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

**Mapeamento genético de marcadores AFLP e de retrotransposons em
cana-de-açúcar (*Saccharum* spp.)**

Alessandra Carolina Palhares

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

**Piracicaba
2010**

Alessandra Carolina Palhares
Bióloga

Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar
(*Saccharum* spp.)

Orientadora:
Prof^a Dra. MARIA LUCIA CARNEIRO VIEIRA

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

Piracicaba
2010

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

Palhares, Alessandra Carolina

Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar
(*Saccharum* spp. / Alessandra Carolina Palhares. - - Piracicaba, 2010.
97 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2010.
Bibliografia.

1. Cana-de-açúcar 2. Genomas 3 Mapeamento genético 4. Marcador genético
Polimorfismo. Título

5

CDD 633.61
P161m

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

Ao Deus de bondade sem fim

Aos meus pais Antonio e Antonieta

A Aline e Ana Paula

Sentido da minha vida...

DEDICO

Agradecimentos

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), particularmente ao Departamento de Genética, pela rica oportunidade.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa concedida para a realização do curso de Mestrado.

À Professora Dra. Maria Lucia Carneiro Vieira, pela orientação, apoio, ensinamentos e confiança ao me propor o desafio que constituiu este trabalho.

Aos Professores Dr. Antonio Augusto Franco Garcia e Dra. Marie-Anne Van Sluys, pelas valiosas colaborações, que me serviram como fonte de aprendizado, força e inspiração.

À empresa CanaVialis S.A., pela proposta de parceria, em especial ao Dr. Walter Maccheroni, e também ao Hamilton Jordão, Rafaela De Gaspari e Fernanda Ferrari, pela amizade e contribuição.

Ao Dr. Luiz Ricardo Hanai, que a todo o momento me ajudou no desenvolvimento deste trabalho, através de seus conhecimentos técnicos, científicos e compartilhando complicados raciocínios sobre a genética de cana-de-açúcar.

Ao Dr. Douglas Silva Domingues que, mesmo distante, sempre se dispôs a cooperar, tirar dúvidas e aconselhar (através de longos e-mails).

Ao Carlos Alberto de Oliveira, pela amizade e pelo importante auxílio técnico durante a realização do trabalho.

Aos colegas do Laboratório de Genética e Estatística da ESALQ/USP Marcelo e Gabriel, por terem colaborado nas análises e me auxiliado a trabalhar no *R* e no *OneMap*.

Aos colegas do GaTELab (IB/USP) Guilherme, Kleber e Sílvia, por ajudarem no preparo e no sequenciamento dos fragmentos para validação.

Aos amigos e companheiros do Laboratório de Biologia Celular e Molecular de Plantas Anselmo, Augusto, Carla, Endson, Guilherme, Helen, Larissa, Luciane, Luciano, Marcelo, Patrícia e Taislene, pela socialização do conhecimento e descontração: cada um com seu modo de ser.

Agradeço de coração.

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	11
1 INTRODUÇÃO.....	13
2 REVISÃO DE LITERATURA.....	15
2.1 Aspectos econômicos da cana-de-açúcar	15
2.2 Classificação botânica e a origem da cana-de-açúcar cultivada	16
2.3 A complexidade do genoma da cana-de-açúcar	18
2.4 Marcadores moleculares e sua aplicação no mapeamento genético da cana-de-açúcar	20
2.4.1 Dosagens alélicas e os padrões de segregação em autopoliploídes.....	22
2.4.2 Mapas genéticos em cana-de-açúcar	25
2.5 Elementos transponíveis.....	28
3 MATERIAL E MÉTODOS	33
3.1 Material vegetal	33
3.2 Extração e quantificação de DNA	33
3.3 Genotipagem com marcadores moleculares	35
3.3.1 Desenvolvimento de marcadores AFLP.....	35
3.3.2 Desenvolvimento de marcadores direcionados a sequências de retrotransposons.....	37
3.3.3 Eletroforese e visualização dos locos.....	43
3.4 Validação molecular dos marcadores direcionados a sequências de retrotransposons	44
3.5 Análise dos dados.....	46
3.5.1 Análise do polimorfismo molecular entre os genitores ‘IAC66-6’ e ‘TUC71-7’	46
3.5.2 Análises de segregação dos locos	47
3.6 Construção do mapa de ligação.....	47
4 RESULTADOS E DISCUSSÃO	49

4.1 Testes e seleção de combinações enzimas de restrição/ <i>primers</i>	49
4.2 Análises da segregação dos locos	56
4.3 Validação molecular dos marcadores direcionados a sequências de retrotransposons	61
4.4 Mapa de ligação	64
5 CONCLUSÕES	77
REFERÊNCIAS	79
ANEXOS	89

RESUMO

Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar (*Saccharum* spp.)

No presente trabalho, AFLPs e marcadores baseados em retrotransposons foram utilizados para a construção de um mapa de ligação integrado em cana-de-açúcar. Dois retrotransposons encontrados no genoma da cana-de-açúcar, denominados *SURE* e *Garapa*, foram estudados. Os princípios da técnica *NBS-profiling* foram usados para gerar marcadores direcionados às sequências desses retrotransposons. Os marcadores foram analisados numa população F₁ de cana-de-açúcar, composta por 188 indivíduos, oriunda do cruzamento entre os genitores 'IAC66-6' e 'TUC71-7'. O mapa integrado foi construído usando-se o *software* OneMap, especialmente desenvolvido para mapear espécies não endogâmicas. Excelentes padrões de AFLP e de marcas direcionadas ao retrotransposon *SURE* foram obtidos; entretanto, para o *Garapa*, ainda são necessários ajustes na técnica. Um total de 600 marcadores de dose única foi obtido a partir de 22 combinações de enzimas de restrição/*primers* de AFLP e seis combinações otimizadas para amplificar marcas direcionadas ao retrotransposon *SURE*. Construiu-se um mapa com 107 grupos de ligação, com tamanho de 4.316,5 cM e densidade de 8,74 cM/marcador. O mapeamento dos marcadores derivados do retrotransposon *SURE* revelou que esse elemento não está uniformemente distribuído nos grupos de ligação e confirmou o seu baixo número de cópias no genoma da cana, conforme foi sugerido na literatura.

Palavras-chave: Cana-de-Açúcar; Retrotransposons; AFLP; Marcadores moleculares; Mapa de ligação

ABSTRACT

Genetic mapping of AFLP and retrotransposon-derived markers in sugarcane (*Saccharum* spp.)

In the present study, AFLPs and retrotransposon-based markers were used for the construction of an integrated linkage map of sugarcane. Two retrotransposons described in the sugarcane genome, named *SURE* and *Garapa* were studied. The principles of NBS-profiling technique were used to generate markers based on these retrotransposon sequences. The markers were analyzed in a F₁-population, composed of 188 individuals, derived from a single cross between the 'IAC66-6' and 'TUC71-7' parents. The integrated genetic map was constructed using the software OneMap, specially designed for mapping outcrossing species. Excellent gel profiles of AFLP and retrotransposon-derived markers were obtained; however, for the *Garapa* element, technical adjustments are still needed. A total of 600 single-dose markers were obtained from 22 AFLP restriction enzyme/primer combinations and six combinations optimized to amplify the *SURE*-based markers. A map with 107 linkage groups was constructed, spanning 4,316.5 cM, with a marker density of 8.74 cM. Mapping of *SURE*-based markers revealed that this element is not uniformly distributed across the linkage groups, and confirmed its low copy number in the sugarcane genome, as suggested in the literature.

Key words: Sugarcane; Retrotransposons; AFLP; Molecular markers; Linkage map

1 INTRODUÇÃO

O Brasil é um país expoente na produção mundial de cana-de-açúcar e de alguns de seus derivados, com destaque para o açúcar e o álcool. Sendo componente essencial da economia nacional ao longo de quatro séculos, a cana-de-açúcar está se tornando matéria-prima de aproveitamento ilimitado e com alto potencial de valor agregado (ANUÁRIO BRASILEIRO DA CANA-DE-AÇÚCAR, 2008).

O contínuo desenvolvimento do setor sucroalcooleiro depende da obtenção de novas variedades que atendam às necessidades do mercado e garantam aumento do lucro às empresas do ramo. São os programas de melhoramento genético da cana-de-açúcar, que atentos às novas demandas, têm sido responsáveis pela mudança no perfil varietal dos clones atuais, por meio de estratégias de cruzamento e seleção diferenciadas (LANDEL; SILVA, 2004).

A cana-de-açúcar é, provavelmente, a cultura mais complexa do ponto de vista genômico em que o melhoramento genético tem sido aplicado. O genoma das cultivares modernas é altamente poliplóide, aneuplóide, com variável número de cromossomos e de origem multiespecífica (GRIVET; ARRUDA, 2002), resultado de cruzamentos e de um processo histórico de melhoramento da cultura, conhecido como “nobilização”. Além disso, o sistema de propagação vegetativa, comumente aplicado, colabora com a manutenção dessa complexidade.

Recentemente, novas ferramentas de análises genéticas têm se tornado acessíveis, o que tem possibilitado refinar o entendimento sobre a complexidade genética e genômica da cana-de-açúcar. A disponibilidade de marcadores moleculares, usados em estudos filogenéticos e para a construção de mapas genéticos tem esclarecido, pelo menos parcialmente, a relação entre espécies de *Saccharum* e a arquitetura genética das cultivares.

Mapas genéticos saturados constituem poderosas ferramentas para a determinação dos efeitos de locos envolvidos no controle de caracteres quantitativos de importância econômica e para a seleção assistida por marcadores, além de possibilitar estudos de herança, a localização de genes, a clonagem gênica e estudos de sintenia (JONES; OUGHAN; THOMAS, 1997; RIPOL et al., 1999). Entretanto, no caso da cana-de-açúcar, a saturação dos mapas é uma tarefa desafiante, e tem exigido abordagens estatísticas para a incorporação de marcadores com os variados padrões de segregação existentes.

Análises sobre o conteúdo genômico da cana-de-açúcar também se tornaram evidentes, como consequência dos esforços para o desenvolvimento e a catalogação de sequências expressas para a cultura, em grande parte detectadas por Vettore et al. (2003). A partir da disponibilidade dessas ESTs, análises sobre o repertório gênico foram realizadas por uma ampla gama de autores, por exemplo, pelo recente trabalho de Papini-Terzi et al. (2009), que encontraram genes diferencialmente expressos envolvidos com o teor de sacarose.

O entendimento sobre a genômica da cultura da cana envolveu também a análise de elementos transponíveis (ROSSI; ARAÚJO; VAN SLUYS, 2001), que correspondem a mais de 50% do genoma das gramíneas. Diante da abundância e diversidade de elementos transponíveis expressos na cana-de-açúcar, novos estudos têm focado na descoberta e caracterização desses elementos, uma vez que eles apresentam importante papel na manutenção e diversificação dos genomas. Nesse contexto, dois elementos do tipo retrotransposons foram recentemente isolados e caracterizados em cana-de-açúcar por Domingues (2009), a partir de dados do projeto SUCEST (VETTORE et al., 2003). Estes retrotransposons apresentam padrões distintos de amplificação no genoma e foram denominados *SURE* (SUGarcane REtrotransposon) e *Garapa*.

Frente aos conhecimentos atualmente disponíveis, a integração das informações genéticas e genômicas se faz necessária na tentativa de esclarecer o complicado genoma da cana-de-açúcar. Assim, o presente trabalho teve como objetivos (1) obter e mapear marcadores moleculares do tipo AFLP, que revelam um alto número de locos polimórficos e, portanto, fornecem um grande número de marcas de DNA, necessário para estabelecer mapas de ligação em cana-de-açúcar usando populações F_1 segregantes, e (2) desenvolver marcadores direcionados a sequências dos retrotransposons *SURE* e *Garapa* para serem alocados no mapa de ligação, de modo a complementar os estudos de Domingues (2009).

2 REVISÃO DE LITERATURA

2.1 Aspectos econômicos da cana-de-açúcar

A cana-de-açúcar é uma importante cultura de regiões tropicais e subtropicais do mundo, sendo plantada em cerca de 20 milhões de hectares (MENOSSE et al., 2008) em mais de 90 países.

Cultivada, principalmente, para a obtenção de sacarose, contribui com 70% da produção de açúcar do mundo (LAKSHMANAN et al., 2005). Em alguns países, como o Brasil e a Índia, a cultura ocupa grandes áreas, sendo que parte de sua biomassa também é usada para a produção de etanol, utilizado como fonte de energia renovável (GRIVET; ARRUDA, 2002).

No Brasil, a indústria da cana-de-açúcar teve forte aumento a partir de 1975, com a criação do Pró-Álcool pelo governo federal, como estratégia para lidar com a crise do petróleo. A partir desse momento, os programas de melhoramento investiram em tecnologias de ponta, aplicados tanto na agricultura como na indústria, tornando a cana-de-açúcar uma cultura bem estabelecida no Brasil (MATSUOKA; FERRO; ARRUDA, 2009).

No país, a atual área plantada com cana-de-açúcar é superior a 7,5 milhões de hectares, o que corresponde a 3,5% da área agricultável disponível. O país atingiu em 2008/09 uma produção de 560 milhões de toneladas, com uma produção de açúcar e de etanol de 32 milhões de toneladas e de 27 bilhões de litros de álcool, respectivamente (JORNAL DA CANA, 2010).

O Estado de São Paulo apresenta destaque no setor, produzindo cerca de 346 milhões de toneladas de cana (safra 2008/09) em uma área de 4,45 milhões de hectares, com uma produção de 19 milhões de toneladas de açúcar e 11 bilhões de litros de álcool, o que corresponde a cerca de 60% tanto da produção nacional de açúcar como de álcool (UNICA, 2010).

O Brasil ainda é um país modelo no desenvolvimento e uso comercial da cana-de-açúcar como biocombustível e para a geração de energia elétrica a partir de sua biomassa. Os variados usos possíveis da cana-de-açúcar como matéria-prima têm gerado aumento do interesse de produtores de vários países pela cultura (LAM et al., 2009).

Estimativas revelam que o atendimento das demandas interna e externa de açúcar e de etanol resultará na necessidade de uma produção de cana-de-açúcar de 685 milhões de toneladas

de cana por ano até 2012/2013 (MACEDO, 2007), um incremento de 20% sobre a produção atual, o que reforça a necessidade do contínuo desenvolvimento do setor e de pesquisas aplicadas à cultura.

2.2 Classificação botânica e a origem da cana-de-açúcar cultivada

A classificação botânica da cana-de-açúcar foi descrita por Cronquist em 1981: a cana pertence à divisão Magnoliophyta, classe Liliopsida, ordem Cyperales, família Poaceae, tribo Andropogoneae, subtribo Saccharininae e gênero *Saccharum* (SCARPARI; BEAUCLAIR, 2008). A família Poaceae também é conhecida como a família das gramíneas, sendo importante porque nela se incluem cereais, tais como milho, trigo, arroz e sorgo, bem como muitas outras culturas forrageiras (TZVELEV, 1989).

O gênero *Saccharum* se divide em seis espécies, sendo que duas são silvestres (*S. spontaneum* e *S. robustum*) e quatro existem apenas em cultivo (*S. officinarum*, *S. barberi*, *S. sinense* e *S. edule*). As duas espécies silvestres e três das espécies cultivadas (com exceção de *S. edule*) são interfertéis. A partir do cruzamento entre elas, se têm produzido os híbridos interespecíficos que constituem as modernas variedades comerciais (IRVINE, 1999).

O gênero *Saccharum*, juntamente com os gêneros *Erianthus*, *Sclerostachya*, *Narenga* e *Miscanthus*, formam um grupo de intercruzamento muito próximo que é denominado “Complexo *Saccharum*”. Este complexo está envolvido na domesticação e formação da cana-de-açúcar, em regiões do continente asiático, no início do processo de domesticação (MATSUOKA; GARCIA; ARIZONO, 2005). A caracterização morfológica dos gêneros do Complexo *Saccharum*, bem como a chave taxonômica para diferenciar as espécies de *Saccharum* e também de *Erianthus* encontram-se descritas no trabalho de Amalraj e Balasundaram (2006).

Com relação à origem da cana-de-açúcar, Daniels et al. (1975) sugerem que ocorreram hibridações entre *Erianthus*, *Sclerostachya* e *Miscanthus*, em variáveis níveis de ploidia e em várias localizações geográficas na Índia, as quais produziram *S. officinarum*, *S. sinense* e *S. barberi*, plantas com variada capacidade de acumular sacarose. Segundo os mesmos autores, *S. officinarum* foi o produto mais eficiente derivado de combinações particularmente favoráveis entre os genomas.

Entretanto, D'Hont, Paulet e Glazmann (2002) investigaram a origem da cana-de-açúcar chinesa e do norte da Índia (*S. barberi* e *S. sinense*, respectivamente) utilizando marcadores RFLP e sondas de GISH (*Genome In Situ Hybridization*), preparadas a partir de DNA genômico de *S. officinarum* e *S. spontaneum*. Como resultado, eles verificaram que *S. barberi* e *S. sinense* são provenientes de hibridações interespecíficas entre *S. officinarum* e *S. spontaneum*, e que nenhum outro gênero está envolvido em sua formação.

Seguindo o mesmo ponto de vista, Grivet et al. (2004) defendem a idéia de que a cana-de-açúcar evoluiu de uma linhagem específica do gênero *Saccharum*, independente dos demais gêneros (*Erianthus* e *Miscanthus*). Neste contexto, as canas cultivadas muito provavelmente emergiram das espécies silvestres de *Saccharum*, sendo as introgressões com outros gêneros pouco frequentes. Os autores defendem a hipótese comumente aceita de Brandes (1958), na qual *S. officinarum* foi domesticada a partir de *S. robustum* na Nova Guiné e, juntamente com as migrações humanas, foi dispersa no Pacífico e no continente asiático, onde hibridizou-se com *S. spontaneum* ali existentes originando as cultivares chinesas e do norte da Índia.

Os primeiros cruzamentos realizados pelo homem iniciaram-se em 1880 e envolveram as espécies *S. officinarum* e *S. spontaneum*, sendo a primeira conhecida como cana “nobre”, caracterizando-se por apresentar folhas largas, talo grande e espesso, alto teor de açúcar e baixo conteúdo de fibras. Esta cana, que foi introduzida como “nobre”, também é conhecida como crioula e é suscetível a várias doenças (BREMER, 1961a). Apesar de *S. spontaneum* apresentar baixo conteúdo de açúcar, possui muitas outras características úteis para o melhoramento, incluindo vigor vegetativo, resistência a estresses abióticos e a vários parasitas. Esta espécie tem ampla distribuição geográfica, podendo ser encontrada do Japão à Nova Guiné e do Mediterrâneo à África (MUKHERJEE, 1957).

Em um século de melhoramento, foi realizada uma série de retrocruzamentos de híbridos interespecíficos com *S. officinarum*, usada como genitor feminino. Este processo de introgressão, chamado “nobilização”, providenciou o maior avanço no melhoramento da cana-de-açúcar, solucionou uma série de problemas relativos a doenças da cana e também providenciou benefícios adicionais, como aumento da colheita e adaptação a estresses abióticos (ROACH, 1972).

Desde então, os melhoristas têm se baseado no intercruzamento das melhores cultivares existentes, seguido pelas fases de seleção e ensaios de competição varietal, nos quais procura-se

aliar a rusticidade e a resistência a moléstias às boas qualidades de rendimento em açúcar (LANDEL; SILVA, 2004).

Assim, as cultivares atuais são clones elites propagados vegetativamente e sua terminologia taxonômica atual é *Saccharum* spp., já que não se cultivam comercialmente canas que não sejam frutos de intercruzamentos (SCARPARI; BEAUCLAIR, 2008).

2.3 A complexidade do genoma da cana-de-açúcar

O genoma da cana-de-açúcar é altamente complexo: poliplóide, aneuplóide e com variável número de cromossomos. Além disso, trata-se de um genoma extremamente grande, o que dificulta a compreensão de sua arquitetura genética. Por exemplo, o conteúdo de DNA de uma célula somática de *S. officinarum* é 7440 Mpb. Como esta espécie é octaplóide, o tamanho do complemento haplóide é 8 vezes menor, cerca de 930 Mpb. Este valor é comparável ao genoma de uma espécie diplóide, como o sorgo, que apresenta cerca de 760 Mpb (GRIVET; ARRUDA, 2002).

Com relação ao número de cromossomos das espécies, *S. spontaneum* apresenta a variação de $2n=40$ a $2n=128$, sendo os genótipos com $2n=64$, 80, 96, 112 e 128 os mais frequentes. Para *S. officinarum*, foi constatado que 92% de seus genótipos apresentam $2n=80$. Já para *S. robustum*, verificou-se variação de $2n=60$ até $2n=200$, mas os valores $2n=60$ e $2n=80$ são os mais comuns. Em *S. edule* são frequentes os genótipos com $2n=60$, 70 e 80 cromossomos. Por fim, *S. sinense* apresenta variação de $2n=111$ a $2n=120$ e *S. barberi* de $2n=81$ a $2n=124$ (BREMER, 1961a; IRVINE, 1999).

Com o objetivo de determinar o número básico de cromossomos das espécies *S. officinarum*, *S. spontaneum* e *S. robustum*, D'Hont et al. (1998) realizaram hibridizações *in situ* de sondas de DNA ribossômico (rDNA) 5S e 45S em vários citotipos de cada espécie. Para ambas as sondas, um total de 8 sítios de hibridização foram identificados para *S. officinarum* ($2n=80$), *S. robustum* ($2n=80$) e *S. spontaneum* ($2n=64$), 6 sítios em *S. robustum* ($2n=60$), e 10 sítios de hibridização em *S. spontaneum* ($2n=80$). Esses resultados foram interpretados como presença dos locos de rDNA de uma família multigênica, repetidos várias vezes, em cada genoma, possibilitando definir o número básico de $x=10$ para *S. officinarum* e *S. robustum* e $x=8$ para *S. spontaneum*.

Dessa forma, ficou evidenciado o alto grau de ploidia nas espécies: *S. officinarum* é octaplóide, *S. robustum* é hexaplóide (em $2n=60$) e octaplóide (em $2n=80$) e, em *S. spontaneum*, o nível de ploidia varia entre 5 a 16 (D'HONT et al., 1998).

Estudando o complemento cromossômico do clone haplóide de *S. spontaneum* 'AP 85-361' ($2n=32$), Ha et al. (1999) confirmaram o número básico $x=8$ para a espécie. Através de métodos de análises de imagens, os autores conseguiram visualizar o padrão de condensação de cada um dos cromossomos, o que possibilitou estabelecer um idiograma para a espécie. No mesmo estudo, determinou-se que *S. spontaneum* é um autopoliploide.

Em se tratando de variedades modernas, a determinação segura do número de cromossomos é dificultada pela presença de elevado número cromossômico ($2n= 100$ a 130) e de mosaicismo. Silvarolla (1989) analisou o número de cromossomos de cinco variedades comerciais de cana-de-açúcar: 'POJ2878', 'Co419', 'NA56-79', 'IAC68-12' e 'IAC68-114'. A partir de diversos pré-tratamentos com inibidores de fuso testados em subclones dessas variedades, regenerados *in vitro*, a morfologia dos cromossomos foi visualizada com clareza. Constatou-se uma média de 114 cromossomos nessas variedades e o predomínio de estabilidade cromossômica. Posteriormente, Silvarolla e Aguiar-Perecin (1994) descreveram uma metodologia que permite a análise de metáfases mitóticas e a investigação do número de cromossomos nas variedades modernas, o que possibilita estudar de forma mais clara a citogenética desses genótipos.

A composição cromossômica da cultivar 'R570' ($2n=115$) foi analisada por D'Hont et al. (1996) através de GISH. Determinou-se que 80% dos cromossomos desta variedade são provenientes de *S. officinarum*, 10% de *S. spontaneum* e os 10% restantes constituem cromossomos recombinantes entre essas espécies. Posteriormente, tais proporções também foram estimadas por hibridização *in situ* nas cultivares 'NCo376' ($2n=112$) (PIPERIDIS; D'HONT, 2001), 'My5514', 'B42231' e 'C236-51' (CUADRADO, 2004); a conclusão geral foi que cerca de 10 a 20% são cromossomos de *S. spontaneum* e que os cromossomos recombinantes constituem de 5 a 17% do genoma da cana.

Bremer (1961b) investigou a causa do elevado número de cromossomos existentes nas variedades de cana-de-açúcar e descobriu que em *S. officinarum* ocorre endoduplicação dos cromossomos durante a megasporogênese. Isso tem implicações quando se utiliza esta espécie

como genitor feminino e esclarece a presença do variado número de cromossomos nos híbridos, além da aneuploidia, que é comum (PRICE, 1963).

No genoma das cultivares modernas evidencia-se a coexistência de duas organizações cromossômicas distintas, devido às diferenças no número básico de cromossomos de *S. officinarum* e *S. spontaneum*. Porém, até o momento, o comportamento dos cromossomos durante os processos de introgressão até a formação dessas cultivares é pouco conhecido, por causa da poliploidia e elevada homeologia dos cromossomos de *Saccharum* (CUADRADO et al., 2004).

Nessas cultivares, o pareamento dos cromossomos durante a meiose é difícil de ser interpretado uma vez que os mesmos podem ter afinidade com seu próprio par ou com os seus homeólogos, formando univalentes, bivalentes ou multivalentes. Isso afeta diretamente o padrão de segregação e, conseqüentemente, induz à variações no número total de cromossomos transmitidos. Na ‘R570’, como exemplo, os estudos sobre a distribuição dos cromossomos na meiose apontou a predominância de bivalentes (JANNOO et al., 2004).

A introgressão aumentou consideravelmente a complexidade genética das cultivares, através dos intercruzamentos de clones de híbridos e seus derivados e, por esse motivo, o nível de heterozigosidade entre as variedades é alto (LU et al., 1994).

De acordo com Cuadrado et al. (2004), a propagação vegetativa dos clones de cana-de-açúcar permitiu o acúmulo de alterações cromossômicas numéricas e estruturais provocadas pelas hibridações interespecíficas, que acabaram “modelando” o genoma da cultura. Tal técnica (através da micropropagação *in vitro*) é comumente utilizada comercialmente para a obtenção de mudas em larga escala, visando à ocupação da área plantada, que é extensa, pelos novos genótipos lançados pelos programas de melhoramento.

2.4 Marcadores moleculares e sua aplicação no mapeamento genético da cana-de-açúcar

Marcadores moleculares são locos cromossômicos, genes ou sequências de DNA que apresentam segregação mendeliana, podem ter ação dominante ou codominante e se prestam para análises genéticas (FUNGARO; VIEIRA, 2001; BORÉM; VIEIRA; COLLI, 2009). Além disso, são seletivamente neutros, ou seja, quando considerados isoladamente não conferem nenhuma vantagem ou desvantagem ao genótipo (RAMALHO; SANTOS, PINTO, 2004).

As técnicas de manipulação do DNA possibilitaram que um número quase ilimitado de marcadores fosse detectado nos genomas, gerados a partir de variadas metodologias que exploram o polimorfismo de DNA (VIEIRA; VELLO; SILVA FILHO, 2004).

A princípio, técnicas como o RFLP (*Restriction Fragment Length Polymorphism* - 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.

Com a disponibilidade cada vez maior de sequências de DNA nos bancos de dados, marcadores associados a regiões alvo foram usados, como o exemplo do SSAP (*Sequence-Specific Amplification Polymorphism* - WAUGH et al., 1997) e dos RGA (*Resistance Gene Analogs* - VAN DER LINDEN et al., 2004), que são derivados do AFLP, mas utilizam motivos de sequências conservadas como âncora.

Recentemente, marcadores funcionais têm sido desenvolvidos a partir de sequências expressas (ESTs), como no caso dos microssatélites inseridos em genes - ilustrados para cana-de-açúcar nos trabalhos de Oliveira et al. (2007) e Maccheroni et al. (2009) - e dos SNPs (*Single Nucleotide Polymorphisms* - RAFALSKI, 2002) adicionando importantes informações para as análises de natureza genética.

Os marcadores moleculares são úteis, entre outras finalidades, para a construção de mapas de ligação, que consistem na representação linear da distribuição das marcas em grupos de ligação, os quais são construídos, basicamente, colocando-se os marcadores em ordem, indicando a distância relativa entre eles com base nas estimativas das frações de recombinação.

Dada a complexidade genômica da cana-de-açúcar, muitas análises genéticas puderam ser realizadas a partir da disponibilidade de marcadores moleculares usados para a construção de mapas de ligação, que têm auxiliado no entendimento de sua estrutura genética. Entretanto, tanto a análise de marcadores como a construção dos mapas genéticos em cana-de-açúcar são afetados pela complexidade de seu genoma. Por isso, são adotadas estratégias peculiares de análises que serão descritas a seguir.

2.4.1 Dosagens alélicas e os padrões de segregação em autopoliplóides

Em locos heterozigóticos de espécies diplóides, obtém-se, no final da meiose, dois alelos que segregam com igual frequência. Em espécies autopoliplóides, por sua vez, existem vários alelos segregando por causa da ploidia e, conseqüentemente, ocorre a formação de diferentes tipos de gametas (WU et al., 1992).

Nos autopoliplóides, para definir os padrões de herança, considera-se haver pareamento bivalente entre os cromossomos e ausência de dupla redução. Para a definição das frequências gaméticas, são calculadas todas as possibilidades de segregação gamética para os vários alelos de um loco. A partir desse cálculo, é possível determinar as segregações esperadas numa progênie. No entanto, quanto maior o nível de ploidia, maior o número de genótipos esperados na população segregante (WU et al., 2001). Isso é evidente quando se compara um autotetraplóide, como a batata, com poliplóides de nível de ploidia superior, como a cana-de-açúcar (Tabela 1).

Tabela 1 – Frequências gaméticas em autopoliplóides com pareamento bivalente e ausência de dupla redução. Adaptado de Sorrells (1992)

Genótipo do genitor	Tipos de gametas					Frequência gamética
Tetraplóide	aa	Aa	AA			A_ : aa
Aaaa	1	1				1:1
AAaa	1	4	1			5:1
AAAa		1	1			
AAAA			1			
Octaplóide	aaaa	Aaaa	AAaa	AAAa	AAAA	A_ _ _ : aaaa
Aaaaaaaa	1	1				1:1
AAaaaaaa	3	8	3			11:3
AAAaaaaa	1	6	6	1		13:1
AAAAaaaa	1	16	36	16	1	69:1
AAAAAaaa		1	6	6	1	
AAAAAAaa			3	8	3	
AAAAAAAa				1	1	

No caso da cana-de-açúcar, os modelos para análises genéticas disponíveis baseiam-se nas premissas de autoploidia e número básico de cromossomos $x=8$. Sendo assim, os locos podem estar em variadas condições (dosagens), que vão desde nuliplex – quando todos os alelos são nulos (aaaaaaaa), simplex – quando apenas um alelo é informativo (Aaaaaaaa), duplex (AAaaaaaa), triplex (AAAaaaaa) até multiplex (DA SILVA; SORRELLS, 1996).

Assim, são formadas 70 combinações alélicas possíveis, a partir das quais se consegue obter as frequências gaméticas esperadas para os tipos de cruzamentos que envolvem genitores com até quatro doses (Tabela 1).

Para fins de mapeamento genético, são utilizados apenas os marcadores em dose única (MDU), tal como descrito por Wu et al. (1992). Tais marcadores são identificados de duas maneiras: pela sua segregação na proporção 1:1, quando o alelo (banda do gel) está presente em um genitor e ausente no outro (ou seja, Aaaaaaaa x aaaaaaaa) e, também, pela segregação 3:1, quando o alelo está presente em ambos os genitores na condição simplex (Aaaaaaaa x Aaaaaaaa) (Figura 1).

Figura 1 – Esquema da segregação gamética de marcadores em dose única na descendência de dois tipos de cruzamentos. A partir do fenótipo molecular, são mostrados os genótipos dos genitores, os gametas formados e o resultado do cruzamento com as respectivas segregações fenotípicas esperadas

A identificação de marcadores com dosagens superiores são efetuadas com base nas frequências gaméticas apresentadas na tabela 1. Para tal, a partir de marcadores polimórficos entre os genitores, é possível verificar as seguintes proporções de marcas de acordo com a dosagem alélica: 11:3 (duplex) e 13:1 (triplex) e 69:1 (quadruplex). No entanto, as estimativas dessas dosagens e a adição desses marcadores nos mapas de ligação requerem estratégias mais sofisticadas sendo, por isso, pouco utilizados até o momento.

Alguns avanços nos métodos estatísticos para o mapeamento genético em autopoliplóides foram obtidos por Ripol e colaboradores (1999), que apresentaram procedimentos para aprimorar as estimativas de dosagem dos fragmentos, com base na proporção de um dado marcador na progênie. Modelos multinomiais também foram utilizados por esses autores para obter estimativas de máxima verossimilhança e realizar testes subsequentes. Porém, os modelos estatísticos ainda estão em estudo.

As diferenças entre as dosagens alélicas dos genótipos podem ser visualizadas através da intensidade das bandas. Este método foi utilizado por Jannoo et al. (2004) para estudos genéticos com a cultivar francesa 'R570', que foi autofecundada para a obtenção de uma população segregante. Analisando marcas de RFLP que segregaram na proporção de 3:1 em F_1 (ou seja, genótipos que apresentam 0, 1 ou 2 doses do alelo) e adotando o padrão de bandas da cultivar 'R570' como referência, esses autores propuseram a análise de marcas duas a duas, na qual as bandas do gel com maior intensidade foram consideradas como apresentando duas doses do alelo, enquanto as menos intensas como apresentando uma dose. Devido ao grande número de comparações possíveis, a dosagem para todos os alelos de todos os indivíduos pôde ser determinada (Figura 2).

Figura 2 – Modelo de interpretação das dosagens alélicas na progênie com base na intensidade de bandas de RFLP da cultivar 'R570', tomada como padrão. Analisando as marcas duas a duas (como exemplo, x e y), a comparação entre as variadas intensidades em relação às da cultivar possibilitou a detecção de alelos nulos (presentes nos indivíduos 3, 4 e 6), alelos com uma dose (indivíduos 1 e 2), com duas doses (indivíduos 1 e 2) e com dosagem duvidosa (indivíduos 4, 5 e 6). Adaptado de Jannoo et al. (2004)

Entretanto, a viabilidade desse método é discutível, uma vez que os genótipos não são facilmente identificados através de seu fenótipo molecular, e algumas ambiguidades puderam ser notadas nesse trabalho com relação a alelos detectados por mais de uma enzima, levando à determinação de uma dosagem duvidosa.

Atualmente, quando se realiza genotipagens com marcadores moleculares em cana-de-açúcar, somente a presença ou a ausência de bandas em um gel são reconhecidas, permanecendo as incertezas sobre as suas dosagens. Os marcadores são, então, analisados para a hipótese de dose única antes da construção do mapa genético. Independente da natureza do marcador molecular utilizado (dominante ou codominante) os locos só são analisados pela presença ou ausência do alelo, desprezando-se o caráter codominante (SORRELLS, 1992).

2.4.2 Mapas genéticos em cana-de-açúcar

Em plantas cultivadas, para a construção de mapas genéticos são utilizadas populações convencionais de mapeamento, como F_2 , derivadas de retrocruzamentos, e linhagens endogâmicas recombinantes (RILs – *Recombinant Inbred Line*). No entanto, em algumas espécies vegetais há limitações para a obtenção dessas populações, por apresentarem mecanismos de auto-incompatibilidade, como os maracujás, ciclo de vida longo, como o eucalipto, poliploidia exacerbada, como a cana-de-açúcar. Nessas espécies, conhecidas como *outcrossing species*, a estratégia do duplo pseudo-cruzamento teste (do inglês, *two-way pseudo-testcross*) foi adotada para se estabelecer os primeiros mapas de ligação utilizando populações F_1 . A análise de ligação é feita para cada genitor separadamente, o que resulta na construção de mapas individuais (GRATTAPAGLIA; SEDEROFF, 1994). Em cana-de-açúcar, os primeiros mapas de ligação foram construídos com essa estratégia utilizando-se os *softwares* JoinMap (VAN OOIJEN; VOORRIPS, 2001) e Mapmaker (LANDER et al., 1987). Em 2002, Wu et al. propuseram novas ferramentas para análises de marcadores genéticos que permitiram aperfeiçoar as estimativas das fases de ligação e das frações de recombinação, simultaneamente, com base em máxima verossimilhança e estatística Bayesiana. Recentemente, os algoritmos propostos por Wu et al. (2002) foram implantados em um *software* público, denominado OneMap (MARGARIDO; GARCIA; SOUZA, 2007), que permite a construção de mapas de ligação para populações F_1 .

Baseados na estratégia de mapeamento de marcadores em dose única, vários mapas puderam ser construídos para as espécies do gênero *Saccharum* e, também, para as cultivares modernas. Porém, uma grande quantidade de marcadores são requeridos para gerarem mapas genéticos com razoável nível de cobertura do genoma (PIPERIDIS et al., 2008) e uma população composta por um grande número de indivíduos é necessária para se detectar e estimar as fases de ligação e as frações de recombinação, considerando-se a complexidade do genoma e o fato de se mapear em F_1 .

Para *Saccharum officinarum* ($2n=80$), muitos mapas de ligação foram construídos, como demonstrado nos trabalhos de Mudge et al. (1996), Ming et al. (1998) e Ming et al. (2002). O mesmo ocorreu com *S. spontaneum*. Assim, Da Silva e colaboradores (1995) construíram um mapa de ligação para o clone SES208 ($2n=64$), usando marcadores RAPD e RFLP. Os autores obtiveram 64 grupos de ligação e 8 grupos homólogos, conforme o esperado para o genótipo. No entanto, os marcadores não foram uniformemente distribuídos nos grupos de ligação e 7 destes não puderam ser alocados nos grupos homólogos, demonstrando as dificuldades encontradas no mapeamento genético da cana. Neste importante trabalho, marcadores em doses dupla e tripla foram adicionados, basicamente, pela sua associação em um mapa “*framework*” estabelecido por marcadores em dose única, o que permitiu alocar os grupos de ligação nos respectivos grupos homólogos e diagnosticar que o clone comportava-se como um autooctaplóide.

Os marcadores em doses múltiplas adicionam importantes informações nos mapas genéticos de poliplóides. Sua utilização possibilita diminuir o número de marcas necessário para a detecção dos grupos de ligação, porque evita que o mapeamento seja limitado a apenas regiões onde os marcadores em dose única são encontrados (AITKEN et al., 2007; DA SILVA et al., 1995).

Dessa forma, Aitken et al. (2007) utilizaram marcas em dose única e em dose dupla para o mapa de *S. officinarum* ‘IJ76-514’, tomando como base as estimativas de dosagens alélicas descritas por Ripol et al. (1999) e um método para adição de doses múltiplas similar ao de Da Silva et al. (1995). A partir das mesmas informações, Edmé, Glynn e Comstock (2006) estudaram a população filial do cruzamento entre *S. officinarum* ‘Green German’ e *S. spontaneum* ‘IND 81-146’. A segregação de microssatélites foi avaliada para as seguintes segregações mendelianas em cana: simplex, duplex, triplex e quadruplex. Os mapas gerados nesse trabalho mostraram que existe recombinação entre os genomas ‘Green German’ e ‘IND

81-146', mas que as diferenças estruturais existentes, tais como translocações e rearranjos, afetam a esterilidade e a performance dos híbridos, conforme observado na progênie analisada em campo experimental.

Por sua vez, Alwala et al. (2008) construíram mapas para *S. officinarum* 'La Striped' e *S. spontaneum* 'SES147B' com marcadores AFLP, SRAP e TRAP. Neste estudo, foram aproveitadas marcas em dose única, dose dupla e marcas com distorção de segregação. Estas foram usadas para investigar a presença de regiões genômicas com alta propensão para distorção de segregação, ou mesmo para esclarecer a ocorrência de comigração de fragmentos não alélicos, que também pode levar a distorções de segregação.

O primeiro mapa construído para uma variedade moderna foi desenvolvido para a cultivar 'R570' a partir da sua autofecundação (GRIVET et al., 1996). Neste estudo, investigou-se a origem das marcas obtidas, usando sondas de RFLP para comparar os genótipos das espécies ancestrais da cultivar. Foi possível investigar a origem de 146 marcas, sendo 80 delas provenientes de *S. officinarum* e 66 de *S. spontaneum*. Como já era conhecido que 80% do genoma das cultivares modernas é proveniente de *S. officinarum* (D'HONT et al., 1996), concluiu-se que a resolução do mapa foi superior na porção do genoma derivada de *S. spontaneum*.

Posteriormente, o mapa dessa cultivar foi incrementado com dados de AFLP (HOARAU et al., 2001), microssatélites e RGAs (ROSSI et al., 2003). Apesar da elevada quantidade e variedade de marcas disponíveis, muitos grupos de ligação permaneceram dissociados dos grupos homólogos e ilustrou-se a dificuldade de acessar grandes porções de cromossomos, particularmente aquelas herdadas de *S. officinarum*.

Outra importante cultivar moderna é a australiana 'Q165' que, depois de cruzada com 'IJ76-514' (*S. officinarum*), teve vários mapas de ligação construídos em F₁, com elevada variedade de marcadores alocados: AFLP, SNPs, RAF e SSRs (AITKEN; JACKSON; McINTYRE, 2005, McINTYRE et al., 2009).

Novas abordagens têm sido desenvolvidas visando aprimorar os mapas de ligação de cana-de-açúcar, tal como proposto por Garcia et al. (2006) e Margarido, Garcia e Souza (2007). Com base num método fundamentado em estatística Bayesiana e máxima verossimilhança, que estima simultaneamente as fases de ligação e as frações de recombinação (WU; MA, 2002), desenvolveu-se um mapa genético integrado para uma população derivada do cruzamento

simples entre duas cultivares brasileiras pré-comerciais: ‘SP80-180’ x ‘SP80-4966’. Para tanto, foi utilizado o *software* OneMap, desenvolvido especificamente para esse fim. A construção do mapa integrado permitiu aumentar a saturação do mapa de ligação. Ressalta-se a importância do desenvolvimento desse *software* devido ao fato de o mapeamento de espécies de cruzamento ser realizado em F₁, como é o caso da cana-de-açúcar (GARCIA et al., 2006), do eucalipto (BRONDANI et al., 2006), da maçã (IGARASHI et al., 2008) e dos maracujás (OLIVEIRA et al., 2008).

A partir de dados gerados pelo SUCEST, Oliveira et al. (2007) desenvolveram marcadores microssatélites desenvolvidos a partir de sequências expressas (ESTs), posteriormente validados e caracterizados em categorias de processos biológicos e de funções moleculares (OLIVEIRA et al., 2009). Esses marcadores funcionais podem ser usados na seleção assistida por marcadores, se associados a um ou mais genes de interesse agrônômico, aumentando a utilidade do mapeamento genético para os programas de melhoramento.

2.5 Elementos transponíveis

Os elementos genéticos móveis foram descobertos na década de 50 por Barbara McClintock, durante seus estudos sobre os diferentes padrões de coloração dos grãos de milho. Ela observou que tais elementos eram capazes de se mover através do genoma. Após meio século da sua descoberta, se verificou que esses elementos apresentam grande influência no desenvolvimento dos organismos. Atualmente, são reconhecidos como segmentos de DNA que contêm em suas sequências todas as informações necessárias para se mover e replicar em diferentes locais do genoma (BIÉMONT; VIEIRA, 2006).

De acordo com Kazazian Jr. (2004), esses elementos podem ser considerados como agentes da evolução dos genomas, por sua atuação em diferentes vias, classificadas como “destrutivas”, quando envolvem inserções e rearranjos (devido à recombinação homóloga desigual), e “construtivas”, quando agem no reparo da dupla hélice do DNA, na regulação da expressão gênica e em outros processos. Por causa dessas suas variadas funções, alguns autores os classificam como “benignos” ou “malignos” mas, de qualquer modo, suas atividades contribuem para o aumento da diversidade genética nos organismos (BIÉMONT; VIEIRA, 2006).

Grande fração do DNA genômico total dos eucariotos é composta por elementos transponíveis. Nas gramíneas, por exemplo, eles respondem por 50 a 80% do tamanho do genoma, sendo que no milho, que é um organismo modelo para análises do impacto desses elementos na estrutura genômica e na expressão gênica, 2/3 de seu genoma são constituídos por elementos transponíveis (SODERLUND et al., 2009).

Apresentam elevada diversidade nas plantas e dividem-se, basicamente, em duas classes, de acordo com a molécula intermediária de transposição, que pode ser RNA (classe I) ou DNA (classe II). No caso dos elementos da classe I, também conhecidos como retrotransposons, uma molécula de RNA é codificada e uma cópia da sequência do elemento é produzida, o que resulta na sua transposição replicativa. Já nos elementos da classe II, o próprio elemento constitui seu intermediário, sendo retirado de um local para ser inserido em outro (FESCHOTTE; JIANG; WESSLER, 2002).

Por causa da elevada abundância e diversidade desses elementos, Wicker et al. (2007) propuseram um novo sistema de nomenclatura e classificação, capaz de identificá-los e distingui-los. No sistema hierárquico sugerido, é possível unificar as informações para todos os elementos transponíveis dos eucariotos, permitindo ainda se alcançar um consenso entre os vários sistemas de classificação previamente empregados.

Os retrotransposons são classificados em duas classes: LTR retrotransposons, que é aquela em que os retroelementos contêm LTRs (do inglês, *Long Terminal Repeats*); e não-LTR retrotransposons, na qual os elementos não possuem as terminações repetidas. Os LTR retrotransposons ainda são subdivididos em duas superfamílias: *Copia* e *Gypsy*, amplamente estudadas por estarem presentes em todos os eucariotos (KUMAR, 1996).

De modo geral, a estrutura de um retrotransposon completo é formada pelas LTRs e por regiões que contêm domínios codificantes (Figura 3). As LTRs possuem em suas sequências as regiões promotoras e terminadoras associadas com a transposição do elemento. Os domínios codificantes produzem proteínas especificadas por dois importantes genes chamados *gag* e *pol*. Inicialmente, as proteínas codificadas por esses genes são sintetizadas como uma poliproteína que posteriormente é clivada em peptídeos funcionais. Os genes *gag* codificam componentes do complexo molecular que estão associados com a maturação, envolvimento e proteção do RNA intermediário. Os genes *pol* codificam várias enzimas: a protease, que cliva as proteínas precursoras e as envolve em uma partícula de montagem; a integrase, que insere o DNA dupla

fita formado dentro de um novo local no genoma hospedeiro; a transcriptase reversa, que é usada para sintetizar uma fita simples de RNA complementar àquela do retrotransposon e a converte para o DNA dupla-fita; e a ribonuclease H, que degrada os híbridos de DNA-RNA obtidos durante a transposição (BIÉMONT; VIEIRA, 2006, KUMAR; BENNETZEN, 1999).

LTR retrotransposons

Figura 3 – Estrutura dos LTRs retrotransposons. As regiões U3, R e U5 das LTRs contêm sinais para início e término da transcrição. PBS (*Primer Binding Site*) e PPT (*Polypurine Tract*) são sítios de anelamento dos iniciadores da primeira e segunda fita de cDNA, respectivamente. CP: proteínas do capsídeo, PR: protease, RT: transcriptase reversa, RNASE H: ribonuclease H, INT: Integrase. Adaptado de Kumar; Bennetzen (1999)

A técnica mais usada para as análises dos retrotransposons é o SSAP, segundo Flavell et al. (1998). Com base nessa metodologia, muitos estudos de diversidade genética foram realizados em variadas espécies, tais como em tomate e pimenta (TAM et al., 2005), castanha-de-caju (SYED et al., 2005) e *Iris missouriensis* (CORNMAN; ARNOLD, 2009). O SSAP também é utilizado para a genotipagem de populações segregantes e construção de mapas genéticos. Isso foi primeiramente realizado em cevada (WAUGH et al., 1997) e depois em outras espécies, como trigo (QUEEN et al., 2004), alface (SYED et al., 2006) e tomate (TAM et al., 2007). Esse método adiciona informações sobre a distribuição de retrotransposons ao longo dos grupos de ligação (WAUGH et al., 1997).

Em cana-de-açúcar, a análise de uma enorme biblioteca de ESTs (237.954 sequências expressas) construída no Brasil, por iniciativa da FAPESP, em um projeto conhecido por SUCEST (VETTORE et al., 2003), apresentou-se como ponto inicial para determinar seu repertório genômico, permitindo avanços significativos no estabelecimento da sua origem evolutiva e estrutura de seu genoma. Usando dados do SUCEST, Rossi et al. (2001) realizaram uma busca de elementos transponíveis, que até então eram desconhecidos na cana-de-açúcar. Obteve-se um total de 276 clones de cDNA que apresentaram elevada homologia com elementos

previamente descritos no GenBank (NCBI - *National Center for Biotechnology Information*). Esses clones foram subdivididos em 21 diferentes famílias de elementos, embora alguns deles pudessem representar novos elementos, o que requeria mais estudos.

A partir desses resultados, Araújo et al. (2005) selecionaram 68 clones, representativos de 11 famílias de elementos transponíveis, para serem analisados através de estratégias combinadas de *macroarrays*, *northern blot* e transformação. Assim, o padrão de expressão dos elementos foi analisada em quatro tecidos de cana-de-açúcar. Como resultado, foi diagnosticado que duas famílias de elementos, *Mutator* e *Hopscotch*, são altamente representadas no transcriptoma da cana. Além disso, identificou-se que as massas de calo são os tecidos onde os elementos transponíveis são altamente expressos e *Hopscotch* constitui, na verdade, um promotor ativo, embora sob forte controle transcricional.

Ainda utilizando a base de dados do SUCEST, dois LTR retrotransposons foram recentemente descritos em cana-de-açúcar por Domingues (2009) e nomeados de *SURE* (SUGarcane REtrotransposons) e *Garapa*. De acordo com o autor, esses elementos são completos e apresentam-se como duas linhagens diferentes da superfamília *Copia*. Foi verificado que *SURE* e *Garapa* possuem padrões distintos de número de cópias, diversidade e ativação transcricional: enquanto *SURE* apresenta baixo número de cópias (entre 40 a 50), baixa diversidade entre cópias e maior controle transcricional, expressando-se em situações específicas e em tecidos pouco diferenciados, *Garapa* é um elemento com alto número de cópias (mais de 1000), diversidade maior que a do *SURE* e é ativo em diferentes situações fisiológicas.

A distribuição desses retrotransposons também foi analisada por Domingues (2009) através de FISH (*Fluorescence In Situ Hybridization*), porém estudos adicionais são necessários para se conhecer a distribuição destes dois elementos nos cromossomos de cana-de-açúcar. Através da técnica de mapeamento genético, pretendeu-se contribuir, no presente estudo, para o esclarecimento desta questão.

3 MATERIAL E MÉTODOS

3.1 Material vegetal

A população utilizada neste trabalho foi constituída por 188 indivíduos escolhidos aleatoriamente de uma progênie F_1 de 400 indivíduos oriundos do cruzamento simples entre o clone ‘IAC 66-6’ e a variedade ‘TUC 71-7’, utilizados como genitores masculino e feminino, respectivamente.

Os genitores apresentam características contrastantes para a resistência ao carvão (*Sporisorium scitamineum*) e para caracteres de produção. Por exemplo, o genitor ‘IAC 66-6’ apresenta menor teor de sacarose, maior diâmetro de colmos e susceptibilidade ao carvão, enquanto o genitor ‘TUC 71-7’ apresenta maior teor de sacarose, menor diâmetro de colmos e é resistente ao carvão.

Tal cruzamento foi realizado na estação experimental Vale das Flores, em Maceió (AL), pertencente à empresa CanaVialis S.A. e as sementes foram germinadas em caixas de plástico para posterior individualização dos *seedlings*. Este material foi transferido para o Pólo Regional de Tecnologia da CanaVialis em Conchal (SP), onde foi plantado em área experimental.

3.2 Extração e quantificação de DNA

Amostras de folhas de cada genitor e dos 188 indivíduos da população filial foram coletadas, identificadas e armazenadas a $-20\text{ }^{\circ}\text{C}$ até a extração de DNA.

O DNA genômico foi extraído pelo método CTAB, baseado no protocolo de Murray e Thompson (1980), com algumas adaptações. Primeiramente, 2g de tecido foliar (sem a nervura central) foram macerados em nitrogênio líquido e transferidos para um tubo tipo *falcon* de 15 mL, ao qual adicionaram-se 9 mL de tampão de extração 2% [2% de brometo de cetiltrimetilamônio (CTAB), 1,4 M de cloreto de sódio, 20 mM de ácido etilenodiaminotetracético (EDTA), 100 mM de Tris-HCl (pH 8,0), 1% de polivinilpirrolidona 40.000 (PVP) e 0,2% de β -mercaptoetanol], previamente aquecido a $65\text{ }^{\circ}\text{C}$. O material foi agitado por 10 s no “vortex” e mantido em banho-maria a $65\text{ }^{\circ}\text{C}$ por 1 h. Após este período, foi realizada a primeira extração orgânica, adicionando-se 4,5 mL de clorofórmio:álcool isoamílico

(24:1) e agitando-se os tubos por inversão durante 5 min. A amostra foi, em seguida, centrifugada por 10 min. a 14.000 rpm à temperatura ambiente. A fase aquosa superior foi transferida para um novo tubo e tratada com 80 µL de RNase (1 mg/mL), sendo incubada por cerca de 60 min. a 37°C em banho-maria. Em seguida, adicionaram-se 700 µL de tampão de extração 10% [10% de brometo de cetiltrimetilamônio (CTAB) e 1,4 M de cloreto de sódio (NaCl)] e homogenizou-se. A amostra foi, então, submetida a uma segunda extração orgânica. Para isso, adicionaram-se ao tubo 4 mL de clorofórmio:álcool isoamílico (24:1) e agitou-se por 5 min., por inversão. O material foi centrifugado por 10 min. a 14.000 rpm e o sobrenadante foi transferido para um novo tubo. O DNA foi precipitado, pela primeira vez, acrescentando-se 2/3 do volume do sobrenadante de isopropanol. A solução foi agitada lentamente, por inversão, e incubada a -20 °C por 30 min., para facilitar a precipitação do DNA. Após este período, a amostra foi centrifugada por 10 min. a 14.000 rpm e o sobrenadante foi descartado. O *pellet* de DNA foi lavado duas vezes com 1 mL de etanol 70%, as lavagens foram seguidas por centrifugação durante 5 min. a 14.000 rpm e pelo descarte do etanol. Após a completa retirada do etanol, o *pellet* foi ressuspensionado em 500 µL de TE 10:1 (10 mM de Tris-HCl e 1 mM de EDTA, pH=8,0). Para a purificação do DNA total, uma segunda precipitação foi realizada com 250 µL de solução de acetato de amônio (7,5 M) e 1875 µL de etanol absoluto, ambos gelados. Para favorecer a precipitação, incubou-se o material por 1 h a -20 °C. Em seguida, centrifugou-se por 10 min. a 14000 rpm, e o sobrenadante foi descartado. O DNA foi lavado pela adição de 1 mL de etanol 70%; nova centrifugação foi conduzida por 5 min. a 14.000 rpm e o álcool foi descartado. A lavagem foi repetida e o *pellet* foi deixado secar à temperatura ambiente. Finalmente, o DNA foi ressuspensionado em 300 µL de TE 10:1 e estocado a 4 °C.

Para a quantificação, uma alíquota do DNA ressuspensionado foi diluída (20X) em solução contendo o corante azul de bromofenol e água ultrapura na proporção 1:4. Um volume de 10 µL desta solução foi aplicado em gel de agarose 1,0% (p/v). Para avaliar a intensidade da fluorescência das bandas, também foi aplicado DNA do fago lambda (Gibco-BRL) íntegro, em concentrações conhecidas, para a posterior comparação visual. A eletroforese foi realizada a 3V/cm durante 1h:30 min. O gel foi corado com solução de brometo de etídeo (10 µg/mL) durante 20 min. e fotografado sob luz ultravioleta no fotodocumentador MultiDoc- It Digital Imaging System® (UVP).

A concentração de DNA de cada amostra no gel foi estimada e o valor obtido foi multiplicado por 2, obtendo-se a concentração de DNA nos estoques em ng/μL. Foram preparadas alíquotas contendo a mesma concentração de DNA diluído em TE para todos os indivíduos.

3.3 Genotipagem com marcadores moleculares

3.3.1 Desenvolvimento de marcadores AFLP

Os genótipos dos genitores do cruzamento e da progênie foram identificados para os locos marcadores (dominantes) do tipo AFLP (VOS et al., 1995) cuja técnica foi adaptada para a cana-de-açúcar no presente estudo, tal como descrito a seguir.

Restrição do DNA

O DNA genômico foi duplamente digerido através da combinação de uma enzima de corte raro com uma enzima de corte frequente, ambas de corte coesivo. Para isso, 250 ng de DNA foram digeridos usando 6 U de cada uma das enzimas de restrição *EcoRI* (corte raro) (Promega®) e *MseI* (corte frequente) (New England BioLabs®) em solução contendo 1X do tampão universal *One Phor All* (Amersham®), 1X de BSA (New England BioLabs®) e água ultrapura esterilizada, num volume final de 25 μL de reação. O material foi incubado no termociclador Gen Amp® PCR System 9700 Thermocycler (Applied Biosystems) por 4 h a 37°C para restrição e, em seguida, por 20 min. a 65 °C para inativar as endonucleases.

Preparo dos adaptadores

O preparo dos adaptadores específicos (Tabela 2) aos terminais clivados por cada enzima ocorreu do seguinte modo:

- Adaptador *EcoRI*: foi preparada uma solução contendo 5 μM de adaptador *EcoRI forward*, 5 μM de adaptador *EcoRI reverse*, 0,5X de tampão *One Phor All* (Amersham®) e água ultrapura esterilizada para o volume final de 200 μL.

- Adaptador *MseI*: foi preparada uma solução composta por 50 μM de adaptador *MseI forward*, 50 μM de adaptador *MseI reverse*, 0,5X de tampão *One Phor All* (Amersham®) e água ultrapura esterilizada para completar o volume final de 200 μL .

Em seguida, as soluções foram submetidas a 65 °C por 10 min., 37 °C por 10 min. e finalmente a 25 °C por 10 min., sendo depois estocadas a -20 °C.

Ligação dos adaptadores

Para ligar os adaptadores foi acrescentado ao material digerido um *mix* de 25 μL contendo 0,25 μM de adaptador *EcoRI*, 2,5 μM de adaptador *MseI*, 1X de tampão para a enzima T4 DNA Ligase, 1,5 U da enzima T4 DNA ligase (New England BioLabs®) e água ultrapura esterilizada para completar o volume. Para que a ligação ocorresse, o material foi mantido por 16 h a 16 °C e a enzima foi inativada a 65 °C por 20 min. O material foi estocado a -20 °C.

Reações de Amplificação

Foram realizadas duas reações consecutivas com *primers* complementares aos adaptadores. Na reação de pré-amplificação foram utilizados *primers* de *EcoRI* e *MseI*, com extensão de 1 nucleotídeo seletivo na extremidade 3'. Os produtos da reação de pré-amplificação foram diluídos e uma alíquota dos mesmos foi utilizada como *template* na reação de amplificação seletiva, na qual foram utilizados 3 nucleotídeos seletivos adicionados na extremidade 3' dos *primers*.

a) Pré-Amplificação

A pré-amplificação foi realizada em um volume final de 20 μL contendo 1,5 mM de MgCl_2 , 0,5 mM de dNTPs, 250 nM de cada *primer* contendo um nucleotídeo seletivo adicionado à extremidade 3': oligo *EcoRI* + A e oligo *MseI* + C (apresentados na Tabela 2), 1X de tampão da enzima *Taq* DNA polimerase, 3 U de *Taq* DNA polimerase (Promega®), 3,0 μL de DNA digerido e ligado e água ultra pura para completar o volume. As condições para a PCR foram: 94°C por 2 min., seguindo-se 26 ciclos de 94 °C por 1 min., 56 °C por 1 min. e 72 °C por 1 min.,

e uma extensão final de 72 °C por 5 min. Os produtos pré-amplificados foram diluídos 5X em água ultrapura esterilizada e armazenados a -20 °C.

b) Amplificação seletiva

A amplificação foi realizada em solução composta por 1,5 mM de MgCl₂, 0,2 mM de dNTPs, 250 nM do oligo *EcoRI* + ANN (Tabela 2), 300 nM do oligo *MseI* + CNN (Tabela 2), 1X de tampão da *Taq* DNA polimerase, 1,6 U de *Taq* DNA polimerase (Fermentas™), 1,5 µL dos produtos pré-amplificados e diluídos e água ultra pura para completar o volume final de 20 µL. A PCR ocorreu nas seguintes condições: desnaturação a 94 °C por 2 min., seguida de 12 ciclos a 94 °C por 30 s, 65 °C por 30 s, 72 °C por 1 min., mais 23 ciclos nas mesmas condições, modificando-se somente a temperatura de anelamento para 56 °C, e uma extensão final de 72 °C por 2 min.

Tabela 2 – Sítios de restrição, sequências de adaptadores e de *primers* usados para as análises de AFLP

Enzima	Sítio de restrição	Especificação	Sequência ¹
<i>EcoRI</i>	5' G AATTC 3'	Adaptador <i>forward</i>	5' – CTC GTA GAC TGC GTA CC – 3'
	3' CTTAA G 5'	Adaptador <i>reverse</i>	5' – AAT TGG TAC GCA GTC TAC – 3'
		<i>Primer E+A</i>	5' – GAC TGC GTA CCA ATT CA – 3'
		<i>Primer E+ANN</i>	5' – GAC TGC GTA CCA ATT CAN N – 3'
<i>MseI</i>	5' T TAA 3'	Adaptador <i>forward</i>	5' – GAC GAT GAG TCC TGA G – 3'
	3' AAT T 5'	Adaptador <i>reverse</i>	5' – TAC TCA GGA CTC AT – 3'
		<i>Primer M+C</i>	5' – GAT GAG TCC TGA GTA AC – 3'
		<i>Primer M+CNN</i>	5' – GAT GAG TCC TGA GTA ACN N – 3'

¹ N: nucleotídeo arbitrário usado na amplificação seletiva.

3.3.2 Desenvolvimento de marcadores direcionados a sequências de retrotransposons

A metodologia para a obtenção de marcadores referentes aos retrotransposons *SURE* e *Garapa* foi desenvolvida a partir da técnica *NBS-profiling* descrita por Van der Linden et al.

(2004) para a detecção de marcadores RGAs (*Resistance Gene Analogs*) e envolve etapas similares às de AFLP.

Inicialmente, o DNA genômico é digerido com apenas uma enzima de restrição, seguida pela ligação de um adaptador específico aos terminais clivados pela enzima. O adaptador é formado por duas fitas: uma fita curta bloqueada na região 3' pela presença de um grupo amino, o que impede a sua extensão pela *Taq* DNA polimerase, e fosforilada na região 5', o que facilita sua ligação aos fragmentos restringidos, e uma fita longa comum.

Na reação de amplificação utiliza-se o *primer* do adaptador e um *primer* de retrotransposon. No primeiro ciclo da PCR, somente as fitas complementares ao *primer* de retrotransposon são amplificadas, devido à estrutura dos adaptadores previamente ligados aos fragmentos de restrição (fase linear). No final desta primeira extensão é que são gerados os sítios complementares ao *primer* do adaptador. Nos demais ciclos, ambos os *primers* se ligam ao DNA e a amplificação ocorre de modo exponencial (veja a Figura 4).

Através dessa técnica são gerados amplicons constituídos obrigatoriamente por sequências parciais de retrotransposons e pela sequência do adaptador, que comportam-se como marcadores dominantes. Assim, duas fontes de polimorfismos são exploradas: o polimorfismo referente à presença ou ausência do retrotransposon e o polimorfismo relacionado ao sítio de restrição das enzimas.

A descrição detalhada da técnica é apresentada a seguir.

Obtenção dos primers complementares a sequências dos retrotransposons

Os *primers* complementares às sequências do retrotransposon *SURE*, utilizados nesse trabalho, foram previamente desenvolvidos por Domingues (2009) e localizam-se em diferentes regiões do elemento (Figura 5). Para o retrotransposon *Garapa*, vários *primers* foram desenhados a partir de sequências consenso das regiões das LTRs e da RT (Figura 5). O programa Primer-3 (ROZEN; SKALETISKY, 2000) foi utilizado para gerar esses *primers*. Para evitar a ocorrência de amplificações espúrias, utilizaram-se dois critérios de estrigência: (1) temperatura de anelamento variando entre 55 e 60 °C e (2) diferença máxima de 3 °C entre as T_M do par *primer* de retrotransposon/*primer* do adaptador. Cada *primer* também foi analisado quanto à possibilidade de auto-anelamento ou de formação de estruturas secundárias.

Figura 4 – Representação esquemática da metodologia desenvolvida para a obtenção de marcadores direcionados a sequências de retrotransposons

Figura 5 – Representação esquemática da estrutura dos retrotransposons estudados indicando as regiões de anelamento dos *primers* (setas). Em a) são mostrados os quatro *primers* do *SURE*: 1 – SSAP1, 2 – GagRev, 3- LTR1 e 4 – LTR2; em b) os dez *primers* do *Garapa*: 1 – LTR-R, 2/3 – LTRrev1 e LTRrev2 (desenhados sobre a mesma região de anelamento), 4/5 – RTfw1 e RTfw2, 6 – LTR116F, 7/8 – LTRfw1 e LTRfw2, 9 – LTR65F e 10 – LTR41F

Restrição do DNA

O DNA genômico foi digerido com diferentes enzimas de corte abrupto. Algumas delas apresentam sítios de restrição apenas na região genômica externa ao retrotransposon e outras reconhecem sítios também na sequência dos elementos (Tabela 3), sendo essas endonucleases de corte frequente (reconhecem sítios de 4 bases) ou de corte raro (reconhecem sítios de 6 a 8 bases). Para a restrição, 500 ng de DNA foram utilizados juntamente com 7,5 U da enzima de restrição, 1X do respectivo tampão da enzima, 1X de BSA (New England BioLabs®) e água ultrapura esterilizada para o volume final de 30 µL. O material foi incubado por 7 h a 37 °C para restrição e, em seguida, por 20 min. a 65 °C para a inativação das enzimas. No caso da enzima *Bse*II, a restrição ocorreu a 65 °C e a inativação a 80 °C.

Preparo do adaptador

O adaptador foi preparado misturando-se 100 µM de cada um dos adaptadores de fita curta e de fita longa (Tabela 4) com água ultrapura esterilizada num volume final de 240 µL. A solução foi incubada a 65 °C por 10 min., 37 °C por 10 min. e a 25 °C por 10 min., sendo depois estocada a -20 °C.

Tabela 3 – Enzimas de restrição utilizadas e o(s) retrotransposon(s) clivado(s)

Enzima	Sítio de restrição ¹	Retrotransposon
<i>DraI</i>	5' TTT AAA 3' 3' AAA TTT 5'	<i>Garapa</i>
<i>RsaI</i>	5' GT AC 3' 3' CA TG 3'	<i>SURE</i> e <i>Garapa</i>
<i>SspI</i>	5' AAT ATT 3' 3' TTA TAA 5'	<i>Garapa</i>
<i>AluI</i>	5' AG CT 3' 3' TC GA 5'	<i>SURE</i> e <i>Garapa</i>
<i>PmeI</i>	5' GTTT AAAC 3' 3' CAAA TTTG 3'	-
<i>Eco105I</i>	5' TAC GTA 3' 3' ATG CAT 5'	<i>SURE</i>
<i>BseII</i>	5' GATNN NNATC 3' 3' CTANN NNTAG 5'	<i>SURE</i>
<i>HpaI</i>	5' GTT AAC 3' 3' CAA TTG 5'	-

¹ N nucleotídeo arbitrário

Ligação do adaptador

A ligação do adaptador ocorreu acrescentando-se ao material digerido um *mix* de 30 µL contendo 1,6 µM de adaptador – no caso de restrição do DNA com uma enzima de corte frequente ou 0,2 µM de adaptador – quando o DNA foi restringido com uma enzima de corte raro, 1X de tampão para a enzima T4 DNA Ligase, 1,5 U de enzima T4 DNA Ligase (New England BioLabs®) e água ultrapura esterilizada para completar o volume. O material foi mantido por 16 h a 16 °C e a enzima foi inativada a 65 °C por 20 min. O produto de ligação foi diluído de modo a padronizar sua concentração final a 5 ng/µL. A solução foi estocada a -20 °C.

Amplificação

A amplificação foi realizada em solução composta por 1,5 mM de MgCl₂, 0,2 mM de dNTPs, 300 nM de *primer* complementar ao retrotransposon (Tabela 4), 300 nM de *primer* complementar ao adaptador (Tabela 4), 1X de tampão da *Taq* DNA polimerase, 2 U de *Taq* DNA polimerase (Fermentas™), 20 ng de DNA digerido e ligado e água ultra pura para completar o volume final de 20 µL. As condições para a PCR foram: desnaturação a 94 °C por 5 min., 8 ciclos a 94 °C por 45 s, 58 °C (- 1 °C por ciclo) durante 50 s e 72 °C por 1h:15 min., seguida de mais 25 ciclos a 94 °C por 45 s, 50 °C por 50 s e 72 °C por 1 min., e uma extensão final a 72 °C por 10 min.

Tabela 4 – *Primers* usados para a amplificação de marcadores direcionados a sequências de retrotransposons

Nome	Sequência ²
Adaptador de fita curta ¹	5' – TGG GAT CTA TAC TT H ₂ N – 3'
Adaptador de fita longa ¹	5' – ACT CGA TTC TCA ACC CGA AAG TAT AGA TCC CA – 3'
<i>Primer</i> do adaptador ¹	5' – ACT CGA TTC TCA ACC CGA AAG – 3'
SURE_LTR1	5' – AGT CCT GCT CCC AGT TAT CA – 3'
SURE_LTR2	5' – GTC GCC TGG GTG TGT TAT C – 3'
SURE_SSAP1	5' – CAA GCC CTT AAT AGC AGA AA – 3'
SURE_GagRev	5' – TCC CTG TAT ACA ACC CTG TC – 3'
GARAPA_LTRfw1	5' – CCA TGW GGC TAA GGC CCA TGW GGC – 3'
GARAPA_LTRfw2	5' – CCA TGW GRC TAG GCC CAT GTG GC – 3'
GARAPA_LTRrev1	5' – GGC TAA CCC TAG AAG GRT AGC – 3'
GARAPA_LTRrev2	5' – ATA CAT GGG CCA CAT GGG – 3'
GARAPA_RTfw1	5' – TYG YGG AYT YTT CMC CAA – 3'
GARAPA_RTfw2	5' – CCT CCC DTC CTC GAC CTT C – 3'
GARAPA_LTR41F	5' – GGG GTG TTG GAG TGT GAT TG – 3'
GARAPA_LTR65F	5' – AGG CCC ATG TGG CCC ATG TAT G – 3'
GARAPA_LTR116F	5' – GGT TAG CCC AAA ATA ATA GAG AAT CA – 3'
GARAPA_LTR-R	5' – GAA CTA CCG GCG CAG CGG AGT TC – 3'

¹Retirados de Van der Linden et al. (2004), ²Código para bases mistas: W = A ou T; R = A ou G; Y = C ou T; M = A ou C; D = A, G ou T.

3.3.3 Eletroforese e visualização dos locos

Separação dos locos amplificados

Aos produtos das amplificações foi adicionado igual volume de tampão desnaturante [98% de formamida, 10 mM de EDTA (pH 8,0), 0,002% de azul de bromofenol e 0,002% de xileno cianol]. Esta mistura foi submetida à desnaturação a 94 °C por 5 min. e em seguida mantida no gelo. Uma alíquota de 3 µL foi aplicada em gel de poliacrilamida 5%, [acrilamida/bisacrilamida (19:1), 7,5 M de uréia e tampão TBE 1X (89 mM de Tris-base, 89 mM de ácido bórico, 2,5 mM EDTA 2Na)] após 1 h de pré-corrida a 70 W. A eletroforese foi realizada a potência constante de 70 W por 4 h, usando um aparato de eletroforese Sequi-Gen® (Bio-Rad). Como marcador de massa molecular utilizaram-se 2 µL de uma mistura dos *Ladders* de 25 e 100 pb (cada um a 40 ng/µL na mistura).

Coloração e visualização dos locos

A coloração foi realizada em solução de nitrato de prata (AgNO_3), seguindo protocolo adaptado de Creste, Tulmann-Neto e Figueira (2001):

a) Fixação – 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 e deionizada, durante 1 min., com agitação.

b) Oxidação – O procedimento de pré-tratamento foi realizado em solução de ácido nítrico 1,5%, agitando durante 3 min., seguindo-se uma nova lavagem do gel com água destilada e deionizada durante 1 min, com agitação.

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

d) Revelação – A revelação se deu em solução de 2 litros contendo 60 g de carbonato de sódio (Na_2CO_3) e 2,0 mL de formaldeído (37%). Nesta etapa, 1 litro da solução foi aplicado sobre o gel até começarem a surgir bandas. Após este passo, a solução foi descartada e o outro

litro adicionado para finalizar o processo de revelação. Após a revelação do gel a solução foi descartada.

e) Bloqueio – O bloqueio da revelação foi realizado em solução de ácido acético 5% durante 4 min., seguindo-se uma nova lavagem com água destilada deionizada durante 1 min.

Após esse passo, o gel foi colocado para secar em posição vertical, em local arejado e sob temperatura ambiente, possibilitando sua secagem e posterior análise dos locos.

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

3.4 Validação molecular dos marcadores direcionados a sequências de retrotransposons

A validação dos fragmentos derivados dos retrotransposons foi feita, basicamente, por sequenciamento, seguido pelo alinhamento das sequências usando o banco de dados GenBank (NCBI). Tal análise foi realizada com o intuito de confirmar se os produtos das ampliações correspondiam a regiões de elementos transponíveis.

Primeiramente, foram escolhidos 10 fragmentos de cada combinação enzima de restrição/*primer* complementar ao retrotransposon. Assim, fragmentos monomórficos (presentes em todos os indivíduos) ou polimórficos (segregantes), que apresentavam boa visualização e tamanhos entre 100 a 1000 pb foram retirados dos géis de poliacrilamida com o auxílio de um bisturi.

Cada fragmento foi mantido submerso em 100 uL de TE 10:1 (10 mM de Tris-HCl e 1 mM de EDTA, pH=8,0) para que o DNA se desprendesse do gel. Após 1 semana sob 4 °C, a solução foi filtrada.

A reamplificação dos fragmentos ocorreu utilizando-se as seguintes concentrações de reagentes: 1,5 mM de MgCl₂, 0,2 µM de dNTPs, 30 nM de *primer* de retrotransposon, 30 nM de *primer* do adaptador, 1X de tampão da *Taq* DNA polimerase, 5 U de *Taq* DNA polimerase (Fermentas™), 5 µL da solução filtrada e água ultrapura esterilizada para o volume final de 50 µL. As condições da PCR foram: desnaturação a 94 °C por 5 min., 30 ciclos a 94 °C por 30 s, 55 °C durante 30 s e 72 °C por 42 s, e extensão final a 72 °C por 10 min.

Os produtos foram submetidos à eletroforese em gel de agarose 1,0% (p/v), acrescido de 0,5X de GelRed™ (Biotium), a 3V/cm durante 2 h. Foi utilizado um marcador de massa

molecular (*Ladder*) de 100 pb para a confirmação do tamanho dos fragmentos. Após a digitalização do gel no fotodocumentador MultiDoc- It Digital Imaging System® (UVP), as amostras foram purificadas da agarose com o kit QIAEX II Gel Extration (QIAGEN).

Os fragmentos purificados foram quantificados aplicando-se 2 µL de cada amostra em gel de agarose 1,0% (p/v), após 1h:30 min. de eletroforese a 3V/cm, e comparando-se a intensidade da fluorescência das bandas com concentrações conhecidas do DNA do fago lambda (Gibco-BRL). A concentração de cada amostra foi ajustada para 2,5 ng/µL.

As etapas seguintes foram realizadas no GaTELab (IB/USP), sob supervisão da Profª Marie-Anne Van Sluys.

Foram preparadas misturas de fragmentos (*pools*) para cada combinação enzima de restrição/*primer* complementar ao retrotransposon, e o trabalho seguiu partindo desse material.

A clonagem dos fragmentos foi realizada com o kit pMOS Blue Blunt Ended Cloning (Amersham/GE). Como o vetor pMOS possui as extremidades abruptas, os fragmentos foram preparados para a ligação ao vetor através da reação de reparo das pontas, realizada seguindo as recomendações do fabricante. Para a ligação, foram acrescentados a cada produto da reação de reparo das pontas 1 µL de vetor pMOS e 1 µL de T4 DNA ligase. O material foi incubado a 22°C por 2 h.

A transformação foi feita em células de *E-coli* eletrocompetentes (DH10B) por eletroporação. O produto da transformação foi deixado crescer em meio SOC líquido, *overnight* e, em seguida, o material foi plaqueado em placas de Petri contendo meio LB com ampicilina (100 µg/mL), 40 µg/mL de X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) e 250 µg/mL de IPTG (isopropyl-beta-D-thiogalactopyranoside), permanecendo a 37 °C por cerca de 16 h.

Após o crescimento, algumas colônias brancas (que contêm os insertos) foram selecionadas e coletadas com auxílio de palito esterilizado, sendo mergulhadas em placas *deep well* contendo meio CircleGrow líquido com ampicilina a 100 µg/mL. Após o crescimento das colônias no meio líquido, foi feita a extração dos plasmídeos (miniprep) pelo método adaptado a partir de Sambrook (2001).

Os produtos da miniprep foram submetidos à eletroforese (3 V/cm) em gel de agarose 1,0% (p/v), corados com solução de brometo de etídeo (10 µg/mL) e fotodocumentados sob luz ultravioleta em um sistema de imagem digital. Esses produtos foram amplificados de acordo com

o método desenvolvido por Sanger, Nicklen e Coulson (1977), utilizando-se o *primer* universal T7 *promoter*. As reações foram purificadas para, então, serem submetidas à corrida eletroforética em sequenciador ABI 3730 (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 existentes no banco de dados do NCBI.

3.5 Análise dos dados

3.5.1 Análise do polimorfismo molecular entre os genitores ‘IAC66-6’ e ‘TUC71-7’

Marcadores AFLP

Um total de 72 combinações de enzimas de restrição/*primers* seletivos foram avaliadas usando o DNA dos genitores do cruzamento. A taxa de polimorfismo foi calculada com base no número de locos polimórficos em relação ao número total de locos amplificados por cada combinação de enzima de restrição/*primer* seletivo.

Combinações que apresentaram padrão uniforme nos géis, elevado número de locos amplificados e cuja taxa de polimorfismo entre os genitores foi $\geq 20\%$ foram escolhidas para genotipagem da população F₁.

Marcadores direcionados a sequências de retrotransposons

Para a obtenção de marcadores direcionados a sequências do retrotransposon *SURE*, foram testadas combinações entre as enzimas de restrição *DraI*, *SspI*, *RsaI* e *AluI* e a série de *primers* desenhados para esse elemento, usando o DNA dos genitores.

Em relação ao *Garapa*, combinações envolvendo todas as enzimas listadas na tabela 3 e os 10 *primers* desse elemento foram analisadas.

As taxas de polimorfismos foram calculadas da mesma forma como para os dados de AFLP. Combinações que apresentaram bom padrão de amplificação e a presença de locos polimórficos foram selecionadas para a genotipagem da população de mapeamento.

3.5.2 Análises de segregação dos locos

A análise genética foi baseada na contagem da presença (1) ou ausência (0) da banda (alelo) por loco de AFLP ou de retrotransposon nos dois genitores e nos 188 indivíduos da população filial.

As proporções foram analisadas pelo *software* R versão 2.10.1 (R DEVELOPMENT CORE TEAM, 2009) usando o teste de qui-quadrado (χ^2), em relação às hipóteses de segregação 1:1 e 3:1, que correspondem às configurações $A0000000 \times 00000000$ (ou vice-versa) e $A0000000 \times A0000000$, respectivamente, tomando como *A* o alelo amplificado (banda) e *0* como o alelo nulo (ausência da banda), tal como proposto por Wu et al. (1992) para o estudo genético de poliplóides, com base na segregação de marcadores de dose única (MDU).

Para minimizar os erros associados com os múltiplos testes, foi aplicada a correção de Bonferroni, na qual o χ^2 crítico foi obtido dividindo-se o valor de α a 5% de probabilidade (0.05) pelo número total de marcadores. As marcas com desvios significativos foram desconsideradas.

3.6 Construção do mapa de ligação

As análises de ligação na população de mapeamento (F_1) e a construção do mapa genético foram realizadas usando-se o *software* OneMap (MARGARIDO; GARCIA; SOUZA, 2007) em sua nova versão, adaptada para realizar análises multiponto.

Esse *software* toma por base os algoritmos propostos por Wu et al. (2002) para a construção de mapas genéticos em populações F_1 segregantes, que estimam, simultaneamente, a fração de recombinação e as fases de ligação entre os marcadores. A abordagem multiponto é realizada por verossimilhança usando o modelo markoviano oculto (LHMM ou *Likelihood Hidden Markov Model*) e o teste de dois pontos entre as marcas é realizado com conversão pelo algoritmo EM (*Expectation Maximization Algorithm*) para o cálculo das frações de recombinação.

Como critérios para estabelecer os grupos de ligação foram usados $LOD\ score \geq 5,0$ e fração de recombinação $\leq 0,35$. A distância expressa em centiMorgans (cM) foi estimada usando a função de Kosambi (1944).

Após a atribuição dos marcadores nos grupos de ligação, esses foram ordenados usando diferentes funções que variavam de acordo com o tamanho de cada grupo. De modo geral, nos grupos de ligação que possuíam até cinco marcadores foi utilizado o comando 'COMPARE'. Já nos grupos com número de marcadores superior a cinco, as marcas foram ordenadas usando diferentes comandos, como 'RCD', 'UG', 'RECORD' e 'ORDER'. Neste caso, os resultados de cada algoritmo foram avaliados buscando-se encontrar aquele que fornecesse a melhor ordem.

Paralelamente, as matrizes gráficas das frações de recombinações e do LOD (*heatmaps*) foram utilizadas para se determinar a ordem final das marcas nos respectivos grupos de ligação. Além disso, a ordem das marcas sugerida pelo OneMap foi verificada com base no algoritmo 'RIPPLE'.

O mapa final foi desenhado pelo programa MapChart 2.2 (VOORRIPS, 2002).

4 RESULTADOS E DISCUSSÃO

4.1 Testes e seleção de combinações enzimas de restrição/*primers*

Marcadores AFLP

Das 72 combinações enzimas de restrição/*primers* seletivos avaliadas, houve amplificação de um total de 7930 locos, dos quais 1276 (16%) foram polimórficos entre os genitores e 6654 monomórficos. Em média, 110 locos foram amplificados por combinação, dos quais 17 (15%) eram polimórficos. Portanto, considerando todas as combinações avaliadas, a taxa de polimorfismo entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ correspondeu a 16%.

Um exemplo ilustrativo dos testes de combinações encontra-se na figura 6. Nos testes, cada genitor foi amplificado com duas repetições para verificar a reprodutibilidade dos resultados e eliminar erros inerentes ao manuseio do material. Todas as combinações avaliadas apresentaram-se claramente reproduzíveis.

Houve diferenças notáveis entre os padrões de amplificação entre as diferentes combinações, sendo que a variação do total para locos amplificados foi de 44 a 174 e, para locos polimórficos, esta variação foi de 4 a 33.

Considerando-se os critérios descritos anteriormente para a escolha das combinações e a boa visualização das bandas, 22 combinações foram selecionadas para a genotipagem da população F₁. As combinações selecionadas foram nomeadas de acordo com o código dos *primers* seletivos de *EcoRI* e *MseI* e encontram-se descritas na tabela 5.

Quatorze dessas combinações já haviam sido utilizadas em trabalhos de mapeamento genético de cana-de-açúcar, tais como nos mapas construídos para a cultivar ‘R570’ (HOARAU et al., 2001), para o clone de *S. officinarum* ‘IJ76-514’ e para as cultivares ‘Q165’ (AITKEN et al., 2005), ‘Q117’ e ‘MQ77-340’ (REFFAY et al., 2005) e para a população derivada do cruzamento ‘SP80-180’ x ‘SP80-4966’ (GARCIA et al., 2006). Essas combinações em comum são importantes porque permitem comparar os diferentes genótipos estudados. Também, existe a possibilidade de as marcas cobrirem uma mesma região genômica e, com isto, se estabelecerem as correspondências entre os grupos de ligação publicados por esses autores.

Figura 6 – AFLP: testes com 13 combinações de enzimas de restrição/*primers* seletivos. As setas indicam alguns locos polimórficos. M: marcador de massa molecular (*Ladder*) de 125 pb; IAC: genitor ‘IAC 66-6’; TUC: genitor ‘TUC 71-7’

A comparação entre os mapas de ligação das cultivares ‘R570’, ‘Q165’, ‘Q117’ e ‘MQ77-340’ já foi realizada por Piperidis et al. (2008), com base em marcas comuns de AFLP e SSRs. Embora esses mapas sejam pouco saturados, foi possível alinhar vários grupos de ligação e associá-los em grupos de homologia, que receberam a mesma nomenclatura, possibilitando que se utilizem esses dados como referência para os demais mapas genéticos de cana-de-açúcar.

Tabela 5 – Combinações de AFLP selecionadas para a genotipagem da população ‘IAC66-6’ x ‘TUC71-7’. São mostrados os nomes dados a cada combinação com base nos nucleotídeos seletivos dos *primers*, o número de locos amplificados e a taxa de polimorfismo calculada

Código	Nucleotídeos seletivos	Total amplificado	Locos polimórficos	Taxa de polimorfismo (%)
E31M49	E + AAA, M+ CAG	73	16	21,9
E31M58	E + AAA, M+ CGT	133	31	23,3
E32M48	E + AAC, M+ CAC	123	29	23,5
E32M50	E + AAC, M+ CAT	91	19	20,8
E32M60	E + AAC, M+ CTC	106	22	20,7
E33M50	E + AAG, M+ CAT	90	18	20,0
E33M58	E + AAG, M+ CGT	105	23	21,9
E33M61	E + AAG, M+ CTG	85	23	27,0
E35M47	E + ACA, M+ CAA	115	23	20,0
E35M48	E + ACA, M+ CAC	79	23	29,1
E35M59	E + ACA, M+ CTA	81	17	20,9
E35M60	E + ACA, M+ CTC	78	20	25,6
E35M62	E + ACA, M+ CTT	102	26	25,4
E36M48	E + ACC, M+ CAC	110	22	20,0
E37M50	E + ACG, M+ CAT	94	20	21,2
E37M58	E + ACG, M+ CGT	106	24	22,6
E37M62	E + ACG, M+ CTT	104	25	24,0
E38M58	E + ACT, M+ CGT	84	19	22,6
E39M60	E + AGA, M+ CTC	92	27	29,3
E40M48	E + AGC, M+ CAC	100	25	25,0
E40M50	E + AGC, M+ CAT	81	17	20,9
E40M49	E + AGG, M+ CAG	136	33	24,2
<i>Total</i>		2168	502	-
<i>Média</i>		98,5	22,8	-
<i>Amplitude</i>		73 - 133	16 - 33	-

Marcadores direcionados a sequências de retrotransposons

Para o retrotransposon *SURE*, do total de 16 combinações enzimas de restrição/*primer* avaliadas houve amplificação de 770 locos, dos quais 125 (16%) foram polimórficos entre os genitores e 645 monomórficos. A média de locos amplificados foi de aproximadamente 48 por combinação, sendo 7 deles polimórficos. A figura 7 mostra as 16 combinações testadas para o *SURE*.

Figura 7 – Retrotransposon *SURE*: padrão de amplificação das 16 combinações de enzimas de restrição/*primers*. As setas indicam alguns locos polimórficos. M: marcador de massa molecular (*Ladder*) de 125 pb; IAC: genitor ‘IAC 66-6’; TUC: genitor ‘TUC 71-7’

De modo similar aos testes com AFLP, o DNA dos genitores foi amplificado, usando três repetições, para cada combinação, cada uma delas contendo: 10, 20 e 30 ng de DNA digerido e ligado. O resultado mostrou que não houve diferenças entre as amplificações com as diferentes quantidades de DNA *template*, assegurando que a técnica é altamente reproduzível nessas condições.

Os *primers* LTR1, LTR2 e SSAP1 resultaram em bons padrões de amplificação, principalmente quando testados com as enzimas *DraI* e *SspI*, que não cortam o elemento *SURE*. As combinações que envolviam esses *primers* e enzimas identificaram maior quantidade de locos polimórficos e apresentaram uma melhor visualização dos locos no gel em comparação às demais.

Somente o *primer* GagRev não resultou em boa amplificação com as enzimas utilizadas. Uma das possibilidades para explicar esse resultado está relacionada com o anelamento desse *primer* no retrotransposon, que é localizado na região gênica *gag*, a 638 pb da região genômica adjacente ao elemento. Assim, as amplificações com esse *primer* predominantemente resultariam em fragmentos de tamanho elevado, que são estendidos com pouca eficiência pela *Taq* DNA polimerase e são mais difíceis de serem diferenciados nos géis.

Também analisando os amplicons obtidos pelos demais *primers* complementares às LTRs (que possuem 250 pb de comprimento), pode-se inferir que eles são constituídos pela sequência da LTR seguida pela sequência adjacente ao elemento, apresentando tamanhos variados de acordo com a localização dos sítios de restrição, conforme esperado pela metodologia. Entretanto, amplificações com tamanhos menores de 200 pb ocorreram, embora inesperadas, porque, provavelmente, as enzimas de restrição reconheceram sítios na LTR, por causa da diversidade de cópias do *SURE* no genoma da cana-de-açúcar. Outra explicação é que tenha ocorrido amplificação de fragmentos inespecíficos, o que poderá ser verificado por sequenciamento.

Os testes com os *primers* LTR1, LTR2 e SSAP1 variaram pouco em comparação aos testes com AFLP, sendo que o total de locos amplificados se estendeu de 44 a 70, e de locos polimórficos de 1 a 17.

A partir desses mesmos resultados, foi estudado o total de locos amplificados para cada genitor separadamente. Para isso, a quantidade de locos monomórficos amplificados em cada combinação foi somada à quantidade de locos polimórficos fornecido pelo respectivo genitor. Desse modo para o genitor ‘IAC 66-6’ foram detectados, em média, 45 locos e para o genitor ‘TUC 71-7’, 46 locos. Esses valores correspondem ao número de cópias do elemento *SURE* nesses genótipos, estando de acordo com as estimativas do número de cópias desse elemento no genoma da cana-de-açúcar, segundo Domingues (2009), que estimou entre 40 e 50 cópias.

Para a genotipagem da população, foram selecionadas seis combinações que apresentaram bom perfil de bandas com a presença de locos polimórficos, mostradas na tabela 6. As combinações foram nomeadas pelas primeiras letras das enzimas de restrição e dos *primers*, com algarismos romanos.

Tabela 6 – Combinações enzimas de restrição/*primers* do retrotransposon *SURE* selecionadas para a genotipagem da população ‘IAC66-6’ x ‘TUC71-7’. O nome dado a cada combinação é indicado junto com os números de locos amplificados e de locos polimórficos

Código	Combinação enzima de restrição/ <i>primer</i>	Total amplificado	Locos polimórficos
DILI	<i>Dra</i> ILTR1	50	12
DILII	<i>Dra</i> ILTR2	70	12
DISI	<i>Dra</i> ISSAP1	50	12
SILI	<i>Ssp</i> ILTR1	48	15
SISI	<i>Ssp</i> ISSAP1	52	13
RISI	<i>Rsa</i> ISSAP1	60	17
<i>Total</i>		330	81
<i>Média</i>		55	13,5
<i>Amplitude</i>		48 - 70	12 - 17

Em relação ao retrotransposon *Garapa*, foram testadas somente combinações envolvendo as enzimas *Dra*I, *Ssp*I, *Rsa*I, *Alu*I e *Bse*JI com os 10 *primers* complementares a esse elemento. Para as demais enzimas (*Hpa*I, *Eco*105I e *Pme*I) que, juntamente com a enzima *Bse*JI, não clivam a sequência do *Garapa*, os testes não puderam ser realizados pois não ocorreu a digestão do DNA genômico. Assim, dos 50 testes realizados, observaram-se dois padrões: (1) ausência de amplificação e (2) amplificação de elevada quantidade de fragmentos.

A ausência de amplificação provavelmente ocorreu pela baixa qualidade de alguns *primers* testados, que apresentavam formação de estruturas secundárias, tais como dímeros e *loop* interno, que prejudicaram a amplificação.

Dentre as combinações que apresentaram amplificação, nenhuma delas forneceu boa visualização dos locos nos géis e nem mesmo revelaram polimorfismos entre os genitores (todos os locos se mostraram monomórficos), tornando inviável a sua utilização para a genotipagem da população. Entretanto, a grande quantidade de locos amplificados permite inferir que esse elemento possui elevado número de cópias no genoma da cana-de-açúcar. Segundo Domingues (2009) estima-se que haja mais de 1000 cópias do *Garapa* no genoma da cana. Os resultados aqui obtidos são melhor explicados considerando-se que os dois retrotransposons estudados consistem em famílias de elementos com as seguintes características:

- a) *SURE* apresenta baixo número de cópias, baixa diversidade entre essas cópias, é transcricionalmente ativo em tecidos pouco diferenciados e encontra-se em expansão no genoma;

- b) *Garapa* possui elevado número de cópias, sua diversidade é maior que a do *SURE* tanto na região codificante quanto na LTR, é ativo em diferentes situações fisiológicas e sua expansão já tem sido retida por geração de variabilidade.

Desse modo, a baixa diversidade do *SURE* possibilitou que se obtivessem bons *primers* para esse elemento e, conseqüentemente, géis com quantidade de locos correspondente à esperada.

Como a família *Garapa* possui mais de 1000 cópias no genoma da cana-de-açúcar, com elevada diversidade e variabilidade entre essas cópias, o desenho dos *primers* para esse elemento foi dificultado. Até mesmo a restrição do DNA genômico com enzimas que não clivam esse retrotransposon foi prejudicada por sua elevada abundância no genoma.

A falta de polimorfismo entre os genitores nas combinações analisadas para o *Garapa*, provavelmente, é consequência do seu alto número de cópias. Assim, a presença desse elemento sob a forma de repetições em tandem dificulta a detecção de polimorfismos moleculares, além do fato de haver comigração de fragmentos, que é muito comum em poliplóides (WU et al., 1992). Esta comigração é altamente esperada neste caso já que os *primers* se anelam em regiões de notável abundância no genoma da cana-de-açúcar.

Uma alternativa para o desenvolvimento de marcadores para esse elemento seria estudar suas subfamílias separadamente. Esta estratégia poderia facilitar o desenho de *primers*, diminuir a quantidade de fragmentos amplificados e, então, detectar a presença de polimorfismos entre os genitores. Além disso, seria possível fornecer uma estimativa do número de cópias de cada subfamília desse elemento no genoma.

Outra possibilidade seria a adição de nucleotídeos seletivos no *primer* do adaptador, tal como é feito na técnica SSAP. Estudando o LTR retrotransposon *BARE-1* em cevada, Waugh e colaboradores (1997) demonstraram uma redução no número de produtos de amplificação como consequência de cada base adicionada. Nas amplificações sem o uso de nucleotídeos seletivos, o padrão de amplificação em cevada foi muito similar ao resultado alcançado nesse trabalho com o retrotransposon *Garapa*.

Queen et al. (2004) ao estudarem com SSAP os retrotransposons *Wis2A-1A* e *BARE-1*, que apresentam mais de 1000 cópias no genoma do trigo, utilizaram até quatro nucleotídeos seletivos para reduzir o número de fragmentos amplificados, conseguindo obter ótimos resultados para a

genotipagem de uma população. Isso indica que essa estratégia também seria muito eficiente para o estudo do polimorfismo molecular associado ao elemento *Garapa*.

4.2 Análises da segregação dos locos

Um total de 3451 locos foram obtidos pela genotipagem dos 188 indivíduos da população usando as 22 combinações de AFLP e 6 combinações direcionadas ao retrotransposon *SURE*. Dentre esses locos, 772 (22,3%) foram segregantes, incluindo 685 (88,7%) marcas provenientes de AFLP e 87 (11,3%) marcas direcionadas para o *SURE*.

Dos 772 marcadores, 562 (72,8%) foram polimórficos entre os genitores e 210 monomórficos (Tabela 7). Os genitores contribuíram de maneira similar para o polimorfismo molecular, sendo 51% dessas marcas provenientes do genitor ‘IAC 66-6’ e 49% do genitor ‘TUC 71-7’.

Tabela 7 – Resultado da genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ com marcadores AFLP e marcas derivadas do retrotransposon *SURE*, segundo as proporções esperadas para marcadores em dose única

		Locos segregantes		
		1:1	3:1	Total
AFLP	3094	488	197	685
<i>SURE</i>	357	74	13	87
<i>Total</i>	3451	562	210	772

Para a leitura dos dados, apenas os marcadores com segregações facilmente identificáveis foram considerados. As marcas de difícil interpretação foram tomadas como dados perdidos para evitar os erros de genotipagem. Nas figuras 8 e 9 são demonstrados os padrões de amplificação obtidos com AFLP e com o retrotransposon *SURE*, respectivamente.

Figura 8 – Genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ utilizando a combinação de AFLP E35M48. São mostrados inúmeros locos segregantes. Em M encontra-se o marcador de massa molecular (*Ladder*) 125 pb. Os números são referentes aos indivíduos da população F_1 . À direita, entre dois marcadores encontram-se os genitores indicados como IAC: ‘IAC 66-6’ e TUC: ‘TUC 71-7’

Figura 9 – Genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ com a combinação SILI direcionada para o retrotransposon *SURE*. Em M encontra-se o marcador de massa molecular (*Ladder*) 125 pb. Os números são referentes aos indivíduos da população F₁. À direita, entre dois marcadores encontram-se os genitores indicados como IAC: ‘IAC 66-6’ e TUC: ‘TUC 71-7’

As análises de segregação das marcas para a hipótese de dose única (χ^2 ao nível de 5% de significância), resultaram em 455 (59%) marcas sem distorção de segregação. Após a correção de Bonferroni, 145 marcas foram recuperadas, resultando em 600 (77,7%) marcas para a construção do mapa genético. Maiores detalhes sobre esses resultados são descritos a seguir e, também, encontram-se nas tabelas 8 e 9.

Para os marcadores AFLP, o número de locos amplificados pelas combinações enzimas de restrição/*primers* variou de 102 a 172 (com média de 140), e o número de locos segregantes variou

de 19 a 48 (média de 31,1). A taxa de polimorfismo se estendeu de 11,8% a 29,8%. De acordo com os critérios escolhidos para as análises de segregação, 535 dos 685 locos segregantes foram considerados marcadores em dose única e as 150 marcas restantes foram consideradas como apresentando distorção, não sendo tomadas para a construção do mapa genético.

Tabela 8 – Genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ com marcadores AFLP segundo a combinação enzima/*primer* e as respectivas análises de segregação para marcadores em dose única (MDU) usando a correção de Bonferroni

(Continua)

Código	Número de locos		Taxa de polimorfismo (%)	MDU		
	amplificados	segregantes		1:1	3:1	Total
E31M49	172	28	16,2	20	2	22
E31M58	120	25	20,8	16	4	20
E32M48	122	26	21,3	16	3	19
E32M50	154	34	22,0	6	8	14
E32M60	137	35	25,5	25	9	34
E33M50	160	19	11,8	8	4	12
E33M58	143	25	17,4	24	8	32
E33M61	148	30	20,2	18	4	22
E35M47	162	31	19,1	21	6	27
E35M48	154	46	29,8	34	6	40
E35M59	123	29	23,5	15	6	21
E35M60	189	48	25,3	23	12	35
E35M62	156	42	26,9	20	13	33
E36M48	128	32	25,0	17	7	24
E37M50	118	34	28,8	22	10	32
E37M58	108	30	27,7	15	5	20
E37M62	147	38	25,8	20	10	30
E38M58	102	19	18,6	10	3	13
E39M60	153	44	28,7	26	7	33
E40M48	123	25	20,3	13	9	22

Tabela 8 – Genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ com marcadores AFLP segundo a combinação enzima/primer e as respectivas análises de segregação para marcadores em dose única (MDU) usando a correção de Bonferroni

(Conclusão)

Código	Número de locos		Taxa de polimorfismo (%)	MDU		
	amplificados	segregantes		1:1	3:1	Total
E40M50	133	25	18,7	8	5	13
E40M49	142	20	14,0	12	5	17
<i>Total</i>	3094	685	-	389	146	535
<i>Média</i>	140,6	31,1	-	17,7	6,6	24,3
<i>Amplitude</i>	102 - 172	19 - 48	-	6 - 34	2 - 13	12 - 40

Para as combinações enzimas de restrição/*primers* direcionadas ao retrotransposon *SURE*, a quantidade de locos amplificados variou de 44 a 74 (média de 59,5) e de locos segregantes de 9 a 27 (média de 14,5), com taxa de polimorfismo entre 16% a 36,5%. Após as análises de segregação, 65 marcas foram consideradas de dose única e as 22 restantes foram descartadas.

Tabela 9 – Genotipagem da população ‘IAC66-6’ x ‘TUC71-7’ com marcadores derivados do retrotransposon *SURE* segundo a combinação enzima/primer e as respectivas análises de segregação para marcadores em dose única (MDU) usando a correção de Bonferroni

Código	Número de locos		Taxa de polimorfismo (%)	MDU		
	amplificados	segregantes		1:1	3:1	Total
DILI	44	14	31,8	12	-	12
DILII	72	11	15,3	8	2	10
DISI	53	16	30,1	8	4	12
SILI	74	27	36,5	14	2	16
SISI	58	10	17,2	6	-	6
RISI	56	9	16,0	7	2	9
<i>Total</i>	357	87	-	55	10	65
<i>Média</i>	59,5	14,5	-	9,1	1,6	10,8
<i>Amplitude</i>	44 - 74	9 - 27	-	6 - 14	0 - 4	6 - 16

Os resultados de segregação apresentados nesse trabalho estão de acordo com aqueles descritos previamente na literatura por Da Silva et al. (1993), Hoarau et al. (2001), Aitken et al. (2005), Oliveira et al. (2007) e Alwala et al. (2008) com relação à constituição genética da cana-de-açúcar, ou seja, 70% dos marcadores correspondem a locos em dose única, quando se analisam populações F_1 derivadas de cruzamentos simples.

4.3 Validação molecular dos marcadores direcionados a sequências de retrotransposons

Os fragmentos selecionados para a validação foram excisados dos géis de poliacrilamida e reamplificados. Posteriormente, seus tamanhos foram confirmados por eletroforese em gel de agarose a 1% (p/v), como exemplificado na figura 10.

Figura 10 – Gel de agarose (1%) mostrando o padrão de amplificação dos fragmentos selecionados para validação. M: marcador de massa molecular (*Ladder*) de 100 pb; nas canaletas 1 a 10 estão os amplicons obtidos usando a combinação DISI (*DraISsapI*): fragmentos com 521, 412, 390, 337, 310, 208, 163, 160, 145 e 138 pb, respectivamente

Devido à mistura de fragmentos (*pools*) preparada para cada uma das combinações enzimas de restrição/*primers* direcionadas ao retrotransposon *SURE*, os marcadores que possuíam tamanho menor foram mais representados que os maiores e, por isso, obtiveram-se várias sequências correspondentes a uma mesma marca. Desse modo, só 27 marcadores foram sequenciados (48,2%).

Todas as 27 sequências obtidas foram analisadas por BLASTn no banco de dados do NCBI (Tabela 10). Similaridades foram observadas com o retrotransposon *SURE* (acesso DQ115032.1) e com outras sequências de *Oryza minuta* (acesso AC216031.1), milho (acesso EU969904.1), sorgo

(acessos AC169373.2 e EF115542.1) e das variedades de cana ‘R570’ (acesso GU080318.1), ‘SP80-3280’ (acesso AE009947.2) e ‘NCo 310’ (acesso AP006714.1). Na tabela 10 encontram-se detalhes sobre esses alinhamentos. No Anexo A estão disponíveis todas as sequências obtidas.

Nove marcadores apresentaram nenhuma similaridade com o banco de dados. Esses foram: SILI182, SILI240, SILI390, SILI412, DILII207, DILII350, DILII374, SISI375 e RISI280.

Tabela 10 – Análise das sequências de marcadores direcionados ao retrotransposon *SURE* por BLAST tomando por base o banco de dados do NCBI
(Continua)

Marcador	Descrição	Identidade (%)	Score (bits)	e-value
DISI138	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	94	93.5	1.10^{-16}
DISI145	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	96	87.9	7.10^{-15}
DISI160	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	97	128	7.10^{-27}
DISI163	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	98	97.1	1.10^{-17}
DISI208	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	98	99	6.10^{-18}
DISI310	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	97	492	5.10^{-136}
	<i>Oryza minuta</i> clone OM_Ba0016E09	84	239	7.10^{-60}
DISI337	<i>Zea mays</i> clone 337091 mRNA sequence	88	132	1.10^{-27}
	<i>Saccharum</i> spp. ‘R570’ clone BAC 148J07	90	102	1.10^{-18}
DISI390	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	99	593	2.10^{-166}
	<i>Oryza minuta</i> clone OM_Ba0016E09	84	274	2.10^{-70}
SISI158	<i>Saccharum</i> spp. ‘SP80-3280’ clone SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like	98	97.1	1.10^{-17}
SISI180	<i>Sorghum bicolor</i> clone SB_BBc0188M08	88	147	2.10^{-32}

Tabela 10 – Análise das sequências de marcadores direcionados ao retrotransposon *SURE* por BLAST tomando por base o banco de dados do NCBI (Conclusão)

Marcador	Descrição	Identidade (%)	Score (bits)	e-value
SISI245	<i>Saccharum</i> spp. ‘SP80-3280’ clone	96	102	6.10^{-19}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
SISI330	<i>Saccharum</i> spp. ‘SP80-3280’ clone	98	97.1	4.10^{-17}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
RISI120	<i>Saccharum</i> spp. ‘SP80-3280’ clone	96	91.6	4.10^{-6}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
RISI195	<i>Saccharum</i> spp. ‘SP80-3280’ clone	96	99	6.10^{-18}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
DILI228	<i>Saccharum</i> spp. ‘SP-80-3280’ chloroplast	100	322	4.10^{-85}
	<i>Saccharum</i> spp. ‘NCo 310’ chloroplast DNA	100	322	4.10^{-85}
	<i>Sorghum bicolor</i> cultivar BTx623 chloroplast	99	316	2.10^{-83}
DILI272	<i>Saccharum</i> spp. ‘SP80-3280’ clone	97	366	3.10^{-98}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
DILI415	<i>Saccharum</i> spp. ‘SP80-3280’ clone	98	363	5.10^{-97}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
DILI605	<i>Saccharum</i> spp. ‘SP80-3280’ clone	96	379	9.10^{-102}
	SCCCCL6002A07 transposon <i>SURE</i> , Tnt1-like			
	<i>Saccharum</i> spp. ‘NCo 310’ chloroplast DNA	100	339	2.10^{-89}
	<i>Sorghum bicolor</i> cultivar BTx623 chloroplast	99	333	7.10^{-88}

Esses resultados são condizentes com o esperado, uma vez que a maior parte das sequências obtidas corresponderam ao próprio *SURE*. Além disso, os alinhamentos com as sequências de *Oryza minuta*, milho, sorgo e com as variedades de cana foram obtidos a partir dos marcadores que apresentavam tamanho superior a 200 pb. Isso indica que essas sequências são referentes às regiões genômicas adjacentes ao *SURE* que, embora desconhecidas, foram similares a regiões de gramíneas (incluindo a cana-de-açúcar).

Um fato interessante foi observado nos casos em que obtiveram-se várias sequências de uma mesma marca. O alinhamento dessas sequências com o banco de dados mostrou diferentes resultados. Frente a isso, elas foram alinhadas entre si utilizando-se o programa *Clustal X* para Windows, versão 1.8 (THOMPSON et al., 1997). Como resultado, confirmou-se que eram de fato fragmentos de mesmo tamanho.

Por exemplo, isso aconteceu com o marcador DILI228, para o qual obtiveram-se duas sequências diferentes (Figura 11). A sequência DILI228a apresentou elevada similaridade com sequências das variedades cana ‘SP80-3280’ e ‘NCo 310’ e com sorgo, como dito anteriormente. Já a sequência DILI228b não apresentou similaridade. O alinhamento pelo *Clustal X* mostrou a existência de dois padrões de sequências, com exceção das extremidades que correspondem aos *primers*. Isso mostra que se trata de marcas distintas.

Figura 11 – Alinhamento das sequências referentes ao fragmento DILI228, realizado pelo programa *Clustal X* (THOMPSON et al., 1997). A sequência DILI228a apresenta um padrão de nucleotídeos diferente da sequência DILI228b, mostrando que se trata de dois fragmentos

É provável que esse fato tenha ocorrido por causa da comigração de fragmentos durante a eletroforese. Outra possibilidade é que essas sequências correspondam a um mesmo loco em cromossomos homeólogos os quais, devido à poliploidia exacerbada, apresentam variações em suas bases nucleotídicas.

4.4 Mapa de ligação

Para as análises de ligação, as 600 marcas em dose única obtidas a partir de AFLP e das marcas direcionadas ao retrotransposon *SURE* foram designadas de acordo com a origem do

polimorfismo nos genitores para o tipo de cruzamento do loco marcador, seguindo a notação de Wu et al. (2002), que também é adotada pelo *software* OneMap. Desse modo, os locos oriundos do genitor ‘IAC 66-6’ foram denominados por D1.13, sendo identificados pela configuração ‘*ao* x *oo*’, em que *a* corresponde ao alelo amplificado e *o* corresponde ao alelo nulo. Os locos oriundos do genitor ‘TUC 71-7’ denominaram-se D2.18, com configuração ‘*oo* x *ao*’. Aqueles provenientes de ambos os genitores (biparentais) denominaram-se C.8, apresentando configuração ‘*ao* x *ao*’. Assim, foi montada uma matriz de genotipagem contendo, para cada loco marcador, a informação alélica dos 188 indivíduos.

Junto a esses dados, foram adicionados 57 marcadores SSRs para as análises de ligação. Estes marcadores foram analisados na população de estudo e fornecidos por T. B. Rodrigues (dados não publicados).

Neste trabalho, nenhum marcador com dados perdidos foi descartado. Uma vez que o tamanho da população estudada é grande, a presença de informação para cerca de 50% dos genótipos ainda é suficiente para realizar as análises de ligação. Além disso, o fato do *software* OneMap realizar análises multiponto através da cadeia de Markov oculta possibilita, entre outras vantagens, a recuperação da informação sobre dados perdidos.

A construção do mapa de ligação iniciou-se com 235 (35,8%) marcadores do tipo D1.13, 250 (38 %) do tipo D2.18 e 172 (26,2%) do tipo C.8. A partir dos critérios estabelecidos para a formação dos grupos, 503 (76,6%) marcadores formaram 91 grupos de ligação e os 154 (23,4%) restantes não foram alocados no mapa (Tabela 11). Desses 91 grupos de ligação, 39 (42,8%) foram constituídos por apenas 2 marcadores, 30 (33%) grupos por 3 a 5 marcadores e 22 (24,2%) grupos por 6 até 45 marcadores.

Esses dados condizem com os demais trabalhos envolvendo a construção de mapas genéticos em cana-de-açúcar, como exemplo dos mapas para as variedades ‘Q117’ e ‘MQ77-340’ construídos por Reffay et al. (2005) e da população F_1 ‘SP80-180’ x ‘SP80-4966’ por Garcia et al. (2006), que relatam a grande quantidade de marcadores não alocados, bem como a ampla variação entre quantidades de marcas presentes nos grupos de ligação. Isso demonstra uma das dificuldades de se estabelecer mapas genéticos para a cana e reforça a necessidade de um elevado número de marcas para a saturação dos grupos de ligação.

Tabela 11 – Número inicial e número de marcadores alocados no mapa da população F_1 derivada do cruzamento entre ‘IAC66-6’ x ‘TUC71-7’. Para cada técnica, é mostrado o número de marcadores para cada configuração, de acordo com a notação de Wu et al. (2002)

	Número inicial de				Número de marcadores			
	marcadores em dose única			Total	alocados no mapa			Total
	D1.13	D2.18	C.8		D1.13	D2.18	C.8	
AFLP	197	192	146	535	156	166	82	404
SURE	23	32	10	65	21	30	5	56
SSR	15	26	16	57	14	20	9	43
Total	235	250	172	657	191	216	96	503

De modo geral, para os 22 grupos que possuíam mais de seis marcadores, o algoritmo de ordenação que resultou nas melhores ordens foi o ‘ORDER’, que forneceu mapas com maior verossimilhança.

Nesses casos, a análise dos *heatmaps*, implementada no *software* OneMap, foi crucial para diagnosticar marcadores que apresentavam ligações espúrias, devido aos seus baixos valores de LOD-scores e altas estimativas de fração de recombinação. Esses marcadores foram removidos dos respectivos grupos, resultando em melhores ordenações entre as marcas remanescentes. Isto ocorreu com os grupos de ligação 19, 26, 64 e 66. Nos grupos de ligação 1, 12, 16 e 39, a remoção de uma marca espúria resultou na subdivisão dos mesmos. Mas, somente no caso do grupo 1, a marca removida pôde ser adicionada em um dos seus subgrupos pelo comando ‘try’. Como resultado dessas análises, nove marcadores foram removidos de sete grupos de ligação e descartados.

Para os grupos de ligação 6, 11, 21, 26, 31 e 35, a matriz gráfica identificou marcas biparentais (tipo C.8) fortemente ligadas a blocos de marcadores D1.13 e D2.18. Apesar das marcas biparentais serem responsáveis pela integração das marcas oriundas de cada genitor, a ordenação final apresentou grupos com *gaps* e de elevado tamanho. Esses grupos foram então divididos manualmente e as marcas biparentais foram mantidas nas subdivisões, sugerindo que elas formam um mesmo grupo, porém, até o presente momento, não há informação suficiente para agrupá-los. Exemplos de *heatmaps* são apresentados na Figura 12.

A identificação dos grupos de ligação divididos por falta de informação consistiu na adição de uma letra ao número do grupo original. As marcas biparentais em comum foram

ligadas por traços. Essa situação também ocorreu em cinco grupos de ligação da população derivada do cruzamento entre as variedades brasileiras ‘SP80-180’ e ‘SP80-4966’, para a qual também foi construído um mapa integrado (OLIVEIRA et al., 2007).

Figura 12 – *Heatmaps* para quatro grupos de ligação da população ‘IAC66-6’ x ‘TUC71-7’. a) matriz gráfica do grupo de ligação 5, mostrando a uniformidade de cores decorrente da melhor ordenação obtida para os marcadores; b) formação de blocos de marcadores D1.13 e D2.18 ligados por uma marca biparental (seta) no grupo de ligação 35; c) marca espúria (seta) que, após removida, resultou na subdivisão do grupo de ligação 39; d) marca espúria (seta) que, após removida, resultou numa melhor ordem final para o grupo 66

O mapa genético final da população em estudo (Figura 13) apresentou 107 grupos de ligação, com um comprimento total de 4.316,5 cM, variando de 0,1 (grupos 30, 54, 57 e 70) a 272,5 (grupo 16a) cM, com média de 40,34 cM por grupo de ligação. A densidade de marcadores foi de 8,74 cM, sendo que o maior intervalo foi de 61,9 cM (grupo 16a).

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 1 a 10). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em *itálico*, são indicados os SSRs e, em **negrito**, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continua)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 11a a 16b). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em itálico, são indicados os SSRs e, em negrito, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continuação)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores 'IAC 66-6' e 'TUC 71-7' (grupos 16c a 26b). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em *itálico*, são indicados os SSRs e, em **negrito**, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continuação)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 27 a 38). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em itálico, são indicados os SSRs e, em negrito, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continuação)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 39a a 57). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em itálico, são indicados os SSRs e, em negrito, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continuação)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 58 a 82). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em itálico, são indicados os SSRs e, em negrito, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Continuação)

Figura 13 – Mapa de ligação integrado para a população derivada do cruzamento entre os genitores ‘IAC 66-6’ e ‘TUC 71-7’ (grupos 83 a 91). À esquerda estão as posições estimadas e à direita o código atribuído aos marcadores. O arcabouço é de AFLP; em itálico, são indicados os SSRs e, em negrito, os marcadores direcionados ao retrotransposon *SURE*. As letras adicionadas aos números dos grupos de ligação indicam que estes foram subdivididos por falta de informação genética suficiente

(Conclusão)

O número de grupos de ligação obtido foi próximo à quantidade de cromossomos existentes nos híbridos de cana-de-açúcar, que apresentam entre 100 a 130 cromossomos (PRICE, 1963). De acordo com Garcia et al. (2006), a existência de *gaps* nos grupos de ligação é esperada, uma vez que apenas os marcadores em dose única foram utilizados para o mapeamento. Isso resulta numa limitação do mapa que tem cobertos apenas regiões onde esses marcadores são encontrados (AITKEN et al., 2007; DA SILVA et al., 1995). Além disso, existe a dificuldade de se acessar regiões cromossômicas herdadas de *S. officinarum*, como verificado no mapa de ligação da cultivar ‘R570’ (ROSSI et al., 2003).

A ligação entre as marcas oriundas dos genitores ‘IAC 66-6’ e ‘TUC 71-7’ (D1.13 e D2.18, respectivamente) por marcadores biparentais (C.8) ocorreu em 10 grupos de ligação: 3, 6c, 12a, 16a, 16d, 17, 23, 28, 40 e 53, além dos grupos 11, 21, 26, 31 e 35 que, como relatado anteriormente, foram divididos manualmente por falta de informação suficiente para agrupá-los. No presente trabalho, a integração das marcas foi mais evidente em relação ao trabalho de Oliveira et al. (2007) com a população ‘SP80-180’ x ‘SP80-4966’, o que pode ser um fato inerente à população estudada, como também ao incremento das análises multiponto pelo *software* OneMap.

Apesar dos marcadores AFLP serem anônimos, eles revelaram uma grande quantidade de locos polimórficos, o que possibilitou a construção do mapa genético apresentado. Com destaque para os marcadores direcionados ao retrotransposon *SURE*, cuja metodologia foi utilizada pela

primeira vez em cana-de-açúcar, neste trabalho, eles foram mapeados em 25 dos 107 grupos de ligação. O percentual de marcas desse retrotransposon alocadas no mapa foi de 86,15%, taxa essa superior às dos demais tipos de marcadores (AFLP e SSR), que tiveram 75,5% de marcas alocadas. Isso prova que os marcadores direcionados a retrotransposons são úteis para se realizar análises de ligação.

O mapeamento genético desses marcadores também é importante pois os retrotransposons exercem papel relevante na constituição dos genomas, sendo frequentemente encontrados em regiões flanqueadoras a genes de plantas. Como se conhece pouco sobre a distribuição dos elementos transponíveis nos genomas, a construção de mapas de ligação incorporando esses marcadores é uma valiosa estratégia para aprimorar esses estudos (WAUGH et al., 1997).

No presente trabalho, como apresentado na figura 13, os marcadores direcionados ao retrotransposon *SURE* não foram uniformemente distribuídos nos grupos de ligação. Por exemplo, em seis grupos (12b, 21a, 23, 26a, 39a, 39b) foram mapeadas mais de uma marca e em oito grupos de ligação (12a, 16a, 26b, 31b, 33, 36, 49 e 89) esses marcadores foram mapeados muito próximos, sugerindo a formação de *clusters*. Essa distribuição é completamente aceitável para o contexto genômico em que se encontra o retrotransposon *SURE*, pois dá uma idéia da sua localização, considerando-se o seu baixo número de cópias.

A metodologia para o desenvolvimento de marcadores direcionados a sequências de retrotransposons apresentada nesse trabalho é muito similar ao SSAP, que também poderia ter sido utilizada para o mesmo fim. No entanto, a técnica aqui proposta apresentou algumas vantagens, como a obtenção de estimativas sobre o número de cópias do elemento nos genótipos e a cobertura de diferentes regiões do mapa com esses marcadores.

Ainda considera-se importante a tarefa de desenvolver marcadores direcionados a sequências do retrotransposon *Garapa*, visando complementar as análises de mapeamento. Pelo fato do retrotransposon *Garapa* apresentar elevado número de cópias no genoma da cana-de-açúcar, acredita-se que, diferentemente do *SURE*, os marcadores direcionados ao *Garapa* apresentem uma ampla distribuição nos grupos de ligação, conforme demonstrado por hibridização *in situ* desse elemento nos cromossomos de cana-de-açúcar (DOMINGUES, 2009).

O estudo do retrotransposon *BARE-1*, realizado em cevada por Waugh et al. (1997) e em trigo por Queen et al. (2004), mostrou que as marcas para esse elemento foram bem distribuídas

ao longo dos mapas. É importante ressaltar que, como o *Garapa*, o retrotransposon *BARE-1* apresenta mais de 1000 cópias nesses genomas. Isso suporta a hipótese sobre a ampla distribuição dos marcadores direcionados ao retrotransposon *Garapa* ao longo dos grupos de ligação, o que também seria útil para aumentar a saturação do mapa.

No que concerne à homologia dos grupos de ligação, tal análise pôde ser realizada a partir dos marcadores derivados de um mesmo loco SSR, como descrito por Da Silva et al. (1993). Assim, com base nos 41 marcadores SSRs alocados no mapa (provenientes de 15 locos diferentes de SSRs), 30 grupos de ligação puderam ser associados em 10 grupos de homologia (Tabela 12).

Tabela 12 – Grupos de ligação da população ‘IAC66-6’ x ‘TUC71-7’ em seus respectivos grupos de homologia

Grupo de homologia	Grupos de ligação
I	19, 26a, 26b, 32
II	9, 16c, 35a, 35b, 69, 81
III	10, 12b, 21a, 21b, 63
IV	10, 44, 76
V	6a, 6c, 8, 12a, 74
VI	28
VII	31a, 39a
VIII	53
IX	90, 91
X	59

Como nas variedades modernas de cana-de-açúcar o conjunto haplóide de cromossomos é comumente igual a 10 (GRIVET; ARRUDA, 2002) pode-se inferir que esse resultado está de acordo com o esperado, embora haja a necessidade de incrementar tal análise. Isso porque poucos marcadores SSRs foram aqui utilizados, amostrando uma quantidade pouco representável de grupos de ligação.

A identificação precisa dos grupos homólogos poderá ser realizada, num próximo trabalho, pela saturação do mapa com mais marcadores de natureza codominante e, também, pela adição de marcas em doses dupla e tripla, que estão presentes em mais de um cromossomo homeólogo.

5 CONCLUSÕES

Marcadores AFLP são fundamentais para o estabelecimento de um mapa de ligação integrado em populações F_1 de cana-de-açúcar, servindo como arcabouço para alocação de outras marcas e contribuindo para o aumento da sua densidade, dado o grande tamanho do genoma da cana;

A metodologia de *NBS-profiling* pode ser adaptada para o desenvolvimento de marcadores direcionados a sequências de retrotransposons, tal como proposto neste trabalho;

O mapeamento de marcadores direcionados a sequências de retrotransposons acrescenta valiosas informações sobre o número de cópias desses elementos no genoma da espécie, bem como sobre a sua distribuição ao longo dos grupos de ligação;

A incorporação da abordagem multiponto no *software* OneMap permite a construção de mapas genéticos integrados com estimativas mais precisas das frações de recombinação e melhor ordenação das marcas.

REFERÊNCIAS

AITKEN, K.; LI, J.; WANG, L.; QING, C.; FAN, Y.H.; JACKSON, P. Characterization of intergeneric hybrids of *Erianthus rockii* and *Saccharum* using molecular markers. **Genetic Resources and Crop Evolution**, Dordrecht, v. 54, p. 1395-1405, 2007.

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**, Berlin, v. 110, p. 789-801, 2005.

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**, Wageningen, v. 164, p. 37-51, 2008.

AMALRAJ, V.A.; BALASUNDARAM, N. On the taxonomy of the members of 'Saccharum complex'. **Genetic Resources and Crop Evolution**, Dordrecht, v. 53, p. 35-41, 2006.

ANUÁRIO BRASILEIRO DA CANA-DE-AÇÚCAR. Santa Cruz do Sul: Editora Gazeta Santa Cruz, 2008. 127 p.

ARAÚJO, P.G.; ROSSI, M.; SACCARO JR, N.L.; KAJIHARA, D.; MASSA, R.; FELIX, J.M.; DRUMMOND, R.D.; FALCO, M.C.; CHABREGAS, S.M.; ULIAN, E.C.; MENOSSI, M.; VAN SLUYS, M.-A. Transcriptionally active transposable elements in recent hybrid sugarcane. **The Plant Journal**, Malden, v. 44, p. 707-717, 2005.

BIÉMONT, C.; VIEIRA, C. Junk DNA as an evolutionary force. **Nature**, New York, v. 443, n. 5, p. 521-524, 2006.

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

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.

BRANDES, E.W. Origin, classification and characteristics. In: ARTSCHWAGER, E.; BRANDES, E.W. (Ed). **Sugarcane** (*Saccharum officinarum* L.). U. S. Department of Agriculture Handbook, 1958. v. 122. p. 1-35.

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.

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.

BRONDANI, R.P.V.; WILLIAMS, E.R.; BRONDANI, C.; GRATTAPAGLIA, D. A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. **BMC Plant Biology**, London, v. 6, p. 20, 2006.

CORNMAN, R.S.; ARNOLD, M.L. Characterization and comparative analysis of sequence-specific amplified polymorphisms based on two subfamilies of IRRE retrotransposons in *Iris missouriensis* (Iridaceae). **Genetica**, Dordrecht, v. 135, n. 1, p. 25-38, 2009.

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.

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. spontaneum* sugarcane cultivars. **Journal of Experimental Botany**, Oxford, v. 55, n. 398, p. 847-854, 2004.

D'HONT, A.; GRIVET, L.; FELDMANN, P.; RAO, S.; BERDING, N., GLAUSZMANN, J-C. Characterization of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. **Molecular and General Genetics**, New York, v. 250, p. 405-413, 1996.

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.

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

DA SILVA, J.A.; SORRELLS, M.E. Linkage analysis in polyploids using molecular markers. In: JAUHAR, P. (Ed). **Methods of Genome Analysis in Plants**. Florida: CRC Press, 1996, p. 211-228.

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.

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

DOMINGUES, D. S. **SURE e GARAPA: Caracterização molecular e distribuição de dois retrotransposons com LTR de cana-de-açúcar**. 2009. 202 p. Tese (Doutorado em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.

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

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.

FESCHOTTE, C.; JIANG, N.; WESSLER, S.R. Plant transposable elements: where genetics meets genomics. **Nature Reviews Genetics**, London, v. 3, p. 329-341, 2002.

FLAVELL, A.; KNOX, M.R.; PEARCE, S.R.; ELLIS, T.H.N. Retrotransposon-based insertion polymorphisms (RBIP) for throughput marker analysis. **The Plant Journal**, Malden, v. 16, n. 5, p. 643-650, 1998.

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.

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, p. 122-127, 2002.

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**, Austin, v. 142, p. 987-1000, 1996.

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**, [s.l.], v. 2, p. 9-17, 2004.

HA, S.; MOORE, P. H.; 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 Molecular Biology**, Dordrecht, v. 39, p. 1165-1173, 1999.

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**, Berlin, v. 103, p. 84-97, 2001.

IGARASHI, M.; ABE, Y.; HATSUYAMA, Y.; UEDA, T.; FUKASAWA-AKADA, T.; KON, T.; KUDO, T.; SATO, T.; SUZUKI, M. Linkage maps of the apple (*Malus x domestica* Borkh.) cvs 'Ralls Janet' and 'Delicious' include newly developed EST markers. **Molecular Breeding**, Dordrecht, v. 22, n. 1, p. 95-118, 2008.

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.

JONES, N.; OUGHAM, H.; THOMAS, H. Markers and mapping: we are all geneticists now. **New Phytologist**, Cambridge, v. 137, p. 165-177, 1997.

JORNAL DA CANA. Disponível em: <http://www.jornalcana.com.br/>. Acesso em: 15 fev. 2010.

KAZAZIAN JR, H. H. Mobile Elements: Drivers of Genome Evolution. **Science**, Washington, v. 303, p. 1626-1632, 2004.

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

KUMAR, A. The adventures of the Ty1-*copia* group of retrotransposons in plants. **Trends in Genetics**, Amsterdam, v. 12, p. 41-43, 1996.

KUMAR, A.; BENNETZEN, J.L. Plant retrotransposons. **Annual Review of Genetics**, Palo Alto, v. 133, p. 479-532, 1999.

LAKSHMANAN, P.; GEIJKES, R.J.; AITKEN, K.S.; GROF, C.L.; BONNET, G.D.; SMITH G.R. Sugarcane biotechnology: the challenges and opportunities. **In Vitro Cellular & Developmental Biology – Plant**, Columbia, v. 41, p. 345-363, 2005.

LAM, E.; SHINE JR, J.; DA SILVA, J.; LAWTON, M.; BONOS, S.; CALVINO, M.; CARRER, H.; SILVA-FILHO, M.C.; GLYNN, N.; HELSEL, Z.; MA, J.; RICHARD JR, E.; SOUZA, G.M.; MING, R. Improving sugarcane for biofuel: engineering for an even better feedstock. **Global Change Newsletter**, Stockholm, v. 1, p. 251-255, 2009.

LANDEL, M.G.A.; SILVA, M.A. As estratégias de seleção da cana em desenvolvimento no Brasil. **Visão Agrícola**, Piracicaba, v. 1, p. 18-23, 2004.

LANDER, E.; GREEN, P.; ABRAHAMSON, J.; BARLON, A.; DALEY, M.; LINCOLN, S.; NEWBURG, L. MAPMAKER: an interactive computer package for constructing primary genetic linkages maps of experimental and natural populations. **Genomics**, San Diego, v. 1, p. 174-181, 1987.

LU, Y.; D'HONT, A.; PAULET, F.; GRIVET, L.; ARNAUD, M.; GLASZMANN, J. C. Molecular diversity and genome structure in modern sugarcane varieties. **Euphytica**, Wageningen, v. 78, p. 217-226, 1994.

MACCHERONI, W.; JORDÃO, H.; DE GASPARI, R.; MOURA, G. L.; MATSUOKA, S. Development of a dependable microsatellite-based. **Sugar cane International**, [s.l.], v. 27, n.2, p. 8-13, 2009.

MACEDO, I.C. **A energia da cana de açúcar - Doze estudos sobre a agroindústria da cana-de-açúcar no Brasil e a sua sustentabilidade**. 2.ed. São Paulo: Berlendis & Vertecchia: UNICA – União da Agroindústria Canavieira do Estado de São Paulo, 2007. 235 p.

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

MATSUOKA, S.; FERRO, J.; ARRUDA, P. The Brazilian experience of sugarcane ethanol industry. **In Vitro Cellular & Developmental Biology – Plant**, Columbia, v. 5, p. 372-281, 2009.

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.

McINTYRE, C.L.; GOODE, M.; MONKS, T.; BONNETT, G.D. The Complex Genetic Structure of Sugarcane Limits Identification of Additional SNP-Defined Simplex Alleles in Microsatellite Loci. **Tropical Plant Biology**, New York, v.2, p. 133-142, 2009.

MENOSSE, M.; SILVA-FILHO, M.C.; VINCENTZ, M.; VAN SLUYS, M.-A.; SOUZA, G.M. Sugarcane functional genomics: gene discovery for agronomic trait development. **International Journal of Plant Genomics**, [s.l.], v. 2008, p. 1-11, 2008. DOI:10.1155/2008/458732

MING, R.; LIU, S.-C.; BOWERS, J.E.; MOORE, P.H.; IRVINE, J.E.; PATERSON, A.H. Construction of a *Saccharum* consensus genetic map from two interspecific crosses. **Crop Science**, Madison, v. 42, p. 570-583, 2002.

- 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**, Austin, v. 150, p. 1663–1682, 1998.
- 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.
- NCBI. **Natinal Center for Biotechnology Information**. Disponível em: <http://blast.ncbi.nlm.nih.gov/>. Acesso em: 10 mar. 2008.
- OLIVEIRA, E.J.; VIEIRA, M.L.C.; GARCIA, A.A.F.; MUNHOZ, C.F.; MARGARIDO, G.R.A.; CONSOLI, L.; MATTA, F.P.; MORAES, M.C.; ZUCCHI, M.I.; FUNGARO, M.H.P. An integrated molecular map of yellow passion fruit based on simultaneous maximum-likelihood estimation of linkage and linkage phases. **Journal of the American Society for Horticultural Science**, Alexandria, v. 133, p. 35-41, 2008.
- 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.
- PAPINI-TERZI, F.S.; ROCHA, F.R.; VÊNCIO, R.Z.N.; FELIX, J.M.; BRANCO, D.S.; WACLAWOVSKY, A.J.; DEL BEM, L.E; LEMBKE, C.G.; COSTA, M.D.L.; NISHIYAMA JR, M.; VICENTINI, R.; VINCENTZ, M.G.; ULIAN, E.C. Sugarcane genes associated with sucrose content. **BMC Genomics**, London, v. 10, p. 120, 2009.
- 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**, [s.l], v. 24, p. 565-566, 2001.
- PIPERIDIS, N.; JACKSON, P.A.; D'HONT, A.; BESSE, P.; HOARAU, J.; OURTOIS, B.; AITKEN, K.S.; McINTYRE, A.L. Comparative genetics in sugarcane enables structured map enhancement and validation of marker-trait associations. **Molecular Breeding**, Dordrecht, v. 21, p. 233-247, 2008.

PRICE, S. Cytogenetics of modern sugar canes. **Economic Botany**, Bronx, v. 17, p. 97-106, 1963.

QUEEN, R.A.; GRIBBON, B.M.; JAMES, C.; JACK, P.; FLAVELL, A.J. Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat. **Molecular Genetics and Genomics**, v. 271, n. 1, p. 91-97, 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 fev. 2010.

RAFALSKI, A. Applications of single nucleotide polymorphisms in crop genetics. **Current Opinion in Plant Biology**, London, v. 5, p. 94-100, 2002.

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

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. Nobilisation of sugarcane. **Proceedings International Society of Sugarcane Technologists**, [s.l.], 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 distribution and characterization of EST-derived resistente gene analogs (RGAs) in sugarcane. **Molecular Genetics and Genomics**, Berlin, v. 269, n. 3, p. 406-419, 2003.

ROSSI, M.; ARAÚJO, P.G.; VAN SLUYS, M.-A. Survey of transposable elements in sugarcane expressed sequence tags (ESTs). **Genetics and Molecular Biology**, Ribeirão Preto, v. 24, p. 147-154, 2001.

ROZEN, S.; SKALETISKY, H.J. Primer3 on the WWW for general users and for biologist programmers. In: KRAWETZ, S.; MISENER, S. **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. Totowa: Humana Press, 2000. p. 365-386.

SAMBROOK, J.; RUSSEL, D.W. **Molecular cloning: a laboratory manual**. 3rd ed. New York: Cold Spring Harbor Laboratory Press, 2001. 999 p.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, 12, v. 74, p. 5463-5467, 1977.

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 Agrônômico, 2008. p. 47-56.

SILVAROLLA, M.B. **Análise citológica de calos e plantas regeneradas in vitro de variedades de cana-de-açúcar** (*Saccharum spp*). 1989. 160 p. Tese (Mestrado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1989.

SILVAROLLA, M.B.; AGUIAR-PERECIN, M.L.R. Evaluation of chromosome number stability in two sugarcane varieties. **Revista Brasileira de Genética**, Ribeirão Preto, v. 17, n. 2, p. 237-342, 1994.

SODERLUND, C.; DESCOUR, A.; KUDRNA, D.; BOMHOFF, M.; BOYD, L.; CURRIE, J.; ANGELOVA, A.; COLLURA, K.; WISSOTSKI, M.; ASHLEY, E.; MORROWS, D.; FERNANDES, J.; WALBOTS, V.; YEISSO, Y. Sequencing, mapping, and analysis of 27,455 maize full-length cDNAs. **PLoS Genetics**, [s.l.], v. 5, n. 11, 13 p., 2009. DOI: 10.1371/journal.pgen.1000740

SORRELS, M. Development and Application of RFLPs in Polyploids. **Crop Science**, Madison, v. 32, p. 1086-1092, 1992.

SYED, N.; SORENSEN, A.; ANTONISE, R.; VAN DE WIEL, C.; VAN DER LINDEN, C.; VAN ‘T WESTENDE, W.; HOOFTMAN, D.A.P.; DEN NIJS, H.C.M.; FLAVELL, A.J. A detailed linkage map of lettuce based on SSAP, AFLP and NBS markers. **Theoretical Applied Genetics**, Berlin, v. 112, n. 3, p. 517–527, 2006.

SYED, N.H.; SURESHSUNDAR, S.; WILKINSON, M.J.; BHAAU, B.S.; CAVALCANTI, J.J.; FLAVELL, A.J. Ty1-copia retrotransposon-based SSAP marker development in cashew (*Anacardium occidentale* L.). **Theoretical Applied Genetics**, Berlin, v. 110, p. 1195-1202, 2005.

TAM, S.M.; CAUSSE, 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. **Journal of evolutionary biology**, Basel, v. 20, n. 3, p. 1056-1072, 2007.

TAM, S.H.; MHIRI, C.; VOGELAAR, A.; KERKVELD, M.; PEARCE, S.R.; GRANDBASTIEN, M-A. Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. **Theoretical and Applied Genetics**, Berlin, v. 110, n. 5, p. 819-831, 2005.

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUNGIN, F.; HIGGINS, DG. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, Oxford, p. 4876-4882, 1997.

TZVELEV, N.N. The system of grasses (Poaceae) and their evolution. **The Botanical Review**, Lancaster, v. 55, n. 3, p. 141-204, 1989.

UNICA. **União da Agroindústria de cana-de-açúcar**. Disponível em: <http://www.portalunica.com.br/>. Acesso em: 15 fev. 2010.

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.

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**, Woodbury, v. 13, n. 12, p. 2725-2735, 2003.

VIEIRA, M.L.C.; VELLO, N.A.; SILVA FILHO, M.C. Genética e Melhoramento Vegetal. In: MIR, L. (Ed). **Genômica**. São Paulo: Atheneu, 2004. p. 677-703.

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.

WAUGH, R.; McLEAN, K.; FLAVELL, A.J.; PEARCE, S.R.; KUMAR, A.; THOMAS, B.B.T.; POWELL, W. Genetic distribution of Bare-1-like retrotransposable elements in barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). **Molecular and General Genetics**, New York, v. 253, p. 687-694, 1997.

WICKER, T.; SABOT, F.; HUA-VAN, A.; BENNETZEN, J.L.; CAPY, P.; CHALHOUB, B.; FLAVELL, A.; LEROY, P.; MORGANTE, M.; PANAUD, O.; PAUX, E.; SANMIGUEL, P.; SCHULMAN, A.H. A unified classification system for eukariotic transposable elements. **Nature Reviews Genetics**, London, v. 8, p. 973-982, 2007.

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.

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

WU, S.S.; WU, R.; MA, C.X.; ZENG, Z.B.; YANG, M.C.; CASELLA, G. A Multivalent Pairing Model of Linkage Analysis in Autotetraploids. **Genetics**, Berlin, v. 159, p. 1339-1350, 2001.

ANEXOS

ANEXO A – Sequências dos marcadores derivados do retrotransposon *SURE*. Letras após o código do marcador só estão presentes nos casos em que obtiveram-se mais de uma sequência de uma mesma marca. As sequências dos *primers* foram removidas.

Pool 1: mistura de fragmentos da combinação *Dra*I^Ssap1

>DISI138a

TATAGATCCCAAATTTGAAGATAAGCTGATCAGTGGAATTGTAGAAGTTAAGTCCTGCTCCCA
GTTATCACTTAGTTATCTGTGCTTATTTATG

>DISI138b

TATAGATCCCACCGAGAACAGATGCCCAACGAACGACGACCAACGTTTCCTGCAAGTGGAGTTGAG
AAACTTCTTCTTTCTTTGTCTTCCTTCTTT

>DISI138c

TATGATCTACTTGAATGAGGATTTCTGAAGAAATCAATTTGCACCTTTTTCTAGTTTCTTTCTGA
TAAACTGAAAATCCTTTTCTGTGGCTGTGGGT

>DISI145a

TATAGATCCCAAATTTGAAGATAAGCTGATCAGTGGAATTGTAGAAGTTAAGTCCTGCTCCCA
GTTATCACTTAGTTATCTGTGCTTATTTGTG

>DISI145b

TATAGATCCCCAAAACCTTTCTGGTTGAAACATTTTTCATTTGGGTTTCATTTAGGACCTGAAATCT
ACATTTCAAATCTGGTAGAGTGAACATAAGGAGAATG

>DISI160

TATAGATCCCCAAAATACCTGTTATAAGCTAGAGCTGTGAACTCGTGAGGAGAAGCCTTATATGT
GTAGAAGTTAAGTCCTGCTCCCAGCTATCACTTAGTTATCTGTGCTTAATTATGTTTCTGCTATT
AAGGCCTTG

>DISI163a

TATAGATCCCAAAAGCAAGGCAAAAGTGGTTAATTATCTTGATTAAGGTAGACCCATTATATGTA
GAAGTTAAGTCCTGCTCCCAGTTATCACTTAGTTATCTGTGCTTAATTATG

>DISI163b

AAGCACCGAGGGCAGTTGGTACAGGATGTAGAATCAAAATATCCAGGCAAGTTGACCGACTAAGC
CAGAACCACCAAGATAAAATGACTCTCAAGTAACCGAGTTTTGGGATCTATA

>DISI208a

TATAGATCCCAAAACAAGGACTCTGCTTGTTCTTAGGAGGCATGTGAAGGTCTCCAACCATTTCC
TATGGTCATTAGCAGTTCTTACTCGAACATGTTATATCTATTATCTGTAGAAGTTAAGTCCTGCT
CCCAGTTATCACTTAGTTATCTGTGCTTAATTATG

>DISI208b

TATAGATCCCAAAAATCAAATATAAGGGTAGAAACATAATTGACCGTTCATGCTGCTGCATTTGA
AATTTTGTATTGAGTGGTTTAAATCAAATTTAAATTAGATTTAAAAATCTTTTGAAAAATAGCTT
TGAAAAATTAGAAAAAGAAACAGAAAAGGAATTC

>DISI310

GTCTTTTTTATAGCAATAGAACCCTATTTTTTCATTGACAGAAGCAAGCATATCAGGATCAAACAAC
ACTTCATCAGAGATTGTGTTGAAGAAGGCAGCATCAAGACTGAGTCCGTCGCCACAAATGATCAA
CTGGCAGACATTCTGACTAAGGCCTTAGGCAAGGCCAAGTTTGAAGACATGAGAAGCAAGATTGG
GATCATGAAGATCAAGGATGGGGGAAGCAGGTCTTAGGGGGAGAACTGTAGAAGTTAAGTCCTGC
TCCCAGTTATCACTTAGTTATCTGTGCTTATTTATG

>DISI337a

TTCAGTCTTGTGTATCTGAAGGGATTATATATGTTGTGTCTTCAGACAACCTTTTTTTTGTGAAT
TGTCTTCAGACAACCTTAGCTTCCTTGCAATTTCCACTTCAGGAACAGCAGAAGCTCCTTTCCGTA
CGAGTCACTTGTATCGCATCCCGAGAGACAATAGTGAGCCTGAGCATAAATAAGTAGCAATATTG
TTAGACGCTGCGAACACGAGGTTGATAAGTACCCTCGTGTATTCGTTTTGTGATGAGTGATTATT
AATGTGATCCGCGAAGTAGTTTTGGGATCTATA

>DISI337b

TATAGATCCCAAACTACCTCGCGGATCACATTAATAATCACTCATCACAAAACGAATACACAAG
GGTACTCATCAACCTCGTGTTTCGCAGCGTCTAACAATATTGCTACTTATTTATGCTCAGGCTCAC
TATTGTCTCTCGGGATGCGATACAAGTGAAGTACGTAAGGAGCTTCTGCTGTTTCTGAAGTG
GAAATGCAAGGAAGCTAAGGTTGTCTGAAGATAATTCACAAAAAAAAGGTTGTCTGAAGACACAA
CATATATAATCCCTTCAGATACACAAGACTGAA

>DISI390

TATAGATCCCATGGAAGTAGTAGAGCTGCGAGTGGATAACAAATCAGCAATAGCACTTGCAAAGA
ACCCTGTTTTTTCATGACAGAAGCAAGCATATCAGGATCAAACAACACTTCATCAGAGATTGTGTT
GAAGAAGGCAGCATCAAGACTGAGTCCGTGCGCCACAAATGATCAACTGGCAGACATTCTGACTAA
GGCCTTAGGCAAGGCCAAGTTTGAAGACATGAGAAGCAAGATTGGGATCATGAAGATCAAGGATG
GGGAAGCAGGTCTTAGGGGGAGAACTGTAGAAGTTAAGTCCTGCTCCCAGTTATCACTTAGTTA
TCTGTGCTTATTTATG

Pool 2: mistura de fragmentos da combinação SspILTR1

>SILI182

GTTCAGGCACTGCACAGTACGCACGCGGCACGCAAGTTTTCTGGCATGAACTTTGTGCGCATGATT
CAGAAACCATAATCAGTTCTCGGCACGCCCCCTGATGCCCTGCCGTTATGCTGGAGCAAAGTGG
GATCTATA

>SILI240

GCGTCGTCGCAGCCTGGGAACTGTTAATAACTCCAGAAAAGAAGACTTTTTTAACCACTTACTTAC
ATATATAATTGCTACTACTATTACTAGTAGTTACGTGCCTTGACAGCCTTGGTAACTAGGAGGAC
CCTGTACTCCCTCCGTATCGAATTTAGAACGTTTTGAACAAGATTTAGGTCAAATTGGGATCTAT
A

>SILI390

TGGCGGCGGGATAATAAGAAAGAGCGCCAAAGAAAACGGTTATCTCACAGTAGGCAGTGAGAAAG
AAAATGAAAGTGAAACCTTTATTTGTGGTAAATAAGATAGATGTGGTAAAGGACAGATTTAAATA
TAGGCTAATCTCAAATTCATTAATATAAAAGATAGGAGAAGGGGAGAGTGGGCAAAGTGCATTGA

GCTCTTAAAAAGGTGAGCCAAGTAAGGGGGAAAAACACAAGAGGGGAGGCCCTAGAAAGAAAAGAG
CAATTTGGGGCGGGGAGGAAATGACCGACCTGAGTGCCGATAAGTGCAGGAAGCGCAATTGGAAA
GAGGATAAGGAGACGGCTAAGTGAAGTGGCCGGAGACAGCAGCCGTATTACTCT

>SILI412

TATAGATCCCAATTGACATGCCATGTATATGTATAAATTGGAGGAGCAATCTGAGCATCTTTGGA
TTCATAGCTCAGGATTTTAGTTCTTGGCCTAGGAGATGGTGCACCTCGACAGACTGATACCCACA
CCATATTCTGTGCAACCATAATAGAGGTATCCACCTTTTGAATTAGCTTTGCAACTAAGCAGAAA
TCTTCTAGTAAAGACAGTCACATTTGCATAATACTATGAGGATGGTCCAGTTGTGATTGAGCGAT
TGAAATGTTGGGCAATTCTTCATTTAGGATTCAGCTGGCTTTTGTTTTTTGATGTTTGCCTAAAT
CTCTTCCACTTTGCATCCTTATCCGTAGCATGACAATGATTGCCCCA

Pool 3: mistura de fragmentos da combinação DraILTR2

>DILII207

TATAGATCCCAAAAAGAAATACTACCTTAGTATCTGGATGGATGCATAAAAGCATAAGTGTCGATA
ACAGTCCTATCGGCCAGAATCATACAAGGGTATCAGTGTTTCATAGAGCGCCGATGCAATAAACA
GGGTTTACAAATTTAGTAAAGATTGGATGGTC

>DILII350

TGGGTCGGTTCTTGGCCGGCGCTGCGGTCACTCTGAGGACTACTAGTCATATGGGATACTCGATT
CTCAACCCGAAAGTATAGATCCCAAAAAGAAATACTACCTTAGTATCTGGATGGATGCATAAAAGC
ATAAGTGTGATAACAGTCCTATCGGCCAGAATCATACAAGGGTATCAGTGTTTCATAGAGCGCC
GATGCAATAAACAGGGTTTACAAATTTAGTAAAGATTGGATGGTC

>DILII374

TCGAGAGTGTGCCGTGTAAGCACACTAAATCGCGAACCCCTATAAGGGAGACGCCCCGATATTAAA
GCTCTGTGACGCGGGAAAACGCGCGCGCGACGCTGTGCGCAGAGAGAAAGGAGAAGGGAAGAGAA
GCGCGAAAAGAGAGCGCGGGCGCGCCTAGGGGCGCTCGTGACGATGATACGGGTCACTGTGCGG
CGTAACCCCCCACACCCCCCGCGCGTATATGGGCCCGCTTCACGGGGGCGCCCGGGGGGGCA
TTTTCCGGGGAGAATGTGGGCGCGAACCCCATATATGTGTTATATTTTTTCAAAAAACATCTACA
ATATATGTATCCCGCGTCAGAGGAACACAAAACCCTGGAAT

Pool 4: mistura de fragmentos da combinação SspISsapI

>SISI158

TATAGATCCCAATTCAAACCTAGTTTTTACCCACAAAGAACATAAATTGTATCTATAGATTGTAG
AAGTTAAGTCCTGCTCCCAGTTATCACTTAGTTATCTGTGCTTAATTATG

>SISI180

TATAGATCCCAATTGTTTTGCTTGGGATTACACTAAAATGCCTGGATTAGACCGATCTATAGTTG
AACATCGGTTGCCTATTAAGTCTGGATTTTCGGCCACATCAGCAGCCAGCACGTCGATGCAATCCT
AATATAC

>SISI245

TATAGATCCCAATTTTAGTAAGGAGGTAGTATTTTGATTCCTTGGCAATCATATAATGAACCTTG
TCAAGTTTAAGTAAGAGAAAAATACAACCTCGTCTCTCGTTAACCGAGTAAAGTATGCATTAATAC
TGAACAAACTGGTGAGCACTGTAGAAGTTAAGTCCTGCTCCCAGTTATCACTTAGTTATCTGTGC
TTAATTATG

>SISI330

TATAGATCCCAATTGCCACCACTCCGACATCAACCTCCTTAATCCTCATCTGCTCTGTACTATGT
TAGTGAGTTAGCTAAGATAAATTATAGACAAATTAAGAACAAATAGAGTGATGGAGTAAAGCGC
TGAACCAGAGTAACTAAAATGCTTTTGAGAAATAACATTCTATTGAACTATCAATTGAGCGTGCA
ACAGTTATATGTAGCATGGCCAGGGCCTGAGGTGTATCTTGTAGAAGTTAAGTCCTGCTCCCAGT
TATCACTTAGTTATCTGTGCTTAATTATG

>SISI375

TCATGAGAGATCAGTTTGGGATAAAACCCAAAGTTAATACTTATTCTTTTCGGACTCCATACCCT
CCAGCATATGATTTAATTCATCTCCCAAATCGGTACAAGGTACTAGATTTACCAAATTTTCTGG
ACAAGATGATACATCAACGATGAAACACGTCAATCGCTTCATTATTCAATATGGAGAGGCAGCTA
GCAGAGATGAATTAAGAGTTTCGATTATTTTCATCATCTTTGTCTGGATCAGCATTCACGTGGTTC
ATATCATTGCCACCAAATTCATTATTACTTGGGCTGATCTAGAAAAACAATTTTCATAAATTGGG
ATCTATA

Pool 5: mistura de fragmentos da combinação RsaISsapI

>RISI120a

TATAGATCCCAACATTAATTGTAGAAGTTAAGTCCCGCTCCCAGTTATCACTTAGTTATCTGTGC
TTAATTATG

>RISI120b

TATAGATCCCAACACGTTAGACCATGGTTCGATTTTGAGTAATGTTGGAGCTACTATTATTCCAC
CAAAGATGG

>RISI195

CATAGATCCCAACACTATTTTCCTCTAAATCTTGCTTTCAATGAATACTTACTTGAGCGTGCTGA
TGTTCTTAAAAAAGAGTAACATGGAAAGGAAATGTAGAAGTTAAGTCCTGCTCCCAGTTATCAC
TTAGTTATCTGTGCTTAATTATG

>RISI280

GAAACAATTTAATGACAAGTTAAATATTAACATGAAGAATGTTTAACTGCATGAAATGATAACAG
ATTAATAGTATTTTCCTTGTAACCTTTTGGTGAGAGTTTTAATAAATAAACTCTACTCTGCATTA
TAAAGGAAATCCTCGCTGAGTCCACCAGGAGGTTTCCACACAAGCGTCAGCTTATCCGCACCTGT
CACCTGCAACAGGGTAAAACAAACCCTAAGTTGGGGATCTATA

Pool 6: mistura de fragmentos da combinação DraILTR1

>DILI228a

TATAGATCCCAAAAGCATATACGTGAGACACAATCTACTAATTTTTTCTTTGATCTATATCTCGT
CTACTAGTATTTATAATACTTCAGGAGCTAATGAACTATTTTAGTATAATTCAATTCTCTTAAT
TCTTCGGCAACCGCGCCAAAACCTCGAGTTCCTTTTGGATTTTCCTTTTGGATCAA

>DILI228b

TATAGATCCCAAAATAGAAATCAATTACCAGAGGTTTCTCAGATGTAGACTTGGAGAAGTCTAGA
TCCGCCAAAGTATAAGCCATAGAGGTAGCACCGATAGGCCGACACTCCCTCCAAACTCTTCCCTT
ACTCTCTCTCTCTATTTGTATACAAAGACTAGATCCTAATAAGGACTAATCTTCTCT

>DILI272

CTTAGTTATCTGTGCTTATTTATGTTTCTGCTATTAAGGCCTTGGGCAGCACAAAGTGAAGTGTGCG
AGTACCCAGGCAGTTAGCTGTTTAGTTAATGCTCTAATGAGCCATCAATGTAATCAATATGCTAC
TTATCTCCTATATATGAAAGAGCAGTCGCCTGGGTGTGTAATCCCTGGCTGAGTAGAAACACTCT
GTTCCCTCCAACAGGAATTTATTTTGAGATCTATACTTTTCGGGGTGAGAATCGAGTA

>DILI415

CTTAGTTATCTGTGCTTAATTATGTTTCTGCTATTAAGGCCTTGGGCAGCGCAAGTGAAGTGTGTG
AGTACCCAGGCAGTTAGCTGTTTAGTTAATGCTCTAATGAGCCATCAATGTAATCAATATGCTAC
TTATCTCCTATATATGAAAGAGCAGTCGCCTGGGTGTGTTATCCCTGGCTGAGTAGAAACACTCT
GTACCTCCAACAGTAATTCCTTTTAAAGTAGCGCATCTGTAGAACAATACTTAAAGTTACTGTACC
TTCTACTTCGTTTGATAGATCTAGTACAAGAAATGCTTTAAGATGATTTTGAATTAGAACAAGGA
ATAAATCCATTTCTTGTTCCCTCAACCCTTATTTTGGGATCTATA

>DILI605

CTCGATTCTCAACCCGAAAGTATAGATCCCAAAATAAGGGTTGAGGAACAAGAAATGGATTTATT
CCTTGTTCTAATTCAAAATCATCTTAAAGCATTTCTTGCTGCTAGATCTATCAAACGAAGTAGAAG
GTACAGTAACTTTAAGTTTTGTTCTACAGATGCGCTACTTAAAAAGAATTACTGTTGGAGGTACA
GAGTGTTTCTACTCAGCCAGGGATAACACACCCAGGCGACTGCTCTTTCATATATAGGAGATAAG
TAGCATATTGATTACATTGATGGCTCATTAGAGCATTAACATAACAGCTAACTGCCTGGGTACTC
ACACAGTCACTTGTGCTGCCCCAAGGCCTTAATAGCAGAAACATAACTAAGCACAGATAACTAAGT
GTAAACTGGGGGAGCAGGACTAGTCCTGCTCCCAGTTATCATTGATCAAAAAGGAAATCCAAAAG
GAACTCGAGTTTTTGGCGCGGTTGCCGAAGAATTAAGAGAATTGAATTATACTAAAAATAGTTTCA
TTAGCTCCTGAAGTATTATAAATACTAGTAGACGAGATATAGATCAAAGAAAAAATTAGTAGATT
GTGTCTCACGTATATGCTTTTGGGATCTATA