concentração de açúcar nos colmos e também manejar a tolerância a condições de estresse biótico e abiótico. Entre eles, podemos citar o aumento da resistência a patógenos.

No século XIX agricultores na ilha de Java e na Índia produziram híbridos interespecíficos entre *Saccharum officinarum* e *Saccharum spontaneum*. A F1 resultante deste cruzamento foi retrocruzadas com *S. officinarum*. Assim plantas oriundas deste retrocruzamento apresentavam algumas características favoráveis para a produção como o alto índice de açúcar e resistência a alguns patógenos. Seus parentais apresentam diferentes

atributos, *S. officinarum* apresenta um alto conteúdo de açúcar e um baixo teor de fibras, já *S. spontaneum* apresenta um baixo teor de açúcar, porém maior índice de resistência e teor de fibras.

Devido a esses cruzamentos, os híbridos de cana-de-açúcar apresentam uma alta complexidade genômica, contendo um genoma poliploide, com um número cromossômico variando de 100-120, sendo 70-80% do genoma oriundo de *S. officinarum*, 10-23% provindo de *S. spontaneum*, e ainda tendo uma pequena porção recombinante (Figura 1) (D'HONT, 2005; PIPERIDIS; PIPERIDIS; D'HONT, 2010)

A complexidade do genoma da cana-de-açúcar torna o estudo genômico dessa planta um grande desafio. Para a maioria das variedades o tamanho total do genoma é desconhecido. Atualmente a variedade R570 foi selecionada para ter seu genoma sequenciado usando a estratégia de biblioteca genômica, e atualmente acredita-se que o tamanho do genoma dessa variedade seja de ~10Gb (D'HONT; GLASZMANN, 2001).

Figura 1: **Complexidade genômica de híbridos de cana-de-açúcar.** Barra amarela, genoma proveniente de *S. officinarum* (70-80%); Barra verde, genoma proveniente de *S. spontaneum* (10-23%); Barra amarela e verde: Porção recombinante proveniente de *S. officinarum* e *S. spontaneum*.

1.2 Variedades R570 e SP803280

Duas variedades foram amplamente utilizadas neste trabalho, sendo elas a variedade R570 e SP803280.

A variedade SP80-3280 está entre os cultivares utilizados nos programas de melhoramento do Brasil. A variedade R570 apresenta 2n: 107-115 e foi desenvolvida na ilha francesa de Réunion com fins de estudos genômicos, elaboração de um mapa genético e compreensão da interação dos complementos cromossômicos dada sua natureza híbrida. Em 2008, durante um encontro internacional entre Austrália, Brasil, Estados Unidos e França,

organizada na FAPESP, esta variedade foi selecionada para dar início ao sequenciamento do genoma. A escolha levou em consideração diversos aspectos agronômicos desta variedade, tais como: Elevada capacidade de desenvolvimento de soqueiras, alto conteúdo de fibras, um bom teor de açúcar, ampla área de adaptação e plantação, uma boa capacidade de germinação e ainda uma boa colheita mecânica (eRcane).

Para as análises genéticas na variedade R570, foi construída uma biblioteca genômica usando BACs (do inglês - *Bacterial Artificial Chromosome*). Foram utilizadas cepas bacterianas de *E. coli* denominadas DH10B e plasmídeos chamados pBeloBac11 (Tomkins, *et al.* 1999).

Já a variedade SP80-3280 foi desenvolvida pelo Centro de Tecnologia Canavieira e ganhou destaque nos estudos realizados. A variedade SP80-3280 tem como sua principal característica o alto teor de sacarose e a produtividade em soqueira. Além disso, esta variedade apresenta resistência a vários patógenos como: carvão, mosaico e ferrugem. Estudos comprovaram ainda um comportamento diferencial desta variedade quando comparada às outras desenvolvidas pelo CTC em relação à aplicação de herbicidas, mostrando um perfil mais tolerante neste aspecto (Ferreira, 2010, Coopercucar União).

1.3 Genes de resistência

As plantas são diariamente desafiadas por uma vasta gama de estresses bióticos e abióticos. O presente estudo foca em genes associados primariamente a respostas para conter estresses bióticos, como insetos, nematoides, infestações por fungos, bactérias ou ainda infecções virais. Para sobreviver a estes estresses, as plantas necessitam de um mecanismo de defesa robusto e bem conservado, ativado através de sinais de reconhecimento do ataque, que possam resultar em proteção da planta (BECKERS; SPOEL, 2006).

Devido à imensa gama de estresses bióticos as plantas estão constantemente desenvolvendo barreiras contra estes agentes, físicas, químicas ou moleculares, motivando uma extensa quantidade de trabalhos dirigidos à compreensão dos mecanismos de defesa das plantas contra estes estresses. Sendo então, a identificação e a caracterização dos genes intimamente envolvidos na defesa contra patógenos, um dos maiores desafios para a genômica das plantas (ARENCIBIA *et al.*, 1997; FALCO *et al.*, 2001 BRAGA *et al.*, 2001).

A interação entre microrganismos e plantas é determinante para o crescimento das plantas. Estas interações entre microrganismo e plantas podem ser patogênica ou não patogênica. Uma interação patogênica ocorre quando a interação resulta em redução de biomassa vegetal e aparecimento de sintomas que afetam o crescimento e desenvolvimento vegetal. Ao contrário, uma interação não patogênica resulta em crescimento/desenvolvimento sem alteração perceptível ou ainda em incremento no crescimento, como no caso das interações benéficas como a fixação de Nitrogênio atmosférico por algumas bactérias ou aporte de fosfato por fungos micorrizos (FRANCHE; LINDSTRÖM; ELMERICH, 2009; VAN DER HEIJDEN et al., 1998).

A primeira etapa de qualquer uma destas interações é o reconhecimento entre os organismos. No caso de interações patogênicas, o processo de defesa da planta é atribuído a genes denominados genes R (*Resistance Genes*). As proteínas transcritas e traduzidas a partir dos genes R podem interagir diretamente ou indiretamente com o transcrito do gene que concede a virulência do patógeno (denominados Avr genes), podendo induzir uma gama de respostas que podem resultar na tolerância ou suscetibilidade na planta e denominada por FLOR, 1971, como interação gene a gene.

Na década de 90, vários experimentos em diferentes grupos de pesquisa alavancaram inúmeros estudos sobre os genes R, aumentando número de genes R isolados e caracterizados em diversas espécies de plantas (Figura 2). A caracterização dos genes R leva em consideração os domínios proteicos contidos na sequência de aminoácidos, assim, foi elaborada uma classificação de acordo com a estrutura proteica dos genes R, sendo três grupos considerados como os principais (SCHWESSINGER; RONALD, 2012): NLR (do inglês - *Nucleotide-binding Leucine-rich Repeat*) (ou NBS-LRR), LRR e Ser/Thr Kinase (Tabela 1 e Figura 2). O primeiro e maior grupo dos genes R codifica proteínas com sítios de ligação nucleotídica (NBS– do inglês *Nucleotide Binding Site*) seguida por uma região com repetições rica em leucina (LRR– do inglês *Leucine Rich Repeat*) e ainda apresentando um domínio contendo uma espiral enrolada (CC – do inglês *Coiled Coil*). O segundo grupo dos genes R codifica proteínas contendo uma região LRR podendo também apresentar outros domínios. Já o terceiro grupo é representado exclusivamente pela família *Pto*, o qual apresenta um domínio quinase de serina-treonina, chamado de Ser/Thr Kinase. Além desses três grupos majoritários, outros grupos foram identificados e caracterizados (DIXON et al.,

1998; HAMMOND-KOSACK; JONES, 1997; MARTIN et al., 2008; SONG et al., 1995)
(Tabela 1 e Figura 2).

Figura 2: **Interação planta patógeno.** LZ/CC: Domínio Zíper de leucina/ Domínio Espiral Enrolada; NBS; Sítio de Ligação Nucleotídica LRR: Repetições ricas em leucina; TIR: Receptor de Interleucina-1 (do inglês – *Toll Interleukin -1 Receptor*); WRKY: Domínio WRKY; TD: Domínio transmembrana. Exemplificação da interação entre *Avr genes* (efetores) liberado pelos patógenos e proteínas de resistência com alguns domínios exemplificados. As proteínas de resistência podem estar situadas totalmente no meio intracelular, totalmente no meio extracelular ou parcialmente localizada no meio extra e intracelular. Alguns exemplos de proteínas de resistência, exemplificando o nome do gene, a planta em que ele foi caracterizado, o patógeno ao qual concede resistência e os domínios encontrados.

Tabela 1: Exemplificação de Genes R em plantas caracterizados até o ano de 2007. Informações como o nome do Gene R, a planta em que o gene foi isolado, o patógeno a qual o gene concede resistência e os domínios proteicos presente nos genes*.

Gene	Planta	Patógeno	Domínio Proteico	Referência
<i>RPS2</i>	<i>Arabidopsis</i>	<i>Pseudomonas syringae</i>	CC-NBS-LRR	BENT et al., 1994
<i>RPS5</i>	<i>Arabidopsis</i>	<i>Pseudomonas syringae</i>	LZ-NBS-LRR	WARREN, 1998
<i>RPP8/HRT</i>	<i>Arabidopsis</i>	<i>Pseudomonas parasitica</i>	CC-NBS-LRR	MCDOWELL, 1998
<i>RPP13</i>	<i>Arabidopsis</i>	<i>Pseudomonas parasitica</i>	CC-NBS-LRR	BITTNER-EDDY et al., 2000
<i>RCY1</i>	<i>Arabidopsis</i>	<i>Cucumber mosaic virus</i>	CC-NBS-LRR	TAKAHASHI et al., 2002
<i>RPP/HRT</i>	<i>Arabidopsis</i>	<i>Turnip crinkle virus</i>	CC-NBS-LRR	COOLEY et al., 2000
<i>RPM1</i>	Tomate	<i>Pseudomonas syringae</i>	CC-NBS-LRR	GRANT et al., 1995
<i>Prf</i>	Tomate	<i>Pseudomonas syringae</i>	CC-NBS-LRR	SALMERON et al., 1996
<i>I2</i>	Tomate	<i>Fusarium oxysporum</i>	CC-NBS-LRR	ORI et al., 1997
<i>Mi-1</i>	Tomate	<i>Meloidogyne javanica</i>	CC-NBS-LRR	MILLIGAN, 1998
<i>Mi-9</i>	Tomate	<i>Meloidogyne javanica</i>	CC-NBS-LRR	JABLONSKA et al., 2006
<i>Sw-5/Mi</i>	Tomate	<i>Topovirus</i>	CC-NBS-LRR	BROMMONSCHENKEL et al., 2000
<i>Rx2</i>	Batata	<i>Potato X virus (PVX)</i>	CC-NBS-LRR	BENDAHDANE, 1999
<i>Gpa2/Rx1</i>	Batata	<i>Globodera pallida</i> /PVX	CC-NBS-LRR	VAN DER VOSSSEN et al., 2000
<i>R1</i>	Batata	<i>Phutophthora infestans</i>	CC-NBS-LRR	BALLVORA et al., 2002
<i>Mla1</i>	Cevada	<i>Blumeria graminis</i>	CC-NBS-LRR	ZHOU et al., 2001

Tabela 1: Exemplificação de Genes R em plantas caracterizadas até o ano de 2007 (continuação).

<i>RPS4</i>	<i>Arabidopsis</i>	<i>Pseudomonas syringae</i>	TIR-NBS-LRR	GASSMANN; HINSCH; STASKAWICZ, 1999
<i>RPP1</i>	<i>Arabidopsis</i>	<i>Peronospora parasitica</i>	TIR-NBS-LRR	BOTELLA et al., 1998
<i>Rpp10</i>	<i>Arabidopsis</i>	<i>Peronospora parasitica</i>	TIR-NBS-LRR	BOTELLA et al., 1998
<i>Rpp14</i>	<i>Arabidopsis</i>	<i>Peronospora parasitica</i>	TIR-NBS-LRR	BOTELLA et al., 1998
<i>L6</i>	Linho	<i>Melanspora lini</i>	TIR-NBS-LRR	LAWRENCE et al., 1995
<i>N</i>	Tabaco	<i>Tobacco mosaic virus</i>	TIR-NBS-LRR	WHITHAM et al., 1994
<i>RRS1-R</i>	<i>Arabidopsis</i>	<i>Ralstonia solanacearum</i>	WRKY**-TIR-NBS-LRR	DESLANDES et al., 2002
<i>Piz-t</i>	Arroz	<i>Magnaporthe grisea</i>	NBS-LRR	ZHOU et al., 2006
<i>Cre1</i>	Trigo	<i>Heterodera avenuae</i>	NBS-LRR	DE MAJNIK et al., 2003
<i>I2C</i>	Tomate	<i>Fusarium oxysporum</i>	NBS-LRR	ORI et al., 1997
<i>Hero</i>	Tomate	<i>Globodera rostochiensis</i>	NBS-LRR	ERNST et al., 2002
<i>Rpg1</i>	Cevada	<i>Puccinia graminis</i>	Protein kinase	BRUEGGEMAN et al., 2002
<i>Pto</i>	Tomate	<i>Pseudomonas syringae</i>	Ser/Thr protein kinase	MARTIN et al., 2008
<i>Xa21</i>	Arroz	<i>Xanthomonas oryzae</i>	LRR-Ser/Thr protein Kinase	SONG et al., 1995
<i>Xa21D</i>	Arroz	<i>Xanthomonas oryzae</i>	LRR receptor-like protein	WANG et al., 1998
<i>Xa26</i>	Arroz	<i>Xanthomonas oryzae</i>	Receptor like protein	SUN et al., 2004
<i>RFO1</i>	<i>Arabidopsis</i>	<i>Fusarium oxysporum</i>	Receptor-like kinase	DIENER; AUSUBEL, 2005

Tabela 1: Exemplificação de Genes R em plantas caracterizados até o ano de 2007 (continuação).

<i>Cf-4</i>	Tomate	<i>Cladosporium fulvum</i>	LRR-TM	THOMAS et al., 1997
<i>Hcr9-4E</i>	Tomate	<i>Cladosporium fulvum</i>	LRR-TM	WESTERINK et al., 2004
<i>Cf-2/5</i>	Tomate	<i>Cladosporium fulvum</i>	LRR-TM	DIXON et al., 1998
<i>Ve</i>	Tomate	<i>Verticillium dahliae</i>	LRR-TM	KAWCHUK et al., 2001
<i>Mlo</i>	Cevada	<i>Erysiphe graminis</i>	TM	BÜSCHGES et al., 1997

* Retirado e modificado de: LIU et al., 2007.

** O domínio WRKY foi classificado posteriormente como uma proteína acessória integrada ao gene *RRS1-R*.

1.3.1 Modelos Gerais de Resistência em Plantas

As plantas ao longo de sua evolução desenvolveram diferentes estratégias de resistência. Três estratégias bem documentadas e estudadas envolvem principalmente a identificação do patógeno pela planta frente a moléculas liberadas pelo micro-organismo. O modelo de resistência PTI (do inglês - *PAMP-triggered immunity*) depende inicialmente do reconhecimento pela planta de padrões de moléculas associadas ao patógeno (PAMPs- *Pathogen-associated molecular patterns*) (FRAITURE; BRUNNER, 2014). A capacidade de infecção é resultado da interferência e/ou inibição da defesa da planta e, portanto, para a defesa da planta ser efetiva é necessário reconhecer as moléculas patogênicas para assim ativar a resposta de defesa contra a infecção.

Em resposta ao sistema PTI, os patógenos desenvolveram uma estratégia de infecção que interfere no reconhecimento das PAMPs interrompendo a cascata de sinalização, deixando, então, a planta suscetível à infecção. Os patógenos utilizam proteínas efetoras que são enviadas ao citoplasma da célula vegetal interrompendo a resposta imune envolvendo as PAMPs, favorecendo a colonização dos tecidos da planta (JONES; DANGL, 2006).

Uma resposta aos efetores foi detectada em algumas plantas, conhecida como ETI (do inglês - *Effector-triggered immunity*). Este modelo utilizado pelas plantas tem como base a identificação dos efetores por proteínas de resistência. Flor (1971) foi quem primeiro idealizou essa estratégia de defesa, e sua teoria abriu um novo campo de estudo. Muitos trabalhos subsequentes envolvendo o ETI foram realizados e descobriu-se que o sistema ETI é mais complexo do que o idealizado por Flor (1971) (SCHWESSINGER; RONALD, 2012).

A resistência envolvendo efetores depende de proteínas de resistência que podem ou não necessitar de iscas ou *decoys*. As iscas são na maioria das vezes proteínas independentes, porém podem estar integradas a uma proteína de resistência. Por um tempo, algumas iscas integradas foram classificadas como domínios proteicos não usuais integradas em proteínas de resistências (JONES; VANCE; DANGL, 2016). Como por exemplo, o gene *RRS1-R* que apresenta um domínio chamado WRKY (Tabela 1 e Figura 2), posteriormente identificado como isca integrado ao gene (JONES; VANCE; DANGL, 2016; MOFFETT, 2016). A seguir, três modelos de resistência são descritos.

I. Modelo PAMP-triggered immunity (PTI)

Neste modelo extensamente estudado relacionado à interação planta patógeno, o microrganismo libera do seu citoplasma moléculas que apresentam um padrão que pode ser reconhecida pela planta, a função de reconhecimento destas moléculas é dada às proteínas receptoras de reconhecimento de padrões, chamadas de PRRs (do inglês -*Pattern Recognition Receptor*), essas proteínas ao se ligarem com as PAMPs e desencadeiam uma cascata de reações que conferem resistência na planta (FRAITURE; BRUNNER, 2014) (Figura 3).

II. Modelo Effector-Triggered immunity (ETI)

Este modelo envolve proteínas efetoras liberadas pelos patógenos que impedem o reconhecimento realizado pelas PRRs, bloqueando a resistência da planta que seria gerada pelo PTI. As plantas, como resposta, desenvolveram proteínas de resistências relacionadas aos

efetores. Geralmente cada efetor produzido por um dado patógeno é reconhecido por uma proteína “R” correspondente na planta. De modo geral, para ocorrer infecção, o patógeno deve conter em seu genoma, um gene efetor, chamado de gene Avr. Já para ocorrer resistência por parte da planta, ela precisa conter um gene de resistência relacionado a este efetor, por exemplo, o gene *Pto* confere resistência ao efetor *avrPto* (MICHELMORE; CHRISTOPOULOU; CALDWELL, 2013; TOBIAS et al., 1999) (Figura 3).

III. Modelo Decoy (Decoy model) e Integrated-Decoy (ID model)

O modelo *Decoy* e o modelo ID seguem a mesma estratégia descrita acima para o modelo ETI, porém necessitam de uma proteína acessória, que tem como a principal função realizar o contato entre o efetor (patógeno) e a proteína de resistência (planta). As proteínas acessórias são modificadas pelas proteínas de resistência, assim, quando ocorre a ligação entre o efetor e a proteína acessória a proteína de resistência é ativada. De modo geral a proteína acessória tem como função mediar a interação entre os efetores (patógeno) e as proteínas efetoras (planta) (JONES; VANCE; DANGL, 2016) (Figura 3).

Algumas proteínas acessórias estão sendo identificadas integradas a transcritos provenientes de genes de resistência. Um exemplo de proteína acessória fusionada com genes de resistência é o já mencionado gene *RRS1-R* descrito originalmente em *Arabidopsis* (DESLANDES et al., 2002) além de os outros genes como *RGA5* e *Pik-1* descritos em *Oryza sativa* (MOFFETT, 2016).

É importante salientar que todos os modelos de resistência acontecem simultaneamente nas células vegetais. Desse modo é cabível pensar que os modelos de resistência citados interagem entre si. Um exemplo de interação envolve os modelos PTI e ETI, onde as proteínas efetoras presentes no modelo ETI bloqueiam a sinalização da PRR presente no modelo PTI. Assim a mesma célula vegetal pode usar diferentes estratégias de resistência, dependendo de qual molécula patogênica ela teve contato.

Figura 3: Exemplificação dos modelos PTI (*PAMP-triggered immunity*), ETI (*Effector-triggered immunity*), modelo *Decoy* e modelo *ID* (*Integrated decoy model*). No modelo PTI o patógeno libera moléculas que são reconhecidas pelas proteínas receptoras de reconhecimento (PRRs) e desencadeiam uma cascata de reações ativando o gene de resistência, transcrevendo a proteína de resistência e concedendo resistência à planta. Já no modelo ETI o patógeno libera efetores que entram na célula causando uma cascata de reações que por fim ativam genes de resistência geralmente específicos para cada efector, as proteínas de resistência concedem resistência à planta. O modelo *Decoy* e o modelo *ID* também envolvem efetores que, devido à cascata de reações, ativam genes que transcrevem proteínas de resistência e proteínas acessórias. As proteínas acessórias podem estar na mesma ORF (do inglês - *Open Reading Frame*) da proteína de resistência. Proteínas acessórias e proteínas de resistência que estiverem na mesma ORF fazem parte do modelo *ID*, já proteínas acessórias e proteínas de resistência que não estão na mesma ORF, fazem parte do modelo *Decoy*.

1.3.2 Cana-de-açúcar e genes de Resistência

Em 2001 foi publicado o banco de dados SUCEST (do inglês – *Sugarcane EST Project*) que disponibilizou uma coleção de mais de 240.000 EST (do inglês – *Expressed Sequence Tag*) (VETTORE et al., 2001) para a comunidade científica. Esta coleção revelou sequências envolvidas em diversas funções celulares, cerca de 10% dos EST estudados foram classificados como pertencentes à função de resposta ao stress, sendo essa categoria definida como sequências envolvidas na defesa da planta contra virulência e doenças (VETTORE et al.,

2003). Uma análise mais aprofundada nos EST envolvidos na resposta ao stress revelou mais de 50 transcritos contendo domínios reconhecidos como RGA (do inglês – Resistance Genes Analogs) (ROSSI et al., 2003).

Os genes de resistência na era da genômica são classificados de acordo com os domínios proteicos que apresentam, sendo que geralmente, cada domínio apresenta uma função específica. Por exemplo, o domínio LRR tem como função o reconhecimento de moléculas patogênicas e o reconhecimento de alguns efetores, já proteínas que apresentam o domínio TD (do inglês - *Transmembrane Domain*) apresenta a função de posicionar a proteína na membrana celular, podendo ter outro domínio no meio extra e/ou intracelular. Um exemplo é o gene *Cf-4*, o qual foi isolado em tomate e confere resistência ao fungo *Cladosporium fulvum* (Tabela 1), o transcrito deste gene apresenta dois domínios, um domínio TD e outro LRR, concedendo a proteína o reconhecimento de moléculas patogênicas presentes no meio extracelular (LIU et al., 2007).

No presente estudo, escolhemos aprofundar o conhecimento sobre três genes R pelos seguintes motivos: 1) cada gene está classificado em um grupo diferente, a saber: NLR, LRR, S/T Kinase (ROSSI et al., 2003); 2) os três genes de resistência já tinham sido identificados em fragmentos genômicos imobilizados em vetores BAC das variedades R570 e SP80-3280 de cana-de-açúcar, disponibilizando material para iniciar o estudo (DE SETTA et al., 2014); 3) os três genes foram mapeados em regiões genômicas específicas da variedade R570, ou seja, os três genes são locos específicos (ROSSI et al., 2003), o que é essencial para aplicar a técnica de RBIP-qPCR.

Os genes escolhidos foram o *I2C-2* (RGA 441), presente no grupo de homologia IV, pertencente ao grupo NLR e com 14 BACs disponíveis; *Xa21* (RGA 386), encontrando-se no grupo de homologia II, pertencente à classe LRR e com oito BACs disponíveis; o terceiro e último gene selecionado é o *Pti1* (RGA 169), presente no grupo de homologia II, classificado como S/T Kinase e com um BAC disponível (DE SETTA et al., 2014; ROSSI et al., 2003).

1.3.2.1 I2C-2

O primeiro gene RGA 441 (denominado neste estudo de *scI2C-2-like*) é representante da família dos genes I2, e chamado de *I2C-2* (ROSSI et al., 2003). Este gene

confere resistência a fungos da espécie *Fusarium oxysporum*, ele foi descrito e caracterizado em tomate (*Solanum lycopersicum*) (WEI et al., 2014; STALL; WALTER, 1965).

O gene *I2C-2* apresenta três domínios proteicos: o domínio chamado de espiral enrolada (CC), também chamado de Zíper de leucina (LZ, do inglês - *Leucine Zipper*) um sítio de ligação NBS e ainda um domínio LRR, a presença destes domínios definem que este gene é do grupo NLR ou NBS-LRR.

Os três domínios apresentam funções específicas. A região NBS é onde ocorre a hidrólise do ATP ou GTP que está envolvido na sinalização que aciona a resistência ao patógeno. O domínio LRR está intimamente envolvido no reconhecimento do organismo patogênico (MICHELMORE; CHRISTOPOULOU; CALDWELL, 2013). Já o domínio CC apresenta uma função relacionada à oligomerização entre uma ou mais cadeias peptídicas (LIU et al., 2007).

O gene *I2C-2* está inserido no modelo ETI, reconhecendo efetores liberados pelo *F. oxysporum*, este gene não necessita de proteína acessória para o reconhecimento do efector (CHISHOLM et al., 2006) (Figura 4).

1.3.2.2 Xa21

O segundo gene a ser estudado pertence ao grupo dos LRR, RGA 386 (denominado neste estudo de *scXa21-like*) (ROSSI et al., 2003). Este gene foi identificado e caracterizado primeiramente em arroz (*Oryza sativa*) e confere resistência a cepas da bactéria *Xanthomonas oryzae* pv. *oryzae* (SCHWESSINGER; RONALD, 2012; SONG et al., 1995). O gene *Xa21* apresenta o domínio LRR e diferentemente do gene *I2C-2*, o *Xa21* não contém o domínio NBS, o que impede o gene de pertencer ao grupo NLR. Além do domínio LRR, o gene *Xa21* apresenta um domínio transmembrana (TD) que localiza a proteína na membrana celular da planta. Na parte extracelular se encontra o domínio LRR, já na parte intracelular, encontra-se um domínio quinase, similar ao encontrado no gene *Pti1* (MICHELMORE; CHRISTOPOULOU; CALDWELL, 2013).

Bem como na proteína *I2C-2*, o domínio LRR apresenta a função de reconhecimento do patógeno, já o domínio quinase é responsável em suprimir a molécula relacionada ao patógeno, promovendo resistência (MICHELMORE; CHRISTOPOULOU; CALDWELL,

2013). *Xa21* é uma PRR, sendo responsável por reconhecer os PAMPs, estando envolvido no modelo de resistência PTI (Figura 4).

1.3.2.3 Pti1

O terceiro gene RGA 169 (denominado neste estudo de *scPti1-like*) é representante do grupo Ser/Thr Kinase e pertence à família *Pto* (ROSSI et al., 2003). O gene *Pti1* foi caracterizado e descrito no tomate selvagem (*Lycopersicon esculentum*) conferindo resistência a cepas da bactéria *Pseudomonas syringae* codificando uma proteína quinase de domínio serina/treonina (CHANG et al., 2002; ZHOU et al., 1995).

A proteína codificada pelo gene *Pti1* tem mostrado uma atividade molecular direta com a molécula liberada pela bactéria patogênica. A quinase apresenta funções de reconhecimento e de inibição da molécula do patógeno resultando na defesa da planta à proteína virulenta (MICHELMORE; CHRISTOPOULOU; CALDWELL, 2013). A atividade direta que esta proteína apresenta a enquadra no modelo de resistência ETI, não necessitando de proteínas assessórias (Figura 4).

Figura 4: **Posicionamento dos RGAs *Xa21*, *I2C-2* e *Pti1* nos modelos de resistência apresentado na figura 3.** *Xa21* é uma PRR (Proteína Receptora de Reconhecimento) presente no modelo PTI; *I2C-2* é uma proteína de resistência presente no modelo ETI; *Pti1* também é uma proteína de resistência presente no modelo ETI.

1.4 Elementos de Transposição

Elementos de Transposição (TE, do inglês - *Transposable Elements*) são sequências de nucleotídeo com capacidade de mobilização em um determinado genoma (BIÉMONT; VIEIRA, 2006).

São descritos diversos domínios associados à capacidade de mobilização dos elementos de transposição, conseqüentemente existe uma grande diversidade estrutural dos elementos de transposição, sendo necessária uma classificação a fim de organizar os elementos de transposição.

A classificação feita por FINNEGAN, 1989, organiza os elementos de transposição em duas grandes classes: elementos de Classe I ou retro elementos, e elementos da Classe II ou transposons. A distinção entre as duas classes ocorre principalmente devido ao mecanismo pelo qual o elemento realiza a sua mobilização dentro do genoma.

A classe dos retroelementos (Classe I) está constituída por elementos denominados de retrotransposons e elementos denominados retroposons. A maioria dos retrotransposons apresenta duas largas terminações repetidas (LTR, do inglês - *Large Terminal Repeat*) que se encontram nas extremidades 5' e 3' do elemento de transposição. Na região interna dessas repetições podem ser encontradas ORFs, iniciando a transcrição dos domínios Gag e Pol, que compõem o mRNA do elemento de transposição (Figura 5). O domínio Gag é responsável por sintetizar proteínas estruturais semelhantes ao capsídeo viral, compreendendo a função de empacotamento de outras proteínas envolvidas no processo de mobilização (GAO et al., 2003). O domínio Pol sintetiza algumas proteínas envolvidas na replicação do elemento de transposição, como por exemplo, proteína aspártica (AP), envolvida no processamento pós-traducional do domínio Pol, transcriptase reversa (RT) e RNaseH, que são encarregados da realização da transcrição reversa do mRNA, e ainda a proteína integrase (IN) que apresenta a função de inserir o novo elemento no genoma (SABOT; SCHULMAN, 2006). Desse modo, os elementos de transposição de Classe I realizam sua transposição através de um intermediário de RNA (WICKER et al., 2007).

Figura 5: **Esquema estrutural de um elemento genérico de Classe I pertencente à Superfamília Copia.** Retângulos mostra regiões codificantes. POL indica região onde é transcrita uma série de proteínas.

Os retroposons presentes na classe dos retro elementos se diferem dos retrotransposons por não apresentarem LTRs, os chamados LINEs (do inglês - *Long Interspersed Nuclear Element*) e SINEs (do inglês - *Short Interspersed Nuclear Element*). Os retroposons são divididos em ordens. A ordem SINE apresenta entre 80 e 400 pares de base (pb) e necessitam da maquinaria de um elemento autônomo para se transpor (OSTERTAG; JR., 2001) sendo considerados elementos não autônomos, por não apresentar a maquinaria necessária para sua transposição.

A ordem LINE é considerada autônoma e diferentemente dos SINEs, os LINEs podem apresentar mais de 1000 pb de comprimento. Os LINEs também codificam proteínas para a mobilização, cada família presente nessa ordem é classificada de acordo com as proteínas transcritas (WICKER et al., 2007).

Os transposons apresentam uma estrutura mais simples quando comparados aos retrotransposons. Os elementos de Classe II apresentam nas extremidades 5' e 3' regiões contendo repetições terminais invertidas (TIRs, do inglês - *Terminal Inverted Repeats*). Em sua região interna, os transposons geralmente apresentam uma única ORF, transcrevendo uma proteína chamada de transposase (Figura 6). A enzima transposase é responsável em promover a atividade catalítica sobre a molécula de DNA permitindo a mobilização do elemento. A transposase reconhece as TIRs do elemento de transposição promovendo a excisão e subsequente inserção em outro lugar no genoma (DE JESUS et al., 2012; WICKER et al., 2007).

Figura 6: Esquema representativo de um elemento genérico pertencente à Classe II.

1.4.1 Retrotransposons e Plantas

WICKER et al., 2007, propuseram um novo sistema de classificação para os elementos mantendo as duas classes já existentes criadas por Finnegan, 1989, organizando os elementos em novos clados inferiores, como subclasse, ordem, superfamília, família e em alguns casos o clado subfamília, facilitando a classificação dos elementos de transposição.

A alta representatividade dos elementos de transposição de Classe I nos genomas de diferentes grupos taxonômicos é reflexo do seu mecanismo de transposição, que resulta no aumento do número de cópias no genoma.

A representatividade dos retrotransposons em plantas é particularmente elevada, chegando a uma porcentagem de 80% no genoma do milho. Por consequência de sua alta representatividade, acabaram por ser usadas como parte de ferramentas moleculares, apresentando diversas finalidades (KALENDAR et al., 2011; MEYERS et al., 2001). Existem várias ferramentas que utilizam elementos de transposição de Classe I como marcadores moleculares. Uma delas se denomina RBIP (do inglês - *Retrotransposon-Based Insertional Polymorphism*), desenvolvida por FLAVELL et al., 1998, e é usada para a verificação de polimorfismo baseado na inserção de um determinado retrotransposon, sendo o único método desenvolvido para detectar diferentes doses baseado na inserção de um retrotransposons íntegro em um locus específico.

1.4.1.1 Modelo RBIP

A metodologia RBIP tem a finalidade de revelar a presença e/ou ausência de um determinado retrotransposon em uma localidade específica no genoma do organismo. Em genomas diploides o uso da metodologia permite informar se o elemento está presente em um locus específico nas duas cópias do genoma, ou ausente nas duas cópias, ou ainda presente em

uma cópia e ausente na outra, em outras palavras a metodologia é capaz de diferenciar estados alélicos no genoma do organismo (KUMAR; HIROCHIKA, 2001).

Para o uso desse modelo é preciso conhecer o elemento de transposição em que a técnica será aplicada, bem como sua região flanqueadora (FLAVELL et al., 1998).

Dessa maneira é possível revelar se o organismo apresenta diferentes doses baseado na inserção do elemento de transposição, porém, um problema é encontrado quando tentamos migrar a utilização desta metodologia para organismos poliploides.

A técnica de RBIP foi desenvolvida visando a aplicação em organismos diploides, a aplicação dela em organismos poliploides apenas responderá se o elemento de transposição está ou não presente no locus, porém ela não responderá quantas regiões estão ocupadas pelo retroelemento, e quantas regiões estão vazias.

Para responder essa pergunta no genoma complexo de cana-de-açúcar, METCALFE et al., 2015) adaptaram a técnica RBIP-qPCR usando o elemento LTR-RT scIvana.

1.4.1.2 RBIP-qPCR

METCALFE et al., 2015, adiciona ao RBIP o PCR quantitativo, podendo assim quantificar a presença e a ausência do elemento de transposição em uma determinada localização no genoma.

Usando a tecnologia Taqman® junto com o modelo RBIP é possível dosar a presença e ausência de elementos de transposição em determinadas regiões cromossômicas de um organismo poliploide.

O modelo RBIP-qPCR utiliza um conjunto de iniciadores de presença e ausência junto com duas sondas. Sendo realizadas duas reações de qPCR para cada região genômica estudada, uma destinada a quantificar a presença do elemento de transposição e outra reação destinada a determinar a ausência do mesmo elemento, permitindo quantificar a presença e ausência elemento em um determinado locus.

1.4.1.2.1 Aplicação da Técnica

A técnica desenvolvida permitiu investigar a inserção do elemento scIvana em três regiões distintas em diversas variedades de cana-de-açúcar do Brasil e da Austrália (Figura 7). A técnica foi capaz de dosar a composição dos loci contendo os elementos scIvana1.1, scIvana1.2 e scIvana1.4, utilizando os elementos de transposição como marcadores moleculares entre as variedades de cana-de-açúcar (Figura 6).

Figura 7: **Gráfico 3D ilustrando a dosagem dos elementos scIvana1.2, scIvana1.1 e scIvana1.4 em diversas variedades de cana-de-açúcar.** Retirado de METCALFE et al., 2015. A posição dos círculos revela a presença dos elementos nos três alelos. Verde, variedades desenvolvidas pela RIDESA; Amarelo, variedades desenvolvidas pelo CTC (Centro de Tecnologia Canavieira); Cinza, *Q canes* do SRA (*Sugar Research Australia*); Azul, população segregante provenientes da RIDESA.

2 CONCLUSÃO

Ao total 34 BACs foram analisados. A média de genes por pares de base corrobora com o que foi encontrado em DE SETTA et al., 2014. Porém, nosso trabalho demonstrou que 30% do espaço dos BACs estão ocupados por elementos de transposição, enquanto que DE SETTA et al., 2014 encontrou um espaço de ocupação perto dos 50%. Esta divergência pode estar relacionada à diferença de BACs analisados entre os trabalhos. Também foram encontrados possíveis elementos de transposição associados a cada conjunto de BACs. Nos BACs com o gene *scXa21-like* foram identificados nove cópias do elemento Reina 2.6 e seis cópias do elemento de transposição Ale 6.1, para os BACs com o gene *scPti-like* foi identificado três cópias do elemento Del 1.1, já para os BACs com o gene *scI2C-2-like* foi identificado cinco cópias do elemento Del 5.1 e cinco cópias do elemento Maximus 1.6, além de um maior número de elementos de classe II.

A genômica comparativa de *I2C-2* revela elevada sintonia entre os BACs de R570 e SP80-3280. Esta elevada conservação na organização genômica sugere que a região está sob forte seleção de manutenção de sua organização. Inclusive, foram identificadas inserções de elementos de transposição compartilhadas pelas duas variedades. Do mesmo modo, foi possível identificar nos genomas de sorgo, milho e arroz a mesma organização dos genes vizinhos (K,D e A).

A análise filogenética revelou a relação das proteínas *I2C-2* e *Xa21* com outras proteínas de resistência. Demonstrando que a proteína *I2C-2* apresenta uma divergência baixa entre as cópias. Já *Xa21* apresentou uma alta divergência proteica entre as cópias, sinalizando, inclusive, proteínas erroneamente classificadas como *Xa21*. As divergências proteicas das duas proteínas refletem à função da proteína quando inseridas no contexto dos modelos de resistência. *Xa21* apresenta uma função mais generalista, sendo uma PRR e com função atribuída ao PTI, já *I2C-2* apresenta uma função mais específica sendo uma proteína de resistência inserida no modelo ETI.

Em especial, o gene glutamato 1-semialdeído aminotransferase apresentou um índice FPKM alto, nas quatro condições analisadas. Este índice elevado corresponde à função do gene, que está relacionada na produção da clorofila.

A reação de ausência revelou que, entre as variedades, o número de alelos sem o elemento de transposição varia, indicando que a ausência do elemento de transposição é diferente entre as variedades de cana-de-açúcar.

A aplicação do RBIP-qPCR sobre o elemento de transposição scDel 5.1 também contribuiu para adicionarmos algumas limitações à aplicação da técnica no genoma de cana-de-açúcar. Uma limitação referente à escolha do elemento e outra referente à síntese de sondas e iniciadores.

A primeira limitação está intimamente ligada número de inserções do elemento de transposição. Compilando dados desta dissertação com dados de METCALFE et al., 2015, constatamos que elementos com mais de 2000 inserções reflete na indeterminação do valor de CT de presença devido à diluição do iniciador da presença em mais de 2000 elementos Del 5.1 distribuídos pelo genoma, já elementos de transposição com menos 600 inserções, como o elemento Ivana, não reflete em valores de CT não determinados.

A segunda limitação está relacionada tanto ao tamanho da região a ser amplificada, quando relacionado à distância entre o iniciador e a sonda. A PCR de ausência demonstrou que a distância entre a sonda e o iniciador equivalente a 80 pares de base não reflete na indeterminação do valor de CT. Já a PCR de presença foi confeccionada em uma distância entre a sonda e o iniciador equivalente a 290 pares de base, podendo contribuir para uma reação de qPCR com valores de CT indeterminados.

Portanto a técnica de RBIP-qPCR apresenta, no mínimo, mais duas limitações para obter êxito na sua aplicação em organismos poliploides.

REFERÊNCIAS

- ALJANABI, S.; FORGET, L.; DOOKUN, A. An improved and rapid protocol for the isolation of polysaccharide-and polyphenol-free sugarcane DNA. **Plant Molecular Biology Reporter**, v. 17, p. 1–8, 1999.
- ARENCIBIA, A et al. Transgenic sugarcane plants resistant to stem borer attack. **Molecular Breeding**, v. 3, n. 4, p. 247–255, 1997.
- BALLVORA, A. et al. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. **Plant Journal**, v. 30, n. 3, p. 361–371, 2002.
- BECKERS, G. J. M.; SPOEL, S. H. Fine-tuning plant defence signalling: Salicylate versus jasmonate. **Plant Biology**, v. 8, n. 1, p. 1–10, 2006.
- BENDAHMANE, A. The Rx Gene from Potato Controls Separate Virus Resistance and Cell Death Responses. **The Plant Cell Online**, v. 11, n. 5, p. 781–792, 1999.
- BENT, F. A. et al. RPS2 of *Arabidopsis thaliana*: A Leucine-Rich Repeat Class of Plant Disease Resistance Genes. **Science**, v. 1, n. 1, p. 1856-1860, 1994.
- BIÉMONT, C.; VIEIRA, C. Genetics: Junk DNA as an evolutionary force. **Nature**, v. 443, n. 7111, p. 521–524, 2006.
- BITTNER-EDDY, P. D. et al. RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. **Plant Journal**, v. 21, n. 2, p. 177–188, 2000.
- BOTELLA, M. A et al. Three genes of the *Arabidopsis* RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. **The Plant cell**, v. 10, n. 11, p. 1847–60, 1998.
- BROMMONSCHENKEL, S. H. et al. The Broad-Spectrum Tospovirus Resistance Gene *Sw-5* of Tomato Is a Homolog of the Root-Knot Nematode Resistance Gene *Mi*. **Molecular Plant-Microbe Interactions**, v. 13, n. 10, p. 1130–1138, 2000.

BRUEGGEMAN, R. et al. The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. **Proceedings of the National Academy of Sciences**, v. 99, n. 14, p. 9328–9333, 2002.

BÜSCHGES, R. et al. The barley Mlo gene: a novel control element of plant pathogen resistance. **Cell**, v. 88, n. 5, p. 695–705, 1997.

CARRUTHER, S.; MILLER, P.; VAUGHAN, C. **Solid biofuels: fuel crops**. In: CARRUTHER, S.; MILLER, P.; VAUGHAN, C. Crops for Industry and Energy. 1. United Kingdom Reading: Centre for Agricultural Strategy, University of Reading, 168–180. 1994.

CHANG, J. H. et al. Functional Analyses of the Pto Resistance Gene Family in Tomato and the Identification of a Minor Resistance Determinant in a Susceptible Haplotype. **Molecular plant-microbe interactions**, v. 15, n. 3, p. 281–291, 2002.

CHISHOLM, S. T. et al. Host-microbe interactions: Shaping the evolution of the plant immune response. **Cell**, v. 124, n. 4, p. 803–814, 2006.

COOLEY, M. B. et al. Members of the Arabidopsis HRT/RPP8 Family of Resistance Genes Confer Resistance to Both Viral and Oomycete Pathogens. **The Plant Cell**, v. 12, n. 5, p. 663, 2000.

D'HONT, A. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. **Cytogenetic and Genome Research**, v. 109, n. 1–3, p. 27–33, 2005.

D'HONT, A.; GLASZMANN, J. C. Sugarcane genome analysis with molecular markers: a first decade of research. **Proc Int Soc Sugar Cane Technol**, v. 24, p. 556–559, 2001.

DE ARAUJO, P. G. et al. Transcriptionally active transposable elements in recent hybrid sugarcane. **Plant Journal**, v. 44, n. 5, p. 707–717, 2005.

DE JESUS, E. M. et al. Diversification of hAT transposase paralogues in the sugarcane genome. **Molecular Genetics and Genomics**, v. 287, n. 3, p. 205–219, 2012.

DE MAJNIK, J. et al. The cre1 and cre3 nematode resistance genes are located at homeologous loci in the wheat genome. **Molecular plant-microbe interactions**, v. 16, n. 12, p. 1129–34, 2003.

DE SETTA, N. et al. Building the sugarcane genome for biotechnology and identifying evolutionary trends. **BMC Genomics**, v. 540, n. 15, 2014.

DESLANDES, L. et al. Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. **Proceedings of the National Academy of Sciences**, v. 99, n. 4, p. 2404–2409, 2002.

DIDENKO, V. V. Dna probes using fluorescence resonance energy transfer (fret): Designs and applications *BioTechniques*. **Biotechniques**, v. 31, n. 31, p. 1106–1121, 2001.

DIENER, A. C.; AUSUBEL, F. M. Resistance to *Fusarium oxysporum* 1, a dominant *Arabidopsis* disease-resistance gene, is not race specific. **Genetics**, v. 171, n. 1, p. 305–321, 2005.

DIXON, M. S. et al. The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. **Plant Cell**, v. 10, n. 11, p. 1915–1925, 1998.

DOMINGUES, D. S. et al. Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns. **BMC Genomics**, v. 137, n. 13, 2012.

ERCANE. Disponível em: <<http://www.ercane.re/>>. Acesso em: 15 Jun. 2015.

ERNST, K. et al. The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. **Plant Journal**, v. 31, n. 2, p. 127–136, 2002.

FALCO, M. C. et al. Mechanisms of sugarcane response to herbivory. **Genetics and Molecular Biology**, v. 24, n. 1–4, p. 113–122, 2001.

FERREIRA R. R. et al. Tolerância diferencial de variedades de cana-de-açúcar a estresse por herbicidas. **Bragantia**, v. 69, p. 395–404, 2010.

FINNEGAN, D. J. Transposable elements and Genome evolution. **Elsevier Science Publisher**, v. 5, n. 4, p. 103–107, 1989.

FLAVELL, A. J. et al. Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. **Plant Journal**, v. 16, n. 5, p. 643–650, 1998.

FLOR, H. H. Current status of the gene-for-gene concept. **Annu. Rev. Phytopathol.**, v. 9, p. 275-296, 1971.

FRAITURE, M.; BRUNNER, F. Killing two birds with one stone: Trans-kingdom suppression of PAMP/MAMP-induced immunity by T3E from enteropathogenic bacteria. **Frontiers in Microbiology**, v. 5, 2014.

FRANCHE, C.; LINDSTRÖM, K.; ELMERICH, C. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. **Plant and Soil**, v. 321, n. 1-2, p. 35-59, 2009.

GAO, X. et al. Translational recoding signals between gag and pol in diverse LTR retrotransposons. **RNA**, v. 9, n. 12, p. 1422-30, 2003.

GASSMANN, W.; HINSCH, M. E.; STASKAWICZ, B. J. The Arabidopsis RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. **Plant Journal**, v. 20, n. 3, p. 265-277, 1999.

GRANDBASTIEN, M. A. et al. Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. **Cytogenetic and Genome Research**, v. 110, n. 1-4, p. 229-241, 2005.

GRANT, M. R. et al. Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. **Science**, v. 269, n. 5225, p. 843-846, 1995.

GYPSY DATABASE 2.0. The Gypsy Database of Mobile Genetic Elements. 2011. Disponível em: <http://www.gydb.org/index.php/Main_Page>. Acesso em 10 jul. 2015.

HALL, B. G. P. Phylogenetic Trees Made Easy: A How-To Manual. 4^o Edition. USA: SINAUER, 2011.

HALTERMAN, D. et al. The MLA6 coiled-coil, NBS-LRR protein confers AvrMla6-dependent resistance specificity to Blumeria graminis f. sp. hordei in barley and wheat. **Plant J**, v. 25, p. 335-348, 2001.

HAMMOND-KOSACK, K. E.; JONES, J. D. G. Plant disease resistance genes. **Annu. Rev. Plant Physiol. Plant Mol. Biol.**, v. 48, p. 575-607, 1997.

JABLONSKA, B. et al. The Mi-9 Gene from Solanum arcanum Conferring Heat-Stable Resistance to Root-Knot Nematodes Is a Homolog of Mi-1. **Plant Physiology**, v. 143, n. 2, p.

1044–1054, 2006.

JONES, J. D. G.; DANGL, J. L. The plant immune system. **Nature**, v. 444, n. 7117, p. 323–329, 2006.

JONES, J. D. G.; VANCE, R. E.; DANGL, J. L. Intracellular innate immune surveillance devices in plants and animals. **Science**, v. 354, n. 6316, 2016.

KALENDAR, R. et al. Analysis of plant diversity with retrotransposon-based molecular markers. **Heredity**, v. 106, p. 520-530, 2011.

KANNANGARA, G. C.; GOUGH, S. P. Biosynthesis of Δ -aminolevulinate in greening barley leaves: Glutamate 1-semialdehyde aminotransferase. **Carlsberg Research Communications**, v. 43, n. 3, p. 185–194, 1978.

KAWCHUK, L. M. et al. Tomato Ve disease resistance genes encode cell surface-like receptors. **Proceedings of the National Academy of Sciences**, v. 98, n. 11, p. 6511–6515, 2001.

KELLOGG, E. A. Evolutionary History of the Grasses. **Plant Physiology**, v. 125, n. 3, p. 1198–1205, 2001.

KOMOR, E. Susceptibility of sugarcane, plantation weeds and grain cereals to infection by Sugarcane yellow leaf virus and selection by sugarcane breeding in Hawaii. **Eur J Plant Pathol**, v. 129, p. 379-388, 2011.

KUMAR, A.; BENNETZEN, J. L. Plant retrotransposons. **Annu, Rev. Genet.**, v. 33, n. 1, p. 479–532, 1999.

KUMAR, A.; HIROCHIKA, H. Applications of retrotransposons as genetic tools in plant biologyTrends in Plant Science. **Trends in plant Science**, v. 6, n. 3, p. 127-134, 2001.

LAWRENCE, G. J. et al. The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. **The Plant Cell**, v. 7, n. 8, p. 1195–1206, 1995.

LIU, J. et al. Recent Progress in Elucidating the Structure, Function and Evolution of Disease Resistance Genes in Plants. **Journal of Genetics and Genomics**, v. 34, n. 9, p. 765-776, 2007.

LOPES, F. C. C. Mapeamento genético de cana-de-açúcar (*Saccharum* spp.) por associação empregando marcadores SRR e AFLP. 2011.

MARTIN, G. B. et al. Map-Based Cloning of a Protein Kinase Gene Conferring Disease Resistance in Tomato. **Advancement Of Science**, v. 262, n. 5138, p. 1432–1436, 2008.

MATSUOKA, S. et al. Energy Cane: Its Concept, Development, Characteristics, and Prospects. **Advances in Botany**, p. 1–13, 2014.

MCDOWELL, J. M. Intragenic Recombination and Diversifying Selection Contribute to the Evolution of Downy Mildew Resistance at the RPP8 Locus of Arabidopsis. **The Plant Cell Online**, v. 10, n. 11, p. 1861–1874, 1998.

METCALFE, C. J. et al. Using quantitative PCR with retrotransposon-based insertion polymorphisms as markers in sugarcane. **Journal of Experimental Botany**, 2015.

MEYERS, B. C. et al. Abundance, Distribution, and Transcriptional Activity of Repetitive Elements in the Maize Genome. **Genome Research**, v. 11, p. 1660–1676, 2001.

MICHELMORE, R. W.; CHRISTOPOULOU, M.; CALDWELL, K. S. Impacts of Resistance Gene Genetics, Function, and Evolution on a Durable Future. **Annual Review of Phytopathology**, v. 51, p. 291–319, 2013.

MILLIGAN, S. B. The Root Knot Nematode Resistance Gene Mi from Tomato Is a Member of the Leucine Zipper, Nucleotide Binding, Leucine-Rich Repeat Family of Plant Genes. **The Plant Cell Online**, v. 10, n. 8, p. 1307–1320, 1998.

MOFFETT, P. Using Decoys to Detect Pathogens: An Integrated Approach. **Trends in Plant Science**, v. 21, n. 5, p. 369–370, 2016.

MORTAZAVI, A. et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, v. 5, n. 7, p. 621–628, 2008.

NEW ENGLAND BIOLABS® INC. New England Biolabs 2013/2014: Catolg & Technical Reference. USA. New England Biolabs, INC. 2013/2014. p. 26–47.

NOVACANA. 2013. Disponível em: < <https://www.novacana.com/>>. Acesso em 22 out. 2016.

ORI, N. et al. The 12C Family from the Wlt Disease Resistance Locus 1 2 Belongs to the Nucleotide Binding, Leucine-Rich Repeat Superfamily of Plant Resistance Genes. **The Plant Cell**, v. 9, p. 521–532, 1997.

OSTERTAG, E. M.; Jr., H. H. K. Biology of Mammalian L1. **Annual Review of Genetics**, v. 35, p. 501–538, 2001.

PIPERIDIS, G.; PIPERIDIS, N.; D'HONT, A. Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. **Molecular Genetics and Genomics**, v. 284, n. 1, p. 65–73, 2010.

PREDIGER, E. Designing PCR primers and probes. 21 out. 2013. Disponível em: < <http://www.idtdna.com/pages/decoded/decoded-articles/pipet-tips/decoded/2013/10/21/designing-pcr-primers-and-probes>>. Acesso em: 15 jun. 2016.

RAMAKERS, C. et al. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, v. 339, n. 1, p. 62–66, 2003.

REPBASE. Genetic Information Research Institute. 2001. Disponível em: < <http://www.girinst.org/rebase/>>. Acesso em: 30 jun. 2015.

ROSSI, M. et al. Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. **Molecular Genetics and Genomics**, v. 249, p. 406–419, 2003.

SABOT, F.; SCHULMAN, A. H. Parasitism and the retrotransposon life cycle in plants: A hitchhiker's guide to the genome. **Heredity**, v. 97, n. 6, p. 381–388, 2006.

SALMERON, J. M. et al. Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. **Cell**, v. 86, n. 1, p. 123–133, 1996.

SCHAKER, P. D. C. et al. RNAseq transcriptional profiling following whip development in sugarcane smut disease. **PLoS ONE**, v. 11, n. 9, p. 1–21, 2016.

SCHWESSINGER, B.; RONALD, P. C. Plant Innate Immunity: Perception of Conserved Microbial Signatures. **Annual Review of Plant Biology**, v. 63, p. 451–482, 2012.

SEKHWAL, M. K. et al. Disease resistance gene analogs (RGAs) in plants. **International Journal of Molecular Sciences**, v. 16, n. 8, p. 19248–19290, 2015.

SIMPSON, D. J.; KNOETZEL, J. **Light-harvesting Complexes of Plants and Algae: Introduction, Survey and Nomenclature**. In: ORT D. R.; YOCUM, C. F. Oxygenic Photosynthesis: The light Reactions. 1. Kluwer Academic Publisher, Printed in The Netherlands, p. 493–506, 1996.

SONG, W. et al. A Receptor Kinase-Like Protein Encoded by the Rice Disease Resistance Gene, Xa21. **Science**, v. 207, n. 5243, p. 1804-1806, 1995.

STALL, R. E.; WALTER, J. M. Selection and inheritance of resistance in Tomato to isolates of races 1 and 2 of the Fusarium wilt organism. **Phytopathology**, v. 55, n. 11, p.1213–1215, 1965.

SUN, X. et al. Xa26, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. **Plant Journal**, v. 37, n. 4, p. 517–527, 2004.

TAKAHASHI, H. et al. RCY1, an *Arabidopsis thaliana* RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. **Plant Journal**, v. 32, n. 5, p. 655–667, 2002.

THOMAS, C. M. et al. Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. **The Plant Cell**, v. 9, n. 12, p. 2209–2224, 1997.

TOBIAS, C. M. et al. Plants expressing the Pto disease resistance gene confer resistance to recombinant PVX containing the avirulence gene AvrPto. **Plant Journal**, v. 17, n. 1, p. 41-50, 1999.

TOMKINS J. P. et al. A bacterial artificial chromosome library for sugarcane. **Theor Appl Genet**. v. 99 p. 419-424, 1999.

TRAPNELL, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nature Biotechnology**, v. 28, n. 5, p. 511–515, 2010.

UNICA. União Da Indústria De Cana-de-açúcar. Disponível em: <<http://www.unicadata.com.br/>>. Acesso em: 28 jun. 2015.

VAN DER HEIJDEN, M. G. A. et al. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. **Nature**, v. 396, n. 6706, p. 69–72, 1998.

VAN DER VOSSEN, E. A. G. et al. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: A virus and a nematode. **Plant Journal**, v. 23, n. 5, p. 567–576, 2000.

VETTORE, A. L. et al. The libraries that made SUCEST. **Genetics and Molecular Biology**, v. 24, n. 1–4, p. 1–7, 2001.

VETTORE, A. L. et al. Analysis and Functional Annotation of an Expressed Sequence Tag Collection for Tropical Crop Sugarcane Analysis and Functional Annotation of an Expressed Sequence Tag Collection for Tropical Crop Sugarcane. **Genome Research**, v. 13, p. 2725–2735, 2003.

VON WETTSTEIN, D.; GOUGH, S.; KANNANGARA, C. G. Chlorophyll Biosynthesis. **The Plant cell**, v. 7, n. 7, p. 1039–1057, 1995.

WANG, G. L. et al. Xa21D encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. **Plant Cell**, v. 10, n. 5, p. 765–779, 1998.

WARREN, R. F. A Mutation within the Leucine-Rich Repeat Domain of the Arabidopsis Disease Resistance Gene RPS5 Partially Suppresses Multiple Bacterial and Downy Mildew Resistance Genes. **The Plant Cell Online**, v. 10, n. 9, p. 1439–1452, 1998.

WEI, C. et al. The I2 resistance gene homologues in Solanum have complex evolutionary patterns and are targeted by miRNAs. **BMC Genomics**, v. 15, n. 743, 2014.

WESTERINK, N. et al. Cladosporium fulvum circumvents the second functional resistance gene homologue at the Cf-4 locus (Hcr9-4E) by secretion of a stable avr4E isoform. **Molecular Microbiology**, v. 54, n. 2, p. 533–545, 2004.

WHITHAM, S. et al. The Product of the Tobacco Mosaic Virus Resistance Gene Al : Similarity to Toll and the Interleukin-1 Receptor. **Cell**, v. 79, p. 1101–1115, 1994.

WICKER, T. et al. A unified classification system for eukaryotic transposable elements. **Nature Reviews Genetics**, v. 8, n. 12, p. 973–982, 2007.

ZHOU, B. et al. The eight amino-acid differences within three leucine-rich repeats between Pi2 and Piz-t resistance proteins determine the resistance specificity to *Magnaporthe grisea*. **Molecular plant-microbe interactions**, v. 19, n. 11, p. 1216–28, 2006.

ZHOU, F. et al. Cell-Autonomous Expression of Barley Mla1 Confers Race-Specific Resistance to the Powdery Mildew Fungus via a Rar1-Independent Signaling Pathway. **The Plant Cell Online**, v. 13, n. 2, p. 337–350, 2001.

ZHOU, J. et al. The tomato gene Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. **Cell**, v. 83, n. 6, p. 925–935, 1995.