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

Mapa de ligação e mapeamento de locos quantitativos de resistência a *Sporisorium scitamineum* em cana-de-açúcar

Taislene Butarello Rodrigues de Morais

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

Piracicaba
2013

Taislene Butarello Rodrigues de Morais
Engenheira Agrônoma

Mapa de ligação e mapeamento de locos quantitativos de resistência a *Sporisorium scitamineum* em cana-de-açúcar

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Profa. Dra. **MARIA LUCIA CARNEIRO VIEIRA**

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

Piracicaba
2013

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

Morais, Taislene Butarello Rodrigues de
Mapa de ligação e mapeamento de locos quantitativos de resistência a *Sporisorium scitamineum* em cana-de-açúcar / Taislene Butarello Rodrigues de Morais.- - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2013.
152 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Cana-de-açúcar 2. Carvão (Doença de planta) 3. Fungos fitopatogênicos 4. Mapeamento genético 5. Marcadores moleculares 6. Melhoramento Genético Vegetal I. Título

CDD 633.61
M827m

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

DEDICATÓRIA

Dedico este trabalho ao meu esposo Alexandre,

e aos meus pais Bernadete e Djalma

AGRADECIMENTOS

A Deus, por toda graça recebida e por permitir mais esta conquista.

À Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ/USP), pela oportunidade de receber excelentes ensinamentos e, em especial, ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas pelo suporte acadêmico oferecido.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, pelo apoio financeiro concedido para a realização do curso.

À professora Maria Lucia Carneiro Vieira, pela oportunidade oferecida e orientação, pelos ensinamentos e confiança em mim depositada, por quem aprendi a ter uma grande admiração.

A todos os professores e funcionários do Departamento de Genética, pelos ensinamentos e a amizade.

Ao meu esposo Alexandre, que com muito amor, carinho, apoio, paciência e companheirismo compartilhou comigo esta conquista.

Ao Carlinhos, pela amizade e otimismo durante a minha estadia no Laboratório de Genética Molecular de Plantas Cultivadas.

Aos amigos do laboratório Alessandra, Anselmo, Augusto, Bruna, Bruno, Carmelice, Cássia, Guilherme Pereira, Helen, Zirlane e também a Larissa Castro, Guilherme Farias, Luciana Chaves, Melina, Pedro Mendonza, Bruno Marco de Lima e Marcelo Mollinari, pela amizade e pelo apoio. Agradeço, em especial, as grandes amigas Luciane Santini e Carla Munhoz, pelo companheirismo, pela amizade, pelas trocas de conhecimento e pelo apoio constante.

Aos amigos Rodrigo Gazaffi e Maria Marta Pastina, pelos conhecimentos compartilhados e pela agradável convivência durante esta etapa.

Aos meus pais, Bernadete e Djalma, meus irmãos, Greici e Elber e aos meus sobrinhos Gustavo e Gabriela, pelo amor e pelos incentivos a prosseguir e lutar pelos meus objetivos.

Ao meu sogro José Lazaro e a minha sogra Maria Elena e meus cunhados Patrícia, Rodrigo e Adriana, pelo carinho e torcida.

A todos, que de certa forma contribuíram para o êxito deste trabalho, meus sinceros agradecimentos.

EPÍGRAFE

“Existem muitas hipóteses em ciência que estão erradas. Isso é perfeitamente aceitável, elas são a abertura para achar as que estão certas.”

Carl Sagan

SUMÁRIO

RESUMO.....11

ABSTRACT.....13

LISTA DE FIGURAS.....15

LISTA DE TABELAS.....19

1 INTRODUÇÃO.....21

2 REVISÃO BIBLIOGRÁFICA.....23

2.1 Aspectos econômicos da cultura da cana-de-açúcar.....23

2.2 A origem do gênero *Saccharum*.....23

2.3 Hibridação interespecífica artificial e estrutura genômica das cultivares modernas de cana-de-açúcar.....24

2.4 Marcadores moleculares em cana-de-açúcar.....26

2.5 Polimorfismo e mapeamento genético em cana-de-açúcar.....29

2.6 O carvão da cana-de-açúcar.....35

2.7 Mapeamento de QTL em cana-de-açúcar.....39

3 MATERIAL E MÉTODOS.....45

3.1 População de mapeamento.....45

3.2 Extração, quantificação de DNA e desenvolvimento de marcadores moleculares.....45

3.2.1 Marcadores EST-SSR.....46

3.2.2 Marcadores TRAP.....48

3.3 Notação das marcas.....50

3.4 Análise de segregação.....51

3.5 Construção do mapa de ligação integrado.....51

3.5.1 Atribuição dos grupos de co-segregação em grupos de homo(eo)logia.....52

3.6 Avaliação da resistência a *S. scitamineum*.....53

3.6.1 Análise dos dados fenotípicos.....54

3.7 Mapeamento de QTLs.....55

3.8 Isolamento e validação dos marcadores TRAP.....59

4 RESULTADOS E DISCUSSÃO.....61

4.1 Desenvolvimento de marcadores moleculares.....61

4.1.1 Seleção dos marcadores EST-SSR.....61

4.1.2 Seleção dos marcadores TRAP.....63

4.2 Genotipagem e análise de segregação dos marcadores.....65

4.3 Construção do mapa de ligação e atribuição dos grupos de co-segregação nos grupos de homo(eo)logia.....70

4.4 Análise dos dados fenotípicos.....87

4.5 Mapeamento de QTL.....91

4.6 Validação das sequências dos marcadores TRAP.....102

5 CONSIDERAÇÕES FINAIS.....107

REFERÊNCIAS.....109

ANEXOS.....129

RESUMO

Mapa de ligação e mapeamento de locos quantitativos de resistência a *Sporisorium scitamineum* em cana-de-açúcar

A cultura da cana-de-açúcar está exposta a diversos patógenos, entre eles, o fungo *Sporisorium scitamineum*, agente causal do carvão, que causa perdas significativas na produção e na qualidade do caldo. O uso de genótipos resistentes é o principal método de controle da doença, o que torna essencial o conhecimento das bases genéticas da resistência, favorecendo a seleção de genótipos superiores. Neste estudo, uma população F_1 segregante, derivada do cruzamento entre 'IAC66-6' e 'TUC71-7', foi genotipada usando marcadores EST-SSR e TRAP, os quais foram alocados em um mapa prévio, visando identificar marcadores associados a locos quantitativos de resistência a esse patógeno. A incidência de carvão foi avaliada em dois ensaios realizados nos anos de 2009 e 2010, em delineamento experimental inteiramente casualizado. O principal sintoma do carvão (número de chicotes) foi tomado ao longo dos ensaios e os dados foram utilizados para o cálculo da área sob a curva de progresso da doença (AUDPC), sendo os valores de AUDPC normalizados e analisados por meio de modelos mistos. O mapeamento de locos quantitativos (QTL) seguiu a abordagem de mapeamento por intervalo composto (CIM). O mapa aqui construído contém 721 marcadores, compilando EST-SSRs, TRAPs, além de AFLPs e marcadores direcionados ao retrotransposon *scIvana_1* já alocados no mapa prévio. Estes foram posicionados em 124 grupos de co-segregação, sendo 65 deles atribuídos em dez grupos de homo(eo)logia, com base em locos EST-SSR em comum. O mapa totaliza 6.223,0 cM e tem uma densidade média de um marcador a cada 8,6 cM. O uso da metodologia CIM permitiu detectar oito locos quantitativos associados à resistência de *S. scitamineum*, sendo quatro deles identificados na primeira avaliação, dois novos locos na segunda avaliação e dois considerando-se o comportamento médio dos genótipos nos dois anos de avaliação. Observou-se a predominância de alelos com efeitos aditivos de resistência, comparativamente aos efeitos de dominância, sendo seis deles relacionados à diminuição da suscetibilidade e quatro com o aumento. Dentre os QTL, cinco segregaram na proporção 1:1, dois 3:1 e um na proporção de 1:2:1. Os QTLs foram capazes de explicar individualmente de 3,4% a 5,7% da variação fenotípica. O QTL que explicou a maior porcentagem da variação, detectado na primeira avaliação (5,7%), foi posicionado a 5,0 cM de outro QTL considerando-se o comportamento médio dos genótipos, sugerindo que estes QTLs localizam-se na mesma região genômica. Interessantemente, um dos QTLs está posicionado em um intervalo que contém uma marca TRAP e outra ancorada no retrotransposon *scIvana*. Espera-se que esses resultados contribuam para estudos futuros sobre genes candidatos envolvidos na resistência, o papel dos retrotransposons na resposta da planta a patógenos, bem como possam dar subsídios a programas de seleção assistida por marcadores.

Palavras-chave: Cana-de-açúcar; *Sporisorium scitamineum*; Mapa de Ligação, Mapeamento Genético; QTL

ABSTRACT

Linkage and mapping quantitative loci for resistance to *Sporisorium scitamineum* in sugarcane

The sugarcane is exposed to several pathogens, including the fungus *Sporisorium scitamineum*, the smut disease´s causal agent, which causes significant losses in yield and juice quality. The use of resistant genotypes is the main strategy to control the disease, therefore it is essential the knowledge on the genetic basis of resistance for selecting superior genotypes. In this study, a F_1 segregating population derived from a cross between 'IAC66-6' and 'TUC71-7' were genotyped using EST-SSR and TRAP markers that were placed on a prior map aiming at to identify markers associated to the smut resistance genes. The incidence of smut was evaluated in two trials performed in 2009 and 2010 in a completely randomized design. The main smut symptom (number of whips) was recorded in both trials, and data were used to calculate the area under the disease-progression curve (AUDPC). The values were normalized and analyzed using mixed models. The mapping of quantitative loci (QTL) followed the composite interval mapping (CIM) method. The map here in constructed contains 721 single dose markers compiling EST-SSRs, TRAPs as well as AFLPs and markers anchored in the retrotransposon *scIvana_1*, both previously allocated. Markers were positioned on 124 co-segregation groups, and 65 of them were assigned to ten groups of homo(eo)logy, based on common EST-SSR loci. The map totalized 6223.0 cM with a marker density of one marker every 8.6 cM. The use of CIM methodology allowed the detection of eight loci associated with quantitative resistance to *S. scitamineum*, four being identified in the first trial, two new QTLs in the second trial, and two considering the average genotype behavior in both years of evaluation. There was predominance of additive effects in comparison to the dominance effects, six of them related to decreased susceptibility and four related to increased susceptibility. Out of the QTLs, five segregated in a 1:1 ratio, two in a 3:1 ratio, and two in a 1:2:1 ratio. QTLs were able to explain individually 3.4% to 5.7% of the phenotypic variation. The QTL that explained the largest percentage of the variation, detected in the first trial (5.7%) was located at a distance of 5.0 cM from other QTL detected when considering the genotype average behavior, suggesting that these QTLs are positioned in the same genomic region. Interestingly, one QTL was located in an interval that contains a TRAP marker and a *scIvana_1*-anchored marker. We expect that our results contribute to future studies on disease resistance candidate genes, the role of retrotransposon in plant responses to pathogens, as well as to add marker-assisted selection programs.

Keywords: Sugarcane; *Sporisorium scitamineum*; Linkage Map; Genetic mapping; QTL

LISTA DE FIGURAS

Figura 1 - Representação esquemática do genoma das cultivares modernas de cana-de-açúcar. Cada barra representa um cromossomo. Barras em branco e cinza correspondem aos cromossomos ou segmentos de cromossomos de *S. officinarum* e *S. spontaneum*, respectivamente. Os cromossomos que estão na mesma linha são homólogos ou homo(eo)logos. Fonte: D'Hont (2005) e Ming et al. (2006).....26

Figura 2 - (a) Representação esquemática da segregação de marcadores em dose única na descendência do cruzamento, de acordo com a metodologia proposta por Wu et al. (1992). Figura de Albino; Creste; Figueira, (2006); (b) Padrão de amplificação de marcadores EST-SSR e respectiva segregação alélica em uma população de mapeamento de cana-de-açúcar. P_1 e P_2 correspondem aos genitores; o alelo 1 está presente em ambos os genitores com segregação esperada na progênie de 3:1; os alelos 3 e 4 estão presentes apenas no genitor P_1, sendo esperada a segregação de 1:1; os alelos 2 e 5 estão presentes em ambos os genitores e não segregam na progênie.....30

Figura 3 - Chicote apical: sintoma típico em plantas de cana-de-açúcar causado por *Sporisorium scitamineum*, após dois meses da inoculação.....36

Figura 4 - Gel de agarose (0,8%) mostrando a quantificação do DNA de parte dos indivíduos da progênie e dos genitores 'IAC66-6' e 'TUC71-7'. As três primeiras canaletas contêm o DNA do fago λ, em concentrações crescentes....46

Figura 5 - Representação do cruzamento entre os genitores *P* e *Q*, com os quatro arranjos para os genótipos dos locos quantitativos $P_m^{\{1,2\}}$, $P_{m+1}^{\{1,2\}}$, $Q_m^{\{1,2\}}$ e $Q_{m+1}^{\{1,2\}}$, sendo os alelos dos marcadores m e m+1, P^1, P^2, Q^1 e Q^2 os alelos dos QTLs.....56

Figura 6 - Resultado dos testes de amplificação para 14 locos de EST-SSR em gel de poliacrilamida desnaturante (5%), corado com nitrato de prata, em que: a) SCB252, b) SCB272, c) SCB225, d) SCB443, e) CV54, f) CV63, g) SCB442, h) SCB243, i) CV53, j) SCB227, k) SCB243, l) SCB329, m) CV12 e n) CV11. A primeira coluna de cada gel corresponde ao genitor 'IAC66-6', a segunda ao 'TUC71-7'; as demais correspondem a quatro indivíduos da progênie F_1.....61

Figura 7 - Genotipagem de 20 indivíduos da progênie F_1 e dos respectivos genitores para o loco CV78. a) Visualização de nove alelos em gel de poliacrilamida desnaturante (5%), corado com nitrato de prata. M: marcador de massa molecular de 10 pb. b) Visualização dos mesmos alelos separados pelo método cromatográfico.....62

Figura 8 - Resultado dos testes com 16 combinações de *primers* fixos e arbitrários relativos à técnica TRAP. As setas indicam alguns dos locos polimórficos entre os genitores e monomórficos entre eles, porém com segregação em pelo menos um dos indivíduos da F_1. A primeira coluna de cada combinação corresponde ao genitor 'IAC66-6', a segunda ao 'TUC71-7' e as demais correspondem a quatro indivíduos da progênie F_1. M: marcador de massa molecular de 125 pb.....64

Figura 9 - Genotipagem em gel de poliacrilamida (5%) da população F_1 com locos de EST-SSR: (a) SCB422 mostrando os alelos de 216 a 249 pb; (b) SCB437 mostrando alelos de 439 a 448 pb. Nas canaletas centrais estão os genitores 'IAC66-6' (I) e 'TUC71-7' (T) em duplicata, flanqueados pelos marcadores de massa molecular de 10 pb (M).....65

Figura 10 - Genotipagem da população F_1 com a combinação KIN169A3. Nas canaletas 2 e 3 está o genitor 'IAC66-6" (I) e nas canaletas 4 e 5, o genitor 'TUC71-7' (T). Na primeira, sexta e última canaletas estão os marcadores de massa molecular de 125 pb.....66

Figura 11 - Matriz gráfica de fração de recombinação e LOD *score* para o CG 5. a) *heatmap* do CG 5 mostrando a marca espúria [174 (D2-E31M49170)] ligada a marcas pertencentes aos dois grupos que, depois de removida, resultou na formação de dois novos CGs (b e c).....72

Figura 12 - Matriz gráfica de fração de recombinação e LOD *score* para o CG 44 a) *heatmap* do CG 44 mostrando a insuficiência de marcas do tipo C.8 para a integração. As marcas 123, 111, 146, 190 e 644 são do tipo D2; as marcas 844, 854, 280, 272, 375, 260, 351, 712, 304, 336 e 258 são do tipo D1 e a marca 429 é do tipo C. b e c) Matrizes gráficas após a divisão manual e o reordenamento do CG 44.....73

Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre 'IAC66-6' e 'TUC71-7'. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no 'IAC 66-6'), D2 (presença no 'TUC 71-7') ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos.....75

Figura 14 - Distribuição dos valores de AUDPC na progênie F_1 de cana-de-açúcar, resultante do cruzamento entre 'IAC66-6' e 'TUC71-7', no primeiro ensaio de avaliação de incidência de carvão.....89

Figura 15 - Distribuição dos valores de AUDPC na progênie F_1 de cana-de-açúcar, resultante do cruzamento entre 'IAC66-6' e 'TUC71-7', no segundo ensaio de avaliação de incidência de carvão.....89

Figura 16 - Grupos de co-segregação que tiveram QTLs mapeados para a resistência ao carvão em cana-de-açúcar, com valores de LOD dispostos graficamente. À direita, estão os gráficos mostrando os QTLs. Em verde claro, a curva do primeiro ano de avaliação (2009), em magenta, do segundo ano de avaliação (2010) e, em preto, o comportamento médio dos genótipos nos anos. A linha tracejada refere-se ao limiar crítico LOD *score* ≥ 3,7.....96

Figura 17 - Análise individual do marcador PALA1350 para o ano de 2009 (verde, à esquerda) e considerando o comportamento médio dos genótipos (preto, à direita). No eixo vertical os valores de BLUPs e, no horizontal, os genótipos da população de mapeamento.....100

Figura 18 - Gel de agarose (1%) mostrando o padrão de amplificação a partir das bandas extraídas do gel de poliacrilamida. Nas canaletas 1 a 8 estão os produtos da reamplificação resultantes das combinações dos *primers* fixos/arbitrários e seus respectivos tamanhos em pares de base (pb): NBS441A1125, KIN169A1228, PALA1350, PALA1363, PR3A1209, LRR149A1233, LRR149A1239 e LRR149A1278; M: marcador de massa molecular de 100 pb.....102

LISTA DE TABELAS

Tabela 1 - *Primers* sintetizados para o desenvolvimento dos marcadores TRAP.....49

Tabela 2 - Número de marcadores EST-SSR e TRAP avaliados e respectiva contribuição de cada genitor.....66

Tabela 3 - Genotipagem da população de mapeamento com locos EST-SSR.....67

Tabela 4 - Genotipagem da população de mapeamento com marcadores TRAP.....70

Tabela 5 - Características dos marcadores usados para genotipar a população de mapeamento.....71

Tabela 6 - Marcadores em dose única alocados no mapa de ligação da população 'IAC 66-6' x 'TUC 71-7'.....74

Tabela 7 - Relação dos locos EST-SSR e respectivos alelos mapeados, localização nos grupos de homo(eo)logia e de co-segregação e sua similaridade com proteínas.....83

Tabela 8 - Mapa de ligação desenvolvido para a cana-de-açúcar, com informações sobre o delineamento genético da população de mapeamento, número de marcas ligadas, número de grupos de co-segregação (CGs), grupos de homologia (HGs) e cobertura total do genoma em centiMorgans (cM).....86

Tabela 9 - Estimativas dos componentes de variância genética ($\hat{\sigma}_G^2$), variância fenotípica ($\hat{\sigma}_F^2$), variância ambiental ($\hat{\sigma}_E^2$) e do coeficiente de herdabilidade no sentido amplo ($\hat{h}_A^2$) para incidência de carvão em cana-de-açúcar.....90

Tabela 10 - Resumo dos QTLs mapeados para a resistência da cana-de-açúcar ao carvão, segundo o modelo CIM e considerando LOD score ≥ 3,7.....92

Tabela 11 - Resumo dos QTLs putativos mapeados para a resistência da cana-de-açúcar ao carvão, porém não significativos, considerando LOD score ≥ 3,7.....101

Tabela 12 - Análise das sequências dos marcadores TRAP, usando a ferramenta BLASTx e tomando por base o banco de dados do NCBI.....105

1 INTRODUÇÃO

A cana-de-açúcar (*Saccharum* spp.) tem ganho destaque no cenário mundial recente, devido a sua importância como fonte de energia renovável. O setor sucroenergético passou por um processo de crescimento impulsionado pelo aumento na demanda por etanol, tanto no mercado interno, em função do avanço da tecnologia dos veículos *flex-fuel*, como pelo interesse mundial crescente na utilização desse biocombustível em mistura à gasolina.

A continuidade do desenvolvimento do setor sucroalcooleiro depende do lançamento de variedades geneticamente superiores, capazes de apresentar alta produtividade, melhor qualidade industrial, alta adaptação a diferentes condições ambientais e resistência a fatores bióticos e abióticos (NOBREGA; DORNELAS, 2006). A cultura apresenta limitações quanto à produtividade em decorrência de problemas fitossanitários. Patógenos como *Sporisorium scitamineum* (também designado *Ustilago scitaminea*), agente causal do carvão da cana-de-açúcar, são capazes de provocar perdas significativas tanto na qualidade industrial como na produção, em todo o mundo (TOKESHI; RAGO, 2005; SUNDAR et al., 2012). Até o momento, o uso de genótipos resistentes é o principal método de controle da doença, o que torna essencial o conhecimento das bases genéticas da resistência, favorecendo a seleção de genótipos superiores.

Alguns trabalhos mostraram que a resistência de cana-de-açúcar a *S. scitamineum* é quantitativa, com herdabilidade no sentido amplo variando de baixa a moderada e influenciada pelas condições ambientais (CHAO; HOY; SAXTON, 1990; SILLS et al., 1995; RABOIN et al., 2003).

Para auxiliar os programas de melhoramento, os marcadores genéticos têm sido amplamente utilizados, inclusive na geração de mapas de ligação de cana-de-açúcar. No entanto, a identificação dos locos e seus respectivos alelos envolvidos na resistência quantitativa em cana-de-açúcar é dificultada devido à complexidade do genoma das cultivares modernas, que apresentam origem multiespecífica, sendo altamente poliploides, aneuploides, com variável número de cromossomos (GRIVET; ARRUDA, 2002). Uma alternativa para elucidar a complexidade genética e obter informações a respeito dos locos que controlam os caracteres quantitativos é obter mapas mais densos, compilando marcadores genéticos capazes de amostrar a variação em regiões transcritas do genoma, tais como microssatélites derivados de sequências expressas (EST-SSR) e polimorfismos de amplificação de regiões alvo (TRAP), possibilitando identificar marcas associadas à expressão de genes alvos em populações segregantes (OLIVEIRA et al., 2007; ALWALA et al., 2008; CRESTE et al.,

2010; OLIVEIRA et al., 2009; ANDRU et al., 2011). Assim, os marcadores moleculares direcionados a sequências expressas agregam informação aos mapas de ligação, inclusive aqueles relacionados à resistência as doenças, podendo facilitar a identificação de QTL (*Quantitative Trait Loci*).

Nesse contexto, os objetivos desse trabalho foram: *i*) desenvolver e obter marcadores do tipo EST-SSR e TRAP e integrá-los ao mapa de ligação previamente desenvolvido para a população derivada do cruzamento entre a variedade 'TUC71-7' e o clone 'IAC66-6' (PALHARES et al., 2012), de modo a aumentar a sua cobertura e densidade, e agregar informações oriundas de marcadores funcionais putativos; *ii*) utilizar o novo mapa gerado para mapear QTL para a resposta da população a *Sporisorium scitamineum*, de modo a fornecer informações aos programas de melhoramento da cana-de-açúcar, visando a resistência ao carvão.

2 REVISÃO BIBLIOGRÁFICA

2.1 Aspectos econômicos da cultura da cana-de-açúcar

A cana-de-açúcar é uma cultura de grande importância econômica, ocupando, no território brasileiro, uma área de 8,6 milhões de hectares (CONAB, 2012). Devido a sua ampla utilidade, pode ser empregada sob a forma de forragem para a alimentação animal, ou como matéria prima para a fabricação de diversos produtos como: aguardente, açúcar e álcool. Ainda, seus resíduos, como a vinhaça, podem ser aproveitados como fertilizantes ou na produção de biogás, e o bagaço, como fonte de energia elétrica.

O Brasil destaca-se como o maior produtor e exportador mundial de cana-de-açúcar. Em 2002, eram 320 milhões de toneladas de cana, contra 571,4 milhões na safra de 2011/2012. A produção acumulada de açúcar, um dos principais subprodutos, atingiu a produção de 21,4 milhões de toneladas na mesma safra, já o volume total de etanol foi de 256 milhões de litros, dos quais 60 milhões de etanol anidro e 196 milhões de etanol hidratado, um acréscimo de três milhões de litros em relação à safra de 2010/2011 (UNICA, 2012; CONAB, 2012). A perspectiva para 2020, é que a produção de cana-de-açúcar no Brasil alcance 846 milhões de toneladas, um crescimento de 3% ao ano, e que a produção de açúcar chegue a 47,5 milhões de toneladas. Já ao excedente líquido de açúcar exportável deve alcançar a produção de 32,9 milhões de toneladas e estima-se um aumento de 43,6 bilhões de litros de etanol em 2020 (BIOSUL, 2012).

2.2 A origem do gênero *Saccharum*

A cana-de-açúcar (*Saccharum* spp.) é uma importante cultura em vários países do mundo, devido aos seus colmos serem ricos em açúcar. É uma planta alógama, membro da família *Poaceae* e da tribo *Andropogoneae*, subtribo *Saccharinae* e gênero *Saccharum* (SCARPARI; BEAUCLAIR, 2008). A maioria dos autores considera que o gênero *Saccharum* é composto por seis espécies, denominadas de *S. officinarum* (2n = 80), *S. spontaneum* (2n = 40 - 128), *S. barberi* (2n = 81 - 124), *S. sinense* (2n = 111 - 120), *S. robustum* (2n = 60, 80) e *S. edule* (2n = 60, 70, 80). O gênero *Saccharum*, juntamente com os gêneros relacionados filogeneticamente, *Erianthus* (seção *Ridium*), *Narenga*, *Sclerostachya* e *Miscanthus* (seção *Diandra* Keng) formam o "Complexo *Saccharum*" (MUKHERJEE, 1957; DANIELS et al., 1975; MATSUOKA; GARCIA; ARIZONO, 2005).

A *S. officinarum*, ou cana nobre, é conhecida pela sua alta produção de açúcar. Há evidências que ela derivou da espécie silvestre *S. robustum*, na Nova Guiné. De lá, clones de *S. officinarum* se espalharam pela Indonésia, China, Índia, Micronésia e Polinésia (BRANDES, 1958; GRIVET et al., 2004). *S. spontaneum* apresenta baixa produção de açúcar e é adaptada a diferentes ambientes e resistente a várias doenças. *S. spontaneum* está distribuída desde o Japão, Indonésia, Papua-Nova Guiné até o Mediterrâneo e a África (DANIELS et al., 1975). *S. sinense*, cultivada na China e na Índia, e *S. barberi*, cultivada no norte da Índia, são duas espécies que provavelmente se originaram da hibridização natural entre *S. officinarum* e *S. spontaneum* (PRICE, 1968; D'HONT et al., 2002; LANDELL; BRESSIANI, 2008). *S. edule* compreende o menor grupo de canas estéreis e é, supostamente, derivada da hibridação intergenérica entre *S. officinarum* ou *S. robustum* com *Miscanthus* (DANIELS; ROACH, 1987).

2.3 Hibridação interespecífica artificial e estrutura genômica das cultivares modernas de cana-de-açúcar

Até o final do século XIX, clones de *S. officinarum* foram cultivadas de maneira extensiva em todo mundo, especialmente devido a sua principal característica - alto teor de sacarose. No início do século XX, com o aparecimento do vírus do mosaico e a alta suscetibilidade de *S. officinarum* ocorreu a hibridação artificial entre *S. officinarum* (autopoliploide, 2n = 80) e clones de *S. spontaneum* (2n = 40 - 128). Posteriormente, sucessivos retrocruzamentos foram realizados, visando obter híbridos com elevada produção de açúcar e resistentes a doenças, num processo denominado de "nobilização" (ROACH, 1972; SREENIVASAN et al., 1987). Entretanto, há registros de que ocorre transmissão desigual de cromossomos no processo de hibridação e no subsequente evento de retrocruzamento, sendo '2n' de *S. officinarum* (genitor feminino) e 'n' de *S. spontaneum* (genitor masculino). Várias hipóteses foram levantadas para explicar este fenômeno, entre elas estão: a ocorrência de uma endoduplicação dos cromossomos durante a megasporogênese, duplicação cromossômica após a primeira divisão meiótica ou fusão dos dois núcleos durante o estágio de tétrade (BREMER, 1961b). Somente a partir do segundo retrocruzamento, é que o genitor *S. officinarum* passa a transmitir 'n' cromossomos (BREMER, 1961a; MING et al., 2006). Como resultado, as cultivares modernas de cana-de-açúcar possuem número de cromossomos variando entre 100 a 130 e alta aneuploidia (GRIVET; ARRUDA, 2002).

Burner e Legendre (1993) verificaram que a transmissão dos cromossomos foi estritamente ‘n + n’ quando foi realizado o cruzamento de clones elites interespecíficos com quatro clones de *S. spontaneum*. Contudo aneuploidia, mosaicismo cromossomal e irregularidades meióticas contribuíram para a produção de gametas desbalanceados. Recentemente, através da citogenética molecular, a transmissão de ‘2n + n’ no cruzamento entre *S. officinarum* e *S. spontaneum*, seguido do primeiro retrocruzamento, foram confirmados por Piperidis, Piperidis e D’Hont (2010), que ainda demonstraram que este tipo de transmissão também pode ocorrer em cruzamentos entre *S. officinarum* e as cultivares modernas.

Ao longo do tempo, tem-se debatido sobre o conjunto básico de cromossomos do gênero *Saccharum* spp. Estima-se que o tamanho do genoma monoplóide de *S. officinarum* (x = 10) seja de 930 Mpb e de *S. spontaneum* (x = 8) seja de 750 Mpb, enquanto o tamanho do genoma da cultivar ‘R570’ (2n = 115) seja de 10.000 Mpb, significativamente maior quando comparado com outras gramíneas como arroz (860 Mpb, 2n = 24), sorgo (760 Mpb, 2n = 20) ou milho (5.500 Mbp, 2n = 20) (D’HONT; GLASZMANN, 2001).

Hibridação cromossômica *in situ* com sondas de dois genes de DNA ribossomal, 18S - 25S e 5S, mostrou que *S. officinarum* é octaploide, com número básico de cromossomo x = 10, enquanto *S. spontaneum* possui x = 8, e que o nível de ploidia dos citotipos desta espécie varia entre 5 a 16 cópias homólogas (D’HONT et al. 1998). Como uma consequência disso, dois conjuntos cromossômicos distintos coexistem nas cultivares modernas. Através da hibridação genômica *in situ* (GISH), D’Hont et al. (1998) demonstraram que 15 a 25% dos cromossomos das cultivares modernas são herdados de *S. spontaneum*, e que há recombinação entre os cromossomos homo(eo)logos. Na cultivar ‘R570’, por exemplo, 10% dos cromossomos são herdados de *S. spontaneum*, 80% são herdados de *S. officinarum* e 10% são resultantes da recombinação entre os cromossomos de seus dois ancestrais. Na cultivar ‘NCo376’ (2n ≈ 112), a distribuição foi de 20% de cromossomos herdados de *S. spontaneum*, 70% de *S. officinarum* e 10% de recombinantes.

Análises adicionais de GISH, em cultivares modernas, mostraram que a proporção da contribuição de *S. spontaneum* permanece entre 10 a 27%, enquanto a proporção de cromossomos recombinantes varia de 5 a 17% (PIPERIDIS; D’HONT, 2001; CUADRADO et al., 2004; PIPERIDIS; PIPERIDIS; D’HONT, 2010)

No seu conjunto, estes dados permitiram D’Hont (2005) propor uma representação esquemática para explicar a organização dos genomas das cultivares modernas (Figura 1), sendo as principais características desse genoma a alta poliploidia, a aneuploidia, a origem

multiespecífica dos cromossomos, a ocorrência de recombinação entre os cromossomos das espécies genitoras e a existência de diferenças estruturais entre os seus cromossomos.

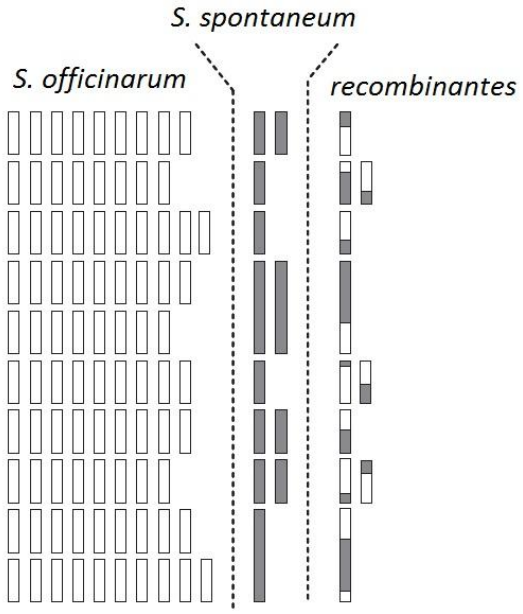


Figura 1 - Representação esquemática do genoma das cultivares modernas de cana-de-açúcar. Cada barra representa um cromossomo. Barras em branco e cinza correspondem aos cromossomos ou segmentos de cromossomos de *S. officinarum* e *S. spontaneum*, respectivamente. Os cromossomos que estão na mesma linha são homólogos ou homo(eo)logos. Fonte: D'Hont (2005) e Ming et al. (2006)

2.4 Marcadores moleculares em cana-de-açúcar

Marcadores moleculares são sequências específicas de DNA que apresentam segregação mendeliana, tem ação dominante ou co-dominante e se prestam para análise genética (FERREIRA; GRATTAPAGLIA, 1998; FUNGARO; VIEIRA, 2001; BORÉM; VIEIRA; COLLI, 2009).

Os diferentes tipos de marcadores moleculares hoje disponíveis diferenciam-se pela tecnologia usada para revelar a variabilidade ao nível de DNA e, assim, variam quanto à habilidade de detectar diferenças entre os indivíduos, ao custo, à facilidade de uso e repetibilidade (MILACH, 1998).

Técnicas como o RFLP (*Restriction Fragment Length Polymorfism* - BOTSTEIN et al., 1980), o RAPD (*Random Amplified Polymorphic DNA* - WILLIAMS et al., 1990), os SSRs (*Simple Sequence Repeats* - JACOB et al., 1991) e o AFLP (*Amplified Fragment Length Polymorphism* - VOS et al., 1995) foram amplamente utilizadas nas diversas culturas, incluindo a cana-de-açúcar.

Nos últimos anos, alguns projetos de sequenciamento de ESTs (*Expressed Sequence Tags*) da cana-de-açúcar foram iniciados na África do Sul (CARSON; BOTHA, 2000), no

Brasil (VETTORE et al., 2001; 2003) e na Austrália (CASU et al. 2001), disponibilizando uma série de sequências que estão sendo exploradas para o desenvolvimento de marcadores funcionais, possibilitando o desenvolvimento de marcadores associados a genes, tais como EST-SSR (*Simple Sequence Repeats derived from Expressed Sequence Tags*) (PINTO et al., 2004; OLIVEIRA et al., 2009; MARCONI et al., 2011); EST-RFLP (*Restriction Fragment Lenght Polymorphism derived from Expressed Sequence Tags*) (OLIVEIRA et al., 2007); SNP (*Single Nucleotide Polymorphism*) (GRIVET et al., 2003; CORDEIRO et al., 2006; McINTYRE et al., 2006; MARCONI, 2011); TRAP (*Target Region Amplification Polymorphism*) (ALWALA et al., 2006a, 2008; ANDRU et al., 2011) e os marcadores ancorados no retrotransposon *scIvana_1* (PALHARES et al., 2012).

Cordeiro et al. (2001) analisaram 8.678 sequências de ESTs de cana-de-açúcar, as quais revelaram aproximadamente 250 microssatélites, a maioria composta por repetições perfeitas de trinucleotídeos, sendo os motivos $(CCG)_n$, $(CGT)_n$ e $(CCT)_n$ os mais comuns. A avaliação quanto ao polimorfismo e à transferibilidade dos locos selecionados revelou que estes são menos polimórficos que os SSRs isolados de bibliotecas genômicas, porém puderam ser transferidos para outros gêneros relacionados, mas o mesmo não foi observado para os SSRs genômicos.

O banco de dados do SUCEST (*Sugarcane Expressed Sequence Taq Project*) tem disponibilizado uma grande coleção de ESTs com potencial para a geração de diferentes classes de marcadores moleculares. A busca no banco de dados do SUCEST revelou um total de 2005 clusters, dos quais 43.141 contém SSRs (PINTO et al., 2004). Desses, 165 são repetições de dinucleotídeos, 613 são de trinucleotídeos e 1227 são de tetranucleotídeos. Entre os motivos de dinucleotídeos, $(CG)_n/(CG)_n$ foram os menos frequentes (12%), já os motivos de trinucleotídeos mais comuns foram $(GCT)_n/(ACC)_n$ com 18,8%, $(CGG)_n/(CCG)_n$ com 16,3%, $(CTC)_n/(GAG)_n$ com 12,4% e $(TCA)_n/(TGA)_n$ com 11,0%. Para os tetranucleotídeos, os motivos mais abundantes foram $(CACG)_n/(CGTG)_n$ com 8,3% e $(AGGC)_n/(GCCT)_n$, $(ACAT)_n/(ATGT)_n$, e $(TCAT)_n/(ATGA)_n$ com aproximadamente 5,0%. O polimorfismo de 23 locos foi testado em 18 cultivares, evidenciando um moderado polimorfismo, com número de alelos variando de dois a 15, com uma média de 6,0 alelos/loco. Em um segundo trabalho, Pinto et al. (2006) compararam 51 locos EST-SSR com 50 locos gerados a partir de bibliotecas genômicas (gSSR) utilizando 18 cultivares modernas, e evidenciaram que a maioria dos locos EST-SSR têm entre quatro a seis alelos enquanto de sete a nove alelos foram observados para locos gSSR. Os autores ressaltam o poder de descriminação dos EST-SSR, além de poderem ser usados com sucesso na análise de diversidade genética a nível

funcional. Mais tarde, Oliveira et al. (2007; 2009) caracterizaram 224 novos marcadores EST-SSR, os quais revelaram locos contendo de dois a 24 alelos, com média de 7,5 alelos/loco.

Recentemente, a partir dos 2005 clusters de SSR contidos no banco de dados do SUCEST, mais 197 EST-SSRs com motivos de trinucleotídeos foram desenvolvidos e avaliados na coleção de 18 genótipos de cana-de-açúcar (MARCONI et al., 2011). Estes locos contêm de dois a 18 alelos, com uma média de 6,0 alelos/loco. Por apresentarem similaridade com proteínas envolvidas em importantes vias metabólicas e mecanismos de defesa da planta, os EST-SSR são valiosas ferramentas para a construção de mapas genéticos, identificação de QTL e possível seleção assistida por marcadores (MAS).

Recentemente, de posse dos dados disponíveis do projeto SUCAST (*Sugarcane Signal Transduction*), 241 locos SNP foram identificados e desenvolvidos por Marconi (2011), e então, usados para genotipar 180 indivíduos derivados do cruzamento entre 'IACSP95-3018' e 'IACSP93-3046' (MARCONI, 2011; MOLLINARI, 2012). A partir dos sinais de intensidade de massa referentes ao loco SNP, fornecida pela plataforma de genotipagem Sequenom iPLEX MassARRAY® (OETH et al., 2007), pôde-se inferir a dosagem de cada um dos alelos presentes no loco e consequentemente tornou-se possível observar diferentes classes genotípicas, tornando os SNPs excelentes marcadores para o estudo genético em poliploides (MOLLINARI, 2012).

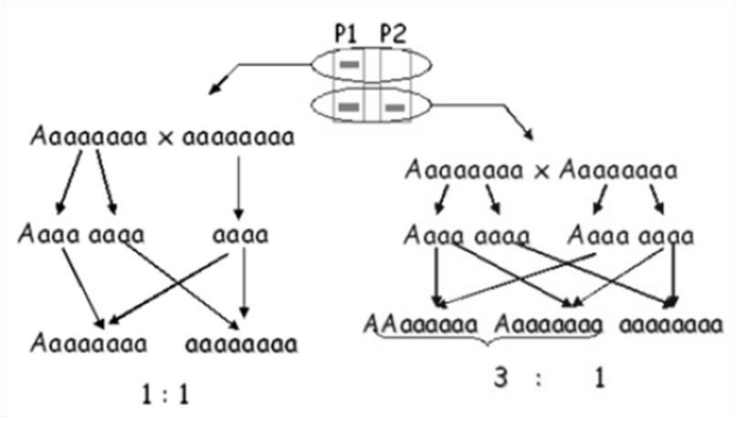
Utilizando a ampla quantidade de informações geradas pelos projetos de sequenciamento de ESTs, Hu e Vick (2003) desenvolveram a técnica TRAP que possibilita a identificação de polimorfismo gerado a partir de dois *primers*: um *primer* fixo, que é desenhado para a sequência de EST depositada nos bancos de dados e um *primer* arbitrário, que é uma sequência arbitrária com um núcleo rico em AT ou GC, complementar a regiões do genoma constituídas de íntrons ou éxons, respectivamente. Assim como os marcadores SRAP (*Sequence-Related Amplified Polymorphism*), os marcadores TRAP detectam polimorfismo intragênico (ALWALA; KIMBENG, 2010).

Sequências de genes candidatos envolvidos no metabolismo da sacarose (ALWALA et al., 2006a; CRESTE et al., 2010), relacionados à via da lignina (ANDRU et al., 2012), tolerância à seca (CRESTE et al., 2010) e ao frio (ALWALA et al., 2008), macho-esterilidade (CHEN et al., 2006) e à resistência a doenças (LIU et al., 2005; MIKLAS et al., 2006) têm servido de base para o desenvolvimento de *primers* usados na metodologia TRAP. Visando confirmar a origem das sequências de fragmentos gerados pela técnica TRAP em cana-de-açúcar, Alwala et al. (2006a) e Andru et al. (2012) compararam estas sequências com as já existentes no banco de dados do NCBI (*National Center for Biotechnology Information*) e

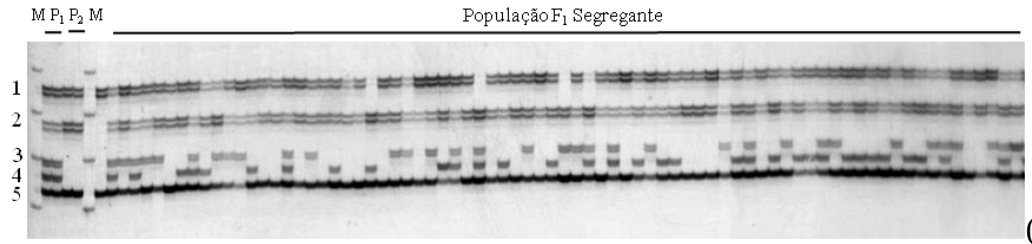
observaram que as sequências geradas apresentaram baixa similaridade com os genes alvo, no entanto, alta similaridade foi identificada com fragmentos genômicos conhecidos de outras espécies de gramíneas incluindo sorgo, milho e arroz. Segundo Andru et al. (2012), isso pode ser atribuído a possível presença de regiões não codantes, imediatamente após o sítio de ligação do *primer*, que resulta na amplificação de fragmentos com baixa similaridade ao gene alvo e/ou com alta similaridade ao fragmento genômico (ou clone de BAC). Por outro lado, LIU et al. (2005) sugerem que a baixa temperatura inicial de anelamento dos *primers* resulta em uma baixa especificidade de anelamento e na amplificação de fragmentos específicos e inespecíficos. Isto significa dizer que a metodologia tem limitações. Entretanto, foi utilizada com sucesso em estudos de diversidade (ALWALA et al., 2006a; CRESTE et al., 2010; ANDRU et al., 2012), para estimar coeficiente de parentesco (ALWALA et al., 2006b), na construção de mapas e detecção de QTL em cana-de-açúcar (ALWALA et al., 2009), trigo (LIU et al., 2005) e no feijoeiro (MIKLAS et al., 2006).

2.5 Polimorfismo e mapeamento genético em cana-de-açúcar

Um método de mapeamento genético, para poliploides, supondo pareamento bivalente na meiose, foi proposto por Wu et al. (1992). Este método é baseado na segregação de marcadores em dose única (MDU). Supondo-se octaploidia, tais marcadores podem apresentar segregação do tipo 1:1, quando em um genitor há somente uma cópia do alelo no loco sob estudo (*Aaaaaaaa* - simplex) e ausência deste alelo no outro genitor (*aaaaaaaa* - nuliplex) e 3:1, quando em ambos os genitores há uma cópia do alelo, ou seja, ambos os genitores são simplex (Figura 2). Desta forma, os indivíduos da progênie originada do cruzamento entre dois genitores, são genotipados com base na presença ou ausência da banda, independentemente da natureza do marcador utilizado. Nesse caso, marcadores co-dominantes, como os microssatélites, perdem esta vantagem e passam a funcionar como marcadores dominantes.



(a)



(b)

Figura 2 - (a) Representação esquemática da segregação de marcadores em dose única na descendência do cruzamento, de acordo com a metodologia proposta por Wu et al. (1992). Figura de Albino; Creste; Figueira, (2006); (b) Padrão de amplificação de marcadores EST-SSR e respectiva segregação alélica em uma população de mapeamento de cana-de-açúcar. P_1 e P_2 correspondem aos genitores; o alelo 1 está presente em ambos os genitores com segregação esperada na progênie de 3:1; os alelos 3 e 4 estão presentes apenas no genitor P_1, sendo esperada a segregação de 1:1; os alelos 2 e 5 estão presentes em ambos os genitores e não segregam na progênie

Devido à complexidade do genoma da cana-de-açúcar, os geneticistas têm utilizado populações F_1 segregantes e/ou progênies provenientes de autofecundação de genitores heterozigóticos (não endogâmicos) para construírem mapas genéticos, pois a obtenção de linhagens nessa espécie torna-se impraticável, devido ao número de gerações necessárias para atingir a homozigose, à consequente depressão causada pela endogamia e às irregularidades meióticas decorrentes de sua origem poliplóide e aneuploidia (ALWALA; KIMBENG, 2010). Devido a isto, não é possível usar populações convencionais de mapeamento que tem por base genótipos homozigóticos e populações deles derivadas.

Parte dos mapas construídos para a cana-de-açúcar é baseada na metodologia *2-way pseudo-testcross*. Esta estratégia foi desenvolvida por Grattapaglia e Sederoff, em 1994, para eucalipto em que, a exemplo da cana-de-açúcar, o mapeamento é realizado na geração F_1 oriunda de cruzamento entre genitores não endogâmicos. Segundo esta metodologia, cada conjunto de marcadores que segregam a partir de um genitor é analisado separadamente, dando origem a mapas individuais dos genitores. Esta metodologia de mapeamento tem sido muito utilizada na construção de mapas genéticos em cana-de-açúcar (AL-JANABI et al.,

1993; AL-JANABI; HONEYCUTT; SOBRAL, 1994; DA SILVA et al., 1993; GUIMARÃES et al., 1999; EDMÉ; GLYNN; COMSTOCK, 2006; ALWALA et al., 2008).

O primeiro estudo de mapeamento em *Saccharum officinarum*, utilizando a metodologia *2-way pseudo testcross*, foi realizado em uma população proveniente do cruzamento entre *S. officinarum* 'LA Purple' e *S. robustum* 'Mol 5829', com base em marcadores RAPD. Como resultado, 50 marcas em dose única e polimórficas foram identificadas no genoma de *S. officinarum* e 33 em *S. robustum* (AL-JANABI; HONEYCUTT; SOBRAL, 1994). Posteriormente, a mesma população foi utilizada na construção de dois diferentes mapas de *S. officinarum.* A geração do primeiro mapa foi realizada usando locos RAPD, sendo possível identificar 279 MDU (ou simplex), dos quais 161 formaram 50 grupos de co-segregação (CGs, do inglês *co-segregation groups*), restando 42% das marcas não ligadas (MUDGE et al. 1996). O segundo mapa foi desenvolvido por Guimarães et al. (1999), utilizando marcadores RFLP, RAPD e AFLP, contendo 283 MDU que formaram 74 CGs para o mapa do genitor 'LA Purple' e 208 MDU alocados em 65 CGs para o mapa do genitor 'MOL5829'.

Mapas individuais, envolvendo cultivares modernas, também foram obtidos. A partir do cruzamento entre dois cultivares australianos 'Q117' x 'MQ77-340', sendo a progênie constituída de 232 indivíduos, formaram-se 93 CGs, a partir de 270 marcas, abrangendo 3.167 cM do genoma de 'Q117'. Já para a MQ77-340', foram obtidos 101 CGs constituídos por 342 MDU abrangendo 3.582 cM (REFFAY et al., 2005). Uma progênie, composta por 192 indivíduos, derivada do cruzamento entre os genitores 'Q117' e '74C42', foi usada para a construção de mapas individuais. Marcadores SSR, AFLP, RGA e RFLP geraram 407 MDU de 'Q117' formando 75 CGs e 445 MDU de '74C42' formando 84 CGs, ambos com seis grupos de homo(eo)logia (HGs do inglês *homo(eo)logous group*) (McINTYRE et al. 2005b).

Poucos mapas genéticos têm sido construídos usando progênies derivadas de autofecundação de clones ("*pseudo-*F_2"). Ao contrário do mapeamento realizado na geração F_1, na qual dois mapas são gerados, a população originada por autofecundação proporciona o desenvolvimento de um único mapa genético do clone ou da cultivar da qual a população foi desenvolvida (ALWALA; KIMBENG, 2010).

O primeiro mapa genético gerado para uma cultivar moderna utilizando progênies de autofecundação foi feito por Grivet et al. (1996). Neste estudo, foram genotipadas 77 progênies da 'R570'. Este mapa foi composto por 408 marcas, geradas por 120 sondas de RFLP e uma isoenzima, as quais foram alocadas em 96 CGs, com comprimento cumulativo de 2.008 cM e distância média entre locos de 6,1 cM. Esses CGs puderam ser agrupados

dentro de dez grupos de homo(eo)logia. A investigação dos produtos de amplificação levou à identificação de 80 marcadores derivados de *S. officinarum* e 66 de *S. spontaneum*. Os autores observaram que a distribuição dos marcadores dentro dos CGs proporcionou melhor cobertura do genoma de *S. spontaneum* do que de *S. officinarum*. Em seguida, Hoarau et al. (2001) desenvolveram marcadores AFLP para esta mesma cultivar, porém usando uma população diferente. Os autores utilizaram 295 indivíduos, o que possibilitou a construção de um mapa composto por 887 AFLPs, distribuídos ao longo de 120 CGs, com uma cobertura genômica de 5.849 cM. A partir de marcadores comuns com o mapa desenvolvido por Grivet et al. (1996), 34 dos CGs puderam ser agrupados em um dos dez HGs. Mais tarde, marcadores de RGAs (*Resistance Gene Analogs*), derivados de ESTs do projeto SUCEST e gSSRs foram adicionados a este mapa, totalizando 128 CGs e sete HGs (ROSSI et al., 2003).

Devido ao tamanho do genoma da cana-de-açúcar, um grande número de marcadores é exigido para a construção de mapas e supõe-se que mesmo o mais saturado dos mapas cobrirá somente dois terços do seu genoma (AITKEN; JACKSON; McINTYRE, 2005). Segundo estes autores, aproximadamente 70% a 84% dos marcadores presentes no genoma da cana-de-açúcar são MDU. Por outro lado, cerca de 16% a 30% das marcas ainda continuam não ligadas, isto é, estariam segregando em outras doses (DA SILVA et al., 1995; AITKEN; JACKSON; McINTYRE, 2005). O uso de marcadores que estão presentes em mais de uma cópia em um dos genitores (multidoses) poderia aumentar a cobertura do mapa em regiões do genoma que apresentam baixo nível de MDU, diminuindo o número de marcadores não ligados (EDMÉ; GLYNN; COMSTOCK, 2006; AITKEN; JACKSON; McINTYRE, 2007).

Edmé, Glynn e Comstock (2006) estudaram a população filial do cruzamento entre *S. officinarum* (‘Green German’) e *S. spontaneum* (‘IND 81-146’). Os autores verificaram que os marcadores SSR segregaram nos modelos mendelianos para doses simples (1:1), duplas (3:1, 7:2 e 11:3), triplas (11:1, 15:1 e 13:1) e quádruplas (64:1 e 69:1). AITKEN; JACKSON; McINTYRE (2007). A partir dessas informações, adicionaram marcadores em doses duplas, incrementando o mapa do clone ‘IJ76-514’ (*S. officinarum*) de 47 para 123 CGs; consequentemente, a cobertura do mapa passou de 2.441 cM, usando marcadores em doses simples, para 4.906 cM quando consideradas as doses duplas.

Alwala et al. (2008) utilizaram marcadores AFLP, SRAP e TRAP em 100 progênies F_1 derivadas do cruzamento interespecífico - *S. officinarum* ‘La Striped’ x *S. spontaneum* ‘SES 147B’. Trezentos e quarenta e quatro marcas polimórficas foram geradas para o genitor *S. officinarum*, das quais 72% segregaram em dose única (1:1), 9% segregaram em doses duplas (3:1, 11:3 e 7:2) e 19% das marcas mostraram segregação diferente da esperada. Já

para o genoma de *S. spontaneum*, foram obtidas 306 marcas, das quais 72% segregam em dose simples, 14% em dose dupla e 14% não apresentaram nenhuma das segregações esperadas. O mapa de ligação de *S. officinarum* foi constituído por 146 marcas distribuídas em 49 CGs, resultando em uma cobertura do genoma de 1.732 cM. Para *S. spontaneum*, foi gerado um mapa com 121 marcas ligadas para formar 45 CGs, abrangendo 1.491 cM. Posteriormente, esta mesma população de mapeamento foi avaliada para características relacionadas com o teor de açúcar como *Brix* e *Pol* (ALWALA et al., 2009).

Seguindo o método proposto por Da Silva (1993) e explicado por Ripol et al. (1999), em que marcadores em doses únicas, duplas e triplas podem ser usados na construção de mapas genéticos, um mapa genético da cultivar de Louisiana 'LCP85-384' foi construído usando a progênie de autofecundação e baseado no polimorfismo de marcadores AFLP, SSR e TRAP, os quais foram testados para a segregação de 3:1 (simplex) e 77:4 (duplex). Esse mapa foi constituído por 718 marcadores simplex e 66 duplex, os quais foram distribuídos em 108 CGs com um comprimento cumulativo de 5.617 cM e uma densidade de 7,16 cM por marcador, estimando uma cobertura de aproximadamente 45,6% do genoma da cultivar 'LCP83-384'. Baseados nas informações dos locos SSR e na fase de ligação em repulsão, nove HGs foram identificados (ANDRU et al., 2011).

A integração de mapas obtidos a partir de cruzamentos bi-parentais é desejável, uma vez que permite aumentar a sua saturação, utilizando diferentes tipos de segregação. Isto se torna possível com a presença de marcadores em heterozigose em ambos os genitores usados para estabelecer as frações de ligação entre os marcadores segregantes (GARCIA et al., 2006). A abordagem proposta por Wu et al. (2002b), baseada em análises de máxima verossimilhança, que permite estimar simultaneamente tanto a fração de recombinação como as fases de ligação entre os locos, tem possibilitado a construção de mapas integrados mais informativos, como demonstrado por Garcia et al. (2006), Oliveira et al. (2007), Pastina et al. (2012) e Palhares et al. (2012). Garcia et al. (2006) desenvolveram um mapa genético integrado para a população de mapeamento derivada do cruzamento entre duas cultivares brasileiras, 'SP80-180' x 'SP80-4966' e constituída por 100 indivíduos. Um total de 1.118 MDU foi utilizado, sendo que 39% das marcas segregaram 1:1 e 61% segregaram 3:1. O mapa final contém 357 marcadores ligados, distribuídos ao longo de 131 CGs, considerando LOD *score* de 5 e fração de recombinação de 37,5 cM. Esse mapa totalizou 2.604,4 cM, com uma densidade de um marcador a cada 7,3 cM. Marcas RFLP e SSR foram utilizadas para a geração de 12 HGs putativos. Todo esse trabalho foi desenvolvido usando o *software OneMap* (MARGARIDO; SOUZA; GARCIA, 2007) que, segundo os autores, apresentou

resultados superiores quando comparado ao do *software* JoinMap (VAN OOIJEN; VOORRIPS, 2001). Em seguida, Oliveira et al. (2007) adicionaram ao mapa marcadores EST-SSR e EST-RFLP resultando na geração de 192 CGs constituídos por 664 MDUs, com comprimento total de 6.261,1 cM. Dos 192 CGs, 120 foram agrupados em 14 HGs putativos.

O recente desenvolvimento e mapeamento de marcadores SNP tem sugerido uma nova organização do genoma da cana-de-açúcar. Os trabalhos desenvolvidos por Marconi (2011) e Mollinari (2012) mostraram que a maioria das segregações observadas não é em dose única e, até agora, apenas uma pequena porção do genoma da cana-de-açúcar vem sendo amostrada. Com esses estudos recentes conseguiu-se estimar a ploidia e a dose de cada loco estudado. Dessa forma, tornou-se possível, também, amostrar locos que segregam em diferentes classes genotípicas como, por exemplo, 1:2:1 e 1:6:6:1 (MARCONI, 2011; MOLLINARI, 2012). A introdução de SNPs no mapa de ligação integrado da população derivada do cruzamento bi-parental entre 'IACSP95-3018' e 'IACSP93-3046', previamente construído a partir de SSRs e AFLPs, resultou em mudanças no arranjo dos CGs e mesmo dentro dos HGs. Isto somente foi possível devido ao trabalho realizado por Mollinari (2012), o qual desenvolveu um novo modelo para a construção de mapas genéticos em autopoliploides, considerando diferentes níveis de ploidia e dosagens, baseado na abordagem dos modelos de Markov ocultos.

Recentemente marcadores AFLP, EST-SSR e marcadores baseados em retrotransposons *scIvana*_1 foram usados na geração de um novo mapa de ligação integrado (PALHARES et al., 2012). Nesse trabalho, uma progênie F_1, constituída de 188 indivíduos provenientes do cruzamento entre os genitores 'IAC66-6' e 'TUC71-7' foram genotipados, sendo identificados 546 marcadores em dose única (1:1 e 3:1), distribuídos em 92 CGs, com tamanho estimado de 4.843,19 cM e uma densidade de 8,87 cM, sendo que 41 CGs puderam ser associados a sete HGs. Os marcadores derivados do retrotransposon estudado (*scIvana_1*) mostraram tendência a formarem *clusters* nos CGs.

Quando observados os mapas de cana-de-açúcar, nota-se que há uma grande variação na distribuição do número de CGs dentro dos HGs, sugerindo haver diferentes ploidias. Por exemplo, o mapa de ligação integrado da população derivada do cruzamento entre 'SP80-180' x 'SP80-4966' (OLIVEIRA et al., 2007) é constituído por 192 CGs, dos quais 120 CGs foram distribuídos dentro de 14 HGs, com grupos contendo de 2 a 24 CGs. Já o mapa genético obtido para a população derivada do cruzamento 'IAC66-6' x 'TUC71-7' (PALHARES et al., 2012) contém 92 CGs, dos quais 41 puderam ser atribuídos dentro de 7 HGs, contendo de 15 a 2 CGs. Vale ressaltar, que ambos os mapas são integrados, no entanto, o mesmo comportamento pode ser verificado quando comparados diferentes mapas desenvolvidos a

partir de populações de mapeamento *pseudo-F_2*. O mapa da cultivar moderna 'R570' publicado por Hoarau et al. (2001) é constituído por 120 CGs, dos quais 34 foram atribuídos em 10 HGs, variando de 2 a 10 CGs, enquanto no mapa da cultivar 'LCP85-384' (ANDRU et al., 2011), 44 CGs foram atribuídos em 9 HGs, variando de 2 a 16.

2.6 O carvão da cana-de-açúcar

Como visto anteriormente, vários patógenos são responsáveis por causar doenças em cana-de-açúcar, entre eles o fungo *Sporisorium scitamineum*, agente causal do carvão. Anteriormente foi denominado *Ustilago scitaminea*, fungo da divisão *Basidiomycota*, classe *Ustilaginomycetes*, ordem *Ustilaginales*, família *Ustilaginacea* (HAUKSWORTH et al., 1995). Recentemente, foi sugerido a designação *S. scitamineum* que foi baseada em dados da sequência LSU rDNA realizada por Piepenbring et al. (2002).

S. scitamineum é um parasita biotrófico que infecta tecidos meristemáticos jovens da planta por meio de teliósporos (esporos). O aparecimento da doença nas plantas de cana-de-açúcar depende da germinação dos teliósporos e da fusão de hifas haploides compatíveis sexualmente (*mating-type*), formando hifas dicarióticas infectivas, que penetram na base das escamas das gemas e atingem o tecido meristemático (ALBERT; SCHENCK, 1996; MACCHERONI; MATSUOKA, 2006).

Diferentes raças de *S. scitamineum* têm sido identificadas em países como Argentina (BRAITHWAITE et al., 2004), Nigéria (ABO; OKUSANYA, 1996), Havaí, Paquistão, Filipinas e Taiwan (FERREIRA e COMSTOCK, 1989; COMSTOCK, 2000). No Brasil, o trabalho realizado por Rago, Casagrande e Massola Júnior (2009) detectou que as populações de *S. scitamineum*, amostradas no Estado de São Paulo, apresentaram diferentes níveis de agressividade, porém não houve alterações de virulência entre elas, o que deve ser levado em consideração para o desenvolvimento de variedades com resistência a *S. scitamineum* nos programas de melhoramento.

A doença tem como principal sintoma "o chicote" (Figura 3), que consiste em uma modificação do meristema apical do colmo, induzida pelo fungo, com tamanho variável. Além do chicote apical, também existe o chicote lateral, o qual tem origem em brotações das gemas laterais infectadas. O chicote é composto por parte do tecido da planta e parte do tecido do fungo contendo milhões de teliósporos, esporos assexuados do patógeno (SANGUINO, 1987).

Reportado primeiramente na África do Sul, em 1877, hoje o patógeno é encontrado em quase todas as áreas cultivadas com cana-de-açúcar. No Brasil, sua primeira identificação foi na década de 40 em variedades suscetíveis, como 'JOJ36' e 'JOJ213', chegando a infectar 100% das touceiras, sendo o problema controlado por vários anos com a utilização de variedades resistentes. No entanto, a partir de 1979, com a ocupação de grandes áreas com a variedade 'NA56-79', de suscetibilidade intermediária, aliadas a outros fatores, o carvão voltou a atingir proporções epidêmicas, tornando-se comum no Estado de São Paulo (BERGAMIN FILHO et al., 1987).

Figura 3 - Chicote apical: sintoma típico em plantas de cana-de-açúcar causado por *Sporisorium scitamineum*, após dois meses da inoculação

Além dos prejuízos diretos, a doença acarreta outros indiretos como: impossibilitar o cultivo de variedades altamente produtivas, em vista de sua suscetibilidade; obrigar a redução de áreas cultivadas com variedades de resistência intermediária; provocar o aumento de custos decorrentes de práticas fitossanitárias, como tratamento térmico e "roguing" (CASAGRANDE; SANGUINO, 1994; SUNDAR et al., 2012).

A maneira mais eficiente de controle da doença é o uso de genótipos resistentes, o que torna essencial o conhecimento das bases genéticas da resistência. Estes estudos são realizados por inoculação artificial da planta com *S. scitamineum*. Os programas de melhoramento genético da cana-de-açúcar no Brasil, e diversos no mundo, utilizam diferentes métodos de inoculação artificial para a avaliação da resistência das progênies. Algumas das técnicas mais comumente reportadas incluem a aspersão de teliósporos (COPERSUCAR, 1995; RAGO; CASAGRANDE; MASSORA JÚNIOR , 2009), imersão de gemas intactas ou

feridas em suspensão de teliósporos (CASAGRANDE; AMORIM; SANGUINO, 1998; MATSUOKA et al., 1986), inoculação por injeção de teliósporos nas gemas (PEROS; BAUDIN, 1983), ou a distribuição de uma pasta de teliósporos sobre as gemas com ou sem ferimento (AMORIM et al., 2000).

Muitos trabalhos têm preferência pela técnica que utiliza ferimentos das gemas, pois esta induz uma maior infecção do que a inoculação por imersão. Porém, esta técnica possui o inconveniente de eliminar a resistência morfológica oferecida pela gema da cana-de-açúcar. Isto pode custar caro para o programa de melhoramento, pois os clones eliminados poderiam ter resistência morfológica. A cultivar 'SP70-1143' é um exemplo disso, pois quando inoculada com ferimento, manifesta a doença, quando inoculada sem ferimento, não mostra nenhum sintoma, evidenciando a função da gema na resistência (AMORIM et al., 2000).

Diversos trabalhos mostraram que a resistência de cana-de-açúcar a *S. scitamineum* é quantitativa e com herdabilidade no sentido amplo variando de baixa a moderada (CHAO; HOY; SAXTON, 1990; SILLS et al., 1995; RABOIN et al., 2003). Ainda, condições de estresse hídrico e calor têm favorecido o aparecimento do chicote em genótipos com resistência ao fungo, demonstrando que há interação genótipo por ambientes (CHAO; HOY; SAXTON, 1990; SANTOS, 2003).

Quanto aos mecanismos de resistência ao carvão da cana-de-açúcar, basicamente eles ocorrem nos níveis morfológicos, fisiológicos e moleculares (ARMAS et al., 2007). Embora a interação cana-de-açúcar *vs S. scitamineum* seja bastante complexa, é certo que o patógeno precisa alcançar os tecidos meristemáticos das gemas do hospedeiro para o sucesso do processo de infecção. Este trajeto pode ser dificultado e até impedido pela resistência das gemas (resistência à infecção) proporcionada por barreiras físicas e/ou bioquímicas (WALLER, 1970; DA GLORIA et al., 1999) e/ou das bainhas (GIGLIOTI; TOKESHI, 1991). Uma vez atingidas as regiões meristemáticas, as hifas do fungo podem ou não infectar a planta (resistência à colonização), o que depende do grau de resistência dos tecidos internos (ELSTON; SIMMONDS, 1988).

Da Gloria et al. (1999) verificaram que a morfologia da gema é fundamental como mecanismo de resistência à infecção. Estes autores evidenciaram que a variedade resistente 'SP70-1143' apresenta maior número de escamas, maior número de tricomas, tecidos mais lignificados e com maior reação a compostos fenólicos, quando comparada à variedade suscetível 'NA56-79.

Ainda, características como idade das gemas e concentração de esporos estão envolvidas na resistência (GIGLIOTI; TOKESHI, 1991; BUENO, 2010). De acordo com

Bueno (2010), há diferenças estatísticas entre os dados de incidência de carvão nas diferentes idades de gemas, de acordo com a variedade testada. Nas variedades 'NA56-79' e 'SP71-1406', as gemas mais novas foram mais suscetíveis que as gemas de idade mediana e velha; por outro lado, na variedade 'SP84-2066', as gemas novas e velhas foram mais suscetíveis que as gemas medianas, o que pode ser explicado pela suscetibilidade da variedade. Certa concentração de esporos é requerida na superfície das gemas para que a infecção ocorra. Sob condições de campo, a proteção das escamas das gemas e bainhas pode determinar a quantidade de inóculo que atinge os tecidos internos. A remoção das bainhas, nas variedades NA56-79 e RB72454, foi seguida de uma maior taxa de infecção nas gemas basais, consequentemente, a evolução do carvão da cana-planta para a soca foi também maior nas parcelas de onde as bainhas foram retiradas (GIGLIOTI; TOKESHI, 1991). Em relação à concentração de inóculo necessária para causar doença, estima-se que $1x10^6$ e $1x10^8$ teliósporos/ml sejam suficientes para produzir gemas infectadas (Bueno, 2010).

Através da resistência dos tecidos internos, a planta pode dificultar ou até mesmo impedir o crescimento da hifa de *S. scitamineum* no interior do meristema de cana-de-açúcar. O estudo molecular da interação cana-de-açúcar vs *S. scitamineum* tem revelado diversos genes diferencialmente expressos. Thokoane e Rutherford (2001) usaram a técnica cDNA-AFLP para detectar a ação de genes diferencialmente expressos em genótipos resistente e suscetível ao carvão. Análise de sequências homólogas mostrou que uma proteína putativa quinase receptor de quitina, uma serina/treonina quinase e um LTR retrotransposon tipo *gypsy* são expressos na variedade resistente em resposta ao carvão. Borrás-Hidalgo et al. (2005) identificaram 11 fragmentos derivados de transcritos (TDFs) diferencialmente expressos após inoculação com *S. scitamineum*, os quais mostraram homologia com genes envolvidos na defesa e/ou transdução de sinais, incluindo um gene que codifica uma proteína com domínio NBS-LRR, dois que codificam serina/treonina quinases e um fenilalanina amônia-liase (PAL). Utilizando cDNA-AFLP, Lao et al. (2008) realizaram estudos de expressão diferencial na interação cana-de-açúcar vs *S. scitamineum* e identificaram 64 TDFs diferencialmente expressos, sendo que 67,2% foram mais expressos na cultivar resistente. Os autores ressalvam a complexa resposta de defesa da planta contra o patógeno, sugerindo um papel fundamental de genes envolvidos na explosão oxidativa e nas vias da lignificação na fase inicial da interação. You-Xiong et al. (2011) identificaram 37 TDFs diferencialmente expressos em resposta a inoculação da cana com *S. scitamineum*, utilizando a mesma técnica, dos quais dois apresentaram homologia com genes envolvidos na defesa da célula e seis na transdução de sinais.

Diante do exposto, fica claro que existem inúmeros fatores como condições ambientais, condições morfológicas e fisiológicas da gema, e há também vários genes envolvidos na resistência da cana-de-açúcar a *S. scitamineum*. Assim, quando se objetiva solucionar o problema de obtenção de plantas resistentes ao carvão, os vários mecanismos de defesa devem ser considerados.

2.7 Mapeamento de QTL em cana-de-açúcar

Para o mapeamento de QTL, é necessário que existam associações entre os fenótipos do caráter quantitativo e os alelos dos marcadores que segregam na população. Para tanto, vários métodos estatísticos foram desenvolvidos, entre eles estão: análise para cada marca individualmente ou marca única (SM – *Single Marker*), baseada em análise de variância e regressão linear simples (LYNCH; WALSH, 1998; STUBER; EDWARDS; WENDEL, 1987); mapeamento por intervalo, (LANDER; BOSTSTEIN, 1989); mapeamento por intervalo composto (ZENG, 1994; JANSEN; STAM, 1994) e mapeamento de múltiplos intervalos (KAO; ZENG; TEASDALE, 1999).

A análise de marca única já foi usada no mapeamento de QTL em cana-de-açúcar e é baseada na comparação entre as médias dos caracteres de diferentes classes genotípicas (SILLS et al., 1995; DAUGROIS et al., 1996; MING et al., 2001; 2002a; 2002b; HOARAU et al., 2002; JORDAN et al., 2004; McINTYRE et al., 2005a; 2005b; DA SILVA; BRESSIANI, 2005; REFFAY et al., 2005; AITKEN; JACKSON; McINTYRE, 2006; RABOIN et al., 2006, 2008; WEI et al., 2006; BUTTERFIELD, 2007; AL-JANABI et al., 2007; PIPERIDIS et al., 2008; AITKEN et al., 2008; GAZAFFI, 2009; PINTO et al., 2010; PASTINA et al., 2012). Quando se utilizam marcas individuais, não é necessária a existência de um mapa de ligação. Porém, os testes estatísticos são viesados, possuindo baixo poder estatístico, pois há um confundimento entre os efeitos do QTL e a sua posição em relação ao marcador, o que pode limitar a sua aplicação no melhoramento assistido (DOERGE, 2002).

O método de mapeamento por intervalo (IM – *Interval Mapping*), abordado por alguns autores para a detecção de QTL em cana-de-açúcar (MING et al., 2001; 2002a; 2002b; AL-JANABI et al., 2007; PIPERIDIS et al., 2008; PASTINA et al., 2012), usa o princípio de máxima verossimilhança para fazer inferência da presença de um QTL entre um par de marcadores adjacentes. Para isso, é necessário um mapa genético, o que possibilita obter a posição do QTL e estimar os seus efeitos. Os resultados do IM são expressos em razão de probabilidades, indicada pelo índice conhecido como LOD *score*. Quando comparado a

análise de marca única, o IM tem maior poder estatístico, porém como neste método as análises são realizadas para intervalos isolados, não considerando a interferência de outros QTL localizados fora do intervalo de mapeamento, podem ocorrer falsos positivos em regiões adjacentes a regiões que contêm QTL (DOERGE, 2002).

Visando resolver os problemas gerados pelo mapeamento por intervalo, Zeng (1994) e Jansen e Stam (1994) desenvolveram o método de mapeamento por intervalo composto (CIM – *Composite Interval Mapping*), o qual inclui marcadores fora do intervalo de mapeamento como cofatores, possibilitando o controle dos efeitos de outros QTL que estão fora do intervalo. O método CIM tem várias vantagens sobre SM e IM, entre elas está a falta de confundimento entre o efeito e a posição do QTL e o alto poder estatístico (DOERGE et al., 1997 *apud* PASTINA et al., 2012). Esta metodologia foi implementada em cana-de-açúcar por Al-Janabi et al. (2007), Aitken et al. (2008) e Gazaffi (2009).

No entanto, CIM também têm limitações, superadas pelo método de mapeamento por intervalo múltiplo (MIM - *Multiple Interval Mapping*). Este método, adaptado para cana-de-açúcar por Margarido (2011), aprimora o conceito do CIM, incorporando ao modelo a análise simultânea de múltiplos intervalos, possibilitando estimar efeitos como epistasia entre QTL e interação QTL x ambiente.

Vários softwares foram utilizados para detectar QTL em cana-de-açúcar, entre eles, MapQTL usado por Daugrois et al. (1996), MAPMAKER/QTL e SAS usados por Ming et al. (2001), Ming et al. (2002a) e Al-Janabi et al. (2007), Map Manager QTXb16 usado por McIntyre et al. (2005b), Genstat usado por Wei et al. (2006), Margarido (2011) e Pastina et al. (2012), bem como o QTL Cartographer, usado por Aitken et al. (2008) e Al-Janabi et al. (2007).

Entre as características mais estudadas em cana-de-açúcar estão: resistência à ferrugem marrom (*Puccinia melanocephala*), resistência ao carvão (*S. scitamineum*), mancha amarela (*Mycovellosiella koepkey)*, prodridão radicular (*Pachymetra chaunorhiza*), tempo de florescimento, produção de açúcar, comprimento, peso, diâmetro e número de colmos e conteúdo de fibra.

Detalhes de trabalhos em que foram identificados QTL para doenças da cana-de-açúcar são descritos a seguir.

O primeiro resultado de mapeamento de QTL em cana-de-açúcar foi obtido por Sills et al. (1995), usando marcadores RAPD e uma progênie de 41 indivíduos obtidos do cruzamento entre *S. officinarum* (‘LA Purple’) e *S. robustum* (‘Mol5829’). A população foi avaliada para características quantitativas importantes para o melhoramento da cana-de-açúcar. A

associação entre 83 locos RAPD em dose única com QTL foi determinada usando ANAVA e regressão múltipla. A análise de QTL revelou 24 marcadores significativamente ($P < 0,05$) associados à expressão de caracteres como: número de colmos, resistência ao carvão, diâmetro médio de colmos, teor de açúcar (POL), conteúdo de fibra, peso de colmos por parcela e colmos com panícula. O modelo utilizado foi capaz de explicar entre 23% a 58% da variação fenotípica total, e 32% a 76% da variação genotípica de várias características.

Em seguida, um marcador molecular ligado à ferrugem marrom (*Puccinia melanocephala* H&P Syd.) foi identificado por Daugrois et al. (1996) na progênie constituída por 77 indivíduos derivados da autofecundação da cultivar 'R570' (GRIVET et al., 1996), utilizando análise de marca única. Uma segregação de 3 (resistente) : 1 (suscetível) foi observada para a resistência à ferrugem marrom, sugerindo herança monogênica. Visando saturar a região em torno deste gene de resistência à ferrugem, Asnaghi et al. (2000) usaram a relação de sintenia entre a cana-de-açúcar e o sorgo, milho e arroz, o que levou à identificação de segmentos de cromossomos homo(eo)logos na extremidade do grupo de ligação D de sorgo, o grupo de ligação 2 de arroz, grupo de ligação 4 e região centromérica do grupo de ligação 5 de milho. Posteriormente, oito marcadores AFLP foram adicionados ao redor do gene de resistência à ferrugem, sendo que duas delas foram alocadas a 1,9 e 2,2 cM do gene (ASNAGHI et al., 2004).

McIntyre et al. (2005a) genotiparam 192 indivíduos derivados do cruzamento entre o 'Q117' e '74C42' com marcadores AFLP, SSR e RFLP e usando a metodologia *2-way pseudo testcross*, obtendo dois mapas . Os genitores 'Q117' e '74C42' foram utilizados como macho e fêmea, respectivamente. Esta população foi avaliada para duas doenças, a podridão radicular causada por *Pachymetra* (PRR) e a ferrugem comum. Pela análise de marca única, duas regiões genômicas, uma do mapa do genitor feminino e outra do mapa do genitor masculino, mais uma marca simplex de origem bi-parental e não ligada, foram identificadas como associadas com PRR para os dois anos de avaliação. Essas regiões explicaram de 6% a 16% da variação fenotípica. Para a ferrugem marrom, duas regiões genômicas foram identificadas, uma em cada mapa, mais uma marca simplex de origem bi-parental e não ligada, que foram capazes de explicar de 7% a 18% da variação fenotípica, para ambos os anos. Marcas para cada ano de fenotipagem também foram associadas às doenças. Posteriormente, 54 sequências de RGAs foram isoladas, caracterizadas e utilizadas para identificar marcadores ligados a locos de resistência à ferrugem (McINTYRE et al., 2005b). Neste estudo, 31 RGAs foram mapeados; destes, três locos estavam significativamente associados com a resistência à ferrugem. Por mapeamento comparativo entre cana-de-açúcar e

sorgo, os três locos de RGA associados à ferrugem foram mapeados no mesmo grupo de ligação no sorgo, sendo um deles mapeado em uma região previamente associada a QTL para resistência à ferrugem nesta espécie, demonstrando o valor desse tipo de estudo para a seleção assistida.

Utilizando AFLPs, SSRs e RFLPs, Raboin et al. (2006) construíram dois mapas a partir de 198 progênies de cruzamentos entre os genótipos 'R570' e 'MQ76-53'. Pela análise de marca única, os autores identificaram dois novos locos em 'MQ76-53', um associado à resistência a ferrugem, segregando na proporção de 1:1 (resistente : suscetível), ligado a 23 cM de um loco AFLP e outro controlando o caráter cor de colmos, segregando 1:1 (vermelho : não vermelho), ligado a 6,5 cM de outro AFLP.

A associação entre AFLPs e SSRs e doenças de cana-de-açúcar como podridão radicular (*Pachymetra chaunorhiza*), escaldadura das folhas (*Xanthomonas albilineans*), doença de Fiji (*Fiji disease virus*) e carvão (*Sporisorium scetamineum*) foram investigadas em uma coleção de 154 clones pela análise de regressão (WEI et al., 2006). A partir deste trabalho, os autores observaram que 59% da variação fenotípica para a resistência ao carvão pôde ser explicada por 11 marcadores, 32% da variação fenotípica para escaldadura das folhas e podridão radicular foram explicadas por quatro marcadores, e 26% da doença de Fiji por cinco marcadores, sugerindo o uso desses marcadores em programas de melhoramento. Verifica-se a partir desse estudo que, majoritariamente, poucos locos (até 11) são identificados em resposta à doença.

Um estudo similar foi realizado por Butterfield (2007) também para identificar marcadores associados à resistência ao carvão e a broca-da-cana (*Eldana saccharina*). Para tanto, o autor genotipou 77 genótipos com 275 RFLPs e 1.056 AFLPs e procurou associação desses com a possível resistência. Por meio da análise de regressão, 64 marcadores foram significativamente associados à resposta ao carvão, sendo que o marcador de maior efeito foi capaz de explicar 16% da variação fenotípica e 115 marcadores associados à *E. saccharina*, onde o marcador de maior efeito foi capaz de explicar 20% da variação fenotípica. Dos 179 marcadores associados com carvão e *E. saccharina*, cinco foram significativamente comuns entre os caracteres de resistência estudados.

Al-Janabi et al. (2007) usaram AFLPs e SSRs para identificar QTL para a resistência à mancha amarela (*Mycovellosiella koepkei*). A partir de 227 progênies do cruzamento bi-parental entre as variedades 'M134/75' (resistente) e 'R570' (suscetível), e aplicando a metodologia de mapeamento *2-way pseudo testcross*, a abordagem de marca única e CIM, os

autores observaram segregação monogênica (3 resistentes : 1 suscetível) na progênie estudada.

Recentemente, Gazaffi (2009) expandiu o modelo proposto por Lin et al. (2003) para o contexto do CIM, incluindo cofatores para controlar a variação causada por QTL localizados fora do intervalo de interesse. Baseado na proposta original do método CIM (ZENG, 1994; JANSEN; STAM, 1994), Gazaffi (2009) faz uso do método da máxima verossimilhança para detectar a presença de QTL ao longo do genoma. O mapa de ligação integrado da população F_1 segregante, derivada do cruzamento entre os genitores 'SP80-180' x 'SP80-4966', foi usado por Gazaffi (2009) para detectar QTLs relacionados à produção de cana-de-açúcar. Através do método CIM foi possível identificar 41 QTLs, sendo 18 para o primeiro corte e 23 para o segundo corte, mostrando a eficiência da abordagem. No geral, nota-se que os trabalhos de mapeamento de QTL para doenças da cana-de-açúcar envolvem cruzamentos bi-parentais, interespecíficos ou não, e/ou progênies obtidas por autofecundação de cultivares. Na grande maioria, a detecção de QTL é realizada por análise de marca única. Assim, o desenvolvimento de mapas genéticos integrados associados à metodologia CIM, adaptada por Gazaffi (2009), traz uma nova abordagem para a detecção de QTLs associados a doenças na cultura, principalmente a resposta ao carvão. Como resultado, a detecção de QTL pode facilitar a aplicação de seleção assistida por marcadores moleculares no melhoramento da cana-de-açúcar logo no início do programa.

Vale ressaltar que quando estudada a organização dos genes de resistência a doenças fúngicas em diferentes genomas, nota-se que esses genes estão em *clusters* e flanqueados por retrotransposons, como demonstrado por Ramakrishna et al. (2002) no genoma do milho, Wei; Wing e Wise (2002) em cevada e Melotto et al. (2004) em feijoeiro. Em cana-de-açúcar, a distribuição genômica de genes análogos de resistência, os quais incluem domínios NBS/LRR, LRR e S/T Kinase, foi investigada por Rossi et al. (2003). Os autores revelaram a presença de dez *clusters* de RGAs no genoma da cultivar 'R570', sendo que um deles, formado por dois locos LRR, foi mapeado a 5,2 cM do gene de resistência à ferrugem marrom.

Existe grande empenho da comunidade científica em identificar *clusters* gênicos e detectar QTL relacionados à resistência a doenças. Neste estudo, pretendeu-se contribuir para o avanço destas questões, de modo a agregar informações ao mapa genético, adicionando marcadores e posicionando QTLs relacionados a locos envolvidos na resistência da planta contra *S. scitamineum*.

3 MATERIAL E MÉTODOS

3.1 População de mapeamento

A população de mapeamento aqui utilizada foi constituída de 188 indivíduos de uma progênie F_1 oriunda do cruzamento entre dois genitores de cana-de-açúcar: a variedade procedente de Tucumán, Argentina, 'TUC71-7' e o clone brasileiro desenvolvido pelo Instituto Agronômico de Campinas, 'IAC66-6' utilizados como genitor feminino e masculino, respectivamente. O genitor 'TUC71-7' é resistente ao carvão (MACHADO, 2001), apresenta maior teor de sacarose e menor diâmetro de colmos, quando comparado ao genitor 'IAC66-6', que é suscetível ao carvão (COPERSUCAR, 1995), apresentando menor teor de sacarose e maior diâmetro de colmos. O cruzamento foi realizado na Estação de Cruzamento Vale das Flores – Maceió (AL), pertencente à CanaVialis/Monsanto, e as sementes germinadas em caixas de plástico para individualização dos "*seedlings*". Posteriormente, esse material foi transferido para a área experimental no Pólo Regional de Tecnologia da CanaVialis/Monsanto em Conchal (SP).

3.2 Extração, quantificação de DNA e desenvolvimento de marcadores moleculares

As amostras de folhas dos 188 indivíduos da progênie bem como dos dois genitores foram coletadas individualmente, identificadas e armazenadas a -20°C até o momento da extração. A extração do DNA total foi realizada utilizando o método CTAB (MURRAY; THOMPSON, 1980).

Cada amostra foi quantificada em espectrofotômetro NanoDrop™ 8000 e então diluída em TE 10:1 (10 mM de Tris-HCl e 1 mM de ESTA, pH = 8,0) para a concentração de 5 ng/µl para ser utilizada nas reações. A diluição final foi confirmada em gel de agarose 0,8% (p/v), acrescido de 0,5 X de *SYBR® Safe DNA gel stain* (Promega®). Para avaliar a intensidade da fluorescência das bandas, foi aplicado no gel DNA do fago λ (lambda) íntegro, em concentrações conhecidas, visando à comparação visual da fluorescência emitida sob luz UV. A concentração de DNA da banda foi, então, estimada e o valor dividido pelo volume de DNA aplicado no gel, obtendo-se a concentração do DNA. A fotodocumentação foi realizada em sistema de imagem *Multi Doc-It Digital Imaging System* – UVP (Figura 4).

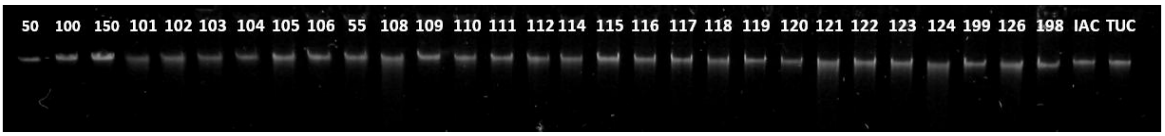


Figura 4 - Gel de agarose (0,8%) mostrando a quantificação do DNA de parte dos indivíduos da progênie e dos genitores 'IAC66-6' e 'TUC71-7'. As três primeiras canaletas contêm o DNA do fago λ, em concentrações crescentes

Quatro tipos de marcadores foram usados para genotipar os genitores e os 188 indivíduos da progênie F_1: AFLP (*Amplified Fragment Length Polymorphism*) e marcadores direcionados ao retrotransposon *scIvana_1*, que resultou na construção de um mapa prévio, como já descrito por Palhares (2010), EST-SSR (*Expressed Sequence Tags-Derived Simple Sequence Repeats*) e TRAP (*Target Region Amplification Polymorphism*), como descrito a seguir.

3.2.1 Marcadores EST-SSR

Neste estudo, foram utilizados 254 pares de *primers*, sendo 72 desenvolvidos por Marconi et al. (2011) e 33 por Oliveira et al. (2009), a partir de 2005 sequências consensos derivadas do banco de dados do SUCEST (http://sucest-fun.org); 134 pares de *primers* desenvolvidos pela CanaVialis/Monsanto e 15 pares de *primers* desenvolvidos por Jordão Junior (2009), sendo três deles publicados por Maccheroni et al. (2009), a partir de 352.122 ESTs de cana-de-açúcar selecionadas dos bancos de dados público (http://www.ncbi.nlm.nih.gov/genbank/) e privado (http://www.alellyx.com/).

Inicialmente, todos os 254 locos foram usados para testar o polimorfismo entre os genitores e uma amostra de quatro indivíduos tomados, aleatoriamente, da população de mapeamento. Locos que exibiram os melhores padrões de amplificação e, que se mostraram polimórficos entre os genitores foram selecionados. Quando monomórficos nos genitores mas que apresentaram indício de segregação na amostra testada, foram igualmente selecionados para a genotipagem da população de mapeamento.

As amplificações foram realizadas usando dois programas de ciclagem, identificados por 1 e 2. Os produtos da PCR também foram analisados por dois métodos, denotados S e F. Os fragmentos foram resolvidos em gel de poliacrilamida desnaturante a 5% (w/v) e corados com nitrato de prata (S), ou em eletroforese capilar (F), usando o sistema de genotipagem MegaBACE 1000® (GE Healthcare Life Sciences), como descrito a seguir.

A PCR foi realizada em um volume final de 10 µl em placas de termociclador com 96 poços. Aproximadamente 20 ng de DNA total foram misturados a 0,25 µM de cada primer forward e reverse; 0,2 mM de cada dNTP; 2,0 mM de $MgCl_2$; 1X Colorless Go Taq buffer e 1 U de Go Taq Flexi DNA Polymerase (Promega®); água Milli-Q autoclavada para completar o volume. A amplificação foi conduzida em termociclador Gene Amp® PCR System 9700 (Applied Biosystems) sob a seguinte condição: uma desnaturação inicial de 94°C por 3 min. seguida por 31 ciclos sendo cada ciclo representado por: desnaturação a 94°C por 60 s anelamento dos primers com temperatura variando por 60 s; extensão a 72°C por 60 s; elongação final a 72°C por 2 min..

Aos produtos amplificados foram adicionados 10 µl de tampão desnaturante (0,2% de azul de bromofenol; 0,2% de xileno cianol; 10 mM EDTA, pH 8,0; 95% de formamida) que foram aquecidos a 94°C por 5 min. e mantidos no gelo durante todo período de aplicação. Uma alíquota de 3 µl de cada amostra foi então aplicada em gel desnaturante de poliacrilamida 5% [acrilamida/bisacrilamida (19:1); 7,5 M de uréia; 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 80 W. A eletroforese foi realizada a potência constante de 80 W por aproximadamente 3 h, usando-se o aparato de eletroforese Sequi-Gen® (Bio-Rad). Ladder de 10 pb foi usado como marcador de massa molecular.

A coloração foi realizada com nitrato de prata ($AgNO_3$), de acordo com Creste, Tulmann Neto e Figueira (2001), com algumas modificações. A fixação do gel foi feita por imersão em solução de etanol 10% e ácido acético 1%, agitando por 10 min., seguida de uma lavagem com água destilada deionizada, durante 1 min., com agitação. O procedimento de pré-tratamento foi realizado com solução de ácido nítrico 1,5%, durante 3 min., seguindo-se uma nova lavagem do gel com água destilada deionizada durante 1 min. A impregnação do gel foi realizada com a solução $AgNO_3$ 0,2% durante 20 min., sob agitação, seguida de duas novas lavagens com água destilada deionizada, com duração de 30 s cada. Para a revelação dos géis foram utilizados 2 l de solução contendo 30 g/l de Na_2CO_3 e 1,0 ml/l de formaldeído (37%). Nesta etapa, 1 litro foi aplicado sobre o gel até começarem a surgirem as bandas. A solução foi então substituída pelo outro litro de solução reveladora, para finalizar o processo de revelação. O gel foi transferido para uma solução de bloqueio contendo ácido acético 5%, onde permaneceu durante 5 min. Em seguida, realizou-se uma lavagem com água destilada deionizada durante 1 min.. Após este passo, para a secagem do gel, a placa foi colocada em posição vertical, em local seco e arejado, sob temperatura ambiente, possibilitando posterior

análise das bandas presentes. Os géis foram analisados sobre uma fonte de luz branca, e digitalizados por um scanner (Color Page - Genius®).

Visando verificar se a tomada dos dados dos produtos amplificados via eletroforese capilar levaria a resultados distintos daqueles obtidos em gel de poliacrilamida, parte dos locos EST-SSR selecionados foram submetidos a um novo teste, utilizando 20 indivíduos tomados aleatoriamente da população F_1 e os dois genitores. Estes foram separados tanto utilizando a técnica de eletroforese em gel de poliacrilamida desnaturante 5%, quanto em capilar. Os alelos separados por ambas as metodologias foram comparados.

A PCR, para esses locos foi realizada do mesmo modo supracitado, porém utilizando primers forward marcados com fluorescência (NED ou 6-FAM, Applied Biosystems). A amplificação foi conduzida em termociclador Gene Amp® PCR System 9700 (Applied Biosystems) sob a seguinte condição: desnaturação inicial a 94 °C por 5 min. seguida por 35 ciclos: desnaturação a 94 °C por 30 s; anelamento dos primers com temperaturas variáveis por 30 s; extensão de 72 °C por 30 s; elongação final a 72°C por 60 min. A eletroforese capilar se deu em sequenciador de DNA MegaBACE1000® (GE Healthcare Life Sciences).

A corrida das reações foi realizada em um arranjo de capilares preenchidos com polímero MegaBACE Long Read Matrix, com tempo de injeção de amostras de 45 s a 3 kV e tempo de corrida de 40 min. a 10 kV e 45°C. Os picos gerados pelo sequenciador foram visualizados com auxílio do software Fragment Profiler version 1.2®. A identificação dos alelos de cada loco foi feita pela inspeção visual e o tamanho dos fragmentos foi estimado pela comparação com o padrão de peso molecular MegaBACE ET550-R Size Standards (GE Healthcare Life Sciences).

3.2.2 Marcadores TRAP

Os *primers* arbitrários foram sintetizados de acordo com Li e Quiros (2001). Os *primers* fixos foram desenhados segundo Hu e Vick (2003), levando em consideração sete sequências obtidas a partir de ESTs de cana-de-açúcar, identificadas por Rossi et al. (2003) como genes análogos de resistência, um fator de transcrição (LIU et al., 2012) relacionado com a ativação da expressão de genes PR (*Pathogenesis Related*) e duas sequências de cana-de-açúcar que codificam as proteínas quitinase e fenilalanina amônia-liase (PAL) obtidas no banco de dados do NCBI, uma vez que estas foram diferencialmente expressos após a inoculação com o patógeno *Sporisorium scitamineum* (LAO et al., 2008 e BORRÁS-HIDALGO et al., 2005). As sequências selecionadas foram introduzidas no programa Primer3

(http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi) (ROZEN; SKALETSKY, 2000) observando-se os parâmetros: temperaturas de *melting* (Tm) ótima de 53°C, máxima de 55°C e mínima de 50°C e, conteúdo de GC entre 40 a 60%. Detalhes das sequências dos *primers* estão apresentados na Tabela 1.

Tabela 1 - *Primers* sintetizados para o desenvolvimento dos marcadores TRAP

Fixo (forward)	**Descrição**	**Sequências (5' - 3')**	**Número de acesso**	**Referência**
LRR173	LRR-Xa21	TCTCAAGCAACAACTTATCC	BQ804007	Rossi et al. (2003)
NBS039	NBS/LRR-PIC17	TCACGATTCCAAGAAGTATG	BQ804022	Rossi et al. (2003)
KIN169	S/T Quinase-Pti1	CAAAGAACACCTGATAATGG	BQ804002	Rossi et al. (2003)
LRR149	LRR-Xa21	GATTGGAGATTGTTCATCAC	BQ804009	Rossi et al. (2003)
NBS152	NBS/LRR-Rp1-D	CAAAAGAAGCAAGGATATG	BQ804044	Rossi et al. (2003)
NBS267	NBS/LRR-RPR1	ATACATGATGCTTTCCAGAC	BQ804047	Rossi et al. (2003)
NBS441	NBS/LRR-I2C-2	GATGATGTATGGAATGAGGA	BQ804028	Rossi et al. (2003)
WRKY	WRKY	CTCGTGCGTCTACTCTACAT	GQ246458	Liu et al. (2012)
PAL	PAL	CCACAGCACAGCATAATG	EF189195	Que et al. (2006)[a]
PR	Quitinase	GCTCTCTTCATCCACATTC	CF576358	Casu et al. (2004)
Arbitrário (reverse)		**Sequências (5' - 3')**		
Arb1		GACTGCGTACGAATTAAT		Li e Quiros (2001)
Arb2		GACTGCGTACGAATTTGC		Li e Quiros (2001)
Arb3		GACTGCGTACGAATTGAC		Li e Quiros (2001)

[a] Sequência submetida diretamente ao banco de dados do NCBI

As reações de amplificação foram realizadas em um volume final de 15 µl, com os seguintes componentes: 20 ng de DNA genômico; 1,5 mM de $MgCl_2$; 0,2 mM de dNTPs; 0,5 µM de *primer* arbitrário e 0,8 µM de *primer* fixo; 1X de tampão de reação Go Taq (Promega®); 1U de Go Taq Flexi DNA Polymerase (Promega®); água Milli-Q autoclavada para completar o volume. A PCR foi realizada em termociclador Gene Amp® PCR System 9700 (Applied Biosystems) e iniciou-se pela desnaturação do DNA a 94°C por 3 min., seguida por duas etapas. A primeira compreendida por 5 ciclos de desnaturação a 94°C por 50 s; anelamento a 35°C por 50 s; extensão a 72°C por 1 min. A segunda etapa foi de 35 ciclos de desnaturação a 94°C por 50 s; anelamento a 50°C por 50 s; extensão 72°C por 1 min., encerrando com uma extensão a 72°C por 7 min.

A cada reação foram adicionados 15 µl de tampão desnaturante (0,2% de azul de bromofenol; 0,2% de xileno cianol; 10 mM EDTA, pH 8,0; 95% de formamida), sendo as amostras desnaturadas a 94°C por 5 min., e mantidas em gelo. Como descrito acima, os produtos de PCR foram resolvidos em eletroforese em gel de poliacrilamida a 5%. A

eletroforese foi realizada por 4 h a uma potência constante de 80 W, utilizando cuba vertical Sequi-Gen® (*Bio-Rad*) 38 x 50 cm. Marcadores de massa molecular (*Ladder*) de 100 e 25 pb foram utilizados para verificar-se o tamanho dos fragmentos amplificados. A revelação dos géis foi realizada pelo método do nitrato de prata (AgNO3) seguindo o protocolo proposto por Creste, Tullman Neto e Figueira (2001).

3.3 Notação das marcas

Os marcadores EST-SSR foram assim nomeados: EST, refere-se aos locos desenvolvidos por Oliveira et al. (2009) e SC àqueles desenvolvidos por Marconi et al. (2011), seguido por uma letra (A para motivos de dinucleotídeos; B para trinucleotídeos e C para tetranucleotídeos) e um código numérico representando o número dos EST-SSR. Os locos desenvolvidos por Jordão Junior (2009), sendo três deles publicados por Maccheroni et al. (2009) e pela CanaVialis/Monsanto, foram nomeados CV seguindo-se um código numérico. Em seguida, para todos os locos estudados, foi adicionado o peso molecular do alelo amplificado, em pares de base. Ao início de cada loco foi adicionada uma letra para denotar a origem do polimorfismo do genitor, de acordo com a metodologia proposta por Wu et al. (2002a).

Os marcadores TRAPs foram identificados de acordo com os domínios caracterizados por Rossi et al. (2003) em: NBS (refere ao domínio NBS-LRR), LRR (domínio LRR) e KIN (domínio S/T KINASE), seguindo-se três dígitos numéricos que representam o RGA caracterizado por Rossi et al. (2003). Os genes envolvidos na resposta de defesa foram nomeados por, PAL (*Phenylalanine ammonia-lyase*), PR (*Chitinase*) e WRKY (referente ao fator de transcrição WRKY). Todos foram seguidos por A1, A2 e A3 correspondendo aos *primers* arbitrários 1, 2 e 3, respectivamente, e pelo peso molecular do fragmento amplificado, em pares de base. Como a exemplo dos marcadores EST-SSR, no início de cada loco foi adicionada a letra correspondente a origem do polimorfismo.

Os marcadores AFLP e os derivados do retrotransposon *scIvana_1* foram notados conforme Palhares et al. (2012), do seguinte modo: para os marcadores AFLP adotou-se a nomenclatura padrão da lista de *primers* de AFLP (Keygene[1]), seguido pelo peso molecular do fragmento, em pares de base. Para os marcadores direcionados ao retrotransposon *scIvana_1*, a notação foi baseada na abreviação da enzima utilizada na reação de digestão (DI

[1] http://www.ncbi.nlm.nih.gov/genomes/PLANTS/AFLP_tbl.html

para *Dra*I, SI para *Ssp*I e RI para *Rsa*I) seguida pelo código do *primer* desenhado para regiões do retrotransposon *scIvana_1* e o peso molecular do fragmento, em pares de base.

Esclarecendo, a origem do polimorfismo foi designada de acordo com a configuração do loco marcador (ou tipo de cruzamento), segundo Wu et al. (2002a). "D1" indica o loco marcador em heterozigose no 'IAC66-6' e homozigose no 'TUC71-7', identificado pela configuração "*ao* x *oo*", segregando na progênie na proporção de 1:1; "D2" corresponde ao loco marcador em homozigose no 'IAC66-6' e heterozigose no 'TUC71-7', identificado pela configuração "*oo* x *ao*", que também segrega na progênie na proporção de 1:1; "C" loco marcador em heterozigose em ambos os genitores com a configuração "*ao* x *ao*", segregando na proporção de 3:1, sendo o alelo *a* dominante (presença da banda) e *o*, o alelo nulo (ausência da banda).

3.4 Análise de segregação

Para a análise de segregação, apenas os marcadores em dose única (WU et al., 1992) foram usados. As bandas que segregaram na progênie da população de mapeamento foram analisadas independentemente como marcadores dominantes, baseado na presença (1) ou ausência (0). Aos dados de genotipagem com marcadores EST-SSR e TRAP foram adicionados aos de AFLP e aos dos marcadores direcionados ao retrotransposon *scIvana_1* obtidos por Palhares (2010). As hipóteses de segregação 1:1 (presença da banda em um genitor e ausência em outro) e 3:1 (presença da banda em ambos os genitores) foram analisadas pelo teste de qui-quadrado (x^2) ao nível de significância de 5% (erro Tipo I). A correção de Bonferroni foi aplicada para limitar o erro experimental associado com o teste múltiplo. O valor do x^2 crítico foi calculado pela divisão do valor de alfa (0,05) pelo número de marcadores. Desta forma, as marcas que não segregaram nas proporções esperadas não foram utilizadas na construção do mapa de ligação. A análise de segregação foi realizada utilizando o programa R versão 2.13.2 (R DEVELOPMENT CORE TEAM, 2011).

3.5 Construção do mapa de ligação integrado

A construção do mapa foi realizada através da estimativa simultânea da fração de recombinação e da fase de ligação entre os marcadores via cadeias de Markov ocultas (HMM), segundo Wu et al. (2002b) e baseada na abordagem multiponto implementada no pacote *OneMap* (MARGARIDO; SOUZA; GARCIA, 2007; MARGARIDO; MOLLINARI;

GARCIA, 2011). Inicialmente, para a determinação dos grupos de ligação foi empregada a análise de dois-pontos (função: '*rf.2pts'*), considerando LOD *score* ≥ 6,0 e fração de recombinação ≤ 0,35. Para os grupos de ligação compostos de até seis marcas, a função '*compare'* foi utilizada para melhor ordená-las dentro de seu respectivo grupo de ligação. Já para os grupos de ligação com sete ou mais marcas foi empregada a função '*order.seq*', possibilitando a construção automatizada dos grupos através da ordenação por busca exaustiva, utilizando os comandos '*compare'* e *'try'*. O primeiro comando forma a sequência base de ordenamento, ordenando os seis marcadores mais informativos do grupo de ligação, enquanto o segundo comando acrescenta e testa os demais marcadores do grupo de ligação para alocá-los na posição de maior verossimilhança. Finalmente, o comando '*ripple*' foi empregado de modo a verificar a ocorrência de inversões locais. As frações de recombinação foram convertidas em distância genética e expressas em centiMorgans (cM), através da função de mapeamento de Kosambi (KOSAMBI, 1944).

A fim de detectar possíveis ligações espúrias, cada grupo de co-segregação foi verificado em um gráfico (*heatmap*) da matriz de fração de recombinação e LOD (função '*rf.graph.table*'). Marcas com pouca evidência de ligação ou que ligaram blocos formados por marcas do tipo "D1" e "D2" foram detectadas a partir dos *heatmaps* e, então, removidas através da função '*drop.marker*'. Em seguida, as demais marcas foram submetidas a um novo ordenamento. Finalmente, as marcas retiradas puderam ser testadas quanto a sua alocação nos subgrupos formados, através da função '*try*'.

Para o desenho do mapa de ligação foi utilizado o software MapChart v. 2.2 (VOORRIPS, 2002).

3.5.1 Atribuição dos grupos de co-segregação em grupos de homo(eo)logia

Inicialmente, os CGs foram atribuídos dentro dos HGs quando pelo menos dois locos de EST-SSR foram compartilhados. Os demais CGs foram acrescentados aos HG pré-estabelecidos quando contiveram um loco EST-SSR em comum (AITKEN; JACKSON; McINTYRE, 2005; OLIVEIRA et al., 2007; PASTINA et al., 2012). Devido à falta de informação decorrente do número limitado de marcadores EST-SSR mapeados, os demais CG, os quais compartilham apenas um alelo do mesmo loco EST-SSR, foram atribuídos dentro de HG putativos (GARCIA et al., 2006; OLIVEIRA et al., 2007; PIPERIDIS et al., 2008). Os HGs foram designados com algarismos romanos, e dentro de cada HG, os CGs mantiverem a sua notação inicial, variando de 1 a 124.

3.6 Avaliação da resistência a *S. scitamineum*

A implementação, condução e fenotipagem da população de mapeamento foi responsabilidade da empresa CanaVialis/Monsanto. Os ensaios de resistência a *S. scitamineum* foram realizados nos anos de 2009 e 2010 na Estação Experimental da CanaVialis/Monsanto, situada na cidade de Conchal, localizada na região Central Oeste do Estado de São Paulo, na altitude de 528 m e pelas coordenadas latitude 22°19'49" e longitude 47°10'21". A implantação dos ensaios ocorreu do modo descrito a seguir.

Os teliósporos do fungo *S. scitamineum* foram obtidos a partir de misturas de diversos "chicotes" coletados de diferentes cultivares no ano de 2008, em Araçatuba, SP. Amostras que apresentaram viabilidade igual ou superior a 95% em ágar-ágar foram utilizadas para o preparo de soluções em pasta, contendo aproximadamente 300 gramas de teliósporos, diluídos em Tween-80 e água MilliQ. Dos 188 indivíduos da população de mapeamento, 176 mais os dois genitores, foram utilizados no primeiro ensaio de inoculação.

Oitenta toletes de cada indivíduo, contendo uma gema, foram inoculados artificialmente através de um ferimento na base da gema, com auxílio de duas agulhas imergidas em solução pastosa de teliósporos. Após inoculação, foram plantados superficialmente em substrato Plantmax®, contido em duas bandejas plásticas, devidamente identificadas e contendo 40 toletes cada. O material foi transferido para a casa de vegetação, em delineamento experimental inteiramente casualizado (DIC), com duas repetições e mantido sob irrigação diária, durante oito meses. Durante este período, realizaram-se tratos culturais como adubação, eliminação manual de plantas daninhas e aplicação de inseticidas.

No primeiro ensaio, a avaliação da incidência de doença iniciou-se aos 74 dias, a contar da data da inoculação (DAI), sendo esta realizada semanalmente durante seis meses (junho a novembro de 2009), totalizando 24 avaliações.

Em cada avaliação os "chicotes" foram retirados das plantas, secos sob luz, peneirados e os teliósporos acondicionados a 4°C em tubos falcon, para então, servirem de inóculo para o próximo experimento, realizado no ano de 2010.

No segundo ensaio, os mesmos critérios de inoculação foram adotados, sendo inoculados 160 indivíduos, dos 188 da população de mapeamento. O delineamento experimental foi o inteiramente casualizado (DIC), com quatro repetições, sendo que cada unidade experimental foi composta de 20 gemas inoculadas. A avaliação da incidência da doença iniciou-se aos 76 dias da inoculação das gemas, sendo esta realizada semanalmente durante cinco meses (julho a novembro de 2010), totalizando 17 avaliações.

Em ambos os ensaios, foram anotados o número de gemas brotadas e este dado foi usado como covariável no modelo estatístico. Em cada uma das avaliaçãoes, foram identificadas, anotadas e extraídas as plantas doentes (“presença de chicote”) de cada repetição, permitindo estabelecer uma curva de progresso da doença no final do ciclo de avaliações. As análises realizadas levaram em conta a área sob a curva de progresso da doença (AUDPC), calculada através da equação (1) proposta por Campbell e Madden (1990):

$$AUDPC = \sum_{i}^{n-1} \left(\frac{y_i + y_{i+1}}{2}\right) * (t_{i+1} - t_i) \quad (1)$$

em que: *n* é o número de avaliações no ciclo e y é o valor de incidência da doença no tempo t. A AUDPC de cada genótipo foi normalizada dividindo-se o respectivo valor pelo tempo total de duração da avaliação.

3.6.1 Análise dos dados fenotípicos

Antes de proceder cada análise, a normalidade dos resíduos foi testada pelo teste de Shapiro-Wilk (SHAPIRO; WILK, 1965). O processo de estimação seguiu a metodologia dos modelos mistos (HENDERSON, 1975), que foi aplicada aos dados de AUDPC. Esta metodologia consiste na obtenção de predições dos valores genotípicos tratados como de efeito aleatório, corrigidos para os demais efeitos fixos contidos no modelo. As estimativas dos componentes de variância foram obtidas pelo método da máxima verossimilhança restrita (*Restricted Maximum Likelihood* – REML) descrito por Patterson e Thompson (1971) e a predição dos valores genotípicos de cada indivíduo foi obtida pelo melhor preditor linear não viesado (*Best Linear Unbiased Prediction* - BLUP).

O modelo estatístico adotado foi, conforme a seguinte equação (2):

$$y_{ijk} = \mu + a_j + ga_{ij} + s_{ijk} + \varepsilon_{ijk} \quad (2)$$

em que:

y_{ijk} é a observação do genótipo *i* no ano *j* na repetição *k*;

μ é a média geral do experimento;

a_j é o efeito fixo do *j*-ésimo ano, sendo *j* = 1, 2;

ga_{ij} é o comportamento do genótipo *i* dentro do ano *j* considerado como efeito aleatório, assumindo que ga ~ N(0, G), em que G é a matriz de variância e covariância não estruturada;

s_{ijk} é o efeito fixo referente ao número de gemas brotadas associado à *ijk*-ésima observação;

ε_{ijk} é o erro associado à *ijk*-ésima observação, assumindo que o erro tem distribuição normal com média 0 e heterogeneidade de variância entre os ambientes (R é uma matriz bloco diagonal)

As estimações dos efeitos fixos e predições dos efeitos aleatórios foram realizadas empregando-se o programa estatístico Genstat 13th edição (PAYNE et al., 2009). Para tanto, foram obtidos BLUPs considerando-se o comportamento dos genótipos dentro de cada ano.

A herdabilidade no sentido amplo de plantas individuais para cada ano foi calculada utilizando a equação (3) proposta em Falconer e Mackay (1996):

$$h_A^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2} \quad (3)$$

em que:

h_A^2 é a herdabilidade no sentido amplo de plantas individuais;

σ_G^2 é a variância genotípica;

σ_E^2 é a variância do resíduo

3.7 Mapeamento de QTLs

O mapa de ligação da população ‘IAC66-6’ x ‘TUC71-7’ foi utilizado para o mapeamento de QTLs associados à resistência ao carvão da cana-de-açúcar. Os dados de BLUPs foram usados para o mapeamento de QTLs, segundo Gazaffi (2009) que propõe o mapeamento por intervalo composto (CIM), porém com base em informação de F_1 segregantes com misturas de tipos de segregação. Esta abordagem também considera os diferentes tipos de segregação que os QTLs podem assumir e suas respectivas fases de ligação com os marcadores.

A partir de uma progênie de irmãos germanos, oriundos do cruzamento entre dois genitores *P* e *Q*, não aparentados e com mapa de ligação conhecido (Figura 5), sendo m e

m+1 marcadores que flanqueiam um intervalo definido, com alelos 1 ou 2, tem-se quatro arranjos para os genótipos destes locos $P_m^{\{1,2\}}$, $P_{m+1}^{\{1,2\}}$, $Q_m^{\{1,2\}}$ e $Q_{m+1}^{\{1,2\}}$, em que {1,2} indicam os possíveis alelos de cada loco. Pressupondo a presença de um QTL neste dado intervalo, cujos alelos são representados por P^1 e P^2, para o genitor P e Q^1 e Q^2, para o genitor Q. Assume-se que os alelos P^1 e Q^1 têm efeito positivo sobre o fenótipo. Tal cruzamento está representado na figura 5.

$$
\begin{array}{ccc}
P & & Q \\
\begin{array}{c|c|c} P_m^1 & & P_m^2 \\ P^1 & & P^2 \\ P_{m+1}^1 & & P_{m+1}^2 \end{array} & \text{x} & \begin{array}{c|c|c} Q_m^1 & & Q_m^2 \\ Q^1 & & Q^2 \\ Q_{m+1}^1 & & Q_{m+1}^2 \end{array}
\end{array}
$$

Figura 5 - Representação do cruzamento entre os genitores P e Q, com os quatro arranjos para os genótipos dos locos quantitativos $P_m^{\{1,2\}}$, $P_{m+1}^{\{1,2\}}$, $Q_m^{\{1,2\}}$ e $Q_{m+1}^{\{1,2\}}$, sendo os alelos dos marcadores m e m+1, P^1, P^2, Q^1 e Q^2 os alelos dos QTLs

As possíveis classes genotípicas para os QTLs são: P^1Q^1, P^1Q^2, P^2Q^1 e P^2Q^2, sendo a respectiva segregação esperada de 1:1:1:1. Sendo assim, é possível definir três contrastes ortogonais entre as médias genotípicas das quatro classes:

$$P^1Q^1 + P^1Q^2 - P^2Q^1 - P^2Q^2$$
$$P^1Q^1 - P^1Q^2 + P^2Q^1 - P^2Q^2$$
$$P^1Q^1 - P^1Q^2 - P^2Q^1 + P^2Q^2$$

Os dois primeiros contrastes referem-se aos efeitos aditivos dos QTLs nos genitores P e Q, respectivamente, e o terceiro refere-se à interação devido à dominância dos efeitos aditivos de cada genitor. Nesta metodologia, o genótipo de cada QTL é inferido com base nos genótipos dos marcadores flanqueadores. Assumiu-se que as probabilidades condicionais para os genótipos dos QTLs são multiponto, sendo que foram obtidas a partir do mapa de ligação construído via HMM, usando o software *OneMap* (MARGARIDO; SOUZA; GARCIA, 2007; MARGARIDO; MOLLINARI; GARCIA, 2011). O mapeamento proposto é explicado pelo seguinte modelo genético-estatístico, eq (4):

$$y_j = Z_j\gamma + \alpha_p^* x_{pj}^* + \alpha_q^* x_{qj}^* + \delta_{pq}^* x_{pj}^* x_{qj}^* + \epsilon_j \qquad (4)$$

em que:

y_j é o valor fenotípico do *j*-ésimo indivíduo (*j* = 1, 2, ..., *n*);

n é o número de indivíduos fenotipados;

Z_j é a *j*-ésima linha da matriz indicadora Z, que contém uma coluna de 1's e variáveis indicadoras dos genótipos dos cofatores selecionados;

γ é o vetor de parâmetros contendo o intercepto (μ) e coeficientes de regressão linear múltipla para os cofatores;

α_p^* é o efeito aditivo para o QTL no genitor P, 'IAC66-6';

α_q^* é o efeito aditivo para o QTL no genitor Q, 'TUC71-7;

δ_{pq}^* é o efeito de dominância entre os alelos do loco considerado;

ϵ_j é o erro experimental com $\epsilon_j \sim N(0, \sigma^2)$;

x_{pj}^* e x_{qj}^* são variáveis indicadores para o genótipo do QTL na progênie, que podem assumir valores 1 ou -1 dependendo da classe genotípica do indivíduo *j*;

Exemplificando, a variável x_{pj}^* assumi valores 1 para as configurações genotípicas P^1Q^1 e P^1Q^2 e -1 para P^2Q^1 e P^2Q^2. Já a variável x_{qj}^* assumi valores de 1 para as configurações genotípicas P^1Q^1 e P^2Q^1 e -1 para P^1Q^2 e P^2Q^2.

Na análise, o genitor *P* refere-se a 'IAC66-6' e o genitor *Q* refere-se a 'TUC71-7'.

O modelo genético-estatístico verifica a existência de QTLs ao longo do genoma, através do teste da razão de verossimilhança (LTR). Para a obtenção das estimativas de máxima verossimilhança foi utilizado o algoritmo EM (DEMPSTER; LAIRD; RUBIN, 1977)

O procedimento de busca por QTL foi realizado através da comparação da verossimilhança dos modelos, considerando a presença de um possível QTL (H_a: pelo menos um efeito diferente de *zero*) *versus* modelo sem QTL (H_0: $\alpha_p^* = \alpha_q^* = \delta_{pq}^* = 0$).

Após a identificação dos QTLs, foram feitas inferências a respeito das fases de ligação dos mesmos, com base nos sinais das estimativas de α_p^* e α_q^*. Na definição dos contrastes, adotou-se a configuração $P_m^1 P^1 P_{m+1}^1 / P_m^2 P^2 P_{m+1}^2$ x $Q_m^1 Q^1 Q_{m+1}^1 / Q_m^2 Q^2 Q_{m+1}^2$, assumindo que os alelos P^1 e Q^1 causam aumento no valor fenotípico. Definiu-se tal configuração como sendo de associação em ambos os genitores, o que implica em estimativas positivas para α_p^* e α_q^*. Caso outras configurações ocorressem, simplesmente, surgiriam mudanças nos sinais das estimativas. No entanto, independente da configuração real, os contrastes definidos para o desenvolvimento do modelo sempre testaram os mesmos efeitos, exceto para os sinais. Neste contexto, a determinação da fase de ligação entre marcadores e QTL pôde ser inferida de

forma simples, com a identificação dos alelos que promovem o aumento ou redução do valor fenotípico, não sendo, portanto necessário incluir as fases de ligação no modelo, como feito por Lin et al. (2003).

Para inferi a segregação do QTL, diversas hipóteses estatísticas apropriadas foram definidas por Gazaffi (2009). Tais hipóteses foram testadas apenas nas posições para as quais QTLs foram mapeados numa primeira etapa, evitando, dessa forma, os problemas de realização de múltiplos testes. Assim, foi estabelecido um procedimento de duas etapas para mapear os QTLs, identificando seus efeitos, a fase de ligação e o padrão de segregação. Inicialmente, fez-se uma busca pelo genoma por possíveis QTLs, testando-se simultaneamente a significância dos três efeitos (α_p^*, α_q^* e δ_{pq}^*), através das hipóteses: H_1: $\alpha_p^* = 0$, H_2: $\alpha_q^* = 0$ e a de dominância H_3: $\delta_{pq}^* = 0$. Se apenas um efeito foi significativo, concluiu-se que há segregação 1:1 e as fases de ligação foram facilmente inferidas. Para os casos em que há mais de um efeito significativo, testes complementares são necessários e são realizados considerando os sinais das estimativas. Para as situações em que há dois efeitos estatisticamente significativos, testes sobre igualdade desses efeitos foram realizados, sendo possível inferir se a segregação era 1:2:1 ou 1:1:1:1. Quando o QTL apresentou todos os efeitos significativos, fez-se necessário testar as três hipóteses complementares (H_4: $|\alpha_p^*| = |\alpha_q^*|$, H_5: $|\alpha_p^*| = |\delta_{pq}^*|$ e H_6: $|\alpha_q^*| = |\delta_{pq}^*|$). No caso em que todas as hipóteses não foram rejeitadas, concluiu-se que a segregação foi 3:1; nas situações em que apenas uma hipótese não foi rejeitada, concluiu-se que a segregação foi 1:2:1. Nos demais casos, concluiu-se que a segregação foi 1:1:1:1. Em todos os casos, as fases de ligação são facilmente deduzidas a partir dos sinais das estimativas (GAZAFFI, 2009).

A seleção dos cofatores foi realizada por regressão linear múltipla entre os fenótipos e os marcadores moleculares, utilizando o critério de informação de Akaike, AIC (AKAIKE, 1974). Com o intuito de evitar a super parametrização do modelo de mapeamento, o número de cofatores foi restringido para até 10 marcadores. Para realizar o mapeamento por intervalo composto (CIM), foram incluídos apenas os marcadores que se encontravam fora do grupo de ligação em estudo. O limiar (*threshold*) utilizado para declarar a presença de QTL foi baseado em 1.000 permutações (CHURCHILL; DOERGE, 1994). Por fim, a variação fenotípica (R^2) explicada pelos QTLs mapeados, foi determinada através do ajuste de um modelo contendo todos os QTLs mapeados pelo modelo CIM, conforme descrito por Gazaffi (2009).

Para a estimativa do valor crítico dos efeitos α_p^*, α_q^* e δ_{pq}^* foi realizado o teste estatístico de qui-quadrado (x^2) para um grau de liberdade, considerando 5% como nível de significância, convertido para LOD s*core*, dividindo pelo fator de correção 4,61.

As análises foram realizadas em ambiente de programação R (R DEVELOPMENT CORE TEAM, 2011), utilizando algoritmo desenvolvido por Gazaffi (2009).

3.8 Isolamento e validação dos marcadores TRAP

Após a genotipagem os fragmentos que segregaram na população foram selecionados e isolados do gel de poliacrilamida. Os fragmentos extraídos foram então eluídos em 100 µl de TE 10:1 (10 mM Tris-HCl, pH 8,0; 1 mM EDTA, pH 8,0) e, posteriormente, transferidos para colunas de filtragem e submetidos a centrifugação a 6000 rpm durante 5 min., em centrífuga de mesa (USA Scientific®). O DNA recuperado foi acondicionado a -20 °C. Baseado em um mapa prévio, 50 fragmentos foram selecionados e reamplificados, purificados e ligados ao vetor de clonagem *pGem-T Easy Vector System* (Promega®), para serem usados na transformação de células competentes de *Escherichia coli* DH5α (HANAHAN, 1983), e então sequenciados.

Tanto a PCR quanto as condições de reamplificação dos fragmentos foram iguais às realizadas para a amplificação dos marcadores TRAP, exceto que 50 ng de DNA genômico foram utilizados como molde, sendo que o volume final da reação foi de 50 µl. Os amplicons foram submetidos à eletroforese em gel de agarose 1,0% (p/v), acrescido de 0,5 X de *SYBR*® *Safe DNA gel stain* (Promega®). Um marcador de massa molecular (*Ladder*) de 100 pb foi utilizado para confirmar o tamanho dos fragmentos. As amostras que apresentaram o tamanho esperado foram purificadas da agarose, com auxílio do kit QIAEX II *Gel Extration* (QIAGEN). Os fragmentos selecionados foram ligados ao vetor de clonagem, conforme as especificações do fabricante do vetor. A ligação foi realizada em volume de 20 µl, contendo: 10 µl de tampão de ligação (2X); 25 ng do vetor; 6 µl dos fragmentos de DNA; 3 U de enzima T4 DNA ligase (Promega®); água destilada deionizada para completar o volume. A ligação ocorreu *overnight* a 4°C.

Para a transformação, 2 µl de reação de ligação foram utilizados para transformar 70 µl de células eletrocompetentes em eletroporador ECM399 (BTX®). As células transformadas foram, imediatamente, ressuspendidas em 1 ml de meio de cultura SOC (SAMBROOK; RUSSELL, 2001) e incubadas a 37°C, 180 rpm, por 1 h. Em seguida, 100 µl do produto da transformação foram planqueados em placas de Petri contendo meio LB, 60 µl de X-GAL 2%

(5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) e 60 µl de IPTG 20% (isopropyl-beta-D-thiogalactopyranoside), as quais foram incubadas a 37°C por 16 h. Logo depois, as placas foram armazenadas a 4°C por 16 h para se fazer a diferenciação das colônias.

Após o crescimento, as colônias de coloração branca (transformadas) foram coletadas com auxílio de palitos estéreis, sendo mergulhados em placas de Elisa contendo meio 200 µl de meio LB líquido com ampicilina a 50 µg/ml. Após o crescimento das colônias no meio líquido, foi feita a extração dos plamídeos (*miniprep*) pelo método de lise alcalina, segundo Sambrook e Russell (2001). A amplificação dos produtos da *miniprep* foi realizada a partir dos *primers* pUC/M13 - 40F e pUC/M13 - 26R, complementares a regiões do vetor de clonagem, de modo a permitir a verificação dos insertos.

As reações foram realizadas em volume de 25 µl, contendo: 2 µl de DNA; 1,5 mM de MgCl2; 0,2 mM de dNTP; 2 µM de cada primer; 1X do tampão de PCR (10 mM Tris-HCl, pH 9,0, 25°C; 50 mM KCl; 0,1% Triton® X-100); 1U da enzima Taq DNA polimerase (*Go Taq Flexi* - Promega®); água destilada deionizada para completar o volume. As amplificações ocorreram com desnaturação inicial a 94°C por 5 min., seguida por 35 ciclos de desnaturação a 94°C por 30 s; anelamento a 55°C por 60 s; e extensão a 72°C por 90 s; extensão final a 72°C por 10 min.

A separação e a verificação do tamanho dos fragmentos foram realizadas por eletroforese em gel de agarose a 1% (p/v), utilizando como padrão de massa molecular o marcador (*Ladder*) 100 pb. Os fragmentos que apresentaram o tamanho esperado foram sequenciados de acordo com o Sanger, Nicklen e Coulson (1977), usando o DYEnamic™ ET Dye Terminator Cycle Sequencing Kit (GE) e o sequenciador ABI PRISM® 3100 Genetic Analyzer (Applied biosystems).

A qualidade das sequências foi avaliada usando-se o pacote computacional *Phred/Phrap/Consed* (EWING et al., 1998), adotando-se como limiar para a aceitação das sequências o valor *Phred/Phrap* igual a 20. As sequências obtidas foram analisadas através do algoritmo BLAST (*Basic Local Alignment*), por comparação com as sequências depositadas no banco de sequência de proteínas não redundante (nr) do NCBI (http://www.ncbi.nlm.nih.gov/).

4 RESULTADOS E DISCUSSÃO

4.1 Desenvolvimento de marcadores moleculares

4.1.1 Seleção dos marcadores EST-SSR

Um total de 254 locos EST-SSR foi utilizado para identificar o polimorfismo entre os genitores e uma amostra de quatro indivíduos tomados aleatoriamente da população F_1 (Figura 6). Baseando-se nos melhores padrões de amplificação e na presença de alelos polimórficos entre os genitores ou, quando monomórficos, os alelos apresentaram indício de segregação na amostra, 60 locos foram selecionados para a genotipagem da população de mapeamento (Anexo A). Vale salientar que a partir desses testes iniciais, foi possível verificar o polimorfismo dos locos, bem como a natureza multialélica dos marcadores EST-SSR.

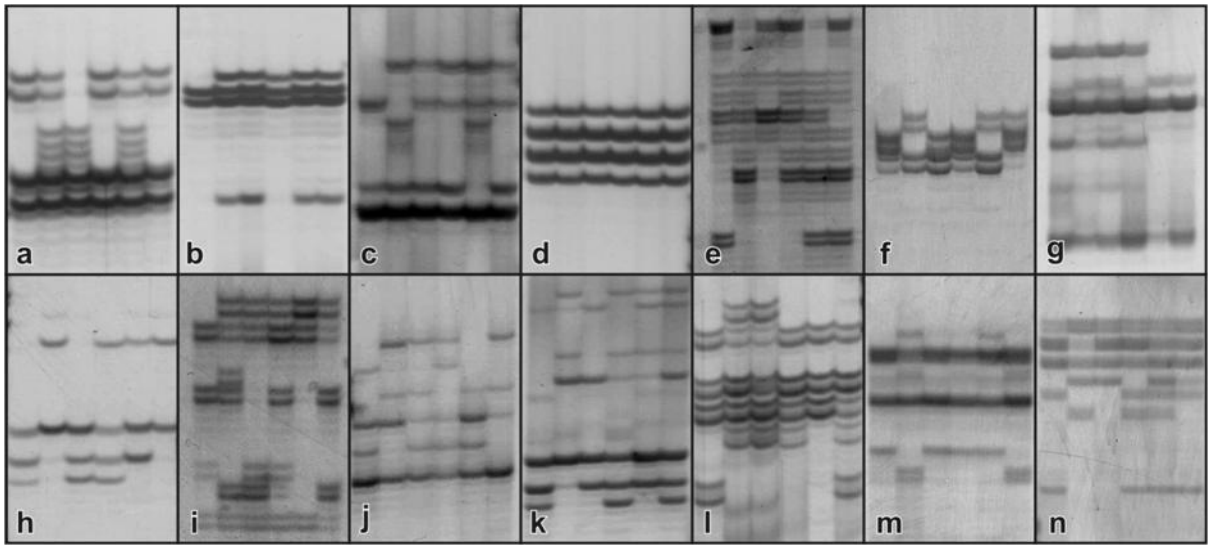


Figura 6 - Resultado dos testes de amplificação para 14 locos de EST-SSR em gel de poliacrilamida desnaturante (5%), corado com nitrato de prata, em que: a) SCB252, b) SCB272, c) SCB225, d) SCB443, e) CV54, f) CV63, g) SCB442, h) SCB243, i) CV53, j) SCB227, k) SCB243, l) SCB329, m) CV12 e n) CV11. A primeira coluna de cada gel corresponde ao genitor 'IAC66-6', a segunda ao 'TUC71-7'; as demais correspondem a quatro indivíduos da progênie F_1

Vinte e seis locos, dos 60 selecionados para a genotipagem, foram submetidos a um novo teste, utilizando 20 indivíduos tomados aleatoriamente da população F_1 e os dois genitores. Isto foi realizado de modo a possibilitar a comparação entre as duas metodologias usadas para a detecção dos alelos e posterior genotipagem da população.

Na figura 7 é apresentado o perfil do loco CV78, cujos alelos foram separados em gel de poliacrilamida desnaturante (5%) ou submetidos à eletroforese capilar. A análise mostra que os mesmos alelos foram detectados em ambas as metodologias de separação, conferindo segurança à leitura dos dados para a genotipagem da população de mapeamento. Em certos

casos, houve alelos anotados como duvidosos e, portanto, não foram considerados na análise de segregação. Exemplificando, há um alelo detectado no gel de poliacrilamida (entre os alelos 173 e 178) que não foi observado na eletroforese capilar e, por isso, foi descartado da análise (Figura 7).

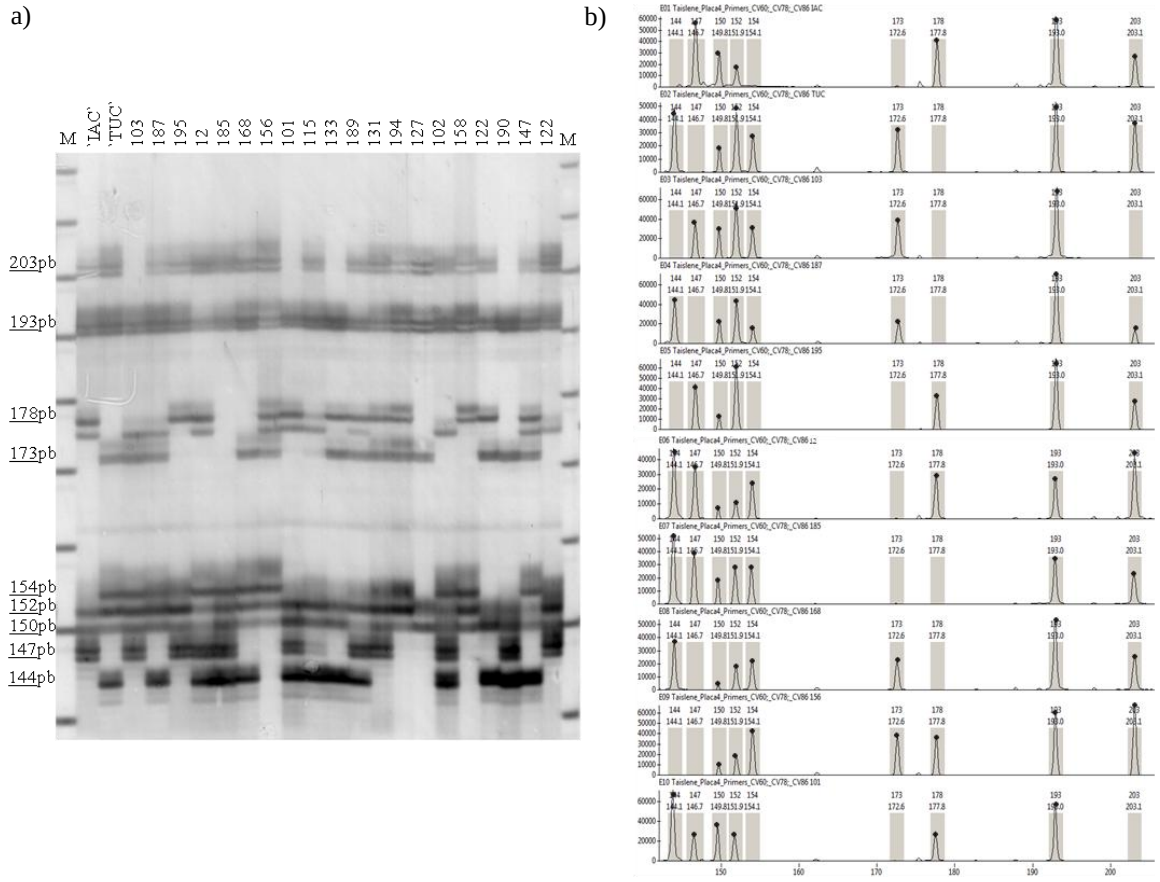


Figura 7 - Genotipagem de 20 indivíduos da progênie F_1 e dos respectivos genitores para o loco CV78. a) Visualização de nove alelos em gel de poliacrilamida desnaturante (5%), corado com nitrato de prata. M: marcador de massa molecular de 10 pb. b) Visualização dos mesmos alelos separados pelo método cromatográfico

O alto grau de variação alélica revelado pelos marcadores microssatélites decorre da variação no número de repetições dos motivos, gerando locos altamente polimórficos que são extremamente úteis para o estudo de mapeamento genético (KALIA et al., 2011). Dentre os locos aqui selecionados, quatro (6,7%) continham repetições de dinucleotídeos, 36 (60,0%) eram de trinucleotídeos, seis (10,0%) de tetranucleotídeos, cinco (8,3%) de pentanucleotídeos, dois (3,3%) de hexanucleotídeos e sete (11,7%) locos foram classificados como compostos. Dentre eles, 34 locos foram caracterizados quanto à função de seus produtos, os quais apresentaram similaridade com diversas proteínas (OLIVEIRA et al., 2009; MARCONI et al.,

2011), enquanto 26 deles não puderam ser divulgados por se tratarem de dados confidenciais da empresa CanaVialis/Monsanto.

4.1.2 Seleção dos marcadores TRAP

No que se refere à técnica TRAP, no total, foram avaliadas 30 combinações de dez *primers* fixos com três *primers* arbitrários, utilizando os genitores e uma amostra de quatro indivíduos da população de mapeamento. Dependendo da combinação, os padrões de amplificação continham de oito a 32 marcas claramente visíveis, com tamanho variando entre 75 a 1.300 pb (Figura 8). A seleção das combinações foi realizada com base no padrão de amplificação e no polimorfismo entre os genitores. Quando os genitores mostraram-se monomórficos, verificou-se se havia indícios de segregação na amostra. Dessa forma, das 30 combinações de *primers* fixos/arbitrários avaliadas, 21 (70,0%) foram selecionadas para a genotipagem. Das combinações, nove correspondem ao *primer reverse* Arb1 com os *primers forward* KIN169, LRR149, LRR173, NBS039, NBS152, NBS267, NBS441, PAL e PR; sete correspondem ao Arb2 com KIN169, LRR149, LRR173, NBS039, NBS 152, NBS 267 e WRKY; e cinco correspondem ao Arb3 com KIN169, LRR149, LRR173, NBS152, NBS267.

A partir dos testes realizados pôde-se associar a detecção do polimorfismo as diferentes combinações dos *primers* arbitrários com os fixos. Por exemplo, o polimorfismo gerado pela combinação do *primer* NBS039 com o *primer* Arb1 foi diferente do polimorfismo produzido pela combinação do mesmo *primer* fixo (NBS039) com os *primers* Arb2 e Arb 3 (Figura 8).

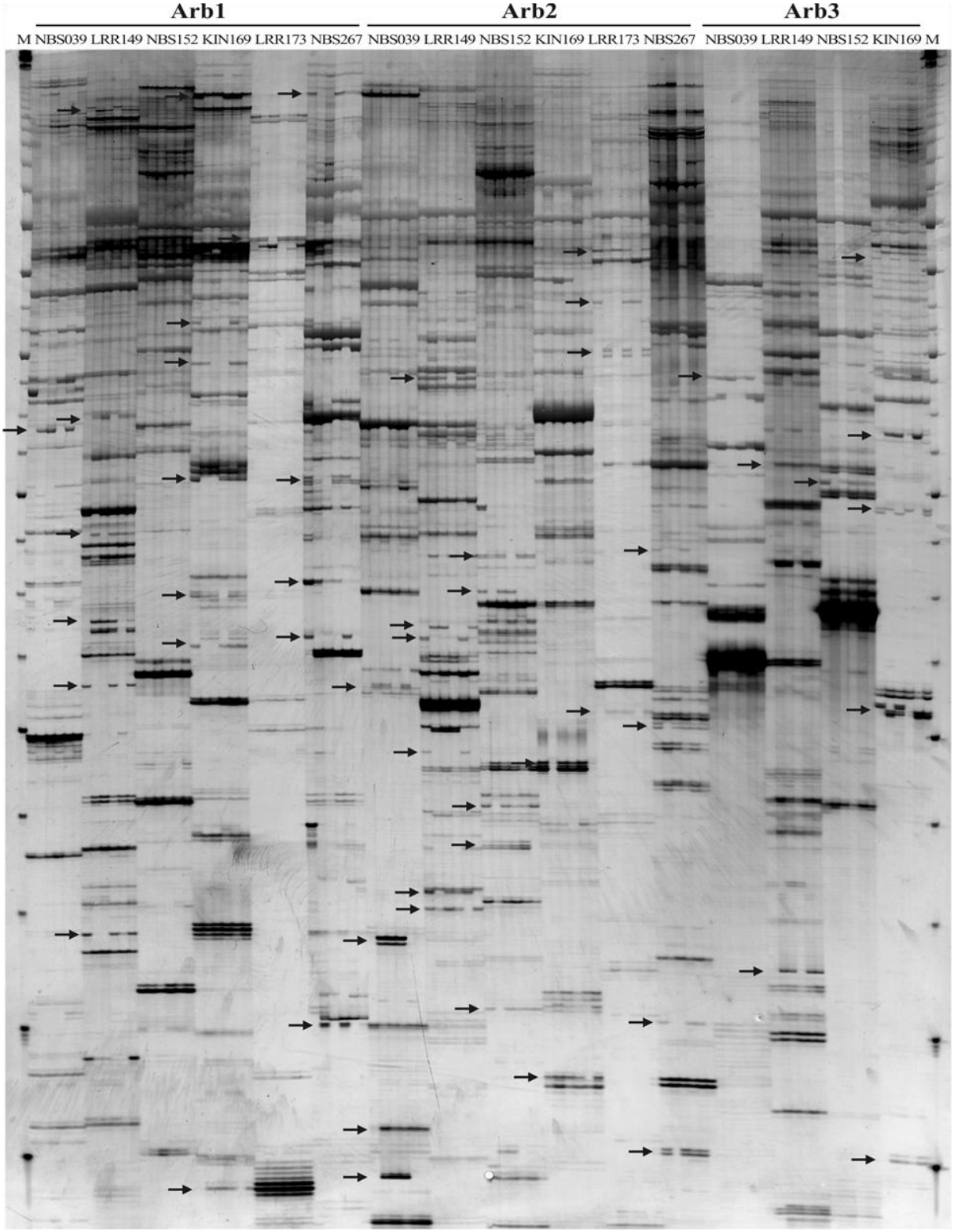


Figura 8 - Resultado dos testes com 16 combinações de *primers* fixos e arbitrários relativos à técnica TRAP. As setas indicam alguns dos locos polimórficos entre os genitores e monomórficos entre eles, porém com segregação em pelo menos um dos indivíduos da F_1. A primeira coluna de cada combinação corresponde ao genitor 'IAC66-6', a segunda ao 'TUC71-7' e as demais correspondem a quatro indivíduos da progênie F_1. M: marcador de massa molecular de 125 pb

4.2 Genotipagem e análise de segregação dos marcadores

A genotipagem da progênie F_1 e dos genitores foi tomada quanto à presença (1) ou ausência (0) das bandas independente do tipo de marcador. Fragmentos gerados pelos locos de EST-SSR SCB422 e SCB437 e pela combinação primer fixo/aleatório KIN169A3 podem ser visualizados nas figuras 9 e 10, respectivamente.

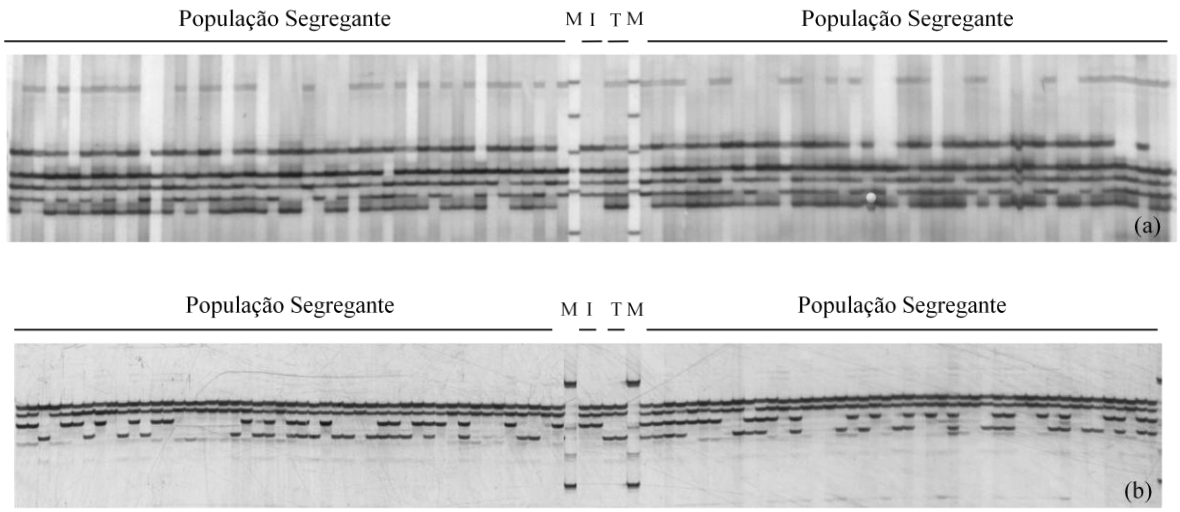


Figura 9 - Genotipagem em gel de poliacrilamida (5%) da população F_1 com locos de EST-SSR: (a) SCB422 mostrando os alelos de 216 a 249 pb; (b) SCB437 mostrando alelos de 439 a 448 pb. Nas canaletas centrais estão os genitores 'IAC66-6' (I) e 'TUC71-7' (T) em duplicata, flanqueados pelos marcadores de massa molecular de 10 pb (M)

Um total de 773 marcas foi detectado após a genotipagem de 188 indivíduos da progênie F_1 com 60 pares de *primers* EST-SSR e 21 combinações TRAP. Das 773 marcas analisadas, 468 (60,5%) segregaram na progênie, sendo 150 (32,1%) marcas oriundas de 'IAC66-6', 182 (38,9%) de 'TUC71-7' e 136 (29,1%), de ambos os genitores. Comparando os dois marcadores, é possível verificar que os EST-SSRs contribuíram com o maior número de alelos segregantes (62,8%). Quando analisado o polimorfismo entre os genitores para cada marcador, individualmente, nota-se que o genitor 'TUC71-7' contribuiu ligeiramente mais que o genitor 'IAC66-6' para o polimorfismo gerado pelos marcadores EST-SSRs, enquanto para os marcadores TRAP essa diferença não pôde ser observada (Tabela 2).

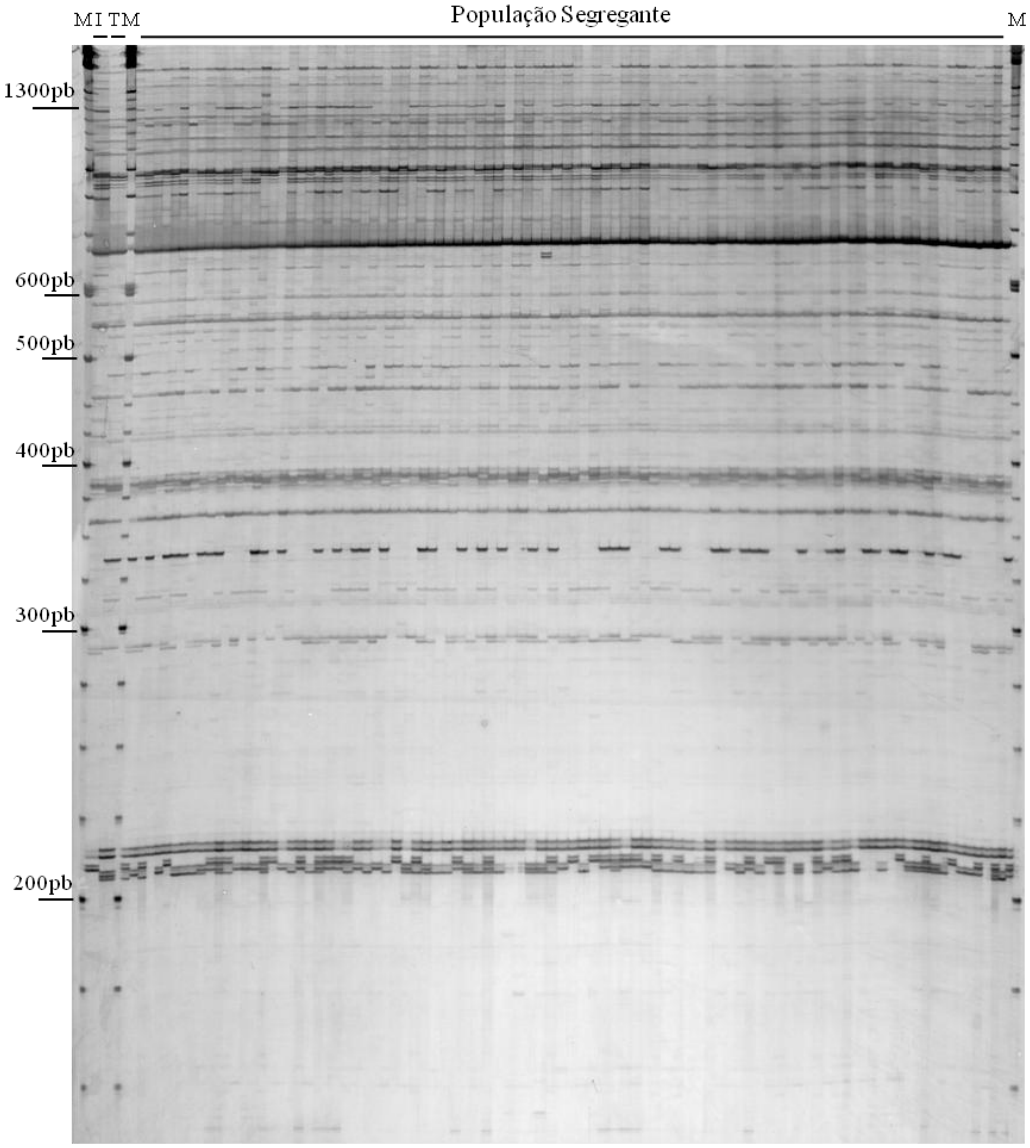


Figura 10 - Genotipagem da população F_1 com a combinação KIN169A3. Nas canaletas 2 e 3 está o genitor 'IAC66-6" (I) e nas canaletas 4 e 5, o genitor 'TUC71-7' (T). Na primeira, sexta e última canaletas estão os marcadores de massa molecular de 125 pb

Tabela 2 – Número de marcadores EST-SSR e TRAP avaliados e respectiva contribuição de cada genitor

Marcadores	EST-SSR (1)	TRAP (1)	Total (1)
Total de marcas amplificadas	382	391	773
Total de marcas segregantes	294 (62,8)	174 (37,2)	468 (60,5) (2)
Marcas segregantes oriundas do genitor 'IAC66-6'	82 (17,5)	68 (14,5)	150 (32,0)
Marcas segregantes oriundas do genitor 'TUC71-7'	117 (25,0)	65 (13,9)	182 (38,9)
Marcas segregantes oriundas de ambos os genitores	95 (20,3)	41 (8,8)	136 (29,1)

(1) Porcentagem em relação ao total de marcas segregantes (468)
(2) Porcentagem em relação ao total de marcas amplificadas (773)

A análise de segregação dos MDU (1:1 e 3:1) foi feita com base no teste de aderência, x^2 ($\alpha = 0,05$), corrigida pelo método de Bonferroni. Detalhes de cada loco EST-SSR e TRAP podem ser visualizados nas tabelas 3 e 4, respectivamente.

Tabela 3 - Genotipagem da população de mapeamento com locos EST-SSR

(continua)

Loco EST-SSR	N° de alelos observados	N° alelos segregantes [1]	Segregação 1:1 [2]		Segregação 3:1 [2]	MDU [2], [3]
			'IAC66-6'	'TUC71-7'		
CV06	3	3 (100,0)	1	2	0	3 (100,0)
CV11	6	6 (100,0)	1	1	0	2 (33,3)
CV22	8	7 (87,5)	1	4	0	5 (71,4)
CV23	9	8 (88,9)	3	1	3	7 (87,5)
CV24	7	4 (57,1)	0	3	0	3 (75,0)
CV29	10	9 (90,0)	1	1	3	5 (55,6)
CV37	6	5 (83,3)	0	2	2	4 (80,0)
CV38	9	8 (88,9)	1	3	2	6 (75,0)
CV46	5	4 (80,0)	0	1	1	2 (50,0)
CV51	3	2 (66,7)	1	0	0	1 (50,0)
CV53	7	7 (100,0)	0	2	1	3 (42,9)
CV58	6	6 (100,0)	0	2	0	2 (33,3)
CV78	9	9 (100,0)	2	3	1	6 (66,7)
CV79	7	6 (85,7)	0	2	1	3 (50,0)
CV86	5	4 (80,0)	0	1	0	1 (25,0)
CV91	8	8 (100,0)	0	2	0	2 (25,0)
CV94	4	3 (75,0)	0	0	1	1 (33,3)
CV100	9	7 (77,8)	2	2	2	6 (85,7)
CV101	11	8 (72,7)	1	2	0	3 (37,5)
CV104	6	6 (100,0)	1	0	3	4 (66,7)
CV106	6	5 (83,3)	1	1	0	2 (40,0)
CV115	6	5 (83,3)	1	0	1	2 (40,0)
CV119	5	4 (80,0)	2	0	1	3 (75,0)
CV128	5	4 (80,0)	0	1	1	2 (50,0)
CV135	6	3 (50,0)	1	0	0	1 (33,3)
CV144	5	3 (60,0)	0	1	0	1 (33,3)
ESTA26	4	3 (75,0)	1	0	0	1 (33,3)
ESTA48	10	9 (90,0)	3	1	1	5 (55,6)
ESTA53	9	7 (77,8)	2	2	1	5 (71,4)
ESTA61	5	4 (80,0)	1	0	1	2 (50,0)
ESTB14	9	4 (44,4)	1	2	0	3 (75,0)
ESTB55	3	2 (66,7)	0	1	0	1 (50,0)
ESTB60	8	5 (62,5)	2	2	0	4 (80,0)
ESTB82	5	3 (60,0)	0	1	0	1 (33,3)
ESTB94	10	8 (80,0)	3	3	0	6 (75,0)
ESTB99	4	3 (75,0)	0	2	0	2 (66,7)
ESTB100	13	11 (84,6)	2	3	2	7 (63,6)
ESTB118	5	3 (60,0)	2	1	0	3 (100,0)
ESTB128	5	4 (80,0)	0	1	1	2 (50,0)
ESTC80	7	5 (71,4)	2	1	1	4 (80,0)
SCB189	5	2 (40,0)	0	1	0	1 (50,0)
SCB191	5	4 (80,0)	0	2	0	2 (50,0)
SCB219	5	5 (100,0)	0	1	1	2 (40,0)

Tabela 3 - Genotipagem da população de mapeamento com locos EST-SSR

(conclusão)

Locos EST-SSR	N° de alelos observados	N° alelos segregantes [1]	Segregação 1:1 [2]		Segregação 3:1 [2]	MDU [2], [3]
			'IAC66-6'	'TUC71-7'		
SCB225	5	4 (80,0)	0	1	1	2 (50,0)
SCB227	6	6 (100,0)	2	1	0	3 (50,0)
SCB243	4	4 (100,0)	1	1	0	2 (50,0)
SCB246	9	6 (66,7)	1	3	1	5 (83,3)
SCB252	4	3 (75,0)	0	2	0	2 (66,7)
SCB272	3	2 (66,7)	0	1	0	1 (50,0)
SCB312	9	5 (55,5)	2	1	1	4 (80,0)
SCB329	6	4 (66,7)	1	2	1	4 (100,0)
SCB344	7	5 (71,4)	2	1	2	5 (100,0)
SCB365	7	3 (42,8)	2	0	0	2 (66,7)
SCB381	8	7 (87,5)	2	0	1	3 (42,9)
SCB397	7	3 (42,8)	0	1	1	2 (66,7)
SCB422	6	4 (66,7)	0	1	1	2 (50,0)
SCB437	4	3 (75,0)	1	1	0	2 (66,7)
SCB442	6	4 (66,7)	0	2	1	3 (75,0)
SCB477	4	3 (75,0)	0	1	0	1 (33,3)
SCB502	4	2 (50,0)	0	1	1	2 (100,0)
Total	382	294 (77,0)	53	81	42	176 (59,9)
Média	6,4	4,9	0,9	1,4	0,7	2,8
Amplitude	3-13	2-11	0 - 3	0 - 4	0 - 3	1-7

[1] Entre parênteses é apresentada a porcentagem de alelos segregantes em relação ao número de alelos observados

[2] Dados obtidos após correção de Bonferroni

[3] Entre parênteses é apresentada a porcentagem de alelos em dose única tomada em relação ao número de alelos segregantes

Nove locos EST-SSR aqui utilizados foram comuns àqueles usados nos trabalhos desenvolvidos por Oliveira et al. (2007) e Pastina et al. (2012), que também estudaram para fins de mapeamento uma população derivada de cruzamento bi-parental. São eles: ESTA48, ESTA53, ESTA61, ESTB60, ESTB94, ESTB99, ESTB100, ESTB118 e ESTC80.

Devido à natureza poliploide da cana-de-açúcar, para muitos dos locos EST-SSR foram detectados três ou mais produtos de amplificação (Figuras 6, 7 e 9; Tabela 3). Os 60 locos utilizados para a genotipagem produziram 382 alelos; observou-se uma amplitude de 3 a 13 alelos e uma média de 6,4 alelos por loco. Das 382 marcas observadas, 294 (77,0%) segregaram na população de mapeamento, ou seja, observou-se uma média de 4,9 alelos segregantes por loco. Destas, 176 alelos (59,9%) foram MDU (1:1 e 3:1), sendo 53 (30,1%) originados do genitor 'IAC66-6', 81 (46,0%) do genitor 'TUC71-7' e 42 (23,9%) originados de ambos os genitores. O número de MDU identificado variou de 1 a 7 por loco (Tabela 3).

Cento e dezoito marcas (40,1%) não segregaram segundo as hipóteses testadas e, portanto, estariam segregando em outras doses (assumindo octaploidia, nas proporções 13:1 e 69:1, quando mais de uma cópia do alelo está presente em apenas um dos genitores e 25:3, 27:1, 139:3, 977:3, 195:1 e 4.899:1, quando mais de uma cópia do alelo está presente em ambos os genitores) e não foram utilizadas na construção do mapa. O trabalho realizado por Oliveira et al. (2007) relatou que apenas 10,0% das marcas EST-SSR não segregaram como MDUs.

Dos 60 locos EST-SSR aqui estudados, 41 já haviam sido utilizados na construção do mapa prévio publicado por Palhares et al. (2012). Quanto ao número médio de alelos amplificados, valores similares foram observados para os seis motivos estudados, sendo que repetições de dinucleotídeos revelaram, em média, 7,0 alelos por loco; trinucleotídeos, 6,2; tetranucleotídeos, 6,5; pentanucleotídeos, 6,8; hexanucleotídeos, 5,5. Já os locos de microssatélites classificados como compostos revelaram média de 6,8 alelos por loco.

Entre as 30 combinações entre *primers* fixos e arbitrários usadas para amplificar marcas TRAP, 21 foram selecionadas para a genotipagem da população F_1. Estas combinações revelaram 391 locos, observou-se uma amplitude de 10 a 32 locos e uma média de 18,6 locos por combinação. Desse total, 174 locos (44,5%) segregaram na população de mapeamento, ou seja, observou-se uma média de 8,3 locos segregantes por combinação. Observaram-se 128 (73,6%) MDUs (1:1 e 3:1), sendo 51 (29,3%) amplicons originados do 'IAC66-6', 55 (31,6%) do 'TUC71-7' e 22 (12,6%) originados de ambos os genitores. O número de MDU identificado variou de 1 a 18 (Tabela 4). Quarenta e seis marcas (26,4%) apresentaram segregações diferentes das esperadas e não foram utilizadas na construção do mapa. A quantidade de marcas TRAP que não segregaram nas proporções esperadas, identificadas neste trabalho, estão de acordo com o que foi observado por Alwala et al. (2008) para a população de mapeamento derivada do cruzamento bi-parental *S. officinarum* 'La Striped' x *S. spontaneum* 'SES 147B'.

Tabela 4 – Genotipagem da população de mapeamento com marcadores TRAP

Combinação de TRAP	N° de locos observados	N° de locos segregantes [1]	Segregação 1:1 [2]		Segregação 3:1 [2]	MDU [2], [3]
			IAC66-6'	TUC71-7'		
KIN169A1	30	11 (36,7)	3	1	2	6 (54,5)
KIN169A2	15	3 (20,0)	0	1	1	2 (66,7)
KIN169A3	32	14 (43,8)	2	5	2	9 (64,3)
LRR149A1	25	11 (44,0)	2	4	0	6 (54,5)
LRR149A2	31	17 (54,8)	5	6	1	12 (70,6)
LRR149A3	23	11 (47,8)	2	6	2	10 (90,9)
LRR173A1	13	4 (30,8)	1	1	1	3 (75,0)
LRR173A2	11	4 (36,4)	1	3	0	4 (100,0)
LRR173A3	10	4 (40,0)	2	2	0	4 (100,0)
NBS039A1	19	7 (36,8)	5	2	0	7 (100,0)
NBS039A2	12	5 (41,7)	2	2	0	4 (80,0)
NBS152A1	16	3 (18,8)	1	0	1	2 (66,7)
NBS152A2	13	8 (61,5)	2	1	1	4 (50,0)
NBS152A3	14	5 (35,7)	3	0	0	3 (60,0)
NBS267A1	20	13 (65,0)	3	3	1	7 (53,8)
NBS267A2	15	8 (53,3)	4	0	0	4 (50,0)
NBS267A3	12	6 (50,0)	2	4	0	6 (100,0)
NBS441A1	21	9 (42,9)	3	2	1	6 (66,7)
PALA1	31	20 (64,5)	6	8	4	18 (90,0)
PRA1	12	4 (33,3)	1	2	1	4 (100,0)
WRKYA2	16	7 (43,8)	1	2	4	7 (100,0)
Total	391	174 (44,5)	51	55	22	128 (73,6)
Média	18,6	8,3	2,4	2,6	1,0	6,1
Amplitude	10-32	3-20	0-6	0-8	0-4	1-18

[1] Entre parênteses é apresentada a porcentagem de locos segregantes em relação ao número de locos observados
[2] Dados obtidos após correção de Bonferroni
[3] Entre parênteses é apresentada a porcentagem de MDU tomada em relação ao número de marcas segregantes

4.3 Construção do mapa de ligação e atribuição dos grupos de co-segregação nos grupos de homo(eo)logia

As 294 marcas EST-SSR e 174 marcas TRAP foram integradas as 685 marcas AFLP e 87 marcas direcionadas ao retrotransposon *scIvana_1* obtidas por Palhares (2010). Assim, um total de 1.240 marcas foi genotipado na população composta por 188 indivíduos F_1 derivada do cruzamento entre 'IAC66-6' e 'TUC71-7'. Estas foram classificadas de acordo com Wu et al. (2002a) em: "D1" (*ao x oo*), quando o alelo (ou banda) está presente apenas no genitor 'IAC66-6'; "D2" (*oo x ao*), quando o alelo (ou banda) está presente apenas no genitor 'TUC71-7' e "C" (*ao x ao*), quando ambos são heterozigóticos para presença dos alelos (ou bandas). Assim, das 1.240 marcas segregantes, 905 (73,0%) MDU, as quais incluem 176 marcas gerados por 60 locos EST-SSR, 128 marcas geradas por 21 combinações *primers*

fixos/arbitrários de TRAP, 536 marcas geradas por 22 combinações enzimas/*primers* de AFLP e 65 marcas geradas por seis combinações enzimas/*primers* direcionadas ao retrotransposon *scIvana_1* (Tabela 5).

Os resultados obtidos quanto a porcentagem esperada de MDU corroboram com os descritos por Da Silva et al. (1995), Aitken, Jackson e McIntyre (2005), Oliveira et al. (2007) e Alwala et al. (2008), quando se avaliam populações derivadas de cruzamentos bi-parentais.

Tabela 5 – Características dos marcadores usados para genotipar a população de mapeamento

Tipo de marcador	EST-SSR	TRAP	AFLP	*scIvana_1*	Total
Número de marcas amplificadas	382	391	3.094	357	4.224
Número de marcas segregantes	294	174	685	87	1.240
Número de marcas polimórficas entre os genitores	199	133	488	74	894
Número de marcas monomórficas entre os genitores	95	41	197	13	346
Número de marcadores em dose única (MDU) [(1)]	176	128	536	65	905
MDU de origem do genitor 'IAC66-6' [D1 (*ao* x *oo*)] [(1), (2)]	53	51	197	23	324
MDU de origem do genitor 'TUC71-7' [D2 (*oo* x *ao*)] [(1), (2)]	81	55	192	32	360
MDU de origem de ambos os genitores [C (*ao* x *ao*)] [(1), (3)]	42	22	147	10	221
Número de marcas que não segregaram nas proporções esperadas	118	46	149	22	335

[(1)] Dados obtidos após correção de Bonferroni
[(2)] Marcas presentes em apenas um dos genitores com a segregação 1:1 na população de mapeamento
[(3)] Marcas presentes em ambos os genitores com a segregação de 3:1 na população de mapeamento

Observando-se as genealogias dos genitores (Anexo B), percebe-se que o genitor ‘IAC66-6’ é oriundo do cruzamento entre as variedades ‘CO419’ x ‘CO356’. Estas, por sua vez, foram obtidas a partir do cruzamento entre as variedades ‘POJ2878’ x ‘CO290’ e ‘POJ2725’ x ‘Sorghumdur’, respectivamente, sendo ‘POJ2878’ e ‘POJ2725’ irmãs germanas, tendo em comum os genitores ‘POJ2364’ e ‘EK28’, apresentando certo grau de endogamia. Isto explica a sua menor contribuição à divergência entre os genitores da população de mapeamento aqui estudada. Exemplificando, para os locos de EST-SSRs o genitor ‘IAC66-6’ contribuiu com 53 alelos enquanto o ‘TUC71-7’ com 81 alelos. Assim também ocorreu para os demais marcadores, com exceção dos AFLPs (Tabela 5).

Como um grande número de marcas foi analisado, os grupos de co-segregação (CGs) foram construídos utilizando LOD *score* ≥ 6,0, de modo a evitar falsas ligações. Inicialmente, das 905 marcas utilizadas na construção do mapa, 729 (80,6%) formaram 110 CGs, sendo que para 176 (19,4%) não foi possível estabelecer a ligação.

Este número é semelhante ao encontrado por Palhares et al. (2012) e Alwala et al. (2008), porém foi menor que o encontrado por Garcia et al. (2006) e Oliveira et al. (2007)

porém maior que o relatado por Hoarau et al. (2001), Aitken, Jackson e McIntyre (2005), Reffay et al. (2005), Raboin et al. (2006) e por Andru et al. (2011).

Dos 110 CGs, três foram desconsiderados por conterem só duas marcas sobrepostas com comprimento de 0,0 cM, não caracterizando, portanto, um CG. Todos os CGs foram individualmente analisados quanto à insuficiência de marcas do tipo C, quanto a presença de ligações espúrias e *gaps*. Para tanto, foi utilizada a matriz gráfica de fração de recombinação e LOD *score* (*heatmap*) disponível no programa computacional *OneMap*, (MARGARIDO; SOUZA; GARCIA, 2007; MARGARIDO; MOLLINARI; GARCIA, 2011), através do comando '*rf.graph.table'*. Baseado nestas observações, nove GCs foram reanalisados, subdivididos e reordenados.

Tomando como exemplo o CG 5, pode-se verificar que ele é formado por dois grupos contendo marcas com valores baixos da fração de recombinação entre si; no entanto, estão ligados devido a uma única marca [174 (D2-E31M49170)], fortemente ligada a marcas pertencentes aos dois grupos (Figura 11). Quando esta marca foi removida, utilizando o comando '*drop.marker*', este GC foi dividido em dois novos grupos: o primeiro grupo formado por sete marcas e o segundo formado por nove marcas. Para cada novo grupo foi realizada uma nova ordenação usando o comando *make.seq*, seguido do *order.seq* para o CG com nove marcas e *compare* para o CG contendo sete marcas. Em seguida, foi executado o comando *try.seq* de modo a recolocar a marca 174, cuja ligação se estabeleceu no primeiro grupo de ligação.

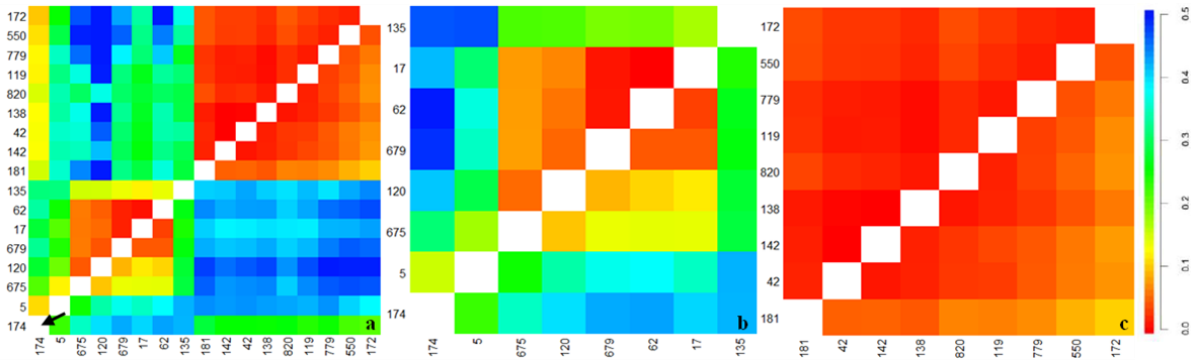


Figura 11 - Matriz gráfica de fração de recombinação e LOD *score* para o CG 5. a) *heatmap* do CG 5 mostrando a marca espúria [174 (D2-E31M49170)] ligada a marcas pertencentes aos dois grupos que, depois de removida, resultou na formação de dois novos CGs (b e c)

As marcas que segregam 3:1 (tipo C) na população são importantes para estabelecer a integração dos mapas construídos pelas marcas originadas de cada genitor [tipo D1 e D2, segundo a metodologia de WU et al. (2002a)]. Assim, um grande número deste tipo de marca é necessário para a integração dos mapas (GARCIA et al. 2006). Na figura 12 está

representada a matriz gráfica do CG 44, na qual a insuficiência de marcas que segregam 3:1 (tipo C) foi detectada, resultando na subdivisão do CG.

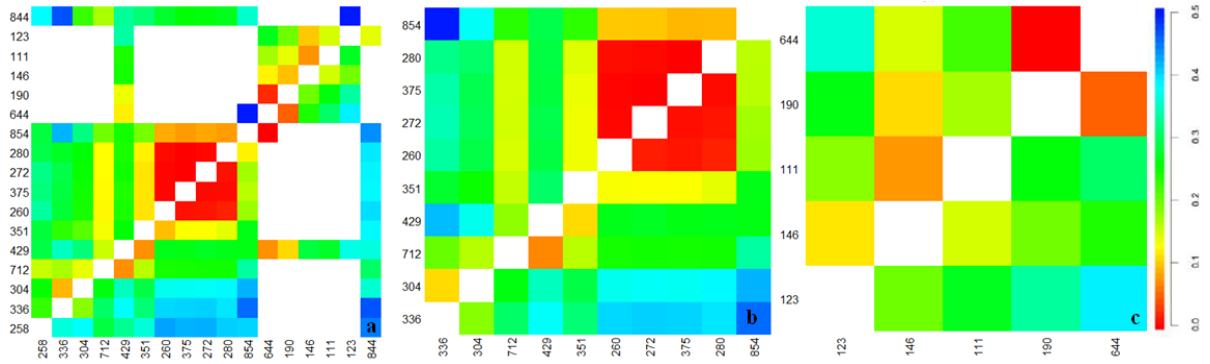


Figura 12 - Matriz gráfica de fração de recombinação e LOD *score* para o CG 44 a) *heatmap* do CG 44 mostrando a insuficiência de marcas do tipo C.8 para a integração. As marcas 123, 111, 146, 190 e 644 são do tipo D2; as marcas 844, 854, 280, 272, 375, 260, 351, 712, 304, 336 e 258 são do tipo D1 e a marca 429 é do tipo C. b e c) Matrizes gráficas após a divisão manual e o reordenamento do CG 44

Assim, após a reanálise de todos os CGs, 721 MDU foram distribuídos em 124 CGs e renumerados de 1 a 124 (Figura 13).

Reanalisando a origem dos 721 MDUs verificou-se que 267 (37,0%) são provenientes do genitor ‘IAC66-6’ (D1), alocados em 60 GCs, 318 (44,1%) são marcas provenientes do genitor ‘TUC71-7’ (D2), alocadas em 73 CGs e o restante, 136 (18,9%) são comuns entre os genitores (C), as quais foram usadas para integrar o mapa de ligação de cada genitor (Tabela 6). Embora os marcadores do tipo C sejam menos informativos geneticamente, já que há mistura de classes na segregação 3:1, atuam como pontes entre os seus genomas, identificando regiões de homologia entre eles.

Utilizando as marcas que segregaram na proporção 3:1 como pontes, 20 CGs foram integrados (2, 5, 6, 11, 20, 23, 24, 25, 30, 32, 36, 42, 45, 47, 51, 55, 56, 68, 75 e 76). O número de marcas por CG variou de dois a 31 (Figura 13). O comprimento dos CGs variou de 1,5 cM (122) a 305,4 cM (36), com uma estimativa média de comprimento de 50,2 cM. O tamanho acumulativo do mapa foi de 6.223,0 cM. A distância média entre as marcas foi de 8,6 cM, sendo a maior distância entre elas igual a 42 cM (23), e a média de marcas por CG igual a 5,8 (Figura 13). Nota-se que o mapa não é saturado, dado ao pequeno tamanho de muitos CGs e aos *gaps* existentes. Quanto ao número de CGs (124) identificados neste estudo, este está dentro da margem para as cultivares modernas que apresentam 2n = 100 – 130.

É importante salientar que o mapa aqui construído é constituído apenas de marcadores em dose única; marcadores mais informativos como os SNPs que possibilitam inferir as ploidias de cada loco e identificar marcas com diferentes segregações (tipo 1:2:1 e 1:6:6:1, considerando autooctaploidia) poderiam cobrir regiões genômicas aqui não detectadas. Quando se utilizam MDUs, cabe comentar que apenas são detectadas ligações entre locos em associação, já que o outro alelo é nulo. Também, o tamanho das populações usadas na construção dos mapas de cana-de-açúcar é um fator que limita as chances de deteção de genótipos raros, reforçando a necessidade do desenvolvimento de metodologias de mapeamento que incluam as demais doses.

Tabela 6 – Marcadores em dose única alocados no mapa de ligação da população ‘IAC66-6’ x ‘TUC71-7’

Tipo de marcadores	EST-SSR	TRAP	AFLP	*scIvana_1*	Total
Total de marcas ligadas	153	95	415	58	721
Número de marcas ligadas de origem do genitor 'IAC66-6' [D1 (*ao* x *oo*)]	49	38	159	21	267
Número de marcas ligadas de origem do genitor 'TUC71-7’ [D2 (*oo* x *ao*)]	70	45	172	31	318
Número de marcas ligadas de origem de ambos os genitores [C (*ao* x *ao*)]	34	12	84	6	136

A análise dos locos EST-SSR revelou que 56 deles geraram 153 marcas que foram distribuídas em 71 CGs, variando de um a nove marcas por CG. Cinco destes CGs foram construídos apenas com EST-SSRs (111, 112, 114, 115 e 116). Do mesmo modo que o observado para os demais marcadores aqui utilizados, os EST-SSRs foram distribuídos desuniformemente nos grupos de co-segregação (Figura 13).

Analisando os marcadores TRAPs direcionados a genes envolvidos na resposta de defesa, verifica-se que eles não foram uniformemente distribuídos nos grupos de co-segregação. Foram mapeados em 58 dos 124 grupos de co-segregação, o que equivale a 74,2% das marcas TRAP alocadas no mapa. Porcentagens semelhantes foram observadas para o marcador AFLP (77,4%); no entanto, os marcadores EST-SSR e *scIvana_1* apresentaram 86,9% e 89,2% das marcas alocadas, respectivamente. A adição de mais marcadores EST-SSR e TRAP ao mapa prévio resultou na recuperação de 20 MDU de AFLPs e dois marcadores direcionados ao retrotransposon *scIvana_1*.

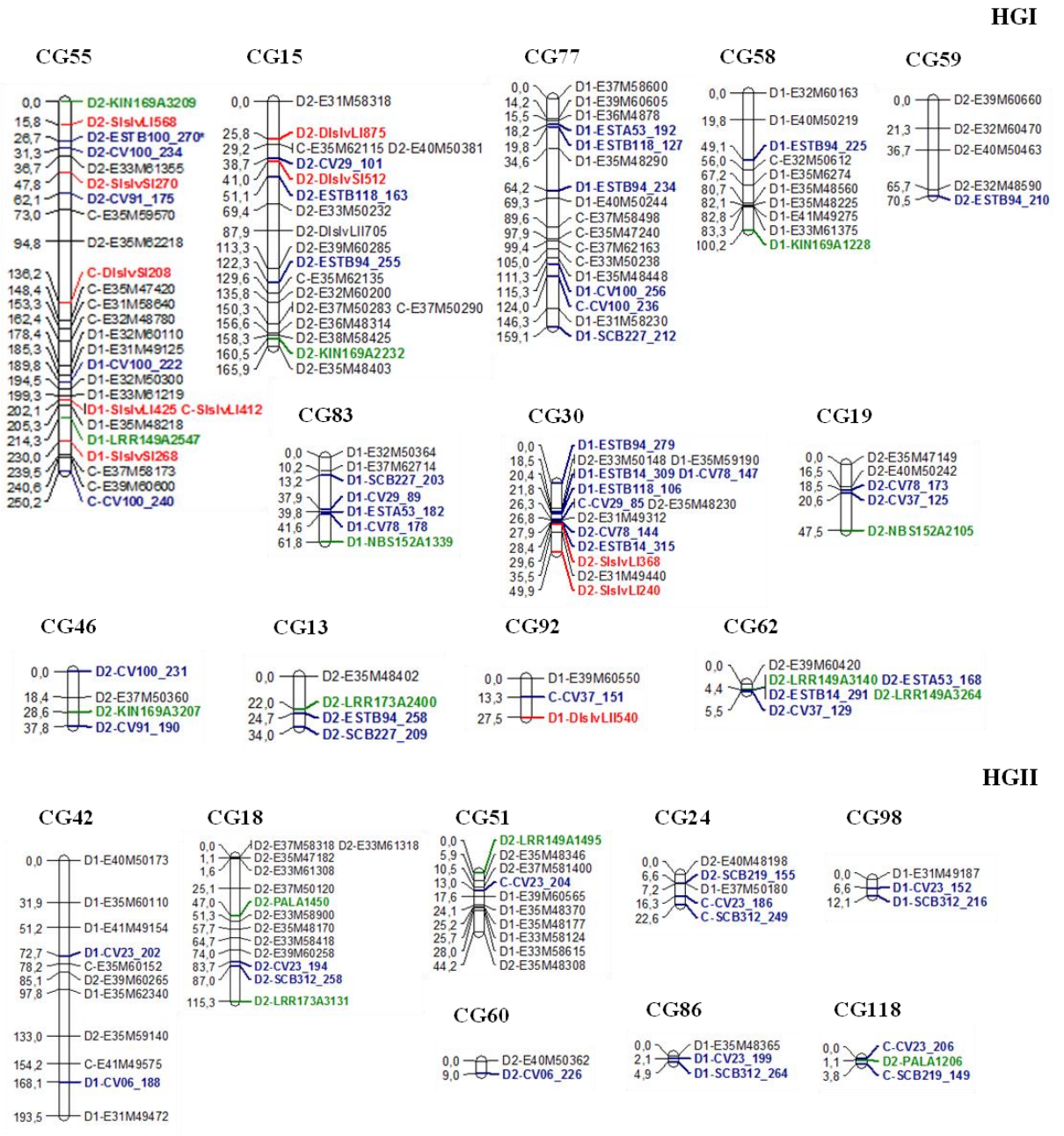


Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre 'IAC66-6' e 'TUC71-7'. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no 'IAC 66-6'), D2 (presença no 'TUC 71-7') ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos

(continua)

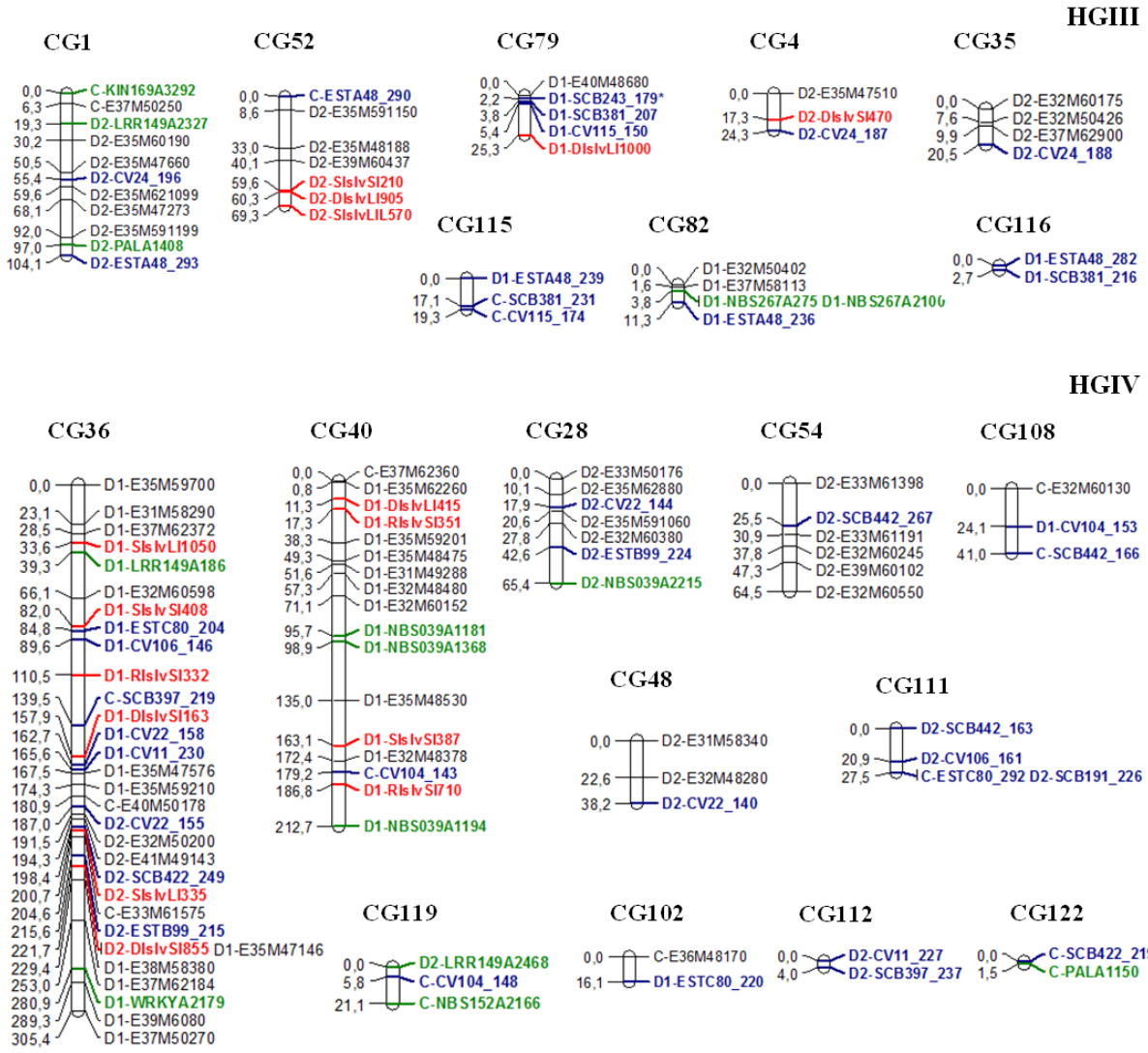


Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre ‘IAC66-6’ e ‘TUC71-7’. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no ‘IAC 66-6’), D2 (presença no ‘TUC 71-7’) ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos

(continuação)

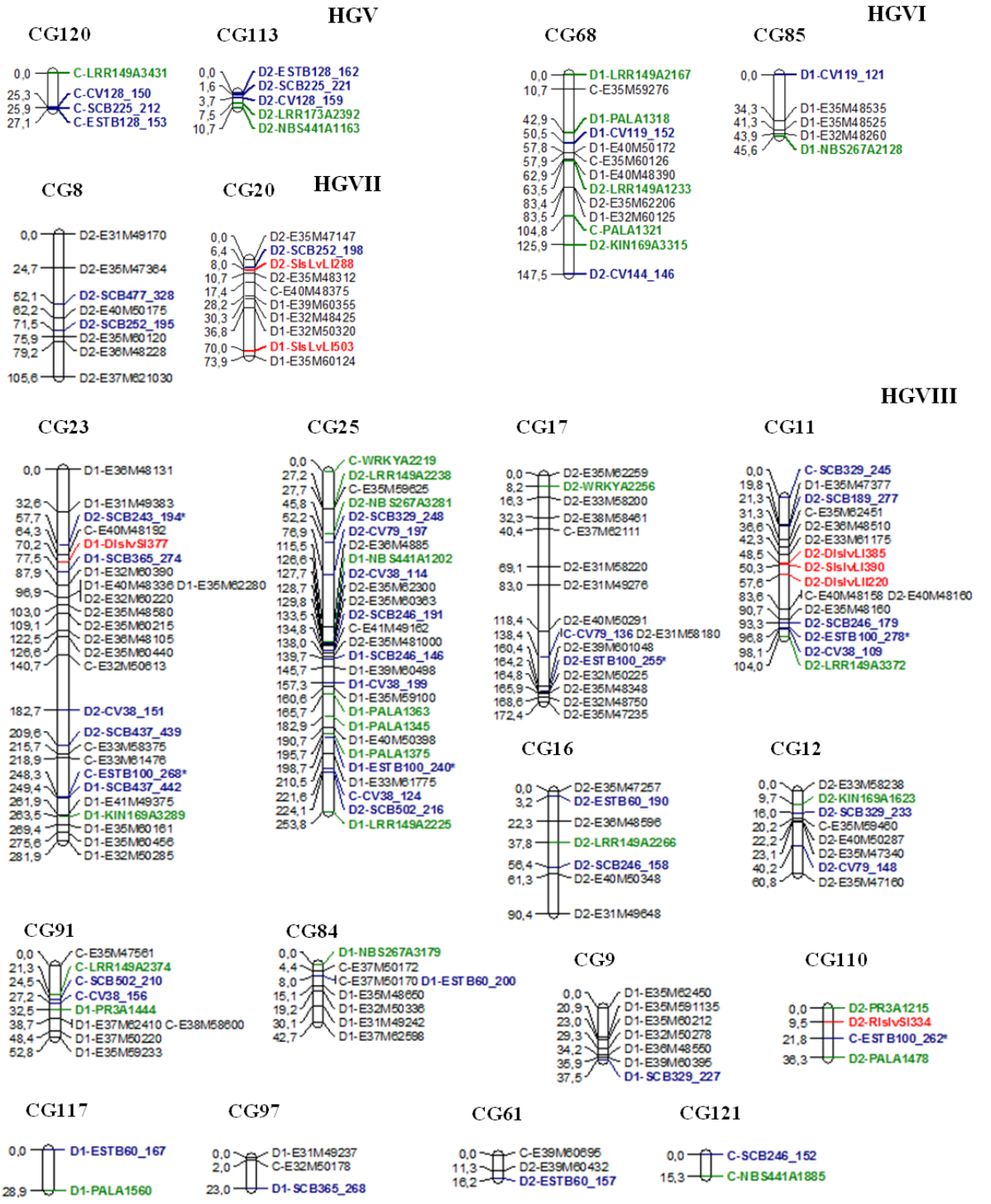


Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre 'IAC66-6' e 'TUC71-7'. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no 'IAC 66-6'), D2 (presença no 'TUC 71-7') ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos

(continuação)

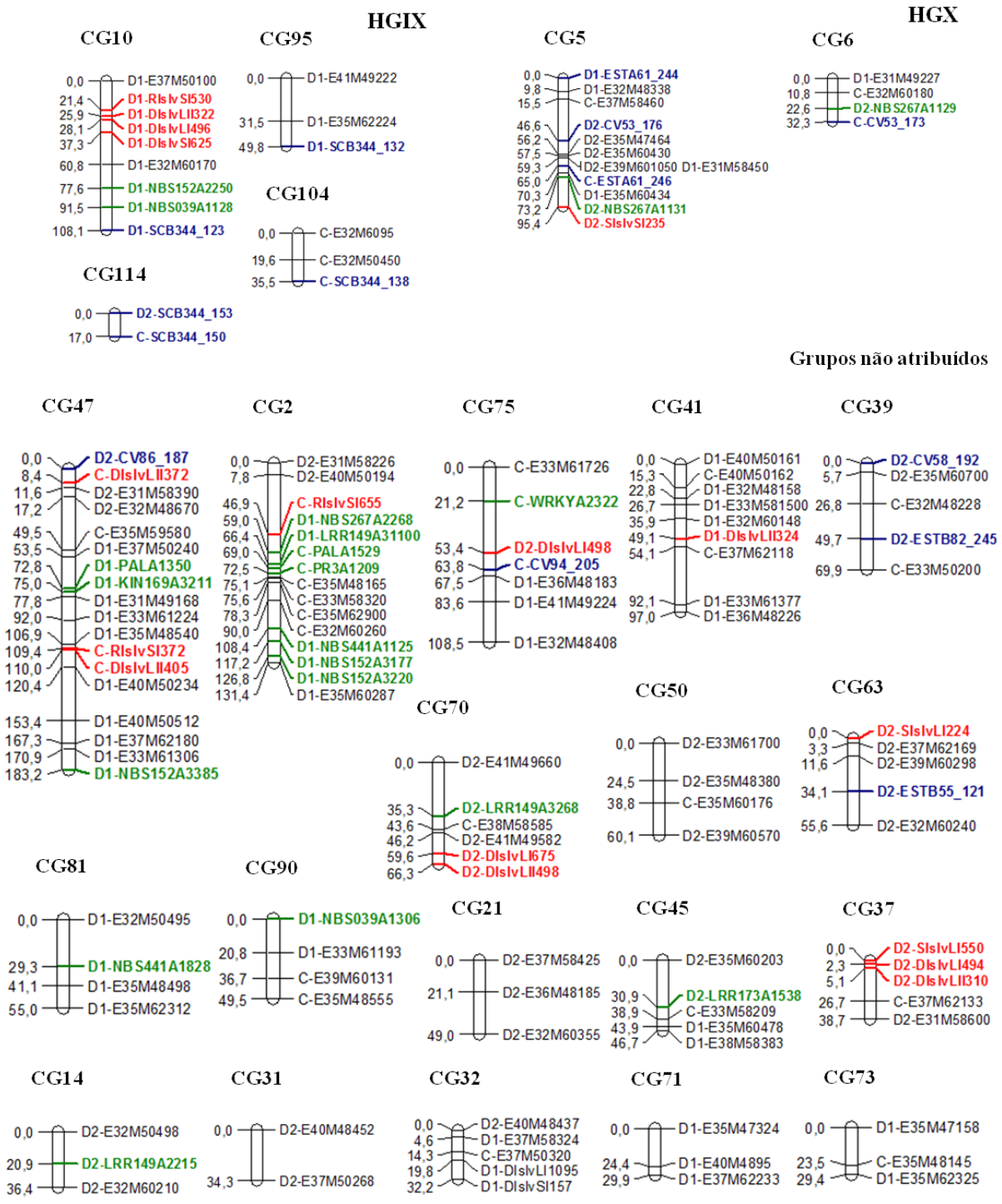


Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre 'IAC66-6' e 'TUC71-7'. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no 'IAC 66-6'), D2 (presença no 'TUC 71-7') ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos

(continuação)

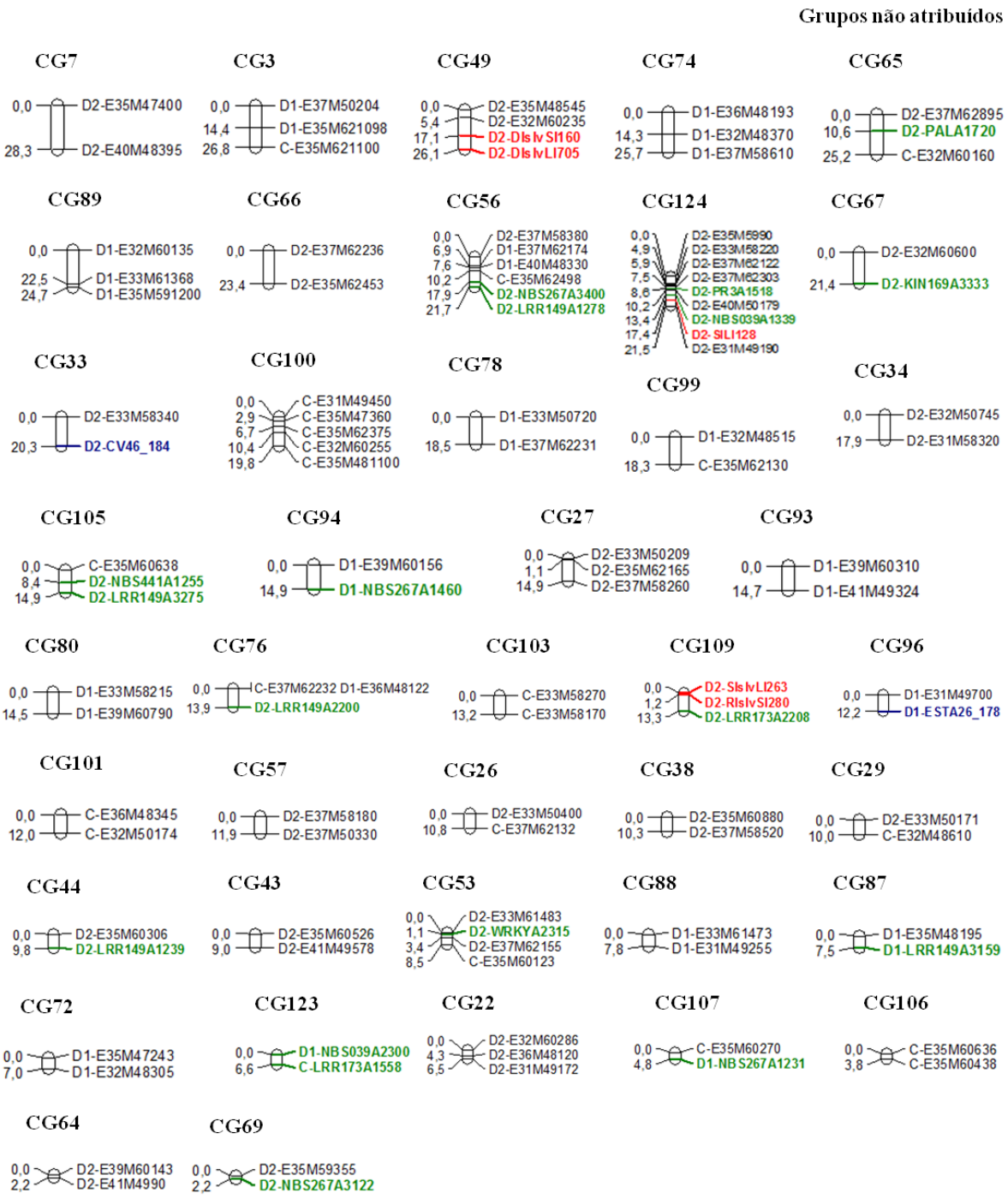


Figura 13 - Mapa de ligação integrado de cana-de-açúcar com base na progênie oriunda do cruzamento entre 'IAC66-6' e 'TUC71-7'. Os números à esquerda correspondem às distâncias do mapa em cM, utilizando função de mapeamento de Kosambi. Os nomes das marcas estão à direita, precedidos pela sigla D1 (presença da marca no 'IAC 66-6'), D2 (presença no 'TUC 71-7') ou C (presente em ambos os genitores). Em preto, estão os marcadores AFLP, em azul os EST-SSR, em verde os marcadores TRAP e em vermelho os retrotransposons *scIvana_1*. Os 65 CGs foram atribuídos em HGs (I ao X), por compartilharem pelo menos um alelo do mesmo loco de EST-SSR. Os demais CGs foram designados de grupos não atribuídos

(conclusão)

Dos 58 CGs onde foram posicionados os marcadores TRAPs, em 19 grupos (1, 2, 10, 18, 25, 36, 40, 47, 56, 62, 68, 82, 91, 105, 110, 113, 119, 123 e 124) foram mapeadas mais de uma marca e, em cinco deles (2, 47, 56, 113 e 124), as marcas estão a uma distância menor ou igual a 5 cM, sugerindo a formação de *clusters*. Essa estrutura aqui observada condiz com dados genéticos e moleculares citados na literatura, que mostram que os genes que determinam a resistência a doenças em plantas são frequentemente encontrados em *clusters* no genoma (MICHELMORE; MEYERS, 1998; ROSSI et al. 2003).

Devido a natureza multialélica e ao seu polimorfismo, os marcadores EST-SSR são utilizados na identificação de HGs em cana-de-açúcar (OLIVEIRA et al., 2007; PASTINA et al., 2012). Neste trabalho, seis HGs (HGI, HGII, HGIII, HGIV, HGV e HGVIII) foram estabelecidos usando CGs que contêm pelo menos dois locos de microssatélites em comum, sendo os demais CGs acrescentados ao HGs pré-estabelecidos quando compartilharam alelos de um loco em comum (AITKEN; JACKSON; McINTYRE, 2005; OLIVEIRA et al., 2007; PASTINA et al., 2012). Os outros quatro HGs putativos (HGVI, HGVII, HGIX e HGX) foram formados por CGs que contêm apenas um alelo do mesmo loco de microssatélite (GARCIA et al., 2006; OLIVEIRA et al., 2007; PIPERIDIS et al., 2008).

Dos 124 CGs, 65 puderam ser reagrupados dentro de dez HGs putativos, com base em 142 marcas geradas por 45 locos EST-SSR. Os locos remanescentes (CV46, CV58, CV86, CV94, CV104, ESTA26, ESTB55, ESTB82, SCB189, SCB191 e SCB477) não contribuíram para atribuir os CGs dentro de HGs. Mapas de ligação prévios de cultivares de cana-de-açúcar têm identificado diferentes números de HG. Rossi et al. (2003) e Palhares et al. (2012) identificaram sete HGs, Aitken, Jackson e McIntyre (2005) oito, Andru et al. (2011) nove, Pastina et al. (2012) 11 e Oliveira et al. (2007) 14 HGs.

O número de CGs aqui atribuídos dentro dos HGs variou de dois a 14, o maior (HGVIII) contém 138 marcas e o menor (HGV) nove. Observou-se que os CGs não foram distribuídos uniformemente ao longo dos diferentes HGs (Figura 13). A grande variação no número de CGs atribuído dentro dos HGs, aqui revelada, também foi observado por Pastina et al. (2012) que construíram um mapa de ligação integrado do cruzamento bi-parental entre ‘SP80-180’ e ‘SP80-4966’, sendo que o maior HG foi composto por 23 CGs e o menor por dois. Dados semelhantes foram relatados por Aitken, Jackson e McIntyre (2005), no mapa da cultivar ‘Q165$^{\varphi}$’, onde o maior HG foi construído por 37 CGs e o menor por três.

Em teoria, o número de grupos de ligação deve corresponder ao valor de *n* da espécie sob estudo, inclusive em plantas poliploides. Assim, no trigo cultivado (*Triticum aestivum*), por exemplo, um alohexaploide composto de três subgenomas, em que 2n = 42 e x = 7, o

número de grupos de ligação nos mapas é igual a 21. A literatura relata que há cromossomos homólogos e homo(eo)logos no genoma do trigo, entretanto, o processo evolutivo, a domesticação e o melhoramento genético propiciaram o isolamento desses subgenomas, de tal sorte que só se formam bivalentes na meiose I e a herança é dissômica. Isto se deve a rearranjos estruturais bem documentados na literatura. Se fossemos identificar, em teoria, os grupos de homo(eo)logia, este valor seria igual a *x*, já que há 7 pares cromossômicos pertencentes a cada um subgenomas, A, B e D.

Como exposto na fundamentação desta tese, no caso da cana, o principal genitor (*S. officinarum*) é um autopoliploide, com 2n = 80 e x = 10, ou seja, há 8 cópias de cada um dos 10 cromossomos. Além disso, há *2n* cromossomos deste genitor na cana cultivada, enquanto há *n* cromossomos do genitor silvestre, *S. spontaneum* (2n = 64, 80, 96, 112 e 128, dependendo do citotipo e x = 8). Embora estudos citogenéticos tenham sugeridos que a maioria dos pareamentos se dá na forma de bivalentes, várias irregularidades têm sido relatadas e a presença de univalentes é usual, sendo que a ocorrência de cromossomos recombinantes já foi descrita (D'HONT et al., 1998; D'HONT, 2005). Isto tem implicações no número de HGs que vem sendo encontrado, e é esperado que variações neste número sejam observadas nos diferentes trabalhos. A real contribuição da espécie silvestre ainda não está definida; somado a isso, há limitações relativas à metodologia usada neste trabalho para fins de construção do mapa, com escassez de marcas SSR, uso das segregações 1:1 e 3:1, desprezando outras, e de marcadores dominantes unicamente.

Das 153 marcas EST-SSR ligados ao mapa, 95 foram usadas para a construção do mapa prévio publicado por Palhares et al. (2012). Assim, 58 EST-SSRs e 95 TRAPs foram incorporadas ao mapa de ligação prévio, formando 21 novos CGs (44, 65, 67, 69, 76, 87, 94, 105, 107, 110, 112, 113, 114, 116, 117, 118, 119, 120, 121, 122 e 123). Devido à adição de mais marcadores do tipo EST-SSR ao mapa prévio, oito CGs, que anteriormente não foram atribuídos dentro de grupos de homo(eo)logia, agora passaram a fazer parte dos HGs, são eles: U-5 atribuído ao HGIV (renomeado 54), U-3 e U-10 atribuídos ao HGVII (renomeados 20 e 8, respectivamente), U-50 atribuídos ao HGVIII (renomeados 97), U-14 e U-25 atribuído ao HGIX (renomeados 95 e 104, respectivaemente) e os CGs U-2 e U-37 atribuídos ao HGX (renomeados 5 e 6).

Os grupos de homo(eo)logia do presente mapa foram organizados da seguinte forma comparativamente ao mapa prévio. O HGI foi estabelecido pela subdivisão do HGI do mapa prévio. O HGII foi constituído pela junção de parte do HGII com o HGVII do mapa prévio e um novo CG. HGIII foi estabelecido pelos mesmos CGs do mapa prévio e um novo GC. O

HGIV foi construído pela junção da outra parte do HGII com o HGIV e por três novos CGs. HGVI foi construído pelos mesmos CGs do mapa prévio. Já o HGVIII foi estabelecido pela junção da outra parte do HGI com o HGV e três novos CGs. Quanto aos demais HGs (V, VII, IX e X) foram formados por CGs novos ou que foram recuperados dos grupos não atribuídos do mapa prévio. Em alguns CGs, a adição de novos marcadores resultou no rearranjo dos marcadores dentro do CG original e até mesmo na sua divisão como demonstrado nas figuras 11 e 12.

A comparação entre o mapa construído no presente estudo e o publicado por Pastina et al. (2012) permite sugerir correspondência entre HGs de ambos os mapas, conforme evidencia a localização dos locos ESTA53, ESTB94, ESTB118 e ESTA48.

Nove locos de EST-SSR mostraram-se duplicados dentro de grupos de co-segregação (CV22, CV38, CV78, CV100, ESTA61, ESTB14, SCB246, SCB344 e SCB437) os quais formam posicionados nos HGI, IV, VIII, IX e X. A ocorrência de regiões genômicas duplicadas tem sido reportada em vários mapas de cana-de-açúcar (MING et al., 1998; AITKEN; JACKSON; McINTYRE, 2005; OLIVEIRA et al., 2007; PIPERIDIS et al., 2008; PASTINA et al., 2012). Além dos locos duplicados, alelos gerados por dois locos EST-SSR (ESTB100 e SCB243) foram mapeados em diferentes grupos de homo(eo)logia (HGI e HGVIII; HGIII e HGVIII, respectivamente) (Figura 13 e Tabela 7). Resultados semelhantes foram observados para o mapa da cultivar ‘R570’ (JANNOO et al., 2004) e no mapa da população derivada do cruzamento entre as cultivares ‘SP80-180’ x ‘SP80-4966’ (OLIVEIRA et al., 2007). Possivelmente a ocorrência de regiões duplicadas é consequência da origem multiespecífica das cultivares modernas e de possíveis rearranjos genômicos.

Tabela 7 – Relação dos locos EST-SSR e respectivos alelos mapeados, localização nos grupos de homo(eo)logia e de co-segregação e sua similaridade com proteínas

(continua)

Loco	Alelos mapeados e respectivos HGs e CGs entre parênteses	Homologia
CV06 [1]	188 (II-42), 226 (II-60)	NC
CV11 [1]	227 (IV-112), 230 (IV-36)	NC
*CV22 [2]	140 (IV-48), 144 (IV-28), 155 (IV-36), 158 (IV-36)	NC
CV23 [1]	152 (II-98), 186 (II-24), 194 (II-18),199 (II-86), 202 (II-42), 204 (II-51), 206 (II-118)	NC
CV24 [1]	187 (III-4), 188 (III-35), 196 (III-1)	NC
CV29 [2], [3]	101 (I-15), 85 (I-30), 89 (I-83)	NC
CV37 [2], [3]	125 (I-125), 129 (I-62), 151 (I-92)	NC
*CV38 [2], [3]	109 (VIII-11), 114 (VIII-25), 124 (VIII-25), 151 (VIII-23), 156 (VIII-91), 199 (VIII-25)	NC
CV46 [1]	184 (33)	NC
CV51 [1]	-	NC
CV53 [1]	173 (X-6), 176 (X-5)	NC
CV58 [1]	192 (39)	NC
*CV78 [2]	144 (I-30), 147 (I-30), 173 (I-19), 178 (I-83)	NC
CV79 [2]	136 (VIII-17), 148 (VIII-12), 197 (VIII-25)	NC
CV86 [1]	187 (47)	NC
CV91 [1]	175 (I-55), 190 (I-46)	NC
CV94 [2]	205 (75)	NC
*CV100 [1]	222 (I-55), 231 (I-46), 234 (I-55), 236 (I-77), 240 (I-55), 256 (I-77)	NC
CV101 [1]	-	NC
CV104 [2]	143 (IV-40), 148 (IV-119), 153 (IV-108)	NC
CV106 [2]	146 (IV-36), 161 (IV-111)	NC
CV115 [1]	150 (III-79), 174 (III-115)	NC
CV119 [1]	121 (VI-85), 152 (VI-68)	NC
CV128 [1]	150 (V-120), 159 (V-113)	NC
CV135 [1]	-	NC
CV144 [1]	146 (VI-68)	NC

Tabela 7 – Relação dos locos EST-SSR e respectivos alelos mapeados, localização nos grupos de homo(eo)logia e de co-segregação e sua similaridade com proteínas

(continuação)

Loco	Alelos mapeados e respectivos HGs e CGs entre parênteses	Homologia
ESTA26 (4)	178 (96)	TPA_inf: WRKY transcription factor 74
ESTA48 (4)	236 (III-82), 239 (III-115), 282 (III-116), 290 (III-52), 293 (III-1)	Chloroplast phytoene synthase 1
ESTA53 (4)	168 (I-62), 182 (I-83), 192 (I-77)	Indeterminate spikelet 1
*ESTA61 (4)	244 (X-5), 246 (X-5)	Lipid transfer protein-like protein
*ESTB14 (4)	291 (I-64), 309 (I-30), 315 (I-30)	RNA-binding protein RZ-1
ESTB55 (4)	121 (63)	Beta adaptin-like
ESTB60 (4)	157 (VIII-61), 167 (VIII-117), 190 (VIII-16), 200 (VIII-84)	OJ000114_01.16
ESTB82 (4)	245 (39)	Putative peroxidase
ESTB94 (4)	210 (I-59), 225 (I-58), 234 (I-77), 255 (I-15), 258 (I-13), 279 (I-30)	No hit
ESTB99 (4)	215 (IV-36), 224 (IV-28)	No hit
**ESTB100 (4)	240 (VIII-25), 255 (VIII-17), 262 (VIII-110), 268 (VIII-23), 270 (I-55), 278 (VIII-11)	22 kDa drought-inducible protein
ESTB118 (4)	106 (I-30), 127 (I-77), 163 (I-15)	Homeobox transcription factor GNARLY1
ESTB128 (4)	153 (V-120), 162 (V-113)	Putative auxin response factor 7a
ESTC80 (4)	204 (IV-36), 220 (IV-102), 292 (IV-111)	Glycosyltransferase
SCB189 (5)	277 (VIII-11)	Putative ubiquitin-specific protease UBP27 [AY133770]
SCB191 (5)	226 (IV-111)	Aspartyl protease family protein [NM_001161138]
SCB219 (5)	149 (II-118), 155 (II-24)	Arsenite transport subunit B [NM_001154652]
SCB225 (5)	212 (V-120), 221 (V-113)	Ethylene-insensitive3-like 1 protein [NM_001158563]
SCB227 (5)	203 (I-83), 209 (I-13), 212 (I-77)	Nucleotide binding protein [NM_001153645]
**SCB243 (5)	179 (III-79), 194 (VIII-23)	Group peroxidase 16 precursor protein [EF990901]
*SCB246 (5)	152 (VIII-121), 158 (VIII-16), 146 (VIII-25), 179 (VIII-11), 191 (VIII-25)	Hypothetical protein [XM_002450508]
SCB252 (5)	195 (VII-8), 198 (VII-20)	ATP binding protein [EU956876]
SCB272 (5)	-	Plant-specific domain TIGR01568 family protein [NM_001153443]
SCB312 (5)	216 (II-98), 249 (II-24), 258 (II-18), 264 (II-86)	Putative auxin efflux carrier (PIN1b) [DQ836240]
SCB329 (5)	227 (VIII-9), 233 (VIII-12), 245 (VIII-11), 248 (VIII-25)	Pi21 gene for Pi21 protein [AB430854]

Tabela 7 – Relação dos locos EST-SSR e respectivos alelos mapeados, localização nos grupos de homo(eo)logia e de co-segregação e sua similaridade com proteínas

(conclusão)

Loco	Alelos mapeados e respectivos HGs e CGs entre parênteses	Homologia
*SCB344 (5)	123 (IX-10), 132 (IX-95), 138 (IX-104), 150 (IX-114), 153 (IX-114)	Hypothetical protein [XM_002440411]
SCB365 (5)	268 (VIII-97), 274 (VIII-23)	No hit
SCB381 (5)	207 (III-79), 216 (III-116), 231 (III-115)	Peroxidase 16 precursor [NP_001152697]
SCB397 (5)	219 (IV-36), 237 (IV-112)	Zinc finger, C2H2 type family protein [NP_001151975]
SCB422 (5)	219 (IV-122), 249 (IV-36)	Hypothetical protein [XP_002443307]
*SCB437 (5)	439 (VIII-23), 442 (VIII-23)	Serine/threonine-protein kinase N1 [NP_058871]
SCB442 (5)	163 (IV-111), 267 (IV-54), 166 (IV-108)	No hit
SCB477 (5)	328 (VII-8)	Hypothetical protein [EEC71447]
SCB502 (5)	210 (VIII-91), 216 (VIII-25)	Associated with HOX family protein [ABA91569]

(1)Desenvolvido por CanaVialis/Monsanto (dados não publicados); (2) Desenvolvido por Jordão Júnior (2009); (3) Publicado por Maccheroni et al. (2009); (4) Desenvolvido por Oliveira et al. (2009); (5) Desenvolvido por Marconi et al. (2011)

* EST-SSR duplicado dentro de CG

** EST-SSR duplicado entre HGs

NC não caracterizado

Na tabela 8 constam alguns dos mapas de cana-de-açúcar disponíveis na literatura. Nota-se que o mapa aqui apresentado incrementa o mapa prévio em 1.379,8 cM, 32 CGs e em quatro novos HGs, no entanto, interessantemente a densidade de marcas (cM/marcas) permanece semelhante (8,6 cM mapa atual e 8,9 cM prévio). O mapa de ligação integrado recentemente publicado por Pastina et al. (2012) contém 317 marcas distribuídas em 96 CGs cobrindo 2.468 cM, com densidade de marcas de 7,5 cM. Embora o número de CGs no mapa publicado por Pastina et al. (2012) seja menor, ele é mais denso.

Andru et al. (2011) construíram um mapa de ligação utilizando uma população proveniente da autofecundação da cultivar 'LCP85-384'. Este mapa contém 784 marcas atribuídas a 108 CGs cobrindo 5.617,0 cM do genoma, com uma densidade de marcas de 7,2 cM. Já o mapa da cultivar '570' tem 887 marcas agrupadas em 120 CGs cobrindo 5.849,0 cM, com densidade de marcas de 6,6 cM (HOARAU et al., 2001). Verifica-se que, neste caso, a cobertura do genoma é semelhante, no entanto o mapa da cultivar 'R570' é mais denso e com maior número de CGs que o mapa da 'LCP85-384'. Provavelmente estas diferenças devem-se ao número de marcas mapeadas, ao polimorfismo intrínseco das populações e ao variável número de cromossomos encontrados nas cultivares ('R570', 2n = ~ 115 e 'LCP85-384', 2n = 106).

Tabela 8 - Mapa de ligação desenvolvido para a cana-de-açúcar, com informações sobre o delineamento genético da população de mapeamento, número de marcas ligadas, número de grupos de co-segregação (CGs), grupos de homologia (HGs) e cobertura total do genoma em centiMorgans (cM)

População de Mapeamento	Número de			Cobertura do mapa (cM)	Metodologia de mapeamento	Referência
	Marcas ligadas	CGs	HGs [1]			
'IJ76-514' x	-	-	-	-	*2-way pseudo testcross*	Aitken, Jackson e McIntyre (2005)
'Q165$^{\varphi}$'	534	136	8 (127)	9.058,3		
'La Striped' x	146	49	-	1.732,0	*2-way pseudo testcross*	Alwala et al. (2008)
'SES 147B'	121	45		1.491,0		
'Q117' x	376	93	10 (38)	3.167,0	*2-way pseudo testcross*	Piperidis et al. (2008)
'MQ77-340'	400	101	11 (52)	3.582,0		
'R570'	887	120	10 (34)	5.849,0	*pseudo-F2*	Hoarau et al. (2001)
'R570'	282	128	7 (66)	-	*pseudo-F2*	Rossi et al. (2003)
'LCP 85-384'	784	108	9 (44)	5.617,0	*pseudo-F2*	Andru et al. (2011)
'SP80-180' x 'SP80-4966'	357	131	12 (42)	2.602,4	mapa integrado	Garcia et al. (2006)
'SP80-180' x 'SP80-4966'	664	192	14 (120)	6.261,1	mapa integrado	Oliveira et al. (2007)
'SP80-180' x 'SP80-4966'	317	96	11 (91)	2.468,1	mapa integrado	Pastina et al. (2012)
'TUC71-7' x 'IAC66-6'	546	92	7 (41)	4.843,2	mapa integrado	Palhares et al. (2012)
'TUC71-7' x 'IAC66-6'	721	124	10 (65)	6.223,0	mapa integrado	Rodrigues-Morais (2012)

[1] Entre parênteses está apresentada a quantidade de CGs atribuída aos HGs

Em síntese, a alocação de mais marcas ao mapa de ligação aqui construído para a progênie do cruzamento ‘IAC66-6’ x ‘TUC71-7’, além de torná-lo mais informativo para o mapeamento de locos relacionados à resposta da cana-de-açúcar ao patógeno *Sporisorium scitamineum*, também resultou em uma reorganização dos marcadores dentro dos CGs, bem como a recuperação de CGs, antes não atribuídos a HGs. No entanto, ressalta-se a necessidade de adicionar mais marcas informativas ao mapa de ligação, como a exemplo dos marcadores co-dominantes SNPs tornando-se possível a detecção de ploidias, bem como a presença de marcas segregando em outras proporções (MOLLINARI, 2012).

4.4 Análise dos dados fenotípicos

Dos 188 indivíduos da progênie F_1 utilizados na construção do mapa de ligação, 176 e 160 indivíduos foram avaliados para incidência de carvão nos anos de 2009 e 2010, respectivamente, juntamente com os genitores. Os ensaios foram conduzidos em casa de vegetação em delineamento inteiramente casualizado (DIC), repetidos em dois ensaios consecutivos e com desbalanceamento. Aplicou-se a metodologia dos modelos mistos para analisar os dados de AUDPC, possibilitando a predição dos valores genotípicos de cada indivíduo, através do melhor preditor linear não viesado ou BLUP. Os dados foram analisados conjuntamente, considerando o comportamento dos genótipos dentro de anos, como efeito aleatório. Ainda, para a análise de incidência de carvão, a variável número de gemas brotadas por parcela foi utilizada como covariável.

O aparecimento de chicote no genitor ‘TUC71-7’ foi observado 140 DAI (Dias Após a Inoculação) no primeiro ensaio e 90 DAI no segundo, enquanto para o genitor ‘IAC66-6’ o chicote foi observado 74 e 76 DAI no primeiro e segundo ensaio, respectivamente. É possível que a diferença entre o aparecimento de chicote entre os genitores esteja relacionada à intensidade e ao tempo de resposta de genes envolvidos no mecanismo de defesa da planta, bem como à presença de alelos de resistência expressos apenas no genitor ‘TUC71-7’, considerado a princípio, o genitor resistente. Borrás-Hidalgo et al. (2005) trabalhando com as variedades ‘Ja S-44’ (resistente a *S. scitamineum*) e ‘Ja 60-5’(suscetível) e usando cDNA-AFLP, encontraram 11 fragmentos diferencialmente expressos (TDF) envolvidos na interação cana-de-açúcar vs *S. scitamineum*, após decorridos dois meses da inoculação. Análises de *Northern blot* indicaram que o nível de mRNA de quadro dos 11 TDFs foram altamente induzidos na variedade resistente, enquanto nenhuma ou pouca expressão foi observada na variedade suscetível.

Os valores de AUDPC dos genitores e da progênie segregante nos dois ensaios aqui realizados estão indicados nas figuras 14 e 15. Verifica-se que os genótipos avaliados mostraram reações ao carvão que variaram de forma contínua dentro de extremos de resistência e suscetibilidade. Para o primeiro ensaio, observa-se que a diferença entre os valores de AUDPC dos genitores 'TUC71-7' e 'IAC66-6' foi de 2,75 e, no segundo, de 0,61. Já na progênie, os valores de AUDPC variaram de 0,13 a 10,98 (amplitude de 10,85) e de 1,26 a 5,50 (amplitude de 4,24) no primeiro e no segundo ensaio, respectivamente. Em ambos os anos, a diferença dos valores de AUDPC da progênie foi maior que a diferença entre os genitores, revelando haver segregação transgressiva na progênie e sugerindo que genes de efeitos menores estão influenciando a resistência. Resultados semelhantes foram obtidos por Iamsupasit et al. (1996) e Skinner (1981) pois estes autores também verificaram a presença de genótipos transgressivos em populações de cana-de-açúcar, quando inoculadas com *S. scitamineum*.

Dentre os genótipos transgressivos os indivíduos 33, 76, 104, 111, 151, 176, 180 e 199 foram estáveis nos anos de condução dos ensaios, apresentando valores de AUDPC (BLUPs) inferiores ao do genitor 'TUC71-7'. Já os indivíduos 3, 6, 8, 10, 12, 15, 16, 17, 18, 22, 23, 30, 31, 34, 35, 36, 37, 40, 44, 46, 47, 48, 50, 52, 53, 59, 64, 65, 67, 69, 72, 74, 80, 82, 83, 84, 89, 98, 99, 101, 103, 105, 106, 109, 113, 114, 115, 116, 117, 118, 119, 120, 123, 124, 129, 132, 135, 137, 139, 144, 145, 155, 157, 158, 164, 167, 168, 169, 173, 182, 192, 194, 198 e 200 apresentaram valores de AUDPC (BLUPs) superiores ao do genitor 'IAC66-6', mostrando-se estáveis nos anos de condução dos ensaios (Anexo C). Apesar das diferenças observadas entre os anos da incidência da doença, os dados mostraram-se consistentes.

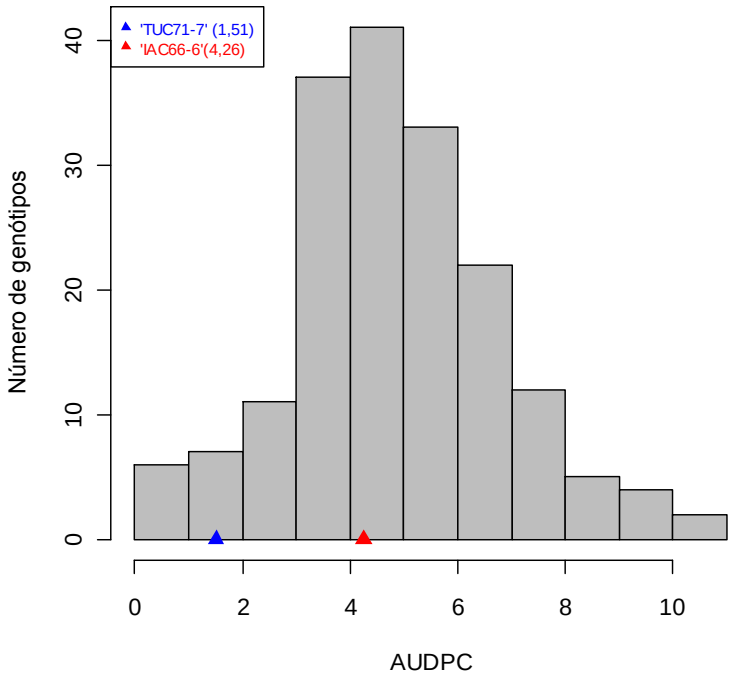


Figura 14 - Distribuição dos valores de AUDPC na progênie F_1 de cana-de-açúcar, resultante do cruzamento entre 'IAC66-6' e 'TUC71-7', no primeiro ensaio de avaliação de incidência de carvão

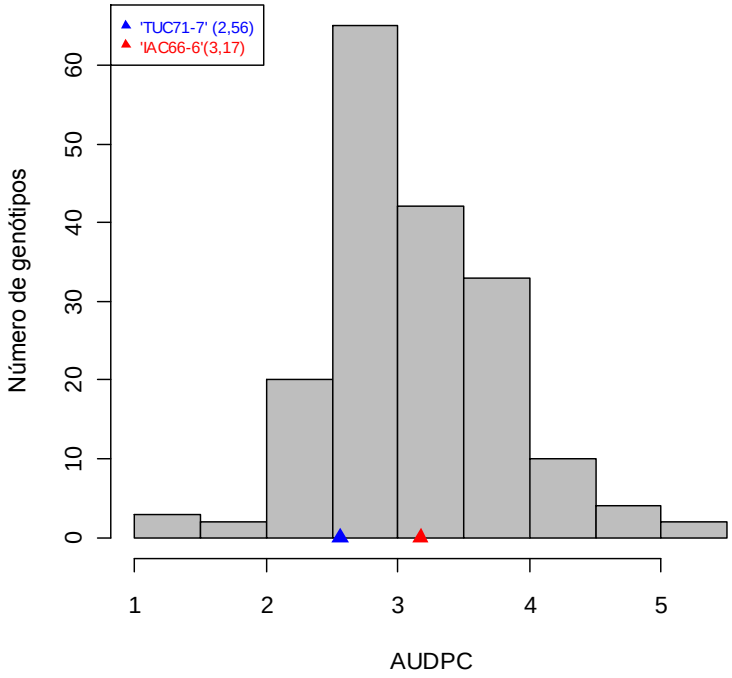


Figura 15 - Distribuição dos valores de AUDPC na progênie F_1 de cana-de-açúcar, resultante do cruzamento entre 'IAC66-6' e 'TUC71-7', no segundo ensaio de avaliação de incidência de carvão

Vale ressaltar que a amplitude de AUDPC observada no primeiro ano de avaliação sugere que a progênie apresentou maior variabilidade na resposta à doença em relação ao

segundo ano. Estudos realizados por Bock (1964), Hoy, Grisham e Benda (1987) e Hoy e Grisham (1988) demonstram que a reação de genótipos de cana-de-açúcar a *S. scitamineum* pode variar de um ano para outro, em função das condições ambientais e da massa de teliósporos utilizada. O mesmo tem sido observado para a resposta de plantas de milho ao ataque do fungo *Sporisorium reiliana*, agente causal do carvão-do-topo (LÜBBERSTEDT et al., 1999; LI et al., 2008; CHEN et al., 2008). Estudos envolvendo o uso da AUDPC para a determinação da resistência de populações de cana-de-açúcar a *S. scitamineum* são difíceis de serem comparados, uma vez que a literatura existente aborda apenas o uso de escalas de nota, não havendo consenso na comunidade científica (WHITTLE; WALKER, 1982; WU; HEINZ; MEYER, 1983; WU; HEINZ; HOGARTH, 1988; CHAO; HOY; SAXTON, 1990; NZIOKI; JAMOZA, 2006).

As estimativas dos componentes de variância genética ($\hat{\sigma}_G^2$) e variância ambiental ($\hat{\sigma}_E^2$) foram obtidas pelo método da máxima verossimilhança restrita (REML) e estão mostradas na tabela 9, juntamente com as estimativas da variância fenotípica ($\hat{\sigma}_F^2$) e do coeficiente de herdabilidade no sentido amplo ($\hat{h}_A^2$), para os dois ensaios. As estimativas da variância genética e ambiental foram significativamente diferentes entre si para os dois ensaios. A magnitude da $\hat{\sigma}_G^2$ obtida no primeiro ensaio foi ligeiramente superior a $\hat{\sigma}_E^2$, sugerindo que os efeitos genéticos tiveram uma maior influência na expressão do caráter. Já para o segundo ensaio, os efeitos ambientais tiveram uma maior influência na expressão do caráter, logo a $\hat{\sigma}_E^2$ apresentou uma maior magnitude em relação à $\hat{\sigma}_G^2$. (Tabela 9).

Tabela 9 - Estimativas dos componentes de variância genética ($\hat{\sigma}_G^2$), variância fenotípica ($\hat{\sigma}_F^2$), variância ambiental ($\hat{\sigma}_E^2$) e do coeficiente de herdabilidade no sentido amplo ($\hat{h}_A^2$) para incidência de carvão em cana-de-açúcar

Componente	**1° ensaio (2009)**	**2° ensaio (2010)**
$\hat{\sigma}_G^2$	5,95**	0,70**
$\hat{\sigma}_E^2$	5,15**	1,31**
$\hat{\sigma}_F^2$	11,10**	2,01**
$\hat{h}_A^2$	0,54	0,35
Média	4,85 (0,13; 10,98)	3,10 (1,26; 5,50)
r (1° ano/2° ano)	0,77[a]	
CV%	46,80	36,86

[a] estimativa da correlação entre o primeiro e o segundo ensaio; ** significância das estimativas a 1% de probabilidade dentro de cada ano

As estimativas da herdabilidade de plantas individuais variaram de um valor moderado (0,54), para o primeiro ano de avaliação, a baixo (0,35) para o segundo ano de avaliação,

indicando que os efeitos genéticos foram mais expressivos no primeiro ano. Estes resultados assemelham-se ao encontrado por Walker (1980), Chao, Hoy e Saxton (1990), Sills et al. (1995) e Raboin et al. (2003) que relataram valores de herdabilidade no sentido amplo variando de baixo a moderado para a resistência ao carvão.

A estimativa da correlação fenotípica (*r*) entre os dois anos de ensaio foi alta, igual a 77% (Tabela 9). A correlação entre anos pode ser influenciada pelas condições ambientais, como temperatura e concentração de teliósporos usadas, tal como relatado por Bock (1964), Grisham e Brend (1987), Hoy e Grisham (1988) e Bueno (2010).

4.5 Mapeamento de QTL

O uso da metodologia de mapeamento por intervalo composto permitiu detectar associações entre os valores da AUDPC e os marcadores posicionados no mapa. Como os dois ensaios foram conduzidos consecutivamente e contêm dados desbalanceados, optou-se pela aplicação da análise de modelos mistos às estimativas da AUDPC, a qual possibilitou obter BLUPs, que representam os valores da AUDPC corrigidos para os efeitos de anos e estande. Estes BLUPs foram, então, usados na detecção dos QTLs para cada ano e para o comportamento médio dos genótipos nos anos, através da abordagem de mapeamento por intervalo composto (CIM), adaptada por Gazaffi (2009) para espécies poliploides. Utilizando esta metodologia, foi possível mapear QTLs a partir do mapa de ligação integrado, construído com marcadores de diferentes padrões de segregação (1:1 e 3:1). Além da localização dos QTLs, também foi possível determinar o padrão de segregação e inferir as fases de ligação entre os QTLs e os marcadores.

No total, foram detectados oito locos quantitativos associados à resistência de *S. scitamineum* em cana-de-açúcar, com LOD *score* $\geq 3,7$. Destes, quatro foram identificados na primeira avaliação, realizada no ano de 2009, os quais estão localizados nos CGs 8, 30, 47 e 81, explicando em conjunto 18,8% da variação fenotípica; dois locos foram detectados na segunda avaliação, conduzida no ano de 2010, localizados nos CGs 25 e 29, explicando 8,6% da variação. Além disso, dois QTLs foram identificados para o comportamento médio dos genótipos nos anos, posicionados no CGs 47 e 55, e capazes de explicar 10,2% da variação (Tabela 10). A seguir uma abordagem detalhada será apresentada.

Tabela 10 - Resumo dos QTLs mapeados para a resistência da cana-de-açúcar ao carvão, segundo o modelo CIM e considerando LOD score ≥ 3,7

	QTL	Marcadores no intervalo	CG	Posição (cM)	LOD	Efeitos						Seg.	Fase	R^2 (%)
						Aditivo 'IAC66-6'	LOD	Aditivo 'TUC71-7'	LOD	Dominância	LOD			
Primeira avaliação (2009)	QTL_1	D2-SCB477_328 / D2-E40M50175	8	60,0	3,73	na	na	0,52	3,73	na	na	1:1	Q^1Q^2	3,7
	QTL_2	D2-E31M49312	30	26,8	4,18	-0,39	2,16	0,24	0,84	-0,29	1,15	3:1	$P^2P^1xQ^1Q^2$	5,2
	QTL_3	D1-E37M50240 / D1-PALA1350	47	72,0	3,82	-0,50	3,81	-0,06	0,01	-0,11	0,04	1:1	P^2P^1	5,7
	QTL_4	D1-E35M48498	81	41,1	4,02	0,52	4,02	na	na	na	na	1:1	P^1P^2	4,2
														18,8
Segunda avaliação (2010)	QTL_5	D2-E36M4885	25	115,5	3,84	-0,05	0,20	-0,09	0,87	0,18	2,41	1:2:1	Q^2Q^1	3,4
	QTL_6	D2-E33M50171	29	0,0	5,02	-0,25	1,81	0,18	3,81	na	na	3:1	$P^2P^1xQ^1Q^2$	5,3
														8,6
Média	QTL_7	D1-E37M50240 / D1-PALA1350	47	67,0	5,16	-0,40	4,26	-0,28	0,70	0,23	0,50	1:1	P^2P^1	5,3
	QTL_8	D1-E35M48218 / D1-LRR149A2547	55	210,0	4,37	-0,34	3,59	0,18	0,33	0,06	0,03	1:1	P^2P^1	4,9
														10,2

CG	Grupo de co-segregação
Posição	Posição do QTL no grupo de co-segregação, distância em relação ao primeiro marcador
——	Efeitos significativo a 5% de probabilidade (LOD *score* ≥ 0,83)
Seg.	Tipo de segregação dos alelos do QTL em consideração
Fase	Apresentados apenas os alelos dos QTL. Por exemplo, P^1P^2 significa configuração $P_m^1P^1P_{m+1}^1/P_m^2P^2P_{m+1}^2$
R^2	Coeficiente de determinação (% da variação fenotípica explicada)
na	Efeito não estimado devido à falta de informação

Resultados relativos à primeira avaliação

Os QTLs detectados para o ano de 2009 foram numerados de 1 a 4, como indicado na tabela 10. O QTL_1 foi localizado no HGVII e CG 8, na posição 60,0 cM e distante 2,2 cM do marcador mais próximo (D2-E40M50175), em uma região contendo dois locos EST-SSR, o D2-SCB477_328 e D2-SCB252_195, caracterizados por Marconi et al. (2011) como proteína hipotética e ligada a ATP, respectivamente (Tabela 7), sendo capaz de explicar 3,7% da variação de AUDPC. O QTL_2 foi mapeado no HGI e CG 30, na posição 26,8 cM, associado à marca D2-E31M49312, explicando 5,2% da variação fenotípica e em uma região que contém dois locos de EST-SSR que apresentaram homologia com o fator de transcrisão homeobox GNARTY1 (D1-ESTB118_106) e proteína ligada ao RNA RZ-1 (D2-ESTB14_315) e, distantes 2,8 cM e 23,1 cM dos locos de retrotranposon D2-SIsIvLI368 e D2-SIsIvLI240, respectivamente. O QTL_3 localizado no CG 47 está na posição 72,0 cM e distante 0,8 cM da marca mais próxima (D1-PALA1350) e a 8,0 cM da marca D1-KIN169A3211, as quais codificam fenilalamina amônia-liase e quitinase. Este QTL foi capaz de explicar 5,7% da variação fenotípica.

Já o QTL_4 foi detectado no CG 81, associado à marca D1-E35M48498, na posição 41,1 cM em uma região onde foi alocada a marca D1-NBS441A1828 derivada de gene análogo de resistência com domínio NBS/LRR, explicando 4,2% da variação. Como dito anteriormente, em conjunto, os QTLs foram capazes de explicar 18,8% da variação fenotípica. Dentre os quatro QTLs mapeados relacionados ao ano de 2009, o maior valor de LOD *score* (4,8) foi detectado para o QTL_2 (Tabela 10 e Figura 16).

Com relação aos efeitos de cada QTL, observa-se que ocorre predominância do efeito aditivo, sendo que para três deles (QTL_2, QTL_3 e QTL_4) foi detectada significância para os alelos do genitor 'IAC66-6' com LOD *score* $\geq$ 0,83. Desses, dois (QTL_2 e QTL_3) apresentaram efeitos negativos sobre o fenótipo, o que implica na diminuição da suscetibilidade à doença. O outro QTL_4 tem efeito positivo, de modo a aumentar a suscetibilidade à doença. No que se refere ao genitor 'TUC71-7', este contribuiu com dois locos quantitativos, cuja detecção foi significativa para o efeito de aditividade (QTL_1 e QTL_2), ambos contribuindo para aumentar a suscetibilidade da população. O efeito de dominância entre os alelos do loco mapeado no CG 30 (QTL_2) foi significativo, com efeito negativo sobre o fenótipo (Tabela 10).

Quanto aos tipos de segregação dos QTLs, nota-se predominância do padrão 1:1 (QTL_1, QTL_3 e QTL_4), uma vez que apenas um dos efeitos aditivos nos genitores foi significativo. Dois deles, QTL_3 e QTL_4, foram detectados no genitor 'IAC66-6' e encontram-

se em fase de repulsão e associação, respectivamente. O outro QTL_1, com este mesmo tipo de segregação, foi detectado no genitor 'TUC71-7', em fase de associação. O QTL_2, detectado no CG 30, apresentou segregação 3:1, uma vez que todos os efeitos genéticos (aditivos e dominância) foram significativos e similares, estando os alelos em repulsão no genitor 'IAC66-6' e, em associação, no genitor 'TUC71-7'.

Resultados relativos à segunda avaliação

No ano de 2010, houve uma redução no número de QTLs mapeados, embora tenham sido detectados dois novos QTLs (Tabela 10 e Figura 16), nomeados 5 e 6. Esse resultado era esperado, dado que o valor da estimativa do coeficiente de herdabilidade e da amplitude da AUDPC, para o ano de 2010, foram menores que os estimados para o ano anterior (Tabela 9; Figuras 14 e 15).

O QTL_5 foi posicionado no HGVIII e CG 25, na posição 115,5 cM e está associado à marca D2-E36M4885 em uma região onde foram alocados marcadores TRAP (NBS441A1202) e EST-SSR (CV38_114 e SCB246_191), sendo o marcador TRAP desenvolvido para o domínio NBS/LRR e o loco SCB246 caracterizado como uma proteína hipotética (Tabela 7). Este QTL explicou 3,4% da variação fenotípica. Já o QTL_6 foi mapeado no CG 29, na posição 0,0 cM, associado à marca D2-E33M50171, explicando a maior parte da variação fenotípica (5,3%) e apresentando o maior LOD *score* (5,0) (Tabela 10). Do mesmo modo que o observado para o ano de 2009, houve significância de efeitos aditivos e de dominância, havendo predominância dos efeitos aditivos oriundos dos alelos de ambos os genitores. O efeito aditivo do QTL_5 foi significativo para o genitor 'TUC71-7', com sinal negativo, de modo que sua contribuição está associada à diminuição da suscetibilidade da população à doença. No entanto, também foi detectado efeito significativo para interação intraloco (dominância) entre os efeitos aditivos dos alelos de cada genitor, aumentando a suscetibilidade da população. O efeito aditivo do QTL_6 foi significativo para o genitor 'IAC66-6' e para o 'TUC71-7', com sinal negativo e positivo, respectivamente.

Baseando-se na presença de dois efeitos estatisticamente significativos e na realização de testes sobre a igualdade desses efeitos, foi possível inferir que o QTL_5 segrega na proporção de 1:2:1, estando na fase de repulsão no genitor 'TUC71-7'. Já o QTL_6 segrega na proporção de 3:1, os alelos estão na fase de repulsão no genitor 'IAC66-6' e em associação no genitor 'TUC71-7'.

Resultados relativos ao comportamento médio de cada genótipo nos dois anos de avaliação

Considerando o comportamento médio de cada genótipo nos dois anos de avaliação foi possível detectar dois QTLs, com LOD *score* ≥ 3,7, denominados 7 e 8 (Tabela 10 e Figura 16). O QTL_7 foi mapeado no CG 47, na posição 67,0 cM e distante a 5,0 cM do QTL_3 detectado para o ano de 2009, sugerindo que o QTL_3 e o QTL_7 correspondem a mesma região cromossômica. Ao se observar o comportamento das curvas, bem como os testes para as estimativas dos efeitos, as fases de ligação e as respectivas segregações, nota-se que esses são o mesmo QTL e, portanto, sugere-se que é um loco cuja expressão é estável (Tabela 10 e Figuras 16). Por este QTL ser estável, marcadores ligados a ele podem aumentar a prossibilidade de sucesso na seleção de genótipos resistentes ao carvão nos programas de melhoramento.

O QTL_8 foi detectado no HGI e CG 55, na posição 210,0 cM e distante a 4,3 cM da marca mais próxima (D1-LRR149A2547) em uma região em que também foram detectados três retrotransposons (D1-SIsvILI425, C-SIsvIvLI412 e D1-SIsIvSI268). Embora o QTL_8 tenha sido detectado apenas na média dos genótipos, quando analisado o perfil da curva gerada para o ano de 2010, observa-se que esta apresenta um comportamento semelhante. Porém, o teste estatístico não foi capaz de detectar a sua presença neste ano de avaliação, uma vez que LOD *score* foi de 3,3, inferior ao limiar crítico utilizado (Figura 16).

Em conjunto, os QTLs controlam 10,2% da variação fenotípica. O QTL_7 foi o que explicou a maior parte da variação (5,3%), sendo que o alelo favorável está ligado em repulsão no cromossomo do genitor ‘IAC66-6’ e segrega na progênie na proporção de 1:1.

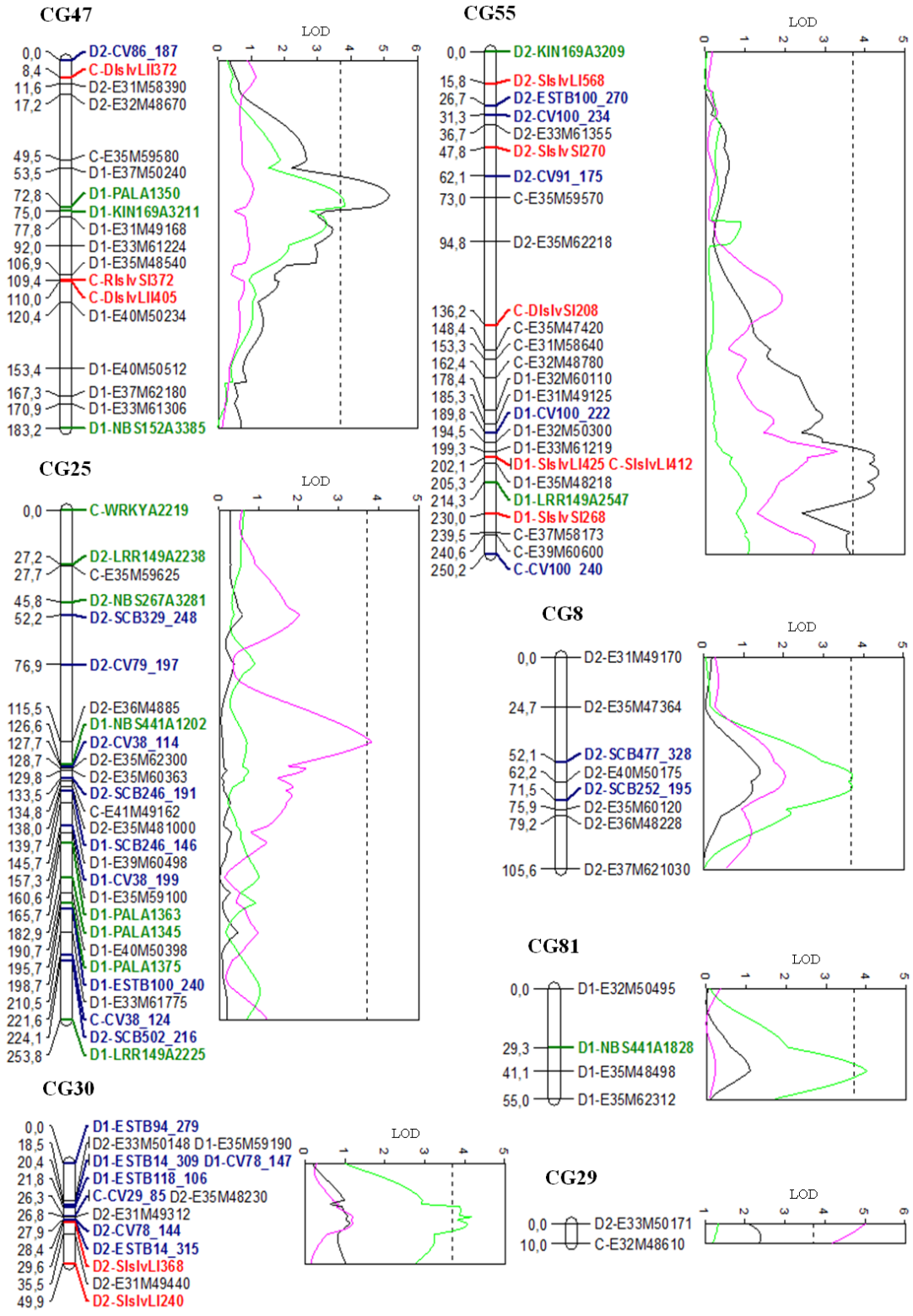


Figura 16 - Grupos de co-segregação que tiveram QTLs mapeados para a resistência ao carvão em cana-de-açúcar, com valores de LOD dispostos graficamente. À direita, estão os gráficos mostrando os QTLs. Em verde claro, a curva do primeiro ano de avaliação (2009), em magenta, do segundo ano de avaliação (2010) e, em preto, o comportamento médio dos genótipos nos anos. A linha tracejada refere-se ao limiar crítico LOD *score* ≥ 3,7

No geral, observou-se que os valores de LOD s*core* obtidos para a resistência ao carvão em cana-de-açúcar variaram entre 3,4 a 5,7. No entanto, alguns QTLs apresentaram o valor de LOD s*core* inferiores ao limiar crítico de 3,7 (Tabela 11) e não foram discutidos nesse trabalho por serem considerados QTLs não significativos. Contudo, estas regiões poderiam, futuramente, ser investigadas quanto à real presença desses QTLs.

Quanto aos efeitos de cada QTL mapeado, nota-se predominância dos efeitos aditivos comparativamente aos efeitos de dominância, indicando que os genitores utilizados neste trabalho foram capazes de fornecer alelos favoráveis para a progênie. Dos dez efeitos aditivos significativos, quatro (40,0%) são referentes a alelos do genitor 'TUC71-7', sendo que apenas um deles (QTL_5) contribui para aumentar a resistência da população. Considerando haver indícios de que o genitor 'TUC71-7' é resistente ao carvão, esperava-se que ele contribuísse com um maior número de efeitos aditivos significativos para aumentar a resistência da população à doença. No entanto, isso não foi observado, sugerindo que: *i*) os locos de resistência estariam localizados em regiões genômicas que não foram cobertas pelos marcadores; *ii*) o número de indivíduos utilizados para construir o mapa pode não ter sido suficiente para se detectar os alelos envolvidos na resistência e *iii*) um maior número de marcadores funcionais, derivados de sequências expressas (EST-SSR e SNPs), seria necessário para aumentar as chances de detecção dos supostos alelos de resistência presentes no genoma do genitor TUC71-7'.

Assim, dentre os QTLs que apresentaram efeito aditivo significativo e que contribuíram para a diminuição da suscetibilidade, cinco deles estão associados a marcas oriundas do genitor 'IAC66-6' e um deles está associado à marca oriunda do genitor 'TUC71-7'. Isto deve ser levado em consideração em estudos futuros, caso se deseje adicionar mais marcadores ao mapa de ligação ou, até mesmo, utilizá-los em seleção assistida por marcadores. Observa-se efeito de dominância completa para dois QTLs (QTL_2 e QTL_5), cujo efeito de dominância foi significativo e que tiveram a mesma magnitude dos efeitos aditivos significativos (Tabela 10).

Em relação à variação fenotípica explicada pelos QTLs, baseada na estimativa R^2 (coeficiente de determinação), observa-se QTLs de pequenos efeitos, explicando individualmente de 3,4% a 5,7% da variação fenotípica. Lübberstedt et al. (1999), utilizando mapeamento por intervalo composto (CIM), também identificaram QTLs de pequenos efeitos para resistência a *Sporisorium reiliana*, em milho. No experimento realizado na China, sob inoculação artificial com o patógeno, os autores relatam a detecção de 8 QTLs com estimativas de R^2 variando de 5,6% a 13,9%. Já para o experimento conduzido na França em

condições naturais de infestação, identificaram-se 3 QTLs explicando individualmente de 6,0% a 7,3% da variação fenotípica. Li et al. (2008) estudando a resistência de milho a *S. reiliana*, em duas localidades e em dois anos consecutivos, sob condições artificiais de inoculação, detectaram 17 QTLs usando a metodologia CIM. Para os experimentos que foram realizados na Província de Jilin, nos anos de 2003 e 2004, os autores detectaram quatro QTLs para o primeiro ano e cinco para o segundo, explicando de 6,6% a 43,7% e 8,1% a 18,9% da variação fenotípica, respectivamente, sendo três deles de expressão estável nos anos de avaliação. Já nos experimentos conduzidos na Província de Heilongjiang, nos anos de 2003 e 2004, foram identificados quatro QTLs para cada ano, explicando de 7,2% a 25,3% e 5,0 a 28,2% da variação, respectivamente, sendo dois deles estáveis. Diferente do patossistema milho – *S. reiliana*, nenhum gene de maior efeito foi detectado em cana-de-açúcar contra *S. scitamineum*. No entanto, a presença de genes de resistência de maiores efeitos não pode ser descartada.

Um fator que poderia contribuir para a detecção de QTLs de maiores efeitos seria a adoção de metodologias mais afinadas para o estudo da interação genótipo *x* ambiente. Para isto, sugere-se avaliar a população sob condições naturais de infecção, a campo, sem interferência humana, usando delineamentos experimentais adequados, contendo várias repetições e em diferentes locais de incidência da doença, levando em conta a variável porcentagem de touceiras doentes, tal como sugerido por Casagrande, Amorim e Sanguino (1998). Segundo esses autores, esta variável tem propiciado melhor distinção dos genótipos quanto à reação ao carvão, mostra associação com os danos provocado pelo patógeno e exige menos mão-de-abra para realizar as avaliações. Contudo, têm-se optado por ensaios utilizando inoculação artificial, com ou sem ferimento nas gemas, e delineamentos simples como o inteiramente casualizado, uma vez que o método supracitado requer áreas de cultivo bastante extensas, vários anos de avaliação e tem custo elevado.

Os QTLs detectados neste trabalho são referentes ao mecanismo de resistência pós-formada da planta, uma vez que a metodologia de inoculação envolveu o ferimento das gemas. No entanto, sabe-se que a morfologia da gema é um importante componente da resistência da cana ao *S. scitamineum*, como demonstrado por Amorim et al. (2000). Estudos que considerem outro método de inoculação − que não por ferimento − devem contribuir para a identificação de QTLs relacionados a resistência pré-formada da planta, aumentando assim o número de QTLs envolvidos na resposta ao fungo.

Até o momento, apenas a abordagem de análise de marcas individuais havia sido usada para associar marcas à resistência ao carvão em cana-de-açúcar (SILLS et al., 1995;

WEI et al., 2006; BUTTERFIELD, 2007), embora esta metodologia seja de fácil implementação e não necessite de mapas de ligação. Por outro lado, os testes estatísticos aplicados são viesados, resultando em baixo poder de detecção, em decorrência do número de marcadores e ao tamanho da amostra, além de que não é possível inferir sobre a localização do QTL no genoma.

Daí a necessidade do uso de metodologias mais apuradas para a detecção de QTLs, principalmente em poliploides, cujo genoma é mais complexo. As metodologias usadas para detectar QTLs em cana-de-açúcar, até o momento, baseavam-se em ferramentas desenvolvidas para espécies diploides. Este é o primeiro trabalho desenvolvido para o mapeamento de QTLs associados à resistência da cana-de-açúcar a *S. scitamineum*, utilizando a metodologia de mapeamento por intervalo composto, fornecendo subsídios para estudos que visem a melhor caracterização das regiões genômicas que apresentam indícios de QTLs relacionados à resistência ao carvão.

Ainda vale salientar que os QTLs, aqui identificados, foram posicionados em regiões do genoma cobertos por marcadores informativos, com exceção do QTL_6 detectado em um intervalo formado por dois AFLPs. Importante comentar que o QTL_8, que apresenta alelo com efeito aditivo negativo oriundo do genitor 'IAC66-6' foi detectado em uma região contendo um marcador TRAP direcionado ao domínio LRR e três retrotransponson *scIvana-1*. Isto corrobora os estudos que têm demonstrado que os retroelementos, importantes componentes da maioria dos genomas de plantas, estão adjacentes a genes de resistência a doenças, como em cevada (WEI; WING; WISE, 2002), citros (YANG et al. 2003), feijoeiro (MELOTTO et al., 2004) e beterraba (KUYKENDALL; SHAO; TRIMMER, 2009).

O uso de ESTs e de marcadores derivados de genes envolvidos na resposta à defesa da planta é uma estratégia que permite o mapeamento direto de sequências de interesse (ROSSI et al., 2003; OLIVEIRA et al., 2007; PINTO et al., 2010). Neste trabalho, o marcador PALA1350 mostrou-se associado a um QTL putativo com efeito negativo sobre o fenótipo, no ano de 2009 (QTL_3) e no comportamento médio dos genótipos nos diferentes anos (QTL_7). Isto é particularmente interessante pois confirma que a fenilalanina amônia liase, ou PAL, está envolvida na resposta a *S. scitamineum*; segundo o estudo de Borrás-Higalgo et al. (2005), este gene é diferencialmente expresso após a inoculação com *S. scitamineum*. Ainda, De Armas et al. (2007) relataram que a cultivar 'Mayari 55-14', considerada resistente a *S. scitamineum*, apresentou alta expressão do gene na presença do patógeno, enquanto a cultivar 'Barbados 42231', suscetível, mostrou baixa expressão. A PAL é a enzima chave na síntese de fenilpropanoides, uma via para a biossíntese de uma gama de produtos que desempenham

um papel chave no desenvolvimento da planta e contra estresses, incluindo a produção de polímeros de lignina, flavonóides, isoflavonóides e fitoalexinas (SANTIAGO et al., 2009). Sugere-se que o marcador PALA1350 tem grande potencial para uso em um programa de seleção assistida.

Quando relaciona-se, graficamente, os valores de AUDPC (BLUPs) e os genótipos observados para o loco PALA1350, marca associada ao QTL_3 e QTL_7, observa-se que os genótipos portadores do alelo nulo (*oo*) mostram as menores AUDPCs (Figura 17), sugerindo que a seleção deva ser conduzida contra a presença da marca (*ao*).

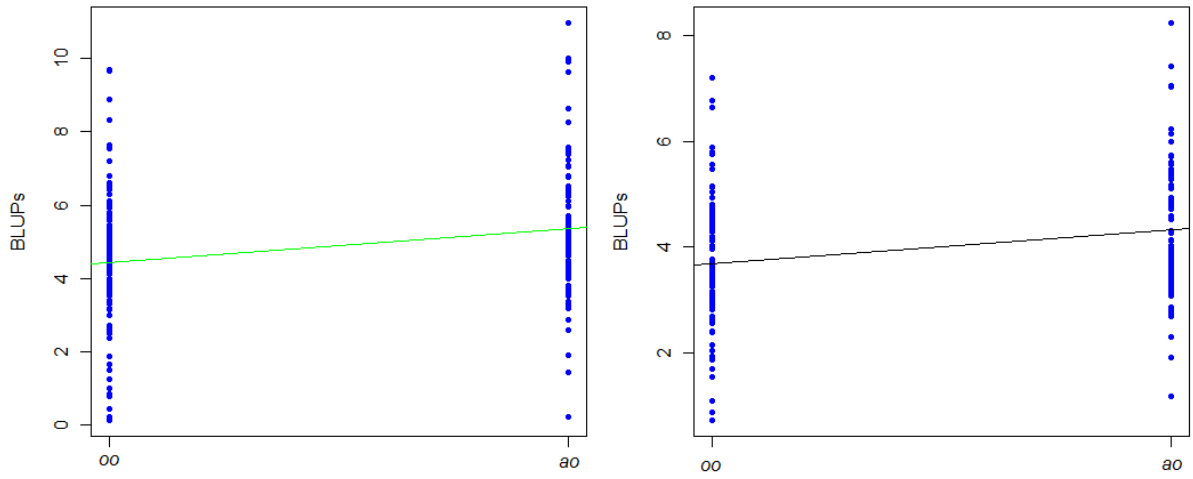


Figura 17 - Análise individual do marcador PALA1350 para o ano de 2009 (verde, à esquerda) e considerando o comportamento médio dos genótipos (preto, à direita). No eixo vertical os valores de BLUPs e, no horizontal, os genótipos da população de mapeamento

Tabela 11 – Resumo dos QTLs putativos mapeados para a resistência da cana-de-açúcar ao carvão, porém não significativos, considerando LOD score ≥ 3,7

	Marcadores no intervalo	CG	Posição (cM)	LOD	Efeitos						Seg.	Fase	R^2 (%)
					Aditivo 'IAC66-6'	LOD	Aditivo 'TUC71-7'	LOD	Dominância	LOD			
Primeira avaliação	D1-ESTB118_127 / D1-E35M48290	77	20,0	3,17	0,48	3,17	na	na	na	na	1:1	P^1P^2	6,8
Segunda avaliação	D1-E31M49472	42	193,5	3,43	-0,13	2,12	0,08	0,39	0,18	1,24	1:2:1	P^2P^1	4,9
	D1-E33M61219	55	199,3	3,33	-0,09	1,00	-0,01	0,00	0,18	1,04	1:2:1	P^2P^1	6,3
Média	D2-E32M60286	22	0,0	3,21	na	na	0,31	3,21	na	na	1:1	Q^1Q^2	3,2

CG — Grupo de co-segregação
Posição — Posição do QTL no grupo de co-segregação, distância em relação ao primeiro marcador
—— — Efeitos significativo a 5% de probabilidade (LOD *score* $\geq$ 0,83)
Seg. — Tipo de segregação dos alelos do QTL em consideração
Fase — Apresentados apenas os alelos dos QTL. Por exemplo, P^1P^2 significa configuração $P_m^1P^1P_{m+1}^1/P_m^2P^2P_{m+1}^2$
R^2 — Coeficiente de determinação (% da variação fenotípica explicada)
na — Efeito não estimado devido à falta de informação

4.6 Validação das sequências dos marcadores TRAP

As bandas obtidas pela técnica TRAP cujos locos correspondentes segregaram na população de mapeamento foram excisadas dos géis de poliacrilamida e armazenadas. Uma amostra de 50 delas foi reamplificada e o seu tamanho foi confirmado por eletroforese em gel de agarose (1%), como exemplificado na figura 18.

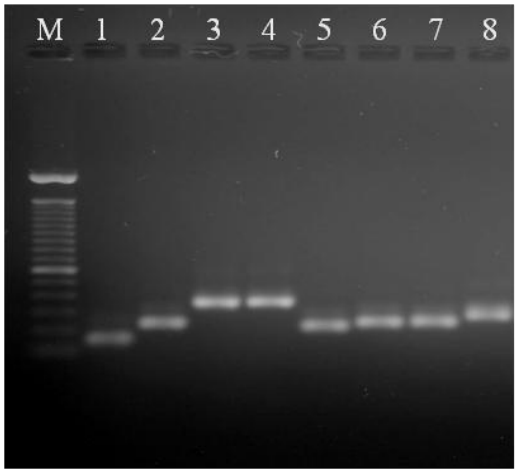


Figura 18 – Gel de agarose (1%) mostrando o padrão de amplificação a partir das bandas extraídas do gel de poliacrilamida. Nas canaletas 1 a 8 estão os produtos da reamplificação resultantes das combinações dos *primers* fixos/arbitrários e seus respectivos tamanhos em pares de base (pb): NBS441A1125, KIN169A1228, PALA1350, PALA1363, PR3A1209, LRR149A1233, LRR149A1239 e LRR149A1278; M: marcador de massa molecular de 100 pb

Os 50 fragmentos foram então clonados, resultando em 41 clones positivos, cujos insertos foram sequenciados. Todas as 41 sequências foram analisadas por BLASTx, destas 30 apresentaram similaridade quando alinhadas com sequências depositadas no banco de dados do NCBI (Tabela 12). A sequência do fragmento LRR173A1558 foi similar a uma proteína de resistência de *Medicago truncatula*, as sequências dos fragmentos LRR149A1278 e NBS152A2250 apresentaram similaridade com uma proteína da família LRR de *Oryza sativa* e *Arabidopsis lyrata*, respectivamente. As sequências dos fragmentos LRR149A2374, LRR149A3372, NBS039A1194, NBS267A3281 e NBS267A3400 foram similares a retrotransposons de *Oryza sativa*. Resultados semelhantes foram encontrados por Van der Linden et al. (2004) que, ao analisarem sequências de RGAs, utilizando a metodologia *NBS profiling*, encontraram similaridade com retrotransposons. Tal como relatado por Thokoane e Rutherford (2001), o retrotransposon Ty3-gypsy mostrou-se diferencialmente expresso em uma variedade resistente ao carvão; no presente estudo, a sequência NBS267A3400

apresentou similaridade com esse retrotransposon, sendo este outro marcador putativo para uso em seleção assistida.

A sequência do fragmento KIN169A3333 mostrou similaridade com a subunidade maior da n-calpaina-1 de *Arabidopsis thaliana*. Segundo Croall e Ersfeld (2007), a calpaina participa dos processos de diferenciação celular, apoptose e transdução de sinais. Já a sequência do fragmento LRR149A3264 foi similar à proteína ADH1 (álcool desidrogenase) de *Saccharum* cultivar R570. Presentes em gimnospermas, angiospermas, leveduras e em algumas bactérias, a enzima ADH faz parte da via de fermentação alcoólica, convertendo piruvato via acetaldeído em etanol. Estudos têm demonstrado que as enzimas ADH1 são essenciais no metabolismo anaeróbico das plantas, sendo que a transcrição do gene *Adh1* aumenta em condições de falta de oxigênio e em resposta a situações de estresse de temperatura e desidratação (CLEGG; CUMMING; DURBIN, 1997), além de ser induzido pelo fitohormômio ABA (DOLFERUS; PEACOCK; DENNIS, 1994; De BRUXELLES et al., 1996). Este gene (*Adh1*) foi localizado no cromossomo 11 do arroz, no cromossomo 1 do milho e no grupo de ligação C do sorgo (PATERSON et al., 1995; TARCHINI et al., 2000). Recentemente 13 cópias foram localizadas no grupo de homo(eo)logia I (HGI) da cultivar de cana-de-açúcar 'R570' (JANNOO et al., 2007; GARSMEUR et al., 2011; D'HONT comunicado pessoal[2]). Interessantemente, o mapa aqui construído contém uma cópia desse gene mapeado no CG 64 atribuído ao HGI. Ainda vale ressaltar que o sequenciamento nos flancos dos genes *Adh1* e *Adh2* do arroz tem revelado a presença de 33 genes putativos, sendo vários deles aparentemente envolvidos em resistência a doenças (TARCHINI et al., 2000).

A sequência do fragmento LRR149A3268 foi similar à proteína da família F-box de *Oryza minuta*; estes genes influenciam numa variedade de processos biológicos, tais como senescência foliar e ramificação (WOO et al., 2001; STIRNBERG; FURNER; OTTOLINE LEYSER, 2007), auto-incompatibilidade (QIAO et al., 2004; TAKAYAMA; ISOGAI, 2005), resposta aos reguladores de crescimento de plantas (DILL et al., 2004; LAI et al., 2004; BADESCU; NAPIER, 2006; BINDER et al., 2007) e respostas a estresses bióticos (KIM; DELANEY, 2002) e abióticos (CALDERON-VILLALOBOS et al., 2007). Já a sequência do fragmento NBS039A2300 apresentou similaridade com a proteína da família dedo de zinco B-box de *Zea mays*, a qual representa um subgrupo das proteínas dedo de Zn que contêm um ou mais domínios B-box, com especialização em estruturas terciárias, estabilizadas pela

[2] Palestra "Evolutionary dynamics of homo(eo)logous chromosome segments within the highly polyploid sugarcane genome", ministrada pela Drª. Angélique D'Hont, no Departamento de Genética da Escola Superior de Agricultura "Luiz de Queiroz", no dia 17 de setembro de 2012.

ligação de íons Zn. É sabido que os domínios B-box estão envolvidos nas regulações transcricionais e também em resposta ao frio (KHANNA et al., 2009). A sequência do fragmento NBS267A2268 foi similar a acetolactato sintase (ALS) de *Ipomoea ochracea*, que é a primeira enzima na via metabólica de biossíntese dos aminoácidos alifáticos de cadeia ramificada, como leucina, valina e isoleucina (EBERLEIN et al., 1997). Já a sequência NBS152A3220 apresentou similaridade com a proteína histona deacetilase, codificada pelo gene *Hcp901* de *Populus trichocarpa*. Esta enzima está envolvida na remoção de grupos acetila das histonas, aumentando afinidade com a molécula de DNA, desenvolvendo um importante papel na repressão da transcrição (WU et al., 2000; TUSKAN et al., 2006). Por outro lado, a sequência do fragmento NBS441T1255 foi similar à proteína de *Medicago truncatula*, denominada E3 ubiquitina proteína ligase CHFR, a qual apresenta o domínio dedo RING. De acordo com Marchler-Bauer et al. (2011), este domínio pode estar envolvido na mediação das interações entre proteínas como transdução de sinais, replicação viral e desenvolvimento vegetal.

As demais sequências (15) apresentaram similaridade com proteínas hipotéticas de gramíneas, *Oryza sativa* e *Sorghum bicolor*. Detalhes da análise dessas sequências podem ser visualizados na tabela 12. Em síntese, 50% dos fragmentos sequenciados são similares a proteínas envolvidas em processo de defesa da planta, contra estresses bióticos e abióticos e na ativação e desativação de processos bioquímicos vegetais.

Tabela 12 – Análise das sequências dos marcadores TRAP, usando a ferramenta BLASTx e tomando por base o banco de dados do NCBI

Sequência	Acesso	Descrição	Organismo	E-value	Identidade (%).	CG
LRR149A1278	AAP54211.2	Proteína da família LRR	*Oryza sativa* Grupo Japônica	9,00E-05	38	56 (U)
NBS152A2250	XP_002879429.1	Proteína da família LRR	*Arabidopsis lyrata* subsp. *lyrata*	6,00E-05	33	10 (IX)
LRR173A1558	XP_003626855.1	Proteína de resistência CC-NBS-LRR	*Medicago truncatula*	1,00E-05	30	123 (U)
LRR149A2374	ABF94719.1	Retrotransposon putativo, não classificado	*Oryza sativa* Grupo Japônica	3,00E-26	64	91 (VIII)
LRR149A3372	ABA97489.1	Retrotransposon putativo, não classificado	*Oryza sativa* Grupo Japônica	8,00E-36	79	11 (VIII)
NBS039A1194	ABA97058.1	Retrotransposon putativo, não classificado	*Oryza sativa* Grupo Japônica	1,00E-08	50	40 (IV)
NBS267A3281	ABA95102.1	Retrotransposon putativo, não classificado	*Oryza sativa* Grupo Japônica	5,00E-37	69	25 (VIII)
NBS267A3400	ABA95128.1	Retrotransposon putativo, subclasse Ty3-gypsy	*Oryza sativa* Grupo Japônica	3,00E-13	64	56 (U)
KIN169A3333	AAG51565.1	Subunidade maior de n-calpaina-1 putativa	*Arabidopsis thaliana*	6,00E-05	32	67 (U)
LRR149A3264	CBB36483.1	ADH1 (álcool desidrogenase)	*Hibrido Saccharum* cultivar R570	2,00E-16	100	62 (I)
LRR149A3268	ACS49611.1	Família-5 F-box	*Oryza minuta*	7,00E-05	38	70 (U)
NBS039A2300	NP_001147455.1	Proteína da família dedo de zinco B-box	*Zea mays*	2,00E-05	41	123 (U)
NBS267A2268	ADO16009.1	Acetolactato sintase	*Ipomoea ochracea*	9,00E-06	24	2 (U)
NBS152A3220	XP_002304453.1	Proteína do complexo histona deacetilase	*Populus trichocarpa*	2,00E-06	32	2 (U)
NBS441A1255	XP_003630425.1	E3 ubiquitina proteína ligase CHFR	*Medicago truncatula*	1,00E-05	31	105 (U)
KIN169A2232	EEC79285.1	Proteína hipotética OsI_20087	*Oryza sativa* Grupo Índica	1,00E-26	53	15 (I)
KIN169A3207	AAR06304.1	Proteína hipotética	*Oryza sativa* Grupo Japônica	7,00E-06	45	46 (I)
LRR149A2231	EEC84293.1	Proteína hipotética OsI_30770	*Oryza sativa* Grupo Índica	6,00E-07	44	-
LRR149A3159	XP_002447289.1	Proteína hipotética SORBIDRAFT_06g032145	*Sorghum bicolor*	9,00E-06	27	87 (U)
NBS152A3177	BAC57715.1	Proteína hipotética	*Oryza sativa* Grupo Japônica	2,00E-05	33	2 (U)
NBS267A3179	XP_002445353.1	Proteína hipotética SORBIDRAFT_07g011900	*Sorghum bicolor*	6,00E-16	89	84 (VIII)
NBS441A1125	EEC82442.1	Proteína hipotética OsI_26865	*Oryza sativa* Grupo Índica	2,00E-04	39	2 (U)
NBS441A1163	EEC72096.1	Proteína hipotética OsI_05056	*Oryza sativa* Grupo Índica	4,00E-05	55	113 (V)
PALA1206	XP_002458210.1	Proteína hipotética SORBIDRAFT_03g029010	*Sorghum bicolor*	5,00E-12	68	118 (II)
PALA1318	XP_002446021.1	Proteína hipotética SORBIDRAFT_06g000577	*Sorghum bicolor*	1,00E-16	50	68 (VI)
PALA1478	XP_002464245.1	Proteína hipotética SORBIDRAFT_01g014840	*Sorghum bicolor*	2,00E-07	32	110 (VIII)
PALA1529	EAY96411.1	Proteína hipotética OsI_18307	*Oryza sativa* Grupo Índica	7,00E-09	71	2 (U)
PR3A1215	EAZ37852.1	Proteína hipotética OsJ_22198	*Oryza sativa* Grupo Japônica	1,00E-27	66	110 (U)
PR3A1231	XP_002453709.1	Proteína hipotética SORBIDRAFT_04g011070	*Sorghum bicolor*	4,00E-29	69	-
PR3A1518	EEC69053.1	Proteína hipotética OsI_37894	*Oryza sativa* Grupo Índica	8,00E-07	33	124 (U)

5 CONSIDERAÇÕES FINAIS

A alocação de novos locos EST-SSR e de marcadores TRAP ao mapa de ligação prévio, construído por Palhares et al. (2012), contribuiu para um acréscimo de 1.379,8 cM no mapa de cana-de-açúcar, bem como para a sua reorganização em 65 grupos de co-segregação, que foram alocados em dez grupos de homo(eo)logia.

A população de mapeamento (F_1 segregante) apresentou diferentes amplitudes de variação da AUDPC (área sob a curva do progresso da doença), em função do ano de condução do ensaio de inoculação com *Sporisorium scitamineum* e, em ambos os casos, essa diferença foi maior que a amplitude entre os valores de AUDPC dos seus genitores (‘IAC66-6’ x ‘TUC71-7’), revelando genótipos transgressivos na progênie.

O uso da metodologia de mapeamento por intervalo composto (CIM) permitiu detectar associações entre os valores da AUDPC e os marcadores posicionados no mapa de ligação, tendo sido identificadas marcas AFLP, TRAP, EST-SSR e marcas ancoradas no retrotransposon *scIvana_1* associadas a locos quantitativos de resistência ao carvão. Foram detectados 8 QTLs, com predominância de alelos com efeitos aditivos, tanto para o aumento como a diminuição da susceptibilidade da população ao patógeno estudado.

As marcas TRAP, aqui sequenciadas, apresentaram similaridade a genes envolvidos no processo de defesa da planta contra estresses bióticos e abióticos e na ativação e desativação de processos bioquímicos vegetais, demonstrando a validade da técnica para a busca de genes candidatos associados à resposta a *S. scitamineum*.

O marcador PALA1350 mostrou-se associado a um QTL putativo com efeito negativo sobre o fenótipo, no primeiro ano de avaliação (QTL_3) e no comportamento médio dos genótipos nos diferentes anos (QTL_7), sugerindo que a enzima fenilalanina amônia liase, ou PAL, está envolvida na resposta da população a *S. scitamineum* e o possível uso deste marcador em seleção assistida.

REFERÊNCIAS

ABO, M.E.; OKUSANYA, B.A. Incidence and variability reaction of sugarcane smut (*Ustilago scitaminea* Syd.) isolates in greenhouse and laboratory test s in Nigeria. **Discovery and Innovation**, Nairobi, v.8, p. 227-231, 1996.

AITKEN, K.S.; JACKSON, P.A.; McINTYRE, C.L. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. **Theoretical Applied Genetics**, New York, v. 110, p. 789-801, 2005.

AITKEN, K.S.; JACKSON, P.A.; McINTYRE, C.L. Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar x *Saccharum officinarum* population. **Theoretical and Applied Genetics**, New York, v. 112, p. 1306-1317, 2006.

AITKEN, K.S.; JACKSON, P.A.; MCINTYRE, C.L. Construction of a genetic linkage map for *Saccharum officinarum* incorporating both simplex and duplex markers to increase genome coverage. **Genome**, Ottawa, v.50, p.742-756, 2007.

AITKEN, K.S.; HERMANN, S.; KARNO, K.; BONNETT, G.D.; McINTYRE, C.L.; JACKSON, P.A. Genetic control of yield related stalk traits in sugarcane. **Theoretical and Applied Genetics**, New York, v. 117, p. 1191-1203, 2008.

AKAIKE, H. A new look at statistical model identification. **The Institute of Electrical and Electronics Engeineers Transactions on Automatic Control**, Notre Dame, v. 19, n. 6, p. 716-723, 1974.

ALBERT, H.H.; SCHENCK, S. PCR amplification from a homolog of the *b*E mating-type gene as a sensitive assay for the presence of *Ustilago scitaminea* DNA. **Plant Disease**, St Paul, v. 80, p. 1189-1192, 1996.

ALBINO, J.C.; CRESTE, S.; FIGUEIRA, A. Mapeamento genético da cana-de-açúcar. **Biotecnologia Ciência & Desenvolvimento**, Brasília, n. 36, p. 82-91, 2006.

AL-JANABI, S.M.; HONEYCUTT, R.J.; SOBRAL, B.W.S. Chromosome assortment in *Saccharum*. **Theoretical Applied Genetics**, New York, v. 89, p. 959-963, 1994.

AL-JANABI, S.M.; HONEYCUTT, R.J.; McCLELLAND, M.; SOBRAL, B.W.S. A genetic linkage map of *Saccharum spontaneum* L. 'SES208'. **Genetics**, Bethesda, v. 134, p. 1249-1260, 1993.

AL-JANABI, S.M.; PARMESSUR, Y.; KROSS, H.; DHAYAN, S.; SAUMTALLY, S.; RAMDOYAL, K.; AUTREY, L.J.C.; DOOKUN-SAUMTALLY, A. Identification of a major quantitative trait locus (QTL) for yellow spot (*Mycovellosiella koepkei*) disease resistance in sugarcane. **Molecular Breeding**, Berlin, v. 19, p. 1-14, 2007.

ALWALA, S.; KIMBENG, C.A. Molecular Genetic Linkage Mapping in Saccharum: Strategies, Resources and Achievements. In: HENRY, R.J.; KOLE, C. **Genetics, genomics and breeding of sugarcane**. New Hampshire: Science Publishers, 2010. chap. 5, p. 69-96.

ALWALA, S.; KIMBENG, C.A.; GRAVOIS, K.A.; BISCHOFF, K.P. TRAP, a new tool for sugarcane breeding: comparison with AFLP and coefficient of parentage. **Journal American Society Sugar Cane Technologists**, [s.l], v. 26, p. 62-86, 2006b.

ALWALA, S.; COLLINS, A.K.; VEREMIS, J.C.; GRAVOIS, K.A. Linkage mapping and genome analysis in a *Saccharum* interspecific cross using AFLP, SRAP and TRAP markers. **Euphytica**, Dordrecht, v. 164, p. 37-51, 2008.

ALWALA, S.; KIMBENG, C.A.; VEREMIS, J.C.; GRAVOIS, K.A. Identification of molecular markers associated with sugar-related traits in a *Saccharum* interspecific cross. **Euphytica**, Dordrecht, v. 167, p. 127–142, 2009.

ALWALA, S.; SUMAN, A.; ARRO, J.A.; VEREMIS, J.C.; KIMBENG, C.A. Target Region Amplification Polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. **Crop Science**, Madison, v. 46, p. 448–455, 2006a.

AMORIM, L.; BERGAMIN-FILHO, A. HAU, B.; BERGER, R.D. Effects of sugarcane wounded-bud inoculation of *Ustilago scitamenea* on the disease process curve and yield. **Fitopatologia Brasileira**, Brasília, v. 25, p. 522-528, 2000.

ANDRU, S.; ALI, K.; ARRO, J.; PARCO, A.; KIMBENG, C.A.; BAISAKH, N. Molecular diversity among members of the *saccharum* complex assessed using TRAP markers based on lignin-related genes**. Bioenergy Research,** New York, v. 5, n.1, p. 197-205, 2012.

ANDRU, S.; PAN, Y.B.; THONGTHAWEE, S.; BURNER, D.M.; KIMBENG, C.A. Genetic analysis of the sugarcane (*Saccharum* spp.) cultivar 'LCP 85-384'. I. Linkage mapping using AFLP, SSR, and TRAP markers. **Theoretical and Applied Genetics**, New York, v. 123, n. 1, p.77-93, 2011.

ARMAS, R.; SANTIAGO, R.; LEGAZ, M.-E.; VICENTE, C. Levels of phenolic compounds and enzyme activity can be used to screen for resistance of sugarcane to smut (*Ustilago scitaminea*). **Australasian Plant Pathology**, Dordrecht, v. 36, 32–38, 2007.

ASNAGHI, C.; PAULET, F.; KAYE, C.; GRIVET, L.; DEU, M.; GLASZMANN, J.C.; D'HONT, A. Application of synteny across Poaceae to determine the map location of a sugarcane rust resistance gene. **Theoretical and Applied Genetics**, New York, v. 101, p. 962–969, 2000.

ASNAGHI, C.; ROQUES, D.; RUFFEL, S.; KAYE, C.; HOARAU, J.Y.; TÉLISMART, H.; GIRARD, J.C.; RABOIN, L.M.; RISTERUCCI, A.M.; GRIVET, L.; D'HONT, A. Targeted mapping of a sugarcane rust resistance gene (Bru1) using bulked segregant analysis and AFLP markers. **Theoretical and Applied Genetics**, New York, v. 108, p. 759-764, 2004

BADESCU, G.O.; NAPIER, R.M. Receptors for auxin: will it all end in TIRs? **Trends Plant Science**, London, v. 11, p. 217–223, 2006.

BERGAMIN-FILHO, A.; AMORIM, L.; CARDOSO, C.O.N.; SANGUINO, A.; IRVINE, J.E.; SILVA, W.M. Carvão da cana-de-açúcar e sua epidemiologia. **Boletim Técnico Copersucar**, Piracicaba, p. 1-23. 1987. Edição especial

BINDER, B.M.; WALKER, J.M.; GAGNE, J.M.; EMBORG, T.J.; HEMMANN, G.; BLEECKER, A.B.; VIERSTRA, R.D. The Arabidopsis EIN3 binding F-box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. **Plant Cell**, Rockville, v. 19, p. 509–523, 2007.

ASSOCIAÇÃO DOS PRODUTORES DE BIOENERGIA DE MATO GROSSO DO SUL. **ISO prevê crescimento de 3% ao ano da produção de cana**. Disponível em: <http://www.biosulms.com.br/perfnoticia.php?not=614>. Acesso em: 27 abr. 2012.

BOCK, K.R. Studies on sugar-cane smut (Ustilago scitaminea) in Kenya. **Transactions of the British Mycologial Society**, London, v. 47, 403-407, 1964

BORÉM, A.; VIEIRA, M.L.C; COLLI, W. **Glossário de Biotecnologia**. 2.ed. Viçosa: UFV, 2009, 186p.

BORRÁS-HIDALGO, O.; THOMMA, B.P.H.J.; 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**, Paris, v. 43, p. 1115–1121, 2005.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics,** Chicago, v. 32, p. 314-331, 1980.

BRAITHWAITE, K.S., BAKKEREN, G., CROFT, B.J., BRUMBLEY, S.M. Genetic variation in a worldwide collection of the sugarcane smut fungus *Ustilago scitaminea*. **Proceedings of the *Australian* Society of *Sugar Cane* Technologists**, Brisbane, v. 26, abstract p 48, 2004.

BRANDES, E.W. Origin, classification and characteristics. In: Artschwager, E.; Brandes, E.W. (Eds) **Sugarcande (*Saccharum officinarum* L.)**. Washington DC: U.S.D. Agriculture Handbook, 1958. v. 122. p. 1-35.

BREMER, G. Problems in breeding and cytology of sugar cane. I - A short history of sugar cane breeding the original forms of *Saccharum*. **Euphytica**, Wageningen, v. 10, p. 59-78, 1961a.

BREMER, G. Problems in breeding and cytology of sugar cane. IV - The origin of the increase of chromosome number in species hybrids of *Saccharum*. **Euphytica**, Wageningen, v. 10, p. 325-342, 1961b.

BUENO, C.R.N.C. **Infecção por *Sporisorium scitamineum* em cana-de-açúcar**: influência de variáveis ambientais e desenvolvimento de método para diagnose precoce. 2010. 68p.Tese (Doutorado em Fitopatologia) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2010.

BURNER, D.M.; LEGENDRE, B.L. Chromosome transmission and meiotic stability of sugarcane (*Saccharum* spp.) hybrid derivatives. **Crop Science**, Madison, v. 33, p.600-606, 1993.

BUTTERFIELD, M.K. **Marker assisted breeding in sugarcane:** a complex polyploidy. 2007. 73p. Tese (Doutorado) - University of Stellenbosch, Department of Agronomy, Faculty of AgriSciences; Matieland, South Africa, 2007.

CALDERON-VILLALOBOS, L.I.; NILL, C.; MARROCCO, K.; KRETSCH, T.; SCHWECHHEIMER, C. The evolutionarily conserved Arabidopsis thaliana F-box protein AtFBP7 is required for efficient translation during temperature stress. **Gene**, Amsterdam, v. 392, p. 106–116, 2007.

CAMPBELL, C.L.; MADDEN, L.V. **Introduction to plant disease epidemiology**. New York: John Wiley, 1990. 532p.

CARSON, D.L.; BOTHA, F.C. Preliminary analysis of expressed sequence tags for sugarcane. **Crop Science**, Madison, v. 40, p. 1769–1779, 2000.

CASAGRANDE, M.V; SANGUINO, A. Perda de produtividade devido a doença do carvão (Ustilago scitaminea Sydow) da cana-de-açúcar. In:SEMINÁRIO DE TECNOLOGIA AGRONÔMICA, 6., 1994.Piracicaba. **Anais...** Piracicaba: Esalq, 1994. p. 249-254.

CASAGRANDE, M.V; AMORIM, L. SANGUINO, A. Relação entre incidência de carvão e danos em diferentes estádios de desenvolvimento da cana-de-açúcar. **Fitopatologia Brasileira**, Brasília, v. 23, p. 374-378, 1998.

CASU, R.; DIMMOCK, C.; THOMAS, M.; BOWER, N.; KNIGHT, D.; GROF, C.; MCINTYRE. L.; JACKSON, P.; JORDAN, D.; WHAN, V.; DRENTH, J.; TAO, YUEZHI.; MANNERS, J. Genetic and expression profiling in sugarcane. **International Society of Sugar Cane Technologists**, Brisbane, v. 24, p. 626–627, 2001.

CASU, R.E.; DIMMOCK, C.M.; CHAPMAN, S.C.; GROF, C.P.L.; McINTYRE, C.L.; BONNETT, G.D.; MANNERS, J.M. Identification of differentially expressed transcripts from maturing stem of sugarcane by *in silico* analysis of stem expressed sequence tags and gene expression profiling. **Plant Molecular Biology**, Netherlands, v.54, p. 503-517, 2004.

CHAO, C.P.; HOY J.W.; SAXTON A.M. Heritability of resistance and repeatability of clone reactions to sugarcane smut in Louisiana. **Phytopathology**, St. Paul, v.80, n. 7, p. 622-626, 1990.

CHEN, J.; HU,J.; VICK, B.A.; JAN, C.C. Molecular mapping of a nuclear male-sterility gene in sunflower (*Helianthus annuus* L.) using TRAP and SSR markers. **Theoretical and Applied Genetics**, New York, v.113, p. 122-127, 2006.

CHEN, Y.; CHAO, Q.; TAN, G.; ZHAO, J.; ZHANG M.; JI, Q.; XU, M. Identification and fine-mapping of a major QTL conferring resistance against head smut in maize. **Theoretical and Applied Genetics**, New York, v. 117, p. 1241–1252, 2008.

CHURCHILL, G.A.; DOERGE, R.W. Empirical threshold values for quantitative trait mapping. **Genetics**, Austin, v. 138, n. 3, p. 963-971, 1994.

CLEGG, M.T.; CUMMING, M.P.; DURBIN, M.L. The evolution of plant nuclear genes. **Proceedings of the National Academy of Science of the United States of America**, Washington, DC , v. 94, p. 7791 – 7798, 1997.

COMPANHIA NACIONAL DE ABASTECIMENTO. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12_04_10_09_19_04_boletim_de_cana.pdf>. Acesso em: 04 maio 2012.

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.

COPERSUCAR. COOPERATIVA CENTRAL DOS PRODUTORES DE CANA, AÇÚCAR E ÁLCOOL DO ESTADO DE SÃO PAULO. **Avaliação de diferentes métodos de inoculação do carvão da cana-de-açúcar (Ustilago scitaminea Sydow)**. Piracicaba: Centro de Tecnologia Copersucar, 1995. 19 p. (Relatório Anual 1994/95)

CORDEIRO, G.M.; CASU, R; MCINTYRE, C.L.; MANNERS, J.M.; HENRY, R.J. Microsatellite markers from sugarcane (*Saccharum* spp) ESTs cross transferable to *Erianthus* and *Sorghum*. **Plant Science**, Ireland, v. 160, p. 1115–1123, 2001.

CORDEIRO, G.M.; ELIOTT, F.; MCINTYRE, C.L.; CASU, R.E.; HENRY, R.J. Characterisation of single nucleotide polymorphisms in sugarcane ESTs. **Theoretical and Applied Genetics**, New York, v. 113, n. 2, p. 331-343, 2006.

CRESTE, S.; TULMANN-NETO, A.; FIGUEIRA, A.V.O. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Report**, Athens, v. 19, p. 299-306, 2001.

CRESTE, S.; ACCORONI, K.A.G.; PINTO, L.R.; VENCOVSKY, R.; GIMENES, M. A.; XAVIER, M. A.; LANDELL, M.G.A. Genetic variability among sugarcane genotypes based on polymorphisms in sucrose metabolism and drought tolerance genes. **Euphytica**, Dordrecht, v. 172, n. 3, p.435-446, 2010.

CROALL, D.E.; ERSFELD, K. The calpains: modular designs and functional diversity. **Genome Biology**, London, v. 8, p. 218.1-218.11, 2007.

CUADRADO, A.; ACEVEDO, R.; MORENO DÍAZ DE LA ESPINA, S.; JOUVE, N.; DE LA TORRE, C. Genome remodelling in three modern *S. officinarum* x *S. spontameum* sugarcane cultivars. **Journal of Experimental Botany**, Oxford, v. 55, n. 398, p. 847-854, 2004.

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

D'HONT, A.; GLASZMANN, J.C. Sugarcane genome analysis with molecular morkers, a first decade of research. **Proceedings International Society of Sugarcane Technologists**, [s.l], v.24, p. 556-559, 2001.

D'HONT, A.; PAULET, F.; GLASZMANN, J.C. Oligoclonal interespecific origin of "North Indian" and "Chinese" sugarcanes. **Chromosome Research**, Oxford, v. 10, p. 253-262, 2002.

D'HONT, A.; ISON, D.; ALIX, K.; ROOUX, C.; GLASZMANN, J.C. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. **Genome**, Ottawa, v. 41, p. 221-225, 1998.

DA GLORIA, B.A.; ALBERNAS, M.C.C.; AMORIM, L.;BERGAMIN-FILHO, A. Morphological characteristics of sugarcane clones, susceptible and resistant to smut (Ustilago scitaminea). In: RAO, G.P.; BERGAMIN-FILHO, A.; MAGAREY, R.C.; AUTREY, L.J.C. **Sugarcane pathology**: Fungal diseases. Enfield:Science Publishers, 1999. p. 167-182.

DA SILVA, J.; BRESSIANI, J.A. Sucrose synthase molecular marker associated with sugar content in elite sugarcane progeny. **Genetics and Molecular Biology**, Ribeirão Preto, v. 28, p. 294-298, 2005.

DA SILVA, J.A.; SORRELLS, M.E.; BURNQUIST, W.; TANKSLEY, S.D. RFLP linkage map of *Saccharum spontaneum*. **Genome**, Ottawa, v. 36, p. 782-791, 1993.

DA SILVA, J.A.G.; HONEYCUTT, R.J.; BURNQUIST, W.; AL-JANABI, S.M.; SORRELLS, M.E.; TANKSLEY, S.D.; SOBRAL, W.S. *Saccharum spontaneum* L. 'SES 208' genetic linkage map combining RFLP and PCR based markers. **Molecular Breeding**, Dordrecht, v. 1, p. 165-179, 1995.

DA SILVA, J.G.A. **A method for genome mapping of autopolyploids and its application to sugarcane *Saccharum* spp**. 1993. 384p. Tese (Doutorado) – Cornell University, New York, Ithaca, 1993.

DANIELS, J.; SMITH, P.; PATON, N.; WILLIAMS, C.A. The origin of the genus Saccharum. **Sugarcane Breeding Newsletter**,Stacks, v. 36, p. 24–39, 1975.

DANIELS, J.; ROACH, B.T. Taxonomy and evolution. In: HEINZ, D.J. (Ed.), **Sugarcane improvement through breeding**. Holanda: Elsevier Press, 1987. p. 7-84.

DAUGROIS, J.H.; GRIVET, L.; ROQUES, D.; HOARAU, J.Y.; LOMBARD, H.; GLASZMANN, J.C.; D'HONT, A. A putative major gene for rust resistance linked with a RFLP marker in sugarcane cultivar 'R570'. **Theoretical and Applied Genetics**, New York, v. 92, p. 1059–1064, 1996.

DE ARMES, R.; SANTIAGO, R.; LEGAZ, M-E.; VICENTE, C. Levels of phenolic compounds and enzyme activity can be used to screen for resitance of sugarcane to smut (*Ustilago scitaminea*). **Australasian Plant Pathology**, Dordrecht, v.36, p. 32-38, 2007.

DE BRUXELLES, G.L.; PEACOCK, W.J.; DENNIS, E.S.; DOLFERUS, R. Abscisic acid induces the alcohol dehydrogenase gene in Arabidopsis. **Plant Physiology**, Rockville, v. 111, p. 381–391, 1996

DEMPSTER, A.P.; LAIR, N.M.; RUBIN, D.B. Maximum likehood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society**, Series B, London, v. 39, p. 1-38, 1977.

DILL, A.; THOMAS, S.G.; HU, J.; STEBER, C.M.; SUN, T.P. The Arabidopsis F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellininduced degradation. **Plant Cell**, Rockville, v. 16, p. 1392–1405, 2004.

DOERGE, R.W. Mapping and analysis of quantitative trait loci in experimental populations. **Nature Reviews Genetics**, New York, v. 3, p. 43-52, 2002.

DOLFERUS, R.; PEACOCK, W.J.; DENNIS, E.S. Differential interactions of promoter elements in stress responses of the Arabidopsis *ADH* gene. **Plant Physiology**, Rockville, v. 105**,** p. 1075–1078, 1994

EBERLEIN, C.; GUTTIERE, M.J.; THILL, D.C.; BAERG, R.J. Altered acetolactate synthase activity in ALS - inhibitor resistant prickly lettuce (*Lactuca serriola*). **Weed Science**, Lawrence, v.45, n.2, p.212-217, 1997.

EDMÉ, S.J.; GLYNN, N.G.; COMSTOCK, J.C. Genetic segregation of microssatellite markers in *Saccharum officinarum* and *S. spontaneum*. **Heredity**, London, v. 97, p. 366-375, 2006.

ELSTON, D.A.; SIMMONDS, N.W. models of sugarcane smut disease and their implications for testing variety resistance. **Journal of Applied Ecology**, Malden, v. 25, p. 319-329, 1988.

EWING, B.; HILLIER, L.; WENDEL, M.; GREEN, P. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Research**, New York, v. 8, p. 175-185, 1998.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to Quantitative Genetics**. 4th ed. Essex, UK: Longman, 1996. 464 p.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: EMBRAPA-CENARGEN, 1998. 220 p.

FERREIRA, S.A.; COMSTOCK, J.C. Smut. In: RICAUD, C.; EGAN, B.T.; GILLASPIE, A.G.; HUGHES, C.G. (Ed.). **Diseases of Sugarcane – Major diseases.** Amsterdam: Elsevier: 1989. p. 211–229.

FUNGARO, M.H.P.; VIEIRA, M.L.C. Marcadores moleculares. In: SERAFINI, L.A.; BARROS, N.M.; AZEVEDO, J.L. **Biotecnologia na agricultura e na agroindústria.** Guaíba: Agropecuária, 2001. p. 153-200.

GARCIA, A.A.F.; KIDO, E.A.; MEZA, A.N.; SOUZA, H.M.B.; PINTO, L.R.; PASTINA, M.M.; LEITE, C.S.; DA SILVA, J.A.G.; ULIAN, E.C.; FIGUEIRA, A.V.O.; SOUZA, A.P. Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. **Theoretical Applied Genetics**, Berlin, v. 112, n. 2, p. 298-314, 2006.

GARSMEUR, O.; CHARRON, C.; BOCS, S.; JOUFFE, V.; SAMAIN, S.; COULOUX, A.; TAN DROC, G.; ZINI, C.; GLASZMANN, J-C.; VAN SLUYS, M-A.; D'HONT, A. High homologous gene conservation despite extreme autopolyploid redundancy in sugarcane. **New Phytologist**, Malden, v. 189, p. 629-642, 2011.

GAZAFFI, R. **Desenvolvimento de modelo genético-estatístico para mapeamento de QTLs em progênie de irmãos completos, com aplicação em cana-de-açúcar**. 2009. 103p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2009.

GIGLIOTI, E.A.; TOKESHI, H. Influência das bainhas e folhas na incidência de carvão da cana-de-açúcar. **Summa Phytopathologica**, Botucatu, v. 17, n.1, p. 17, 1991.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic Linkage Maps of Eucalyptus grandis and Eucalyptus urophylla Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers. **Genetics,** Austin, v. 137, p. 1121-1137, 1994.

GRIVET, L.; ARRUDA, P. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Current Opinion in Plant Biology**, London, v. 5, n. 2, p. 122-127, 2002.

GRIVET, L.; DANIELS, C.; GLASZMANN, J.C.; D'HONT, A. A review of recent molecular genetics evidence for sugarcane evolution and domestication. **Ethnobotany Research & Applications**, Fort Worth, v. 2, p. 9-17, 2004.

GRIVET, L.; GLASZMANN, J.C.; VINCENTZ, M.; DA SILVA, F.; ARRUDA, P. ESTs as a source for sequence polymorphism discovery in sugarcane: example of the *Adh* genes. **Theoretical and Applied Genetics**, New York , v. 106, n. 2, p. 190-197, 2003.

GRIVET, L.; D'HONT, A., ROQUES, D.; FELDMANN, P.; LANAUD, C.; GLASZMANN, J.C. RFLP mapping in cultivated sugarcane *(Saccharum* spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid. **Genetics**, Bethesda, v. 142, p. 987-1000, 1996.

GUIMARÃES, C.T.; HONEYCUTTR.J.; SILLS, G.R.; SOBRAL, B.W.S. Genetic maps of S*accharum officinarum* L. and S*accharum robustum* Brandes & Jew. ex grassl. **Genetic and Molecular Biology**, Ribeirão Preto, v. 22, p. 125-132, 1999.

HANAHAN, D. Studies on the transformation of *Escherichia coli* with plasmids. **Journal of Molecular Biology**, London, v. 166, p. 557-580, 1983.

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.

HENDERSON, C. H. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics**, Malden, v. 31, n. 2, p. 423-447, 1975.

HOARAU, J.Y.; GRIVET, L.; OFFMANN, B.; RABOIN, L.M.; DIORFLAR, J.P.; PAYET, J.; HELLMANN, M.; D'HONT, A.; GLASZMANN, J.C. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). II. Detection of QTLs for yield components. **Theoretical and Applied Genetics**, New York, v. 105, p. 1027-1037, 2002.

HOARAU, J.Y.; OFFMANN, B.; D´HONT, A.; RISTERUCCI, A.M.; ROQUES, D., GLASMANN, J.C.; GRIVET, L. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. **Theoretical Applied Genetics**, New York, v. 103, p. 84-97, 2001.

HOY, J.W.; GRISHAM, M.P.; BRENDA, G.T.A. Effect of Louisiana growing conditions on the overwintering of smut-infected sugarcane. **Sugar Cane**, Washington, D.C., v. 3, p. 11-15, 1987.

HOY, J.W.; GRISHAM, M.P. Spread and increase of sugarcane smut in Louisiana. **Phytopathology**, St. Paul, v. 78, p. 1371-1376, 1988.

HU, J.; VICK, B.A. Target region amplification polymorphism: a novel marker technique for plant genotyping. **Plant Molecular Biology Reporter**, New York, v. 21, p. 289-294, 2003.

IAMSUPASIT, N.; PLIANSINCHAI, U.; LEABWAN, U.; PA-OBLEK, S.; LAIRUNGREONG, C. Selection for smut resistance in two sugarcane population. In: WILSON, J.R.; HOGARTH, D.M.; CAMPBELL, J.A.; GARSIDE, A.L. (Ed.). **Sugarcane :** Research Towards Efficient and Sustainable Production. Brisbane: CSIRO, 1996. p. 149-150.

IRVINE, J.E. *Saccharum* species as horticultural classes. **Theoretical Applied Genetics**, Berlin, v. 98, p. 186-194, 1999.

JACOB, H. J.; LINDPAINTNER, K.; LINCOLN, S. E.; KUSUMI, K.; BUNKER, R. K.; MAO, Y.-P.; GANTEN, D.; DZAU, V. J.; LANDER, E. S. Genetic mapping of a gene causing hypertensive rat. **Cell**, Cambridge, v. 67, p. 213-224, 1991.

JANNOO, N.; GRIVET, L.; DAVID, J.; D'HONT, A.; GLASZMANN, J. Differential chromosome pairing affinities at meiosis in polyploid sugarcane revealed by molecular markers. **Heredity**, London, v. 93, n. 5, p. 460-467, 2004.

JANNOO, N.; GRIVET, L.; CHANTRET, N.; GARSMEUR, O.; GLASZMANN, J.C.; ARRUDA, P.; D'HONT, A. Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome. **Plant Journal,** Malden, v. 50, n. 4, p. 574-585, 2007.

JANNOO, N.; GRIVET, L.; SEGUIN, M.; PAULET, F.; DOMAINGUE, R.; RAO, P.S.; DOOKUN, A.; D'HONT, A., GLASZMANN, J.C. Molecular investigation of the genetic base of sugarcane cultivars. **Theoretical and Applied Genetics**, New York, v.99, p. 171-184, 1999.

JANSEN, R.C.; STAM, P. High resolution of quantitative traits into multiple loci via interval mapping. **Genetics**, Bethesda, v. 136, p. 1447-1455, 1994.

JORDAN, D.R.; CASU, R.E.; BESSE, P.; CARROLL, B.C.; BERDING, N.; McINTYRE, C.L. Markers associated with stalk number and suckering in sugarcane colocate with tillering and rhizomatousness QTLs in sorghum. **Genome**, Ottawa, v. 47, p. 988-993, 2004.

JORDÃO Jr., H. **Desenvolvimento de um sistema baseado em marcadores moleculares de DNA do tipo microssatélites para identificação de variedades de cana-de-açúcar**. 2009. 106p. Dissertação (Mestrado em Biologia Funcional e Molecular) – Universidade Estadual de Campinas, Campinas, 2009.

KALIA, R.K.; RAI, M.K.; KALIA, S.; SINGH, R.; DHAWAN, A.K. Microsatellite markers: an overview of the recent progress in plants. **Euphytica**, Dordrecht, v. 177, p. 309-334, 2011.

KAO, C.H., ZENG, Z.B.; TEASDALE, R D. Multiple interval mapping for quantitative trait loci. **Genetics**, Bethesda, v. 152, p. 1203-1216, 1999.

KHANNA, R.; KRONMILLER, B.; MASZLE, D.R.; COUPLAND, G.; HOLM, M.; MIZUNO, T.; WU, S-H. The Arabidopsis B-Box Zinc Finger Family. **The Plant Cell**, Waterbury, v. 21, p. 3416–3420, 2009.

KIM, H.S.; DELANEY, T.P. Arabidopsis SON1 is an F-box protein that regulates a novel induced defense response independent of both salicylic acid and systemic acquired resistance. **Plant Cell**, Rockville, v. 14, p. 1469–1482, 2002.

KOSAMBI, D. D. The estimation of map distance from recombination values. **Annuaire Eugenetic**, Washington, v. 12, p. 172-175, 1944.

KUYKENDALL, D.; SHAO, J.; TRIMMER, K. A Nest of LTR Retrotransposons Adjacent the Disease Resistance-Priming Gene NPR1 in Beta vulgaris L. U.S. Hybrid H20. **International Journal of Plant Genomics**, New York, v. 2009, Article ID 576742, 2009.

LAI, C.P.; LEE, C.L.; CHEN, P.H.; WU, S.H.; YANG, C.C.; SHAW, J.F. Molecular analyses of the Arabidopsis TUBBY-like protein gene family. **Plant Physiology**, Rockville, v. 134, p. 1586–1597, 2004.

LANDELL, M.G.A.; BRESSIANI, J.A. Melhoramento genético, caracterização e manejo varietal. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A (Ed.). **Cana-de-açúcar**. Campinas: Instituto Agronômico, 2008. p. 101-156.

LANDER; E.S.; BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Bethesda, v. 121, p. 185-199, 1989.

LAO, M.; ARENCIBIA, A.D.; CARMONA, E.R.; ACEVEDO, R.; RODRÍGUEZ, E.; LEÓN, O.; SANTANA, I. Differential expression analysis by cDNA-AFLP of *Saccharum* spp. after inoculation with the host pathogen *Sporisorium scitamineum*. **Plant Cell Reports**, New York, v. 27, p. 1103–1111, 2008.

LE CUNFF, L.; GARSMEUR, O; RABOIN,L.M ; PAUQUET, J.; TELISMART, H.; SELVI, A.; GRIVET, L.; PHILIPPE, R.; BEGUM, D.; DEU, M.; COSTET, L.; WING, R.; GLASZMANN, J.C.; D'HONT, A. Diploid/Polyploid Syntenic Shuttle Mapping and Haplotype-Specific Chromosome Walking Toward a Rust Resistance Gene (Bru1) in Highly Polyploid Sugarcane (2n ~ 12x ~ 115). **Genetics**, Bethesda, v. 180, p. 649–660, 2008.

LI, G.; QUIROS, C.F. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. **Theoretical and Applied Genetics**, New York, v. 103, p. 455-461, 2001.

LI, X.H.; WANG, Z.H.; GAO, S.R.; SHI, H.L.; ZHANG, S.H.; GEORGE, M.L.C.; LI, M.S.; XIE, C.X. Analysis of QTL for resistance to head smut (*Sporisorium reiliana*) in maize. **Field Crops Research**, Amsterdam, v. 106, p. 148–155, 2008.

LIN, M.; LOU, X.Y.; CHANG, M.; WU, R. A general statistical framework for mapping quantitative trait loci in nonmodel systems: issue for characterizing linkage phases. **Genetics**, Bethesda, v. 165, p. 901-9013, 2003.

LIU, J.; QUE, Y.; GUO, J.; XU, L.; WU, J.; CHEN, R. Molecular cloning and expression analysis of a WRKY transcription factor in sugarcane. **African Journal of Biotechnology**, Nairobi, v.11, p. 6434-6444, 2012.

LIU, Z.H.; ANDERSON, J.A.; HU, J.; FRIESEN, T.L.; RASMUSSEN, J.B.; FARIS, J.D. A wheat intervarietal genetic linkage map based on microsatellite and target region amplified polymorphism markers and its utility for detecting quantitative trait loci. **Theoretical and Applied Genetics**, New York, v. 111, p. 782-794, 2005.

LÜBBERSTEDT, T.; XIA, X.C.; TAN, G.; LIU, X.; MELCHINGER, A.E. QTL mapping of resistance to Sporisorium reiliana in maize. **Theoretical Applied Genetics**, New York, v. 99, p. 593–598, 1999.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 980p.

MACCHERONI, W.; MATSUOKA, S. Manejo das principais doenças da cana-de-açúcar. In: SEGATO, S.V.; PINTO, A.S.; JENDIROBA, E.; NOBREGA, J.C.M. **Atualização em produção de cana-de-acúcar**. Piracicaba: CP 2, 2006. cap. 14, p. 239-256.

MACCHERONI, W.; JORDÃO, H.; De GASPARI, R.; De MOURA, G.L.; MATSUOKA, S. Development of a dependable microsatellite-based fingerprinting system for sugarcane. **Sugar Cane International**, Amsterdam, v. 27, n. 2, p. 8-13, 2009.

MACHADO, G.R. **Sugarcane variety notes**: na international directory. 7th.ed. Piracicaba, Brazil: Guilherme Rossi Machado, 2001. 132p.

MARCHLER-BAUER, A.; LU, S.; ANDERSON, J.B.; CHITSAZ, F.; DERBYSHIRE, M.K.; DEWEESE-SCOTT, C.; FONG, J.H.; GEER, L.Y.; GEER, R.C.; GONZALES, N.R.; GWADZ, M.; HURWITZ, D.I.; JACKSON, J.D.; KE, Z.; LANCZYCKI, C.J.; LU, F.; MARCHLER, G.H.; MULLOKANDOV, M.; OMELCHENKO, M.V.; ROBERTSON, C.L.; SONG, J.S.; THANKI, N,; YAMASHITA, R.A.; ZHANG, D.; ZHANG, N.; ZHENG, C.; BRYANT, S.H. CDD: a Conserved Domain Database for the functional annotation of proteins. **Nucleic Acids Research**, Oxford, v.39, p. 225-229, 2011.

MARCONI, T.G. **Mapa Funcional em cana-de-açúcar utilizando marcadores moleculares baseados em SSR e SNP**. 2011. 147 p. Tese (Doutorado em Genética e Biologia Molecular) – curso de Genética e Biologia Molecular, Universidade de Campinas, Campinas, 2011.

MARCONI, T.G.; COSTA, E.A.; MIRANDA, H.R.C.A.N.; MANCINI, M.C.; CARDOSO-SILVA, C.B.; OLIVEIRA, K.M.; PINTO, L.R.; MOLLINARI, M.; GARCIA, A.A.F.; SOUZA, A.P. Functional markers for gene mapping and genetic diversity studies in sugarcane**. BMC Research Notes**, London, v. 4, p. 264, 2011.

MARGARIDO, G.R.A. **Mapeamento de QTLs em múltiplos caracteres e ambientes em um cruzamento comercial de cana-de-açúcar usando modelos mistos**. 2011. 107p. Tese (Doutoraro em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2011.

MARGARIDO, G.R.A.; SOUZA, A.P.; GARCIA, A.A.F. OneMap: software for genetic mapping in outcrossing species. **Hereditas**, Malden, v. 144, p. 78-79, 2007.

MARGARIDO, G.R.A.; MOLLINARI, M.;GARCIA, A.A.F. OneMap Tutorial. Disponível em: <http://brieger.esalq.usp.br/CRAN/web/packages/onemap/index.html>. Acesso em 02 mar. 2011.

MATHER, K.; JINKS, L.L. **Introdução à genética biométrica**. Ribeirão Preto: Sociedade Brasileira de Genética, 1984. 242p.

MATSUOKA, S.; GARCIA, A.A.F.; ARIZONO, H. Melhoramento da cana-de-açúcar. In: BORÉM, A. (Ed). **Melhoramento de espécies cultivadas**. 2.ed. Viçosa: UFV, 2005. p. 225-274.

MATSUOKA, S.; SORGI, R.A.; MASUDA, Y.; ARIZONO, H. Realiability and feasibility of the needle-bud puncture method for rapid identification of smut susceptible sugarcane clones. **Proceedings International Society Cane Technology**, Guatemala, v. 19, p. 375-385, 1986.

McINTYRE, C.L.; CASU, R.E.; DRENTH, J.; KNIGHT, D.; WHAM, V.A.; CROFT, B.J.; JORDAN, D.R.; MANNERS, J.M. Resistance gene analogues in sugarcane and sorghum and their association with quantitative trait loci for rust resistance. **Genome**, Ottawa, v. 48, p. 391-400, 2005b.

MCINTYRE, C.L.; JACKSON, M.; CORDEIRO, G. M.; AMOUYAL, O.; HERMANN, S.; AITKEN, K. S.; ELIOTT, F.; HENRY, R. J.; CASU, R. E.; BONNETT, G. D. The identification and characterisation of alleles of sucrose phosphate synthase gene family III in sugarcane. **Molecular Breeding**, Dordrecht, v. 18, n. 1, p. 39-50, 2006.

McINTYRE, C.L.; WHAN, V.A.; CROFT, B.; MAGAREY, R.; SMITH, G.R. Identification and validation of molecular markers associated with Pachymetra root rot and brown rust resitance in sugarcane using map- and association-based approaches. **Molecular Breeding**, Berlin, v. 16, p. 151-161, 2005a.

MELOTTO, M.; COELHO, M.F.; PEDROSA-HARAND, A.; KELLY, J.D.; CAMARGO, L.E.A. The anthracnose resistance locus Co-4 of common bean is located on chromosome 3 and contains putative disease resistance-related genes. **Theoretical and Applied Genetics**, New York , v. 109, p. 690–699, 2004.

MICHELMORE, R.W.; MEYERS, B.C**.** Clusters of resistance genes in plants evolve by divergent selectionand a birth-and-death process. **Genome Research**, Woodbury, v. 8, p. 1113–1130, 1998.

MIKLAS, P.N.; HU, J. GRÜNWALD, N.J. LARSEN, K.M. Potential application of TRAP (target region amplification polymorphism) markers for mapping and tagging disease resistance traits in common bean. **Crop Science,** Madison, v.46, p. 910-916, 2006.

MILACH, S.C.K. Marcadores de DNA: aplicações no melhoramento de plantas. **Biotecnologia Ciência e Desenvolvimento**, Brasília, v. 5, p. 14-17, 1998.

MING, R.; LIU, S.C.; MOORE, P.H.; IRVINE, J.E.; PATERSON, A.H. QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. **Genome Research**, Cold Spring Harbor, v. 11, p. 2075–2084, 2001.

MING, R.; DEL MONTE, T.A.; HERNANDEZ, E.; MOORE, P.H.; IRVINE, J.E.; PATERSON, A.H. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes. **Genome**, Ottawa, v. 45, p. 794-803, 2002b.

MING, R.; LIU, S.C.; LIN, Y.R.; DA SILVA, J.A.G.; WILSON, W.; BRAGA, D.; VAN DEYNZE, A.; WENSLAFF, T.F.; WU, K.K.; MOORE, P.H.; BURNQUIST, W.; SORRELLS, M.E.; IRVINE, J.E.; PATERSON, A.H. Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. **Genetics**, Bethesda, v. 150, p. 1663–1682, 1998.

MING, R.; MOORE, P.H.; WU, K.K.; D'HONT, A.; GLASZMANN, J.C.; TEW, T.L. Sugarcane improvement through breeding and biotechnology. **Plant Breeding Reviews**, New Jersey, v. 27, p. 15–118, 2006.

MING, R.; WANG, Y.W.; DRAYE, X.; MOORE, P.H.; IRVINE, J.E.; PATERSON, A.H. Molecular dissection of complex traits in autopolyploids: mapping QTLs affecting sugar yield and related traits in sugarcane. **Theoretical and Applied Genetics**, New York, v. 105, p. 332-345, 2002a.

MOLLINARI, M. **Desenvolvimento de um modelo para construção de mapas genéticos em autopoliploides, com aplicação em cana-de-açúcar**. 2012. 98p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2012.

MUDGE, J.; ANDERSON, W.R.; KEHRER, R.L.; FAIRBANKS, D.J. A RAPD genetic map of *Saccharum officinarum*. **Crop Science**, Madison, v. 36, p. 1362-1366, 1996.

MUKHERJEE, S.K. Origin and distribution of *Saccharum*. **Botanical Gazette**, Chicago, v. 119, p. 55–61, 1957.

MURRAY, M.G.; THOMPSON, W.F. Rapid isolation of high molecular weight plant DNA. **Nucleic Acids Research**, Oxford, v. 8, n. 19, p. 4321-4325, 1980.

NATINAL CENTER FOR BIOTECHNOLOGY INFORMATION. Disponível em: <http://blast.ncbi.nlm.nih.gov/>. Acesso em: 23 fev. 2012.

NOBREGA, J.C.M.; DORNELAS, M.C. Biotecnologia e melhoramento da cana-de-açúcar. In: SEGATO, S.V.; PINTO, A.S.; JENDIROBA, E.; NOBREGA, J.C.M. **Atualização em produção de cana-de-açúcar.** Piracicaba: Livroceres, 2006. cap. 3, p. 39-56.

NZIOKI, H.S.; JAMOZA, J.E. Assessment of yield loss due to sugarcane smut (*Ustilago scitaminea*) infection in kenya. In: RONO, J.K.; WAWIRE, N.W.; KAHORA, F.; JUMA, F.; AMOLO, R. **Kenya Sugar Research Foundation.** Nairobi: KENLIB, 2006. 105 p. (Technical Bulletin, 1).

OETH, P.; BEAULIEU, M.; PARK, C.; KOSMAN, D.; MISTRO, G. del.; VAN DEN BOOM, D.; JURINKE, C. **iPLEX™ Assay**: Increased Plexing Efficiency and Flexibility for MassARRAY® System Through Single Base Primer Extension with Mass-Modified Terminators. San Diego: Sequenom, 2007. 12 p.

OLIVEIRA, K.M.; PINTO, L.R.; MARCONI, T.G.; MARGARIDO, G.R.A.; PASTINA, M.M.; TEIXEIRA, L.H.M.; FIGUEIRA, A.V.O.; ULIAN, E.C.; GARCIA, A.A.F.; SOUZA, A.P. Funcional integred genetic linkage map base on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding**, Dordrecht, v. 20, p. 189-208, 2007.

OLIVEIRA, K.M.; PINTO, L.R.; MARCONI, T.G.; MOLLINARI, M.; ULIAN, E.C.; CHABREGAS, S.M.; FALCO, M.C.; BURNQUIST, W.; GARCIA, A.A.F.; SOUZA, A.P. Characterization of new polymorphic functional markers for sugarcane. **Genome**, Ottawa, v. 52, p. 191–209, 2009.

PALHARES, A.C. **Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar (*Saccharum* spp.)**. 2010. 98p. Dissertação (Mestrado em Genética e Melhoramento) – curso de Genética e Melhoramento de Plantas, Escola Superir de Agricultura "Luiz de Queiróz", Universidade de São Paulo, Piracicaba, 2010.

PALHARES, A.C.; RODRIGUES-MORAIS, T.B.; VAN SLUYS, M-A.; DOMINGUES, D.S.; MACCHERONI, JR.W.; JORDÃO JR. H.; MARCONI, T.G.; SOUZA, A.P.; MOLLINARI, M.; GAZAFFI, R.; GARCIA, A.A.F.; VIEIRA, L.M.C. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers, **BMC Genetics**, London, v. 13, p. 1-16, 2012. doi:10.1186/1471-2156-13-51.

PASTINA, M.M.; MALOSETTI, M.; GAZAFFI, R.; MOLLINARI, M.; MARGARIDO, G.R.A.; OLIVEIRA, K.M.; PINTO, L.R.; SOUZA, A.P.; VAN EEUWIJK, F.A.; GARCIA, A.A.F. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theoretical and Applied Genetics**, New York, v. 124, p. 835–849, 2012.

PATERSON, A.H.; LIN, Y.R.; LI, Z.; SCHERTZ, K.F.; DOEBLEY, J.F.; PINSON, S.R.M.; LIU, S.C.; STANSEL, J.W.; IRVINE, J.E. Convergent domestication of cereal crops by indiendent mutations at corresponding genetic loci. **Science**, Washington DC, v. 269, p. 1714-1718, 1995.

PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, Oxford, v. 58, p. 545-554, 1971.

PAYNE, R.W.; MURRAY, D.A.; HARDING, S.A.; BAIRD, D.B.; SOUTAR, D.M. Introduction to GenStat 12 for Windows. **VSN International**, Hemel Hempstead, 2009. 37 p.

PEROS, J.P.; BAUDIN, P. Etude de la variabilité d' *Ustilago scitaminea* Syd., agent du carbón de la canne à sucre. **L'Agronomie Tropicale**, Montpellier, v. 38, p.234-240, 1983.

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

PINTO, L.R.; GARCIA, A.A.F.; PASTINA, M.M.; TEIXEIRA, L.H.M.; BRESSIANI, J.A.; ULIAN, E.C.; BIDOIA, M.A.P.; SOUZA, A.P. Analysis of genomic and functional RFLP derived markers associated with sucrose content, fiber and yield QTLs in a sugarcane (*Saccharum* spp.) commercial cross. **Euphytica**, Wageningen, v. 172, p. 313-327, 2010.

PINTO, L.R.; OLIVEIRA, K.M.; MARCONI, T.G.; GARCIA, A.A.F.; ULIAN, E.C.; SOUZA, A.P. Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. **Plant Breeding,** Berlin**,** v. 125, p. 378–384, 2006.

PINTO, L.R.; OLIVEIRA, K.M.; ULIAN, E.C.; GARCIA, A.A.F.; SOUZA, A.P. Survey in the expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome**, Ottawa, v. 47, p. 795–804, 2004.

PIPERIDIS G.; D'HONT A. Chromosome composition analysis of various *Saccharum* interspecific hybrids by genomic *in situ* hybridization (GISH). **Proceedings International Society of Sugarcane Technologists**, Brisbane, v. 24, p. 565-566, 2001.

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

PIPERIDIS, N.; JACKSON, P.A.; D'HONT, A.; BESSE, P.; HOARAU, J.Y.; COURTOIS, B.; AITKEN, K.S.; McINTYRE, C.L. Comparative genetics in sugarcane enables structured map enhancement and validation of marker-trait associations. **Molecular Breeding**, Berlin, v. 21, p. 233-247, 2008.

PRICE, S. Cytology of Chinese and North Indian sugarcanes. **Economic Botany** , New York, v. 22, p.155–164, 1968.

QIAO, H.; WANG, F.; ZHAO, L.; ZHOU, J.; LAI, Z.; ZHANG, Y.; ROBBINS, T.P.; XUE, Y. The F-box protein AhSLF-S2 controls the pollen function of S-RNase based self-incompatibility. **Plant Cell**, Rockville, v. 16, p. 2307–2322, 2004.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, 2009. Disponível em: <http://www.r-project.org> Acesso: 18 maio 2011.

RABOIN, L.M.; HOARAU, J.Y.; COSTET, L.; TELISMART, H.; GLASZMANN J C.; D'HONT A. Progress in genetic mapping of sugarcane smut resistance. **Proceedings of the South African Sugar Technologists Association**, Durban, v.77, p. 134-141, 2003.

RABOIN, L.M.; OLIVEIRA, K.M.; LECUNFF, L.; TELISMART, H.; ROQUES, D.; BUTTERFIELD, M.; HOARAU, J.Y.; D'HONT, A. Genetic mapping in sugarcane, a high polyploid, using bi-parental progeny: identification of a gene controlling stalk colour and A new rust resistance gene. **Theoretical and Applied Genetics**, New York, v. 112, n. 7, p. 1382-1391, 2006.

RABOIN, L.M.; PAUQUET, J.; BUTTERFIELD, M.; D'HONT, A.; GLASZMANN, J.C. Analysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. **Theoretical and Applied Genetics**, New York, v. 116, p. 701-714, 2008.

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

RAMAKRISHNA, W.; EMBERTON, J.; OGDEN, M.; SANMIGUEL, P.; BENNETZEN, J.L. Structural Analysis of the Maize *Rp1* Complex Reveals Numerous Sites and Unexpected Mechanisms of Local Rearrangement. **The Plant Cell**, Waterbury, v. 14, p. 3213–3223, 2002.

RAMALHO, M.A.; SANTOS, J.B.; PINTO, C.A.B.P. **Genética na agropecuária**. 3. ed. Lavras: UFLA, 2004. 472 p.

RAMSEY, J.; SCHEMSKE, D.W. Neopolyploidy in flowering plants. **Annual Review of Ecology and Systematics**, Palo Alto, v.33, p. 589-639, 2002.

REFFAY, N.; JACKSON, P.A.; AITKEN, K.S.; HOARAU, J.-H.; D'HONT, A.; BESSE, P.; McINTYRE, C.L. Characterisation of genome regions incorporated from an important wild relative into Australian sugarcane. **Molecular Breeding**, Dordrecht, v. 15, p. 367-381, 2005.

RIPOL, M.I.; CHURCHILL, G.A.; DA SILVA, J.A.G.; SORRELLS, M. Statistical aspects of genetic mapping in autopolyploids. **Gene**, Amsterdam, v. 235, p. 31-41, 1999.

ROACH, B.T. Nobilisation of sugarcane. **Proceedings International Society of Sugarcane Technologists**, New Orleans, v. 14, p. 206-216, 1972.

ROSSI, M.; ARAUJO, P.G.; PAULET, F.; GARSMEUR, O.; DIAS, V.M.; CHEN, H. VAN SLUYS, M.–A.; D´HONT, A. Genomic distribuition and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. **Molecular Genetics and Genomics**, Heidelberg, v. 269, n. 3, p. 406-419, 2003.

ROZEN, S.; SKALETSKY, H.J. **Primer3**: pick primers from a DNA sequence. 2000. Disponível em: <http://frodo.wi.mit.edu/primer3/input.htm>. Acesso em: 25 nov. 2008.

RUTHERFORD, R.S.; THOKOANE, L.N. cDNA-AFLP differential display of sugarcane (*Saccharum* spp. hybrids) genes induced by challenge with the fungal pathogen *Ustilago scitaminea* (sugarcane smut). **Proceedings of the South African Sugar Technologists Association**, Durban, v. 74, p. 104-107, 2001.

SAMBROOK, J.; RUSSELL, D.W. **Molecular cloning**: a laboratory manual. New York: Cold Spring Harbor Laboratory Press, 2001. 999p.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of the United States of America**, Washington, v.74, p. 5463-5467, 1977.

SANGUINO, A. Principais moléstias da cana-de-açúcar. In: PARANHOS, S.B. (Ed.). **Cana-de-açúcar**: cultivo e utilização. Campinas: Fundação Cargill, 1987. p. 741-757.

SANTIAGO, R.; DE ARMAS, R.; FONTANIELLA, B.; VICENTE, C.; LEGAZ, M-E. Changes in soluble and cell wall-bound hydroxycinnamic and hydroxybenzoic acids in sugarcane cultivars inoculated with *Sporisorium scitamineum* sporidia. **European Journal of Plant Pathology**, Dordrecht, v. 124, p. 439-450, 2009.

SANTOS, A.S. Doenças causadas por fungos e bactérias em cana-de-açúcar. In: IX REUNIÃO ITINERANTE DE FITOSSANIDADE DO INSTITUTO BIOLÓGICO, 9., 2003, Catanduva. **Anais**... Catanduva: RIFIB, 2003. p. 11-17.

SCARPARI, M.S.; BEAUCLAIR, E.G.F. Anatomia e botânica. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A (Ed). **Cana-de-açúcar**. Campinas: Instituto Agronômico, 2008. p. 47-56.

SHAPIRO, S.S.; WILK, M.B. An analysis of variance test for normality (Complete Samples). **Biometrika,** London, v. 42, n. 3/4, p. 591-611, 1965.

SILLS, G.R.; BRIDGES, W.; AL-JANABI, S.M.; SOBRAL, B.W.S. Genetic analysis of agronomic traits in a cross between sugarcane (*Saccharum officinarum* L.) and its presumed progenitor (*S. robustum* Brandes & Jesw. ex Grassl). **Molecular Breeding**, Berlin, v. 1, p. 355-363, 1995.

SKINNER, J.C. application of quantitative genetics to breeding of vegetatively reproduced crops. **Journal Australian Institute of Agricultural Science**, Curtin, v. 47, p. 82-83, 1981.

SREENIVASAN, T.V.; AHLOOWALIA, B.S.; HEINZ, D.J. Cytogenetics. In: HEINZ, D.J. **Sugarcane improvement through breeding**. Amsterdam: Elsevier, 1987. p. 211-253.

STIRNBERG, P.; FURNER, I.J.; OTTOLINE-LEYSER, H.M. MAX2 participates in an SCF complex which acts locally at the node to suppress shoot branching. **Plant Journal,** Malden, v. 50, p. 80–94, 2007.

STUBER, C.W.; EDWARDS, M.D.; WENDEL, J.F. Molecular-marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. **Crop Science**, Madison, v. 27, p. 639-648, 1987.

SUNDAR, A.R.; BARNABAS, E.L.; MALATHI, P.; VISWANATHAN, R. A Mini-Review on Smut Disease of Sugarcane Caused by *Sporisorium scitamineum*. In: MWORIA, J.K. **Botany**. Rijeka: InTech, 2012. cap. 5, p. 107-128.

TAKAYAMA, S; ISOGAI, A. Self-incompatibility in plants. **Annual Review of Plant Biology**, Palo Alto, *v.* 56, p. 467–489, 2005.

TANKSLEY, S.D. Mapping polygenes. **Annual Review of Genetics**, Palo Alto, v. 27, p. 205-234, 1993.

TARCHINI, R.; BIDDLE, P.; WINELAND, R.; TINGEY, S.; RAFALSKI, A. The Complete Sequence of 340 kb of DNA around the Rice *Adh1–Adh2* Region Reveals Interrupted Colinearity with Maize Chromosome 4. **The Plant Cell**, Waterbury, v. 12, p. 381–391, 2000.

THOKOANE, L.N.; RUTHERFORD, R.S. cDNA-AFLP differential display of sugarcane (*Saccharum* spp. Hybrids) genes induced by challenge with the fungal pathogen *Ustilago scitaminea* (Sugarcane smut). **Proceedings of the South African Sugar Technologists Association**, Durban, v.74, p. 104-107, 2001.

TOKESHI, H.; RAGO, A.M. Doenças da cana-de-açúcar. In: KIMATI, H.; AMORIM, L.; REZENDE, J.A.M.; BERGAMIN FILHO, A.; CAMARGO, L.E.A. (Ed.). **Manual de Fitopatologia**: Doenças das plantas cultivadas. São Paulo: Ceres, 2005. v. 2, cap. 21, p.185-196.

TUSKAN, G.A.; DIFAZIO, S.; JANSSON, S.; BOHLMANN, J.; GRIGORIEV, I.; HELLSTEN, U.; PUTNAM, N.; RALPH, S.; ROMBAUTS, S.; SALAMOV, A.; SCHEIN, J.; STERCK, L.; AERTS, A.; BHALERAO, R.R.; BHALERAO, R.P.; BLAUDEZ, D.; BOERJAN, W.; BRUN, A.; BRUNNER, A.; BUSOV, V.; CAMPBELL, M.; CARLSON, J.; CHALOT, M.; CHAPMAN, J.; CHEN, G.L.; COOPER, D.; COUTINHO, P.M.; COUTURIER, J.; COVERT, S.; CRONK, Q.; CUNNINGHAM, R.; DAVIS, J.; DEGROEVE, S.; DÉJARDIN, A.; DEPAMPHILIS, C.; DETTER, J.; DIRKS, B.; DUBCHAK, I.; DUPLESSIS, S.; EHLTING, J.; ELLIS, B.; GENDLER, K.; GOODSTEIN, D.; GRIBSKOV, M.; GRIMWOOD, J.; GROOVER, A.; GUNTER, L.; HAMBERGER, B.; HEINZE, B.; HELARIUTTA, Y.; HENRISSAT, B.; HOLLIGAN, D.; HOLT, R.; HUANG, W.; ISLAM-FARIDI, N.; JONES, S.; JONES-RHOADES, M.; JORGENSEN, R.; JOSHI, C.; KANGASJÄRVI, J.; KARLSSON, J.; KELLEHER, C.; KIRKPATRICK, R.; KIRST, M.; KOHLER, A.; KALLURI, U.; LARIMER, F.; LEEBENS-MACK, J.; LEPLÉ, J.C.; LOCASCIO, P.; LOU, Y.; LUCAS, S.; MARTIN, F.; MONTANINI, B.; NAPOLI, C.; NELSON, D.R.; NELSON, C.; NIEMINEN, K.; NILSSON, O.; PEREDA, V.; PETER, G.; PHILIPPE, R.; PILATE, G.; POLIAKOV, A.; RAZUMOVSKAYA, J.; RICHARDSON, P.; RINALDI, C.; RITLAND, K.; ROUZÉ, P.; RYABOY, D.; SCHMUTZ, J.; SCHRADER, J.; SEGERMAN, B.; SHIN, H.; SIDDIQUI, A.; STERKY, F.; TERRY, A.; TSAI, C.J.; UBERBACHER, E.; UNNEBERG, P.; VAHALA, J.; WALL, K.; WESSLER, S.; YANG ,G.; YIN, T.; DOUGLAS, C.; MARRA, M.; SANDBERG, G.; VAN DE PEER, Y.; ROKHSAR, D. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). **Science**, Washington DC, v. 313, n. 5793, p. 1596-1604, 2006.

UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR. Disponível em: <http://www.unica.com.br/noticias/show.asp?nwsCode=BFBD146D-7EFF-4CF3-9DD7-1F35061BE779>. Acesso em: 04 maio 2012.

VAN DER LINDEN, C.G.; WOUTERS, D.C.A.E.; MIHALKA, V.; KOCHIEVA, E.Z.; SMULDERS, M.J.M.; VOSMAN, B. Efficient targeting of plant disease resistance loci using NBS Profiling. **Theoretical Applied Genetics**, Berlin, v. 209, p. 384-393, 2004.

VAN OOIJEN, J.W.; VOORRIPS, R.E. **JoinMap 3.0:** Software for the calculations of genetic linkage maps. Wageningen: Plant Research International. 2001. 1 CD ROM.

VETTORE, A.L.; DA SILVA, F.R.; KEMPER, E.L.; ARRUDA, P. The libraries that made SUCEST. **Genetics and Molecular Biology**, Ribeirão Preto, v. 24, p. 1-7, 2001.

VETTORE, A.L.; DA SILVA, F.R.; KEMPER, E.L.; SOUZA, G.M.; FERRO, T.; HENRIQUE-SILVA, F.; GICLIOTI, E.A.; LEMOS, M.V.F.; COUTINHO, L.L.; NOBREGA, M.P.; CARRER, H.; FRANÇA, S.C.; BACCI Jr, M.; GOLDMAN, M.H.S.; GOMES, S.L.; NUNES, L.R.; CAMARGO, L.E.A.; SIQUEIRA, W.J.; VAN SLUYS, M.A.; THIEMANN, O.H.; KURAMAE, E.E.; SANTELLI, R.V.; MARINO, C.L.; TARGON, M.L.P.N.; FERRO, J.A.; SILVEIRA, H.C.S.; MARINI, D.C.; LEMOS, E.G.M.; MONTEIRO VITORELLO, C.B.; TAMBOR, J.H.M.; CARRARO, D.M.; ROBERTO, P.G.; MARTINS, V.G.; GOLDMAN, G.H.; OLIVEIRA, R.C.; TRUFFI, D.; COLOMBO, C.A.; ROSSI, M.; ARAÚJO, P.G.; SCULACCIO, S.A.; ANGELLA, A.; LIMA, M.M.A.; ROSA, V.E.; SIVIERO, F.; COSCRATO, V.E.; MACHADO, M.A.; GRIVET, L.; MAURO, S.M.Z.; NOBREGA, F.G.; MENCK, C.S.F.M.; BRAGA, M.D.V.; TELLES, G.P.; CARA, F.A.A.; PEDROSA, G.; MEIDANIS, J.; ARRUDA, P. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome Research**, London, v. 13, n. 12, p. 2725-2735, 2003.

VOORRIPS, R.E. MapChart: software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, Washington, v. 93, p. 77-78, 2002.

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, n. 21, p. 4407-4414, 1995.

WALKER, D.I.T. Heritability of smut resistance. **ISSCT Sugar Cane Breeding**, [s.l], v. 43, p. 19-23, 1980.

WALLER, J.M. Sugarcane smut (*Ustilago scitaminea*) in Kenia. II. Infection and resistance. **Transactions British Mycological Society**, London, v. 54, p. 405-414, 1970.

WEI, F.; WING, R.A.; WISE, R.P. Genome dynamics and evolution of the Mla (powdery mildew) resistance locus in barley. **Plant Cell**, Rockville, v. 14, p. 1903–1917, 2002.

WEI, X.; JACKSON, P.A.; McINTYRE, C.L.; AITKEN, K.S.; CROFT, B. Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. **Theoretical and Applied Genetics**, New York, v. 114, p. 155-164, 2006.

WHITTLE, A.M.; WALKER, D.I.T. Interpretation of sugarcane smut susceptibility trials. **Tropical Pest Management**, London, v.28, p. 228-237, 1982.

WILLIAMS, J. G. K.; KUBELIK, A. R.; LIVAK, K. J.; RAFALSKI, J. A.; TINGEY, S. V. DNA polymorphisms amplified by arbitrary primers are useful as genetic-markers. **Nucleic Acids Research,** Oxford, v. 18, p. 6531-6535, 1990.

WOO, H.R.; CHUNG, K.M.; PARK, J.H.; OH, S.A.; AHN, T.; HONG, S.H.; JANG, S.K.; NAM, H.G. ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. **Plant Cell**, Rockville, v. 13, p. 1779–1790, 2001.

WU, K.; TIAN, L.; MALIK, K.; BROWN, D.; MIKI, B. Functional analysis of HD2 histone deacetylase homologues in *Arabidopsis thaliana*. **Plant Journal**, Malden, v. 22, p. 19-27, 2000.

WU, K.K.; BURNQUIST, W.; SORRELLS, M.E.; TEW, T.L.; MOORE, P.H. The detection and estimation of linkage in polyploids using single-dose restriction fragments. **Theoretical Applied Genetics**, Berlin, v. 83, p. 294-300, 1992.

WU, K.K.; HEINZ, D.J.; HOGARTH, D.M. Association and heritability of sugarcane smut resistance to races A and B in Hawaii. **Theoretical Population Biology**, New York, v 75, p. 754-760, 1988.

WU, K.K.; HEINZ, D.J.; MAYER, H.K. Heritability of sugarcane smut resistance and correlation between smut grade and yield components. **Crop Science**, Malden, v.23, p. 54-56, 1983.

WU, R.; MA, C.X.; WU, S.S.; ZENG, Z.B. Linkage mapping of sex-specific differences. **Genetical Research,** London, v. 79, p. 85-96, 2002b.

WU, R.; MA, C.X.; PAINTER, I.; ZENG, Z.B. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. **Theoretical Population Biology**, New York, v. 61. p. 349-363, 2002a.

YANG, Z-N.; YE, X-R.; MOLINA, J.; ROOSE, M.L.; MIRKOV, T.E. Sequence analysis of a 282-kilobase region surrounding the citrus tristeza vírus resistance gene (*Ctv*) locus in Poncirus tifoliata L. Raf. **Plant Physiology**, Rockville, v. 131, p. 482-492, 2003.

YOUNG, N.D. QTL mapping and quantitative disease resistance in plants. **Annual Review of Phytopathology**, Palo Alto, v. 34, p. 479–501, 1996.

YOU-XIONG, Q.; JIAN-WEI, L.; XIAN-XIAN, S.; LI-PING, X.; RU-KAI, C. Differential Gene Expression in Sugarcane in Response to Challenge by Fungal Pathogen Ustilago scitaminea Revealed by cDNA-AFLP. **Journal of Biomedicine and Biotechnology**, New York, v. 2011, p. 1-10, 2011.

ZENG, Z.B. Precision mapping of quantitative trait loci. **Genetics**, Bethesda, v. 136, p. 1457-1468, 1994.

ANEXOS

Anexo A – Detalhes dos locos microssatélites de cana-de-açúcar derivados de etiquetas de sequências expressas (ESTs)

(continua)

Código do marcador	Motivo	Primer forward (5' - 3')	Primer reverse (5' - 3')	PCR [5]	AT [6]	D [7]	Número de alelo e amplitude [8]	
CV06 [1]	$(AATT)_{13}$	TCTCAAGCTTCGCCAGCTA	TGGCTCGGCTGTAGGAATTA	2	60.0	F	3	188-230
CV11 [1]	$(GAA)_6$	TGGCATGTGTCATAGCCAAT	CCCCAACTGGGACTTTTACA	2	60.0	F	6	227-242
CV22 [2]	$(AGGG)_5$	CACTACTCGCCCCGATTTC	CGAGTGCTTCTCCATCTGC	2	64.0	F	8	140-166
CV23 [2]	$(GGAA)_7/(AGG)_6$	GAACTGCTCACTGGCTCCTC	GTAGAAGTCCGTCGCCGTAA	2	64.0	F	9	150-206
CV24 [1]	$(CCAA)_5/(CACCT)_4$	TCGGAGAAGTTGACCGAGTT	GGTTTAGAGTTGGGGCCTTC	2	60.0	F	7	187-205
CV29 [2, 3]	$(ATCT)_{14}$	TCGCGTCCACCAATGTAACC	GCGTGCATCGCTTGTGTCTT	2	64.0	F	10	85-133
CV37 [2, 3]	$(TTTC)_{15}$	GGATGGACGACGTGTCCTGG	ATAAAGTGGCCGCTTGGATTGA	2	64.0	F	6	117-155
CV38 [2, 3]	$(CTTTT)_{18}$	GAAGCAGGGGCCTCAAGTTG	GTCAAACAGGCGATCTGGCTC	2	64.0	F	9	109-199
CV46 [1]	$(GGTAA)_{11}$	TGTTCCAAGTTCATGCGCTCC	ATGCATGCAGGTTCAAAAGCAG	2	64.0	F	5	146-188
CV51 [2]	$(TGT)_{13}$	CTACCCCAACTTGCTTGGGAC	GACTGGAACAAAGACGGACTG	2	64.0	F	3	147-160
CV53 [1]	$(AAAAT)_5/(TTTAT)_6$	CCCCACCGTAGCTTGTGCAT	AAACGTGCACATGCTTGTATGC	2	64.0	F	7	160-183
CV58 [2]	$(ATAGAT)_{10}$	CGGGTAGTTAGGAGGAGATGG	GTCATCCATTTTGGAACGAATGG	2	64.0	F	6	153-195
CV78 [2]	$(CTGTG)_9$	ACGAGGCCACCATAGAACATG	GCAATTGGGAGGAGAGGAATG	2	64.0	F	9	144-203
CV79 [2]	$(CTATAT)_{11}/(TATAGA)_6$	GGCACTGCTGGTGGTTGATTG	TCCCACATCAAGAGGCAGCTA	2	64.0	F	7	136-197
CV86 [2]	$(AATT)_8$	CCTCAGCAGCCCAAAGTCCT	GTCGGAATCAGCCGGATTAGC	2	64.0	F	5	159-187
CV91 [1]	$(GCC)_6/(GCA)_6$	AAAGGAAATCGCCCTCCGTCT	CCGATGATGAGCCAGCAATCC	2	64.0	F	8	175-197
CV94 [2]	$(AAAAAG)_5/(CGT)_5$	GGCAGGCCAAGATGAATGAAG	AGCACAGCGGAGGGTACGG	2	64.0	F	4	187-205
CV100 [1]	$(GAG)_{13}$	CTGTTGAGGAGCCGGATGAG	CTCTTCCGATGGCTCGGTCT	2	64.0	F	9	222-256
CV101 [1]	$(ATC)_{23}$	GTCGTCGTCGTCACGATCATC	AGTTGACGGCATGGTTCTTGC	2	64.0	F	11	111-180
CV104 [2]	$(TCCTG)_5$	GATTTTCGACTGTGCGCTTGG	AAGTTCTCTGCCGGAGCAAAC	2	64.0	F	6	133-158
CV106 [2]	$(GGC)_8$	AAACAGAGCATACTCGAGGCC	ACGTTGCTGACGAGGTTTTCC	2	64.0	F	6	146-161
CV115 [1]	$(TCACAG)_{10}/(GTA)_6/(AGA)_5$	GTCCATGTCCATCCATGATCC	GGAGCTCCGTCTTCTTGTTAC	2	60.0	F	6	150-174
CV119 [1]	$(AAAAC)_7$	TATCTCTCCTTGGTTTGGATGG	CACCCTACCAAATACCACAACA	2	64.0	F	5	121-175

Anexo A – Detalhes dos locos microssatélites de cana-de-açúcar derivados de etiquetas de sequências expressas (ESTs)

(continuação)

Código do marcador	Motivo	Primer forward (5' - 3')	Primer reverse (5' - 3')	PCR [5]	AT [6]	D [7]	Número de alelo e amplitude [8]	
CV128 [2]	$(GCA)_{13}$	AGGGCAACGGAGTCTTCGAC	CTGAACTCCGATGTGCTGGTG	2	60.0	F	5	147-168
CV135 [2]	$(AAG)_{16}$	AGCAAAACCAGCCTTCCCTTC	CTGTTTGTTTCTGCTTGCTTGC	2	64.0	F	6	129-159
CV144 [1]	$(TCTCCG)_5$	GCGCCTCCGTGGATAAGAATC	CCTTCCCCTACAGCGCCTAC	2	64.0	F	5	146-164
ESTA26 [4]	$(TG)_{11}$	GGCAGCCCCACATCTTCCT	GGGCACAAGCATCCGAACC	1	56.0	S	4	172-186
ESTA48 [4]	$(CA)_8$	AGCAACTCCGGCCTCTCCTG	CTTTCTGTTTTGCTCCTCCGTCTG	1	62.7	S	10	233-295
ESTA53 [4]	$(TG)_8$	TGGAAATGGCAGCTGGTCTCGT	ATGCACGTACCAGAGGGAGATTTG	1	58.9	S	9	168-192
ESTA61 [4]	$(AT)_{12}$	ACCTCAGTCTCCTCCTCAACC	TATACTACACATGCACAGGCTACG	1	56.4	S	5	236-246
ESTB14 [4]	$(CGT)_8$	TGAGGGAATGAATGGACTGG	CCACCACCACCATACCTGTC	1	52.0	S	9	285-315
ESTB55 [4]	$(CCA)_5$	CTTCTTGGCCTTGGCGTTACTGA	GCTAGCTGGCCCCATTTCCTCT	1	60.0	S	3	118-124
ESTB60 [4]	$(TTG)_{10}$	AGCCGCAATGAATCCAACTG	CTCTAGCTCCGACGATGATACCTC	1	61.0	S	8	157-206
ESTB82 [4]	$(CGT)_9$	CGTCGATCGAGATGAAGAAGG	GAAGCAGTCGTGGAAGTGGAG	1	62.7	S	5	245-263
ESTB94 [4]	$(CTT)_9$	GAGGCAGCCAGGCAGGTCAC	GGTGGCAGTGTTCAGGCAGATG	1	61.0	S	10	210-279
ESTB99 [4]	$(TCG)_5$	GAGGTCCTTCTTGTAGTTGTATGC	GTGCCGGAGGATTTGATG	1	64.7	S	4	215-224
ESTB100 [4]	$(TCG)_6$	CCACGGGCGAGGACGAGTA	GGGTCCTTCTTCGCCTCGTG	1	64.7	S	13	240-278
ESTB118 [4]	$(TTC)_6$	CTTGGCTAGGGTTTCTTGAGTCGT	CATGGCTTTTGGCTTGCTTCT	1	61.0	S	5	106-163
ESTB128 [4]	(ACC)6	GCAGGCGCTCATCTACAAGT	GGAGCACCGCCACTTCTT	1	55.0	S	5	150-171
ESTC80 [4]	$(ATTC)_3$	ATTCTTTCTCCCCCTGTTGTGC	GTCGCCAGATCGCTTTCGTT	1	58.9	S	7	188-292
SCB189 [5]	$(TCA)_{10}$	GTAAGGAAGAAGCAACAAACAACAG	GATTCGATGCAACTCTCCTGTAAA	1	60.0	S	5	261-280
SCB191 [5]	$(GCT)_5$	GCGCCATCAGGGAAGCCAAAAC	GCGCGTGCGAGCAGATGAAC	1	60.0	S	5	213-226
SCB219 [5]	$(CTT)_6$	CTCCATGGCCATCGCTGCTCTG	TTCTTTGGTGGTTGGTGGCTGTGG	1	62.0	S	5	146-158
SCB225 [5]	$(GCA)_6$	GCCTAATGCCATGCCCCAGAG	AGAACCGAACCTGAACTCCGATGTG	1	56.0	S	5	209-230
SCB227 [5]	$(GAT)_9$	GGGGGAGGACGACGACGAGGAC	CACCCTCGCCATCATCCTCCATTT	1	64.0	S	6	200-215
SCB243 [5]	$(TAC)_6$	ATATGGAGCTCCGTCTTCTTGTTA	CCGTAGCTGGGGGTTGAGA	1	63.0	S	4	179-194
SCB246 [5]	$(GAT)_6$	CACCAGAAACCGATACAACAAAGAC	AGCTCATCAATTCGTCCTATCACAC	1	61.0	S	9	140-191
SCB252 [5]	$(GCT)_5$	CCCTTTTTGCTGCTTTCTCACTTT	AGTCTCCTCCCGCTGTGATTG	1	50.0	S	4	192-201

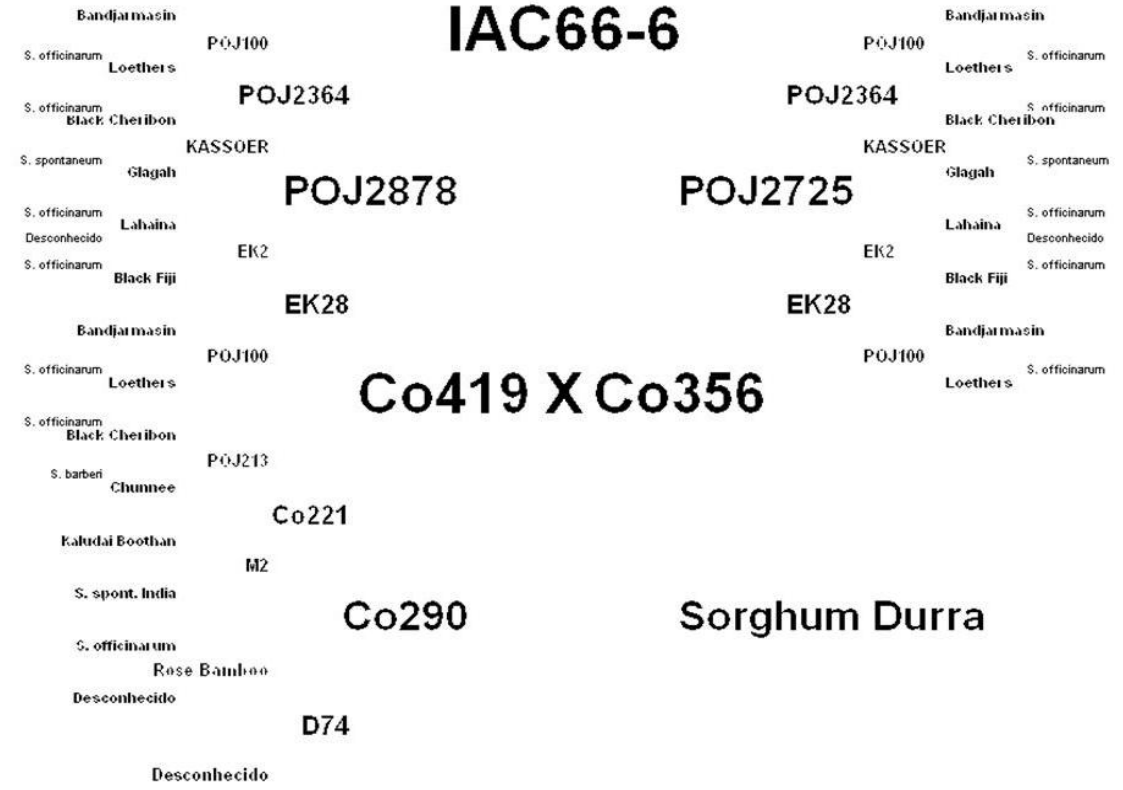
Anexo A – Detalhes dos locos microssatélites de cana-de-açúcar derivados de etiquetas de sequências expressas (ESTs)

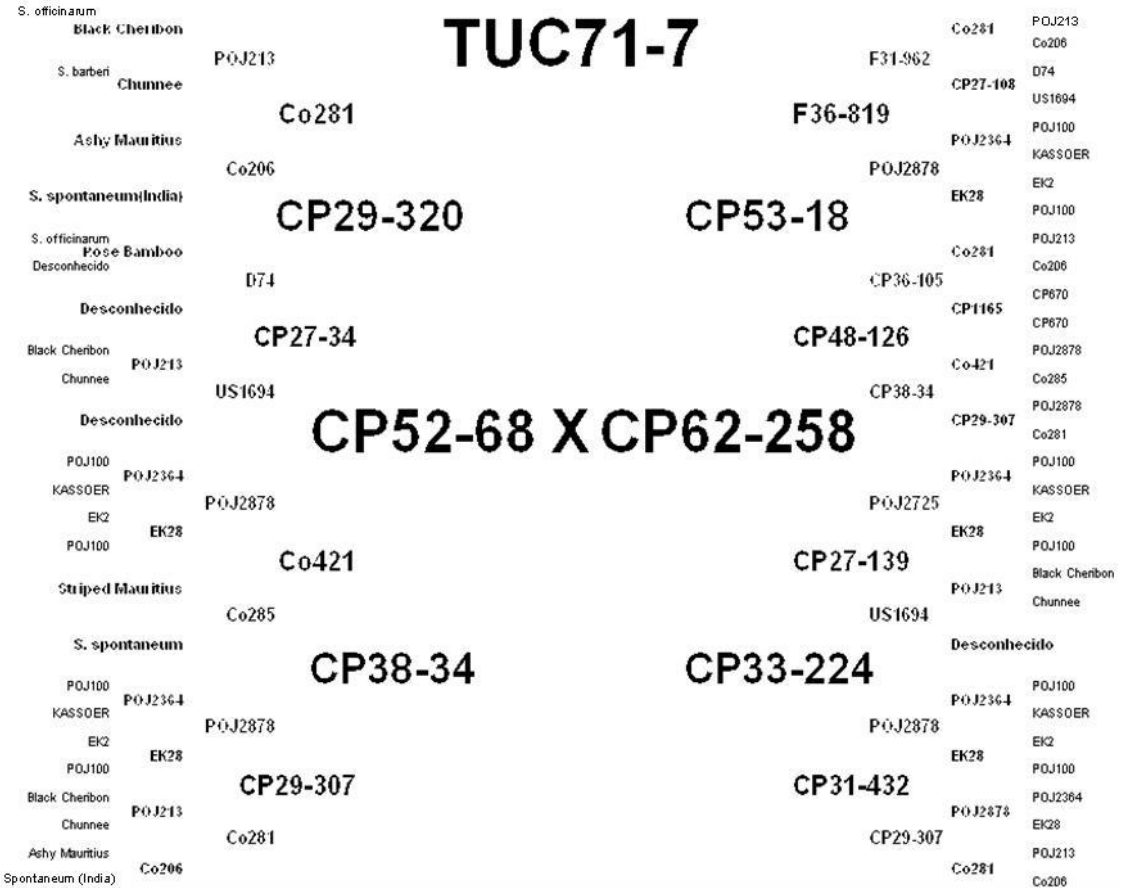
(conclusão)

Código do marcador	Motivo	Primer forward (5' - 3')	Primer reverse (5' - 3')	PCR [5]	AT [6]	D [7]	Número de alelo e amplitude [8]	
SCB272 [5]	$(GTG)_5$	GTGCAAAGACGAGGATGAGAA	ACTGCCGCGTCAACCAC	1	63.0	S	3	143-155
SCB312 [5]	$(TCC)_7$	AGTCCGTCGCCGTAATCATCTTG	GCACCTCCTCCTTTCCTTCCTTATT	1	65.0	S	9	213-264
SCB329 [5]	$(CCG)_8$	CGCCACCGGAAAAACC	CAGACTGCAAGAAAGGAACCA	1	52.0	S	6	227-248
SCB344 [5]	$(ATG)_{13}$	ATGCGGCTCGATATCTTGTTGG	CTCCTCCTCCTTCTTATTCACCTCA	1	61.0	S	7	123-153
SCB365 [5]	$(CAG)_5$	GTACGGAGCGCAGTGTAGGT	CACCGCCTTGTACTCCAAAT	1	61.0	S	7	223-274
SCB381 [5]	$(TAC)_8$	TGGAGCTCCGTCTTCTTGTT	GCTAGCCCGTACATTGGGTA	1	61.0	S	8	207-234
SCB397 [5]	$(GTG)_8$	CTGTGGGACTACTCGCCTTC	GGAGGAGGCTGTGATTAGCA	1	61.0	S	7	219-237
SCB422 [5]	$(GAT)_7$	GCCTGCTATTTGCTCGATTC	TTTGGAGCACCCTCTCAATC	1	61.0	S	6	216-249
SCB437 [5]	$(CGG)_{12}$	ATAGAACTTCCCCTCTGACC	GACTGTTACGGGGTCTCTTC	1	50.0	S	4	439-448
SCB442 [5]	$(AGG)_5$	ACACACACCTCGTCTTGCAG	CCATGTTTGGGCTATCACCT	1	61.0	S	6	160-288
SCB477 [5]	$(GCC)_{10}$	CCCCCTGCCTCCTACAGT	GTGGTGGGTGTCAGTTTGC	1	61.0	S	4	313-328
SCB502 [5]	$(CTT)_5$	GGAGGAAGCAGGCAGGAGAGGG	AGGGGCAGGATGAGCGAGGAAG	1	64.0	S	4	201-216

[1] Desenvolvido por CanaVialis/Monsanto (dados não publicados); [2] Desenvolvido por Jordão Júnior (2009); [3] Publicado por Maccheroni et al. (2009); [4] Desenvolvido por Oliveira et al. (2009); [5] Desenvolvido por Marconi et al. (2011); [5] Programa da PCR como descrito em Material e Métodos; [6] Temperatura de anelamento na reação de amplificação; [7] D: Detecção dos alelos EST-SSR através das metodologias: eletroforese em gel de poliacrilamida 5% corado com nitrato de prata (S) ou eletroforese capilar baseado em fluorescência (F); [8] Número de alelos observados por loco e sua amplitude em pares de base (bp).

Anexo B - Genealogia dos genitores ‘IAC66-6’ e ‘TUC71-7’





Anexo C – Valores de AUDPC (BLUPs) para os genitores e os indivíduos da progênie, considerando os diferentes anos de condução dos experimentos

Genót.	BLUPs		Genót.	BLUPs		Genót.	BLUPs		Genót.	BLUPs		Genót.	BLUPs	
	2009	2010		2009	2010		2009	2010		2009	2010		2009	2010
'IAC66-6'	4,26	3,17	41	4,25	2,28	83	8,28	3,70	131	5,46	3,05	173	7,47	3,96
'TUC71-7'	1,51	2,56	42	4,27	2,40	84	9,91	4,18	132	7,41	4,04	174	2,63	2,70
1	4,01	3,20	43	3,17	2,76	85	3,92	2,88	133	4,51	3,04	175	3,69	3,70
2	3,90	2,73	44	4,73	3,21	86	4,33	3,00	134	6,52	2,96	176	0,20	1,26
3	7,59	3,37	45	3,98	3,00	87	3,79	2,86	135	5,72	3,71	177	4,34	2,81
4	4,58	1,55	46	5,55	3,62	88	2,69	2,50	136	3,61	2,73	178	3,64	2,51
5	4,53	2,08	47	9,71	4,70	89	6,79	3,52	137	-	-	179	2,50	2,26
6	5,03	3,23	48	7,24	3,67	90	3,17	2,56	138	3,99	2,86	180	0,44	1,29
7	3,36	2,30	49	5,67	2,63	91	4,49	3,04	139	6,00	3,51	181	3,52	2,81
8	6,10	3,21	50	6,77	3,93	92	4,83	2,75	140	5,93	3,08	182	5,97	3,63
9	2,71	2,66	51	4,30	2,46	93	2,60	2,53	141	3,80	3,11	183	6,29	2,71
10	6,58	4,53	52	7,59	3,64	94	4,81	2,71	142	3,85	2,99	184	3,98	2,69
97	5,14	2,70	53	4,44	3,29	101	5,61	4,01	143	4,24	3,10	185	1,91	2,69
98	6,23	4,12	54	4,49	3,04	102	3,32	2,92	144	7,64	3,95	19	2,86	2,53
13	2,60	2,87	100	3,90	3,22	103	4,81	3,43	145	7,08	4,03	187	4,15	3,09
14	4,11	2,87	56	3,62	2,90	104	1,01	2,08	146	5,54	3,03	188	6,34	3,16
15	-	-	57	3,58	2,05	105	-	-	147	4,06	2,94	189	5,22	3,02
16	-	-	58	3,28	2,42	106	5,27	4,45	200	7,55	3,55	190	3,78	2,77
17	5,43	4,45	59	-	-	55	6,44	2,94	149	3,31	2,36	191	3,14	2,59
18	5,22	3,82	60	2,66	2,44	108	3,71	2,73	150	9,62	5,22	192	8,90	4,64
99	4,97	3,42	61	3,97	2,94	109	5,45	3,20	151	1,42	2,39	193	2,37	1,50
20	5,81	2,89	62	3,52	2,34	110	5,34	2,72	152	2,57	2,63	194	4,53	3,43
21	4,18	2,50	63	4,21	2,64	111	0,83	2,26	153	4,51	2,96			
22	5,60	3,57	64	7,19	3,77	112	4,86	3,09	154	3,72	3,05			
23	8,63	3,65	65	10,01	4,04	113	6,50	3,58	155	7,40	3,57			
24	5,98	3,13	66	4,25	2,99	114	7,05	3,58	156	4,19	2,70			
25	3,59	3,97	67	8,27	4,17	115	7,65	3,88	157	10,98	5,50			
26	4,34	3,00	68	6,36	2,81	116	5,04	3,18	158	4,63	3,97			
27	4,90	3,10	69	6,03	3,31	117	6,79	4,00	159	4,95	2,88			
28	1,65	2,63	70	5,19	3,14	118	5,59	3,27	160	3,73	2,50			
29	3,01	3,08	71	6,55	2,60	119	5,73	3,76	11	3,37	2,17			
30	6,60	3,66	72	5,39	3,23	120	9,67	3,60	162	4,15	2,59			
31	6,46	4,44	73	5,29	3,06	121	4,91	3,13	163	4,36	2,82			
32	3,83	2,67	74	5,25	4,53	122	3,00	3,18	164	4,67	3,30			
33	0,21	2,15	75	4,71	2,89	123	-	-	165	3,66	2,68			
34	5,38	3,36	76	0,13	2,04	124	6,78	3,83	12	6,12	3,38			
35	6,30	3,58	77	3,16	2,63	199	1,24	2,51	167	6,76	3,82			
36	5,03	3,56	78	5,11	2,96	126	5,22	2,99	168	5,57	3,56			
37	6,76	3,48	79	3,40	2,81	198	-	-	169	6,02	3,39			
38	4,69	2,85	80	8,32	3,43	128	3,69	2,87	170	-	1,56			
39	0,79	2,61	81	4,31	2,56	129	-	-	171	1,86	2,94			
40	5,61	3,84	82	-	-	195	3,61	2,97	172	4,36	2,89			

Anexo D - Artigo científico publicado no periódico *BMC Genetics*:

PALHARES, A.C.; RODRIGUES-MORAIS, T.B.; VAN SLUYS, M-A.; DOMINGUES, D.S.; MACCHERONI, JR.W.; JORDÃO JR. H.; MARCONI, T.G.; SOUZA, A.P.; MOLLINARI, M.; GAZAFFI, R.; GARCIA, A.A.F.; VIEIRA, L.M.C. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers, **BMC Genetics**, London, v. 13, p. 1-16, 2012. doi:10.1186/1471-2156-13-51.

Palhares et al. BMC Genetics 2012, 13:51
http://www.biomedcentral.com/1471-2156/13/51

BMC Genetics

RESEARCH ARTICLE Open Access

A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers

Alessandra C Palhares[1], Taislene B Rodrigues-Morais[1], Marie-Anne Van Sluys[2], Douglas S Domingues[2,3], Walter Maccheroni Jr[4,5], Hamilton Jordão Jr[4,5], Anete P Souza[6], Thiago G Marconi[6], Marcelo Mollinari[1], Rodrigo Gazaffi[1], Antonio Augusto F Garcia[1] and Maria Lucia Carneiro Vieira[1*]

Abstract

Background: The development of sugarcane as a sustainable crop has unlimited applications. The crop is one of the most economically viable for renewable energy production, and CO_2 balance. Linkage maps are valuable tools for understanding genetic and genomic organization, particularly in sugarcane due to its complex polyploid genome of multispecific origins. The overall objective of our study was to construct a novel sugarcane linkage map, compiling AFLP and EST-SSR markers, and to generate data on the distribution of markers anchored to sequences of *sclvana_1*, a complete sugarcane transposable element, and member of the *Copia* superfamily.

Results: The mapping population parents ('IAC66-6' and 'TUC71-7') contributed equally to polymorphisms, independent of marker type, and generated markers that were distributed into nearly the same number of co-segregation groups (or CGs). Bi-parentally inherited alleles provided the integration of 19 CGs. The marker number per CG ranged from two to 39. The total map length was 4,843.19 cM, with a marker density of 8.87 cM. Markers were assembled into 92 CGs that ranged in length from 1.14 to 404.72 cM, with an estimated average length of 52.64 cM. The greatest distance between two adjacent markers was 48.25 cM. The *sclvana_1*-based markers (56) were positioned on 21 CGs, but were not regularly distributed. Interestingly, the distance between adjacent *sclvana_1*-based markers was less than 5 cM, and was observed on five CGs, suggesting a clustered organization.

Conclusions: Results indicated the use of a NBS-profiling technique was efficient to develop retrotransposon-based markers in sugarcane. The simultaneous maximum-likelihood estimates of linkage and linkage phase based strategies confirmed the suitability of its approach to estimate linkage, and construct the linkage map. Interestingly, using our genetic data it was possible to calculate the number of retrotransposon *sclvana_1* (~60) copies in the sugarcane genome, confirming previously reported molecular results. In addition, this research possibly will have indirect implications in crop economics e.g., productivity enhancement via QTL studies, as the mapping population parents differ in response to an important fungal disease.

Keywords: *Saccharum* spp, AFLP, EST-SSR, Retrotransposon-based markers, Single-dose markers, Integrated genetic map, Marker distribution

Background

Sugarcane is a crop of unquestionable importance for tropical and subtropical regions of the world, where it occupied 24 million hectares in 2009 [1]. Sugarcane is a cost effective renewable resource, with alternative production in food, feed, fiber, and energy e.g. sugar, animal feeds, alcohols, and fertilizers. Brazil is one of the greatest producers and exporters of sugar and ethanol from sugarcane, where cane production reached approximately 625 million tons in 2010, and the sugarcane industry generated a gross annual income of approximately US$ 23 billion [2].

Sugarcane exhibits the most complex genome of any genetically bred crop. Selection based on scientific approaches began in 1888; the first hybridizations were conducted in Java and Barbados, between *Saccharum*

* Correspondence: mlcvieir@usp.br
[1]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900, Piracicaba, Brazil
Full list of author information is available at the end of the article

BioMed Central

© 2012 Palhares et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

officinarum and *S. spontaneum*. *S. officinarum*, known as 'noble' canes, are high in sucrose with juicy, thick stalks; low in fiber content, and susceptible to several diseases [3]. *S. spontaneum* is low in sucrose content, but robust, and resistant to abiotic stresses and pests [4]. One hundred years of interspecific hybrid backcrossings with *S. officinarum* (used as the maternal parent), a process called 'nobilization', lead breeders to obtain more productive varieties, with ratooning ability, and increased resistance to biotic and abiotic stresses [5,6]. Subsequently, all modern sugarcanes derive largely from intercrossing these canes, followed by intensive selection [7,8]. Therefore, currently grown cultivars are denoted *Saccharum* spp. Varieties are evaluated for rusticity, pest resistance, and high sugar yield prior to release, which requires 12 to 15 years.

The contemporary sugarcane cultivars have a large (10 Gb) and complex genome structure that is highly polyploid, aneuploid (2n = 100 to 130), and have multispecific origins [9] with a complete set of homo(eo)logous genes ranging from 8 to 10 alleles. Classic cytological works as well as fluorescent *in situ* hybridization determined that *S. officinarum* is an octaploid species (2n = 8*x* = 80) that experienced few aneuploid events, and the ploidy level of *S. spontaneum* varies from 5 to 16 (2n = 40 to 128) [10,11]. Genome *in situ* hybridization assays reveal that the 'R570' cultivar (2n = 115) shares 80% of its chromosomes with *S. officinarum*, 10% with *S. spontaneum*, and 10% are recombinant chromosomes [12]. These results clearly indicate that two chromosome sets coexist in the sugarcane genome. Vegetative propagation resulted in sugarcane clones with variable chromosome numbers cultivated in field plantations, and numerical and structural alterations continued to accumulate.

Linkage maps are valuable tools to elucidate genetic and genomic organization, particularly in sugarcane due to its increased ploidy levels. However, high inbreeding depression caused by endogamy limits the production of experimental mapping populations, such as F_2, BCs, RILs, and DH lines. The *S. spontaneum* 'SES 208' (2n = 64) linkage map published by Al-Janabi *et al.* [13] was the first map constructed directly from a complex polyploid species, based on single-dose markers (or SDMs) proposed by Wu *et al.* [14], which considers the use of simplex (single allele copy from one parent) to obtain the genetic map. Al-Janabi *et al.* [13] used progeny from a cross between 'SES 208', and a diploidized haploid derived from anther cultured 'SES 208' and 'ADP 85-0068' to estimate linkage. This strategy facilitated direct meiotic analysis in 'SES 208', and gametic segregation ratios to be observed. Results showed autopolyploid chromosomal behavior in 'SES 208', and the proportion of SDMs to higher dose markers (multiple alleles) fit the assumption of auto-octaploidy, with the absence of repulsion phase linkages. Subsequently, da Silva *et al.* [15] integrated the map of Al-Janabi *et al.* [13] with the simplex-based map of Sobral *et al.* [16]. SDMs linkage relationships were determined using MapMaker software [17].

Later, Grivet *et al.* [18] used selfed progeny to estimate linkage for the elite cultivar 'R570'. To date, self-fertilized sugarcane progeny are used to map simplex and duplex markers on co-segregation groups (CGs); for instance, Andru *et al.* [19] using the software JoinMap 3.0 [20] constructed a map for the cultivar Louisiana 'LCP 85-384'. Alternatively, crossing unrelated heterozygous genotypes generates a segregating sibling population (F_1), which can be valuable for constructing outcrossing species maps, including mapping in sugarcane. From the genetic configurations expected in a segregating F_1, and denoting gel band presence as *A*, and its absence as *O*, we can make use of the designation *AO x OO* (in a diploid species), or Simplex *x* Nuliplex (in sugarcane) to construct individual maps, one for each parent. This approach, known as the pseudo testcross strategy [21] uses two sets of dominant markers that segregate in a 1:1 ratio. It was applied to estimate linkage in interspecific crosses, where *S. officinarum* was the female ('Green German', 'IJ 76-514', 'La Striped'), and *S. spontaneum* the male parent ('IND 81-146' and 'SES 147B') [22-26]. The Australian cultivars 'MQ77-340', 'Q165', and 'MQ76-53' were mapped similarly [27-29]. Each of these studies determined a comparable number of linkage groups (LGs) and map lengths; e.g. 'La Striped' (2n = 80): 49 LGs, 1,732 cM; 'SES 147B' (2n = 64): 45 LGs, 1,491 cM [26].

Garcia *et al.* [30] constructed a single integrated genetic map based on simultaneous maximum-likelihood linkage estimates and linkage phase methodology [31], based on a population derived from a cross between two pre-commercial sugarcane cultivars ('SP80-180' *x* 'SP80-4966'). A total of 1,118 single-dose markers were identified; 39% were derived from a testcross configuration between parents segregating in a 1:1 ratio, and 61% segregated in 3:1 ratio, representing heterozygous loci in both parentals with identical genotypes. The final map was comprised of 357 linked markers, including RFLPs, SSRs, and AFLPs assigned to 131 CGs, with a LOD score of 5.0, recombination fraction of 37.5 cM. Authors indicated the simultaneous maximum-likelihood estimates of linkage, and linkage phases were appropriate to generate an integrated genetic map of sugarcane [30]. Then, to enhance existing map resolution, and identify putative functional polymorphic gene loci, Oliveira *et al.* [32] screened EST-SSRs and EST-RFLPs in the same mapping population. Markers analyzed in the previous map were added to 2,303 newly generated polymorphic

markers, including 1,669 (72.5%) SDMs; 664 (40%) were scattered on 192 CGs, with a total estimated length of 6,261 cM.

The current development of expressed sequence-based markers such as EST-SSRs, genic SNPs, and TRAPs enrich the genetic data that comprise linkage maps. In sugarcane, due to the necessity of mapping a large number of markers to guarantee a reasonable coverage of its genome with many chromosomes [33], a novel and potentially useful approach is to compile anonymous and putative functional markers. Earlier, the Brazilian Sugarcane Expressed Sequence Tag (or SUCEST) project [34,35] generated 237,954 high-quality ESTs organized into 43,141 putative unique sugarcane transcripts referred to as sugarcane assembled sequences. Based on SUCEST data, Rossi *et al.* [36] developed RFLPs using probes derived from NBS-LRR and LRR conserved domains, and S-T Kinase type resistance genes, and positioned them on the 'R570' map. Besides, Rossi *et al.* [37] conducted a transposable element (TE) search, revealing a surprising high number of expressed TE homologues, and found all major transposon families were represented in sugarcane. *Mutator* and *Hopscotch* were later reported as the most represented TE families in the sugarcane transcriptome [38]. The SUCEST database was used to describe two LTR retrotransposon families, which were denoted as *scIvana_1* and *scAle_1* [39]. Both were reported as complete elements, and different members of the *Copia* superfamily. The *scIvana_1* shows low copy numbers (40 to 50) and diversity among copies, and is expressed under specific conditions in low-differentiated tissues; *scAle_1* exhibits high copy numbers in the sugarcane genome (> 1000), is more diversified compared to *scIvana_1*, and active under varied physiological conditions.

Retrotransposon-based markers have been developed using several approaches. In Poaceae species, for example, SSAP was first used to study the distribution of *BARE-1*-like retrotransposable elements in barley genome [40]. In brief, SSAP uses two restriction enzymes to generate a large number of DNA fragments; after that, a retrotransposon-anchored PCR is used to perform a selective amplification. In the 90's, two new techniques were developed to exploit the polymorphism generated by *BARE-1* genome integration, named REMAP and IRAP [41]. Patterns indicate that although the *BARE-1* family of retrotransposons is dispersed, these elements are clustered or nested locally, and often found near microsatellite sequences. Later, both procedures were reported as useful to screen insertional polymorphisms in populations of *Spartina anglica*, an allopolyploid involved in natural and artificial invasions [42]. These methods are dominant and multiplex, and generate anonymous marker bands. IRAP is based on the PCR amplification of genomic DNA fragments which lie between two retrotransposon insertion sites, and REMAP is based on the amplification of fragments which lie between a retrotransposon insertion site and a microsatellite site. Polymorphism is detected by the presence or absence of the PCR product in both techniques. Lack of amplification indicates the absence of the retrotransposon at the particular locus. In contrast, the RBIP [43] and ISBP [44] score individual loci and are used to search for insertional polymorphisms. RBIP, for example, was used to address the issue of evolution of rice varieties [45] and ISBP was used to analyze diversity in wheat [46]. The RBIP method exploits knowledge of the sequence flanking a TE to design the primers while the ISBP method uses one primer in the element and the other in the flanking DNA sequence.

The overall objective of our study was to generate data on *scIvana_1* and *scAle_1*-based marker distribution to a novel sugarcane linkage map based on a compilation of AFLPs and EST-SSRs.

Results

Genotyping and segregation analyses

Excellent AFLPs, EST-SSRs, and *scIvana_1*-based marker banding profiles were obtained [see Additional file 1], despite the size and complexity of the sugarcane genome. The 72 enzyme-selective primer combinations tested provided a range of AFLP band numbers per gel (44 to 174), and polymorphic loci (four to 33). Subsequently, 22 combinations were selected as optimal for genotyping the F_1 population [Additional file 2]. The combinations generated 102 to 172 AFLP bands per gel, and 19 to 48 polymorphic loci, which revealed 22.1% (685/3,094) segregating loci. Among the segregating loci, 71.2% (488/685) segregated in only one parent, and 28.8% (197/685) segregated in both parents (Table 1). The 'IAC66-6' clone and 'TUC71-7' variety contributed a respective 52% (254/488) and 48% (234/488) of the loci that segregated in only one parent. The average number of amplified bands and segregating loci per enzyme-primer combination were 140.6 (3,094/22) and 31.1 (685/22), respectively.

From the 184 EST-SSR loci initially investigated, 22.3% (41) were selected for genotyping [Additional file 2]. These loci revealed 273 alleles with an average of 6.7 alleles per locus; 80.6% (220) segregated in the F_1 population, 68.6% (151/220) segregating in only one parent, and 31.4% (69/220) in both parents (Table 1). The 'IAC66-6' clone and 'TUC71-7' variety contributed a respective 43% (65/151) and 57% (86/151) of the alleles that segregated in only one parent.

Among the 16 restriction enzyme-primer combinations used to amplify the retrotransposon *scIvana_1*-based markers, six were selected for genotyping the F_1 population.

Table 1 Marker polymorphisms used for mapping, and distribution of the different markers according to the cross type (D1, D2 and C)

Marker type	AFLP	EST-SSR	*scIvana_1*	Total
Number of scorable bands (evaluated in total)[a]	3,094	273	357	3,724
Number of segregating markers (genotyped)	685	220	87	992
Number of polymorphic markers between parents	488	151	74	713
Number of monomorphic markers between parents	197	69	13	279
Number of single dose markers (SDMs)[b]	535	130	65	730
SDMs of origin from 'IAC66-6' [D1][c]	197	41	23	261
SDMs of origin from 'TUC71-7' [D2][c]	192	60	32	284
SDMs of origin from both parents [C][d]	146	29	10	185
Total number of linked markers on the map	395	95	56	546
Number of linked markers of origin from 'IAC66-6' [D1]	154	33	21	208
Number of linked markers of origin from 'TUC71-7' [D2]	164	46	30	240
Number of linked markers of origin from both parents [C]	77	16	5	98

[a] AFLPs generated from 22 enzyme-selective primer combinations, EST-SSR alleles generated from 41 loci and *scIvana_1*-based markers generated from 6 restriction enzyme-primer combinations.
[b] Data obtained after Bonferroni's correction.
[c] Markers present in only one parent with a 1:1 segregation ratio in the mapping population.
[d] Markers present in both parents with a 3:1 segregation ratio in the mapping population.

These combinations revealed 357 loci; 24.4% (87) behaved as segregating loci. Among them, 85.1% (74/87) and 14.9% (13/87) segregated in only one parent and in both parents, respectively (Table 1). The average number of amplified bands and segregating loci per enzyme-primer combination were 59.5 (357/6) and 14.5 (87/6), respectively [Additional file 2]. The male and female parents contributed a respective 44.6% (33/74) and 55.4% (41/74) of loci that segregated in only one parent. Sixteen retrotransposon *scAle_1* combinations revealed gel profiles; however, amplicon absence, or amplifications associated with non-specific polymorphism prevented profile use in genotyping.

AFLP and *scIvana_1*-based loci exhibited similar levels of segregating alleles in the mapping population (~25%). Notably, both techniques reveal polymorphisms at restriction enzyme cleavage sites. On the other hand, EST-SSR loci showed high levels of segregating alleles in the mapping population (~80%). The polymorphisms observed at EST-SSR loci are due to differences in the size of multiple alleles; we cannot predict if any allele will be fixed in a sugarcane cultivar, which is expected to be highly heterozygous. Both parents contributed equally to polymorphisms i.e. 49.4% ± 3.7% and 50.6% ± 3.8% of the amplicons derived from the male ('IAC66-6') and female ('TUC71-7') parent, respectively, independent of marker type.

It is important to clarify that we organized our segregation data assuming that homo(eo)logous chromosomes paired faithfully during meiosis, leading to regular bivalent formation as well as normal gametes. It is imperative to emphasize that sugarcane is an artificial genome, highly polyploid, aneuploid, and has interspecific origins, which impedes our capacity to designate co-dominant markers at any locus. Consequently, loci were divided into heterozygous in one parent and null in the other (simplex *x* nuliplex), and heterozygous in both parents (simplex *x* simplex). Based on this model, all markers were scored as dominant (or binary), and were assigned to the expected segregation ratios i.e. 1:1 and 3:1 (Table 1).

A total of 992 segregating loci were genotyped in the mapping population; 685 AFLPs (generated from 22 enzyme-selective primer combinations), 220 EST-SSRs (derived from 41 loci), and 87 *scIvana_1*-based loci (obtained from six enzyme-primer combinations). The expected segregation ratios at each locus (992) were checked using Chi-square tests, adjusting for multiple tests using Bonferroni correction. Then, 730 (73.6%) loci were safely used to build the map, being 535 AFLPs, 130 EST-SSRs, and 65 *scIvana_1*-based loci. The global level of significance used to determine the validity of the segregation ratios of 1:1 and 3:1 was 5.04e-05 (alpha = 0.05/992).

The genetic map and marker distribution within co-segregation groups

The marker number used to perform linkage analyses was 730; 261 were derived from 'IAC66-6' (here indicated as D1), 284 from 'TUC71-7' (or D2), and 185 were present in both parents (or C). The final sugarcane map was comprised of 546 markers assembled into 92 co-segregation groups (CGs), and 184 markers not assigned to any CG. Coincidently, D1 (208) and D2 (240) markers

were distributed into nearly the same number of CGs (51 and 55, respectively). By using the loci that segregated in a 3:1 fashion (98) as bridges, we provided the integration of 19 CGs (I-1, I-2, I-3, I-4, I-6, I-8, I-9, I-12, II-1, II-3, III-1, IV-1, VII-1, U-1, U-2, U-3, U-8, U-13, and U-33). The marker number per CG ranged from two to 39 (Figure 1). The total map length was 4,843.19 cM, with a marker density of 8.87 cM. The CG length covered a range from 1.14 to 404.72 cM, with an estimated average of 52.64 cM. The greatest distance between two adjacent markers was 48.25 cM (CG I-3).

Forty-one CGs (44.6%) were assembled into seven putative independent homo(eo)logous groups (HGs), which were recognized based on 82 alleles generated by 24 EST-SSR loci. The remaining loci (ESTB189, ESTB191, CV11, and CV86) did not contribute to CG assembly; only one allele was positioned on the respective HG. The CG number assembled into HGs ranged from two to 15, the largest (HGI) contained 210 markers, and the smallest (HGVII) six markers, which was clearly a very non-uniform distribution (Figure 1). Eight previously mapped EST-SSR loci [32] were placed here on HGI (ESTA53, ESTB94, ESTB100, and ESTB118), HGIII (ESTA48), HGIV (ESTC80 and ESTB99), and HGV (ESTB60).

The mapped proportion of *scIvana_1*-based loci was 86.2% (56/65), greater than AFLPs (73.8%, 395/535) and EST-SSRs (73.1%, 95/130). The *scIvana_1*-based loci (56) were positioned on 21 CGs, which were represented in four HGs. However, there is some evidence that these markers may not be regularly distributed (Figure 1); only one marker was placed in groups I-6, I-13, III-3, III-4, U-1, U-2, U-4, U-7 and U-8, and more than five markers were placed in groups I-1, I-3, and IV-1. The evidence for clustering of *scIvana_1*-based markers was verified by the Chi-square goodness-of-fit test that results in a p-value 4.64e-04. This model has been used for testing if marker distribution deviates significantly from a random distribution in genetic maps i.e. segregating as expected in a Poisson distribution [47]. The number of regions with two (6) and three (1) adjacent markers was higher than expected only by chance (2 and 0.09, respectively). Interestingly, distances between some of the adjacent *scIvana_1*-based markers were lower than 5 cM, notably in groups, I-1, I-2, III-2, IV-1, and U-51 (Table 2). This value was the same used by Rossi *et al.* [36] to define NBS/LRR RGA clusters.

We subsequently obtained the total *scIvana_1*-based marker numbers (357), divided by the number of enzyme-primer combinations used to obtain the amplicons (6), and the result was ~60. We propose this as the number of retrotransposon *scIvana_1* copies in the sugarcane genome. Similarly, by obtaining the total number of *scIvana_1*-based markers each parent contributed separately, the following results correspond to the number of *scIvana_1* copies in the respective 'IAC66-6' and 'TUC71-7' genomes: 53 (316/6) and 54 (324/6).

However, the total number of *scAle_1*-based fragments was 1008. Nine restriction enzyme-primer combinations were used to amplify the *scAle_1*-based fragments, an average of 112 per combination. The results did not facilitate the identification of any clear polymorphisms in the gels. Consequently, the data were not used for mapping studies.

Molecular validation of *scIvana_1*-based fragments

Twenty-five *scIvana_1*-based markers selected for genotyping were sequenced; and 64% (16/25) showed homology with known nucleotide sequences deposited in GenBank (Table 3). Most (13 sequences) showed homology with *scIvana_1* sequences, and others showed homology to *Zea mays* (DIsIvSI337), *Oryza minuta* (DIsIvSI390, DIsIvSI412), and *Sorghum bicolor* (SIsIvSI180) sequences. The DIsIvLI228 fragment also exhibited homology with a *Saccharum* chloroplast sequence. Six of the sequences were mapped, DIsIvSI163, DIsIvSI208, DIsIvSI160, SIsIvLI240, SIsIvLI412, and DIsIvLI415, and two were tightly linked with EST-SSRs. The DIsIvSI337 fragment was submitted for tblastx search, and revealed similarity with a hypothetical, highly conserved protein of unknown function in *Arabidopsis thaliana*, and other species. Similarly, DIsIvSI390 and DIsIvSI412 fragments were aligned with an *Oryza minuta* sequence; tblastx indicated similarity with a hypothetical protein of unknown function. The SIsIvSI180 fragment exhibited similarity to two *Sorghum* sequences, one to a retrotransposon of *S. bicolor* [48]. Finally, the DIsIvLI228 fragment exhibited a certain identity with a ribosomal protein chloroplast sequence of *Saccharum* 'SP80-3280' (Table 3).

Discussion

AFLPs have been used to assess genetic diversity in germplasm collections of sugarcane and close relatives [49-52], and to build linkage maps [19,53]. In light of these studies, AFLPs have been informative in generating a substantial amount of unambiguous polymorphic markers. For example, Andru *et al.* [19] reported 64 AFLP restriction enzyme-primer combinations, and detected 816 polymorphic loci; in the present study, 22 combinations revealed 685 polymorphic loci. We suggest the differences were the result of the genotyping population, the first with increased homozygosity (S_1 progeny) relative to a segregating F_1 population. In both studies, AFLPs were a viable marker to build a scaffold for other marker types, including expressed sequence-based markers, which could be positioned on the scaffold. Furthermore, this scaffold is particularly important in

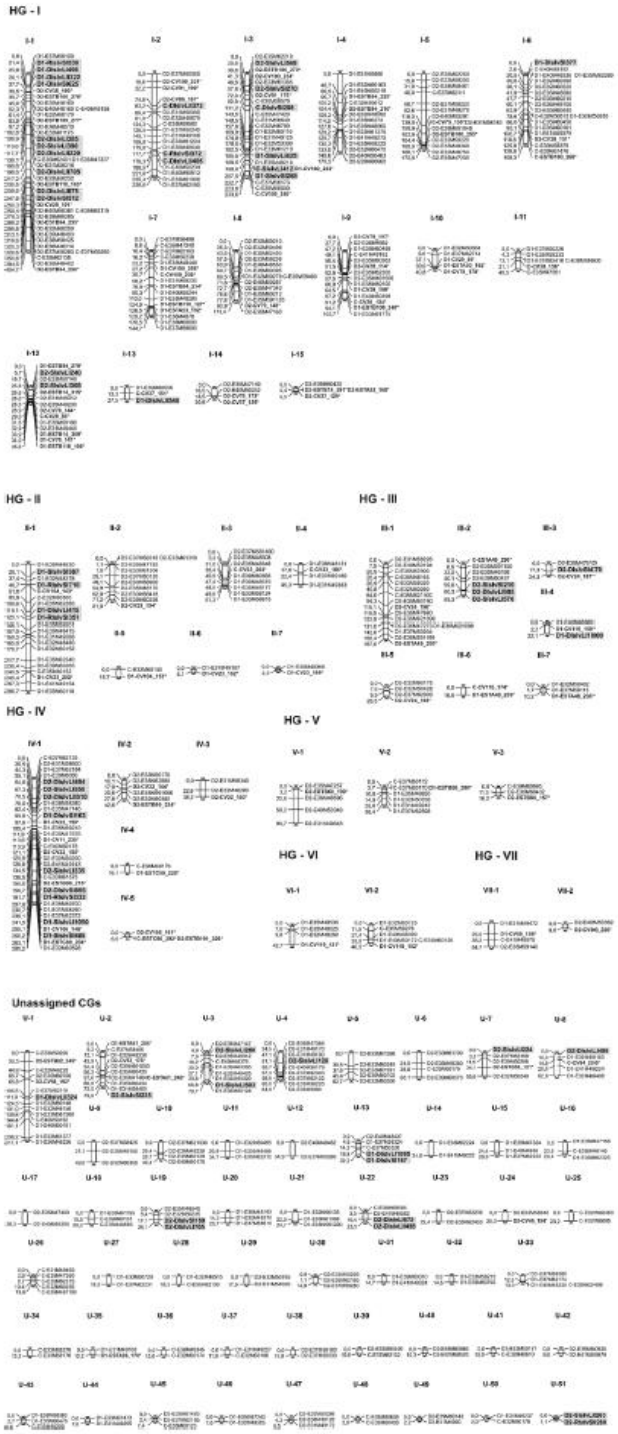


Figure 1 (See legend on next page.)

(See figure on previous page.)
Figure 1 Linkage map of the population 'IAC66-6' x 'TUC71-7'. Genetic distances between adjacent markers are shown on the left of each co-segregation group (CG). AFLPs constitute the map scaffold; EST-SSR loci appear in bold with asterisk symbols, and *scIvana_1*-based markers are depicted in bold in a gray box. The final map was constructed with 546 markers associated with 92 CGs (in Arabic numerals). Forty-one CGs (42.7%) were assembled into seven putative independent HGs (in Roman numerals). Other CGs (51) were denoted as unassigned groups (U). Note the clustered tendency of some *scIvana_1*-based markers.

sugarcane, as expressed sequences are physically too far apart.

From the 22 enzyme-selective primer combinations selected in this study to detect AFLPs, 14 were previously tested to build maps for the 'R570' cultivar [53], the *S. officinarum* 'IJ 76-514' clone, and 'Q165' [28], 'Q117' and 'MQ77-340' [27] cultivars as well as for the F_1 population 'SP80-180' *x* 'SP80-4966' [30]. Based on shared AFLPs and SSRs, the map comparisons of 'R570', 'Q165', 'Q117', and 'MQ77-340' cultivars were done [33]; several co-segregation groups (CGs) were aligned, and homo(eo)logous groups (HGs) associated, which received the same designation. As several common alleles were positioned, therefore suitable data should be available to construct a reference map for sugarcane commercial varieties, despite the pedigree complexity. For instance, a comparison between the map built based on the present study and the one published by Pastina *et al.* [54] is possible. Both are integrated linkage maps that share some common SSRs. In both maps, ESTA53, ESTB94, ESTB100 and ESTB118 were assigned to HGI as well as ESTA48 was mapped in the same grouping, HGIII. In addition, HGV contains three CGs that share the ESTB60 locus, which was assigned to HGVII of Pastina's map [54]. This suggests a possible correspondence between these HGs.

Several authors have applied SSR markers to estimate linkage in sugarcane [19,30,32,33]. Due to the multiallelic nature and relative abundance of SSRs in plant genomes, they have utility to identify HGs in polyploid species [33]. Based on this principle, Rossi *et al.* [36] identified 66 CGs (of 128) assembled into seven HGs from the French cultivar 'R570' linkage map. Similarly, Aitken *et al.* [28] identified 136 CGs assembled into eight HGs from the Australian cultivar 'Q165' linkage map. Oliveira *et al.* [32] identified 120 CGs (of 192) assembled into 14 HGs from a map of the progeny derived from a single cross between 'SP80-180' and 'SP80-4966'.

Here, we compiled 41 CGs (of 92) into seven putative HGs. Interestingly, six EST-SSR loci were duplicated within chromosomes CV22, CV38, CV78, CV100, ESTB14, and ESTB94, which were positioned on HG I and IV (Figure 1). Duplicated genomic regions were reported to occur in various sugarcane maps [23,28,32,33], and are possibly a consequence of the multispecific origins of the modern cultivars. Structural genomic rearrangements, including the movement of transposable elements (TEs) may also explain the duplications [55].

We used an innovative approach by mapping transposon-based markers in sugarcane using the NBS-profiling technique. In other plant species, TEs have been mapped using SSAP. This marker system was applied in barley (namely the TE *BARE-1*) [40], wheat (TEs, *Wis2A-1 A* and *BARE-1*) [56], lettuce, (*Tls1* and *Tls2*) [57], and tomato (*ToRTL1*, *T265* and *Tnt1*) [58]. Due to the advanced knowledge in tomato genetics, it was possible to determine that polymorphic insertions were primarily located in the centromeric regions. Both the above-mentioned approaches increase the available information regarding retrotransposon distribution over plant linkage groups. The NBS-profiling protocol efficiently targeted *scIvana_1* retrotransposon sequences and, at the same time, produced a polymorphic multilocus marker profile that was enriched for these sequences. Both the SSAP approach and the NBS-profiling technique investigate polymorphic restriction sites and the presence or absence of the retrotransposon sequence. SSAP uses two restriction enzymes (one with frequent cut sites and the other with rare cut sites), generating a large number of DNA fragments before performing the selective amplification. PCR results from the use of a primer that is complementary to the adaptor sequence and other complementary to the retrotransposon sequence. Since we have to choose two enzymes which have no recognition sites for restriction in the retrotransposon sequence, it reduces the number of combinations (enzyme/primer) to be tested. The NBS-profiling technique only uses one restriction enzyme (with frequent or rare cut sites), generating a small number of DNA fragments to be selected and consequently be stained and visualized separately in the gel. This is especially important considering the large genome size of sugarcane. Using enzymes that have no recognition sequences in the retrotransposon *scIvana_1* (combined with primers complementary to its sequence), it was possible to estimate the number of copies of this element in the genome.

Among the 16 restriction enzyme-primer combinations used to amplify *scIvana_1*, the enzymes *Dra*I and *Ssp*I resulted in an increased number of scorable bands and polymorphic loci compared to *Alu*I and *Rsa*I. Earlier studies have shown rare-cutting enzymes such as *Dra*I and *Ssp*I are more suitable for restricting sugarcane DNA due to its large size. However, enzymes that

Table 2 Retrotransposon clustering and its position into the sugarcane genetic map

Cluster[a]	Markers	Cluster size [a] (cM)	Map position [b]	
			HG	CG
1	D1-RIslvSI530, D1-DIslvLI496, D1-DIslvLII322	6.64	I	I-1
2	D2-DIslvLI385, D2-SIslvLI390	1.74	I	I-1
3	C-RIslvSI372, C-DIslvLII405	0.61	I	I-2
4	D2-SIslvSI210, D2-DIslvLI905	0.7	III	III-2
5	D2-DIslvLI494, D2-SIslvLI550, D2-DIslvLII310	5.66	IV	IV-1
6	D2-SIslvLI263, D2-RIslvSI280	1.14	-	U-51

[a] Defined by the distance between the flanking markers.
[b] HG, homo(eo)logous groups; CG, co-segregation groups.

frequently cut sugarcane DNA have the potential to generate an enormous number of fragments, and consequently affect other protocol steps, and subsequent results. All enzyme-primer combinations resulted in non-amplification for *scAle_1*. Primer design was challenging due to sequence diversity among *scAle_1* copies. When amplicons were obtained, all combinations resulted in ~ 112 bands per gel, which prevented polymorphism identification. Alternatively, a reduction in the number of bands per gel can be reached by adding selective bases at the 3'-end of PCR primers, as previously demonstrated in barley [40]. Queen *et al.* [56] used SSAP to study the elements *Wis2A-1A* and *BARE-1* in wheat, and four selective nucleotides were added as an attempt to reduce the number of amplified fragments. Both elements are known to have 1,000 copies in the wheat genome, and good results were obtained for genotyping and mapping when applying this strategy. Besides this, we should try to produce markers derived from *scAle_1* subfamilies, therefore having an estimate of the number of copies of each subfamily in the sugarcane genome.

The mean number of amplicons obtained by restriction enzyme-primer combinations was very similar to the number estimated by molecular methods; 40 to 50 copies of *scIvana_1* were detected in the sugarcane genome, and *scAle_1* exceeded 1,000 copies [39]. These results were congruent with 56 *scIvana_1*-based loci

Table 3 *scIvana_1*-based fragments with homology to nucleotide sequences deposited in GenBank (e-value < e^{-5})

Marker code	Size (bp)	GenBank accession no.	Homology	E-value [a]
DIslvSI	138	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-27
DIslvSI	160	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	8e-27
DIslvSI	208	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	5e-29
DIslvSI	310	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-28
DIslvSI	337	EU969904.1	*Zea mays* clone 337091 mRNA sequence	6e-28
DIslvSI	390	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	5e-178
		AC216031.1	*Oryza minuta* clone OM__Ba0016E09, complete sequence	1e-70
DIslvSI	412	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-147
		AC216031.1	*Oryza minuta* clone OM__Ba0016E09, complete sequence	4e-60
SIslvSI	158	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-28
SIslvSI	180	AC169373.2	*Sorghum bicolor* clone SB_BBc0188M08, complete sequence	3e-36
		FN431662.1	*Sorghum bicolor* BAC contig 24P17c, cultivar Btx623	1e-34
SIslvSI	245	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-28
SIslvSI	330	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-28
RIslvSI	195	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	5e-29
DIslvLI	228	AE009947.2	*Saccharum* 'SP-80-3280' chloroplast, complete genome	3e-90
DIslvLI	272	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	9e-106
DIslvLI	415	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-93
DIslvLI	605	DQ115032.1	*Saccharum* 'SP80-3280' clone SCCCCL6002A07 Tnt1-like, partial sequence	2e-102

[a] All are standard nucleotide-nucleotide BLASTn scores.

positioned on the linkage map, exhibiting preferential cluster distribution along 21 CGs. In large-genome cereals, Bennetzen [59] reported retrotransposon distribution as nested insertions in highly heterochromatic transposon clusters. Later, authors reported there appears to be some clustering of TE *BARE-1/Wis-2-1 A*-based markers on the linkage map of wheat [56], and transposon cluster interference with recombination machinery acting in adjacent euchromatic regions in maize [60]. When additional genomic data is available for sugarcane, it will be interesting to investigate if genes adjacent to retrotransposon clusters are less recombinogenic. Dooner and He [60] suggested the more condensed chromatin state of retrotransposon clusters in maize might interfere with recombination machinery access in adjacent euchromatic regions. Additionally, in a recent review published by Kalendar *et al.* [61] authors indicated that at least in cereals and citrus retrotransposons are often locally nested one into another and in extensive domains that have been referred to as 'retrotransposon seas' surrounding gene islands.

We are possibly facing an association between clustered retrotransposon sequences, the inhibition of DNA recombination, an explanation of the small map distance between adjacent retrotransposon-based markers, and the element copy number in plant the genome. This should explain *scIvana_1* properties, such as low copy numbers (~60) with expression and mobility under strict control, conversely against the properties of *scAle_1* retrotransposons. Therefore, mapping *scAle_1* element is of great interest, as well as the location of these two elements in chromosome regions.

The segregation results presented here independently indicated that AFLPs, EST-SSRs, or *scIvana_1*-based loci were consistent with the outcome of former studies [26,28,32,53,62] i.e. most markers (~70%) were SDMs. Furthermore, a substantial number were unassigned markers, in addition to variation in the marker number per CG.

The genetic map constructed here ('IAC66-6' *x* 'TUC71-7') has 546 SDMs covering 4,843 cM that were ordered in 92 CGs, with a marker density of 8.87 cM. The genetic map recently published by Pastina *et al.* ('SP80-180' *x* 'SP80-4966') has 317 markers covering 2,468 cM that were ordered in 96 CGs, with a marker density of 7.5 cM [54]. These are both integrated maps constructed based on segregating F_1 populations. Note that the number of CGs is somewhat high in Pastina's map, but it is shorter and denser. Cultivar maps are established using self-fertilized populations, and therefore are not comparable to other maps built based on different backgrounds. For instance, the 'LCP 85-384' map has 784 markers covering 5,617 cM that were assigned to 108 CGs, with a marker density of 7.16 cM [19], and 'Q165' map has 910 markers covering 9,058 cM that were assigned to 116 CGs, with a marker density of 9.95 cM [28]. Note that, in this case, the shorter and denser map is the one with a low number of CGs.

Enhancement of sugarcane genetic maps should include additional segregation ratios in mapping analyses, and an increased number of informative SNP- and SSR-loci segregating in larger populations. In addition, there is a need for meiotic studies that it is an important component of future studies in deciphering the genetic configuration of sugarcane genotypes.

Finally, it is important to note that the parents of the mapping population differ in response to the *Sporisorium scitamineum* infection; therefore we expect that offspring segregate for this trait. Consequently, the genetic map established here should be used to localize quantitative loci. It will certainly contribute to a better view on the genetic architecture of smut resistance in sugarcane, as little is known on this subject [63,64]. As recently shown in Pastina *et al.* [54] integrated genetic maps are useful for mapping QTLs. Based on interval mapping and mixed models, authors map QTL effects on a segregating progeny from a cross between two pre-commercial cultivars. The same approach should be interesting to be applied using the present map that includes retrotransposon-based markers. Moreover, we should improve McNeil *et al.*'s strategy [65] by aligning marker sequences tightly linked to QTLs for smut resistance with data from the sugarcane genome sequencing project currently underway [66].

Conclusions

The results of this study showed that AFLPs are a viable marker to create a scaffold for a linkage map, where other marker types can be positioned including expressed sequence-based markers. Results indicated the use of a NBS-profiling technique was efficient to develop retrotransposon-based markers in sugarcane. The simultaneous maximum-likelihood estimates of linkage and linkage phase based strategies confirmed the suitability of its approach to estimate linkage, and construct the linkage map. Interestingly, using our genetic data it was possible to calculate the number of retrotransposon *scIvana_1* (~60) copies in the sugarcane genome, confirming previously reported molecular results. In addition, this research possibly will have indirect implications in crop economics e.g., productivity enhancement via QTL studies, as the mapping population parents differ in response to an important fungal disease.

Methods

Plant material and genomic DNA extraction

The mapping population was composed of 188 individuals derived from a single cross between 'IAC66-6' and 'TUC71-7'. The male parent 'IAC66-6' is a clone with

low sucrose content, large diameter stems, and is susceptible to sugarcane smut, a fungal disease caused by *Sporisorium scitamineum*; the female parent, the Argentinean variety 'TUC71-7', exhibits a higher sucrose content, lower diameter stems, and resistance to smut disease. This disease limits the use of recent high-yielding sugarcane varieties developed in Brazil. The cross was made under field conditions at the CanaVialis/Monsanto Company experimental station, located in the northeastern state of Alagoas, Brazil (S 09°39'57"; W 35°44'07"). Sugarcane successfully flowers and sets seed at the locality due to light period duration (photoperiod), and plants were therefore cultivated at this site. Seeds were harvested, germinated in plastic boxes, and transported to the southeastern state of São Paulo (S 22°19'49"; W 47°10'21") for field cultivation.

DNA was isolated from young leaves of F_1-progeny and parental plants using the CTAB-based extraction procedure [67], with minor modifications. DNA concentrations were carefully estimated following electrophoresis on ethidium bromide-stained agarose gels using molecular weight standards; aliquots of 50 ng/μl were prepared following quantification.

Generation of AFLP profiles

AFLPs were amplified based on the protocol described by Vos *et al.* [68] and applied to sugarcane. Briefly, 250 ng of genomic DNA was double digested with 6U of *Eco*RI (Promega) and *Mse*I (NE Biolabs) in a 25-μl reaction mixture (10 mM Tris–acetate pH 7.5, 10 mM Mg-acetate, 50 mM K-acetate, 5 mM DTT, 1 X BSA) for 4 h at 37°C. Restrictions were terminated by heat inactivation for 20 min at 65°C. The resulting fragments were ligated to adapter sequences by addition of an equal volume of ligation solution comprised of 0.25 μM *Eco*RI adapter, 2.5 μM *Mse*I adapter, 1 X enzyme reaction buffer, and 67 U of T4 DNA ligase (400 units/μl, NE Biolabs). Incubations were performed for 16 h at 16°C, and reactions were terminated by heat inactivation. The adapter-ligated DNA (3 μl) was used for pre-selective amplification with primers based on the adapter sequences with one selective nucleotide at the 3' end (*Eco*RI + A and *Mse*I + C). Pre-selective amplification was performed in a 20 μl reaction mixture containing 1.5 mM $MgCl_2$, 0.5 mM each dNTP, 250 nM each primer, 1 X enzyme reaction buffer, and 3 U *Taq* DNA polymerase (Promega). Amplifications were conducted under the following conditions: 94°C for 2 min; 26 cycles of 94°C for 60 s, 56°C for 60 s, 72°C for 60 s; and a final elongation at 72°C for 5 min. For the selective step, 1.5 μl of a 5-fold water diluted pre-selected PCR product was used as DNA template. The 20 μl reaction mixture contained 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 250 nM *Eco*RI + ANN, 300 nM oligo *Mse*I + CNN, 1.6 U *Taq* DNA polymerase (Fermentas), and 1 X enzyme reaction buffer. Selective amplification was conducted under the following conditions: 94°C for 2 min; 12 cycles of 94°C for 30 s, 65°C for 30 s, 72°C for 60 s; the final 23 cycles had similar conditions with the exception of a 56°C primer annealing temperature, and a final elongation at 72°C for 2 min. Following PCR, the amplified products were mixed with an equal volume of denaturing buffer containing 95% formamide, 10 mM EDTA (pH 8.0), 0.2% bromophenol blue, and 0.2% xylene cyanol. Samples (3 μL) were loaded into 5% (w/v) polyacrylamide gels (acrylamide/bis-acrylamide, 19:1). Electrophoresis was performed at a constant power of 70 W for 4 h, using a Sequi-Gen® GT (Bio Rad) apparatus. Gels were silver-stained according to the protocol described by Creste *et al.* [69].

Seventy-two different restriction enzyme and selective primer combinations were examined using DNA from both parents in duplicate reactions. Combinations that exhibited good profiles, and revealed a large number of loci and polymorphism rates (≥ 20%) were selected for genotyping the F_1 population. The polymorphism rate between parents was calculated by assessing the number of bands present in one parent and absent in the other, in relationship to the total number of amplified bands.

EST-SSRs amplification

In analyzing the SUCEST database, Pinto *et al.* [70] identified 2005 clusters containing SSRs. Primer sets were subsequently developed from these data, and used in polymorphism analyses [70-73]. In addition, Maccheroni *et al.* [74] analyzed 352 and 122 sugarcane ESTs available in both public [75] and private [76] databases to establish sugarcane SSRs. Primer sets were developed from these sequences. In the present study, we used published [72-74], and non-published primer sets developed by CanaVialis/Monsanto.

EST-SSRs amplification was performed in a final volume of 10 μl in 96-well thermocycler plates. Approximately 20 ng of template DNA was mixed in a solution of 0.25 μM of each forward and reverse primer, 0.2 mM each dNTP, 2.0 mM $MgCl_2$, 1X Colorless Go *Taq* buffer, and 1.0 U Go *Taq* Flexi DNA Polymerase (Promega). Amplifications were performed using two thermal cycling programs. The first program was conducted under the following conditions: 94°C for 3 min; 31 cycles of 94°C for 60 s; primer annealing at varied temperatures for 60 s; extension at 72°C for 60 s; and a final elongation at 72°C for 2 min. The second was conducted under the following cycle parameters: an initial denaturation step at 94°C for 5 min; followed by 35 cycles of 94°C for 30 s; primer annealing at varied temperatures for 30 s; extension at

72°C for 30 s; and a final elongation at 72°C for 60 min. PCR products were analyzed by two methods, denoted S and F. Amplicons were resolved by 5% (w/v) denaturing polyacrylamide gel electrophoresis, and silver stained (S) as above described; or capillary electrophoresis using a MegaBACE 1000® genotyping system (GE Healthcare Life Sciences). For capillary electrophoresis, forward primers were labeled with fluorescent dyes (F) (fluorophore NED or 6-FAM, Applied Biosystems), and fragments were verified with the Fragment Profiler version 1.2®.

Polymorphisms between parental genotypes were assessed by amplifying 184 EST-SSRs using DNA from 'IAC66-6', 'TUC71-7', and a sample from the mapping population (F_1). The data included 33 EST-SSRs developed by Oliveira *et al.* [72], two from Marconi *et al.* [73], and three from Maccheroni *et al.* [74]. In addition, 146 EST-SSRs were available from CanaVialis/Monsanto (unpublished data). Results showed 41 polymorphic loci, which did segregate in the progeny sample; therefore, these loci were applied to genotype the mapping population. Details on these sugarcane EST-SSRs are presented in Table 4.

Marker generation based on sugarcane retrotransposon sequences

The principle NBS-profiling technique [77] was applied according to Hanai *et al.* [78] to generate markers based on two retrotransposons named *scIvana_1* (GenBank Accession Number JN800016) and *scAle_1* (GenBank Accession Number JN800006). Approximately 500 ng of genomic DNA were digested with *Alu*I, *Dra*I, *Ssp*I, or *Rsa*I (NE BioLabs). Digestions were performed in a final volume of 30 μl using 7.5 U of enzyme for 7 h at 37°C, according to the manufacturer's recommendations. Reactions were terminated by heat inactivation (20 min at 65°C). Adapters were prepared by incubating equimolar amounts of LA (long arm) and SA (short arm) oligonucleotides at 65°C for 10 min, and respectively cooled to 37°C and 25°C (10 min each). The SA oligonucleotide 3' end was blocked for *Taq* DNA polymerase extension by the addition of an amino group, but phosphorylated at the 5' end, which results in an adapter primer-annealing site only following the first PCR cycle. Subsequently, the digested material and a solution containing a 1.6 μM adapter (when restricted with *Alu*I or *Rsa*I), or a 0.2 μM adapter (when *Dra*I or *Ssp*I was used), 1 X ligation buffer (NE BioLabs), and 67 U T4 DNA ligase (400 units/μl; NE BioLabs) were mixed in equal volumes (30 μl), totaling 60 μl. Ligation was performed at 16°C for 16 h, and terminated by heat inactivation at 65°C for 20 min. The ligation products (diluted to 5 ng/μL) as template DNA were used to amplify selected fragments anchored to the retrotransposon sequence. A final volume 20 μl reaction mixture contained 4 μl of ligation products, 300 nM of each primer (a primer complementary to the adapter, and a primer complementary to the retrotransposon sequence, Table 5, Figure 2), 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 1 X buffer enzyme, and 2 U *Taq* DNA polymerase (Fermentas). PCR was conducted under the following conditions: an initial denaturation step at 94°C for 5 min; followed by 8 cycles at 94°C for 45 s, 58°C (– 1°C per cycle) during 50 s, and 72°C for 1 h:15 min; 25 cycles at 94°C for 45 s, 50°C for 50 s, and 72°C for 1 min; and a final extension at 72°C for 10 min. After PCR, the protocol for preparing AFLP gels was completed, followed by electrophoresis, and gel staining.

Sixteen restriction enzyme and retrotransposon complementary primer combinations (for each sugarcane retrotransposon) were examined using the DNA from both parents in duplicate reactions. Combinations that exhibited clear band distribution over the gels, and revealed polymorphism rates ≥ 15% were selected for genotyping the F_1 population.

Marker nomenclature, genotyping and segregation analyses

The nomenclature adopted for AFLP primers followed the Keygene standard primer list [80] followed by the corresponding molecular size (in bp) of the band. A primer code was adopted for EST-SSR loci (ESTA, ESTB, ESTC and CV) followed by the molecular size of the allele (in bp). The retrotransposon-based locus nomenclature was an abbreviation that indicated the enzymes used in the digesting reaction (AI for *Alu*I, DI for *Dra*I, SI for *Ssp*I, and RI for *Rsa*I) followed by the primer code (Table 5), and the molecular size of the fragment (in bp). Wu's [31] loci-segregation pattern notations follow all marker abbreviations e.g., D1-E35M47510 (Figure 1). We assumed the presence of a band, denoted by *A*, dominant over all *0* or null alleles (or simplex configuration with eight homo(eo)logous copies in the sugarcane genome), independent of marker type. Loci were denoted as "D1" when 'IAC66-6' was heterozygous for band presence, and the other parent was homozygous for band absence (*A0000000 x 00000000*). Loci were denoted as "D2" when 'TUC71-7' was heterozygous and 'IAC66-6' was homozygous. These loci were expected to segregate in a 1:1 fashion in the F_1 population. "C" loci were heterozygous in both parents (*A0000000 x A0000000*), and exhibited a 3:1 segregation ratio. Differences between observed and expected proportions were compared using Chi-square test, assuming a polyploid model based on single-dose markers (SDM) for analyzing segregation in outcrossing species [14]. For minimizing problems caused by multiple comparisons, the Bonferroni correction was performed. Chi-square tests and Bonferroni adjustments for the effects of multiple

Table 4 Details on the sugarcane microsatellite loci derived from expressed sequence tags (ESTs)

Marker code	Repeat motif	Forward primer (5'→3')	Reverse primer (5'→3')	PCR [e]	AT [f]	D [g]	Allele number and size range [h]	
ESTA26 [a]	$(TG)_{11}$	GGCAGCCCCACATCTTCCT	GGGCACAAGCATCCGAACC	1	56.0	S	4	172-186
ESTA48 [a]	$(CA)_8$	AGCAACTCCGGCCTCTCCTG	CTTTCTGTTTTGCTCCTCCGTCTG	1	62.7	S	10	233-295
ESTA53 [a]	$(TG)_8$	TGGAAATGGCAGCTGGTCTCGT	ATGCACGTACCAGAGGGAGATTTG	1	58.9	S	9	168-192
ESTA61 [a]	$(AT)_{12}$	ACCTCAGTCTCCTCCTCAACC	TATACTACACATGCACAGGCTACG	1	56.4	S	5	236-246
ESTB14 [a]	$(CGT)_8$	TGAGGGAATGAATGGACTGG	CCACCACCACCATACCTGTC	1	52.0	S	9	285-315
ESTB55 [a]	$(CCA)_5$	CTTCTTGGCCTTGGCGTTACTGA	GCTAGCTGGCCCCATTTCCTCT	1	60.0	S	3	118-124
ESTB60 [a]	$(TTG)_{10}$	AGCCGCAATGAATCCAACTG	CTCTAGCTCCGACGATGATACCTC	1	61.0	S	8	157-206
ESTB82 [a]	$(CGT)_9$	CGTCGATCGAGATGAAGAAGG	GAAGCAGTCGTGGAAGTGGAG	1	62.7	S	5	245-263
ESTB94 [a]	$(CTT)_9$	GAGGCAGCCAGGCAGGTCAC	GGTGGCAGTGTTCAGGCAGATG	1	61.0	S	10	210-279
ESTB99 [a]	$(TCG)_5$	GAGGTCCTTCTTGTAGTTGTATGC	GTGCCGGAGGATTTGATG	1	64.7	S	4	215-224
ESTB100 [a]	$(TCG)_6$	CCACGGGCGAGGACGAGTA	GGGTCCTTCTTCGCCTCGTG	1	64.7	S	13	240-278
ESTB118 [a]	$(TTC)_6$	CTTGGCTAGGGTTTCTTGAGTCGT	CATGGCTTTTGGCTTGCTTCT	1	61.0	S	5	106-163
ESTB189 [b]	$(TCA)_{10}$	GTAAGGAAGAAGCAACAAACAACAG	GATTCGATGCAACTCTCCTGTAAA	1	60.0	S	5	261-280
ESTB191 [b]	$(GCT)_5$	GCGCCATCAGGGAAGCCAAAAC	GCGCGTGCGAGCAGATGAAC	1	60.0	S	5	213-226
ESTC80 [a]	$(ATTC)_3$	ATTCTTTCTCCCCCTGTTGTGC	GTCGCCAGATCGCTTTCGTT	1	58.9	S	7	188-292
CV06 [c]	$(AATT)_{13}$	TCTCAAGCTTCGCCAGCTA	TGGCTCGGCTGTAGGAATTA	2	60.0	S	3	188-230
CV11 [c]	$(GAA)_6$	TGGCATGTGTCATAGCCAAT	CCCCAACTGGGACTTTTACA	2	60.0	S	6	227-242
CV22 [c]	(AGGG)5	CACTACTCGCCCCGATTTC	CGAGTGCTTCTCCATCTGC	2	64.0	F	8	140-166
CV23 [c]	$(GGAA)_7/(AGG)_6$	GAACTGCTCACTGGCTCCTC	GTAGAAGTCCGTCGCCGTAA	2	64.0	F	9	150-206
CV24 [c]	$(CCAA)_5/(CACCT)_4$	TCGGAGAAGTTGACCGAGTT	GGTTTAGAGTTGGGGCCTTC	2	60.0	F	7	187-205
CV29 [d]	$(ATCT)_{14}$	TCGCGTCCACCAATGTAACC	GCGTGCATCGCTTGTGTCTT	2	64.0	F	10	85-133
CV37 [d]	$(TTTC)_{15}$	GGATGGACGACGTGTCCTGG	ATAAAGTGGCCGCTTGGATTGA	2	64.0	F	6	117-155
CV38 [d]	$(CTTTT)_{18}$	GAAGCAGGGGCCTCAAGTTG	GTCAAACAGGCGATCTGGCTC	2	64.0	F	9	109-199
CV46 [c]	$(GGTAA)_{11}$	TGTTCCAAGTTCATGCGCTCC	ATGCATGCAGGTTCAAAAGCAG	2	64.0	F	5	146-188
CV51 [c]	$(TGT)_{13}$	CTACCCCAACTTGCTTGGGAC	GACTGGAACAAAGACGGACTG	2	64.0	F	3	147-160
CV53 [c]	$(AAAAT)_5/(TTTAT)_6$	CCCCACCGTAGCTTGTGCAT	AAACGTGCACATGCTTGTATGC	2	64.0	F	7	160-183
CV58 [c]	$(ATAGAT)_{10}$	CGGGTAGTTAGGAGGAGATGG	GTCATCCATTTTGGAACGAATGG	2	64.0	F	6	153-195
CV78 [c]	$(CTGTG)_9$	ACGAGGCCACCATAGAACATG	GCAATTGGGAGGAGAGGAATG	2	64.0	F	9	144-203
CV79 [c]	$(CTATAT)_{11}/(TATAGA)_6$	GGCACTGCTGGTGGTTGATTG	TCCCACATCAAGAGGCAGCTA	2	64.0	F	7	136-197
CV86 [c]	$(AATT)_8$	CCTCAGCAGCCCAAAGTCCT	GTCGGAATCAGCCGGATTAGC	2	64.0	F	5	159-187
CV91 [c]	$(GCC)_6/(GCA)_6$	AAAGGAAATCGCCCTCCGTCT	CCGATGATGAGCCAGCAATCC	2	64.0	F	8	175-197
CV94 [c]	$(AAAAAG)_5/(CGT)_5$	GGCAGGCCAAGATGAATGAAG	AGCACAGCGGAGGGTACGG	2	64.0	F	4	187-205
CV100 [c]	$(GAG)_{13}$	CTGTTGAGGAGCCGGATGAG	CTCTTCCGATGGCTCGGTCT	2	64.0	F	9	222-256
CV101 [c]	$(ATC)_{23}$	GTCGTCGTCGTCACGATCATC	AGTTGACGGCATGGTTCTTGC	2	64.0	F	11	111-180
CV104 [c]	$(TCCTG)_5$	GATTTTCGACTGTGCGCTTGG	AAGTTCTCTGCCGGAGCAAAC	2	64.0	F	6	133-158
CV106 [c]	$(GGC)_8$	AAACAGAGCATACTCGAGGCC	ACGTTGCTGACGAGGTTTTCC	2	64.0	F	6	146-161
CV115 [c]	$(TCACAG)_{10}/(GTA)_6/(AGA)_5$	GTCCATGTCCATCCATGATCC	GGAGCTCCGTCTTCTTGTTAC	2	60.0	S	6	150-174
CV119 [c]	$(AAAAC)_7$	TATCTCTCCTTGGTTTGGATGG	CACCCTACCAAATACCACAACA	2	64.0	F	5	121-175
CV128 [c]	$(GCA)_{13}$	AGGGCAACGGAGTCTTCGAC	CTGAACTCCGATGTGCTGGTG	2	60.0	F	5	147-168
CV135 [c]	$(AAG)_{16}$	AGCAAAACCAGCCTTCCCTTC	CTGTTTGTTTCTGCTTGCTTGC	2	64.0	F	6	129-159
CV144 [c]	$(TCTCCG)_5$	GCGCCTCCGTGGATAAGAATC	CCTTCCCCTACAGCGCCTAC	2	64.0	F	5	146-164

[a, b, c, d] Develop by Oliveira *et al.* [72], Marconi *et al.* [73], CanaVialis (unpublished); Maccheroni *et al.* [74], respectively.
[e] PCR program as described in Material and Methods.
[f] Annealing temperature in the amplification reaction.
[g] D: Silver-stained polyacrylamide gel electrophoresis (S) or fluorescence-based automated capillary electrophoresis (F) for the detection of EST-SSR alleles.
[h] Observed number of alleles per locus and their size ranges in bp.

Table 5 Primer sequences used for the generation of sugarcane retrotransposons-based markers

Primer	Sequence 5'→3' [a]
Short arm oligonucleotide (AS)	TGGGATCTATACTT - H_2N
Long arm oligonucleotide (LA)	ACTCGATTCTCAACCCGAAAGTATAGATCCCA
Adapter primer (AP)	ACTCGATTCTCAACCCGAAAG
*scIvana1*_SSAP1 (slvSI)	CAAGCCCTTAATAGCAGAAA
*scIvana1*_GagRev (slvGR)	TCCCTGTATACAACCCTGTC
*scIvana1*_LTR1 (slvLI)	AGTCCTGCTCCCAGTTATCA
*scIvana1*_LTR2 (slvLII)	GTCGCCTGGGTGTGTTATC
*scAle1*_LTRr (sAlLr)	ATACATGGGCCACATGGG
*scAle1*_RT (sAlRT)	CCTCCCDTCCTCGACCTTC
*scAle1*_LTR1 (sAlLI)	CCATGWGRCTAGGCCCATGTGGC
*scAle1*_LTR2 (sAlLI)	GGGGTGTTGGAGTGTGATTG

[a] D = A, G or T; R = A or G; W = A or T.

comparisons were performed using the software R, v. 2.13.0 [81]. Loci with segregation distortion were not included in linkage analysis.

Linkage analyses, map construction, and identification of sugarcane homo(eo)logous groups

All linkage analyses were performed using OneMap software [82]. Version 2.0-1 was preferred to construct a multipoint maximum likelihood linkage map applying a Hidden Markov Model approach [83].

Firstly, co-segregation groups (CGs) were established using a LOD score ≥ 6.0, and a recombination fraction ≤ 0.35. For groups with six or less markers, the best order was obtained by comparing all possible orders choosing the one with highest likelihood using the algorithm implemented in the command named "compare". To obtain the best order for larger groups (more than six markers), the command "order.seq" was applied. In this case, the likelihood was the criteria used to place the markers along the CGs under a multipoint approach, as validated by Mollinari *et al.* [84]. Additionally, the "ripple" command was used to check for alternative orders, as well a visual inspection on the matrix containing the pairwise recombination fractions and LOD scores (heatmaps) for the CGs. The commands "compare", "order.seq" and "ripple" were similar to those in the MAPMAKER/EXP software [17]. Finally, multipoint estimates of recombination fractions were calculated and converted into linkage distances using the Kosambi map function [85]. Map drawings were generated using MapChart 2.2 [86].

Due to the multiallelic nature and known polymorphisms, EST-SSR loci are valuable in recognizing homo(eo)logous groups (HGs) in sugarcane. Initially, CGs were assigned to HGs if they contained at least two of the same EST-SSRs. In addition, CGs were putatively added if they contained an EST-SSR locus in common with the HG [28,30,32,54]. Using this practice, two HGs were established, I and IV. Then due to the insufficient amount of SSRs, in a number of instances only one locus was used to assign CGs to the following putative HGs, II, III, V, VI and VII [30,32]. We applied Roman numerals to denote HGs; within each HG, CGs were classified in a descending order according to size (in cM). The unassigned groups were designated U, and also classified according to their size.

Finally, to test if *scIvana_1*-based markers have a tendency to be clustered along the genome, we used an approach similar to the one presented by Echt *et al.* [87]. The genetic map was divided in 10 cM bins and the number of *scIvana_1*-based markers in each interval was recorded. If markers were randomly distributed, they would follow a Poisson distribution [47], defined as $P(x) = \lambda^x e^{-\lambda}/x!$, where P(x) is the probability function; x is the number of markers observed in the intervals (ranging from 0 to 3), λ is the distribution parameter calculated as average number of markers per interval in the map. A Chi-square goodness-of-fit test was performed, with 2

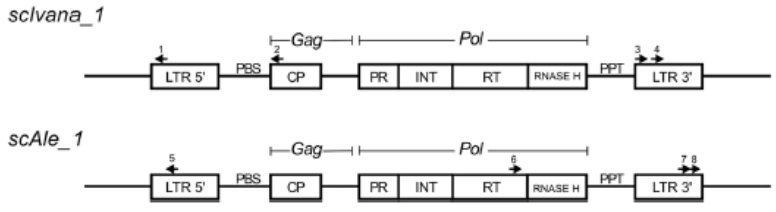


Figure 2 Structure of sugarcane retrotransposons *scIvana_1* and *scAle_1*. Retrotransposons are LTR (long terminal repeats) consisting of elements within transcription initiation and termination sequences and detected as Gag, Pol, and Int domains that code for CP (capsid-like proteins), PR (protease), RT (reverse transcriptase), RNAase-H (ribonuclease H), and INT (integrase). Other sequences featured are PBS (primer binding sites), and PPT (polypurine tracts). Arrows indicate the primers designed for amplifying each of the elements, and synthesis direction. 1: *scIvana_1*-SSAP1; 2: *scIvana_1*-GagRev; 3: *scIvana_1*-LTR1; 4: *scIvana_1*-LTR2; 5: *scAle_1*-LTRr; 6: *scAle_1*-RT; 7: *scAle_1*-LTR1; 8: *scAle_1*-LTR2. Figures were not drawn to scale and were adapted from Kumar and Bennetzen [79].

degrees of freedom (df = c −1- r, where c is the number of classes and r is the number of estimated parameters, r = 1).

Molecular validation of *sclvana_1*-based fragments

Verification that amplified DNA fragments were derived from retrotransposon templates was conducted by excising DNA fragments from polyacrylamide gels, eluting DNA in a TE solution (10:1), and re-amplifying the DNA fragments. Five μl of the diluted DNA mixture was added to 50 μL of the same solution used for retrotransposon-based marker generation, however primer concentration was 30 nM, and *Taq* DNA polymerase was 5 U. The PCR program was simplified and conducted under the following conditions: an initial denaturation step at 94°C for 5 min; 30 cycles of 94°C for 30 s, 55°C for 30 s; 72°C for 42 s; and a final extension at 72°C for 10 min. For sequencing, PCR fragments were resolved on agarose gels, purified with the QIAEX II Gel Extraction kit (QIAGEN), and cloned into pMOS Blue Blunt-Ended Cloning Kit (GE Healthcare Life Sciences). Inserts were sequenced in the forward direction. Sequencing reactions were performed according to Sanger *et al.* [88] using DYEnamicTM ET Dye Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech, Inc.) on an ABI 3730 system (Applied Biosystems). Sequence quality was examined using the Phred/Phrap/Consed package [89]. Nucleotide sequences were compared to reference data available at Genbank by BLAST analysis [90].

Additional files

Additional file 1: Amplification patterns obtained from AFLP, EST-SSR, and *sdvana_1*-based markers for the sugarcane mapping population. Several segregating alleles are shown, as well as molecular weight standard (lane M) fragment sizes. Codes correspond to parental and F_1-progeny genotypes.

Additional file 2: Genetic information provided by the AFLP, EST-SSR and the retrotransposon *sclvana_1*-based loci.

Abbreviations

AFLP: Amplified Fragment Length Polymorphism; BC: Backcross; CG: Co-segregation Group; DH: Double Haploid; EST: Expressed Sequence Tag; EST-RFLP: Expressed Sequence Tag-Restriction Fragment Length Polymorphism; EST-SSR: Expressed Sequence Tag-Simple Sequence Repeat; HG: Homo(eo)logous group; IRAP: Inter-Retrotransposon Amplified Polymorphism; ISBP: Insertion Site-Based Polymorphism; LG: Linkage Group; LRR: Leucine Rich Repeat; NBS-LRR: Nucleotide Binding Site–Leucine Rich Repeat; QTL: Quantitative Trait Loci; RBIP: Retrotransposon-Based Insertion Polymorphism; REMAP: Retrotransposon Microsatellite Amplified Polymorphism; RFLP: Restriction Fragment Length Polymorphism; RIL: Recombinant Inbreed Line; SDM: Single Dose Marker; SNP: Single Nucleotide Polymorphism; SSAP: Sequence Specific Amplified Polymorphism; SSR: Simple Sequence Repeat (or Microsatellite); S-T Kinase: Serine-Threonine Kinase; SUCEST: Sugarcane Expressed Sequence Tag; TE: Transposable Element; TRAP: Target Region Amplification Polymorphism.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ACP and TBRM ran AFLP, SSR and *sclvana_1*-based markers, created the figures and tables, and drafted the manuscript; MM and RG constructed the linkage map, and applied all statistical tests cited under and the supervision of AAFG; and MAVS and DSD described sugarcane *sclvana_1* and *scAle_1* retrotransposons. MLCV conceived the study, and wrote the final manuscript. APS and TM developed some of the SSRs used here. WMJr and HJJr provided the plant material as well as developed most of the SSRs. All authors read and approved the manuscript.

Acknowledgements

Authors are grateful to the following Brazilian funding institutions: Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP, grant 2010/51708-5), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Ensino Superior (CAPES).

Author details

[1]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900, Piracicaba, Brazil. [2]Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, 05508-090, São Paulo, Brazil. [3]Present address: Laboratório de Biotecnologia Vegetal, Instituto Agronômico do Paraná, 86047-902, Londrina, Brazil. [4]CanaVialis/Monsanto Co, Condomínio Techno Park, 13069-380, Campinas, Brazil. [5]Present address: Companhia Vale do Rio Doce, 20020-900, Rio de Janeiro, Brazil. [6]Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, 13083-875, Campinas, Brazil.

Received: 21 January 2012 Accepted: 13 June 2012
Published: 28 June 2012

References

1. *Food and Agriculture Organization of the United Nations*. http://faostat.fao.org.
2. *Companhia Nacional de Abastecimento*. http://www.conab.gov.br/OlalaCMS/uploads/arquivos/11_08_30_13_41_19_boletim_cana_portugues_-_agosto_2011_2o_lev.pdf.
3. Bremer G: **Problems in breeding and cytology of sugar cane. I - A short history of sugar cane breeding; the original forms of *Saccharum*.** *Euphytica* 1961, **10**:59–78.
4. Mukherjee SK: **Origin and distribution of *Saccharum*.** *Bot Gaz* 1957, **119**:55–61.
5. Roach BT: **Nobilisation of sugarcane.** *Proc Int Soc Sugar Cane Technol* 1972, **14**:206–216.
6. Sreenivasan TV, Ahloowalia BS, Heinz DJ: **Cytogenetics.** In *Sugarcane improvement through breeding*. Edited by Heinz DJ. Amsterdam: Elsevier; 1987:211–253.
7. Kimbeng CA, Cox MC: **Early generation selection of sugarcane families and clones in Australia: A review.** *J Am Soc Sugar Cane Technol* 2003, **23**:20–39.
8. Landel MGA, Silva MA: **As estratégias de seleção da cana em desenvolvimento no Brasil.** *Visão Agrícola* 2004, **1**:18–23.
9. D'Hont A, Glaszmann JC: **Sugarcane genome analysis with molecular markers, a first decade of research.** *Proc Int Soc Sugarcane Technol* 2001, **24**:556–559.
10. D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC: **Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes.** *Genome* 1998, **41**:221–225.
11. Ha S, Moore PH, Heinz D, Kato S, Ohmido N, Fukui K: **Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence in situ hybridization and imaging methods.** *Plant Mol Biol* 1999, **39**:1165–1173.
12. D'Hont A, Grivet L, Feldmann P, Rao PS, Berding N, Glauszmann JC: **Characterisation of the double genome structure of modern sugarcane cultivars (Saccharum spp.) by molecular cytogenetics.** *Mol Gen Genet* 1996, **250**:405–413.
13. Al-Janabi SM, Honeycutt RJ, McClelland M, Sobral BWS: **A genetic linkage map of *Saccharum spontaneum* L. 'SES 208'.** *Genetics* 1993, **134**:1249–1260.

14. Wu KK, Burnquist W, Sorrells ME, Tew TL, Moore PH, Tanksley SD: **The detection and estimation of linkage in polyploids using single-dose restriction fragments.** *Theor Appl Genet* 1992, **83**:294–300.
15. Da Silva J, Honeycutt RJ, Burnquist W, Al-Janabi SM, Sorrells ME, Tanksley SD, Sobral BWS: ***Saccharum spontaneum* L. 'SES 208' genetic linkage map combining RFLP- and PCR-based markers.** *Mol Breed* 1995, **1**:165–179.
16. Sobral BWS, Honeycutt RJ: **High output genetic mapping of polyploids using PCR-generated markers.** *Theor Appl Genet* 1993, **86**:105–112.
17. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1**:174–181.
18. Grivet L, D'Hont A, Roques D, Feldmann P, Lanaud C, Glaszmann JC: **RFLP mapping in cultivated sugarcane (Saccharum spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid.** *Genetics* 1996, **142**:987–1000.
19. Andru S, Pan Y-B, Thongthawee S, Burner DM, Kimbeng CA: **Genetic analysis of the sugarcane (Saccharum spp.) cultivar 'LCP 85–384'. I. Linkage mapping using AFLP, SSR, and TRAP markers.** *Theor Appl Genet* 2011, **123**:77–93.
20. van Ooijen JW, Voorrips RE: *JoinMap 3.0 Software for the calculation of genetic linkage maps.* Wageningen: Plant Research International; 2001.
21. Grattapaglia D, Sederoff R: **Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers.** *Genetics* 1994, **137**:1121–1137.
22. Mudge J, Andersen WR, Kehrer RL, Fairbanks DJ: **A RAPD genetic map of *Saccharum officinarum*.** *Crop Sci* 1996, **36**:1362–1366.
23. Ming R, Liu S-C, Lin Y-R, da Silva J, Wilson W, Braga D, van Deynze A, Wenslaff TF, Wu KK, Moore PH, Burnquist W, Sorrells ME, Irvine JE, Paterson AH: **Detailed alignment of *Saccharum* and *Sorghum* chromosomes: Comparative organization of closely related diploid and polyploid genomes.** *Genetics* 1998, **150**:1663–1682.
24. Edmé SJ, Glynn NC, Comstock JC: **Genetic segregation of microsatellite markers in *Saccharum officinarum* and *S. spontaneum*.** *Heredity* 2006, **97**:366–375.
25. Aitken KS, Jackson PA, McIntyre CL: **Construction of a genetic linkage map for *Saccharum officinarum* incorporating both simplex and duplex markers to increase genome coverage.** *Genome* 2007, **50**:742–756.
26. Alwala S, Kimbeng CA, Veremis JC, Gravois KA: **Linkage mapping and genome analysis in a *Saccharum* interspecific cross using AFLP, SRAP and TRAP markers.** *Euphytica* 2008, **164**:37–51.
27. Reffay N, Jackson PA, Aitken KS, Hoarau J-Y, D'Hont A, Besse P, McIntyre CL: **Characterisation of genome regions incorporated from an important wild relative into Australian sugarcane.** *Mol Breeding* 2005, **15**:367–381.
28. Aitken KS, Jackson PA, McIntyre CL: **A combination of AFLP and SSR markers provides extensive map coverage and identification of homo (eo)logous linkage groups in a sugarcane cultivar.** *Theor Appl Genet* 2005, **110**:789–801.
29. Raboin L-M, Oliveira KM, Lecunff L, Telismart H, Roques D, Butterfield M, Hoarau J-Y, D'Hont A: **Genetic mapping in sugarcane, a high polyploid, using bi-parental progeny: identification of a gene controlling stalk colour and a new rust resistance gene.** *Theor Appl Genet* 2006, **112**:1382–1391.
30. Garcia AAF, Kido EA, Meza AN, Souza HMB, Pinto LR, Pastina MM, Leite CS, da Silva JAG, Ulian EC, Figueira A, Souza AP: **Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases.** *Theor Appl Genet* 2006, **112**:298–314.
31. Wu R, Ma C-X, Painter I, Zeng Z-B: **Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species.** *Theor Popul Biol* 2002, **61**:349–363.
32. Oliveira KM, Pinto LR, Marconi TG, Margarido GRA, Pastina MM, Teixeira LHM, Figueira AV, Ulian EC, Garcia AAF, Souza AP: **Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross.** *Mol Breeding* 2007, **20**:189–208.
33. Piperidis N, Jackson PA, D'Hont A, Besse P, Hoarau J-Y, Courtois B, Aitken KS, McIntyre CL: **Comparative genetics in sugarcane enables structured map enhancement and validation of marker-trait associations.** *Mol Breeding* 2008, **21**:233–247.
34. *The Sugarcane EST Project.* http://sucest-fun.org.
35. Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MIT, Henrique-Silva F, Giglioti EA, Lemos MVF, Coutinho LL, Nobrega MP, Carrer H, França SC, Bacci M Jr, Goldman MHS, Gomes SL, Nunes LR, Camargo LEA, Siqueira WJ, Van Sluys M-A, Thiemann OH, Kuramae EE, Santelli RV, Marino CL, Targon MLPN, Ferro JA, Silveira HCS, Marini DC, Lemos EGM, Monteiro-Vitorello CB, *et al*: **Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane.** *Genome Res* 2003, **13**:2725–2735.
36. Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias VM, Chen H, Van Sluys M–A, D'Hont A: **Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane.** *Mol Gen Genomics* 2003, **269**:406–419.
37. Rossi M, Araujo PG, Van Sluys M-A: **Survey of transposable elements in sugarcane expressed sequence tags (ESTs).** *Genet Mol Biol* 2001, **24**:147–154.
38. Araujo PG, Rossi M, Jesus EM, Saccaro NL Jr, Kajihara D, Massa R, Felix JM, Drummond RD, Falco MC, Chabregas SM, Ulian EC, Menossi M, Van Sluys M-A: **Transcriptionally active transposable elements in recent hybrid sugarcane.** *Plant J* 2005, **44**:707–717.
39. Domingues DS, Cruz GMQ, Metcalfe CJ, Nogueira FTS, Vicentini R, Alves CS, Van Sluys M-A: **Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns.** *BMC Genomics* 2012, **13**:137.
40. Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W: **Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP).** *Mol Gen Genet* 1997, **253**:687–694.
41. Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A: **IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques.** *Theor Appl Genet* 1999, **98**:704–711.
42. Baumel A, Ainouche M, Kalendar R, Schulman AH: **Retrotransposons and genomic stability in populations of the young allopolyploid species Spartina anglica C.E. Hubbard (Poaceae).** *Mol Biol Evol* 2002, **19**:1218–1227.
43. Flavell AJ, Knox MR, Pearce SR, Ellis THN: **Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis.** *Plant J* 1998, **16**:643–650.
44. Paux E, Roger D, Badaeva E, Gay G, Bernard M, Sourdille P, Feuillet C: **Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B.** *Plant J* 2006, **48**:463–474.
45. Vitte C, Ishii T, Lamy F, Brar D, Panaud O: **Genomic paleontology provides evidence for two distinct origins of Asian rice (Oryza sativa L.).** *Mol Gen Genomics* 2004, **272**:504–511.
46. Paux E, Faure S, Choulet F, Roger D, Gauthier V, Martinant J-P, Sourdille P, Balfourier F, Le Paslier M-C, Chauveau A, Cakir M, Gandon B, Feuillet C: **Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat.** *Plant Biotechnol J* 2010, **8**:196–210.
47. Basten CJ, Weir BS, Zeng ZB: *QTL Cartographer: Version 1.17.* Raleigh: Department of Statistics, North Carolina State University, North Carolina; 2003.
48. Garsmeur O, Charron C, Bocs S, Jouffe V, Samain S, Couloux A, Droc G, Zini C, Glaszmann J-C, Van Sluys M-A, D'Hont A: **High homologous gene conservation despite extreme autopolyploid redundancy in sugarcane.** *New Phytol* 2011, **189**:629–642.
49. Besse P, Taylor G, Carroll B, Berding N, Burner D, McIntyre CL: **Assessing genetic diversity in a sugarcane germplasm collection using an automated AFLP analysis.** *Genetica* 1998, **104**:143–153.
50. Lima MLA, Garcia AAF, Oliveira KM, Matsuoka S, Arizono H, Souza CL Jr, Souza AP: **Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum* spp.).** *Theor Appl Genet* 2002, **104**:30–38.
51. Aitken KS, Li JC, Jackson P, Piperidis G, McIntyre CL: **AFLP analysis of genetic diversity within *Saccharum officinarum* and comparison with sugarcane cultivars.** *Aust J Agric Res* 2006, **57**:1167–1184.
52. Selvi A, Nair NV, Noyer JL, Singh NK, Balasundaram N, Bansal KC, Koundal KR, Mohapatra T: **AFLP analysis of the phenetic organization and genetic diversity in the sugarcane complex, *Saccharum* and *Erianthus*.** *Gen Res Crop Evol* 2006, **53**:831–842.
53. Hoarau J-Y, Offmann B, D'Hont A, Risterucci A-M, Roques D, Glaszmann J-Ch, Grivet L: **Genetic dissection of a modern sugarcane cultivar**

(*Saccharum* spp.). I. Genome mapping with AFLP markers. *Theor Appl Genet* 2001, **103**:84–97.
54. Pastina MM, Malosetti M, Gazaffi R, Mollinari M, Margarido GRA, Oliveira KM, Pinto LR, Souza AP, van Eeuwijk FA, Garcia AAF: **A mixed model QTL analysis for sugarcane multiple-harvest-location trial data.** *Theor Appl Genet* 2012, **124**:835–849.
55. Jannoo N, Grivet L, David J, D'Hont A, Glaszmann J-C: **Differential chromosome pairing affinities at meiosis in polyploid sugarcane revealed by molecular markers.** *Heredity* 2004, **93**:460–467.
56. Queen RA, Gribbon BM, James C, Jack P, Flavell AJ: **Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat.** *Mol Gen Genomics* 2004, **271**:91–97.
57. Syed NH, Sorensen AP, Antonise R, van de Wiel C, van der Linden CG, van 't Westende W, Hooftman DAP, den Nijs HCM, Flavell AJ: **A detailed linkage map of lettuce based on SSAP, AFLP and NBS markers.** *Theor Appl Genet* 2006, **112**:517–527.
58. Tam SM, Causse M, Garchery C, Burck H, Mhiri C, Grandbastien M-A: **The distribution of copia-type retrotransposons and the evolutionary history of tomato and related wild species.** *J Evol Biol* 2007, **20**:1056–1072.
59. Bennetzen JL: **Transposable element contributions to plant gene and genome evolution.** *Plant Mol Biol* 2000, **42**:251–269.
60. Dooner HK, He L: **Maize genome structure variation: Interplay between retrotransposon polymorphisms and genic recombination.** *Plant Cell* 2008, **20**:249–258.
61. Kalendar R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH: **Analysis of plant diversity with retrotransposon-based molecular markers.** *Heredity* 2011, **106**:520–530.
62. Da Silva JAG, Sorrells ME, Burnquist WL, Tanksley SD: **RFLP linkage map and genome analysis of *Saccharum spontaneum*.** *Genome* 1993, **36**:782–791.
63. Wei X, Jackson PA, McIntyre CL, Aitken KS, Croft B: **Associations between DNA markers and resistance to diseases in sugarcane, and effects of population substructure.** *Theor Appl Genet* 2006, **114**:155–164.
64. Butterfield MK: *Marker assisted breeding in sugarcane: a complex polyploid. PhD Thesis*. Matieland, South Africa: University of Stellenbosch, Department of Agronomy, Faculty of AgriSciences; 2007.
65. McNeil MD, Hermann S, Jackson PA, Aitken KS: **Conversion of AFLP markers to high-throughput markers in a complex polyploid, sugarcane.** *Mol Breeding* 2011, **27**:395–407.
66. Souza GM, Berges H, Bocs S, Casu R, D'Hont A, Ferreira JE, Henry R, Ming R, Potier B, Van Sluys M-A, Vincentz M, Paterson AH: **The Sugarcane Genome Challenge: Strategies for Sequencing a Highly Complex Genome.** *Tropical Plant Biol* 2011, **4**:145–156.
67. Murray MG, Thompson WF: **Rapid isolation of high molecular weight plant DNA.** *Nucleic Acids Res* 1980, **8**:4321–4325.
68. 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 Res* 1995, **23**:4407–4414.
69. Creste S, Tulmann-Neto A, Figueira A: **Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining.** *Plant Mol Biol Rep* 2001, **19**:299–306.
70. Pinto LR, Oliveira KM, Ulian EC, Garcia AAF, Souza AP: **Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats.** *Genome* 2004, **47**:795–804.
71. Pinto LR, Oliveira KM, Marconi T, Garcia AAF, Ulian EC, Souza AP: **Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs.** *Plant Breed* 2006, **125**:378–384.
72. Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas SM, Falco MC, Burnquist W, Garcia AAF, Souza AP: **Characterization of new polymorphic functional markers for sugarcane.** *Genome* 2009, **52**:191–209.
73. Marconi TG, Costa EA, Miranda HRCAN, Mancini MC, Cardoso-Silva CB, Oliveira KM, Pinto LR, Mollinari M, Garcia AAF, Souza AP: **Functional markers for gene mapping and genetic diversity studies in sugarcane.** *BMC Res Notes* 2011, **4**:264.
74. Maccheroni W, Jordão H, de Gaspari R, de Moura GL, Matsuoka S: **Development of a dependable microsatellite-based fingerprinting system for sugarcane.** *Sugar Cane Intl* 2009, **27**:8–13.
75. *National Center for Biotechnology Information*. http://www.ncbi.nlm.nih.gov/genbank/.
76. *Alellyx Applied Genomics*. http://www.alellyx.com/.
77. van der Linden CG, Wouters DCAE, Mihalka V, Kochieva EZ, Smulders MJM, Vosman B: **Efficient targeting of plant disease resistance loci using NBS profiling.** *Theor Appl Genet* 2004, **109**:384–393.
78. Hanai LR, Santini L, Camargo LEA, Fungaro MEP, Gepts P, Tsai SM, Vieira MLC: **Extension of the core map of common bean with EST-SSR, RGA, AFLP, and putative functional markers.** *Mol Breeding* 2010, **25**:25–45.
79. Kumar A, Bennetzen JL: **Plant retrotransposons.** *Annu Rev Genet* 1999, **33**:479–532.
80. *Standard nomenclature for AFLP® primers*. http://www.ncbi.nlm.nih.gov/genomes/PLANTS/AFLP_tbl.html.
81. *The R Project for Statistical Computing*. http://www.r-project.org.
82. Margarido GRA, Garcia AAF, Souza AP: **OneMap: software for genetic mapping in outcrossing species.** *Hereditas* 2007, **144**:78–79.
83. Wu RL, Ma CX, Wu SS, Zeng ZB: **Linkage mapping of sex-specific differences.** *Genet Res* 2002, **79**:85–96.
84. Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF: **Evaluation of algorithms used to order markers on genetic maps.** *Heredity* 2009, **103**:494–502.
85. Kosambi DD: **The estimation of map distances from recombination values.** *Ann Eugen* 1943, **12**:172–175.
86. Voorrips RE: **MapChart: Software for the graphical presentation of linkage maps and QTLs.** *J Hered* 2002, **93**:77–78.
87. Echt CS, Saha S, Krutovsky KV, Wimalanathan K, Erpelding JE, Liang C, Nelson CD: **An annotated genetic map of loblolly pine based on microsatellite and cDNA markers.** *BMC Genet* 2011, **12**:17.
88. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors.** *Proc Natl Acad Sci U S A* 1977, **74**:5463–5467.
89. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using Phred.** *I. Accuracy assessment. Genome Res* 1998, **8**:175–185.
90. *National Center for Biotechnology Information*. http://blast.ncbi.nlm.nih.gov/Blast.cgi.

doi:10.1186/1471-2156-13-51
Cite this article as: Palhares *et al.*: A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. *BMC Genetics* 2012 **13**:51.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central