

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS (TEDE) NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico:**      ☐ **Dissertação**      ☒ **Tese**

### 2. Identificação da Tese ou Dissertação

|                                                                                                                     |                                                                                                                                                  |     |    |        |      |
|---------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|-----|----|--------|------|
| Autor (a):                                                                                                          | <b>CAMILA DE MARILLAC COSTA NUNES</b>                                                                                                            |     |    |        |      |
| E-mail:                                                                                                             | camilamarillac@hotmail.com                                                                                                                       |     |    |        |      |
| Seu e-mail pode ser disponibilizado na página? <input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não |                                                                                                                                                  |     |    |        |      |
| Vínculo empregatício do autor:                                                                                      | Bolsista                                                                                                                                         |     |    |        |      |
| Agência de fomento:                                                                                                 | Conselho Nacional de Desenvolvimento Científico e Tecnológico                                                                                    |     |    | Sigla: | CNPq |
| País:                                                                                                               | BRASIL                                                                                                                                           | UF: | GO | CNPJ:  |      |
| Título:                                                                                                             | MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR ( <i>Saccharum</i> spp.) UTILIZANDO MARCADORES DArT ( <i>DIVERSITY ARRAYS TECHNOLOGY</i> ) E MICROSSATÉLITES |     |    |        |      |
| Palavras-chave:    Poliplóide, <i>framework</i> , marcadores moleculares.                                           |                                                                                                                                                  |     |    |        |      |
| Título em outra língua:                                                                                             | QTL MAPPING IN SUGARCANE ( <i>Saccharum</i> spp.) USING MICROSATELLITE AND DArT ( <i>DIVERSITY ARRAYS TECHNOLOGY</i> ) MARKERS                   |     |    |        |      |
| Palavras-chave em outra língua:                                                                                     | Polyploid, <i>framework</i> , molecular markers.                                                                                                 |     |    |        |      |
| Área de concentração:                                                                                               | Genética e Melhoramento de Plantas                                                                                                               |     |    |        |      |
| Data defesa: (dd/mm/aaaa)                                                                                           | 30/04/2013                                                                                                                                       |     |    |        |      |
| Programa de Pós-Graduação:                                                                                          | Agronomia                                                                                                                                        |     |    |        |      |
| Orientador (a):                                                                                                     | Alexandre Siqueira Guedes Coelho                                                                                                                 |     |    |        |      |
| E-mail:                                                                                                             | coelho@agro.ufg.br                                                                                                                               |     |    |        |      |
| Co-orientador (a):*                                                                                                 |                                                                                                                                                  |     |    |        |      |
| E-mail:                                                                                                             |                                                                                                                                                  |     |    |        |      |

\*Necessita do CPF quando não constar no SisPG

### 3. Informações de acesso ao documento:

Concorda com a liberação total do documento    ☒ SIM      ☐ NÃO<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC da tese ou dissertação.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses e ou dissertações, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

\_\_\_\_\_  
Assinatura do (a) autor (a)

Data: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

<sup>1</sup>Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

**CAMILA DE MARILLAC COSTA NUNES**

**MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR (*Saccharum* spp.)  
UTILIZANDO MARCADORES DArT (*DIVERSITY ARRAYS TECHNOLOGY*) E  
MICROSSATÉLITES**

Tese apresentada ao Programa de Pós-Graduação em Agronomia, da Universidade Federal de Goiás, como requisito parcial à obtenção do título de Doutor em Agronomia, área de concentração: Genética e Melhoramento de Plantas.

Orientador:

**Prof. Dr. Alexandre Siqueira  
Guedes Coelho**

Goiânia, GO – Brasil  
2013

**Dados Internacionais de Catalogação na Publicação (CIP)  
GPT/BC/UFG**

|              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>N972m</b> | <p><b>Nunes, Camila de Marillac Costa.</b></p> <p><b>Mapeamento de QTL em cana-de-açúcar (<i>Saccharum</i> spp.) utilizando marcadores DArT (<i>Diversity Arrays Technology</i>) e Microsatélites [manuscrito] / Camila de Marillac Costa Nunes. - 2013.</b></p> <p><b>189 f. : figs., tabs.</b></p> <p><b>Orientador: Prof. Dr. Alexandre Siqueira Guedes Coelho.</b></p> <p><b>Tese (Doutorado) – Universidade Federal de Goiás, Escola de Agronomia, 2013.</b></p> <p><b>Bibliografia.</b></p> <p><b>1. Cana-de-açúcar – Poliploide. 2. Cana-de-açúcar – Marcadores moleculares. I. Título.</b></p> <p><b>CDU: 633.61:575.116.4</b></p> |
|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**CAMILA DE MARILLAC COSTA NUNES**

**MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR (*Saccharum* spp.)  
UTILIZANDO MARCADORES DArT (*DIVERSITY ARRAYS TECHNOLOGY*) E  
SSR**

Tese DEFENDIDA e APROVADA, em 30 de abril de 2013, pela Banca Examinadora constituída pelos membros:

---

Prof. Dr. Evandro Novaes  
- Membro Externo -

---

Prof. Dr. Monalisa Sampaio Carneiro  
- Membro Externo -

---

Dr. Tereza Cristina de Oliveira Borba  
- Membro Externo -

---

Prof. Dr. João Batista Duarte  
- Membro Interno -

---

Prof. Dr. Alexandre Siqueira Guedes Coelho  
Orientador – EA/UFG

Goiânia, Goiás  
Brasil

*“Justificados, pois, mediante a fé, temos paz com Deus por meio de nosso Senhor Jesus Cristo; por intermédio de quem obtivemos igualmente acesso, pela fé, a esta graça na qual estamos firmes. E não somente isto, mas também nos gloriamos nas próprias tribulações, sabendo que **a tribulação produz perseverança; e a perseverança, experiência; e a experiência, esperança.**” (Ro 5:1-4).*

**OFEREÇO** este trabalho aos meus pais, Luiza e Arivaldo, pela confiança e esperança que sempre depositaram em mim. Ao meu marido Gustavo, pelo amor e apoio revelados sobre a forma de uma paciência incondicional. À Cecília, presente de Deus, que tem dado sentido às nossas vidas e que tem transformado meus dias em uma alegria indescritível!

## **AGRADECIMENTOS**

À Escola de Agronomia da Universidade Federal de Goiás, pela oportunidade e incentivo.

Ao Programa de Pós-Graduação em Agronomia, por todo auxílio oferecido.

Ao Conselho Nacional de Científico e Tecnológico – CNPq, pela concessão da Bolsa de Estudos.

À Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa), pela concessão de recursos para realização dos experimentos.

Aos funcionários da Ridesa, especialmente ao Carlão, à Manu e ao Danilo, pelo auxílio na condução do experimento de campo.

À Centroálcool, por disponibilizar a área para a condução do experimento e mão-de-obra para auxílio durante o plantio e as avaliações fenotípicas.

Aos professores do Programa de Pós-graduação em Genética e Melhoramento de Plantas da EA-UFG, pelos conhecimentos compartilhados.

Aos professores da Escola de Agronomia, Prof. Dr. Evandro Novaes e às Prof. Dra. Priscila Rangel e Prof. Dra. Rita Ganga, pelas palavras de incentivo durante os nossos almoços e pela disposição em sempre ajudar.

Aos funcionários de todos os Setores e Unidades da UFG, pelo auxílio prestado, especialmente à Agnalva, por sempre se preocupar em propiciar um ambiente limpo e agradável.

Ao Professor Dr. Alexandre Siqueira Guedes Coelho, pela orientação, pelo auxílio e pela paciência incondicional frente às minhas dificuldades.

À Carol e ao Jarênio, por termos, de fato, adquirido a virtude do trabalho em equipe, e pela amizade e bom relacionamento que construímos ao longo de uma jornada de muitas dificuldades, mas também de muitas palavras de incentivo, troca de experiências, conversas, consolos e diversão.

À minha querida amiga Keylinha, pela amizade, pela confiança, pelo apoio e pela disposição incondicional em sempre me ajudar das mais diferentes formas: Muito Obrigada!

À Milena e à Clistiane, pelo companheirismo e ajuda incondicional na avaliação fenotípica. Sem mulheres fortes e dispostas como vocês eu jamais teria conseguido!

A todos os colegas do Laboratório de Genética e Genômica de Plantas da EA-UFG: Jarênio, Carol, Keyla, Lud, Milena, Isabela, Athur, Clistiane, Priscila, Juliane, Ivone, Bianca, Jocasta e demais estagiários e colegas da Escola de Agronomia da UFG, muito obrigada pela disposição em ajudar nos momentos de muito trabalho.

## SUMÁRIO

|                                                                               |           |
|-------------------------------------------------------------------------------|-----------|
| <b>RESUMO GERAL .....</b>                                                     | <b>9</b>  |
| <b>GENERAL ABSTRACT .....</b>                                                 | <b>11</b> |
| <b>1 INTRODUÇÃO GERAL .....</b>                                               | <b>13</b> |
| <b>2 REVISÃO DE LITERATURA .....</b>                                          | <b>15</b> |
| 2.1 A CULTURA DA CANA-DE-AÇÚCAR ( <i>Saccharum</i> spp.) .....                | 15        |
| 2.1.1 Aspectos gerais.....                                                    | 15        |
| 2.1.2 Importância econômica .....                                             | 18        |
| 2.1.3 Estratégias utilizadas no melhoramento genético da cana-de-açúcar ..... | 21        |
| 2.1.3.1 Nobilização e os primeiros híbridos interespecíficos .....            | 21        |
| 2.1.3.2 Melhoramento genético da cana-de-açúcar no Brasil .....               | 24        |
| 2.2 CARACTERIZAÇÃO GENÉTICA DA CANA-DE-AÇÚCAR .....                           | 28        |
| 2.2.1 Comportamento cromossômico em poliploides .....                         | 28        |
| 2.2.2 Detecção de polimorfismo .....                                          | 31        |
| 2.2.3 Marcadores DArT ( <i>Diversity Arrays Technology</i> ) .....            | 34        |
| 2.3 MAPEAMENTO GENÉTICO .....                                                 | 36        |
| 2.3.1 Importância e princípios do mapeamento genético .....                   | 36        |
| 2.3.1.1 Análise da segregação Mendeliana .....                                | 40        |
| 2.3.1.2 Análise de ligação .....                                              | 41        |
| 2.3.1.3 Obtenção dos grupos de ligação .....                                  | 43        |
| 2.3.1.4 Análise de ordenação .....                                            | 44        |
| 2.3.2 Mapeamento genético em cana-de-açúcar .....                             | 46        |
| 2.4 MAPEAMENTO DE QTL ( <i>Quantitative Trait Loci</i> ) .....                | 60        |
| 2.4.1 Princípios do mapeamento de QTL .....                                   | 60        |
| 2.4.1.1 Mapeamento por intervalo (MI) .....                                   | 64        |
| 2.4.1.2 Mapeamento por intervalo composto (MIC) .....                         | 65        |
| 2.4.2 Mapeamento de QTL em cana-de-açúcar .....                               | 66        |



|          |                                                                                                                                                                    |            |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>3</b> | <b>MAPEAMENTO GENÉTICO DE MARCADORES DArT<br/>(<i>Diversity Arrays Technology</i>) E SSR (Microssatélites) EM<br/>CANA-DE-AÇÚCAR (<i>Saccharum</i> spp.) .....</b> | <b>77</b>  |
|          | RESUMO .....                                                                                                                                                       | 77         |
|          | ABSTRACT .....                                                                                                                                                     | 78         |
| 3.1      | INTRODUÇÃO .....                                                                                                                                                   | 78         |
| 3.2      | MATERIAL E MÉTODOS .....                                                                                                                                           | 81         |
| 3.2.1    | Material vegetal .....                                                                                                                                             | 81         |
| 3.2.2    | Extração e quantificação do DNA genômico .....                                                                                                                     | 81         |
| 3.2.3    | Genotipagem dos marcadores SSR .....                                                                                                                               | 81         |
| 3.2.4    | Genotipagem dos marcadores DArT .....                                                                                                                              | 83         |
| 3.2.5    | Análise da segregação Mendeliana .....                                                                                                                             | 84         |
| 3.2.6    | Construção dos mapas de ligação .....                                                                                                                              | 84         |
| 3.2.7    | Estimação do tamanho do genoma do genótipo-elite<br>RB97327 .....                                                                                                  | 85         |
| 3.3      | RESULTADOS E DISCUSSÃO .....                                                                                                                                       | 86         |
| 3.3.1    | Genotipagem .....                                                                                                                                                  | 86         |
| 3.3.2    | Mapas genéticos de ligação .....                                                                                                                                   | 89         |
| 3.3.2.1  | Mapa de ligação do genitor feminino RB97327 .....                                                                                                                  | 90         |
| 3.3.2.2  | Mapa de ligação integrado .....                                                                                                                                    | 93         |
| 3.3.2.3  | Comparação entre os mapas do genitor RB97327 e integrado                                                                                                           | 96         |
| 3.4      | CONCLUSÕES.....                                                                                                                                                    | 102        |
| <b>4</b> | <b>MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR<br/>(<i>Saccharum</i> spp.) UTILIZANDO MARCADORES DArT<br/>(<i>Diversity Arrays Technology</i>) E SSR .....</b>             | <b>103</b> |
|          | RESUMO .....                                                                                                                                                       | 103        |
|          | ABSTRACT .....                                                                                                                                                     | 103        |

|         |                                                                   |     |
|---------|-------------------------------------------------------------------|-----|
| 4.1     | INTRODUÇÃO .....                                                  | 104 |
| 4.2     | MATERIAL E MÉTODOS .....                                          | 105 |
| 4.2.1   | <b>Material vegetal</b> .....                                     | 105 |
| 4.2.2   | <b>Dados moleculares</b> .....                                    | 106 |
| 4.2.3   | <b>Obtenção dos mapas <i>framework</i></b> .....                  | 107 |
| 4.2.4   | <b>Fenotipagem</b> .....                                          | 108 |
| 4.2.5   | <b>Análises estatísticas dos dados fenotípicos</b> .....          | 109 |
| 4.2.6   | <b>Análises visando à detecção de QTL</b> .....                   | 110 |
| 4.3     | RESULTADOS E DISCUSSÃO .....                                      | 111 |
| 4.3.1   | <b>Análises fenotípicas</b> .....                                 | 111 |
| 4.3.2   | <b>Mapas <i>framework</i></b> .....                               | 118 |
| 4.3.3   | <b>Detecção de QTL</b> .....                                      | 119 |
| 4.3.3.1 | População derivada da autofecundação de RB97327 .....             | 119 |
| 4.3.3.2 | População proveniente do cruzamento entre RB97327 e RB72454 ..... | 126 |
| 4.3.3.3 | Considerações gerais sobre os QTL identificados .....             | 132 |
| 4.4     | CONCLUSÕES .....                                                  | 135 |
| 5       | <b>CONCLUSÕES GERAIS</b> .....                                    | 136 |
| 6       | <b>REFERÊNCIAS</b> .....                                          | 137 |
|         | <b>APÊNDICES</b> .....                                            | 150 |

## RESUMO GERAL

NUNES, C. M. C. **MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR (*Saccharum* spp.) UTILIZANDO MARCADORES DArT (*Diversity Arrays Technology*) E MICROSSATÉLITES**. 2013. 190 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas)-Escola de Agronomia, Universidade Federal de Goiás, Goiânia, 2013.<sup>1</sup>

Os primeiros trabalhos de melhoramento genético em cana-de-açúcar envolveram cruzamentos entre espécies poliplóides, *Saccharum spontaneum* L. e *Saccharum officinarum* L., os quais originaram híbridos interespecíficos que foram sucessivamente retrocruzados com *S. officinarum*. Essa estratégia resultou em considerável aumento da complexidade do genoma, de modo que as variedades atuais apresentem elevados níveis de ploidia e heterozigose, além de aneuploidias. Tais características dificultam a compreensão do genoma da cana-de-açúcar e, conseqüentemente, representam um desafio para o desenvolvimento de estudos genéticos com esta cultura. Dentre os diferentes estudos de caracterização genética, o desenvolvimento de mapas genéticos é importante por fornecer informações acerca da estrutura do genoma de uma espécie. Neste trabalho foram obtidos os primeiros mapas de ligação para cana-de-açúcar utilizando marcadores DArT (*Diversity Arrays Technology*) e SSR (*Single Sequence Repeat*). Além disso, foram identificados marcadores significativamente associados aos caracteres envolvidos na produção de açúcar. Para a obtenção dos mapas foram utilizadas duas populações, sendo uma constituída por 81 genótipos derivados da autofecundação de uma planta da cultivar RB97327, e outra constituída por 91 genótipos oriundos do cruzamento RB97327 x RB72454. O DNA genômico foi extraído de gemas axilares. A genotipagem foi realizada a partir de vinte pares de *primers* SSR e 7.680 marcadores DArT. A análise de segregação mendeliana permitiu a distinção de 392 marcas DArT e 57 marcas SSR polimórficas na população de autofecundação, e 632 DArT e 79 marcas SSR polimórficas na população de fecundação cruzada, com segregação *single-dose*. Ambos os mapas foram obtidos através do *software* OneMap utilizando-se um valor crítico de *LOD-score* igual a 3,5 e de fração de recombinação igual a 0,3. No mapa associado à população de autofecundação, os 449 marcadores DArT e SSR polimórficos com segregação 3:1 foram utilizados para originar 95 grupos de ligação referentes à variedade RB97327. Esse mapa apresentou um comprimento total de 1.217,2 cM. O tamanho estimado do genoma de RB97327 foi de 10.540,9 cM, o que permite afirmar que o mapa obtido apresentou baixa cobertura (11,5%). Para a população derivada de cruzamento, os 711 marcadores DArT e SSR polimórficos com segregação 3:1 e 1:1 originaram 136 grupos de ligação e o mapa apresentou um comprimento total de 2.722,2 cM. Os marcadores SSR também possibilitaram a identificação de seis possíveis grupos de homeologia no mapa referente ao genitor feminino RB97327 e nove no mapa integrado. Para cada população, foram obtidos os mapas *framework* nos quais foram identificados marcadores DArT e SSR associados aos caracteres envolvidos na produção de açúcar. Em ambas as populações procedeu-se à identificação de QTL a partir de

---

<sup>1</sup>Orientador: Prof. Dr. Alexandre Siqueira Guedes Coelho. EA-UFG.

análises de marcas simples e de mapeamento por intervalo composto. Na população derivada de autofecundação foram identificados, pela análise de marcas simples, 63 marcadores significativamente associados às seis variáveis avaliadas: número de entrenós, número de colmos por planta, comprimento de colmos, diâmetro de colmo, peso médio de colmo e teor de sólidos solúveis (brix). Pelo mapeamento por intervalo composto, três QTL relacionados a diâmetro de colmo, comprimento de colmo e brix foram identificados. Na população proveniente do cruzamento RB97327 x RB724554, foram identificados pela análise de marcas simples, 60 marcadores significativamente associados às seis variáveis. Pelo mapeamento por intervalo composto identificou-se dois QTL relacionados ao diâmetro e ao comprimento de colmo.

*Palavras-chave:* poliploide, *framework*, marcadores moleculares.

## GENERAL ABSTRACT

NUNES, C. M. C. **QTL MAPPING IN SUGARCANE (*Saccharum* spp.) USING MICROSATELLITE AND DArT (Diversity Arrays Technology) MARKERS.** 2013.190 f. Thesis (PhD in Agronomy: Genetics and Plant Breeding)–Escola de Agronomia, Universidade Federal de Goiás, Goiânia, 2013.<sup>1</sup>

The first efforts in sugarcane breeding involved crosses between polyploid species, *Saccharum spontaneum* L. and *Saccharum officinarum* L.. These crosses produced interspecific hybrids that were successively backcrossed to *S. officinarum*. This strategy resulted in a considerable increase in the sugarcane genome complexity. Current varieties exhibit high levels of ploidy and heterozygosity, besides varying levels of aneuploidy. These properties make the understanding of sugarcane genome more difficult; and therefore present a challenge to the development of genetic studies with this culture. Among the different approaches to perform the genetic characterization of a species, the development of genetic maps is useful in providing information about its genomic structure. In this work, we report the first linkage maps for sugarcane using both DArT (*Diversity Arrays Technology*) and SSR (*Single Sequence Repeat*) markers. We identified markers significantly associated to characters involved in sugar production. Maps were obtained using two populations: one, consisting of 81 genotypes, was derived from the selfing of a single RB97327 plant; the other, consisting of 91 genotypes, was derived from the crossing RB97327 x RB72454. Genomic DNA was extracted from axillary buds. Genotypes for twenty pairs of SSR primers and 7680 DArT markers were identified. Using mendelian segregation analysis a total of 392 DArT and 57 SSR polymorphic markers, in the population of selfing, and 632 DArT and 79 SSR polymorphic markers, in the outcrossing population, were detected to be segregating as single-dose markers. Both maps were obtained using the OneMap software. Critical values for *LOD*-score of 3.5 and recombination fraction of 0.3 were chosen. In the map obtained with the selfing population, 449 polymorphic markers with 3:1 segregation were used to originate 95 linkage groups for the variety RB97327. This map had a total length of 1217.2 cM. The estimated size of the genome of RB97327 was 10540.9 cM, which suggests that the obtained coverage (11.5%) is still low. For the population derived from crossing, the 711 polymorphic markers with 3:1 and 1:1 segregation originated 136 linkage groups. The map showed a total length of 2722.2 cM. The SSR markers allowed the identification of six possible homeology groups for the female parent RB97327, and nine homeology groups for the integrated map. For each population, framework maps were produced which were then used to investigate putative associations between markers and characters involved in sugar production. QTL were found both using single marker analysis and composite interval mapping. In the population derived from selfing, using single marker analysis, 63 markers were significantly associated to six variables: number of internodes, number of stems per plant, stem length, stem diameter, stem weight and percentage of soluble solids (°Brix). Using composite interval

---

<sup>1</sup>Adviser: Prof. Dr. Alexandre Siqueira Guedes Coelho. EA-UFG.

mapping, three QTL related to stem diameter, length of stem and °Brix were identified. In the population derived from the cross RB97327 x RB724554, using single marker analysis, 60 markers were significantly associated with the same six variables. Using composite interval mapping, two QTL related to diameter and length of stem were detected.

*Key words:* polyploid, framework, molecular markers.

## 1 INTRODUÇÃO GERAL

A cana-de-açúcar (*Saccharum* spp.) é cultivada nas regiões tropicais e subtropicais e está entre as mais importantes espécies cultivadas por ser a principal fonte de sacarose e etanol, além de inúmeros subprodutos obtidos a partir do seu processamento. Atualmente o Brasil é o maior produtor mundial de cana-de-açúcar, sendo o setor sucroalcooleiro responsável por 2,4% do PIB brasileiro (Procana, 2011).

O crescente interesse pela cultura da cana-de-açúcar como matéria-prima para geração de bioenergia tem exigido estratégias que impulsionem a sua produção. Para isso, faz-se necessário promover incrementos da área plantada com cana-de-açúcar ou desenvolver variedades que apresentem características agronômicas e industriais cada vez mais favoráveis. Neste contexto, o melhoramento genético apresenta função preponderante por ser capaz de estabelecer estratégias que aumentem a produtividade por meio do lançamento de variedades selecionadas que atendam às necessidades vigentes (Cesnick & Mioque, 2004).

Resumidamente, o melhoramento genético da cana-de-açúcar tem sido baseado na seleção e clonagem de genótipos superiores, obtidos a partir de cruzamentos realizados entre genitores que, na maioria das vezes, constituem variedades comerciais. Apesar de esta metodologia ser relativamente simples, os caracteres de importância agronômica e comercial, normalmente avaliados durante a seleção fenotípica, são de natureza quantitativa, os quais, por resultarem em um padrão contínuo de variação, dificultam a identificação dos genótipos superiores, fazendo com que o processo de liberação de uma nova variedade, em média, leve de dez a doze anos (Matsouka et al., 2005).

O desenvolvimento de metodologias capazes de possibilitar a seleção precoce e eficiente de genótipos superiores, como as baseadas em marcadores moleculares, pode contribuir notadamente para a redução do tempo envolvido no lançamento de novas variedades e para um melhor entendimento da

complexidade genética que é marcante na cana-de-açúcar. Entre as estratégias a serem estabelecidas para que esses objetivos sejam alcançados estão a construção de mapas genéticos de ligação e a identificação de marcadores associados às regiões genômicas que controlam os locos de interesse agrônomo (QTL – *Quantitative Trait Loci*) (Liu, 1998).

Nesse contexto, com o intuito de fornecer informações que ampliem a compreensão sobre a forma em que o genoma da cana-de-açúcar está estruturado, e, conseqüentemente, minimizar o tempo envolvido no processo de seleção de genótipos superiores, o presente estudo teve como objetivos: (1) construir os primeiros mapas de ligação para cana-de-açúcar utilizando marcadores DArT (*Diversity Arrays Technology*) e SSR (*Single Sequence Repeat*), a partir de populações derivadas de autofecundação e de fecundação cruzada; (2) estimar o tamanho do genoma do genótipo-elite da Ridesa RB97327; (3) identificar marcadores DarT e SSR associados à produção de açúcar.



## **2 REVISÃO DE LITERATURA**

### **2.1 A CULTURA DA CANA-DE-AÇÚCAR (*Saccharum* spp.)**

#### **2.1.1 Aspectos gerais**

A teoria mais aceita sobre a origem da cana-de-açúcar considera que seja nativa das ilhas do arquipélago da Polinésia, sendo o seu centro de diversidade a Nova Guiné, onde se teve o primeiro contato com uma planta “associada intimamente com a essência da doçura”. Posteriormente, a cultura da cana-de-açúcar foi difundida para as ilhas do Pacífico, Índia e China, tendo sido desenvolvidas, nestes dois últimos locais, as técnicas de extração do açúcar. Daí em diante, a cana-de-açúcar passou a ser cultivada na Pérsia, no norte da África e nas ilhas do Atlântico (Daniels & Roach, 1987).

Nas Américas, a cana-de-açúcar foi introduzida na segunda expedição de Cristóvão Colombo, em 1493, tendo ocorrido a implementação da cultura primeiramente no Haiti. Desde aquela época, seu cultivo tem sido estendido às demais regiões tropicais e subtropicais do mundo (Machado, 2000).

No Brasil, as primeiras mudas de cana-de-açúcar, provenientes da Ilha da Madeira, foram introduzidas em 1532, na Capitania de São Vicente, por Martim Afonso de Souza. Era o princípio da indústria açucareira que, por ter encontrado no Brasil solos férteis e condições climáticas favoráveis, expandiu-se rapidamente (Miocque, 1993).

Após um início de tentativas frustrantes e de grandes dificuldades, a produção de açúcar prosperou no Brasil, sobretudo nas regiões que atualmente compreendem a Bahia e os demais estados do Nordeste, e, após menos de cinquenta anos, o país já detinha o monopólio mundial de produção. Mas, no século XVII, com o início da produção açucareira pelos ingleses e franceses no Caribe e com a descoberta do ouro em Minas Gerais, o açúcar deixou de ocupar o primeiro lugar na geração de riquezas, tendo sua produção retraída até o final do século XIX (Machado, 2000).

Somente após a abolição da escravatura, com o incentivo à imigração europeia para o interior paulista, foi que os senhores das glebas, principalmente de origem italiana, optaram pela produção de aguardente a partir da cana-de-açúcar, produto este de comércio fácil e altamente rentável. A partir daí, inúmeros engenhos concentraram-se em diversas regiões brasileiras, como Campinas, Piracicaba e Ribeirão Preto. Rapidamente o sudeste do Brasil tornou-se o maior centro produtor de açúcar do país. Concomitante à expansão paulista ocorreu também o surgimento de grande quantidade de usinas sucro-alcooleiras no Nordeste, principalmente nos estados de Alagoas e Pernambuco (Machado, 2000).

A cana-de-açúcar é uma planta alógama, semi-perene, de metabolismo C4 e adaptada a regiões subtropicais. Esta cultura desenvolve-se em forma de touceiras, sendo a parte aérea constituída por colmos que crescem próximos entre si, folhas e inflorescências. Os colmos apresentam nós, entrenós e gemas axilares, sendo estes órgãos importantes na identificação de variedades. A parte subterrânea é formada por raízes fasciculadas e por rizomas (Cesnick & Miocque, 2004).

A propagação comercial da cana-de-açúcar é realizada de forma vegetativa, através de toletes obtidos pelos cortes dos colmos. Este método, embora facilite a multiplicação de materiais, é também responsável por fazer com que, após longos ciclos de replicação das variedades comerciais, seja observada uma redução considerável no vigor e produtividade, exigindo periodicamente o seu rejuvenescimento ou, mesmo a substituição das variedades em uso (Silva et al., 2008).

Durante a maturação é possível, por vezes, observar a ocorrência de floração, fenômeno indesejável em cultivos comerciais em razão de promover considerável redução das reservas energéticas e, consequentemente, da produtividade. A ocorrência de floração em cana-de-açúcar depende de uma série de condições abióticas, como temperatura, luminosidade, altitude e pluviosidade, por exemplo, e constitui um importante evento para o melhoramento genético de plantas, pois possibilita a obtenção de sementes a partir da realização de cruzamentos entre genótipos que apresentam características desejáveis (Matsouka et al., 2005).

De acordo com Daniels & Roach (1987), a cana-de-açúcar pertence à família das Gramíneas (Poaceae), tribo Andropogoneae e gênero *Saccharum*. A família Poaceae é uma das maiores dentre todas as famílias de monocotiledôneas e desenvolveu-se há mais de 50 milhões de anos, compreendendo, aproximadamente, 10 mil espécies distribuídas em 650 gêneros (Glaszmann et al., 1997; Matsuoka et al., 1999).

Brandes (1958) afirmava que os híbridos de cana-de-açúcar cultivados foram obtidos a partir do cruzamento realizado entre seis espécies do gênero *Saccharum*, sendo duas selvagens, *Saccharum robustum* Brandes e Jeswiet ex Grassl ( $2n=60-205$ ) e *Saccharum spontaneum* L. ( $2n=40-128$ ); e quatro domesticadas, *Saccharum officinarum* L. ( $2n=80$ ), *Saccharum barberi* Jeswiet ( $2n=81-124$ ), *Saccharum sinense* Roxb. ( $2n=111-120$ ) e *Saccharum edule* Hassk. ( $2n = 60-80$ ). De acordo com Mukherjee<sup>1</sup>, citado por Daniels & Roach (1987), os gêneros *Saccharum*, *Erianthus*, *Sclerostachya* e *Narenga* constituem grupos inter cruzantes com os que originaram os híbridos de cana-de-açúcar e, por este motivo, cunhou-se o termo “Complexo *Saccharum*” referido-se a este grupo. Esta, no entanto, não é uma nomenclatura científica. Posteriormente, os gêneros *Miscanthus* e *Ripidium* foram também adicionados a este “Complexo”.

Estudos filogenéticos e citológicos permitiram observar em *S. officinarum* cromossomos em comum com *Miscanthus*, *Erianthus* e *Ripidium*, revelando que a cana-de-açúcar pode ter emergido da hibridização envolvendo diferentes gêneros, uma vez que já se observou também homologia cromossômica destes gêneros com *Sclerostachya* (Daniels & Roach, 1987). Para Brandes (1958), *S. officinarum* foi domesticada a partir de *S. robustum*, na Nova Guiné, e foi dispersa para hibridizar com *S. spontaneum*, originando assim as cultivares chinesas e do norte da Índia.

Estudos realizados com marcadores moleculares (D'Hont et al., 2002; Grivet et al., 2006), entretanto, têm contestado a hipótese de que as variedades atuais apresentem contribuições de outros gêneros, reforçando a ideia de que a cana-de-açúcar é originária apenas do gênero *Saccharum*. Alix et al. (1999), também verificaram ausência de hibridação de sondas específicas de *Miscanthus*

---

<sup>1</sup> MUKHERJEE, S. K. Origin and distribution of *Saccharum*. **Botanical Gazette**, Chicago, v. 119, p. 55-61, 1957.

e *Erianthus* sobre o DNA de diferentes cultivares atuais de cana-de-açúcar. Para D'Hont et al. (1996) e Besse et al. (1997), *Saccharum*, *Erianthus* e *Miscanthus* podem ser seguramente diferenciados e estes dois últimos não parecem ter contribuído para originar as variedades atuais de cana-de-açúcar.

### **2.1.2 Importância econômica**

A cana-de-açúcar é cultivada comercialmente em mais de oitenta países, sendo o Brasil, seguido respectivamente por Índia e China, responsável por 20% da produção mundial, o que garante a liderança na exportação dos seus principais subprodutos, o açúcar e o álcool. Na safra 2010/2011, em área de 8.033,6 mil hectares foram colhidas 625 milhões de toneladas de cana, das quais, aproximadamente, 46,2% foram destinadas à produção de açúcar e 53,8% à produção de álcool. Nessa mesma safra foram produzidos, aproximadamente, 38,6 milhões de toneladas de açúcar e 27,7 bilhões de litros de álcool, o que corresponde a, aproximadamente, por 37% do mercado internacional de açúcar e etanol (Conab, 2011).

Embora o setor sucroalcooleiro brasileiro ocupe posição de destaque, da década de 70 até os dias atuais, grandes dificuldades econômicas e políticas também já foram enfrentadas, passando por fases de “altos e baixos”. O Proálcool, Programa Nacional do Álcool, criado em 1975/1976, teve seu auge no ano de 1985, quando 96% da frota brasileira de automóveis era movida a álcool. As safras entre os anos de 1991 a 1996 permitiram que a produção de etanol atingisse 12,3 bilhões de litros, mas, posteriormente, o programa foi abandonado fazendo com que a produção de etanol e o incentivo ao cultivo da cana-de-açúcar declinassem consideravelmente (BNDES, 2008).

Nos últimos anos observa-se novamente aumento crescente, tanto na quantidade de usinas e destilarias, quanto na quantidade de cana-de-açúcar processada pelas usinas brasileiras. De acordo com Procana (2011), 420 destilarias e usinas compõem o setor sucro-alcooleiro brasileiro, além de 40 novos projetos em andamento. Na safra 1990/1991 foram processadas 222,5 milhões de toneladas de cana-de-açúcar, sendo, aproximadamente 60% dessa produção proveniente apenas do Estado de São Paulo, e 23,5% da região

Norte-Nordeste. Para a safra 2012/2013, estima-se que sejam processadas 595 milhões de toneladas de cana, quase o triplo do que era processado nos anos de 1990 (Conab, 2012).

Diversos fatores têm colaborado para incentivar a expansão da cultura da cana-de-açúcar no mundo. Entre eles, pode-se citar a conscientização de que o petróleo é uma fonte de energia esgotável, o apoio econômico-político para a utilização do álcool como biocombustível e o desenvolvimento de tecnologias que permitem o aproveitamento de praticamente “todos” os subprodutos da cultura. Atualmente, após o processamento da cana-de-açúcar, o bagaço, subproduto da moagem dos colmos, pode ser utilizado não apenas na alimentação animal, mas também como fonte de energia renovável e no aquecimento de caldeiras. Pode ainda ser utilizado como fonte de celulose e de fibra, como matéria-prima para madeira composta e na síntese de fibras de carbono. As cinzas, resultantes da queima do bagaço, têm sido utilizadas nas usinas como adubo. A vinhaça, subproduto da produção de etanol, é utilizada para irrigar canaviais funcionando como importante fonte de matéria orgânica e nutrientes inorgânicos (BNDES, 2008; Silva & Moraes, 2008).

O crescente interesse pela cultura da cana-de-açúcar como matéria-prima para geração de bioenergia tem exigido estratégias que impulsionem a sua produção. No Brasil, atualmente, a cana-de-açúcar está presente em quase todos os estados brasileiros e ocupa cerca de 9% da superfície agrícola do país, sendo o seu cultivo o terceiro mais importante em superfície ocupada, depois de soja e milho.

No Brasil, ainda que o Estado de São Paulo continue sendo o maior produtor de cana-de-açúcar, contribuindo com aproximadamente 54% da área total cultivada no país, a área plantada com a cultura em outros estados também tem colaborado significativamente à produção nacional. Destacam-se estados como Minas Gerais com 8,0%, Goiás com 7,5%, Paraná com 7,2%, Alagoas com 5,5%, Mato Grosso do Sul com 5,0% e Pernambuco com 4,2% (Conab, 2011). Diante disso é válido considerar que, embora o Brasil lidere a produção mundial de cana-de-açúcar e apresente áreas com significativo potencial de expansão, o aumento da produção desta cultura nos últimos anos tem ocorrido muito mais em função da expansão da fronteira agrícola do que do aumento da produtividade

média, estimada em 69,864 kg ha<sup>-1</sup> na safra 2012/2013. Para a safra de 2012/2013 a área cultivada com cana-de-açúcar destinada à atividade sucroalcooleira está estimada em 8.520,5 mil ha, o que representa um aumento de produção em função da expansão de área plantada de 2% em relação à safra anterior (Conab, 2012).

De acordo com Silva (2006), para atender à demanda projetada, o setor sucroalcooleiro deve continuar em crescimento, mas, para isso, a área agrícola necessita ampliar a capacidade de produção, tanto pela expansão de área, quanto pela obtenção de variedades melhoradas capazes de incrementar a produtividade. De acordo com Dal-Bianco et al. (2011), os ganhos em produtividade ocorridos nos últimos 35 anos foram de aproximadamente 66% e deveram-se a fatores como a liberação de variedades mais produtivas e à adequação das práticas agronômicas utilizadas na cultura.

De acordo com Dal-Bianco et al. (2011), recentes avaliações dos programas de melhoramento indicaram que, embora os ganhos em produtividade tenham sido alcançados, nota-se considerável aumento do número de variedades cultivadas, provavelmente em função de uma maior quantidade de ambientes explorados, e uma redução, cada vez mais pronunciada, dos ganhos em teores de sacarose, que têm variado entre 1,0% e 1,5% ao ano. Em 1995, cinco variedades ocupavam 70% das áreas brasileiras cultivadas com cana-de-açúcar. Em 2010, esse número praticamente dobrou.

Apesar disso, também se percebe que, da mesma forma que o número de variedades de cana-de-açúcar em uso tem aumentado, a similaridade genética entre tais variedades também. Tal fato evidencia a necessidade de aumento da base genética dos programas de melhoramento desta cultura, uma vez que a baixa diversidade genética entre as variedades atuais pode ser responsável pela redução de incremento no teor de sacarose nas variedades cultivadas nos últimos anos (Dal-Bianco et al., 2011).

Além do que já foi mencionado, o Brasil apresenta outras características que o tornam um país ainda mais promissor para o setor sucroalcooleiro. Entre elas estão o clima e o relevo favoráveis, boas perspectivas pluviométricas, localização favorável ao escoamento da produção e ampla área agricultável exigida pela cultura. Assim, manter a liderança mundial em

produtividade e exportação e garantir a sustentabilidade do agronegócio da cana-de-açúcar tem exigido, não apenas a implantação de novas unidades de usinas, mas também investimentos que visem aumentar o rendimento médio da cultura. Nesse contexto, a utilização de ferramentas biotecnológicas constitui uma eficiente estratégia a ser implementada junto ao melhoramento convencional da cana-de-açúcar, uma vez que estas ferramentas podem auxiliar em diversas etapas do desenvolvimento de novas variedades, além de permitirem um melhor conhecimento sobre a genética e a genômica da espécie (Bertrand & Mackill, 2008).

### **2.1.3 Estratégias utilizadas no melhoramento genético da cana-de-açúcar**

#### **2.1.3.1 Nobilitação e os primeiros híbridos interespecíficos**

Os primeiros híbridos interespecíficos de cana-de-açúcar foram obtidos em Java entre os anos de 1887 e 1893, a partir de cruzamentos realizados entre clones de *S. spontaneum* (Glagah x Loethers) e, também entre Kassoer, híbrido natural entre *S. officinarum* e *S. spontaneum*, e Black Cheribon, um clone de *S. officinarum*. Híbridos como POJ100, S.W.3, S.W.111 e E.K.28 puderam ser obtidos e, posteriormente, foram usados como genitores em diversos cruzamentos por terem promovido ganhos em produtividade e resistência a doenças (Bremer, 1961). Estes resultados demonstraram a importância da realização de cruzamentos controlados, viabilizando o melhoramento genético da cana-de-açúcar.

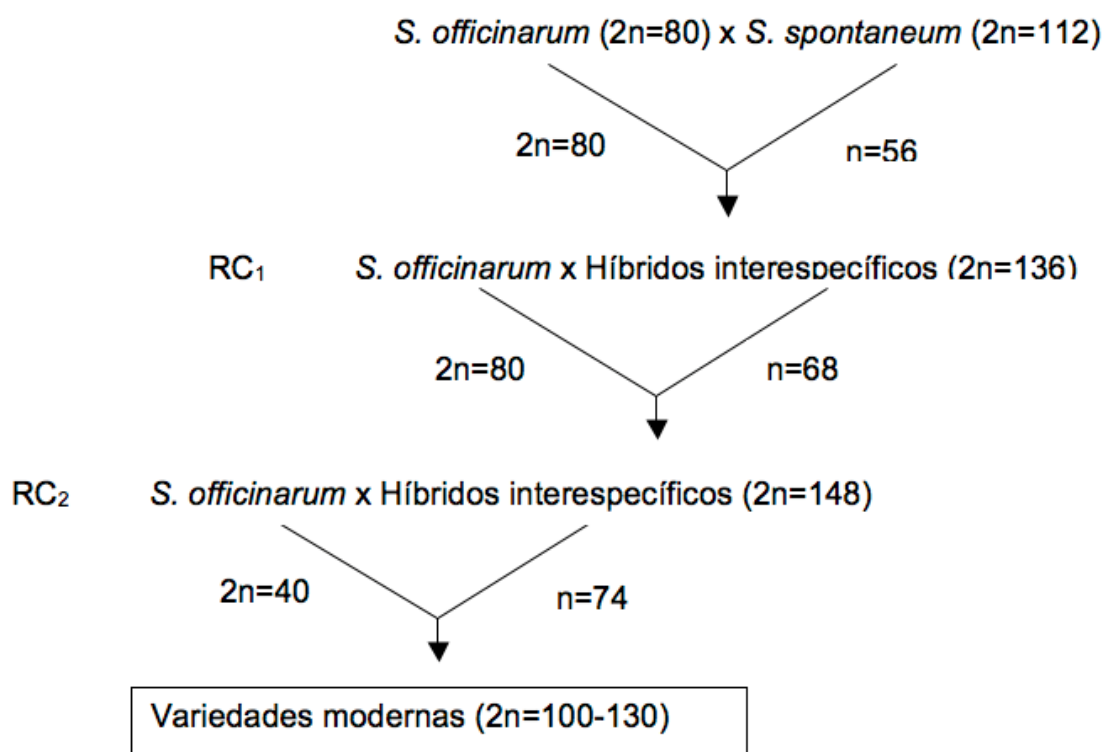
Diante desse histórico é possível perceber que tanto *S. officinarum* como *S. spontaneum* estavam envolvidas na obtenção dos primeiros híbridos de cana-de-açúcar e, por este motivo, são encontradas na genealogia de grande parte das variedades cultivadas atualmente. Progressos genéticos consideráveis no que se refere a resistência a doenças, ganhos em produtividade e adaptabilidade de crescimento a condições de estresse, foram obtidos. Isto porque *S. officinarum*, popularmente conhecida como “cana nobre” ou produtora de açúcar, apresenta qualidades agroindustriais satisfatórias, como colmos espessos, alto teor de sacarose e baixos conteúdos de fibras e amido. *S.*

*spontaneum*, por sua vez, possui considerável resistência a pragas e doenças, boa capacidade de rebrota em soqueira, vigor e adaptabilidade. Este processo de promover o cruzamento de um clone “nobre”, rico em açúcar, com um clone de uma espécie geneticamente próxima que apresentasse alguma vantagem como, por exemplo, resistência a alguma doença, ficou conhecido como “nobilização” (Bremer, 1961).

Era possível perceber nos híbridos provenientes dos eventos de “nobilização” uma maior semelhança à *S. officinarum* do que com *S. spontaneum*, no que se referia ao teor de sacarose, bem como aumento da quantidade de cromossomos presentes nos híbridos em relação aos genitores. Na tentativa de justificar este elevado número de cromossomos, Bhat & Gill (1985) realizaram um extenso relato no qual afirmaram que *S. officinarum*, ao ser utilizada como progenitor feminino, transmitiu seu número somático de cromossomos ( $2n=80$ ) no primeiro cruzamento, enquanto *S. spontaneum*, utilizada como progenitor masculino, transmitiu apenas o seu número haplóide ( $n=56$ ) de cromossomos. Este fenômeno provavelmente persistiu até a segunda geração de nobilização, correspondente ao primeiro retrocruzamento ( $RC_1$ ) com *S. officinarum* e, somente a partir da terceira nobilização ( $RC_2$ ), o genitor *S. officinarum* passou a transmitir “n” cromossomos (Figura 2.1).

Observa-se assim que, nas duas primeiras etapas da nobilização, o mecanismo de transmissão dos cromossomos consiste na doação de gametas que contém “ $2n + n$ ” cromossomos. Bremmer (1961) elucidou que a peculiaridade citológica da transmissão de óvulos gaméticos “ $2n$ ”, transmitidos por *S. officinarum*, ocorre após a primeira divisão meiótica por endoduplicação ou por fusão dos núcleos celulares haplóides após a segunda divisão meiótica. No caso da endoduplicação ocorre recombinação normal entre genes durante o *crossing-over* e uma subsequente duplicação, que resulta em 100% de gametas homozigotos. Já no caso da fusão dos núcleos haplóides, os gametas diplóides são geneticamente equivalentes aos produtos da primeira divisão meiótica, na qual todos os genes entre o centrômero e o local de ocorrência de um primeiro *crossing-over* são homozigotos e todos os genes entre o primeiro e o segundo *crossing-over* são heterozigotos.





**Figura 2.1.** Representação esquemática do mecanismo de transmissão cromossômica de *S. officinarum* e *S. spontaneum* durante o processo de “nobilização” para originar as primeiras variedades modernas de cana-de-açúcar ( $RC_1$  corresponde ao primeiro retrocruzamento realizado entre os híbridos interespecíficos e *S. officinarum*, e  $RC_2$  ao segundo deste retrocruzamento).

Embora o mecanismo de transmissão cromossômica “ $2n + n$ ” dos híbridos de cana-de-açúcar modernos venha há muito tempo sendo elucidado na literatura, para alguns autores o comportamento cromossômico durante a fecundação ainda não foi definitivamente esclarecido e as informações sobre a estrutura e organização do genoma são meramente especulativas (Oliveira, 2007). Roach (1969) já havia contestado o mecanismo de transmissão “ $2n+n$ ” cromossomos ao observar em híbridos interespecíficos e em diferentes citótipos de *S. spontaneum* a ocorrência de transmissão “ $n + n$ ”. Todavia, o entendimento da evolução do genoma tem se tornado mais claro, devido à possibilidade do uso de ferramentas moleculares e citogenéticas.

D’Hont et al. (1996), utilizando DNA genômico hibridizado *in situ* observou que em híbridos, provenientes do cruzamento entre *S. officinarum* e *S. spontaneum*, 39 cromossomos eram provenientes de *S. officinarum* e 32 cromossomos eram provenientes de *S. spontaneum*, indicando a possibilidade de

transmissão de “n+n” cromossomos pelos progenitores. Esse trabalho, além de contestar a explicação histórica que visa justificar o elevado número de cromossomos dos híbridos atuais, também revela que há ocorrência de recombinação cromossômica entre *S. officinarum* e *S. spontaneum*. Estes resultados foram posteriormente confirmados por Grivet et al. (1996). Todavia Hoarau et al. (2001), utilizando marcadores AFLP, também contestaram os resultados de D'Hont et al. (1996) ao afirmarem que a técnica de hibridização *in situ* é eficiente para identificar porções do genoma provenientes de *S. spontaneum* em híbridos, mas possui baixa resolução para identificar regiões cromossômicas recombinantes.

#### 2.1.3.2 Melhoramento genético da cana-de-açúcar no Brasil

No Brasil, até meados do século XVIII, as plantações de cana-de-açúcar eram provenientes de um único clone, ou de um pequeno número de clones, denominados de canas creoulas. As canas ceroulas eram clones de *S. barberi* ou de um híbrido desta espécie com *S. officinarum*. Muito difundidos, estes clones eram bastante rústicos e suscetíveis a doenças, fazendo com que mais tarde fossem substituídos pelo clone “Bourbon”, e este, por sua vez, pelo clone “Cheribon” (Matsuoka et al., 2005).

Até o final do século XIX acreditava-se que a cana-de-açúcar era uma planta totalmente estéril. Somente em 1858, em Barbados, observou-se a ocorrência de plântulas de cana-de-açúcar provenientes de sementes naturais. A partir de 1888, a capacidade das plantas de cana-de-açúcar de gerar sementes passou a ser explorada para a produção de novas variedades, pela realização de cruzamentos intra-específicos entre as canas “nobres” (Machado, 2000). Todavia, a grande suscetibilidade a doenças, observada nas variedades resultantes dos cruzamentos intraespecíficos de *S. officinarum*, foi o que impulsionou a realização de cruzamentos interespecíficos, os quais, ao serem realizados com variedades selvagens, proporcionaram a obtenção de híbridos resistentes a uma grande quantidade de doenças (Lima, 2000).

No Brasil, somente no início do século XX, identificou-se a necessidade de se organizar a experimentação canavieira com a finalidade de substituir as

variedades comerciais por materiais mais estáveis e geneticamente superiores. As “conferências açucareiras” passaram a ser realizadas e foram fundamentais para a implantação das estações experimentais no nordeste e no sudeste do país, e para o surgimento dos atuais programas de melhoramento (Matsuoka et al., 1999).

Em 1968, a Cooperativa Central dos Produtores de Açúcar e Alcool do Estado de São Paulo (Coopersucar) deu início a um importante programa de melhoramento genético de cana-de-açúcar, incorporando a Cooperativa dos Usineiros do Oeste do Estado de São Paulo (Copereste) à sua estação experimental. Na década de 70, também foram criados o Instituto do Açúcar e do Alcool (IAA) e o Plano Nacional de Melhoramento da Cana-de-açúcar (Planalsucar), ambos de abrangência nacional e que muito colaboraram para o desenvolvimento científico e econômico brasileiro (Miocque, 1993). Todavia, em 1990 o Planalsucar foi extinto e novos programas foram estabelecidos, como o ProCana, resultante da reorganização institucional do Instituto Agrônomo de Campinas (IAC), e a Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro (Ridesa) que, contando com a estrutura do Planalsucar, passou a ser conduzido por algumas das universidades federais brasileiras (Machado, 2000).

Atualmente no Brasil existem vigentes quatro programas de melhoramento genético de cana-de-açúcar: Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa) Instituto Agrônomo de Campinas (IAC), Centro de Tecnologia Canavieiro Coopersucar (CTC) e CanaVialis. A Ridesa, por ser composta por onze universidades federais, constitui o maior destes programas e é responsável pelas variedades de cana-de-açúcar RB (República do Brasil) (BNDES, 2008).

Embora cada um destes programas apresentem suas peculiaridades, em essência, todos baseiam-se na seleção e clonagem dos genótipos superiores presentes em populações segregantes provenientes de cruzamentos inicialmente realizados. Sendo assim, como o objetivo é maximizar a probabilidade de se obter genótipos que poderão ser liberados como variedades comerciais, é conveniente que genótipos que apresentam bom desempenho em caracteres economicamente importantes sejam utilizados nos cruzamentos (Landell et al., 2005).

De acordo com Lima (2000), a escolha dos genitores é fundamental para a obtenção da população segregante, uma vez que a variabilidade genética a ser disponibilizada para as etapas de seleção posteriores dependerá das características genéticas dos genitores. Sendo assim, é importante que critérios como divergência genética, complementaridade de caracteres e capacidade de combinação entre os genitores sejam considerados.

O programa de melhoramento da Ridesa, que é enfatizado no presente trabalho, possui em seu banco de germoplasma, aproximadamente, 2.000 acessos, dentre os quais se encontram variedades comerciais utilizadas no país e nas demais regiões canavieiras do mundo; e clones e espécies relacionadas ao gênero *Saccharum*, que são utilizados como genitores em cruzamentos biparentais, policruzamentos e policruzamentos interespecíficos. Para obtenção das sementes que constituem as populações segregantes, os cruzamentos da Ridesa têm sido realizados na Estação de Cruzamentos de Serra do Ouro, em Murici, estado de Alagoas. Esta estação de cruzamentos está localizada a uma longitude de 9°13'S, altitude de 450-500 m, local que apresenta pluviosidade média de 2.000 mm anuais, com temperaturas médias de 19,5°C a 26,5°C. Tais condições são propícias ao florescimento e viabilidade dos grãos de pólen da maioria dos genótipos de cana-de-açúcar (Matsuoka et al., 2005).

Obtidas as sementes estas são colhidas, dessecadas e encaminhadas às universidades vinculadas à Rede para que os experimentos possam então ser estabelecidos. A semeadura normalmente é realizada em casa-de-vegetação e, após algumas semanas, as plântulas obtidas da semeadura são individualizadas em copos plásticos, contendo substrato composto, e são mantidas em casa-de-vegetação até pleno enraizamento (Landell et al., 2005).

Após a realização das etapas iniciais de obtenção das plântulas, segue-se com o transplântio das mudas obtidas em campo. Como nesta fase cada plântula constitui um único genótipo sem repetição, é comum que a avaliação experimental seja conduzida no delineamento de blocos aumentados (Federer, 1956). A primeira touceira, T<sub>1</sub>, emerge normalmente três meses após ter sido transplantada. Passados cinco a sete meses, a cana-planta pode ser cortada para que se minimize a desuniformidade inerente a esta fase, sem que tenha sido realizada seleção (Cesnik & Miocque, 2004). A rebrota das touceiras após o corte

constitui uma das primeiras características a serem avaliadas ao longo das várias etapas de seleção e, os genótipos capazes de persistir constituirão a segunda fase de seleção ( $T_2$ ).

Landell et al. (2005) destacam que a intensidade de seleção empregada nas etapas iniciais varia de acordo com a população e com a forma que a seleção é praticada. Como nesta etapa ainda não é possível avaliar genótipos em experimentos com várias repetições e em vários locais, deve-se utilizar uma maior intensidade de seleção apenas para os caracteres de alta herdabilidade, minimizando assim a possibilidade de se descartar genótipos superiores.

Na fase  $T_2$  cada clone é também avaliado com uma única repetição e, geralmente se avaliam os seguintes caracteres: peso total da parcela, número de colmos por touceira, capacidade de rebrota, diâmetro do colmo, número de entrenós por colmo, comprimento do colmo e teor de sólidos solúveis. Com exceção dos dois primeiros caracteres, a avaliação é realizada considerando-se uma amostra dos colmos adultos por touceira (Matsuoka et al., 2005).

Os clones que apresentam melhor desempenho ao longo destas etapas de seleção passam a ser avaliados em vários locais, com repetições e em sucessivos cortes. Na Ridesa tem-se grande envolvimento do setor produtivo ao longo das etapas de seleção, uma vez que os experimentos são instalados em usinas conveniadas ao programa. Assim que os caracteres avaliados permitirem recomendações com razoável segurança, as novas variedades podem ser usadas comercialmente (Landell et al., 2005; Matsuoka et al., 2005).

A liberação de variedades de cana-de-açúcar, quando comparado ao de outras culturas, é consideravelmente moroso, pois requer entre dez e doze anos. O seu sucesso depende de uma série de decisões a serem tomadas desde a escolha adequada dos genitores, até às fases finais de recomendação de variedades. A adoção de determinados cuidados pode maximizar consideravelmente o sucesso na liberação de uma variedade, em menor intervalo de tempo. Entre este estão: estabelecimento de controle dos efeitos de endogamia; presença de complementaridade genética entre os genitores; avaliação da herdabilidade; correlação entre os caracteres de interesse; ponderação na intensidade de seleção para diferentes caracteres; avaliação da

interação de genótipos x ambientes; e consideração das épocas de avaliação (Matsuoka et al., 2005).

## 2.2 CARACTERIZAÇÃO GENÉTICA DA CANA-DE-AÇÚCAR

### 2.2.1 Comportamento cromossômico em poliploides

A poliploidia é de ocorrência comum em plantas, sendo que, das espécies cultivadas, aproximadamente 40% são poliploides (Schifino-Wittmann, 2004). Uma planta é dita poliploide se dentro de suas células somáticas existirem mais de dois cromossomos pertencentes a um mesmo grupo de homologia. Se o número total de cromossomos é múltiplo do número básico de cromossomos da espécie, a planta é dita euploide, caso contrário é dita aneuploide. Os poliploides podem ser autopoliploides, caso sejam provenientes da duplicação de um mesmo genoma, ou podem ser oriundos da duplicação de genomas diferentes, normalmente após um evento de hibridação, sendo neste caso denominados alopoliploides (Allard, 1971).

Nos alopoliploides, como o genoma é uma combinação de genomas de espécies distintas, reunidos por hibridações interespecíficas, há tendência de ocorrer pareamento cromossômico preferencial entre os homólogos durante a meiose, caracterizando assim uma herança dissômica. De modo contrário, como nas espécies autopoliploides, como o genoma é proveniente da duplicação de genomas homólogos, durante a meiose os cromossomos tendem a se parear de forma aleatória dentro de um mesmo grupo de homologia, caracterizando uma herança polissômica. Nestes casos, pode ainda ocorrer pareamento simultâneo entre mais de dois cromossomos de um mesmo grupo de homologia, formando multivalentes (Qu & Hancock, 2001).

Nos casos de herança polissômica, como parte dos genes presentes em um genoma está repetida nos demais genomas, os padrões de segregação mendeliana se tornam ainda mais complexos, dificultando as análises e interpretações genéticas. Assim, considerando os diferentes comportamentos cromossômicos durante a meiose de organismos poliploides, Albino et al. (2006) afirmam que os organismos alopoliploides apresentam tendência de herança

dissômica, enquanto os organismos autopoliploides tendem possuir herança polissômica.

A constituição genômica de muitos poliploides com alto nível de ploidia ainda não é completamente entendida. Mas, uma forma de se obter este conhecimento é através do estudo do comportamento dos cromossomos durante a meiose. Na cana-de-açúcar a poliploidia atinge todas as espécies que compõem o gênero *Saccharum*, tornando o genoma desta planta um dos maiores e mais complexos dentre todas as espécies cultivadas. De acordo com D'Hont et al. (1998), o número básico de cromossomos no gênero *Saccharum* já foi estudado por diversos autores, que chegaram a conclusões divergentes. Os mais prováveis são de  $x=10$  para *S. officinarum* e  $x=8$  para *S. spontaneum*. Quanto ao tamanho do genoma, Grivet & Arruda (2001) estimam que o tamanho do genoma básico de *Saccharum* é similar ao genoma gamético de sorgo, que contém cerca de 760 Mpb, cerca de duas vezes maior que o genoma de arroz (430 Mpb), e significativamente menor que o genoma de milho (2500 Mpb).

Da Silva et al. (1993) e Al-Janabi et al. (1993), buscando avaliar o comportamento dos cromossomos de *S. officinarum* e de *S. spontaneum*, durante a meiose, verificaram que *S. spontaneum* ( $2n=64$ ) comporta-se como um autopoliplóide com herança polissômica, apesar de ter sido observado também a ocorrência de alguns pareamentos preferenciais. De acordo com Bremer (1961), embora os clones das diferentes espécies do gênero *Saccharum* tenham meiose geralmente normal, com formação de bivalentes, irregularidades meióticas como a formação de univalentes ou até algumas multivalentes podem ser comumente observadas. Estas são mais frequentes em *S. spontaneum* que em *S. officinarum*. Diante disso, pode-se afirmar que o modo preciso de pareamento dos cromossomos nas espécies do gênero *Saccharum* durante a meiose ainda não é conhecido. Sabe-se, no entanto que cromossomos recombinantes já foram observados em variedades comerciais.

Aitken et al. (2005) e Piperidis et al. (2008; 2010) constataram que, nos híbridos cultivados atualmente, 10% a 28% do genoma é proveniente de *S. spontaneum* e 67% a 87% é proveniente de *S. officinarum*. O restante, aproximadamente 13%, é atribuído a recombinações ocorridas entre os cromossomos de *S. officinarum* e de *S. spontaneum*. Segundo estes autores, o

genoma de *S. officinarum* é o responsável pela diversidade encontrada nas variedades modernas.

D'Hont et al. (1996), usando DNA hibridizado *in situ*, distinguiram em clones da variedade R570 os cromossomos doados por *S. officinarum* e por *S. spontaneum*. Além disso, perceberam que 10% dos cromossomos foram doados por *S. spontaneum* e 10% eram produtos de recombinação entre *S. officinarum* e *S. spontaneum*, confirmando que pode haver recombinação entre os cromossomos das duas espécies.

Al-Janabi et al. (1993) e Ming et al. (1998) também investigaram o comportamento do pareamento de cromossomos em cana-de-açúcar, através de análises de repulsão entre marcadores mapeados. Verificaram em clones de *S. robustum* que na metade dos grupos de ligação previamente formados houve elevada proporção de pareamento preferencial. No clone de *S. officinarum*, alguns pareamentos preferenciais puderam ser observados e em *S. spontaneum* nenhum pareamento preferencial foi verificado.

Grivet et al. (1996) observaram em progênie obtida de autofecundação da variedade R570 ocorrência de pareamentos aleatórios, com poucos casos de pareamento preferencial. Este resultado também foi verificado por Hoarau et al. (2001), que, usando um maior número de marcadores e de genótipos oriundos de autofecundação da mesma variedade, também observaram poucos pareamentos preferenciais, sugerindo que, as variedades modernas apresentem tendência de herança polissômica.

Diante do exposto, observa-se que a natureza poliploide associada aos recorrentes eventos de aneuploidia em cana-de-açúcar dificulta a compreensão do comportamento cromossômico durante a meiose, fazendo com que, entre as espécies cultivadas, esta seja a que apresenta a genética mais complexa. Embora as informações obtidas até o presente momento não permitam conclusões claras e generalizáveis para todas as espécies do gênero *Saccharum*, é importante considerar que estes tipos de estudo, ainda assim, representa a chave para o entendimento de como o genoma da cana-de-açúcar está estruturado.



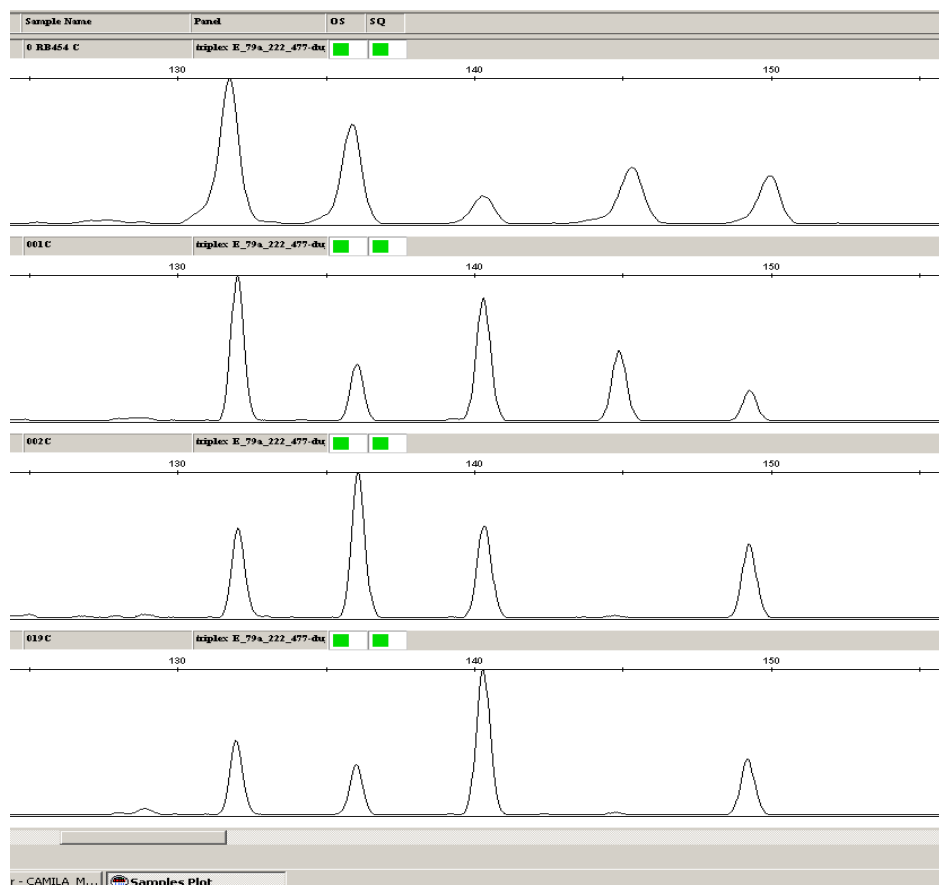
### 2.2.2 Detecção de polimorfismo

A característica poliploide das plantas de cana-de-açúcar apresenta implicações não apenas no entendimento do comportamento cromossômico durante a meiose, mas também na detecção de polimorfismo em estudos nos quais são utilizados marcadores moleculares. Sabe-se que a disponibilidade de grande quantidade desses marcadores permitiu a realização de diversos tipos de estudos genéticos, entre eles, a construção de mapas genéticos (Ferreira & Grattapaglia, 1998).

Quando se trata de organismos poliploides, a complexidade envolvida no uso de marcadores aumenta. Isso porque nestes casos é obtida grande quantidade de fragmentos, os quais representam os diferentes alelos de um mesmo loco, em diferentes dosagens, nos vários cromossomos homólogos envolvidos (Aitken et al., 2007). Esse tipo de perfil pode ser visualizado em análises de fragmentos de DNA de genótipos de cana-de-açúcar, em que foram avaliados locos SSR (*Simple Sequence Repeats*), como mostra o cromatograma da Figura 2.2, obtido via *software* GeneMapper® 3.0 (Life Technologies).

Wu et al. (1992) propuseram uma metodologia baseada na dosagem dos fragmentos presentes no genoma para identificar o polimorfismo em poliploides. O estabelecimento desta metodologia facilitou consideravelmente o desenvolvimento de mapas genéticos para tais tipos de espécies, especialmente cana-de-açúcar, possibilitando com isso uma melhor compreensão de genomas assim tão complexos (Alwala & Kimbeng, 2010).

É necessário considerar que propriedades como praticidade, transferibilidade, boa capacidade de multiplexagem e de discriminar genótipos, custo, tempo e praticidade inerente aos procedimentos laboratoriais são fundamentais para a obtenção de resultados satisfatórios em análises moleculares de plantas geneticamente complexas como a cana-de-açúcar. De acordo com os critérios definidos por Wu et al. (1992), marcadores como SSR, ao serem utilizados em análises genômicas de cana-de-açúcar, devem ser interpretados como marcadores dominantes, fazendo com que percam uma de suas maiores vantagens reportadas na literatura que é a sua natureza codominante.



**Figura 2.2.** Cromatograma obtido a partir de fragmentos de DNA de genótipos de cana-de-açúcar derivados do cruzamento entre RB97327 e RB72454, amplificados via *primers* microssatélites.

Segundo a metodologia de Wu et al. (1992), marcadores que estiverem presentes uma vez no genoma de um dos genitores e ausente no outro são denominados de marcadores em dose única ou *single-dose* (SD), e segregam em uma geração  $F_1$  na proporção de 1:1 (Aaaaaaaa x aaaaaaaa, ou vice-versa, admitindo-se  $2n=8x$ ). Alternativas à segregação 1:1 também podem ser observadas. Por exemplo, ao invés do alelo estar presente uma vez no genoma de um dos genitores, como descrito anteriormente, o mesmo alelo pode estar presente uma vez no genoma de cada um dos dois genitores (Aaaaaaaa x Aaaaaaaa). Marcadores que se comportam dessa forma são também denominados de SD, embora segreguem, na geração  $F_1$  na proporção de 3:1 (Wu et al., 1992).

Aplicações práticas de marcadores SD que segregam na proporção 3:1 são comumente observadas na literatura. Hoarau et al. (2001) desenvolveram um mapa de ligação a partir de população obtida por autofecundação da cultivar de

cana R570, utilizando marcadores com segregação 3:1. Estes marcadores também têm sido utilizados para identificar grupos de homeologia (Edmé et al., 2006), assim como para fazer inferências acerca de regiões do genoma conservadas entre as diferentes espécies do gênero *Saccharum* (Alwala et al., 2008).

Aitken et al. (2005) ressaltam ainda que, como na cana-de-açúcar a poliploidia atinge todas as espécies do gênero *Saccharum*, outras alternativas às segregações 1:1 e 3:1 podem ser aproveitadas. Exemplo disso ocorre nos casos em que um alelo se encontra presente duas vezes no genoma de um genitor e ausente no outro, e segrega na progênie na proporção 11:3, se o genoma monoplóide do genitor for  $x=8$ ; ou segrega na progênie na proporção 7:2, se o genoma monoplóide do genitor for  $x=10$ . Marcadores que apresentam esse comportamento são denominados de dose supla ou *double-dose* (DD).

Marcadores DD são de fundamental importância para mapeamento, pois, ao serem adicionados a um mapa de ligação previamente construído com marcadores SD, aumentam consideravelmente a saturação do mapa e facilitam a identificação dos grupos de ligação pertencentes a um mesmo grupo de homeologia (Da Silva et al. , 1995; Aitken et al., 2007).

Palhares (2010) considerou que, no caso da cana-de-açúcar, se as análises genéticas se basearem nas premissas de autopoliploidia e número básico de cromossomos  $x=8$ , os locos podem apresentar dosagens alternativas às SD e DD, como por exemplo, quando não houver a presença de alelo dominante em nenhum dos genitores (nuliplex - aaaaaaaaa), ou quando houver mais de dois alelos informativos em pelo menos um dos genitores (multiplex – AAAaaaaa, AAAAaaaa, AAAAAaaa, AAAAAAaa e AAAAAAAa). Estes marcadores são denominados *multi-dose* (MD) e permitem a obtenção de até mais de setenta combinações alélicas, sendo possível estimar as frequências gaméticas em cruzamentos que envolvem genitores com até quatro doses de um alelo. De acordo com Aitken et al. (2007), os marcadores MD podem acrescentar informações importantes a um mapa de ligação, por facilitarem a identificação dos grupos de homeologia.

### 2.2.3 Marcadores DArT (*Diversity Arrays Technology*)

A tecnologia dos marcadores DArT foi inicialmente desenvolvida para a cultura do arroz por Jaccoud et al. (2001), e consiste em um método de identificação de polimorfismo de DNA ao longo de todo o genoma de um organismo, via detecção da presença/ausência de fragmentos individuais em representações genômicas imobilizadas em microarranjos. Esta metodologia permite a obtenção do perfil de todo o genoma sem a necessidade de informações prévias sobre uma sequência específica de oligonucleotídeos ou sobre a sequência do DNA.

A obtenção dos marcadores DArT pode resumidamente ser dividida em duas etapas: desenvolvimento dos painéis de diversidade e genotipagem. Na primeira etapa o DNA das amostras a serem estudadas é agrupado e clivado por enzimas de restrição, sendo uma de corte raro e outra de corte frequente, e, posteriormente, aos fragmentos digeridos são ligados adaptadores às extremidades de corte raro. Em seguida, uma reação em cadeia da polimerase (PCR) é realizada para amplificação dos fragmentos gerados utilizando-se *primers* complementares à sequência dos adaptadores. A complexidade do genoma é reduzida, pois somente fragmentos com as duas extremidades ligadas aos adaptadores são amplificados (Jaccoud et al., 2001).

Heller-Uszynska et al. (2011) afirmam que a etapa de redução da complexidade do genoma determina a eficiência da identificação dos polimorfismos na técnica de DArT, especialmente em espécies que apresentam genomas extensos e complexos como a cana-de-açúcar. Estes autores testaram doze combinações de enzimas de restrição, utilizando dez metodologias diferentes aplicadas a dezesseis genótipos representantes da diversidade genética do gênero *Saccharum*. Encontraram combinações como *PstI/TaqI* associadas a *BstNI* ou *HaeIII*, que podem ser usadas com sucesso em rotinas laboratoriais para gerar representações genômicas em cana-de-açúcar.

Ainda nessa primeira etapa, os fragmentos amplificados via PCR são clonados em um vetor. Os insertos são novamente amplificados, purificados e arranjados em um suporte sólido (*microarray*), gerando assim um painel de diversidade através da impressão dos *microarrays* contendo as representações

genômicas em lâminas de vidro, também denominadas *slides* (Jaccoud et al., 2001).

A segunda etapa de obtenção dos marcadores DArT consiste na genotipagem dos indivíduos no painel de diversidade obtido na primeira etapa. Para isso, o DNA genômico de cada amostra a ser genotipada é marcado com fluorescência tendo como sinal de referência o fragmento do vetor de clonagem. Em seguida, promove-se a hibridização do DNA das amostras, com os fragmentos de referência presentes no *microarray*. Os clones polimórficos apresentarão intensidade de sinais de hibridização variáveis para diferentes amostras conforme os genótipos. Os *slides* são escaneados e as imagens obtidas são analisadas computacionalmente utilizando-se o *software* DArTSoft (DArT Pty/Ltd.). Este aplicativo é capaz de definir todos os *clusters* centrais, identificar os pontos inseridos dentro destes e calcular a probabilidade de cada clone pertencer a um *cluster*, gerando desta forma uma tabela de dados binários baseada nos valores de adesão de cada clone a um determinado *cluster* (Jaccoud et al., 2001).

Atualmente essa tecnologia tem sido realizada pela empresa DArT Pty/Ltd., e utilizada com sucesso para a construção de mapas genéticos, estudos de caracterização de coleções de germoplasma e de diversidade genética de várias espécies de importância agrônômica, como cevada (Wenzl et al., 2004), trigo (Akbari et al., 2006), arroz (Xie et al., 2006), sorgo (Mace et al., 2008), aveia (Cover, 2010), cana-de-açúcar (Heller-Uszynska et al., 2011) e eucalipto (Sansaloni, 2011). Isso porque a metodologia utilizada pelos marcadores DArT para identificação de polimorfismo de DNA oferece um sistema rápido, robusto e que requer quantidades mínimas de DNA e que elimina a subjetividade dos géis de eletroforese por permitir a automatização dos dados.

É importante salientar ainda que um fator limitante da utilização de marcadores moleculares em programas de melhoramento tem sido a velocidade e os custos envolvidos nas análises. Isto, sobretudo para a obtenção de resultados expressivos em populações grandes ou em espécies que apresentam genomas complexos. Os marcadores SNP e SSR, por exemplo, representam avanços consideráveis no que se refere à quantidade de informações geradas. Todavia, a metodologia envolvida na obtenção de suas informações ainda depende do

conhecimento de sequências do genoma, limitando a sua utilização em espécies que não possuem informações prévias acerca do genoma ou em espécies poliploides como a cana-de-açúcar, cujo sequenciamento do genoma é tecnicamente oneroso e demorado (Heller-Hszynska et al., 2011)

De acordo com Heller-Hszynska et al. (2011), em espécies como a cana-de-açúcar por exemplo, como as variedades modernas apresentam aproximadamente 110 cromossomos, seriam necessários entre 5.000 e 7.000 marcadores para obtenção de um mapa genético com cobertura adequada às suas diferentes aplicações. Neste caso, os marcadores DArT, por combinarem baixo custo ao elevado rendimento e precisão das informações obtidas, revelam-se promissores para auxiliar na melhor compreensão desses genomas.

Acredita-se que a obtenção de bibliotecas geradas a partir de melhores métodos de redução da complexidade do genoma permitirá construir um “chip de DNA de cana”, possibilitando ampla “varredura” no genoma. Dessa forma, milhares de marcadores em centenas de genótipos poderão ser gerados. Isso seria um passo inicial para a realização da estratégia da seleção genômica ampla (*Whole Genome Selection*) em cana-de-açúcar, a qual permitiria o lançamento de variedades superiores em tempo muito mais curto do que o exigido pelos métodos de melhoramento tradicionais (Heller-Hszynska et al., 2011).

## 2.3 MAPEAMENTO GENÉTICO

### 2.3.1 Importância e princípios do mapeamento genético

O mapeamento genético consiste na distribuição relativa dos marcadores moleculares ao longo do genoma, isto é, nos seus respectivos grupos de ligação, bem como na determinação da distância genética entre eles. Nos últimos anos a disponibilidade de um grande número de marcadores moleculares, altamente polimórficos e neutros, aliada a uma série de procedimentos estatístico-computacionais, permitiu a construção de mapas genéticos para a maioria das espécies vegetais de importância agrônômica, constituindo esta uma das aplicações de médio-longo prazo de maior impacto da tecnologia dos marcadores moleculares no melhoramento de plantas (Oliveira, 2006).

De acordo com Coelho (2000), os mapas genéticos são capazes de gerar uma grande quantidade de informações básicas sobre a estrutura e organização do genoma de uma espécie, além da obtenção de informações capazes de auxiliar os programas de melhoramento genético de diversas formas. Entre estas informações Liu (1998) cita, por exemplo, que os mapas genéticos podem permitir a localização e a avaliação da magnitude do efeito das regiões genômicas associadas aos caracteres de importância agrônômica (QTL – *Quantitative Trait Loci*); realizar estudos evolutivos e de mapeamento comparativo; clonar genes a partir de suas posições previamente identificadas nos mapas; e realizar seleção assistida por marcadores (SAM).

A possibilidade de se analisar a estrutura geral do genoma das diversas espécies cultivadas e, ainda, estabelecer comparações entre genomas de espécies geneticamente relacionadas revelam uma importante aplicação dos mapas genéticos. O mapeamento comparativo constitui uma boa estratégia para aumentar a cobertura dos mapas de espécies cujo genoma apresenta poucas informações e/ou cujos estudos apresentam dificuldades para serem conduzidos. Esta metodologia contribui para o entendimento da evolução dos genomas e permite que informações obtidas a partir de genomas simples sejam extrapoladas para espécies relacionadas e com genomas complexos (Cordeiro et al., 2003).

De acordo com Liu (1998), a possibilidade de se clonar genes de interesse através da identificação de sua posição nos mapas de ligação constitui a mais audaciosa aplicação dos mapas genéticos. Nesse procedimento, tendo sido obtidos o efeito fenotípico do QTL sobre o caráter de interesse e a sua posição cromossômica, torna-se possível identificá-lo através de marcadores presentes próximos a ele. A identificação individual de um QTL permite a sua clonagem, ou seja, a transferência de sua sequência para algum vetor. De posse de um clone, torna-se possível estudar o QTL em diferentes tecidos e condições, isolar sequências que sejam homólogas às de outras espécies e, ainda, inserir esta sequência em outras espécies (transgenia).

A obtenção de mapas genéticos também é capaz de auxiliar a seleção baseada em informações oriundas de marcadores moleculares que representa um grande avanço biotecnológico utilizado em diversas etapas do melhoramento. Bearzoti (2000) destaca que a SAM, pode, inclusive ser realizada sem o

mapeamento de QTL e pode ser útil no direcionamento de cruzamentos específicos, pois permite monitorar a presença de alelos favoráveis, maximizando, tanto o número de QTL com alelos favoráveis, como o valor genotípico global.

Embora ainda se observe certa resistência na utilização de SAM em diversos programas tradicionais de melhoramento, é válido considerar que o seu emprego pode ser estratégico, principalmente quando a determinação do fenótipo é complexa e/ou possui custos elevados; quando a análise requer a destruição da planta; ou quando se objetiva realizar a introdução de várias características, simultaneamente, em determinado genótipo. Na tomada de decisão acerca do uso ou não de SAM é importante considerar alguns fatores, como a natureza genética dos caracteres que se almeja selecionar, o modo de ação gênica, o efeito dos genes na expressão do fenótipo, a complexidade da avaliação fenotípica e a eficiência com que o marcador discrimina as características de interesse (Albino et al., 2006). Hotta et al. (2010) ratificam também a necessidade de que as informações geradas por estas ferramentas biotecnológicas sejam disponibilizadas aos programas de seleção, de forma mais simples, rápida e pouco onerosa para utilização em larga escala, de modo a facilitar a sua inserção na rotina do melhorista.

De acordo com Coelho (2000), para a construção de mapas genéticos, independentemente da espécie para a qual se almeja construir um mapa, é necessário considerar: os genótipos dos genitores a serem utilizados para obtenção da população de mapeamento, as características e a quantidade de marcadores que devem ser mapeados e o número de indivíduos a serem avaliados. Se o objetivo for utilizar o mapa de ligação para posteriormente identificar QTL, há de se considerar, ainda, a necessidade de obtenção dos dados fenotípicos, bem como a associação destes aos dados moleculares de cada indivíduo. Schuster & Cruz (2008) afirmam que esta etapa inicial de estabelecimento dos critérios a serem utilizados antes de se proceder às análises de mapeamento é crítica, pois, depende dos objetivos a serem alcançados com a elaboração do mapa e de fatores como a disponibilidade de tempo e recursos.

Independentemente do tipo de população, para a elaboração de um bom mapa genético, isto é, que agregue boa quantidade de marcas informativas, é fundamental que a população de mapeamento apresente máximo desequilíbrio



de ligação (Schuster & Cruz, 2008). O desequilíbrio de ligação de interesse, neste caso, é atribuído à ligação física entre locos e é proveniente do desequilíbrio gamético, o qual decorre da baixa fração de recombinação entre locos situados próximos em um cromossomo. Este fenômeno também pode ser entendido como a associação não casual entre alelos de diferentes locos, isto é, quando o alelo de um loco mantém-se associado ao alelo de outro loco numa frequência diferente daquela esperada se eles se combinassem independentemente (Nordborg & Tavaré, 2002).

Liu (1998) define os mapas genéticos como modelos abstratos nos quais um grupo de marcadores é organizado com base nas frequências de recombinação que ocorrem entre cromossomos homólogos durante a meiose. Estas recombinações são resultado das permutações, isto é, da ocorrência de trocas de segmentos entre cromátides não-irmãs pertencentes a cromossomos homólogos (*crossing-over*).

Os eventos de recombinação durante a meiose ocorrem ao acaso entre diferentes locos e estão associados à distância física entre os locos. Assim, quanto mais próximos estiverem dois marcadores em um cromossomo, menor será a possibilidade de ocorrer um evento de recombinação entre eles. Desta forma, a construção de mapas de ligação baseia-se em estimativas de frequências de recombinação para que, obtendo-se a distância genética entre os marcadores, possa se estabelecer sua ordem relativa ao longo dos grupos de ligação (Coelho, 2000).

Os marcadores podem ser arranjados em grupos de ligação, os quais devem corresponder ao número de cromossomos característico da espécie, embora isto exija a utilização de uma grande quantidade de marcadores de forma a se obter uma ampla cobertura do genoma. O número de marcadores necessários para construção de um mapa genético depende, no entanto, das características inerentes ao marcador, do número de cromossomos característico da espécie e do objetivo para o qual o mapa será construído (Heller-Uszynska, 2011).

Xu (2010) considera que bons marcadores devem apresentar características como: neutralidade, praticidade, reprodutibilidade, elevado nível de polimorfismo genético, padrão de herança codominante, boa capacidade em

distinguir genótipos e ampla distribuição ao longo do genoma. A influência do número de marcadores a serem mapeados está relacionada à sua resolução, uma vez que a cobertura do genoma e a densidade do mapa são funções diretas do número de marcadores utilizados.

Outro fator preponderante refere-se ao número de indivíduos a ser utilizado, uma vez que é este valor que estabelece o nível de resolução máximo quando um número muito grande de marcadores moleculares é utilizado (Coelho, 2000). Pastina et al. (2010) alertam ainda que, além do tamanho da população utilizada para obtenção do mapa, a complexidade e o tamanho do genoma da espécie também influenciam na qualidade do mapa obtido. Em cana-de-açúcar, por exemplo, como o genoma é grande e apresenta elevado nível de ploidia, nenhum dos mapas construídos até o momento mostrou-se completo, contrariamente a outras espécies como milho, soja, sorgo para as quais mapas consideravelmente saturados já foram obtidos.

Observados os aspectos anteriormente descritos, que precedem a construção dos mapas de ligação, deve-se dar início à sua execução através da realização de uma série de análises, a saber: análise da segregação Mendeliana, análise de ligação, análise de agrupamento e análise de ordenação. Realizados todos esses passos pode-se proceder à representação gráfica do mapa, sob a forma dos seus grupos de ligação.

#### 2.3.1.1 Análise da segregação Mendeliana

Na primeira etapa de execução do mapeamento genético é necessário selecionar os marcadores, presentes nos genitores e na população de mapeamento, que serão de fato utilizados na construção do mapa; isto é aqueles que apresentam polimorfismo e que segregam de acordo com as proporções Mendelianas esperadas. A ausência de um padrão mendeliano é chamada de distorção da segregação mendeliana e pode ser causada, por exemplo, por efeito de seleção ou viabilidade diferencial de gametas (Bearzoti, 2000).

O critério mais imediato para se realizar a seleção dos marcadores a serem incluídas na análise pode ser feito comparando-se as quantidades de marcas observadas na população segregante com as frequências esperadas,

através de testes estatísticos simples como o qui-quadrado ( $X^2$ ). Considerando que para cada marcador deve-se realizar um teste  $X^2$ , uma das questões a serem consideradas ao se realizam testes conjuntos de segregação é o estabelecimento do nível de significância ( $\alpha$ ).

Isso porque se em um único teste realizado para um dado marcador, a probabilidade de não se cometer o erro tipo I (rejeitar  $H_0$  quando esta é verdadeira) é dada por  $P=1-\alpha$ , em múltiplos testes, referentes aos múltiplos marcadores a serem testados, esta probabilidade passa a ser  $P=(1-\alpha)^n$ ; em que “n” corresponde ao número de testes. Logo, a probabilidade de se cometer pelo menos um erro tipo I, quando forem realizados “n” testes independentes pode ser dada por  $\alpha^*=1-(1-\alpha)^n$ . Significa que a probabilidade de se cometer pelo menos um erro tipo I é crescente à medida que o número de testes aumenta, fazendo com que, nestes casos, uma grande quantidade de falsos positivos sejam obtidos na análise (Schuster & Cruz, 2008).

Por esse motivo, quando múltiplos testes  $X^2$  são realizados em um experimento, a adoção de um critério de proteção do erro tipo I faz-se necessária, e, nestes casos pode-se usar, por exemplo, a correção de Bonferroni ou o critério FDR (*False Discovery Rate*) (Liu, 1998). A utilização de critérios de proteção do erro tipo I permite uma análise mais criteriosa dos marcadores que apresentam distorções em relação às segregações esperadas para a população de mapeamento. Para os marcadores mantidos na análise, o próximo passo é testar a eventual ligação entre eles através da análise de ligação.

### 2.3.1.2 Análise de ligação

Dispondo-se de vários marcadores em uma população, é necessário se estimar as fases de ligação e as frações de recombinação entre eles. Fase de ligação é o termo usado para denotar a forma de ocorrência das associações de alelos de locos ligados. Quando dois alelos (dominantes ou recessivos) estiverem localizados em um mesmo cromossomo diz-se que estão ligados em fase de acoplamento ou *cis*; e, se em um mesmo cromossomo estiverem presentes um alelo dominante de um gene e um alelo recessivo de outro gene diz-se que eles estão ligados em fase de repulsão ou *trans* (Schuster & Cruz, 2008).

Em cruzamentos controlados derivados de linhagens puras as fases de ligação são conhecidas pela análise dos progenitores. Já naqueles derivados de genitores não-endogâmicos as fases de ligação são inferidas estatisticamente (Liu, 1998). Aitken et al. (2007) afirmam que a obtenção de informações sobre as fases de ligação entre os alelos são importantes pois, além de auxiliar a identificação dos grupos de ligação, permite a interpretação do modo de pareamento cromossômico durante a meiose. Qu & Hancock (2001) revelam que informações sobre as fases de ligação também podem ser usadas para distinguir aloploiploides de autopoliploides, permitindo com isso uma melhor compreensão do comportamento cromossômico durante a meiose.

A recombinação genética é mensurada através da fração de recombinação, que, de acordo com Ott (1991), é a proporção de gametas recombinantes produzidos por um dado genitor. Diversos procedimentos estatísticos podem ser utilizados para a sua estimação, como o método da máxima verossimilhança, o método dos momentos, o método dos quadrados mínimos, dentre outros (Bearzoti, 2000). Ritter et al. (1990) estabeleceram uma metodologia para se estimar simultaneamente a fração de recombinação e as fases de ligação entre locos em populações provenientes do cruzamento entre indivíduos não-endogâmicos.

Independentemente do método utilizado para tal estimação é importante salientar que os valores obtidos não são utilizados diretamente como medidas de distância genética entre os marcadores pois não se tratam de medidas aditivas, como ilustrado por Lynch & Walsh (1998). A aditividade se faz necessária, pois, caso contrário, ao se atribuir novos marcadores ao mapa seria necessário ajustar as distâncias já estabelecidas (Carneiro & Vieira, 2002).

Para satisfazer a condição de aditividade, funções de mapeamento podem ser utilizadas com o objetivo de converter as frações de recombinação, dadas em porcentagem, em centiMorgans (cM) (Liu, 1998). De acordo com Weir (1996), uma unidade de medida de distância genética em um mapa é igual à distância entre dois locos ao longo da qual se espera que ocorra um *crossing-over* por gameta, por geração. Esta medida é expressa em cM.

Entre as funções de mapeamento mais utilizadas para vegetais, destacam-se as de Haldane (1919) e Kosambi (1944). Ambas convertem as

frequências de recombinantes entre dois locos quaisquer em distâncias genéticas, dadas em cM. O diferencial entre elas é que a primeira admite a ocorrência de permutas próximas como eventos independentes (sem interferência) e segue uma distribuição de Poisson. Já a segunda admite a ocorrência de permutas próximas como eventos não independentes, isto é, com interferência, de forma que esta interferência decresce para locos mais distantes e se iguala a zero para locos independentes (Bearzoti, 2000).

Ferreira & Grattapaglia (1998) ressaltam que é importante considerar que a correlação entre as distâncias genéticas calculadas a partir das frações de recombinação e as distâncias físicas expressas em número de pares de bases é baixa. Isto porque existem pequenas regiões cromossômicas, com algumas dezenas de milhares de pares de bases, em que a probabilidade de ocorrer um evento de recombinação é alta (*hotspots* de recombinação), enquanto outras regiões, com alguns milhões de pares de bases, a ocorrência de recombinação é praticamente nula.

Após a obtenção das distâncias entre os marcadores, a construção de um mapa de ligação deve prosseguir com a análise de ligação. Nesta, as marcas são testadas duas a duas considerando a hipótese de nulidade ( $H_0$ ), de que a frequência de recombinação entre elas é igual a 0,5, isto é, admitindo-se que as marcas segregam independentemente; ou considerando a hipótese alternativa ( $H_1$ ) de que as marcas não segregam conjuntamente por estarem ligadas. Este teste deve ser realizado para todos os possíveis pares de marcadores.

#### 2.3.1.3 Obtenção dos grupos de ligação

Partindo-se das distâncias genéticas entre os marcadores, é necessário construir os grupos de ligação (GL) através da análise de agrupamento. Ott (1991) afirma que um grupo de ligação pode ser definido como um conjunto de genes em que cada um deve estar ligado a, pelo menos, outro gene do mesmo conjunto. Para Liu (1998) um conjunto de locos herdados juntos constitui um grupo de ligação.

O número de grupos de homologia deve ser proporcional ao número básico de cromossomos da espécie. Em cana-de-açúcar, como “*a priori*” não se

tem conhecimento sobre o modo preciso do pareamento entre os vários cromossomos homólogos durante a meiose, afirma-se a existência de grupos de homeologia (HG), ao invés de grupos de homologia, e espera-se que sejam identificados entre oito e dez destes grupos. Todavia, Garcia et al. (2006) comentam sobre a dificuldade de obtenção desta quantidade de HG nos mapas de ligação construídos para cana-de-açúcar. Isto decorre da possibilidade de ocorrência de pareamentos preferenciais entre determinados cromossomos durante a meiose e da heterogeneidade do genoma das variedades atuais, o que inclui duplicações cromossômicas e rearranjos estruturais resultantes de translocações ou associações a *transposons*.

Para demonstrar que dois locos estão ligados, estatísticas implementadas nos *softwares* desenvolvidos para fins de mapeamento genético, como Mapmaker/Exp (Lander et al., 1987), OneMap (Margarido et al., 2007), JoinMap (Van Ooijen & Voorrips, 2001), dentre outros, têm sido baseadas na razão de verossimilhança conhecida como LODscore (*Likelihood Ratio*) ou, simplesmente, LOD. De acordo com Ott (1991), o valor do LOD pode ser obtido pela razão entre as verossimilhanças da hipótese  $H_1$ , de que os locos são considerados ligados, e a hipótese  $H_0$ , de que eles não estão ligados. Assim, se para o cálculo do valor de LOD for utilizado um logaritmo na base 10, um valor LOD=3 indicará que  $H_1$  é 1000 vezes mais verossímil que  $H_0$ , ou seja, que a hipótese de ligação é mil vezes mais provável que a de segregação independente. De acordo com Carneiro & Vieira (2002), o estabelecimento dos grupos de ligação deve ser feito com o LOD variando entre 4 e 6 pois, valores muito baixos podem ocasionar ligações espúrias, enquanto valores muito altos resultam na fragmentação dos grupos de ligação.

#### 2.3.1.4 Análise de ordenação

Tendo sido os grupos de ligação formados, deve-se então proceder à análise de ordenação a fim de se determinar qual a melhor ordem das marcas dentro de cada grupo. Vários critérios podem ser utilizados para esta finalidade, como:

- a) SAR - Soma das frações de recombinação adjacentes (*Sum of Adjacent Recombination fractions*): determina que a melhor ordem das marcas num GL é aquela que possui menor valor de SAR, segundo a fórmula de *Trow*. Todavia, precauções devem ser tomadas ao se utilizar este critério, pois como somente as estimativas das frações de recombinação são conhecidas, a melhor ordem estabelecida pode não ser a verdadeira em função de erro amostral. Além disso, se a análise envolver um grande número de locos, a utilização deste critério pode se tornar impraticável, haja vista que um número muito grande de ordens possíveis precisaria ser avaliado (Liu, 1998).
- b) DRC - Delineação Rápida em Cadeia: este método consiste na obtenção de uma ordem preliminar para os locos, tendo por base a matriz das frações de recombinações de todos os pares de marcas. Após ter sido obtida esta ordem preliminar, é necessário testar sucessivas inversões considerando as marcas três-a-três, a fim de se obter o menor valor de SAR, LOD ou verossimilhança (Carneiro & Vieira, 2002).
- c) Sieriação: por este método, proposto por Buetow & Chakravarti (1987), em uma tabela que relacione todas as marcas com todas as distâncias entre elas, a melhor distância entre as marcas é a de menor valor, isto é, a que apresenta menor valor de SAR. Este método, embora seja relativamente simples, não assegura a melhor ordenação, sobretudo nos casos em que as matrizes de frações de recombinantes não são monotônicas (Weir, 1996).
- d) Verossimilhança: este critério gera a ordem de máxima verossimilhança, isto é, a ordem de maior probabilidade dos dados terem sido gerados. Para proceder a análise utilizando este critério é necessário, inicialmente, avaliar a ordem e as frações de recombinação que maximizam a função de verossimilhança. Como muitas ordens são possíveis, há um crescimento exponencial do número de ordens prováveis com o aumento do número de marcadores, não sendo este critério, portanto, recomendado se este número for muito elevado (Bearzoti, 2000).
- e) Ramos e Conexões: este método, proposto por Thompson (1987), também conhecido com *Branch and Bound*, é baseado na estrutura de uma “árvore” na qual os ramos são os locos e as conexões entre eles são os pontos alternativos para a inserção de novos locos. Este método garante encontrar a melhor ordem

dos locos, mesmo que não se tenha avaliado todas as ordens possíveis (Weir, 1996).

f) *Simulated Annealing*: por este método, proposto por Kirkpatrick et al. (1983) a ordem dos locos apresenta determinado valor de SAR e, à medida que esta ordem sofre permutações, uma nova ordem é obtida com novo valor de SAR. Se a segunda ordem for menor que a primeira então ela é aceita. Após várias permutações, a melhor ordem é a dada quando o menor valor de SAR é alcançado com uma dada probabilidade. Este método é um dos que analisa o maior número de ordens, embora não exista garantia de que a melhor ordem será encontrada (Weir, 1996).

### 2.3.2 Mapeamento genético em cana-de-açúcar

Observa-se na literatura que há anos, várias equipes tentam estabelecer mapas genéticos para a cana-de-açúcar. Todavia, desafios inerentes à complexidade genética da planta, como a grande extensão do genoma decorrente da sua natureza poliploide, fizeram com que a quantidade e a saturação dos mapas gerados ficassem muito aquém do almejado pelos programas de melhoramento genético (Pastina et al., 2010).

Para a cana-de-açúcar e muitas outras espécies alógamas para as quais não há linhagens endogâmicas disponíveis, as populações segregantes que tradicionalmente são usadas para fins de mapeamento não podem ser obtidas com facilidade, em função dos longos ciclos necessários para se realizar cruzamentos e da forte depressão por endogamia gerada nos casos de autofecundações. Sendo assim, as populações segregantes a serem usadas para a construção de mapas genéticos de cana-de-açúcar são as de primeira geração ( $F_1$ ), obtidas a partir do cruzamento entre genótipos não-endogâmicos, isto é, com elevada heterozigosidade (Garcia et al., 2006).

Para estes tipos de espécies, conhecidas como *outcrossing species*, uma alternativa inicialmente utilizada para a construção de mapas genéticos foi a metodologia do duplo *pseudo-testcross*, que permitiu a construção dos primeiros mapas genéticos para *Eucalyptus grandis* e *Eucalyptus urophylla*, e, posteriormente foi também utilizada para construção de mapas de outras



espécies. Nesta metodologia a análise de ligação é feita para cada genitor, separadamente, resultando em dois mapas individuais, através da identificação dos marcadores de dose única presentes em cada um deles.

Nesse contexto é importante salientar que, diferentemente de espécies endogâmicas, nas quais há a necessidade de considerável divergência entre os genitores, principalmente no que se refere aos caracteres de interesse, quando populações  $F_1$  provenientes do cruzamento entre genótipos não-endogâmicos são usadas para fins de mapeamento, o contraste entre os genitores não agrega qualquer tipo de informação. Isto porque pela estratégia do *pseudo-testcross* a divergência entre os genitores não é mapeável, sendo que, o que se faz necessário nestes casos é a existência de uma condição heterozigótica de cada genitor, de forma que as frequências dos genótipos que constituirão a população segregante sejam informativas em relação à frequência de gametas recombinantes, permitindo, assim a estimação da distância genética entre os locos (Alwala & Kimbeng, 2010). Logo, nestes casos a informação usada para o mapeamento resulta do contraste entre os alelos de cada genitor e não do contraste entre os genitores, como melhor ilustrado por Maliepaard et al. (1997).

Em cana-de-açúcar, os primeiros mapas de ligação foram construídos utilizando estratégias como a do duplo *pseudo-testcross*, implementada em *softwares* como Mapmaker (Lander et al., 1987) e JoinMap (Van Ooijen & Voorrips, 2001). No início da década de 1990 foram obtidos os primeiros mapas construídos partir da análise direta do genoma de uma espécie poliploide complexa (Al-Janabi et al., 1993; Da Silva et al., 1993); pois, até esta época, os mapas genéticos obtidos para poliploides eram construídos somente através de análises de cruzamentos que envolviam espécies diploides relacionadas (Dufour et al., 1997). A Tabela 2.1, adaptada de Alwala et al. (2010) e Pastina et al. (2010) ilustra os mapas genéticos obtidos para clones de *S. spontaneum* que, posteriormente, serviram como referência para a construção dos mapas de ligação obtidos para *S. spontaneum*, *S. officinarum*, *S. robustum* e para variedades comerciais.

**Tabela 2.1.** Mapas genéticos obtidos para cana-de-açúcar (*Saccharum* spp.).

| Espécies mapeadas     | População de mapeamento                            | Tamanho da progênie | Tipo de marcador   | Dosagem dos marcadores   |                          |                  | Marcadores ligados | Grupos de ligação                      | Comprimento total do mapa em cM | Referência                   |
|-----------------------|----------------------------------------------------|---------------------|--------------------|--------------------------|--------------------------|------------------|--------------------|----------------------------------------|---------------------------------|------------------------------|
|                       |                                                    |                     |                    | SD <sup>1</sup><br>(1:1) | SD <sup>2</sup><br>(3:1) | DD <sup>3</sup>  |                    |                                        |                                 |                              |
| <i>S. spontaneum</i>  | ADP85-068 x SES208                                 | 90                  | RFLP, AP-PCR, RAPD | 216                      | -                        | -                | 188                | 44                                     | 1.361                           | Da Silva et al. (1993; 1995) |
| <i>S. spontaneum</i>  | ADP85-068 x SES208                                 | 88                  | RAPD               | 208                      | -                        | -                | 176                | 42                                     | 1.500                           | Al-Janabi et al. (1993)      |
|                       | SP701006                                           | 32                  | RFLP               | 253                      | -                        | -                | 94                 | 25                                     | -                               | D'Hont et al. (1994)         |
| Variedade comercial   | R570                                               | 77                  | RFLP, Isoenzimas   | 505                      | -                        | -                | 408                | 96                                     | 2.008                           | Grivet et al. (1996)         |
| <i>S. officinarum</i> | La Purple x Mol5829                                | 84                  | RAPD               | 279                      | -                        | -                | 161                | 50                                     | 1.152                           | Mudge et al. (1996)          |
| <i>S. spontaneum</i>  | Green German <sup>a</sup> x IND81-146 <sup>b</sup> | 85                  | RFLP               | 418 <sup>a</sup>         | -                        | 131 <sup>a</sup> | 270 <sup>a</sup>   | 72 <sup>a,c</sup><br>69 <sup>b,d</sup> | 2.304 <sup>a</sup>              | Ming et al. (1998)           |
|                       | Muntok Java <sup>c</sup> x PIN84-1 <sup>d</sup>    | 85                  |                    | 385 <sup>b</sup>         | -                        | 49 <sup>b</sup>  | 248 <sup>b</sup>   |                                        | 2.063 <sup>b</sup>              |                              |
|                       |                                                    |                     |                    | 355 <sup>c</sup>         | -                        | 156 <sup>c</sup> | 206 <sup>c</sup>   |                                        | 1.143 <sup>c</sup>              |                              |
| <i>S. officinarum</i> | La Purple <sup>a</sup> x Mol5829 <sup>b</sup>      | 100                 | RAPD, RFLP, AFLP   | 355 <sup>c</sup>         | -                        | 84 <sup>d</sup>  | 206 <sup>c</sup>   | 74 <sup>a</sup><br>65 <sup>b</sup>     | 1.303 <sup>d</sup>              | Guimarães et al. (1999)      |
|                       |                                                    |                     |                    | 297 <sup>d</sup>         | -                        |                  | 182 <sup>d</sup>   |                                        | 1.181 <sup>a</sup>              |                              |
| Variedade comercial   | R570                                               | 295                 | AFLP               | 341 <sup>a</sup>         | -                        | -                | 283 <sup>a</sup>   |                                        | 1.189 <sup>b</sup>              | Hoarau et al. (2001)         |
|                       | Green German <sup>a</sup> x IND81-146 <sup>b</sup> |                     |                    | 301 <sup>b</sup>         | -                        | -                | 208 <sup>b</sup>   |                                        |                                 |                              |
| <i>S. spontaneum</i>  | PIN84-1 <sup>c</sup> x Muntok Java <sup>d</sup>    | 85                  | RFLP               | 939                      | -                        | -                | 887                | 120                                    | 5.849                           | Ming et al. (2002)           |
|                       |                                                    |                     |                    | 434 <sup>a</sup>         | 132 <sup>a</sup>         |                  | 289 <sup>a</sup>   | 75 <sup>a</sup>                        | 2.446 <sup>a</sup>              |                              |
|                       |                                                    |                     |                    | 395 <sup>b</sup>         | 54 <sup>b</sup>          |                  | 257 <sup>b</sup>   | 70 <sup>b</sup>                        | 2.172 <sup>b</sup>              |                              |
|                       |                                                    |                     |                    | 308 <sup>c</sup>         | 86 <sup>c</sup>          |                  | 194 <sup>c</sup>   | 71 <sup>c</sup>                        | 1.395 <sup>c</sup>              |                              |
| Variedade comercial   | R570                                               | 112                 | RFLP, SSR          | 359 <sup>d</sup>         | 159 <sup>d</sup>         |                  | 214 <sup>d</sup>   | 73 <sup>d</sup>                        | 1.472 <sup>d</sup>              |                              |
| Variedade comercial   | IJ76-514 x Q165                                    | 227                 | AFLP, RAF, SSR     | 347                      | -                        | -                | 282                | 128                                    | -                               | Rossi et al. (2003)          |
| Variedade comercial   | IJ76-514 x Q165                                    | 227                 | AFLP, SSR          | 967                      | 36                       | 123              | 1074               | 136                                    | 9.058,3                         | Aitken et al. (2005)         |
| Variedade comercial   | IJ76-514 x Q165                                    | 227                 | AFLP, SSR          | 240                      | -                        | 234              | 534                | 123                                    | 4.906,4                         | Aitken et al. (2007)         |
| Variedade comercial   | Q117 x MQ77-340                                    | 232                 | AFLP, SSR          | 395                      | 58                       | -                | 342                | 101                                    | 3.582                           | Reffay et al. (2005)         |

Continua...

**Tabela 2.1.** Continuação.

| Espécies mapeadas    | População de mapeamento                            | Tamanho da progênie | Tipo de marcador                                   | Dosagem dos marcadores               |                          |                                    | Marcadores ligados                   | Grupos de ligação                   | Comprimento total do mapa em cM          | Referência              |
|----------------------|----------------------------------------------------|---------------------|----------------------------------------------------|--------------------------------------|--------------------------|------------------------------------|--------------------------------------|-------------------------------------|------------------------------------------|-------------------------|
|                      |                                                    |                     |                                                    | SD <sup>1</sup><br>(1:1)             | SD <sup>2</sup><br>(3:1) | DD <sup>3</sup>                    |                                      |                                     |                                          |                         |
| Variedade comercial  | SP80-180 <sup>a</sup> x SP80-4966 <sup>b</sup>     | 100                 | RFLP, SSR, AFLP                                    | 441                                  | 677                      | -                                  | 357 <sup>a</sup><br>217 <sup>b</sup> | 131 <sup>a</sup><br>98 <sup>b</sup> | 2.602 <sup>a</sup><br>1.340 <sup>b</sup> | Garcia et al. (2006)    |
| Variedade comercial  | R570 <sup>a</sup> x MQ76-53 <sup>b</sup>           | 198                 | AFLP, SSR, RFLP                                    | 1057                                 | -                        | -                                  | 424 <sup>a</sup><br>536 <sup>b</sup> | 86 <sup>a</sup><br>105 <sup>b</sup> | 1.144 <sup>a</sup><br>4.329 <sup>b</sup> | Raboin et al. (2006)    |
| <i>S. spontaneum</i> | Green German <sup>a</sup> x IND81-146 <sup>b</sup> | 169                 | SSR                                                | 125                                  | 26                       | -                                  | 91 <sup>a</sup><br>46 <sup>b</sup>   | 10 <sup>a</sup><br>25 <sup>b</sup>  | 1.180 <sup>a</sup><br>614 <sup>b</sup>   | Edmé et al. (2006)      |
| Variedade comercial  | M134/75 x R570                                     | 227                 | AFLP, SSR                                          | 557                                  | -                        | 79                                 | 474                                  | 95                                  | 6.200                                    | Al-Janabi et al. (2007) |
| Variedade comercial  | SP80-180 x SP80-4966                               | 100                 | AFLP, SSR, EST-SSR, RFLP, EST-RFLP                 | 800                                  | 869                      | -                                  | 664                                  | 192                                 | 6.261                                    | Oliveira et al. (2007)  |
| <i>S. spontaneum</i> | La Striped <sup>a</sup> x SES147B <sup>b</sup>     | 100                 | AFLP, SRAP, TRAP<br>AFLP e marcadores direcionados | 247 <sup>a</sup><br>221 <sup>b</sup> | -                        | 33 <sup>a</sup><br>43 <sup>b</sup> | 146 <sup>a</sup><br>121 <sup>b</sup> | 49 <sup>a</sup><br>45 <sup>b</sup>  | 1.732 <sup>a</sup><br>1.491 <sup>b</sup> | Alwala et al. (2008)    |
| Variedade comercial  | IAC66-6 x TUC71-7                                  | 188                 | <sup>a</sup> sequências de retrotransposons        | 485                                  | 172                      | -                                  | 503                                  | 91                                  | 4.316                                    | Palhares et al. (2010)  |
| Variedade comercial  | LCP85-384                                          | 300                 | AFLP, SSR, TRAP                                    | -                                    | 773                      | 182                                | 784                                  | 108                                 | 5.617                                    | Andru et al. (2011)     |

<sup>1</sup>SD(1:1): marcadores *sigle-dose* com segregação 1:1, presentes somente uma vez no genoma de um dos genitores. <sup>2</sup>SD(3:1): marcadores *sigle-dose* com segregação 3:1, presentes uma vez no genoma de cada um dos genitores. <sup>3</sup>DD: marcadores *double-dose*, presentes duas vezes no genoma dos genitores, podendo apresentar segregação 11:3 (x=8) ou 7:2 (x=10). a, b, c, d: refere-se às informações do mapa do respectivo genitor.

Da Silva et al. (1993) obtiveram um mapa de ligação para o mesmo clone, todavia, além de noventa genótipos provenientes do cruzamento entre ADP-850068 x SES208, estes autores utilizaram também uma população constituída por oitenta genótipos haploides derivados de cultura de anteras, cujo objetivo era avaliar o comportamento cromossômico de SES208 durante a meiose. Os resultados foram semelhantes aos de Al-Janabi et al. (1993), tendo sido obtido, a partir da análise com 116 sondas RFLP, um mapa com 44 GL e a confirmação de que SES208 de fato comportava-se como um autopoliploide.

Al-Janabi et al. (1993) utilizaram 88 genótipos provenientes do cruzamento entre um clone de *S. spontaneum*, SES208, e o seu duplo-haplóide ADP-850068, para construir um mapa para SES208. Nesse trabalho foram utilizados 127 *primers* RAPD, os quais geraram 2.160 marcas, sendo que 279 destas estavam presentes na variedade SES208 e ausentes em ADP-850068. O estudo permitiu a obtenção de informações importantes sobre o comportamento cromossômico da variedade SES208, a qual mostrou ser autopoliploide.

Tanto o mapa de Al-Janabi et al. (1993), quanto o obtido por Da Silva et al. (1993), por se constituírem estudos iniciais, apresentaram baixa cobertura do genoma e, entre as principais causas deste fato, pode-se citar: o pequeno número de genótipos analisados, a utilização de apenas um único tipo de marcador molecular e de um único tipo de segregação (SD 1:1). Com o intuito de amenizar estas deficiências, Da Silva et al. (1995) construíram novamente o mapa genético do clone SES208, mas combinando marcadores RFLP e RAPD. Dessa vez incluíram, além de 442 marcadores SD, 41 marcadores DD e um marcador MD. A adição destes tipos de segregações permitiu a formação total de 64 grupos de ligação, os quais foram organizados em 8 grupos de homeologia. Além disso, foi possível observar que SES208 apresentou segregação polissômica, confirmando ser este clone um autopoliploide.

Ainda utilizando marcadores RFLP, mapas relativamente mais saturados foram obtidos por Ming et al. (1998). Estes exploraram dois cruzamentos interespecíficos entre clones de *S. officinarum* e *S. spontaneum*: Green German x IND 81-146 e Muntok Java x PIN84-1. Para cada um destes clones obteve-se um mapa utilizando-se informações fornecidas por marcadores SD. Para os clones de *S. officinarum*, Green German e Muntok Java, 72 grupos

de ligação foram obtidos, sendo que para o primeiro o mapa apresentou extensão de 2.304 cM e para o segundo 1.443 cM. Para os clones de *S. spontaneum*, IND81-46 e PIN84-1, foram obtidos 69 grupos de ligação apresentando os mapas 2.063 cM e 1.103 cM, respectivamente.

Reffay et al. (2005), buscando avaliar a contribuição do clone Mandalay de *S. spontaneum* para as variedades australianas, utilizaram 232 indivíduos provenientes do cruzamento biparental entre Q117 x MQ77-340, para construir um mapa de ligação com informações de marcadores AFLP e SSR. O mapa foi construído a partir de 400 marcadores polimórficos, os quais formaram 101 grupos de ligação e geraram um mapa de 3.852 cM. Foi possível observar que, embora o mapa tivesse apresentado baixa saturação, informações importantes referentes à contribuição do genoma de Mandalay aos seus descendentes puderam ser obtidas. Verificou-se que em MQ77-340 25% dos marcadores eram provenientes de Mandalay, e 11% dos grupos de ligação demonstraram ser produtos da recombinação interespecífica entre *S. spontaneum* e *S. officinarum*.

Edmé et al. (2006) utilizaram uma população F<sub>1</sub> constituída por 169 genótipos provenientes do cruzamento entre os clones Green German (*S. officinarum*) e IND81-146 (*S. spontaneum*), para construir mapas de ligação para estas variedades. Cento e noventa e três marcadores SSR foram utilizados, embora destes, apenas 46 apresentaram segregação Mendeliana e foram incluídos na análise. Para Green German, 91 marcadores foram agrupados em 10 grupos de ligação, apresentando um mapa com extensão de 1.180 cM. Para IND81-146, 46 marcadores permitiram a formação de 25 grupos de ligação e mapa com 614 cM. Embora o autor afirme que a quantidade de marcadores utilizada tenha sido insuficiente para abranger a maior parte do genoma de ambas as espécies, pôde-se perceber que uma maior cobertura do genoma de *S. officinarum* foi obtida, e que rearranjos cromossômicos e eventos de translocação são recorrentes nos genomas de ambos os genitores.

O primeiro mapa obtido para uma variedade comercial foi construído por Grivet et al. (1996), a partir de 77 genótipos provenientes de autofecundação da variedade francesa R570. Um conjunto de 128 sondas RFLP foi utilizado e permitiu a detecção de 408 locos polimórficos, os quais foram suficientes para formar 96 grupos de ligação, gerando um mapa com extensão total de 2.008 cM.

Os resultados do estudo também permitiram a obtenção de informações referentes à origem da variedade R570, a qual apresentou, aproximadamente, 55% das marcas provenientes de *S. officinarum* e 45% provenientes de *S. spontaneum*.

Dufour et al. (1997), a partir de duas populações de linhas puras recombinantes de sorgo construíram um mapa de ligação composto usando 188 sondas RFLP, já mapeadas em milho e sorgo, e compararam o mapa obtido ao mapa construído por Grivet et al. (1996) para a variedade R570 de cana-de-açúcar. Ao realizarem este trabalho, os resultados previamente encontrados por D'Hont et al. (1996), sobre a colinearidade entre os genomas de cana-de-açúcar e sorgo, foram novamente confirmados. De acordo com os autores a utilização de genomas diploides geneticamente próximos ao da cana-de-açúcar, como o do sorgo, pode facilitar a compreensão da estruturação desse genoma, bem como ser utilizado para a identificação de genes homólogos de interesse agrônomo.

Posteriormente, Hoarau et al. (2001), utilizando marcadores AFLP e diferentes genótipos pertencentes à mesma progênie utilizada por Grivet et al. (1996), obtiveram, para a mesma variedade R570, um mapa mais saturado e com maior extensão. A partir de 37 pares de *primers* AFLP foi possível detectar 1.185 marcadores polimórficos, dos quais 939 apresentaram segregação 3:1. Destes, 887 foram distribuídos em 120 grupos de ligação, passando o mapa a apresentar 5.849 cM, extensão esta que representou um terço do comprimento total do genoma da cana-de-açúcar.

Apesar da obtenção de grande quantidade de marcas nesse trabalho, foi possível observar que muitos grupos de ligação permaneceram dissociados dos grupos de homeologia, revelando a dificuldade de se acessar, de maneira uniforme, todas as regiões cromossômicas do genoma da cana-de-açúcar. Além disso, verificou-se que 11 dos 120 grupos de ligação obtidos para a R570 resultaram de recombinações interespecíficas entre os cromossomos de espécies ancestrais.

À medida que os mapas de ligação foram sendo construídos, publicados e complementados com diferentes populações, marcadores e segregações, observou-se maior saturação e melhor caracterização do genoma pôde ser obtida, tornando-se comum a associação de mais de um tipo de

marcador molecular na construção dos mapas. Raboin et al. (2006) utilizaram uma progênie de 198 genótipos do cruzamento entre as variedades MQ76-53 e R570 e construíram mapas individuais para cada uma delas. Um total de 1.666 marcadores polimórficos foram obtidos a partir de 37 combinações de *primers* AFLP, 46 SSR e 9 sondas RFLP. Destes, 1.057 marcadores de dose única foram usados. Para a variedade R570, 424 marcadores específicos foram distribuídos em 86 grupos de ligação, apresentando o mapa uma extensão total de 3.144 cM. Quinhentos e oitenta e quatro marcadores específicos de MQ76-53 foram alocados em 105 grupos de ligação, e este mapa apresentou comprimento de 4.329 cM.

Raboin et al. (2006) afirmaram que, embora a quantidade de marcadores AFLP utilizada em ambos os trabalhos tenha sido a mesma, a cobertura do mapa obtido por eles para a variedade R570 foi menor do que a do mapa construído por Hoarau et al. (2001), sendo o tipo de progênie utilizada o provável motivo deste fato. Para estes autores, quando o objetivo é mapear um genitor, progênies oriundas de cruzamentos biparentais são menos informativas do que as progênies oriundas de autofecundação, já que marcadores presentes em ambos os genitores (SD 1:1) são pouco informativos.

Guimarães et al. (1999) analisaram cem genótipos provenientes do cruzamento entre um clone de *S. officinarum*, LA Purple, e outro de *S. robustum*, Mol 5829, para construírem mapas para cada genitor combinando 642 marcadores entre RAPD, RFLP e AFLP. O mapa obtido para LA Purple foi composto de 341 marcadores, distribuídos em 74 grupos de ligação e apresentou 1.881 cM, enquanto o mapa de Mol 5829 apresentou 301 marcadores distribuídos em 65 grupos de ligação e comprimento de 1.189 cM. Mudge et al. (1996) já haviam obtido um mapa de 1.152 cM para o clone LA Purple. Mas, como haviam sido usados apenas 160 marcadores RAPD, somente 51 grupos de ligação tinham sido obtidos, comprovando que a adoção de mais de um tipo de marcador em análises que visam o mapeamento de poliploides contribui consideravelmente para aumentar a saturação e a cobertura do genoma.

Aitken et al. (2005; 2007), utilizando uma progênie de 227 genótipos o cruzamento entre o clone IJ76-514 de *S. officinarum* e a variedade comercial australiana Q165, construíram o mapa mais extenso para cana-de-açúcar até o

presente momento. Nesse trabalho foram utilizadas 40 combinações de *primers* AFLP, 5 RAPD e 72 SSR. Para a variedade Q165 foram detectados 1.365 locos polimórficos, dos quais 910 foram distribuídos em 116 grupos de ligação, permitindo a obtenção de um mapa com extensão total de 9.058,3 cM, contendo 8 grupos de homeologia. Para o clone IJ76-514 foram geradas 595 marcas polimórficas, das quais, apenas 240 apresentaram segregação SD e puderam ser distribuídas em 47 grupos de ligação.

Tanto no mapa construído para a variedade Q165, como no mapa do clone IJ76-514 marcadores SD descartados nas análises iniciais e marcadores que apresentaram segregação DD e SD (3:1) foram posteriormente considerados nas análises. Com isso, 136 GL passaram a ser observados para a variedade Q165, enquanto para o clone IJ76-514, além de um acréscimo de 4.906 cM no comprimento total do mapa, 123 GL foram formados comprovando que, de fato, quando marcadores que apresentam segregações alternativas à SD (1:1) são considerados, uma maior saturação e cobertura do genoma são alcançadas.

Apesar da superação de inúmeros desafios relacionados às análises estatísticas e aos esforços realizados para obtenção de mapas saturados para cana-de-açúcar, ainda era desejável que se obtivesse a integração das informações contidas nos mapas individuais, a fim de facilitar a localização de QTL. E, para que isso fosse realizado, era necessária a presença de marcadores em heterozigose em ambos os genitores, os quais são utilizados para estabelecer relações de ligação entre os marcadores que segregam individualmente em cada genitor (Wu et al., 2002).

Contudo, em cana-de-açúcar a análise de ligação em genitores heterozigotos seria dificultada uma vez que, como diferentes números de alelos podem existir para cada loco, as fases de ligação nos genitores são desconhecidas *a priori*, dificultando a detecção dos eventos de recombinação.

Diante disto, Wu et al. (2002) desenvolveram o algoritmo EM (*Expectation Maximization*), baseado em máxima verossimilhança, que permitiu a construção de mapas genéticos integrados, uma vez que as frações de recombinação e as fases de ligação são estimadas simultaneamente. Através desse algoritmo, diferentes tipos de segregações passaram a ser consideradas (1:1 e 3:1) e analisadas em um único mapa integrado.



Essa metodologia permitiu o desenvolvimento de *softwares*, como, por exemplo, o OneMap (Margarido et al., 2007), capazes de corrigir as dificuldades inerentes ao desconhecimento das fases de ligação *a priori* e permitir a obtenção de um mapa integrado. Então Garcia et al. (2006) utilizaram este *software* para construir o mapa integrado para as variedades brasileiras pré-comerciais SP80-180 e SP80-4966. A partir da combinação de 57 sondas RFLP, 64 *primers* SSR e 236 *primers* AFLP, 1.118 marcas polimórficas puderam ser identificadas, tendo sido suficientes para formar 131 grupos de ligação em um mapa com extensão de 2.602,4 cM e densidade de 7,3 cM. Esses autores também analisaram o mesmo conjunto de dados utilizando o *software* JoinMap, desenvolvido para análise de espécies diploides e populações de mapeamento constituídas por irmãos completos.

Observaram que um maior número de marcas ficou desligado, resultando, conseqüentemente, em um mapa com menor quantidade de grupos de ligação (98), menor densidade (6,2 cM) e comprimento (1.340 cM). Neste caso, foi possível constatar que a metodologia proposta por Wu et al. (2002), baseada em máxima verossimilhança e implementada no *software* OneMap, foi o método mais eficiente para gerar um mapa integrado.

O trabalho de Garcia et al. (2006) foi estendido por Oliveira et al. (2007), através da incorporação de mais 149 marcadores SSR e 10 sondas RFLP, ambos derivados de EST (*Expressed Sequence Tag*). A adição destes marcadores gerou 1.669 marcas SD, das quais 664 (40%) foram distribuídas em 192 GL, originando um mapa com 6.261,1 cM. Portanto, a utilização dos marcadores provenientes de sequências expressas aumentou consideravelmente a cobertura e a densidade do mapa.

De acordo com Oliveira (2007) e Pinto et al. (2010), o uso de marcadores derivados de sequências expressas, ao invés de sequências anônimas de DNA, apresenta um grande potencial na construção de mapas funcionais e de QTL, pois possibilitam inúmeras vantagens como, por exemplo, o fato de que se um marcador EST estiver geneticamente associado a uma característica de interesse, é possível que o gene mapeado afete diretamente tal característica. Além disso, há grande tendência de que estas sequências apresentem homologia a genes candidatos ou a genes que estejam

diferencialmente expressos no tecido de interesse, facilitando assim suas identificações.

De acordo com Pastina et al. (2010), as sequências dos EST foram originalmente concebidas com o objetivo de identificar transcritos gênicos, mas tem sido utilizados como uma ferramenta para identificar genes, bem como informações referentes a seus mecanismos de regulação e expressão. A identificação destas sequências tem avançado muito rapidamente, estando hoje, disponíveis nos bancos de dados aproximadamente 6 milhões de EST.

Teixeira (2006), visando mapear EST relacionados a genes de interesse em cana-de-açúcar, utilizou a mesma progênie que Garcia et al. (2006). Quarenta sondas EST, provenientes do projeto Sucest (*Sugarcane Expressed Sequence Tag*), e 83 sondas genômicas obtidas aleatoriamente de um banco genômico de *Saccharum* spp. foram utilizadas. O mapa obtido apresentou 33 grupos de ligação e extensão de 574 cM. Para a autora, os marcadores derivados de EST são mais eficientes do que os obtidos por sondas genômicas, tanto para a detecção de polimorfismo, quanto para a identificação de prováveis QTL. Em seu trabalho, observou-se que 90% das sondas EST apresentaram polimorfismo e geraram 512 fragmentos polimórficos entre os genitores, enquanto apenas 66% das sondas genômicas apresentaram polimorfismo e geraram apenas 112 fragmentos.

Alwala et al. (2008) utilizaram uma população constituída por cem genótipos provenientes do cruzamento entre o clone Louisiana Striped de *S. officinarum* e o clone SES147B de *S. spontaneum*, e, através de marcadores SRAP, AFLP e TRAP construíram um mapa para cada genitor. Um total de 344 marcadores polimórficos foi obtido a partir de Louisiana Striped, dos quais 72% apresentaram segregação SD e 9% apresentaram segregação DD. Para o clone SES147B, 306 marcadores polimórficos foram gerados, dos quais 72% apresentaram segregação SD e 14% segregação DD. O mapa para o clone de *S. officinarum* foi obtido a partir de 146 marcadores alocados em 49 GL e apresentou comprimento de 1.732 cM. Para o clone de *S. spontaneum*, 121 marcadores formaram 45 grupos de ligação e o mapa total apresentou extensão de 1.491 cM.

De acordo com Alwala et al. (2008), os mapas obtidos para os clones de *S. officinarum* e *S. spontaneum* cobriram, respectivamente, 69% e 46% dos

genomas destas espécies, e permitiram a confirmação de que a primeira espécie comporta-se como um auto-alopoliploide, enquanto a segunda comporta-se como um autopoliploide. Nesse estudo também foi possível observar que dentre os marcadores utilizados, AFLP foi mais eficiente para gerar maior número de marcas, enquanto marcadores SRAP e TRAP foram capazes de marcar regiões do genoma ricas em genes, podendo, portanto, serem utilizados com eficiência quando o objetivo for a construção de um mapa de QTL.

Palhares (2010) utilizou AFLP e marcadores baseados em *retrotransposons* para construir um mapa de ligação integrado para as variedades comerciais IAC66-6 e TUC71-7. Um total de 600 marcadores SD foi obtido a partir de 22 combinações de enzimas de restrição de AFLP, assim como seis combinações otimizadas para amplificar marcas direcionadas a dois *retrotransposons*. De acordo com a autora, o mapeamento de marcadores derivados de *retrotransposons* é importante, pois normalmente estes estão próximos a regiões flangeadoras de genes em plantas. O mapa obtido apresentou 91 grupos de ligação e uma extensão de 4.316,5 cM, tendo revelado que *retrotransposons* não estão distribuídos uniformemente nestes grupos. Além disso, um número de cópias relativamente baixo ao longo de todo o genoma das variedades de cana-de-açúcar analisadas.

Andru et al. (2011) construíram um mapa de ligação para a popular variedade da Louisiana LCP 85-384, utilizando 64 marcadores AFLP, 19 SSR e 12 TRAP, em uma população constituída por trezentos genótipos provenientes de autofecundação. A combinação desses três tipos de marcadores permitiu a obtenção de 1.111 marcas polimórficas das quais, 773 apresentaram segregação SD e 182 segregação DD. Como o genoma de LCP 85-384 foi estimado em 12.313 cM, o mapa obtido, com comprimento total de 5.617 cM, foi capaz de cobrir apenas 45,6% do genoma. No entanto, informações fornecidas pelos marcadores SSR que apresentaram segregação DD, e as referentes às fases de repulsão foram úteis para permitir a atribuição de 44 dos 108 grupos de ligação a 9 grupos de homeologia. De acordo com os autores este, assim como a maioria dos mapas obtidos para cana-de-açúcar até o momento, apresentou baixa densidade, de 7,16 cM, e a principal razão disso é a complexidade genética e a grande extensão do genoma da planta.

A partir de uma breve análise dos mapas genéticos publicados na literatura para cana-de-açúcar percebe-se que, até o presente momento, apesar dos grandes avanços alcançados quanto à possibilidade de se construir um mapa integrado e à utilização de informações provenientes de diferentes tipos de marcadores e populações, a obtenção de mapas saturados que possuam ampla cobertura do genoma ainda se constitui num desafio. A complexidade genômica da cana-de-açúcar, decorrente de sua condição poliploide e os recorrentes eventos de aneuploidia, associados ao descarte dos marcadores que não apresentam as segregações comumente utilizadas nas análises de mapeamento, tem contribuído consideravelmente para que a obtenção de um mapa saturado para esta planta ainda não tenha sido alcançada.

Difícilmente marcadores que apresentam segregações DD ou MD têm sido incorporados com frequência aos mapas de ligação. Ao desconsiderar estas segregações, além da necessidade de se utilizar maior quantidade de marcadores, visando aumentar a probabilidade de ocorrência dos que apresentam segregação SD, maiores dificuldades também são encontradas nas análises de ligação e ordenação, resultando em mapas pouco saturados e com reduzida cobertura do genoma (Garcia et al., 2006).

Wu et al. (1992) explicam que os marcadores SD são os menos informativos e que a ordenação e a estimação da distância entre marcadores SD em fase *trans* é possível somente através da utilização de grande população de mapeamento. Mas, Da Silva et al. (1995) sugerem que a incorporação de marcadores DD ou MD pode evitar a necessidade de se trabalhar com grande grupos de homeologia, a identificação indireta dos marcadores SD em *trans* pode ser facilitada, aumentando assim a saturação dos mapas.

Estima-se que um terço do genoma da cana de açúcar tenha sido representado nos mapas, embora para as variedades comerciais a cobertura seja desigual, com os cromossomos de *S. spontaneum* sendo mais densamente cobertos que os de *S. officinarum*. De acordo Heller-Uszynska et al. (2011), o elevado número de cromossomos e a grande extensão desse genoma tornam necessárias quantidades extensivas de marcadores moleculares (entre 5.000 e 7.000) para que sejam construídos mapas genéticos com cobertura e densidade consideráveis.

Mais recentemente, a identificação e desenvolvimento de marcadores SNP (*Single Nucleotide Polymorphism*) derivados de EST de cana-de-açúcar, representou uma valiosa ferramenta a ser utilizada para obtenção de ampla cobertura desse genoma. Marcadores SNP são, geralmente, de natureza bi-alélica, abundantes no genoma e podem ocorrer tanto em regiões expressas quanto em regiões não-expressas. De acordo com Bundock et al. (2009), os SNP são altamente frequentes no genoma de cana-de-açúcar, apresentando uma frequência de um a cada 35 pb, sendo este tipo de polimorfismo fundamental para obtenção de mapas genéticos de alta resolução, bem como a realização da seleção genômica ampla.

Cordeiro et al. (2006) identificaram diversos SNP derivados de EST (EST-SNP) disponíveis em bancos de dados de cana-de-açúcar e revelaram que pode ser obtida uma média de 9 SNP por EST ou 1 SNP a cada 50 pb, embora variações nesta frequência tenham sido observadas entre as variedades australianas e as espécies selvagens. Além disso, foi possível observar em uma população de mapeamento entre as variedades de cana-de-açúcar australianas IJ76-514 e Q165, que os marcadores EST-SNP apresentaram as segregações esperadas, mostrando que SNP com ocorrência em um ou dois cromossomos de cana-de-açúcar apresentam possibilidade de serem mapeados.

Para Pastina et al. (2010), a abundância de SSR e SNP em regiões transcritas do genoma e o elevado grau de polimorfismo destes marcadores tem tornado as bibliotecas de EST um recurso valioso, capaz de fornecer marcadores genéticos que muito poderão contribuir para a compreensão do genoma da cana-de-açúcar.

A utilização dos marcadores DArT (*Diversity Arrays Technology*), proposta inicialmente por Jaccoud et al. (2001), também tem se apresentado como uma alternativa eficiente na obtenção de ampla cobertura do genoma, por ser capaz de gerar uma grande quantidade de informações de vários locos simultaneamente, sem a necessidade de informações prévias das sequências de DNA ou de sequências específicas de oligonucleotídeos. Os marcadores DArT permitem identificar paralelamente entre centenas e milhares de sequências polimórficas do tipo SNP ou *indel* e são equivalentes aos marcadores RFLP; porém de forma reversa e com um desempenho muito maior, uma vez que as

sondas são imobilizadas em um microarranjo no qual o DNA a ser analisado é hibridizado.

Os marcadores DArT têm sido utilizados em diversas culturas diploides de importância agrônômica, com as mais diversas finalidades, como em estudos de diversidade genética, caracterização de bancos de germoplasma e construção de mapas genéticos. Em cana-de-açúcar, os marcadores DArT foram utilizados com o objetivo de avaliar o efeito dos diversos métodos de redução da complexidade do genoma repetitivo (Heller-Uszynska et al., 2011), e também para a obtenção de um mapa de ligação integrado para as variedades RB97327 e RB72454 (Silva, 2012).

## 2.4 MAPEAMENTO DE QTL (*Quantitative Trait Loci*)

### 2.4.1 Princípios do mapeamento de QTL

As características agrônômicas de importância econômica que são selecionadas nos programas de melhoramento genético são, em sua grande maioria, poligênicas ou de herança quantitativa; isto é, são resultantes da ação conjunta de vários locos que podem apresentar efeitos variáveis e cuja expressão fenotípica recebe forte influência do ambiente. As regiões cromossômicas que contém estes locos constituem os *Quantitative Trait Loci* (QTL) (Tanksley, 1993).

De acordo com Coelho (2000), os caracteres quantitativos são muito estudados, porém pouco compreendidos, devido à sua natureza complexa e importância que apresentam. Esta complexidade está relacionada ao fato de estes caracteres serem controlados por muitos locos e por serem fortemente influenciados pelo ambiente.

Quando um caráter é quantitativo as diferenças nas expressões fenotípicas dos genótipos são mínimas, dificultando as seleções feitas exclusivamente em campo. Neste caso, os marcadores moleculares, por acessarem diretamente o genótipo, constituem importantes ferramentas a serem utilizadas na construção de mapas de ligação, pois, associados às informações fenotípicas, possibilitam a construção dos mapas de QTL (Ferreira & Grattapaglia, 1998).

Bearzoti (2000) afirma que mapear QTL significa fazer inferências em todo o genoma sobre as relações entre o genótipo e o fenótipo de caracteres quantitativos. Essas inferências incluem informações sobre número e posição dos locos no genoma, efeitos dos locos sobre a característica de interesse, modo de ação gênica (efeitos de aditividade, dominância, heterose e epistasia) e a possibilidade de decompor a interação genótipo x ambiente em nível de cada QTL. Logo, pode-se afirmar que o mapeamento de QTL tem por objetivo não apenas identificar os locos de interesse, mas também permitir a compreensão da forma como os caracteres quantitativos são arquitetados para que assim possam ser estabelecidas estratégias de seleção que maximizem a eficiência dos programas de melhoramento genético (Bertrand & Mackill, 2008).

A identificação de QTL permite ainda a sua clonagem, isto é, a transferência de uma sequência de DNA para um plasmídeo ou algum outro vetor a fim de que seja possível estudá-lo em diferentes tecidos e condições, compará-lo a sequências homólogas em outras espécies e ainda inserí-lo em diferentes espécies. É importante considerar ainda que, embora a seleção assistida por marcadores (SAM) seja possível sem o mapeamento de QTL, este pode ser útil no direcionamento de cruzamentos específicos, maximizando tanto o número de QTL com alelos favoráveis, como o valor genotípico global, incluindo também efeitos epistáticos (Bearzoti, 2000).

Hotta et al. (2010) afirmam que, embora as aplicações dos mapas de QTL sejam promissoras, é fato que a sua utilização prática ainda tem se mostrado limitada, mesmo que diversos métodos de detecção de QTL relacionados às mais diferentes características tenham sido desenvolvidos. Entre os prováveis motivos que justificam este fato destacam-se os estudos normalmente realizados em populações muito pequenas, com baixa precisão das avaliações fenotípicas, poucos *backgrounds* genéticos testados e baixa precisão da posição e magnitude dos QTL identificados.

Liu (1998) afirma que QTL são genes hipotéticos encontrados com base em inferências baseadas em análises estatísticas e cujos efeitos no mapeamento podem ter um significado biológico muito pequeno. Além disso, relata também que os modelos adotados para a construção de mapas de QTL não são acurados, sendo que a maioria das informações contidas nos mapas não

são adequadas, pois as metodologias estatísticas utilizadas não conseguem lidar com os altos níveis de interações epistáticas. Para Coelho (2000), os modelos existentes para explicar o controle dos caracteres quantitativos representam “supersimplificações” face aos intrincados e complexos caminhos do metabolismo que levam à expressão dos respectivos genes, principalmente se considerados os efeitos que o ambiente exerce sobre estes mecanismos.

A detecção de QTL numa determinada população pode exigir a construção prévia de um mapa de ligação ou não. Procedimentos mais simples, como o método BSA - *Bulked Segregant Analysis* (Michelmore et al., 1991) e a análise baseada em um único marcador – *Single Marker Analysis* (Lynch & Walsh, 1998), também conhecida como análise de marcas simples, não exigem um mapa genético para que QTL sejam detectados.

A estratégia utilizada pelo método BSA é considerada a mais direta na associação de um marcador a um QTL e consiste na avaliação, para um determinado carácter quantitativo, de dois grupos de indivíduos obtidos de pontos extremos de uma população segregante e na ocorrência de polimorfismo em locos de marcadores entre os grupos. Espera-se que os marcadores que estiverem associados aos alelos que participam do controle do QTL estejam distribuídos de forma desigual nos dois grupos, permitindo assim a sua detecção. Marcadores não associados aos alelos do QTL em questão estarão presentes igualmente nos dois grupos (Coelho, 2000).

Como BSA requer uma concentração dos alelos relacionados ao QTL nos pontos extremos da distribuição do carácter, este método não é eficiente para detectar QTL de efeito reduzido e exige que sejam utilizados cruzamentos a partir de linhagens endogâmicas ou progênes de populações segregantes; já que, da mesma forma que outros métodos, este também necessita da presença de desequilíbrio de ligação entre os marcadores e o QTL. Esta metodologia, embora promova uma considerável economia de tempo e recursos e seja eficiente na detecção de ligação entre marcador e QTL, é pouco eficiente na determinação dos efeitos individuais do QTL e quando os indivíduos com fenótipos extremos não são facilmente distinguíveis (Lynch & Walsh, 1998).

A análise baseada em um único marcador, proposta por Lynch & Walsh (1998), ou análise de marcas simples, permite determinar o modo de herança do



marcador e a sua localização no genoma. Através de testes estatísticos simples, como teste  $t$ , análise de variância, regressão linear simples ou teste da razão de verossimilhanças, busca-se estabelecer associações significativas entre os valores fenotípicos médios e os valores genotípicos das diferentes classes de marcadores. Assim, se o resultado destas associações for significativo, indicará ligação entre o marcador e o QTL. Estas comparações são realizadas para cada loco marcador separadamente, sem considerar as informações dos outros locos; daí o nome “análise baseada em um único marcador”. É importante ressaltar que esta análise não necessariamente mapeia um QTL, mas permite associar marcas a ele (Tanksley, 1993).

Entre as principais vantagens da análise de marcas simples estão a simplicidade de sua execução, pois não exige conhecimentos avançados de estatística, podendo ser realizadas em *softwares* genéricos destinados a análises estatísticas, como SAS (SAS Institute, 1989) e R (R Development Core Team, 2012). A sua rapidez e o fato de não haver necessidade de obtenção prévia de um mapa de ligação são outras destas vantagens. No entanto, uma desvantagem do método resume-se no fato de que, se os marcadores não saturarem completamente o genoma e/ou as amostras forem pequenas, quanto maior a distância entre QTL e marcador, menor será a probabilidade de detectá-lo estatisticamente. Isso devido à possibilidade de recombinação entre o marcador e o QTL, reduzindo consideravelmente o poder de detecção do QTL (Liu, 1998).

Além disso, a magnitude do efeito do QTL nem sempre pode ser discernida na análise de marcas simples, ou seja, um QTL de grande efeito pode não ser distinguido caso esteja próximo a QTL de pequenos efeitos. Ambas as desvantagens podem ser minimizadas se o número de marcadores utilizados na população segregante permitir elevada densidade e ampla cobertura do genoma, de maneira que a distância entre um QTL e um marcador seja a mínima possível (Tanksley, 1993).

Observa-se na literatura que grande quantidade dos mapas de QTL construídos para espécies diploides, a partir de populações provenientes de linhagens endogâmicas, foram obtidos por análises que requerem a construção prévia de um mapa de ligação, como o mapeamento por intervalo (Lander & Botstein, 1989), o mapeamento por intervalo composto (Zeng, 1993; Jansen,

1993) e o mapeamento por intervalo múltiplo (Kao, 1999). É importante considerar que embora estes métodos de mapeamento estejam bem implementados em *softwares* como Mapmaker/Exp (Lander et al., 1987), QTL-Cartographer (Basten et al., 2005), MapQTL (Van Ooijen, 2009), dentre outros, ainda há limitações às suas utilizações pelos programas de melhoramento, uma vez que todos eles foram delineados para mapeamento de QTL relacionados a apenas um caráter e em um único ambiente.

Métodos que visam identificar QTL com efeitos pleiotrópicos e que consideram a interação genótipo x ambiente também já foram desenvolvidos e, com isso, melhorias na precisão da estimação dos parâmetros puderam ser obtidas, uma vez que possibilitaram maior poder para a detecção de QTL cujas características são correlacionadas foi possibilitado (Jiang & Zeng, 1995). Outras abordagens que visam o mapeamento de múltiplos caracteres, simultaneamente, também foram desenvolvidas (Knott & Haley, 2000). Mas, provavelmente, dificuldades inerentes à interpretação dos resultados obtidos justificam ainda suas poucas aplicações.

#### 2.4.4.1 Mapeamento por intervalo (MI)

O mapeamento por intervalo (MI), proposto por Lander & Botstein (1989), ofereceu uma alternativa para se mapear QTL através de análises que visam associar duas marcas adjacentes a um QTL, na forma de um intervalo. Esta abordagem permite estimar o efeito e a posição dos QTL e tem como base informações de segregação de pares de marcadores adjacentes, como unidades de análise (Liu, 1998).

No método MI, Lander & Botstein (1989) propuseram que a fração de recombinação e a magnitude do efeito do QTL no intervalo entre dois marcadores deve ser estimada por máxima verossimilhança. Neste caso, o teste de hipótese utiliza o valor de LOD (*Logarithm of Odds*), o qual indicará quantas vezes é mais verossímil existir um QTL no intervalo sob análise, em relação à hipótese de sua ausência. Desta forma, a essência deste método é, portanto, testar a possibilidade de ocorrência de um QTL numa série de posições dentro de um intervalo definido por dois marcadores (Lander & Botstein, 1989).

Haley & Knott (1992) afirmam que MI também pode ser realizado com base em análise de regressão. De acordo com esses autores, a regressão do valor fenotípico observado de determinado caráter sobre um conjunto de possíveis genótipos para determinado QTL possibilita estimar a ocorrência do QTL em um intervalo delimitado por dois marcadores. Assim, para cada possível genótipo, um QTL possui uma probabilidade de ocorrência em função de sua localização no intervalo e dos genótipos dos marcadores.

Embora MI baseado em análise de regressão seja uma alternativa ao de máxima verossimilhança, proposto por Lander & Botstein (1989), observa-se na literatura que este é muito mais utilizado, uma vez que a informação completa da distribuição da associação marcador-caráter é utilizada. Todavia, se o modelo incluir muitos QTL, este método pode se tornar inviável, pois, as análises computacionais requeridas necessitam de muitas etapas de métodos iterativos para maximizar a máxima verossimilhança (Schuster & Cruz, 2008).

Diante do exposto, observa-se que MI apresenta vantagens em relação à análise baseada em um único marcador, como, por exemplo, o fato de permitir a obtenção de mapas que apresentam baixa densidade com melhor resolução, e, em casos nos quais há grande quantidade de dados perdidos e as análises não serem muito prejudicadas. Todavia, em MI os intervalos são considerados independentes, ficando a análise sujeita à ocorrência de falsos positivos, caso existam QTL de grandes efeitos em regiões adjacentes. Além disso, como todos os QTL fora do intervalo não são considerados, toda a variação devida a eles é considerada resíduo, diminuindo desta forma a precisão do teste (Tanksley, 1993).

#### 2.4.1.2 Mapeamento por intervalo composto (MIC)

Jansen (1992; 1993) e Zeng (1993;1994) propuseram, separadamente, uma metodologia que combina um modelo de regressão múltipla ao mapeamento por intervalo, a qual foi denominada de mapeamento por intervalo composto (MIC). Esta metodologia foi proposta com o objetivo de corrigir a principal desvantagem do método MI, que é a detecção de falsos positivos em função da presença de QTL de grandes efeitos em intervalos adjacentes. Além disso, Liu

(1998) complementa ainda que em MIC os marcadores limitam a provável posição em que o QTL se encontra, melhorando consideravelmente a resolução deste local e garantindo a obtenção de maior informatividade sobre o loco.

No método MIC considera-se a informação de múltiplos locos de cada indivíduo e os marcadores adicionais são incluídos na análise como co-fatores, reduzindo a variância residual, a qual, normalmente, é resultante da influência de QTL fora do intervalo em análise. Nesse caso, diferentes marcadores são considerados, como os que estão: ao redor do QTL, ligados dentro da região do QTL, ligados fora da região do QTL, não ligados dentro da região do QTL e não ligados fora da região do QTL (Liu, 1998).

Definir quantos e quais destes marcadores devem ser utilizados como co-fatores na regressão não é, porém, um procedimento simples. Jansen (1993) propôs que, para isso, uma análise de regressão múltipla seja inicialmente realizada com todos os marcadores com efeito significativo e que se utilize como co-fatores os marcadores que se apresentam como significativo mas que não estão ligados ao principal intervalo em estudo. O método MIC apresenta algumas desvantagens como: o fato de uma distribuição irregular dos marcadores afetar a análise, a dificuldade de se estimar a contribuição conjunta de múltiplos QTL ligados e a não inclusão dos efeitos epistáticos no modelo.

A busca simultânea de QTL em todo o genoma, permitindo detectar tanto QTL individuais, como realizar a análise de epistasia entre múltiplos QTL foi posteriormente proposta por Kao et al. (1999). Este método ficou conhecido como mapeamento por intervalo múltiplo (MIM) e permitiu uma maior eficiência e precisão na identificação do QTL, melhor entendimento e identificação dos locos epistáticos e a estimação sem viés dos efeitos dos QTL.

#### **2.4.2 Mapeamento de QTL em cana-de-açúcar**

O melhoramento de plantas contribuiu significativamente para promover incrementos de produtividade em várias culturas por meio da seleção de genótipos que apresentam os caracteres fenotípicos que refletem o valor genotípico dos indivíduos para os genes que controlam o caráter sob seleção. No entanto, os caracteres que apresentam importância agronômica, que em cana-de-

açúcar são aqueles relacionados à produtividade de açúcar e de cana, e à resistência a doenças, são controlados por muitos genes e fortemente influenciados pelo ambiente (Raboin et al., 2006).

Diante disso, a utilização das informações fenotípicas associadas àquelas fornecidas pelos marcadores moleculares tem permitido um melhor conhecimento do número, da localização, do mecanismo de ação, dos efeitos e das interações dos locos que contribuem para os caracteres quantitativos, através do mapeamento de QTL. Além de todas estas informações, o mapeamento de QTL também permite identificar a existência de correlações genéticas entre características e assim estabelecer esquemas de seleção para mantê-las ou interrompê-las, dependentemente de suas vantagens. Tendo sido os QTL mapeados, é possível também identificar qual ambiente é mais favorável à expressão de determinados QTL, bem como obter um melhor entendimento de como o genoma da espécie está estruturado (Pastina et al., 2010). Diante de todos esses benefícios que o mapeamento de QTL pode oferecer ao melhoramento convencional, dada a crescente importância econômica da cultura da cana-de-açúcar e as dificuldades associadas à obtenção de variedades que atendam à produtividade almejada pelos métodos convencionais de melhoramento, acredita-se que esta metodologia possui grande potencial para aumentar a eficiência dos programas de melhoramento genético desta cultura (Ming et al., 2002a).

Os primeiros mapas de QTL em cana-de-açúcar surgiram aproximadamente na mesma época em que foram publicados os primeiros mapas de ligação. Foram obtidos a partir dos métodos convencionalmente utilizados para diploides, como análises baseadas em um único marcador e MI. Na maioria dos mapas obtidos foram identificados QTL relacionados aos caracteres envolvidos com a produtividade de açúcar e seus componentes, como teor de açúcar (Pol), brix, toneladas de açúcar por hectare, toneladas de cana por hectare, teor de fibras, peso, número, comprimento e diâmetro de colmos, florescimento (Sills et al., 1995; Guimarães et al., 1999; Ming et al., 2001, 2002a, 2002b; Hoarau et al., 2002; Rossi et al., 2003; Da Silva & Bressiani, 2005; Reffay et al., 2005; Aitken et al., 2006, 2008; Piperidis et al., 2008; Alwala et al., 2009; Pinto et al., 2010; Pastina et al., 2012), e resistência a doenças (Sills et al., 1995; Guimarães et al.,

1999; Daugrois et al., 1996; Asnaghi et al., 2000, 2004; McIntyre et al., 2005; Wei et al., 2006; Raboin et al., 2006; Al-Janabi et al., 2007).

Sills et al. (1995), através da análise de marcas simples, identificaram, na mesma população estudada por Mudge et al. (1996), 24 associações significativas entre marcadores e número de colmos, porcentagem de colmos ramificados, resistência ao carvão (*Ustilgo scitaminea*), diâmetro de colmos, teor de sacarose e de fibras, e produtividade. Três interações epistáticas significativas também foram identificadas entre as características sob avaliação. Posteriormente, Guimarães et al. (1999), utilizando a mesma progênie, encontraram as mesmas associações significativas verificadas por Sills et al. (1995).

Daugrois et al. (1996) utilizaram a população de estudo avaliada por Grivet et al. (1996), proveniente da autofecundação da variedade R570, para avaliarem o padrão de herança da resistência à ferrugem marrom, causada pelo fungo *Puccinia melanocephala*. Através da análise de marcas simples, foi possível verificar um marcador de grande efeito relacionado à resistência à doença, ligado a 10 cM de uma sonda RFLP. Este trabalho documenta o primeiro gene com herança monogênica, denominado Bru1, a ser detectado no genoma da cana-de-açúcar.

Dando prosseguimento ao trabalho de Daugrois et al. (1996), Asnaghi et al. (2000) buscaram verificar relações de sintenia através da comparação de mapas de cana, milho, sorgo e arroz. A comparação destes mapas permitiu a identificação de segmentos cromossômicos próximos ao gene Bru1 entre todas as culturas, e, de acordo com os autores, esta elevada colinearidade do genoma de cana-de-açúcar com genomas menores e mais simples, especialmente o do sorgo, é promissora para o desenvolvimento de recursos moleculares capazes de auxiliar a identificação de genes de importância agrônômica em cana-de-açúcar.

Posteriormente, Asnaghi et al. (2004), buscando confirmar a presença do gene Bru1 no genoma da variedade R570, utilizaram 443 marcadores AFLP em uma população constituída por 695 indivíduos, derivados de autofecundação, e em 2.102 indivíduos oriundos de cruzamentos biparentais, nos quais a R570 foi utilizada como genitor masculino, em combinação com quatro diferentes progenitores femininos não comerciais (M555, Triton, BT 64/134 e R94).

Utilizando o método BSA, oito marcadores AFLP foram identificados próximos ao gene Bru1. No entanto, os autores afirmam que, embora os marcadores AFLP tenham garantido ampla saturação do genoma, por gerarem grande quantidade de fragmentos, para que sejam eficientes em análises de mapeamento, um perfil de melhor qualidade é exigido.

Estudos adicionais de QTL para populações oriundas da autofecundação da variedade R570 também foram realizados por Hoarau et al. (2002) e por Rossi et al. (2003). Estes, a partir do mapa de ligação previamente construído por Hoarau et al. (2001). Em uma progênie de 295 indivíduos foram encontrados quarenta prováveis QTA (*Quantitative Trait Alleles*), cujos efeitos individuais observados foram baixos, variando entre 3% e 7%, relacionados a quatro componentes de produção para cana-de-açúcar: altura da planta observou-se ainda, através de análises digênicas, a ocorrência de 41 interações epistáticas (Hoarau et al., 2002).

De acordo com Hoarau et al. (2002), em cana-de-açúcar a abordagem baseada no mapeamento de alelos com efeitos quantitativos é utilizada na tentativa de se encontrar associações significativas entre as variações fenotípicas observadas nos caracteres de interesse e a presença de alelos que podem estar segregando para um loco específico. Contudo, espera-se que os efeitos destes QTA em organismos poliploides sejam, de fato, menores do que QTL observados em espécies diploides, devido principalmente ao elevado número de alelos segregantes por loco.

O mapa inicialmente construído por Grivet et al. (1996) e estendido por Hoarau et al. (2001) para a variedade R570 foi complementado por Rossi et al. (2003), os quais conseguiram identificar 88 genes análogos de resistência (RGA) no banco de dados do Sucest, com base em sua homologia típica para genes de resistência, incluindo três dos maiores grupos de genes de resistência. Esses genes foram utilizados como sondas e geraram um total de 148 marcadores RFLP SD, os quais foram integrados ao mapa genético da variedade R570 previamente complementado por Hoarau et al. (2001).

Raboin et al. (2006), utilizando marcadores AFLP e SSR, construíram, a partir de 198 genótipos provenientes do cruzamento entre R570 x MQ-76-53, um mapa de ligação para R570 com 3.144 cM e um mapa de ligação para MQ-76-

53 com 4.329 cM . Cento e sessenta e seis destes genótipos foram avaliados quanto à resistência/suscetibilidade à ferrugem marrom, entre os anos de 2003 e 2006, em condições naturais. Através da análise de marcas simples, pôde-se observar, na progênie, marcadores AFLP flanqueando o gene de resistência à ferrugem marrom, Bru1, previamente identificado em R570 por Daugrois et al. (1996) e Asnaghi et al. (2004).

QTL relacionados à resistência à mancha amarela também foram identificados por Al-Janabi et al. (2007) em estudo conduzido com população segregante em que a variedade R570 também foi utilizada como progenitor. Neste caso, utilizaram uma progênie proveniente do cruzamento entre M134/75 x R570 e através de MIC e de informações fenotípicas referentes a dois locais e dois anos, dois QTL associados à esta doença foram identificados, sendo que um destes explicou 23,8% da variação fenotípica.

De acordo com Ming et al. (2002a), o mapeamento de QTL em autopoliploides é ainda mais complicado que em aloploiploides, uma vez que há ausência de pareamento preferencial. No entanto, um subconjunto de alelos polimórficos que apresentam segregação SD pode facilitar a identificação de QTL nestes casos. Partindo deste princípio, Ming et al. (2001) utilizaram os mapas previamente obtidos por Ming et al. (1998), relativos a duas populações interespecíficas provenientes do cruzamento entre clones de *S. officinarum* e *S. spontaneum*, para identificar QTL referentes ao conteúdo de açúcar, com o intuito de avaliar a base genética da variação desta característica entre genótipos autopoliploides de cana-de-açúcar. Através de MI, para o cruzamento entre PIN 84-1 x Muntok Java, puderam ser identificadas 22 associações significativas; e para Green German x IND 81-146, quatorze associações foram encontradas para locos relacionados aos teores de sacarose.

Em seguida, Ming et al. (2002a), utilizando a mesma progênie que Ming et al. (2001), mas adicionando informações fenotípicas de três anos consecutivos, identificaram, em análises separadas, 61 QTL relacionados ao teor de açúcar, dos quais 50 foram aglomerados em 12 regiões genômicas. Observaram que plantas que possuíam elevado teor de sacarose apresentavam um maior número de QTL positivos para os componentes relacionados à produção de açúcar, indicando que a SAM pode representar estratégia eficiente a



ser adotada nos programas de melhoramento genético de cana-de-açúcar quando uma quantidade de informações suficientes sobre a população em estudo já foi obtida.

Essa mesma população foi utilizada por Ming et al. (2002b) para verificar a possibilidade de associações entre marcadores e altura e florescimento. Sessenta e cinco associações significativas foram observadas, das quais trinta puderam ser alocadas nos grupos de ligação previamente obtidos em mapas genéticos. Todos os dados genotípicos para esta população foram provenientes de marcadores RFLP.

Jordan et al. (2004)<sup>3</sup>, citados por Pastina et al. (2010) e Margarido (2011), a partir de estudo em uma população obtida pelo cruzamento entre os clones australianos Q117 x 74C42, em dois locais e durante dois anos, identificaram por análise de marcas simples sete marcadores RFLP associados ao número de colmos e seis associados ao número de perfilhos. As treze associações mostraram-se consistentes em ambos os locais e anos, e três marcadores apresentaram associações significativas aos dois caracteres, revelando que é possível, no melhoramento da cana-de-açúcar, selecionar concomitantemente genótipos para elevado número de colmos e baixo número de perfilhos.

Da Silva & Bressiani (2005), visando mapear QTL relacionados ao teor de sacarose e de fibras, utilizaram progênie do cruzamento entre as variedades SP80-180 x SP80-4966 e, através de análises de marcas simples, identificaram um marcador EST-RFLP inversamente associado ao teor de sacarose, indicando que os genótipos que apresentaram este marcador apresentaram níveis baixos de sacarose. No entanto, o marcador foi verificado apenas no genitor masculino SP80-4966, o qual apresenta níveis de sacarose considerados bons, sugerindo que, nesta variedade, provavelmente outros alelos contribuem para a produção de sacarose.

Pinto et al. (2010) deram continuidade a este trabalho e utilizaram sondas de DNA genômico e EST, que apresentavam homologia com genes envolvidos no metabolismo de carboidratos. Esses autores conduziram ensaios

---

<sup>3</sup>JORDAN, D. R.; CASU, R. E.; BESSE, P.; CARROL, B. C.; BERDING, N.; McINTYRE, C. L. Markers associated with stalk number and suckering in sugarcane collocate with tillering and rhizomatousness QTLs in sorghum. **Genome**, Ottawa, v. 47, p. 988-993, 2004.

com marcadores RFLP visando identificar QTL relacionados a teor de fibras, teor e produtividade de açúcar e produtividade de cana, bem como possíveis interações epistáticas. Foi possível identificar 120 associações entre marcadores e QTL, das quais 26 ocorreram tanto na avaliação de cana-planta, quanto em cana-soca, sendo 32 observadas apenas nesta última. Os autores também verificaram que um marcador derivado da sacarose sintase foi associado a um possível QTL com elevado efeito negativo sobre a produtividade e efeito positivo sobre o teor de açúcar, em ambos os estágios da cultura. Para os quatro caracteres avaliados, observou-se também a ocorrência de cinquenta interações epistáticas, das quais apenas duas foram observadas em ambos os estágios de desenvolvimento das plantas.

Reffay et al. (2005) obtiveram mapas de ligação para os clones MQ77-340 e Q117, através da utilização de marcadores AFLP e SSR e dados fenotípicos referentes a dois anos de avaliações, para teor de açúcar e brix, e um ano de avaliação para teor de fibra, peso e produtividade. Quarenta e uma associações, tanto positivas quanto negativas, foram identificadas, embora estas tenham sido obtidas considerando-se os dados fenotípicos de cada ano separadamente.

Piperidis et al. (2008), buscando validar QTL encontrados em cana-de-açúcar, compararam os QTL previamente identificados por Reffay et al. (2005), para as variedades MQ77-340 e Q117, aos identificados nas variedades R570 e Q165. O estudo permitiu confirmar consistência entre os QTL identificados nas variedades em estudo, bem como observar que em dois dos oito grupos de homeologia uma maior quantidade de marcadores associados a brix foi observada. Além disso, observaram que todos os 10 grupos de homeologia da variedade Q117 e os 11 HG de MQ77-340 poderiam ser reatribuídos a sete dos oito HG esperados para cana, revelando que um grupo de homeologia completo ainda não foi representado nos mapas e que os sete restantes ainda necessitam de maior cobertura.

McIntyre et al. (2005) utilizaram indivíduos provenientes do cruzamento entre Q117 e 74C42 e, através de informações obtidas por marcadores AFLP, RFLP e SSR, identificaram, por análise de marcas simples sete QTL associados à resistência a podridão radicular e quatro associados à resistência a ferrugem

marrom, os quais explicaram entre 6% e 18% da variação fenotípica. Estes resultados foram posteriormente confirmados por Wei et al. (2006), que também verificaram, em estudo de mapeamento associativo conduzido com 154 clones de uma coleção, os mesmos marcadores associados à resistência a ferrugem marrom.

Wei et al. (2006) afirmam que foi possível obter dados referentes às quatro doenças mais importantes que acometem as variedades de cana-de-açúcar australianas: podridão radicular, escaldadura das folhas, doença de Fiji e carvão. Para identificação dos QTL utilizaram 19 marcadores AFLP e 14 SSR. As análises foram baseadas em regressão simples, o que permitiu concluir que 59% da variação fenotípica referente ao carvão puderam ser explicadas por onze marcadores, 32% da variação referente à escaldadura e à podridão radicular foram explicados por quatro marcadores, e 26% da variação associada à doença de Fiji foram explicados por cinco marcadores.

Aitken et al. (2006), a partir do mapa obtido por Aitken et al. (2005), identificaram 37 associações significativas para brix e teor de sacarose, através de análises de marcas simples, realizadas com informações fenotípicas obtidas em dois anos consecutivos. Cada QTL explicou de 3% a 9% da variação fenotípica observada e, juntos entre 37% e 66%, tendo sido observados tanto efeitos positivos como negativos. De acordo com os autores, a alocação da maior parte dos QTL num mesmo HG sugere que, o que se identificou como QTL tratava-se de formas alélicas de um mesmo gene. Neste estudo, 97 interações epistáticas também foram identificadas, das quais nove foram consistentes nos dois anos de avaliação fenotípica. Estas interações explicaram, individualmente, entre 11% e 13% da variação fenotípica e, quando foram consideradas em conjunto, explicaram entre 24% e 36% dessa variação.

A mesma progênie avaliada por Aitken et al. (2006) foi utilizada por Aitken et al. (2008) em estudo que buscou mapear QTL, através de MIC, para caracteres relacionados à produtividade, utilizando informações de campo obtidas durante três anos consecutivos. No estudo identificaram 18 QTL associados ao peso, 8 associados à altura, 7 ao número de colmos e 19 ao diâmetro do colmo, além de 6 QTL relacionados à produtividade. De acordo com os autores, estes QTL explicaram entre 2% e 9% da variação fenotípica e esta pequena magnitude

dos efeitos dos QTL foi justificada pelo fato de que a seleção recorrente pode ter favorecido a concentração de alelos relacionados ao teor de sacarose, à medida que diminuiu o contraste entre QTL individuais.

Alwala et al. (2009) associaram ao mapa de ligação obtido por Alwala et al. (2008) informações referentes aos caracteres envolvidos na produção de açúcar. Para o clone de *S. officinarum*, 30 QTL puderam ser observados, os quais explicaram entre 22% e 48% da variação fenotípica, dependendo da característica. Para o clone de *S. spontaneum* 11 QTL foram identificados tendo explicado entre 9% e 43% da variação fenotípica.

Embora se observe que muitos QTL relacionados à produtividade e à resistência a doenças tenham sido identificados em cana-de-açúcar, poucos exemplos de suas aplicações práticas têm sido relatados na literatura. De acordo com Pastina et al. (2010, 2012), os QTL identificados em cana-de-açúcar até o presente momento explicam apenas de 2% a 22,6% da variação fenotípica observada, indicando que as metodologias utilizadas para identificação de QTL ainda não contemplam situações próximas das práticas de melhoramento, tornando-se, em muitas situações, falhas.

Observa-se que na maioria dos trabalhos a identificação de QTL em cana-de-açúcar foi realizada por meio da análise de marcas simples ou através de MI, sendo raros os QTL obtidos por MIC. Estes tipos de análises são muito sujeitas à identificação de QTL “fantasmas”, em função da ocorrência de falsos positivos próximos aos QTL alvos do estudo o que reduz muito o poder de detecção de QTL reais. Além disso, Pastina et al. (2010) afirmam que nenhum destes métodos deveria ser utilizado para a construção de mapas de QTL em cana-de-açúcar, uma vez que todos foram desenvolvidos para análise de populações oriundas de cruzamentos endogâmicos, não sendo assim adequadas para populações como as de cana-de-açúcar.

De acordo com Hoarau et al. (2002), em poliploides como a cana-de-açúcar, como vários alelos coexistem num mesmo loco, para que o efeito de um alelo de determinado loco seja percebido é necessário que o seu efeito exceda não apenas o efeito de um outro alelo qualquer, como em diplóides, mas também o efeito médio dos demais alelos segregantes para o mesmo loco; além de considerar ainda o complexo padrão de interações epistáticas e dominantes

existente entre e dentro dos locos. Nos mapas de QTL de cana-de-açúcar, quando interações epistáticas foram estudadas, a abordagem utilizada foi a digênica, a qual não inclui a influência de outros QTL, sejam eles epistáticos ou não.

Nesse contexto é importante considerar que, como a cana-de-açúcar é uma espécie semi-perene, durante a realização dos vários cortes, as plantas ficam sujeitas a diferentes condições ambientais. Sendo assim, faz-se necessária a identificação de QTL que se expressem continuamente ao longo de vários anos, cortes e diferentes condições ambientais, a fim de que, de fato, esta ferramenta possa ser útil ao melhoramento. Nos mapas de QTL em cana-de-açúcar que utilizaram dados referentes a vários cortes e/ou locais, os métodos estatísticos foram aplicados isoladamente para cada ano, não sendo esta metodologia apropriada para se concluir se o mesmo QTL foi expresso em diferentes estágios da planta.

Pastina et al. (2012) propuseram uma estratégia para detecção de QTL em múltiplos locais x safra x cortes, baseada em MI e modelos mistos. A abordagem foi utilizada em população constituída por cem genótipos oriundos do cruzamento biparental entre as cultivares brasileiras pré-comerciais SP80-180 x SP80-4966, avaliada em dois locais e três cortes consecutivos. Foram avaliados caracteres envolvidas na produção de açúcar, o que permitiu a identificação de 46 QTL, além de informações referentes à posição, estabilidade e interações entre QTL. Pôde-se perceber, através da utilização desta nova abordagem, que as interações QTL x safra, QTL x local e QTL x corte foram significativas para todas as características avaliadas, indicando a necessidade de que informações fenotípicas referentes a várias características e anos simultaneamente sejam consideradas.

Outra observação refere-se à qualidade dos mapas de ligação utilizados para posterior identificação de QTL. Grande parte destes mapas foram obtidos através das metodologias convencionais, anteriores à proposta por Wu et al. (2002), em que os mapas eram construídos para cada um dos genitores separadamente. Notadamente, estes mapas apresentam desvantagens que reduzem sua capacidade de detecção de QTL, como, por exemplo, o fato de as populações normalmente utilizadas para a construção dos mapas serem

pequenas, de apenas um só tipo de marcador e/ou apenas os que apresentaram segregação SD terem sido incluídos na análise, resultando em mapas pouco saturados e com baixa cobertura do genoma (Aitken et al., 2008; Pastina et al., 2012).

Além disso, a maior parte dos estudos foi conduzida com populações provenientes de cruzamentos biparentais, que, na maioria dos casos, envolveram os mesmos genitores. Em poucos trabalhos foram utilizadas progênies obtidas por autofecundação e sabe-se que as informações obtidas com este tipo de população são fundamentais para a compreensão da estruturação do genoma da cana-de-açúcar. Em alguns casos, quando produtos da autofecundação foram usados para construção de mapas genéticos e identificação de QTL, trataram-se, quase sempre de progênies provenientes da autofecundação da variedade francesa R570, não permitindo avaliar se o comportamento do genoma dessa variedade pode ser extrapolado para as demais.

Diante do exposto, observa-se que, ainda que os progressos referentes à compreensão da estruturação do genoma da cana-de-açúcar possam parecer lentos é válido afirmar que eles têm ocorrido. E mesmo que inúmeros desafios ainda precisem ser superados, a necessidade de se obter variedades capazes de responder à crescente demanda da cultura há de fazer com que esforços extras sejam dedicados à inovação de metodologias desenvolvidas exclusivamente para a análise do genoma particular da cana-de-açúcar.

### 3 MAPEAMENTO GENÉTICO DE MARCADORES DArT (*Diversity Arrays Technology*) E SSR (microssatélites) EM CANA-DE-AÇÚCAR (*Saccharum spp.*)

#### RESUMO

Visando auxiliar a compreensão da forma como as características genéticas estão arquitetadas no genoma da cana-de-açúcar, no presente estudo foram obtidos os primeiros mapas genéticos integrados utilizando marcadores DArT (*Diversity Arrays Technology*) e SSR (microssatélites) para cana-de-açúcar, avaliados em uma população constituída por 91 genótipos derivados da autofecundação do genótipo-elite RB97327, e em outra população constituída por 81 genótipos oriundos do cruzamento RB97327 x RB72454. O DNA genômico foi extraído de gemas axilares e após a quantificação do DNA das amostras, realizou-se a genotipagem dos marcadores SSR e dos marcadores DArT. Para a genotipagem de SSR, vinte pares de *primers* foram avaliados em ambas as populações, tendo sido gerado um total de 115 marcas polimórficas na população de autofecundação e 132 marcas na população de fecundação cruzada. Para a genotipagem dos marcadores DArT, o DNA extraído, quantificado e homogeneizado foi enviado à empresa DArT P/L para a geração dos respectivos marcadores. Um total de 7680 marcadores foi analisado, os quais geraram 471 marcas polimórficas na população de autofecundação e 850 marcas polimórficas na população de fecundação cruzada. A análise de segregação mendeliana permitiu a distinção de 392 marcas DArT e 57 marcas SSR presentes na população de autofecundação, e 632 DArT e 79 marcas SSR polimórficas na população de fecundação cruzada, com segregação *single-dose*. As análises de ligação foram realizadas nas duas populações separadamente. O *software* OneMap foi utilizado para a construção dos mapas. Os critérios estabelecidos para proceder as análises de ligação foram LOD-score maior ou igual 3,5 e fração de recombinação menor ou igual 0,3. No mapa derivado da população de autofecundação, as 449 marcas DArT e SSR polimórficas com segregação 3:1 foram utilizadas para originar 95 grupos de ligação referentes à variedade RB97327. Esse mapa apresentou um comprimento total de 1.217,2 cM. O tamanho estimado do genoma de RB97327 foi de 10.540,9 cM, o que permite afirmar que o mapa obtido apresentou baixa cobertura (11,5%). Para a população derivada de cruzamento, as 711 marcas DArT e SSR polimórficas com segregação 3:1 e 1:1 originaram 136 grupos de ligação e o mapa apresentou um comprimento total de 2.722,2 cM. Embora uma pequena quantidade de marcas SSR tenha sido obtida, estas possibilitaram a identificação de seis possíveis grupos de homeologia no mapa referente ao genitor feminino RB97327 e nove no mapa integrado. Apesar do baixo polimorfismo, os marcadores DArT mostraram-se eficientes para genotipagem de progênies de cana-de-açúcar, pois, centenas de marcas polimórficas foram geradas em apenas um ensaio, com dois métodos de redução da complexidade genômica. Além disso, a utilização das sequências destes marcadores deverá contribuir para o desenvolvimento do *assembly* de um genoma de referência de cana-de-açúcar.

*Palavras-chave:* mapa de ligação, marcadores moleculares, poliploide.

## ABSTRACT

As a contribution to our understanding of how genetic characters are organized in the sugarcane genome, in the present study we report the first genetic maps using DArT (Diversity Arrays Technology) and SSR (microsatellite) markers. The maps were obtained using two populations: one consisting of 91 genotypes derived from selfing of elite-genotype RB97327, and the other consisting of 81 genotypes derived from the crossing RB97327 x RB72454. Genomic DNA was extracted from auxillary buds and, after quantification, was used to obtain genotypes for SSR and DArT markers. Twenty pairs of SSR primers were evaluated in both populations. A total of 115 SSR polymorphic markers in the selfing population and 132 SSR markers in the outcrossing population were generated. To generate the DArT markers, DNA samples were obtained, checked for quality and sent to the DArT P/L. A total of 7680 markers was produced, which generated 471 polymorphic markers in the selfing population and 850 polymorphic markers in the full-sib population. Mendelian segregation analysis allowed the detection of 392 DArT and 57 SSR markers in the population of selfing, and 632 DArT and 79 SSR polymorphic markers in the outcrossing population, segregating as single-dose markers. The linkage analyses were performed separately for the two populations. The OneMap software was used to construct the maps. Linkage analyses were carried out using 3.5 and 0.3 as critical values for LOD-score and recombination fraction, respectively. In the map derived from the selfing population, the 449 polymorphic markers with 3:1 segregation were used to originate 95 linkage groups for the variety RB97327. This map had a total length of 1217.2 cM. The estimated size of the genome of RB97327 was 10540.9 cM, suggesting that the obtained map is still of low coverage (11.5%). In the population derived from crossing, 711 polymorphic loci with 1:1 and 3:1 segregation patterns originated 136 linkage groups. The map showed an overall length of 2722.2 cM. Although only a small number of SSR markers was used, these markers allowed the identification of six possible groups of homeology in the map for the female parent and nine homeology groups in the integrated map. Despite the low polymorphism, DArT markers showed to be effective for genotyping sugarcane progenies. Hundreds of polymorphic markers were generated in only one trial, with two methods of reduction of genome complexity. The use of sequences of these markers should contribute to the development of a reference genome assembly for sugarcane.

*Key words:* linkage map, molecular markers, polyploid.

## 3.1 INTRODUÇÃO

A cana-de-açúcar está entre as gramíneas de grande importância econômica nos países tropicais, sendo as variedades cultivadas atualmente híbridos interespecíficos derivados do cruzamento entre *Saccharum officinarum* L. ( $2n=70-140$ ) e *Saccharum spontaneum* L. ( $2n=36-128$ ). A utilização de



ferramentas citológicas e moleculares revelou que essas variedades são organismos poliploides, amplamente heterozigóticos e que apresentam recorrentes aneuploidias (Grivet & Arruda et al., 2001).

Tais características fazem com que a complexidade genética inerente ao genoma da cana-de-açúcar represente um desafio para os programas de pesquisa em genética e melhoramento da cultura. O melhoramento genético da cana-de-açúcar apresenta importante relevância por permitir a seleção de variedades com caracteres agrônômicos e industriais favoráveis. Nesse contexto, as informações fornecidas por marcadores moleculares disponibilizam novas estratégias de análise genética capazes de auxiliar a obtenção de novas variedades (Albino et al., 2006).

No início da década de 1990, a utilização das informações fornecidas por diferentes tipos de marcadores, associada a uma série de procedimentos estatístico-computacionais, possibilitou o desenvolvimento dos primeiros mapas de ligação para clones de cana-de-açúcar. Estes mapas iniciais foram construídos para *S. spontaneum* (Al-Janabi et al., 1993; Da Siva et al., 1993; Ming et al., 1998) e foram obtidos a partir de informações de populações oriundas de cruzamentos biparentais. Posteriormente, tais mapas serviram como referencial para a construção de mapas de ligação para *S. officinarum* (Mudge et al., 1996; Ming et al., 1998; Guimarães et al., 1999; Edmé et al., 2006; Aitken et al., 2007; Alwala et al., 2008), *S. robustum* (Guimarães et al., 1999) e para as variedades cultivadas (Grivet et al., 1996; Dufour et al., 1997; Hoarau et al., 2001; Aitken et al., 2005, 2007; Refray et al., 2005; Raboin et al., 2006; Garcia et al., 2006; Oliveira et al., 2007; Palhares, 2010; Andru et al., 2011).

Os mapas genéticos são capazes de gerar grande quantidade de informações básicas sobre a estrutura e organização do genoma de uma espécie, além de permitir a obtenção de informações que auxiliam os programas de melhoramento genético, por exemplo, na localização e avaliação da magnitude do efeito das regiões genômicas associadas aos caracteres de importância agrônômica (QTL – *Quantitative Trait Loci*); na realização de estudos evolutivos e de mapeamento comparativo; na clonagem de genes a partir de suas posições previamente identificadas nos mapas; e ainda na realização de seleção assistida por marcadores (SAM) (Liu, 1998; Coelho et al., 2000).

Os mapas de ligação para cana-de-açúcar obtidos por Grivet et al. (1996) e Hoarau et al. (2001) foram os primeiros a serem construídos a partir de populações derivadas de autofecundação. De acordo com esses autores, a utilização deste tipo de população permitiu desvendar as contribuições genômicas relacionadas às espécies ancestrais, bem como melhor compreender o comportamento cromossômico em espécies geneticamente complexas como a cana-de-açúcar.

A maioria dos mapas inicialmente construídos foi obtida a partir de populações oriundas de cruzamentos biparentais, utilizando-se a estratégia do duplo *pseudo-testcross*, na qual mapas individuais são elaborados para cada um dos genitores. Apesar da superação de inúmeros desafios relacionados às metodologias de genotipagem e às análises estatísticas, ainda era desejável que se obtivesse a integração das informações contidas nos mapas individuais de cada genitor, bem como mapas com considerável saturação, a fim de facilitar a localização de QTL.

Diante disso, Wu et al. (2002) desenvolveram metodologia que permite a construção de um mapa genético integrado, no qual diferentes tipos de segregação gerados por vários tipos de marcadores passaram a ser consideradas e analisadas em um único mapa integrado. De acordo com Garcia et al. (2006) e Oliveira et al. (2007), a integração dos mapas revela-se amplamente promissora aos programas de melhoramento genético de cana-de-açúcar, pois permite uma melhor caracterização do polimorfismo, tornando mais precisa a identificação de QTL.

Com o intuito de revelar a eficiência do polimorfismo repetitivo e do polimorfismo gerado por vários locos simultaneamente, distribuídos ao longo de todo o genoma, bem como compreender a complexa estrutura genômica da cana-de-açúcar, o presente estudo teve como objetivo a construção dos primeiros mapas de ligação integrados para cana-de-açúcar, utilizando marcadores DArT e SSR, a partir de populações derivadas de autofecundação e de fecundação cruzada, buscando estimar também o tamanho do genoma estudado.

## 3.2 MATERIAL E MÉTODOS

### 3.2.1 Material vegetal

Foi utilizada uma população constituída por 91 genótipos, oriunda da autofecundação de um genótipo-élite da Ridesa RB97327 e outra, constituída por 81 genótipos, proveniente do cruzamento entre RB97327 e a variedade comercial RB72454. O genótipo RB97327 possui como um de seus genitores a variedade H641881, a qual tem sido utilizada em um grande número de cruzamentos pelo programa de melhoramento da Ridesa, por apresentar elevados valores de toneladas de cana/hectare, teor de fibra, brix e pureza (Moraes et al., 2010).

A variedade RB72454 constitui uma variedade plantada há mais de 30 anos no Brasil e representa um dos genitores mais utilizados pelos programas brasileiros de melhoramento genético de cana-de-açúcar. Destaca-se por apresentar elevado teor de sacarose e produtividade e por possuir considerável sanidade a um número considerável de patologias que comumente afetam a produtividade da cana-de-açúcar (Ferreira et al., 2005).

### 3.2.2 Extração e quantificação do DNA genômico

O DNA genômico total foi obtido pela extração realizada a partir da coleta de duas gemas laterais dos colmos de cada genótipo, a qual foi feita seguindo o protocolo de Al-Janabi et al. (1999) com adaptações para microtubos de 1,5 mL. No total realizou-se a extração de DNA de 172 plantas e dos genitores envolvidos no cruzamento em estudo. A quantificação do DNA extraído foi realizada através do espectrofotômetro Qubit (*Invitrogen*®).

### 3.2.3 Genotipagem dos marcadores SSR

Na genotipagem dos marcadores SSR, foram testados 35 pares de *primers* SSR, selecionados com base na literatura. Destes, vinte apresentaram polimorfismo entre os genitores e foram utilizados na genotipagem dos marcadores SSR (Tabela 3.1).

Dos vinte pares de *primers* utilizados, dezessete foram sintetizados diretamente com a sequência *forward* marcada com fluorocromo específico (6FAM, HEX e NED), de forma a permitir a montagem de cinco painéis triplex e um duplex. O triplex restante foi sintetizado de acordo com a metodologia descrita por Missiaglia & Grattapaglia (2006), denominada “*multiple-tailed*”, segundo a qual, *primers* previamente marcados com alguma fluorescência se hibridizam especificamente à extremidade 5' de um dos três *primers* presentes na reação.

**Tabela 3.1.** Referências dos marcadores SSR utilizados na genotipagem de 91 genótipos, provenientes da autofecundação de RB97327, e de 81 genótipos oriundos do cruzamento entre as variedades RB97327 e RB72454.

| Marcador SSR | Referências            |
|--------------|------------------------|
| mSSCIR11     | Cordeiro et al. (2000) |
| mSSCIR12     | Cordeiro et al. (2000) |
| mSSCIR14     | Cordeiro et al. (2000) |
| mSSCIR21     | Cordeiro et al. (2000) |
| mSSCIR24     | Cordeiro et al. (2000) |
| mSSCIR25     | Cordeiro et al. (2000) |
| mSSCIR26     | Cordeiro et al. (2000) |
| mSSCIR35     | Cordeiro et al. (2000) |
| mSSCIR55     | Cordeiro et al. (2000) |
| mSSCIR56     | Cordeiro et al. (2000) |
| mSSCIR58     | Cordeiro et al. (2000) |
| mSSCIR67     | Cordeiro et al. (2000) |
| SMC222CG     | Cordeiro et al. (2000) |
| SMC477CG     | Cordeiro et al. (2000) |
| MCSA042E08   | Cordeiro et al. (2001) |
| MCSA053C10   | Cordeiro et al. (2001) |
| MCSA223B07   | Cordeiro et al. (2001) |
| SEGMS79a     | Parida et al. (2009)   |
| SEGMS200a    | Parida et al. (2009)   |
| SEGMS792     | Parida et al. (2009)   |

O DNA dos genótipos em estudo foi diluído para a concentração de 3 ng  $\mu\text{L}^{-1}$  para a realização das reações para amplificação dos marcadores SSR. As reações foram feitas utilizando o kit Multiplex PCR Kit (*Qiagen*), seguindo as recomendações do fabricante; mas, reduzindo-se o volume final de cada reação para 5  $\mu\text{L}$ , utilizando, desta forma, 1/10 da quantidade de cada componente recomendada pelo fabricante. As reações foram realizadas no termociclador

GeneAmp PCR System 3700 (*Applied Biosystems*), o qual foi programado para realizar uma etapa inicial de 95°C/15 min., seguido de PCR *touchdown*, com temperatura inicial de 60°C para anelamento dos *primers*, durante os cinco primeiros ciclos, e redução de 0,5°C por ciclo nos trinta ciclos seguintes.

Os produtos da PCR foram revelados no analisador automático ABI PRISM 3100 (*Life Technologies*), a partir de mistura contendo 1,0 µL do produto da reação de PCR; 8,85 µL de formamida Hi-Di (*Life Technologies*); e 0,15 µL de solução contendo os fragmentos de tamanho padrão (GS500-ROX, *Life Technologies*). Os cromatogramas resultantes foram analisados utilizando-se o *software* GenneMapper® 3.0 (*Life Technologies*).

A genotipagem dos marcadores SSR, tanto na população de autofecundação, quanto na população oriunda de cruzamento, foi realizada seguindo a metodologia proposta por Wu et al. (1992), segundo a qual cada fragmento deve ser considerado como um marcador, codificando-se a sua presença com “1” e a sua ausência com “0”. Os alelos derivados da amplificação dos marcadores SSR foram identificadas pelo nome do marcador, seguido do número indicativo do tamanho do fragmento gerado.

### **3.2.4 Genotipagem dos marcadores DArT**

Os dados referentes aos marcadores DArT foram gerados pela empresa *Diversity Arrays Technology PTy Ltd* (DArT P/L), localizada na cidade de Canberra, Austrália. Para isso, o DNA genômico dos 174 genótipos foi diluído para 60 ng µL<sup>-1</sup> e 20 µL deste foi distribuído em duas placas de 96 poços, as quais foram lacradas e enviadas à prestadora dos serviços, a temperatura ambiente.

Na obtenção das marcas DArT foram usadas duas enzimas de restrição para redução da complexidade genômica (scPth e scPtb), ambas indicadas na identificação dos marcadores. A partir desse procedimento, obteve-se os dados de genotipagem de 7680 marcadores referentes aos 174 genótipos em estudo, os quais também foram codificados com “1” para presença e “0” para ausência.

### 3.2.5 Análise da segregação Mendeliana

Os dados referentes aos marcadores SSR foram reunidos aos de marcadores DArT, de modo que todas as análises de mapeamento genético fossem realizadas para cada uma das duas populações, separadamente. Assim, com o objetivo de selecionar quais marcadores, em cada população, deveriam ser incluídos nas análises, realizou-se inicialmente a análise da segregação Mendeliana.

No presente estudo foram considerados marcadores *single-dose* (SD) que apresentaram segregação 1:1, referente a cada genitor (*ao* x *oo* e *oo* x *ao*), e marcadores que apresentaram segregação 3:1 (*ao* x *ao*). As frequências observadas foram submetidas ao teste de aderência de qui-quadrado, sob a hipótese  $H_0$  de ausência de distorção da segregação Mendeliana. Para correção da taxa de erro tipo I, foi utilizado o critério FDR (*False Discovery Rate*), com  $\alpha=0,01$ . Os marcadores que apresentaram outros tipos de segregação, que não as mencionadas anteriormente, foram desconsideradas das análises de construção dos mapas de ligação.

### 3.2.6 Construção dos mapas de ligação

Diante da existência de duas populações de mapeamento, procedeu-se, separadamente, à construção de dois mapas genéticos de ligação. Ambos contendo todas as marcas ligadas, sendo um primeiro referente ao genitor feminino RB97327, construído a partir dos dados que apresentaram segregação 3:1, oriundos dos 81 genótipos da população de autofecundação; e um segundo mapa integrado, construído a partir dos dados referentes aos 91 genótipos resultantes do cruzamento entre RB97327 e RB72454.

A construção dos mapas foi realizada utilizando-se o pacote OneMap (Margarido et al., 2007) desenvolvido para ambiente computacional R (*R Development Core Team*, 2012). Este *software* possibilita a estimação simultânea das frações de recombinação e das fases de ligação entre os marcadores, bem como a análise de diversos tipos de marcadores contendo diferentes padrões de segregação (Wu et al., 2002). A abordagem multiponto é realizada por

verossimilhança, usando o modelo markoviano oculto (*Likelihood Hidden Markov Model*), e o teste de dois pontos entre as marcas é feito através da conversão realizada pelo algoritmo EM (*Expectation Maximization Algorithm*) para o cálculo das frações de recombinação.

Ambos os mapas foram obtidos utilizando-se a melhor ordem dos marcadores sugerida pelo *software* nos grupos de ligação. Os critérios adotados para a construção de ambos os mapas foram  $LOD \geq 3,5$  e fração de recombinação  $fr \leq 0,3$ . Em ambos os mapas os valores das frações de recombinação entre as marcas foram convertidos em centiMorgans (cM), utilizando a função de mapeamento proposta por Kosambi (1944). Os desenhos dos grupos de ligação foram obtidos no *software* MapChart 2.2 (Voorrips, 2002).

### 3.2.7 Estimação do tamanho do genoma do genótipo-elite RB97327

Para obtenção da porcentagem de cobertura do mapa de ligação construído para o genótipo-elite RB97327, estimou-se o tamanho do genoma de acordo com a metodologia de Hulbert et al. (1988). Para o mapa integrado este dado não foi obtido, em razão de que a metodologia não ter sido desenvolvida para se considerar informações referentes a mais de um tipo de segregação. Esta metodologia permite a estimação do número total de cM do genoma de um indivíduo, a partir de dados de mapeamento, de forma que o tamanho do genoma ( $G$ ), em cM, é estimado por:

$$G = \frac{n(n-1)}{2} \cdot \frac{2X}{K}$$

em que:  $n$  corresponde ao número de marcadores utilizados na construção do mapa;  $X$  é a distância correspondente ao valor de LOD estabelecido como limite; e  $K$  é o número de pares de marcadores com valores de LOD igual ou superior ao limite estabelecido. Os valores de  $X$  e  $K$  foram obtidos pela análise dos dados da matriz que relaciona as frações de recombinação e os valores de LOD para os pares de marcadores utilizados.

### 3.3 RESULTADOS E DISCUSSÃO

#### 3.3.1 Genotipagem

Os resultados referentes à análise da segregação Mendeliana dos marcadores DArT e SSR utilizados nas populações em estudo estão ilustrados na Tabela 3.2. Para a população resultante da autofecundação do genótipo RB97327, dos 7.680 locos DArT genotipados, 471 (6,1%) apresentaram polimorfismo, sendo que destes, 392 (83,2%) apresentaram segregação 3:1. Nessa mesma população, os vinte *primers* SSR analisados geraram um total de 115 marcas polimórficas, das quais 57 (36,7%) apresentaram segregação 3:1.

**Tabela 3.2.** Resultados das análises de segregação Mendeliana dos marcadores DArT e SSR utilizados em uma população de cana-de-açúcar oriunda de autofecundação do genótipo RB97327 e em outra resultante do cruzamento RB97327 x RB72454.

| Total de marcadores DArT e SSR analisados | População de autofecundação |     | Total | População de fecundação cruzada |     | Total |
|-------------------------------------------|-----------------------------|-----|-------|---------------------------------|-----|-------|
|                                           | DArT                        | SSR |       | DArT                            | SSR |       |
| Marcadores genotipados                    | 7680                        | 155 | 7835  | 7680                            | 155 | 7835  |
| Marcas polimórficas                       | 471                         | 115 | 586   | 850                             | 132 | 982   |
| segregação 1:1                            | -                           | -   | -     | 470                             | 31  | 501   |
| segregação 3:1                            | 392                         | 57  | 449   | 162                             | 48  | 210   |
| Marcas com outras segregações             | 79                          | 58  | 137   | 218                             | 53  | 271   |

Não se observa na literatura relatos da taxa de polimorfismo de marcadores DArT em populações provenientes de autofecundação em cana-de-açúcar. Todavia, quando comparado aos valores observados para outras classes de marcadores utilizadas nesse tipo de população, como RFLP (Grivet et al., 1996), AFLP (Hoarau et al., 2001; Andru et al., 2011) e SSR (Andru et al., 2011), observa-se que a taxa de polimorfismo dos marcadores DArT (6,1%) pode ser considerada baixa. Hoarau et al. (2001) verificaram em uma população de 295 genótipos oriundos de autofecundação da variedade R570, que de um total de 4.611 marcas, geradas a partir de 37 combinações de AFLP, 1.180 (26%) eram polimórficas. Andru et al. (2011) utilizando marcadores AFLP, SSR e TRAP em uma progênie de 300 genótipos resultantes da autofecundação da variedade LCP



85-384, obtiveram 1.111 marcas polimórficas, das quais 955 (85,9%) foram usadas para a construção de um mapa *framework* para esta variedade.

Os *primers* SSR avaliados na população resultante de autofecundação geraram média de 5,75 marcas polimórficas/par de *primer*. Estes dados diferem dos encontrados por Andru et al. (2011), que ao utilizarem 19 *primers* SSR nesse mesmo tipo de população, observaram média de 2,2 marcas polimórficas/par de *primer*. Segundo Hoarau et al. (2001), várias razões podem justificar o polimorfismo encontrado numa determinada população, incluindo diferenças no nível de heterozigosidade e números de cromossomos dos genitores que originaram a população de mapeamento.

A quantidade de marcas *single-dose* (SD) obtida pelos marcadores DArT no presente estudo assemelha-se às verificadas por Hoarau et al. (2001) e Andru et al. (2011), os quais afirmam que marcadores SD são abundantes em poliplóides e representaram entre 70-80% das marcas polimórficas verificadas em seus estudos. Embora a taxa de polimorfismo para os marcadores DArT seja relativamente baixa, é importante salientar que uma maior quantidade de informações geradas por estes marcadores foi utilizada para a obtenção do mapa referente ao genitor RB97327. Este fato reforça a eficiência dos marcadores DArT, especialmente em genomas grandes e complexos como o da cana-de-açúcar, por combinar não somente rapidez, baixo custo (R\$ 0,04 / genótipo / indivíduo) e automatização dos dados, mas também elevado rendimento e precisão das informações obtidas (Heller-Uszinska et al., 2011).

Das 586 marcas polimórficas presentes na população derivada de autofecundação, 137 (23,3%), incluindo 79 marcas DArT e 58 marcas SSR, não foram consideradas nas análises de construção do mapa de ligação, por não se comportarem como marcas SD, apresentando possivelmente com outros tipos de segregação. A análise da segregação Mendeliana permitiu incluir na análise para obtenção do mapa de ligação do genitor RB97327, 449 marcas SD com segregação 3:1, das quais 392 (87,3%) são marcas DArT e 57 (12,7%) são marcas SSR.

Para a população de fecundação cruzada, embora tenha sido observada uma taxa de polimorfismo dos marcadores DArT de 11,1%, superior à da população oriunda de autofecundação (6,1%) ainda considera-se baixo o

polimorfismo observado para estes marcadores, quando comparado aos SSR, por exemplo. De acordo com Heller-Uszinska et al. (2011), o baixo índice de polimorfismo gerado por marcadores DArT está relacionado ao número de métodos de redução da complexidade do genoma, pois isto determina a eficiência na identificação do polimorfismo. Neste caso, a utilização de mais de duas enzimas de restrição no desenvolvimento dos marcadores revela-se necessária para alcançar um número satisfatório de locos polimórficos.

De modo geral, verifica-se que tanto para marcadores DArT, quanto para marcadores SSR, é possível observar um maior número de marcas polimórficas na população de fecundação cruzada (982) do que na população de autofecundação (586). Este evento está relacionado ao fato de que ao se considerar as marcas SD presentes em uma população de cruzamento para obtenção de mapa integrado, tanto as marcas presentes exclusivamente em cada genitor (*ao x oo / oo x ao*), que apresentam segregação 1:1, quanto as marcas presentes em ambos genitores (*ao x ao*), com segregação 3:1, são considerados. Assim, a quantidade de marcas a serem incluídos nas análises é aumentada.

Entre as 711 marcas polimórficas presentes na população de fecundação cruzada submetidas à análise de ligação, 301 (42,3%) constituem marcas SD com segregação 1:1 presentes exclusivamente no genoma do genitor masculino (*oo x ao*), a variedade comercial RB72454. Este fato sugere que, para as regiões do genoma amostradas, este genitor é mais polimórfico que o genitor feminino, o genótipo RB97327.

Um nível mais elevado de polimorfismo na variedade RB72454 talvez justifique a sua utilização em grande número de cruzamentos realizados nos programas de melhoramento e a sua recomendação durante aproximadamente, 30 anos para plantio comercial. Edmé et al. (2006), ao utilizarem 193 marcadores SSR em uma população proveniente do cruzamento biparental entre os clones de *S. spontaneum* (IND81-146) e *S. officinarum* (Green German), também verificaram que um dos genitores, neste caso IND81-146, apresentava maior taxa de polimorfismo do que o genitor masculino.

Quanto aos marcadores SSR analisados na população de fecundação cruzada, os vinte *primers* genotipados originaram um total de 132 marcas polimórficas, resultando numa média de 6,6 marcas/par de *primer*. Esta

quantidade é semelhante à verificada por Garcia et al. (2006) em estudo com 52 *primers* SSR em população de cem genótipos oriundos do cruzamento biparental entre as variedades pré-comerciais SP80-180 x SP80-4966. Encontraram uma média de 6,8 marcas/par de *primer*.

O número de marcas DArT e SSR que apresentaram segregação 1:1 (72,8%) na população de fecundação cruzada também foi semelhante ao encontrado em outros estudos de mapeamento genético em cana-de-açúcar, com o mesmo tipo de população. Al-Janabi et al. (1993) encontraram 74,6%; Aitken et al. (2005), 71%; Alwala et al. (2008), 78%; e Heller-Uszinska et al. (2011), 73%. Para esses autores, os resultados sugerem que em populações provenientes de cruzamento biparental de primeira geração ( $F_1$ ), 70-75% das marcas SD apresentam segregação 1:1, enquanto 25-30% apresenta segregação 3:1.

As marcas polimórficas que não foram incluídas nas análises posteriores de construção dos mapas de ligação, por apresentarem outros tipos de segregação, foi superior na população de fecundação cruzada (27,5%) quando comparadas às identificadas na população de autofecundação (23,3%). De acordo com Andru et al. (2011), híbridos interesepecíficos em cana-de-açúcar possuem tendência de exibirem maior proporção de marcas distorcidas que populações de autofecundação, uma vez que os desvios podem estar relacionados com as dimensões significativamente diferentes do tamanho e do número de cromossomos nos genomas parentais.

### 3.3.2 Mapas genéticos de ligação

Os mapas genéticos construídos neste estudo (Apêndice A) são os primeiros desenvolvidos para cana-de-açúcar, combinando marcadores SSR e DArT. No mapa de ligação obtido a partir de uma progênie derivada de autofecundação do genótipo RB97327 os grupos de ligação (GL) estão identificados por um número seguido de “s” (“*selfing*”). Já no mapa integrado, obtido a partir da análise de progênie de fecundação cruzada (RB97327 e RB72454), os grupos de ligação estão identificados por um número seguido de “c” (“*cross*”).

### 3.3.2.1 Mapa de ligação do genitor feminino RB97327

Para a obtenção do mapa de ligação referente ao genitor feminino RB97327, um total de 449 marcas que apresentaram segregação 3:1, sendo 392 DArT e 57 SSR, foi utilizado na análise de ligação. Destas, 375 (83,5%) apresentaram-se ligadas com  $LOD \geq 3,5$  e fração de recombinação  $fr \leq 0,3$ . O mapa obtido consistiu de 95 grupos de ligação (GL) e apresentou um comprimento total de 1.217,2 cM.

No que se refere ao número de GL, resultados semelhantes foram observados por Grivet et al. (1996) e Andru et al. (2011), utilizando esse mesmo tipo de população, porém marcadores diferentes. Grivet et al. (1996) com 128 *primers* RFLP em população constituída por 77 genótipos provenientes de autofecundação da variedade R570, obtiveram um mapa com 2.008 cM e 96 GL. Andru et al. (2011) construíram um mapa para a popular variedade da Louisiana LCP 85-384, combinando 64 *primers* AFLP, 19 SSR e 12 TRAP em uma população de 300 genótipos resultantes de autofecundação. O mapa apresentou 5.617 cM e 108 GL.

Por outro lado, Hoarau et al. (2001), utilizando 37 *primers* AFLP em população de 295 genótipos, também oriundos de autofecundação da variedade R570, obtiveram um mapa de 5.849 cM e 120 GL para a mesma variedade. De acordo com os autores, a obtenção de 24 GL adicionais em relação ao trabalho similar de Grivet et al. (1996), promoveu não apenas o aumento de 3.842 cM na extensão total do mapa, mas também ganho na extensão média dos grupos. Confirmou-se assim que o significativo aumento na cobertura dos mapas é diretamente dependente da quantidade de marcadores e de genótipos utilizados, bem como do tipo de marcador escolhido para realização das análises.

Quanto à extensão total do mapa obtido neste estudo, observa-se que foi inferior à verificada por Grivet et al. (1996), Hoarau et al. (2001) e Andru et al. (2011). Este fato, provavelmente, decorre da grande quantidade de grupos de ligação com tamanho reduzido, que apresentam comprimento entre 0 cM e 15 cM. No presente estudo observou-se que, dos 95 GL formados, 84,2% apresentaram-se com menos de sete marcas, sendo 41 GL compostos por apenas duas marcas e 39 GL por três a seis marcas. Somente 15 GL apresentaram sete ou mais

marcas. O tamanho dos grupos de ligação variou de 0 a 91 cM, com média de 12,8 cM.

Até mesmo Hoarau et al. (2001), que construíram um mapa com maior extensão, revelaram dificuldades de se obter GL de grandes extensões. Em seu trabalho, 33% dos GL apresentaram comprimento inferior a 20 cM, e 63% foram menores que 40 cM. De acordo com D'Hont et al. (1998) e Piperidis et al. (2010), esses resultados decorrem do fato de que os cromossomos apresentados pelas espécies do gênero *Saccharum* e por variedades modernas são fisicamente de tamanhos bastante reduzidos. No entanto, é importante salientar que outros fatores, como os valores críticos de LOD e de fração de recombinação, também influenciam no tamanho dos GL formados.

Na maioria dos mapas obtidos para cana-de-açúcar tem sido utilizados valores de LOD  $\geq 3,0$  e fração de recombinação entre 0,25-0,45 (Da Silva et al., 1993; Al-Janabi et al., 1993; Grivet et al., 1996; Aitken et al., 2005; Andru et al., 2011). A utilização de valores de LOD muito altos na tentativa de se evitar ligações espúrias pode resultar na fragmentação dos grupos de ligação que, de outro modo, permaneceriam conectados entre si formando grupos maiores.

No mapa obtido para a variedade RB97327, quinze grupos de ligação foram formados por apenas duas marcas DArT, ligadas a uma distância de 0 cM. Em função da extensão do genoma da cana-de-açúcar e do reconhecimento da necessidade de quantidade extensiva de marcadores para obtenção de mapas saturados (Garcia et al., 2006; Heller-Uszinska et al., 2011), a ocorrência de *gaps* em mapas de ligação obtidos para esta cultura é esperada.

Dos quinze grupos de ligação constituídos por apenas duas marcas DarT ligadas à distância de 0 cM, quatro (GL17s, GL83s, GL90s e GL92s) foram formados pelas mesmas marcas DarT obtidas por diferentes métodos de redução da complexidade. A ocorrência deste tipo de marca decorre do fato de que elas foram originadas de uma mesma sonda do *array*, derivadas de diferentes métodos da redução de complexidade e que, de fato, correspondem à mesma região genômica.

De acordo com Heller-Uszinska et al. (2011), a existência desse tipo de redundância é esperada numa frequência de aproximadamente 10-15% para o genoma da cana-de-açúcar; e a obtenção das sequências de todos os

marcadores DArT permitirá reduzir não apenas suas ocorrências, como também ancorar informações genômicas de espécies geneticamente relacionadas.

Entre as 449 marcas SD que apresentaram segregação 3:1 e que foram utilizadas para realização da análise de ligação, 74 (16,5%) não se ligaram a qualquer outra marca, e, portanto, não foram incluídas nas análises para obtenção do mapa referente ao genitor RB97327. Esta quantidade foi semelhante à encontrada no mapa obtido por Andru et al. (2011), de 14,1%. Estes valores podem ser considerados baixos quando comparados aos 35,3% de marcas derivadas de sondas RFLP não ligadas verificadas por Grivet et al. (1996). No entanto, valores significativamente inferiores de marcas não ligadas (5%) foram observados por Hoarau et al. (2001).

De acordo com Garcia et al. (2006), elevadas taxas de marcas não ligadas possivelmente estão relacionadas com altos valores críticos de LOD adotados e com o tipo de população utilizada para se obter o mapa de ligação. Cruzamentos entre variedades comerciais de cana-de-açúcar originárias de híbridos interespecíficos resultam em progênies geneticamente complexas, envolvendo ocorrências de aneuploidias e dificuldades de pareamentos cromossômicos frequentes durante a meiose.

Considerando-se que o tamanho estimado do genoma do genótipo RB97327, de acordo com a metodologia proposta por Hulbert et al. (1988), seja de 10.540,9 cM, o presente mapa, cobriu aproximadamente 11,5% desse genoma. Observa-se que a cobertura do respectivo mapa é inferior à dos demais mapas obtidos para cana-de-açúcar relatados na literatura; em que para *S. spontaneum*, os mapas obtidos até o presente momento alcançaram cobertura máxima de 64,6% (Al-Janabi et al., 1993; Da Silva et al., 1993; Da Silva et al., 1995; Alwala et al., 2008), e para *S. officinarum*, a máxima cobertura foi de 69% (Alwala et al., 2008).

A baixa cobertura obtida no mapa do genitor RB97327, bem como as descritas na literatura, evidenciam a dificuldade de se obter mapas saturados para um genoma tão extenso como o da cana-de-açúcar. Além disso, sugere que o número de marcadores, bem como os tamanhos das populações utilizadas nas análises de mapeamento foi reduzido. De acordo com Heller-Uszynska et al. (2011), para se alcançar uma cobertura adequada de um genoma tão extenso de

aproximadamente 17.000 cM, seriam necessários entre 5000-7000 marcas. Diante disso, metodologias com potencial para genotipar milhares de amostras, para milhares ou dezenas de milhares de marcadores a custos de poucas dezenas de dólares, como a abordagem GbS (*Genotype-by-Sequencing*), revelam-se promissoras para estudos de mapeamento em cana-de-açúcar (Elshire et al., 2011).

### 3.3.2.2 Mapa de ligação integrado

No mapa de ligação integrado, obtido a partir dos dados de 81 genótipos provenientes do cruzamento RB97327 x RB72454, um total de 711 marcas polimórficas, incluindo 632 DArT e 79 SSR, com segregações 1:1 e 3:1, foram utilizadas na análise de ligação. Destas, 604 (84,9%) apresentaram-se ligadas com  $LOD \geq 3,5$  e fração de recombinação ( $fr$ )  $\leq 0,3$ . O mapa integrado consistiu de 136 grupos de ligação e apresentou comprimento total de 2.722,2 cM.

De forma geral, observa-se uma variação no número grupos de ligação presente nos mapas obtidos a partir de populações derivadas de cruzamentos biparentais para as variedades comerciais de cana-de-açúcar. Em geral, esta quantidade varia de 96 GL (Grivet et al., 1996) a 136 GL (Aitken et al., 2005). No entanto, quantidades menores de GL também foram descritas: Edmé et al. (2006), 10 GL e 25 GL; Al-Janabi et al. (1993), 42 GL; Alwala et al. (2008), 45 e 49 GL; Da Silva et al. (1995), 64 GL; e Raboin et al. (2006), 86 GL.

O número de grupos de ligação identificado no mapa integrado (136) foi igual ao encontrado por Aitken et al. (2005) e, se considerar que a quantidade de cromossomos esperada para as variedades de cana-de-açúcar varia de 100 a 130 cromossomos (D'Hont et al., 1996), essa quantidade está próxima do esperado. Todavia, é importante salientar que, de acordo com a metodologia utilizada neste estudo, proposta por Wu et al. (2002), os grupos de ligação mapeados constituem grupos de homólogos e não de cromossomos individuais como verificado nos demais mapas de ligação descritos na literatura.

Para Garcia et al. (2006), um mapa integrado que inclui diferentes tipos de marcadores fornece inúmeras vantagens em relação à estratégia do duplo *pseudo-testcross*. Pode-se citar a maior saturação, melhor caracterização do

polimorfismo em função de sua distribuição mais homogênea, mapeamento dos cromossomos homólogos e possibilidade de identificação dos grupos de homeologia. Além disso, a localização de locos relacionados a caracteres quantitativos (QTL) também pode ser facilitada se as informações de ambos os genitores estiverem integradas.

Quanto à extensão total do mapa obtido neste estudo, observa-se que mesmo tendo sido semelhante à do mapa obtido por Garcia et al. (2006), com 2.602,4cM e sob a mesma metodologia para obtenção de mapa integrado, o tamanho acumulado pode ser considerado baixo se comparado aos mapas de outros trabalhos. Aitken et al. (2005) obtiveram um mapa de 9.058,3 cM; Reffay et al. (2005), de 3.852cM; Raboin et al. (2006), de 3.144cM; Oliveira et al. (2007), de 6.261,1cM; e Palhares (2010), de 4.316cM.

Dos 136 GL obtidos, 110 (80,8%) apresentaram-se com menos de sete marcas, sendo 33 com apenas duas marcas e 77 com três a seis marcas. Somente 26 GL apresentaram sete ou mais marcas. Dos 33 GL constituídos por apenas duas marcas, quinze apresentaram distância de 0 cM, sendo que destes, seis (GL98c, GL116c, GL121c, GL130c, GL131c e GL136c) foram originados de uma mesma sonda do *array*, com diferentes métodos de redução de complexidade. O tamanho dos GL variou de 0 a 87,7 cM, com média de 20 cM.

Garcia et al. (2006) também observaram que, dos 131 grupos de ligação obtidos, 90 (70%) foram formados por apenas duas marcas. Para estes autores, o grande número de grupos de ligação reduzidos nos mapas de cana-de-açúcar representam partes desconexas de grupos de ligação maiores. Isso decorre da necessidade de utilização de uma maior quantidade de marcadores, para que uma maior cobertura do genoma seja alcançada. É importante considerar também, os elevados valores de LOD adotados, os quais fazem com que um grande número de marcas que apresentam a segregação requerida não sejam incluídas no mapa.

Trinta e um grupos de ligação apresentaram intervalos entre duas marcas maiores que 20 cM, embora a distância média entre estas foi de 4,5 cM. O maior grupo de ligação obtido foi GL38c, o qual apresentou 87,7 cM. O maior *gap* está localizado no GL21c, com distância igual a 42,4 cM, adjacente aos marcadores scPtb-189316, scPtb-429322 e scPth-366899 (Apêndice A).



Na literatura observa-se que a ocorrência de *gaps* nos mapas de cana-de-açúcar são comuns, como verificado por Da Silva et al. (1995), que detectaram até quatro *gaps* de 30-38 cM no mapa obtido para a variedade SES208. Garcia et al. (2006) observaram 30% dos intervalos entre dois marcadores maiores que 20 cM, e Oliveira et al. (2007), que deram continuidade a este trabalho, aumentando a quantidade e os tipos de marcadores, ainda verificaram presença de *gaps* na maioria dos grupos de ligação obtidos. De acordo com Ming et al. (1998), a presença de *gaps* nos mapas de cana-de-açúcar é esperada devido ao fato de que, na maioria dos estudos somente marcadores SD forma considerados, limitando os mapas apenas às regiões onde estes tipos de marcadores são encontrados.

Das 711 marcas utilizadas na análise de ligação, 107 (15%) não se ligaram a qualquer outra marca, e, portanto, não foram incluídas nas análises seguintes envolvidas na obtenção do mapa. Esta quantidade foi superior à porcentagem média de 5% de marcas não ligadas verificadas nos mapas obtidos por Da Silva et al. (1995), Aitken et al. (2005) e Raboin et al. (2006), e semelhante à média de 15% encontrada nos trabalhos de Al-Janabi et al. (1993) e Reffay et al. (2005).

A quantidade de marcas não ligadas observadas neste presente estudo, entretanto, pode ser considerada baixa se comparada às encontradas por Edmé et al. (2006) e Palhares et al. (2010), de 22% e 23,4%, respectivamente; ou mesmo aos 68% verificados por Garcia et al. (2006) e aos 60% de Oliveira et al. (2007) e Alwala et al. (2008). De acordo com Jannoo et al. (1999) e Oliveira et al. (2007), nos mapas genéticos obtidos de populações derivadas de cruzamentos entre variedades comerciais, os descendentes constituem híbridos interespecíficos e, portanto, apresentam estrutura genética complexa, com frequentes ocorrências de aneuploidias que, por serem caracterizadas por grande quantidade de cromossomos desemparelhados durante a meiose, dificulta a ligação entre marcas, resultando na perda de marcas informativas para o mapeamento.

### 3.3.2.3 Comparação entre os mapas do genitor RB97327 e integrado

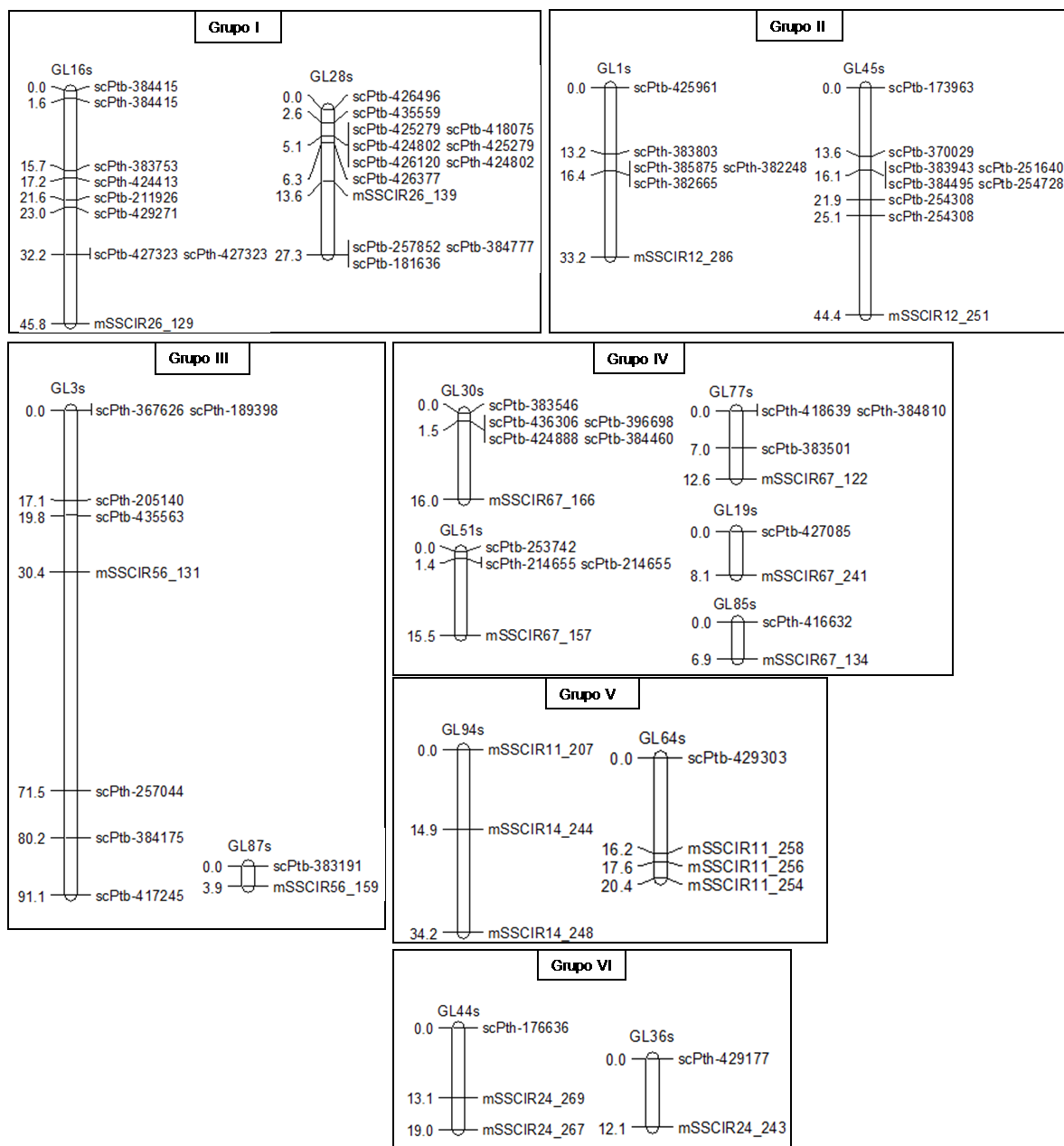
O mapa integrado obtido neste estudo apresentou 41 grupos de ligação e 1.505 cM adicionais, em relação ao mapa obtido para o genitor RB97327. Conforme comentado anteriormente, a maior saturação observada no mapa integrado é decorrente do fato de que neste, 229 marcas adicionais foram consideradas nas análises.

Dos 136 grupos de ligação constituintes do mapa integrado, observa-se em 14 a presença de marcas exclusivas em cada genitor (*ao x oo* ou *oo x ao*), unidas através de marcas comuns a ambos genitores (*ao x ao*): GL 3c, GL 11c, GL 13c, GL 17c, GL 22c, GL 28c, GL 38c, GL 40c, GL 47c, GL 54c, GL 59c, GL 64c, GL 67c e GL 69c. De acordo com Garcia et al. (2006), na obtenção de um mapa integrado é importante a presença de grande número de marcas com segregação 3:1. Isso porque nesse tipo de mapa um GL representa uma meiose que ocorreu nos pais e por isso não se espera que marcas derivadas de um genitor estejam representadas em um GL do mapa integrado junto a marcas derivadas de outro genitor; a menos que neste mesmo GL esteja presente pelo menos uma marca comum ao genoma de ambos os genitores.

De acordo com Grivet & Arruda et al. (2001), marcadores SSR ou RFLP derivados de um mesmo loco ou sonda permitem a identificação dos grupos de ligação que pertencem a um mesmo grupo de homeologia, o qual está associado ao número básico de cromossomos do gênero. A Figura 3.1 ilustra a identificação de seis marcadores SSR que amplificaram um mesmo loco no mapa obtido para o genitor RB97327, sugerindo, neste caso, a formação de seis possíveis grupos de homeologia (identificados por algarismos romanos de I a VI). Para o mapa integrado, nove possíveis grupos de homeologia foram identificados e estão ilustrados na Figura 3.2, numerados de I a IX.

Na literatura observa-se que já foram identificados entre oito (Rossi et al., 2003; Aitken et al., 2005) e quatorze (Oliveira et al., 2007) grupos de homeologia para cana-de-açúcar. No entanto, nos mapas destes estudos os grupos de homeologia foram estabelecidos separadamente para cada genitor. De acordo com Butterfield et al. (2001), as diferenças estruturais e comportamentais dos cromossomos dos genitores sugerem que os híbridos de cana-de-açúcar

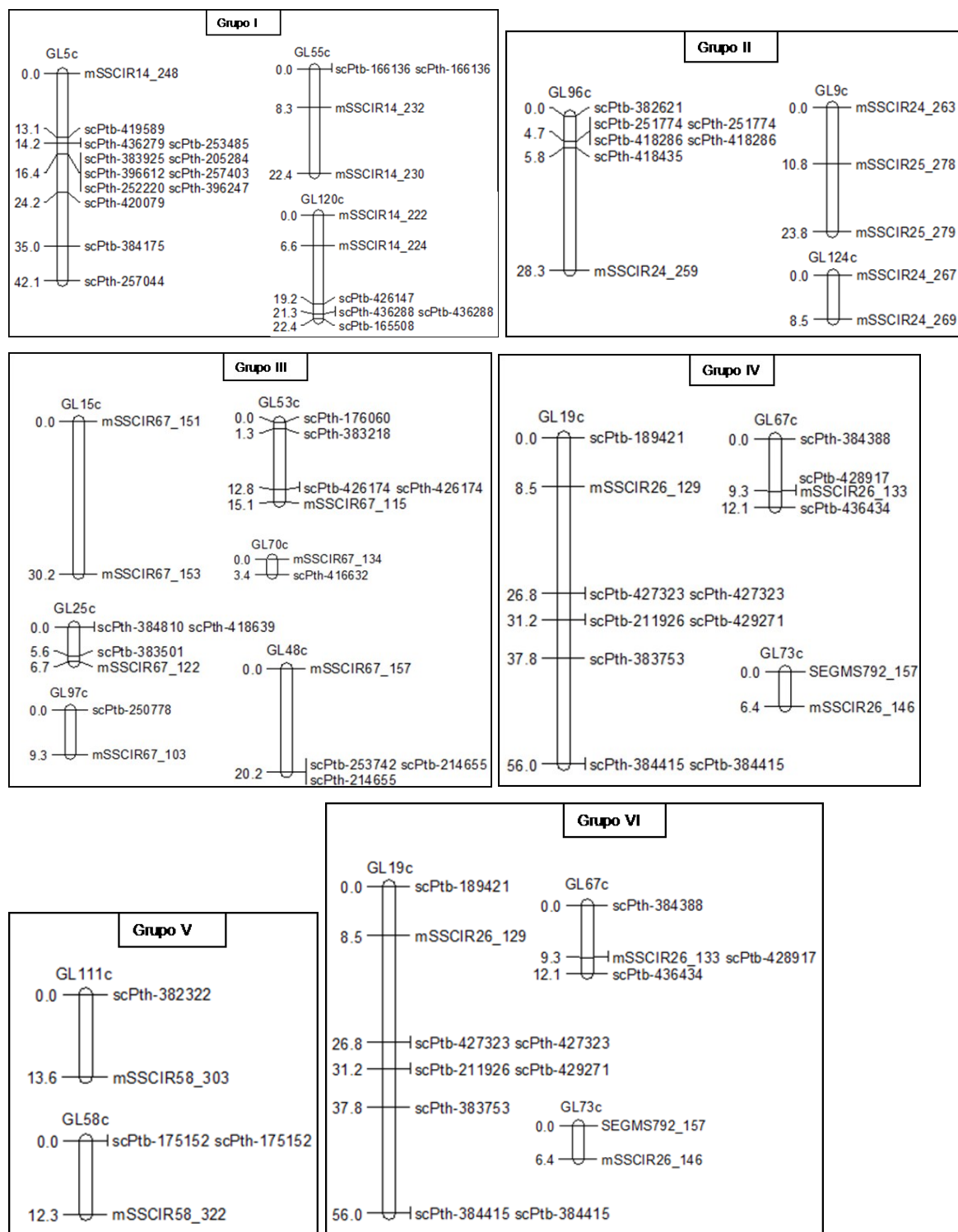
apresentem dez ou mais cromossomos. Diante disso, a quantidade de grupos de homeologia identificada no mapa integrado do presente estudo é próxima da estimada.



**Figura 3.1.** Grupos de homeologia identificados a partir dos marcadores SSR (mSSCIR) no mapa obtido para o genitor RB97327.

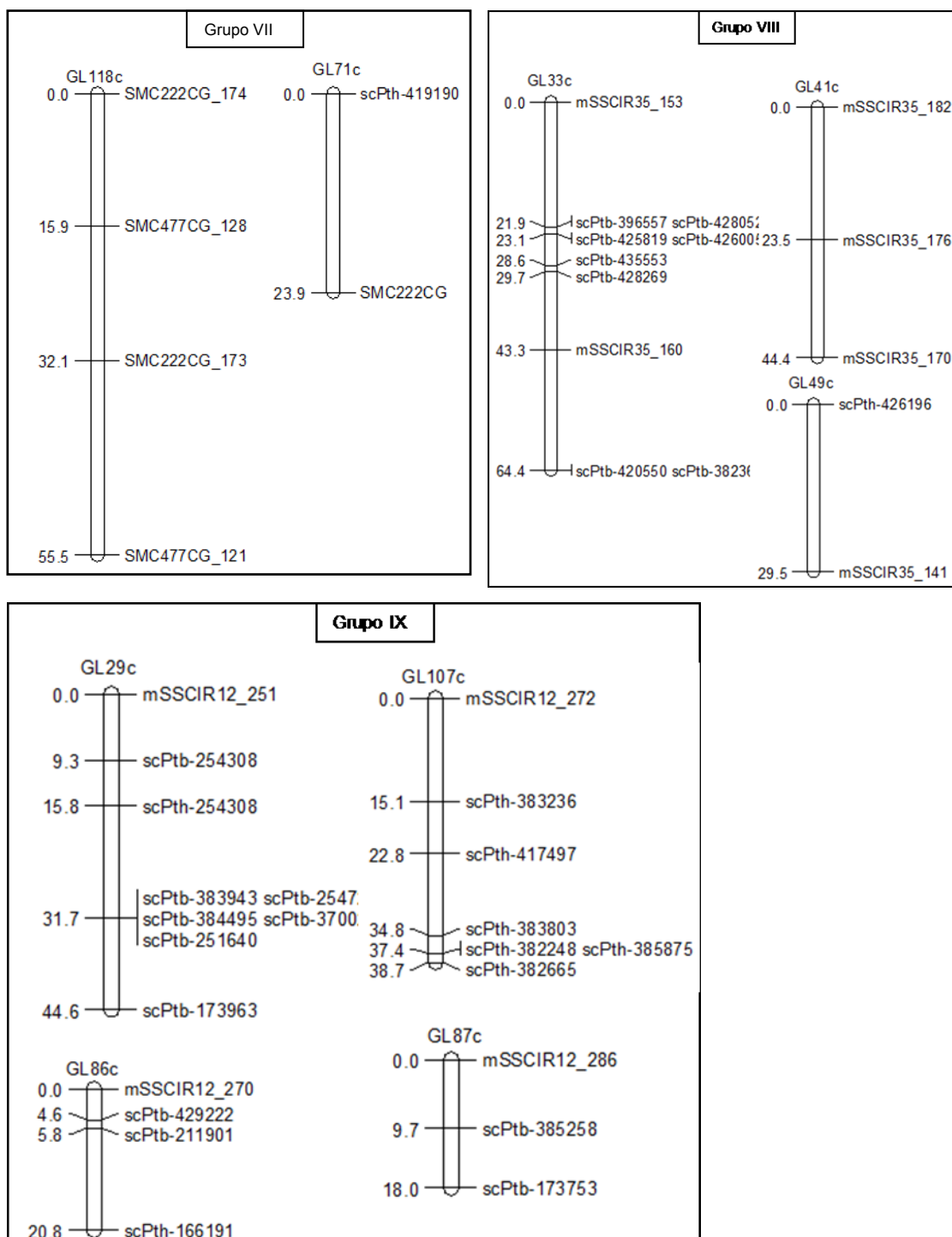
Dos 95 grupos de ligação presentes no mapa do genitor RB97327, somente 15 puderam ser atribuídos a algum grupo de homeologia (Figura 3.1). Com exceção do grupo de homeologia IV, que foi constituído por cinco grupos de ligação, os demais grupos de homeologia foram constituídos por apenas dois

grupos de ligação. Para o mapa integrado, dos 136 grupos de ligação, apenas 29 foram atribuídos a algum dos nove grupos de homeologia.



**Figura 3.2.** Grupos de homeologia identificados a partir dos marcadores SSR (mSSCIR e SMC) no mapa integrado (continua).

Figura 3.2. Continuação.



Embora tenha sido observada uma distribuição relativamente homogênea dos grupos de ligação nos grupos de homeologia identificados neste mapa, também verificou-se que a maior parte destes grupos foi formada por dois ou três grupos de ligação. Para Hoarau et al. (2001), o fato de muitos grupos de

ligação permaneceram dissociados dos grupos de homeologia, revela a dificuldade de se acessar, de maneira uniforme, todas as regiões cromossômicas do genoma da cana-de-açúcar.

A ocorrência de distribuição desigual de grupos de ligação nos grupos de homeologia foi verificada por Oliveira et al. (2007). Num mapa integrado, eles observaram que, entre 14 grupos de homeologia, a maior parte foi constituída por grupos de ligação pequenos (com apenas duas marcas).

Para Oliveira et al. (2007), a incorporação de mais marcadores SSR provavelmente permitiria ancorar grupos de homeologia que foram fragmentados, bem como a identificação de novos grupos. No entanto, é importante considerar também que, devido aos recorrentes eventos de aneuploidias no genoma da cana-de-açúcar, é esperada a ocorrência de um número desigual de cromossomos nos grupos de homeologia.

No mapa integrado, 47 marcas SSR foram identificadas em 32 grupos de ligação, e no mapa do genitor RB97327, 28 marcas SSR estiveram presentes em 21 grupos de ligação. Comparando-se as marcas SSR verificadas em ambos os mapas é possível perceber que 16 delas são comuns.

Um levantamento na literatura sobre os marcadores SSR revelou que alguns dos marcadores mSSCIR identificados neste estudo também foram mapeados em outros trabalhos: mSSCIR12, mSSCIR14, mSSCIR21, mSSCIR24, mSSCIR25, mSSCIR26 e mSSCIR35 (Aitken et al., 2005); mSSCIR14, mSSCIR26, mSSCIR55, mSSCIR67 (Rossi et al., 2003); mSSCIR21 e mSSCIR26 (Reffay et al., 2005), mSSCIR67 e mSSCIR14 (Garcia et al., 2006); e mSSCIR14 e mSSCIR35 (Edmé et al., 2006). Contudo, não se verificou nos mapas publicados na literatura os marcadores mSSCIR11, mSSCIR56 e mSSCIR58, bem como os pertencentes às séries MCSA, SMC e SEGMS, presentes em ambos os mapas de ligação aqui produzidos. Desta forma, as marcas identificadas em associação com estes marcadores podem ser consideradas inéditas.

No mapa integrado, observa-se ainda que 37 grupos de ligação compartilharam as mesmas marcas, DArT e SSR, com 38 grupos de ligação presentes no mapa obtido para o genitor RB37327 (Apêndice A). Isso representa, portanto, grupos de ligação comuns a ambos os mapas. No mapa integrado, 45

grupos de ligação apresentaram pelo menos uma marca comum a 46 grupos de ligação presentes no mapa do genitor RB37327.

A ocorrência de marcas com posições invertidas pode ser observada nos grupos que compartilharam apenas algumas marcas. Este fato também foi verificado por Aitken et al. (2005), e, de acordo com estes autores, essa inversão normalmente é observada em marcas localizadas muito próximas nos grupos de ligação, às quais normalmente é associada maior dificuldade de ordenação, devido à ocorrência de dados perdidos.

Dos 95 grupos de ligação obtidos para o genitor RB97327, somente 11 (GL 6s, GL 15s, GL 19s, GL 28s, GL 30s, GL 36s, GL 39s, GL 42s, GL 57s, GL 67s e GL 68s) não apresentaram marca comum aos grupos de ligação obtidos no mapa integrado. Portanto, para os grupos de ligação que permaneceram isolados no mapa do genitor RB97327, pode-se afirmar que, dos 1.217,2 cM apresentados pelo mapa, 119 cM (9,7%) não está representada no mapa integrado. Diante disso, mesmo que os mapas obtidos não possam ser considerados saturados, a identificação das mesmas marcas SSR previamente reportados na literatura, bem como a identificação de grande proporção (1.098,2 cM) e quantidade de grupos de ligação (84GL) compartilhados entre o mapa do genitor RB97327 e o mapa integrado, revelam a estabilidade dos mapas obtidos.

A elevada correspondência entre os grupos de ligação verificados em ambos os mapas, acrescida dos fatos de que no mapa do genitor RB97327 todas as marcas apresentaram-se ligadas em *cis*, e no mapa integrado raramente em *trans*, sugere que os genitores utilizados no presente estudo apresentam tendência de herança polissômica na meiose; mesmo que a ocorrência de determinados pareamentos específicos também tenha sido verificada. Este fato também foi constatado por Da Silva et al. (1993), Al-Janabi et al. (1993), Grivet et al. (1996) e Hoarau et al. (2001), os quais já haviam afirmado que as variedades modernas apresentam uma tendência a apresentarem herança polissômica.

Diante do exposto, pode-se afirmar que a natureza poliploide da cana-de-açúcar, associada à ocorrência de pareamentos cromossômicos aleatórios, dificulta a compreensão do comportamento cromossômico durante a meiose, fazendo com que, das espécies cultivadas, talvez esta seja a que apresente a genética mais complexa. Embora as informações obtidas até o presente momento

ainda não permitam conclusões claras e que possam ser generalizadas a todas as espécies do gênero *Saccharum*, é importante considerar que estes estudos constituem passos iniciais para o entendimento de como o genoma da cana-de-açúcar está estruturado.

### 3.4 CONCLUSÕES

- Inaugura-se com este estudo a obtenção dos primeiros mapas de ligação integrado e para o genótipo-elite da Ridesa RB97327, utilizando marcadores DArT e SSR.
- Apesar da baixa cobertura alcançada (11,5%), a identificação de marcas SSR comuns a de mapas publicados na literatura e o compartilhamento de 90,2% com o mapa integrado, revela a estabilidade do mapa obtido.
- Os marcadores SSR, embora com pequeno número de marcas polimórficas revelam-se importantes em estudos de mapeamento genético para cana-de-açúcar, pois permitem a identificação de possíveis grupos de homeologia, fundamentais para a compreensão da forma como o genoma da planta está estruturado.
- Os marcadores DArT por sua vez, revelam-se amplamente eficientes para a genotipagem, pois são capazes de gerar centenas de marcas polimórficas em apenas um ensaio, com dois métodos de redução da complexidade.



#### **4 MAPEAMENTO DE QTL EM CANA-DE-AÇÚCAR (*Saccharum* spp.) UTILIZANDO MARCADORES DArT (*Diversity Arrays Technology*) E MICROSSATÉLITES (SSR)**

##### **RESUMO**

A cana-de-açúcar está entre as culturas de maior importância comercial, sendo o Brasil o maior produtor e exportador mundial dos seus principais produtos, o açúcar e o álcool. Nos últimos anos observa-se que, devido a questões econômico-ambientais sobre estes produtos, o interesse sobre esta cultura tem incentivado o melhoramento genético para obter variedades mais produtivas em menor intervalo de tempo. Para tanto, a utilização de metodologias baseadas nas informações fornecidas por marcadores moleculares, como a identificação precoce de genótipos que apresentam QTL que contribuem para incrementar as médias dos caracteres relacionados à produção, constitui estratégia eficiente a ser implementada nos programas de melhoramento. Diante disso, no presente estudo foram identificados em mapas *framework* obtidos a partir de marcadores DArT e SSR, QTL associados a variáveis relacionadas à produção de cana-de-açúcar, utilizando-se dois tipos de populações, uma proveniente de autofecundação do genótipo-elite RB97327 e outra de cruzamento entre RB97327 e RB72454. Em ambas as populações procedeu-se à identificação de QTL a partir de análises de marcas simples e mapeamento por intervalo composto. Na população derivada de autofecundação foram identificados, pela análise de marcas simples, 63 marcadores significativamente associados às seis variáveis avaliadas: número de entrenós, número de colmos por planta, comprimento de colmos, diâmetro de colmo, peso médio de colmo e teor de sólidos solúveis (brix). Pelo mapeamento por intervalo composto, três QTL relacionados a diâmetro de colmo, comprimento de colmo e brix foram identificados. Na população proveniente do cruzamento entre RB97327 e RB72454, foram identificados pela análise de marcas simples, 60 marcadores significativamente associados às seis variáveis. Pelo mapeamento por intervalo composto identificou-se dois QTL relacionados ao diâmetro e ao comprimento de colmo.

*Palavras-chave:* caracteres quantitativos, poliploides, marcadores moleculares.

##### **ABSTRACT**

Sugarcane is among the most important commercial crops in the world. Brazil is the largest exporter and producers of its main products, sugar and alcohol. In recent years, the economic and environmental issues on these products, has encouraged the breeding efforts for the production of more productive varieties in a shorter period of time. The use of biotechnologies based on information provided by molecular markers, such as the early identification of genotypes that have

favorable QTL alleles that contribute to increase the means of traits related to production, is considered an efficient strategy to be added to traditional breeding programs. In this study we report, using framework maps obtained from DArT and SSR markers, QTL associated with variables related to sugarcane yield, using two types of populations, one derived from selfing the elite-genotype RB97327, and the other derived from the cross between the RB97327 and RB72454 genotypes. QTL analyses were performed using single marker and composite interval mapping approaches. In the selfing population, using single marker analysis, 63 markers were significantly associated with six variables: number of internodes, number of stems per plant, stem length, stem diameter, stem weight and percentage of soluble solids (°Brix). Using composite interval mapping, three QTL related to stem diameter and length of stem were identified. In the population derived from outcrossing, using single marker analysis, 60 markers were significantly associated with the six variables. Composite interval mapping identified two QTL related to the diameter and length of stem.

*Key words:* quantitative characters, polyploids, molecular markers.

#### 4.1 INTRODUÇÃO

A cana-de-açúcar é cultivada comercialmente em mais de oitenta países. O Brasil, com 20%, lidera a produção mundial desta cultura e exportação de seus principais subprodutos, o açúcar e o álcool. O crescente interesse pela cana-de-açúcar como matéria-prima para geração de bioenergia tem exigido estratégias que impulsionem a sua produção, tanto pelo aumento da área plantada, quanto pelo incremento em produtividade. Neste contexto, o melhoramento genético apresenta importante relevância por permitir a seleção de variedades com caracteres agrônômicos e industriais favoráveis (Cesnick & Mioque, 2004).

Atualmente, o melhoramento da cana-de-açúcar baseia-se sobretudo na seleção fenotípica de características de importância comercial e agrônômica, presentes em populações segregantes oriundas de cruzamentos biparentais (Matsuoka et al., 2005). Como em geral as características fenotípicas são de natureza quantitativa e as populações nas quais estas características são avaliadas apresentam elevada variabilidade, a utilização de marcadores moleculares em estudos relacionados à forma com que tais caracteres estão estruturados pode beneficiar os programas de melhoramento, através da seleção precoce e eficiente dos genótipos superiores (Pastina et al., 2012).

Os primeiros mapas de QTL para cana-de-açúcar surgiram,

aproximadamente, na mesma época em que foram publicados os primeiros mapas de ligação. Nesses mapas foram identificados QTL relacionados aos caracteres envolvidos em produtividade de açúcar e seus componentes, a saber: teor de açúcar (Pol), brix, toneladas de açúcar por hectare, toneladas de cana por hectare, teor de fibras, peso, número, comprimento e diâmetro de colmos (Sills et al., 1995; Guimarães et al., 1999). Todavia, os mapas de QTL inicialmente publicados foram construídos a partir de mapas de ligação obtidos por métodos convencionais, desenvolvidas para diploides, e a partir de análises baseadas em um único marcador e mapeamento por intervalo.

Notadamente esses métodos resultaram em mapas de QTL pouco saturados e com pequena cobertura do genoma. Além disso, Pastina et al. (2012) ressaltam que é importante considerar que a maior parte desses mapas foi obtida a partir de informações oriundas de populações provenientes de cruzamentos biparentais, sendo poucos os trabalhos nos quais tenham sido utilizadas progênies obtidas por autofecundação. Também sabe-se que as informações obtidas por este tipo de população são fundamentais para a compreensão da estruturação do genoma da cana-de-açúcar. Diante disso, o presente estudo teve por objetivo mapear QTL em uma população derivada da autofecundação de um genótipo-elite, RB97327, bem como em uma população proveniente do cruzamento entre RB97327 x RB72454.

## 4.2 MATERIAL E MÉTODOS

### 4.2.1 Material vegetal

Dados moleculares e fenotípicos foram obtidos a partir de uma população constituída por 91 genótipos, derivada da autofecundação do genótipo-elite RB97327, e a partir de outra, constituída por 81 genótipos, proveniente do cruzamento entre RB97327 e a variedade comercial RB72454, constituída por 81 genótipos. Este cruzamento representa um experimento típico do Programa de Melhoramento da Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa) que, por apresentar uma progênie com grande número de indivíduos e com média de produtividade e variabilidade elevadas, foi

escolhido para este estudo.

#### 4.2.2 Dados moleculares

Para obtenção do DNA genômico total, foi realizada a coleta de duas gemas laterais dos colmos de cada genótipo. A extração de DNA foi feita seguindo o protocolo de Al-Janabi et al. (1999), com adaptações para microtubos de 1,5 mL. No total realizou-se a extração de DNA de 172 plântulas mais os genitores envolvidos no estudo. A quantificação do DNA extraído foi realizada através do espectrofotômetro Qubit (*Invitrogen*®).

Para genotipagem dos marcadores SSR, vinte pares de *primers* SSR, obtidos a partir de Cordeiro et al. (2000; 2001) e Parida et al. (2009), polimórficos entre os genitores, foram utilizados de forma a permitir a montagem de cinco painéis triplex e um duplex. Para a realização das reações de amplificação dos locos SSR, o DNA dos genótipos em estudo foi diluído para a concentração de 3 ng/μL. As reações foram feitas utilizando o kit Multiplex PCR Kit (Qiagen), seguindo as recomendações do fabricante, mas reduzindo-se o volume final de cada reação para 5 μL; assim tem-se 1/10 da quantidade de cada componente. Para a realização das reações foi utilizado termociclador GeneAmp PCR System 3700 (*Applied Biosystems*), o qual foi programado para realizar uma etapa inicial de 95°C/15 min., seguido de uma PCR *touchdown* com temperatura inicial de 60°C, para anelamento dos *primers*, durante os 5 primeiros ciclos, e redução de 0,5°C por ciclo nos trinta ciclos seguintes.

Os produtos da PCR foram revelados no analisador automático ABI PRISM 3100 (*Life Technologies*), a partir de uma mistura contendo: 1,0 μL do produto da reação de PCR, 8,85 μL de formamida Hi-Di (*Life Technologies*) e 0,15 μL de solução com os fragmentos de tamanho padrão (GS500-ROX, *Life Technologies*). Os cromatogramas resultantes foram analisados utilizando-se o *software* GenneMapper® 3.0 (*Life Technologies*).

A genotipagem dos marcadores SSR, tanto na população de autofecundação, quanto na população oriunda de cruzamento, foi realizada seguindo a metodologia proposta por Wu et al. (1992), segundo a qual cada fragmento deve ser considerado como um marcador, codificando-se a sua

presença com “1” e a sua ausência com “0”. As marcas SSR foram identificadas pelo nome do marcador, seguido do número indicativo do tamanho do fragmento gerado.

Os dados referentes aos marcadores DArT foram gerados pela empresa *Diversity Arrays Technology PTy Ltd* (DArT P/L), localizada na cidade de Canberra, Austrália. Para isso, o DNA genômico dos 174 genótipos foi diluído para a concentração de  $60 \text{ ng } \mu\text{L}^{-1}$ , e 20  $\mu\text{L}$  deste material foi distribuído em duas placas de 96 poços, as quais foram lacradas e enviadas, à temperatura ambiente, à prestadora dos serviços.

Na obtenção das marcas DArT foram usadas duas enzimas de restrição para redução da complexidade genômica (scPth e scPtb), ambas indicadas na identificação dos marcadores. A partir deste procedimento, obtiveram-se os dados de genotipagem de 7.680 marcadores, referentes aos 172 genótipos em estudo, os quais também foram codificados com “1” para presença e “0” para ausência.

#### 4.2.3 Obtenção dos mapas *framework*

Os dados referentes aos marcadores SSR foram reunidos aos de marcadores DArT, de modo que todas as análises envolvidas na obtenção dos mapas *framework* fossem realizadas separadamente para cada uma das populações (de autofecundação e de fecundação cruzada). Inicialmente para seleção dos marcadores *single-dose* (SD) 1:1 e 3:1, realizou-se a análise da segregação Mendeliana. As frequências observadas foram submetidas ao teste de aderência de qui-quadrado, sob a hipótese  $H_0$  de distorção desta segregação. Para correção da taxa de erro tipo I, foi utilizado o critério FDR (*False Discovery Rate*) com  $\alpha=0,01$ .

Para obtenção do mapa *framework* referente ao genitor feminino RB97327, foram considerados apenas os marcadores *single-dose* (SD) com segregação 3:1; enquanto para o mapa integrado, obtido a partir da população derivada do cruzamento biparental foram considerados marcadores SD que apresentavam segregação 3:1 (*ao x ao*), bem como aqueles com segregação 1:1 (*ao x oo* e *oo x ao*). Os marcadores com outros tipos de segregação foram

desconsiderados nas análises seguintes, envolvidas na construção dos mapas *framework*.

A construção de ambos os mapas foi executada através do pacote OneMap (Margarido et al., 2007), disponível no *software* R (*R Development Core Team*, 2012), com base em abordagem multiponto. Para a determinação dos grupos de ligação (GL), uma análise de dois pontos foi inicialmente realizada, considerando-se  $LOD \geq 3,5$  e fração de recombinação  $fr \leq 0,3$ . Para os grupos de ligação que apresentaram seis ou menos marcas, todas as possíveis ordens foram comparadas, e a escolha da melhor ordem foi feita quando esta revelou-se, no mínimo, 100 vezes mais verossímil do que a segunda melhor ordem ( $\Delta LOD=2$ ). Para os GL que apresentaram mais de seis marcadores, o algoritmo “*order*” foi iniciado com cinco marcadores, os quais foram ordenados através da comparação de todas as possíveis ordens. Posteriormente, os marcadores restantes foram sequencialmente incluídos com base num critério de  $\Delta LOD=2$ . Com o intuito de se verificar a ocorrência de inversões locais, o comando “*ripple*” foi executado.

Em ambos os mapas, os valores das frações de recombinação entre marcas foram convertidos em centiMorgans (cM) utilizando a função de mapeamento proposta por Kosambi (1944). Os desenhos dos grupos de ligação foram obtidos no *software* MapChart 2.2 (Voorrips, 2002).

#### 4.2.4 Fenotipagem

As populações de mapeamento utilizadas no estudo fazem parte de uma geração  $F_1$  constituída por 792 genótipos. Para a obtenção dos dados fenotípicos, os genótipos foram plantados em novembro de 2010, no campo experimental da Usina Centroálcool, localizada no município de Inhumas-GO ( $16^{\circ}20'50''S$ ,  $49^{\circ}29'2''W$ ), vinculada ao PMGCA-UFG.

O experimento apresentou área total de 978,75 m<sup>2</sup> e o delineamento utilizado foi o de blocos aumentados (Federer, 1956). O experimento foi constituído por nove linhas de 67,5 m cada, tendo sido cada uma destas subdividida em 15 blocos. Cada bloco foi constituído por oito parcelas (touceiras), das quais duas eram testemunhas comuns a todos os blocos. As variedades

comerciais utilizadas como testemunhas foram a RB99395 e RB98710. O de cada parcela, bem como o utilizado entre as plantas na linha foi de 0,5m, apresentando, portanto, cada bloco, 4,5 m de comprimento.

Em julho de 2012, os colmos pertencentes aos 792 genótipos foram identificados, coletados, amarrados em feixes e transportados para a Escola de Agronomia da Universidade Federal de Goiás, onde foi realizada a avaliação fenotípica dos seguintes caracteres: peso total da parcela em kg (PESO), número de colmos por parcela (NCOLM), diâmetro médio do colmo em mm (DIAM), comprimento do colmo em m (COMP), número de entrenós (NENT) e brix em °Brix (BRIX). Com exceção dos dois primeiros caracteres, os demais foram avaliados tomando-se seis colmos ao acaso em cada touceira.

Para a obtenção do peso total da parcela, os colmos cortados foram amarrados em feixes e pesados em balança digital. NCOL foi obtido pela contagem direta do número de colmos presentes em cada feixe; DIAM foi obtido por paquímetro, tendo sido a medição realizada no centro do entrenó localizado na porção mediana do colmo; COMP foi obtido com trena, considerando-se a medida da base do colmo até o último entrenó; e NENT pela contagem direta do número de internódios existentes desde a base do colmo até à altura da emissão das últimas folhas. Para a avaliação do teor de sólidos solúveis totais presentes nos colmos (°BRIX), estes foram cortados em sua porção mediana, no centro de um entrenó, e esmagados para obtenção de uma amostra de “suco”. A leitura do valor referente ao brix, diretamente relacionada à quantidade de sacarose nos colmos, foi realizada em refratômetro analógico portátil.

#### **4.2.5 Análises estatísticas dos dados fenotípicos**

Para cada variável fenotípica avaliada (PESO, NCOLM, DIAM, COMP, NENT e BRIX), foram estimados os parâmetros estatísticos descritivos relacionados aos valores mínimos e máximos, média, variância e coeficiente de variação. Posteriormente, as variáveis avaliadas em nível de parcelas foram submetidas à análise estatística, utilizando-se o seguinte modelo linear, adaptado de Duarte (2000):

$$Y_{ij} = \mu + g_i + b_j + e_{ij}$$

em que:

$Y_{ij}$ : corresponde ao valor observado da variável na parcela correspondente ao genótipo  $i$ , localizada no bloco  $j$ ;

$\mu$ : é a constante comum a todas as observações;

$g_i$ : é o efeito do genótipo  $i$ , assumido fixo e nulo se o genótipo for uma testemunha, ou aleatório com distribuição  $N(0, \sigma_g^2)$  independente, se o genótipo for um tratamento adicional;

$b_j$ : é o efeito fixo do bloco  $j$ ;

$e_{ij}$  é o efeito do erro aleatório associado à observação da parcela correspondente ao genótipo  $i$ , no bloco  $j$ .

Com base neste modelo, os componentes de variância de interesse foram estimados e os melhores preditores lineares não-viesados empíricos (*eBLUPs*) foram obtidos para cada genótipo. Todas as análises foram executadas no programa de análises estatísticas R (R Development Core Team, 2012).

#### 4.2.6 Análises visando à detecção de QTL

A detecção de QTL foi realizada por análises de marcas simples e de mapeamento por intervalo composto (MIC) utilizando o *software* QTL Catographer versão 2.5 (Basten et al., 2005) para realização das. Para execução destas análises dois arquivos foram construídos para cada população (de autofecundação e de fecundação cruzada). Um deles continha os dados moleculares e os valores dos *eBLUPs* de cada genótipo (referentes às variáveis fenotípicas PESO, NCOLM, DIAM, NENT, COMP e BRIX). O outro, as informações referentes aos grupos de ligação, com seus respectivos marcadores e distâncias, integrantes dos mapas *framework*. Visando agregar informações sobre os marcadores, aqueles que permaneceram não ligados após as análises para obtenção dos mapas *framework* também foram considerados nas análises de QTL. Cada um destes marcadores foi considerado constituindo um grupo de ligação.

É importante considerar que para a população derivada de autofecundação, embora todos os marcadores cuja segregação 3:1 tenham sido considerados nas análises visando a detecção de QTL, na população proveniente



de fecundação cruzada somente os marcadores com segregação 1:1 (*ao x oo* e *oo x ao*) foram considerados. No modelo utilizado para estas análises, os marcadores foram considerados como de efeitos fixos e os *eBLUPs* como funções lineares dos genótipos de cada um dos marcadores analisados.

Na análise de marcas simples (Lynch & Walsh, 1998), os dados foram ajustados para o seguinte modelo de regressão linear simples:

$$Y_i = b_0 + b_1 x_i + e_i$$

em que:

$Y_i$ : representa o valor do *eBLUP* do genótipo ou indivíduo  $i$  para a variável;

$b_0$ : é uma constante;

$b_1$ : representa o efeito associado à presença do alelo sob análise;

$x_i$ : é a variável indicadora da presença do alelo sob análise no indivíduo  $i$ ;

$e_i$ : é o efeito do erro aleatório associado à observação realizada no indivíduo  $i$ , com média igual a “0” e variância igual a  $\sigma^2$ .

Na análise de mapeamento por intervalo composto (Zeng et al., 1994), utilizou-se o modelo 6 (“*Standard Model*”) do QTL Cartographer. A hipótese de presença de QTL foi avaliada a cada 0,5 cM, com uma distância de 5 cM para controlar a interferência de múltiplos QTL por meio de modelos de regressão “*Backward* e *Forward*”. A localização mais provável do QTL foi estimada pela razão de verossimilhança e o limiar para se aceitar um QTL foi obtido a partir do valor de LOD=2,5, tendo sido o pico do LOD score adotado como estimativa da posição do QTL no mapa de ligação.

## 4.3 RESULTADOS E DISCUSSÃO

### 4.3.1 Análises fenotípicas

Um resumo das estatísticas descritivas relacionadas a cada um dos caracteres avaliados (PESO, NCOLM, COMP, DIAM, NENT e BRIX) nas populações derivadas de autofecundação e de fecundação cruzada (B) é apresentado na Tabela 4.1. Foram estimados os valores mínimo (Mín.) e máximo (Máx.), a média, a variância ( $s^2$ ) e o coeficiente de variação (CV) para cada um deles. Dentre as seis variáveis, PESO e NCOLM apresentaram variabilidade

fenotípica superior na população proveniente de fecundação cruzada (B) em comparação à população de autofecundação (A) (Tabela 4.1, Figura 4.1). A superioridade dos valores referentes às variâncias para a população de fecundação cruzada corrobora com este fato. Diante disso, observa-se que a fecundação cruzada é fundamental para a geração de variabilidade nos locos responsáveis por controlar estas características.

**Tabela 4.1.** Valores mínimos (Mín.), máximos (Máx.) e médio, com seus respectivos valores de variância ( $s^2$ ) e coeficientes de variação fenotípica (CV), das seis variáveis fenotípicas avaliadas ao nível de parcelas, na progênie derivada da autofecundação do genótipo-elite RB97327 (A) e na progênie oriunda do cruzamento RB97327 e RB72454 (B).

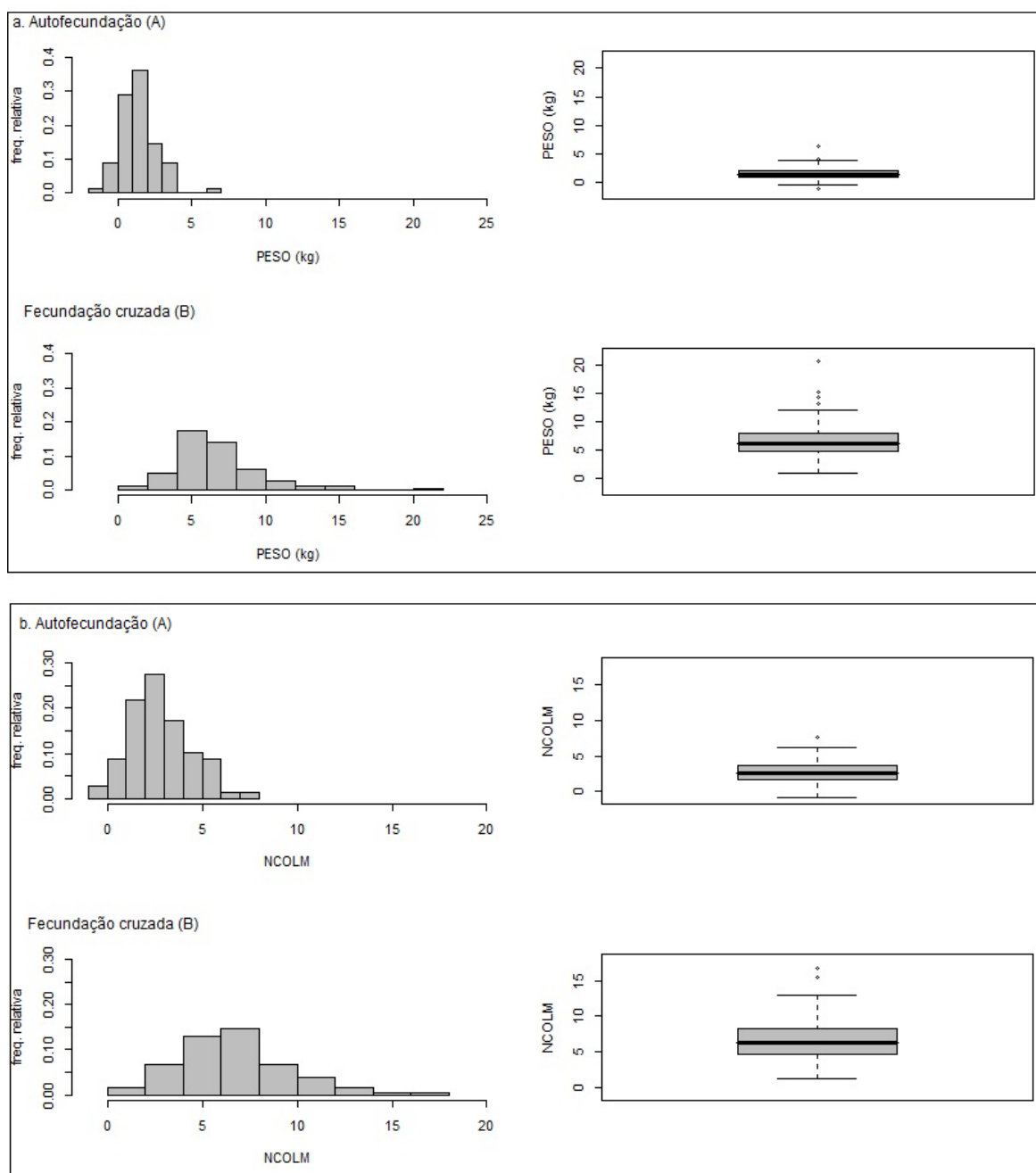
| Variável <sup>1</sup> | Mín.  |       | Máx   |       | Média <sup>2</sup> |        | $s^2$ |      | CV(%) |       |
|-----------------------|-------|-------|-------|-------|--------------------|--------|-------|------|-------|-------|
|                       | A     | B     | A     | B     | A                  | B      | A     | B    | A     | B     |
| PESO                  | -1,26 | 0,93  | 6,33  | 20,62 | 1,42a              | 6,76b  | 1,53  | 9,38 | 86,61 | 45,26 |
| NCOLM                 | -0,84 | 1,33  | 7,56  | 16,67 | 2,79a              | 6,68b  | 2,62  | 8,90 | 57,70 | 44,61 |
| COMP                  | 1,21  | 1,63  | 2,08  | 2,52  | 1,66a              | 2,03b  | 0,03  | 0,02 | 10,24 | 6,89  |
| DIAM                  | 19,38 | 23,65 | 21,55 | 26,40 | 20,46a             | 25,28b | 0,22  | 0,26 | 2,24  | 1,97  |
| NENT                  | 26,75 | 25,79 | 37,96 | 38,27 | 32,06a             | 31,16b | 6,43  | 4,22 | 7,89  | 6,57  |
| BRIX                  | 15,83 | 15,60 | 22,13 | 22,05 | 19,73a             | 19,96a | 1,45  | 1,27 | 6,08  | 5,61  |

<sup>1</sup> PESO: peso médio da parcela em kg; NCOLM: número de colmos da parcela; COMP: comprimento do colmo em m; DIAM: diâmetro da parcela em mm; NENT: número de entrenós/colmo e BRIX: teor de sólidos solúveis em °brix.

<sup>2</sup> a e b: médias na mesma linha com diferenças significativas ( $p < 0,05$ ) pelo teste de Tukey.

De acordo com McIntyre & Jacson (2001), plantas de cana-de-açúcar obtidas por autofecundação apresentam-se menos vigorosas do que as progênies derivadas de cruzamentos, indicando que, de fato, o desempenho dos genótipos, em relação a algumas variáveis pode reduzir consideravelmente. Tal fato pode ser observado neste estudo, no qual a autofecundação promoveu, para as variáveis PESO e NCOLM, efeito endogâmico, reduzindo consideravelmente a variabilidade fenotípica (Figura 4.1a e 4.1b).

Os maiores valores dos coeficientes de variação fenotípica na população derivada de autofecundação revelam o efeito da depressão por endogamia nesta população. Observa-se também que as médias referentes à população proveniente de autofecundação foram significativamente inferiores ( $p < 0,05$ ) às médias da população derivada de fecundação (Tabela 4.1).



**Figura 4.1.** Histogramas e *boxplots* referentes ao peso (PESO: peso médio da parcela em kg) (a) e número de colmos (NCOLM) (b), avaliados na população oriunda da autofecundação (A) de RB97327 e na população proveniente da fecundação cruzada (B) entre RB97327 x RB72454.

Para as variáveis PESO e NCOLM, observa-se ainda que em ambas as populações se observaram genótipos com médias superiores à das variedades comerciais RB99395 e RB98710, usadas como testemunhas neste estudo, as quais apresentaram médias de 4,74 kg/planta e 3,31 colmos/planta, para PESO e NCOLM, respectivamente. Entretanto, a população de fecundação cruzada foi a

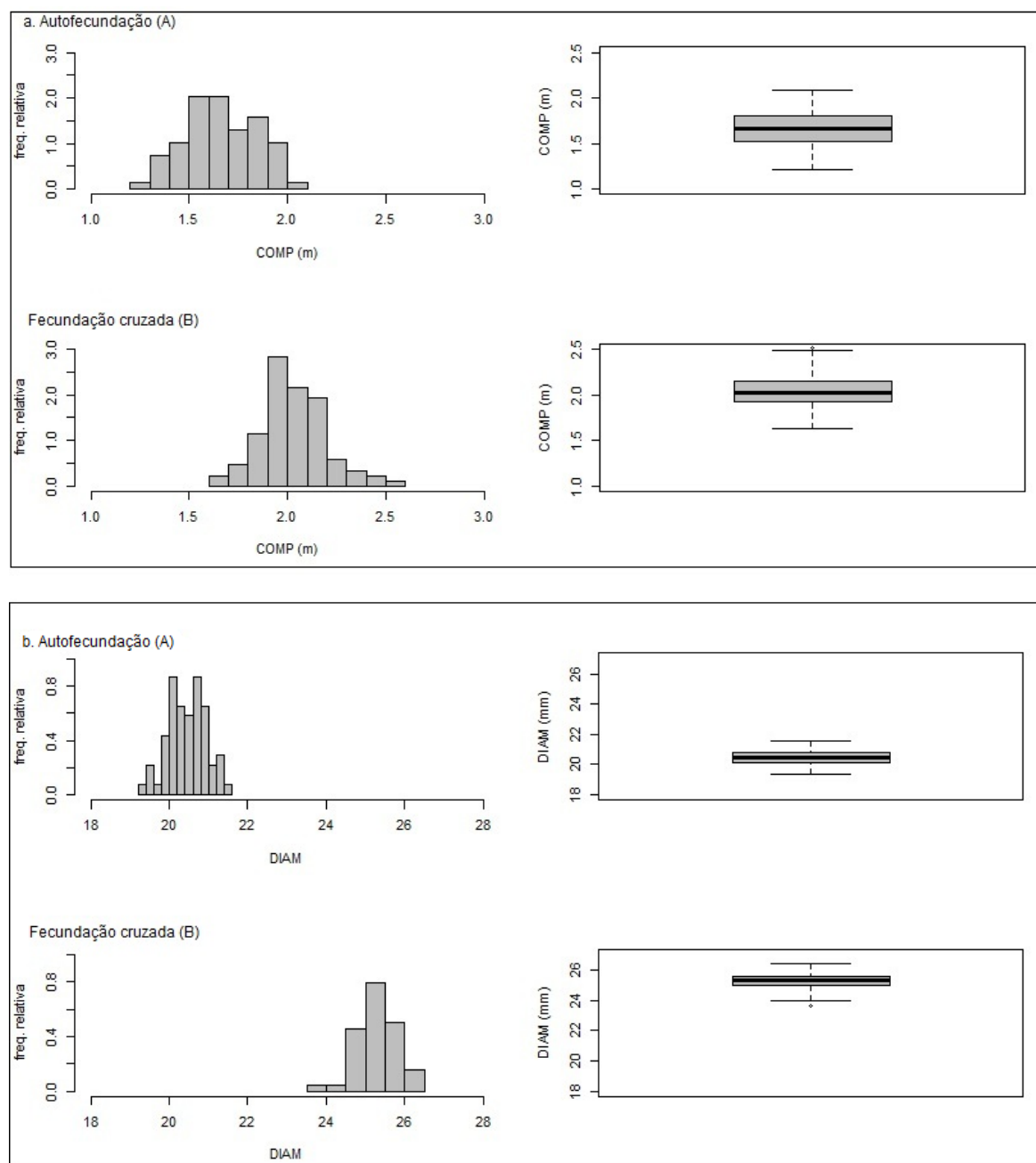
que contemplou os genótipos superiores para estas variáveis, bem como o genótipo de valor máximo. Quanto a variável PESO destacou-se o genótipo 235, com média de 20,62 kg/planta; e para a variável NCOLM, o genótipo 53, com média de 16,67 colmos/planta.

Quanto ao COMP é possível observar que a amplitude apresenta a mesma magnitude em ambas as populações, A e B (Figura 4.2a). No entanto, nos genótipos provenientes de autofecundação observa-se média significativamente inferior ( $p < 0,05$ ), de 1,66 m, em relação à população de fecundação cruzada, de 2,03 m (Tabela 4.1). Para esta variável, também se observa genótipos com médias superiores à das testemunhas (1,81 m) em ambas as populações, embora o valor máximo encontrado para esta variável tenha sido o do genótipo 168 (2,52 m), oriundo de fecundação cruzada (Tabela 4.1).

No que se refere à variável DIAM, as populações A e B não diferiram em relação à magnitude da amplitude. A média estimada da população de autofecundação foi significativamente inferior (20,46 mm) à da população de fecundação cruzada (25,28 mm), revelando também efeito de endogamia para este caráter (Tabela 4.1). Especificamente para esta variável observa-se que o efeito da autofecundação promoveu o distanciamento dos intervalos fenotípicos de cada população sem sobreposição de valores (Figura 4.2b). Além disso, em ambas as populações, todos os genótipos apresentaram valores médios inferiores à média das testemunhas (26,99 mm), tendo sido o genótipo 188 o que apresentou a maior média (26,40 mm). Logo, pode-se afirmar que, tanto a população de fecundação cruzada, quanto a população de autofecundação carecem de genótipos superiores para esta variável.

Diferentemente dos demais caracteres, observa-se que para as variáveis NENT e BRIX, ambas populações apresentaram, separadamente, intervalos similares para estas variáveis (Figura 4.3). Para número de entrenós (NENT) a população de autofecundação apresentou média ligeiramente superior em relação à população de fecundação cruzada, e para BRIX, não houve diferença entre as médias (Tabela 4.1). Resultados semelhantes quanto à variável brix foram reportados na literatura por Hogarth (1980), Ferreira et al. (2005) e Silva & Gonçalves (2011), os quais sugerem ser esta uma variável cuja depressão endogâmica apresenta menor magnitude em populações derivadas de

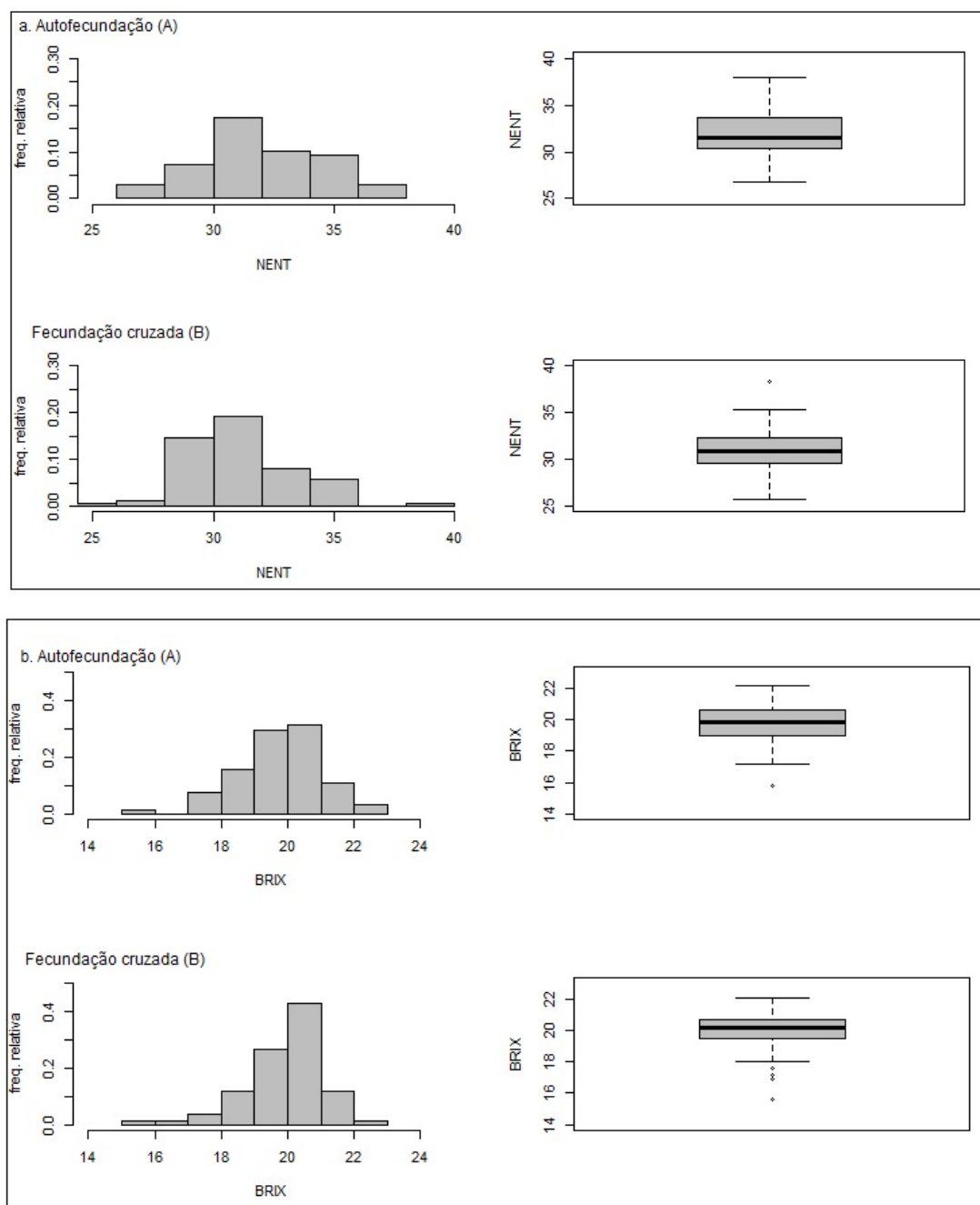
autofecundação. Isso indica que, provavelmente, a variação genética para brix, em sua maior parte, é de efeito aditivo.



**Figura 4.2.** Histogramas e *boxplots* referentes ao comprimento (COMP: comprimento do colmo em m) (a) e diâmetro (DIAM: diâmetro do colmo em mm) (b), avaliados na população oriunda da autofecundação (A) de RB97327 e na população proveniente da fecundação cruzada (B) entre RB97327 e RB72454.

Do ponto de vista do melhoramento convencional, caracteres que apresentam maior parte da variação devida a efeitos aditivos podem ser favorecidas por seleção recorrente. Todavia, os resultados observados no

presente estudo para a variável brix sugerem a necessidade de precauções adicionais ao longo das etapas de seleção envolvidas no melhoramento convencional, quando o objetivo for realizar seleção entre genótipos derivados de cruzamentos superiores para vários caracteres.



**Figura 4.3.** Histogramas e *boxplots* referentes ao número de entrenós (NENT: número de entrenós) (a) e brix (BRIX: teor de sólidos solúveis solúveis em °brix) (b), avaliados na população oriunda da autofecundação (A) de RB97327 e na população proveniente da fecundação cruzada (B) entre RB97327 e RB72454.

Isso porque, embora genótipos derivados de autofecundação demonstraram não apresentar depressão por endogamia para esta variável, observa-se que para as demais a redução do vigor foi significativa.

De forma geral, os resultados revelam que a população oriunda de fecundação cruzada apresentou tendência de maior variabilidade que a população derivada de autofecundação; e, também incluiu uma maior quantidade de genótipos com valores médios superiores aos dos genótipos comerciais utilizados como testemunhas. Todavia, quanto à variável DIAM é importante considerar que a média estimada para a população de fecundação cruzada foi ligeiramente inferior à das testemunhas e nenhum genótipo desta população apresentou valor médio superior aqueles de cada testemunha, separadamente. Diante do exposto, sugere-se que se o objetivo do programa de melhoramento for utilizar uma das duas populações em estudo para a seleção de genótipos, deve-se priorizar a população de fecundação cruzada, devido à maior variabilidade e ao melhor desempenho destes em relação aos genótipos da população de autofecundação.

A abordagem de que genótipos superiores não podem ser obtidos através de autofecundação de variedades de cana-de-açúcar tem sido controversa. Silva & Gonçalves (2011) avaliaram genótipos oriundos da autofecundação de 26 variedades, quanto a brix, massa e número de colmos, comprimento e diâmetro de colmos e observaram que a autofecundação produziu depressão endogâmica somente para massa e comprimento de colmos. Por esse motivo, os autores afirmam ser possível obter clones de cana-de-açúcar derivados de autofecundação com considerável vigor híbrido. Assim recomendam a utilização destes em cruzamentos ou, nos casos em que os clones obtidos apresentarem elevados níveis de heterozigosidade, para realização de posteriores autofecundações com o objetivo de se avaliar a intensidade do efeito endogâmico em diferentes genótipos, ao longo de sucessivas gerações de autofecundações.

De acordo com Ferreira et al. (2005), a intensidade da depressão endogâmica não depende somente da característica avaliada, mas também do genótipo autofecundado. Em seu estudo, constataram, por exemplo, que quando as variedades RB72454 e SP811763 foram autofecundadas, não se observou a

mesma intensidade de depressão endogâmica em comparação a outras variedades. Assim, sustentaram que maior frequência de alelos favoráveis tende a apresentar menor depressão endogâmica. Todavia, os autores sugerem a utilização de linhagens puras derivadas de autofecundação em cana-de-açúcar como alternativa para a formação de populações-base em programas de seleção recorrente recíproca; pois, conforme afirmado também por Hogarth (1980), a utilização de autofecundação no melhoramento da cana-de-açúcar é mais importante para expressar a capacidade específica de combinação de um dado genótipo do que para expressar a sua capacidade geral de combinação.

#### 4.3.2 Mapas *framework*

Um total de 449 marcadores com segregação 3:1, sendo 392 DArT e 57 SSR, foi utilizado nas análises envolvidas na obtenção do mapa *framework* referente ao genitor feminino RB97327 (Apêndice B). Destes, 375 (83,5%) apresentaram ligação com  $\text{LOD} \geq 3,5$  e fração de recombinação ( $fr$ )  $\leq 0,3$ . O mapa obtido consistiu de 95 grupos de ligação e apresentou comprimento total de 1.019,69 cM. O tamanho dos grupos de ligação variou de 0 cM a 40,58 cM, com média de 10,73 cM.

Nesse mapa, 16 grupos de ligação foram formados por apenas duas marcas, ligadas a uma distância de 0 cM. Das 449 marcas SD que apresentaram segregação 3:1 e que foram utilizadas para a análise de ligação, 74 (16,5%) não se ligaram a qualquer outra marca, e, portanto, não foram incluídas nas análises de obtenção do mapa. Cento e cinquenta e uma marcas não puderam ser ordenadas nos grupos de ligação. O mapa *framework* referente ao genitor feminino RB97327 apresentou 224 marcas.

Para obtenção do mapa *framework* integrado (Apêndice C), obtido a partir dos genótipos provenientes de fecundação cruzada, um total de 711 marcas polimórficas, incluindo 632 DArT e 79 SSR, com segregação 1:1 e 3:1, foram utilizadas para análise de ligação. Destas, 604 (84,9%) apresentaram ligação com  $\text{LOD} \geq 3,5$  e fração de recombinação ( $fr$ )  $\leq 0,3$ . O mapa integrado consistiu de 136 GL e apresentou comprimento total de 2.237,2 cM. O tamanho dos GL variou de 0 cM a 76,7 cM, com média de 16,45 cM. Dezoito grupos de ligação foram



formados por apenas duas marcas ligadas a uma distância de 0 cM. Das 711 marcas utilizadas na análise de ligação, 107 (15%) não se ligaram a qualquer outra marca e, portanto, não foram incluídas nas análises para obtenção do mapa. Duzentas e sessenta e sete marcas não puderam ser ordenadas nos grupos de ligação, tendo sido incluídas neste mapa *framework* somente 337 marcas.

### 4.3.3 Detecção de QTL

#### 4.3.3.1 População derivada da autofecundação de RB97327

Os resultados referentes à análise de marcas simples para os genótipos derivados da autofecundação de RB97327 estão ilustrados na Tabela 4.2. Nesta também são apresentadas a proporção da variação genotípica predita pelo marcador ( $R^2$ ) e a estimativa do efeito do marcador identificado em associação às variáveis em estudo ( $b_1$ ), sobre a média da variável ( $b_0$ ).

Observa-se que, com um nível de significância igual a 5%, pode-se verificar associações significativas entre 63 marcadores e as seis variáveis fenotípicas avaliadas. É importante considerar que este estudo constitui o primeiro a mapear QTL relacionados à produtividade em cana-e-açúcar em genótipos derivados de autofecundação, utilizando marcadores DarT e SSR. Não há relatos na literatura do mapeamento de QTL relacionados à produção em cana-de-açúcar a partir de marcadores DarT.

Quanto aos marcadores SSR, observa-se que todos aqueles significativamente associados a alguma das variáveis avaliadas já haviam sido reportados em mapas genéticos e/ou em mapas de QTL obtidos para cana-de-açúcar (Rossi et al., 2003; Reffay et al., 2005; Aitken et al., 2005; 2006; 2008; Garcia et al., 2006; Edmé et al., 2006; Piperidis et al., 2008). Todavia, é importante considerar que, com exceção do primeiro destes trabalhos, nos demais os marcadores SSR foram utilizados em estudos que envolveram populações derivadas de fecundação cruzada.

No que se refere a PESO, sete marcadores, sendo seis DarT e um SSR, apresentaram-se significativamente associados a esta variável, explicando

de 6% a 9% da variação genotípica predita. As estimativas dos efeitos dos marcadores scPtb-428069, scPtb-418674, scPtb-425167 e mSSCIR25\_107 sobre o peso evidenciam que suas presenças contribuem para o incremento da expressão da variável. O marcador mSSCIR25 já havia sido identificado nos mapas obtidos por Rossi et al. (2003) e Aitken et al. (2005), embora, nesses trabalhos o marcador não foi associado a qualquer caráter quantitativo. De modo contrário, os marcadores scPtb-175248, scPtb-384723 e scPtb-251155 demonstram contribuir negativamente para o peso de colmos dos genótipos em estudo, estando suas presenças associadas à redução na média de até 0,6 kg/parcela.

**Tabela 4.2.** Marcadores DarT e SSR significativamente associados às seis variáveis fenotípicas avaliadas na população derivada da autofecundação do genótipo RB97327, pela análise de marcas simples. Os grupos de ligação (GL) a que pertencem estes marcadores, bem como o p-valor, o efeito da sua presença (b1) sobre a média da variável (b0) e o coeficiente de determinação ( $R^2$ ) de cada associação também são apresentados.

| Variável | GL  | Marcador     | p-valor | b0    | b1     | $R^2$  |
|----------|-----|--------------|---------|-------|--------|--------|
| PESO     | 9   | scPtb-175248 | 0,0330  | 1,477 | -0,501 | 0,0625 |
| PESO     | 38  | scPtb-428069 | 0,0412  | 1,402 | 0,498  | 0,0561 |
| PESO     | 52  | scPtb-384723 | 0,0133  | 1,402 | -0,661 | 0,0953 |
| PESO     | 56  | scPtb-418674 | 0,0362  | 1,424 | 0,555  | 0,0542 |
| PESO     | 130 | scPtb-251155 | 0,0277  | 1,474 | -0,519 | 0,0701 |
| PESO     | 136 | scPtb-425167 | 0,0385  | 1,425 | 0,533  | 0,0620 |
| PESO     | 144 | mSSCIR25_107 | 0,0209  | 1,526 | 0,758  | 0,0766 |
| NCOLM    | 12  | scPth-396599 | 0,0153  | 2,895 | -0,738 | 0,0704 |
| NCOLM    | 12  | scPth-426123 | 0,0286  | 2,869 | -0,678 | 0,0624 |
| NCOLM    | 31  | scPth-128704 | 0,0208  | 2,837 | -0,737 | 0,0728 |
| NCOLM    | 32  | scPth-384764 | 0,0492  | 2,802 | -0,648 | 0,0522 |
| NCOLM    | 36  | mSSCIR24_243 | 0,0199  | 2,858 | 0,851  | 0,0949 |
| NCOLM    | 39  | scPth-253873 | 0,0356  | 2,734 | 0,647  | 0,0546 |
| NCOLM    | 39  | scPth-426439 | 0,0356  | 2,734 | 0,647  | 0,0612 |
| NCOLM    | 42  | scPth-429055 | 0,0428  | 2,865 | -0,633 | 0,0598 |
| NCOLM    | 45  | mSSCIR12_251 | 0,0432  | 2,765 | 0,665  | 0,0328 |
| NCOLM    | 48  | scPtb-418615 | 0,0179  | 2,831 | -0,754 | 0,0812 |
| NCOLM    | 48  | scPth-418615 | 0,0206  | 2,821 | -0,753 | 0,0845 |
| NCOLM    | 52  | scPtb-384723 | 0,0217  | 2,770 | -0,804 | 0,0814 |
| NCOLM    | 56  | scPtb-418674 | 0,0273  | 2,797 | 0,763  | 0,0683 |

Continua...

**Tabela 4.2.** Continuação.

| Variável | GL  | Marcador     | p-valor | b0     | b1     | R <sup>2</sup> |
|----------|-----|--------------|---------|--------|--------|----------------|
| NCOLM    | 67  | scPtb-205691 | 0,0288  | 2,889  | 0,87   | 0,0555         |
| NCOLM    | 86  | scPtb-182090 | 0,0107  | 2,818  | -0,82  | 0,1026         |
| NCOLM    | 91  | scPtb-425985 | 0,0147  | 2,796  | 0,815  | 0,085          |
| NCOLM    | 141 | scPth-417124 | 0,0373  | 2,786  | -0,72  | 0,0652         |
| NCOLM    | 156 | mSSCIR14_232 | 0,0376  | 2,699  | -0,9   | 0,0683         |
| COMP     | 9   | scPth-250581 | 0,0433  | 1,669  | -0,074 | 0,0479         |
| COMP     | 16  | mSSCIR26_129 | 0,0228  | 1,676  | 0,11   | 0,0575         |
| COMP     | 16  | scPth-424413 | 0,0325  | 1,672  | 0,092  | 0,0534         |
| COMP     | 58  | scPth-251334 | 0,0142  | 1,667  | -0,09  | 0,0843         |
| COMP     | 16  | scPtb-429271 | 0,0334  | 1,672  | 0,092  | 0,0400         |
| COMP     | 58  | scPth-427777 | 0,0310  | 1,665  | -0,082 | 0,0720         |
| COMP     | 67  | scPtb-428031 | 0,0255  | 1,648  | -0,112 | 0,0756         |
| DIAM     | 1   | mSSCIR12_286 | 0,0133  | 20,454 | -0,246 | 0,0820         |
| DIAM     | 1   | scPth-383803 | 0,0176  | 20,434 | -0,266 | 0,0557         |
| DIAM     | 3   | mSSCIR56_131 | 0,0238  | 20,484 | 0,262  | 0,0790         |
| DIAM     | 16  | scPth-424413 | 0,0165  | 20,482 | 0,263  | 0,0705         |
| DIAM     | 16  | scPth-384415 | 0,0352  | 20,487 | 0,253  | 0,0451         |
| DIAM     | 16  | scPtb-427323 | 0,0371  | 20,488 | 0,257  | 0,0493         |
| DIAM     | 44  | mSSCIR24_269 | 0,0182  | 20,483 | -0,203 | 0,0799         |
| DIAM     | 44  | mSSCIR24_267 | 0,0327  | 20,472 | -0,195 | 0,0576         |
| DIAM     | 116 | scPtb-257282 | 0,0267  | 20,458 | 0,221  | 0,0690         |
| DIAM     | 120 | scPtb-419043 | 0,0422  | 20,495 | 0,273  | 0,0590         |
| DIAM     | 126 | scPth-166003 | 0,0109  | 20,425 | -0,308 | 0,0914         |
| DIAM     | 131 | scPtb-424529 | 0,0416  | 20,454 | -0,208 | 0,0605         |
| NENT     | 27  | scPth-417576 | 0,0103  | 32,116 | -1,271 | 0,0917         |
| NENT     | 27  | scPth-165652 | 0,0220  | 32,111 | -1,143 | 0,0584         |
| NENT     | 27  | scPtb-257223 | 0,0237  | 32,111 | -1,130 | 0,0446         |
| NENT     | 37  | scPtb-190474 | 0,0459  | 32,041 | -1,078 | 0,0492         |
| NENT     | 54  | mSSCIR35_141 | 0,0332  | 32,182 | -1,000 | 0,0679         |
| NENT     | 82  | scPtb-396589 | 0,0267  | 32,070 | -1,155 | 0,0697         |
| NENT     | 86  | scPtb-382123 | 0,0406  | 32,040 | 1,061  | 0,0607         |
| NENT     | 86  | scPtb-182090 | 0,0460  | 32,028 | 1,015  | 0,0550         |
| NENT     | 113 | scPtb-436278 | 0,0287  | 32,032 | 1,136  | 0,0651         |
| NENT     | 155 | mSSCIR11_282 | 0,0287  | 32,103 | -1,153 | 0,0685         |
| BRIX     | 12  | scPth-258257 | 0,0313  | 19,775 | -0,525 | 0,0612         |
| BRIX     | 23  | scPth-425555 | 0,0143  | 19,664 | -0,735 | 0,0916         |
| BRIX     | 23  | scPtb-416459 | 0,0144  | 19,675 | -0,711 | 0,1015         |
| BRIX     | 23  | scPth-425402 | 0,0197  | 19,696 | -0,643 | 0,0838         |
| BRIX     | 23  | scPtb-429941 | 0,0441  | 19,711 | -0,536 | 0,0676         |
| BRIX     | 54  | scPth-251623 | 0,0263  | 19,724 | 0,578  | 0,0667         |
| BRIX     | 74  | scPth-369680 | 0,0363  | 19,763 | 0,584  | 0,0871         |
| BRIX     | 74  | scPth-254449 | 0,0370  | 19,762 | 0,581  | 0,0759         |
| BRIX     | 160 | mSSCIR56_167 | 0,0357  | 19,718 | -0,571 | 0,0687         |

Da mesma forma, para a variável COMP também se verificou associação significativa a sete marcadores, os quais explicaram de 4% a 8% da variação genotípica predita. Dentre estes, somente os marcadores mSSCIR26\_129, scPth-424413 e scPth-429271 contribuíram para o incremento da expressão do comprimento de colmos, estando os demais associados negativamente ao caráter. O marcador mSSCIR26 já havia sido identificado nos mapas obtidos nos trabalhos de Rossi et al. (2003) e Reffay et al. (2005), embora nenhuma associação a QTL tenha sido reportada para este marcador nesses estudos.

Dentre as seis variáveis fenotípicas avaliadas, a que apresentou um maior número de marcadores associados foi NCOLM, tendo sido identificadas 18 associações significativas a esta variável. Os marcadores significativamente associados a NCOLM explicaram uma pequena parte da variação genotípica (3% a 10%) e em sua maioria (onze) demonstrou contribuir negativamente para a expressão do caráter.

O marcador mSSCIR24\_243, presente no GL 36, foi o que apresentou maior contribuição positiva (0,851) para a média do caráter número de colmos, sendo a sua presença capaz de explicar 9% da respectiva variação genotípica predita. Este marcador já havia sido identificado no mapa obtido para a variedade australiana de cana-de-açúcar Q165, por Aitken et al. (2005). Todavia em estudo de QTL, trata-se da primeira vez que o marcador é descrito associado com um caráter quantitativo. Em contrapartida, o marcador mSSCIR14\_232, localizado no GL 156 foi o que apresentou maior contribuição negativa para a variável, estando a sua presença associada a uma redução média de 0,9 colmos/parcela. Não há relatos na literatura da associação deste marcador a algum caráter fenotípico avaliado em genótipos derivados de autofecundação. Todavia, em estudos com genótipos derivados de fecundação cruzada, o marcador foi identificado nos trabalhos de Aitken et al. (2006) e Piperidis et al. (2008), associados à variável brix; e no trabalho de Santana (2011), associado também ao número de colmos, mas promovendo incrementos da média em até 0,42 colmos/parcela.

Para a variável DIAM, doze marcadores, sendo oito DarT e quatro SSR, apresentaram-se significativamente associados à variável. Eles explicaram de 4% a 9% de sua variação genotípica predita. Os marcadores mSSCIR56\_131,

scPth-424413, scPth-384415, scPtb-427323, scPtb-257282 e scPtb-419043 contribuíram de maneira positiva e aproximadamente na mesma proporção, de 0,2 mm, para o incremento na expressão do diâmetro de colmos. De modo contrário, mSSCIR12\_286, scPth-383803, mSSCIR24\_269, mSSCIR24\_267, scPth-166003 e scPtb-424529 demonstram contribuir negativamente para a média do caráter, estando a presença destes associada à redução de até 0,3 mm/colmo.

Nos grupos de ligação GL16, GL52 e GL56 é possível verificar a presença de um só marcador significativamente associado a diferentes caracteres, e com contribuição de magnitude similar. De acordo com Ferreira et al. (2005), a ocorrência deste fato já era, contudo, esperada, haja vista que todos os caracteres avaliados no presente estudo são amplamente correlacionados.

No GL16 observa-se que o marcador scPth-424413 está significativamente associado a COMP e DIAM, contribuindo positivamente para incrementos das médias destas variáveis. Da mesma forma, no grupo de ligação GL56, o marcador scPtb-418674 apresenta-se associado a PESO e NCOLM, também promovendo incremento nas médias destes caracteres. Neste caso, na obtenção de um ideótipo, a seleção poderia ser baseada na presença destes marcadores, de forma a reunir o maior número possível de características favoráveis em um único genótipo.

O marcador scPtb-384723, presente no GL52, mostrou-se significativamente associado, tanto a PESO, quanto a NCOLM, contribuindo negativamente para incrementos das médias de ambas as variáveis. Desta forma, sugere-se que incrementos na média destas variáveis podem ser obtidos a partir da seleção de genótipos que apresentem ausência deste marcador.

Quanto ao número de entrenós (NENT), observa-se que dez marcadores apresentaram-se significativamente associados ao caráter. Estes explicaram de 4% a 9% da variação genotípica predita. Considerando-se que a estimativa do efeito do marcador associada às variáveis permite-nos inferir o quanto a média do genótipo pode ser aumentada ou reduzida pela presença ou ausência de determinado marcador, pode-se propor, a partir dos marcadores, um ideótipo para cada carácter. Assim, um genótipo ideal no que se refere aos caracteres relacionados à produção de açúcar deveria apresentar as marcas que

contribuíram positivamente para a expressão do carácter. Todavia, para a variável NENT, um genótipo ideal deveria apresentar os marcadores scPth-417576, scPth-165652, scPtb-257223, scPtb-190474, mSSCIR35\_141, scPtb-396589 e mSSCIR11\_282, os quais contribuíram negativamente para este carácter. Isto porque foi possível observar durante a avaliação fenotípica que, normalmente, os colmos menos vigorosos eram os apresentavam maior quantidade de entrenós, sugerindo que incrementos na média desta variável pode não constituir uma vantagem quando o objetivo for selecionar materiais com melhor desempenho para os caracteres relacionados à produção de açúcar.

Para a variável BRIX, nove marcadores apresentaram associação significativa a esta variável, explicando entre 6% e 10% da variação genotípica predita. Entre estes, os marcadores scPth-251623, scPth-369680 e scPth-254449 contribuíram para o incremento da média em aproximadamente 0,6° brix. Piperidis et al. (2008) demonstram, em estudos realizados a partir genótipos derivados de fecundação cruzada, envolvendo as variedades R570, MQ77-340 e Q117, a associação de diversos marcadores SSR ao teor de brix, de forma positiva. Todavia, no presente estudo, somente o marcador SSR mSSCIR56\_167, previamente identificado no mapa obtido por Rossi et al. (2003), apresentou-se significativamente associado ao BRIX de forma negativa, estando a sua presença associada a uma redução na média desta variável, em até 0,57° brix.

Pela abordagem tradicional de mapeamento de QTL, a identificação dos locos relacionados aos caracteres quantitativos só poderia ser realizada após a obtenção de um mapa de ligação, no qual associações seriam estabelecidas via análise de marca simples, mapeamento por intervalo (simples ou composto) ou mapeamento por intervalo múltiplo. Todavia, no presente estudo os marcadores que não foram incluídos nas análises de obtenção dos mapas *framework*, também foram considerados nas análises visando à identificação de QTL, e, conforme apresentado na Tabela 4.3, referem-se a todos os marcadores que pertencem aos grupos de ligação identificados como GL96 ou superior.

Das 63 associações significativas obtidas para essa população, via análise de marcas simples, 19% (12) correspondem a de marcas que não participaram de qualquer grupo de ligação pertencente ao mapa *framework*. Entre estes marcadores, verifica-se a presença dos tipos SSR e DarT, isto é, ambos

contribuem tanto de forma positiva, quanto de forma negativa, para incrementos nas médias dos caracteres avaliados. Somente para a variável COMP não se constatou associação significativa com esses marcadores. Desta forma, recomenda-se em estudos posteriores, visando a identificação de QTL em cana-de-açúcar, que para os caracteres relacionados à produção, sejam usadas informações aqui relatadas sobre tais marcadores.

**Tabela 4.3.** Grupos de ligação (GL) contendo os três QTL identificados na população derivada da autofecundação do genótipo RB97327 através da análise de mapeamento por intervalo composto.

| GL | QTL | Posição (cM) <sup>1</sup> | Variável <sup>2</sup> | Marcador associado ao QTL | LOD <sup>3</sup> | (R <sup>2</sup> ) <sup>4</sup> | Marcadores flancuadores      |
|----|-----|---------------------------|-----------------------|---------------------------|------------------|--------------------------------|------------------------------|
| 16 | 1   | 2,7                       | DIAM                  | scPtb-427323              | 4,12             | 0,38                           | scPtb-211926                 |
|    | 2   | 3,2                       | COMP                  |                           | 2,5              | 0,23                           | mSSCIR26_129                 |
| 23 | 3   | 7,1                       | BRIX                  | scPth-425402              | 2,80             | 0,22                           | scPtb-429941<br>scPth-425555 |

<sup>1</sup> Posição (cM): Posição do QTL no seu respectivo grupo de ligação;

<sup>2</sup> Variável: variáveis fenotípicas associadas aos QTL. DIAM: diâmetro da parcela em mm; COMP: comprimento do colmo em m; BRIX: teor de sólidos solúveis em °brix;

<sup>3</sup> LOD: Razão entre a hipótese H<sub>1</sub> (presença de QTL) e H<sub>0</sub> (ausência de QTL);

<sup>4</sup> (R<sup>2</sup>): Proporção da variação genotípica predita explicada pela variação do marcador.

Os resultados referentes à análise de mapeamento por intervalo composto (MIC) permitiram verificar a presença de três QTL (1, 2 e 3), sendo um relacionado a diâmetro dos colmos (DIAM), um a comprimento dos colmos (COMP) e outro ao teor de brix (BRIX) (Tabela 6). Os QTL relacionados a DIAM e COMP estão presentes no grupo de ligação GL16, enquanto o QTL relacionado a BRIX foi verificado no grupo de ligação GL23. Estes QTL estão indicados nos seus respectivos grupos de ligação no mapa *framework* referente ao genitor RB97327, conforme ilustrado no Apêndice B.

Os QTL 1 e 2, relacionados respectivamente a DIAM e COMP, estão associados ao marcador scPtb-427323, o qual explicou 38% da variação genotípica predita para DIAM, e 23% da variação genotípica predita para COMP. A associação de scPtb-427323 ao diâmetro de colmos já havia sido constatada na análise de marcas simples, porém, explicando 5% da variação genotípica para este caráter. Contudo, por este mesmo tipo de análise não se verificou associação significativa entre o marcador scPtb-427323 e COMP, mas sim entre o marcador que o flancueia, mSSCIR26\_129 e COMP. Este, pela análise de

marcas simples, revelou contribuir para incrementos na média do comprimento de colmos.

Para o marcador scPth-425402, presente no grupo de ligação GL23, bem como para os marcadores que o flanqueiam (scPth-425555 e scPtb-429941), verificou-se associação significativa com BRIX, tanto através da análise de marcas simples, quanto através da análise de mapeamento por intervalo composto. A identificação de marcadores associados a QTL relacionados às mesmas variáveis, através de mais de um tipo de análise, garante maior acurácia das associações identificadas, reduzindo a possibilidade de identificação de falsos positivos.

Através da análise de mapeamento por intervalo composto, observa-se que 22% da variação genotípica predita para a variável BRIX foi explicada pela variação do marcador scPth-425402; enquanto na análise de marcas simples a presença deste mesmo marcador explicou 8% dessa mesma variação genotípica. De acordo com Liu (1998), no MIC os marcadores limitam a provável posição em que o QTL se encontra, melhorando consideravelmente a resolução deste local, garantindo assim a obtenção de maior informatividade sobre o loco.

Quanto aos marcadores associados à variável BRIX é importante resaltar que a análise de marcas simples revelou que tanto o marcador scPth-425402, quanto os marcadores que o flanqueiam, contribuem de forma negativa para incrementos da média deste caráter. Diante disso, mesmo que avaliações fenotípicas posteriores sejam necessárias para que os QTL associados aos marcadores scPtb-427323 e scPth-425402 sejam validados, os resultados deste estudo permitem constatar que no genótipo-elite RB97327 há tanto QTL que contribuem para incrementos das médias dos caracteres COMP e DIAM, quanto QTL que contribuem para decréscimos da média de BRIX.

#### 4.3.3.2 População proveniente do cruzamento entre RB97327 e RB72454

Os resultados referentes à análise de marcas simples para os genótipos provenientes do cruzamento entre as variedades RB97327 e RB72454 são apresentados na Tabela 4.4. Verificaram-se, com nível de significância de 5%, associações significativas entre 60 marcadores e as seis variáveis fenotípicas



avaliadas. A variável PESO foi a que apresentou maior número de marcadores associados, tendo sido identificadas treze associações significativas, as quais explicaram entre 3% e 6% da variação genotípica predita.

**Tabela 4.4.** Marcadores DarT e SSR significativamente associados às seis variáveis fenotípicas avaliadas na população proveniente do cruzamento entre RB97327 x RB72454 pela análise de marcas simples. Os grupos de ligação (GL) a que pertencem estes marcadores, bem como o p-valor, o efeito da sua presença (b1) sobre a média do genótipo (b0) e o coeficiente de determinação ( $R^2$ ) de cada associação também são apresentados.

| Variável | GL  | Marcador     | p-valor | b0    | b1     | $R^2$  |
|----------|-----|--------------|---------|-------|--------|--------|
| PESO     | 6   | scPtb-427371 | 0,0257  | 7,350 | -1,514 | 0,0461 |
| PESO     | 45  | scPtb-427354 | 0,0409  | 7,364 | -1,335 | 0,0361 |
| PESO     | 49  | scPth-426196 | 0,0324  | 6,168 | 1,441  | 0,0549 |
| PESO     | 55  | scPh-166136  | 0,0154  | 7,450 | -1,586 | 0,0662 |
| PESO     | 63  | scPth-427226 | 0,0226  | 7,563 | -1,489 | 0,0590 |
| PESO     | 81  | scPtb-436552 | 0,0181  | 7,717 | -1,581 | 0,0631 |
| PESO     | 96  | mSSCIR24_259 | 0,0126  | 6,017 | 1,622  | 0,0623 |
| PESO     | 96  | scPtb-382621 | 0,0381  | 6,163 | 1,385  | 0,0511 |
| PESO     | 97  | mSSCIR67_103 | 0,0463  | 6,075 | 1,312  | 0,0426 |
| PESO     | 99  | scPtb-385759 | 0,0467  | 7,320 | -1,305 | 0,0454 |
| PESO     | 107 | scPh-383236  | 0,0256  | 6,088 | 1,477  | 0,0565 |
| PESO     | -   | mSSCIR24_243 | 0,0478  | 7,253 | -1,335 | 0,0447 |
| PESO     | -   | mSSCIR56_131 | 0,0288  | 7,336 | -1,474 | 0,0543 |
| NCOLM    | 45  | scPtb-427354 | 0,0098  | 7,425 | -1,631 | 0,0576 |
| NCOLM    | 45  | scPtb-190474 | 0,0118  | 7,375 | -1,606 | 0,0826 |
| NCOLM    | 49  | scPth-426196 | 0,0439  | 6,144 | 1,324  | 0,0462 |
| NCOLM    | 68  | scPtb-430009 | 0,0248  | 6,170 | 1,475  | 0,0490 |
| NCOLM    | 88  | scPtb-166529 | 0,0046  | 5,682 | 1,798  | 0,0894 |
| NCOLM    | 99  | scPtb-385759 | 0,0483  | 7,230 | -1,262 | 0,0492 |
| NCOLM    | 107 | scPth-383236 | 0,0248  | 6,030 | 1,446  | 0,0572 |
| NCOLM    | 108 | scPtb-417627 | 0,0135  | 5,923 | 1,582  | 0,0685 |
| COMP     | 6   | scPtb-417245 | 0,0228  | 2,068 | -0,084 | 0,0561 |
| COMP     | 6   | scPth-367708 | 0,0247  | 2,067 | -0,083 | 0,0572 |
| COMP     | 48  | mSSCIR67_157 | 0,0229  | 2,001 | 0,084  | 0,0587 |
| COMP     | 65  | scPth-435948 | 0,0432  | 2,070 | -0,074 | 0,0466 |
| COMP     | 74  | scPth-426018 | 0,0319  | 1,995 | 0,079  | 0,0525 |
| COMP     | 81  | scPth-396535 | 0,0412  | 2,080 | -0,076 | 0,0390 |
| COMP     | -   | scPtb-254099 | 0,0124  | 2,080 | -0,092 | 0,0704 |
| COMP     | -   | mSSCIR24_263 | 0,0336  | 2,007 | 0,08   | 0,0514 |
| COMP     | -   | mSSCIR24_243 | 0,0218  | 2,067 | -0,085 | 0,0596 |

Continua...

**Tabela 4.4.** Continuação.

| Variável | GL  | Marcador     | p-valor | b0     | b1     | R <sup>2</sup> |
|----------|-----|--------------|---------|--------|--------|----------------|
| COMP     | -   | scPth-166003 | 0,0388  | 2,080  | -0,077 | 0,0486         |
| COMP     | -   | mSSCIR56_171 | 0,0183  | 2,086  | -0,088 | 0,0630         |
| COMP     | -   | scPth-382489 | 0,0380  | 2,078  | -0,076 | 0,0491         |
| DIAM     | 6   | scPth-367708 | 0,0142  | 25,385 | -0,274 | 0,0678         |
| DIAM     | 6   | scPtb-424764 | 0,0157  | 25,398 | -0,273 | 0,0519         |
| DIAM     | 26  | scPtb-205242 | 0,0162  | 25,435 | -0,266 | 0,0653         |
| DIAM     | 27  | scPtb-254331 | 0,0217  | 25,139 | 0,260  | 0,0706         |
| DIAM     | 44  | scPth-258257 | 0,0227  | 25,168 | 0,252  | 0,0542         |
| DIAM     | 55  | scPth-166136 | 0,0118  | 25,403 | -0,277 | 0,0713         |
| DIAM     | 65  | scPth-435948 | 0,0440  | 25,387 | -0,224 | 0,0463         |
| DIAM     | 81  | scPth-429526 | 0,0181  | 25,432 | -0,265 | 0,0628         |
| DIAM     | 96  | scPtb-382621 | 0,0234  | 25,173 | 0,254  | 0,0695         |
| DIAM     | -   | scPtb-417627 | 0,0322  | 25,167 | 0,238  | 0,0500         |
| DIAM     | -   | scPtb-206567 | 0,0152  | 25,169 | 0,266  | 0,0665         |
| NENT     | 1   | scPtb-417297 | 0,0216  | 31,547 | -1,051 | 0,0491         |
| NENT     | 12  | scPth-426651 | 0,0403  | 31,579 | -0,903 | 0,0484         |
| NENT     | 90  | scPth-250985 | 0,0122  | 31,764 | -1,101 | 0,0588         |
| NENT     | 99  | scPtb-385759 | 0,0369  | 30,765 | 0,917  | 0,0494         |
| NENT     | 112 | scPth-417296 | 0,0344  | 31,626 | -0,926 | 0,0464         |
| NENT     | -   | mSSCIR24_243 | 0,0287  | 31,522 | -0,988 | 0,0544         |
| NENT     | -   | mSSCIR56_171 | 0,0187  | 31,770 | -1,061 | 0,0626         |
| NENT     | -   | scPth-382489 | 0,0137  | 31,765 | -1,084 | 0,0685         |
| BRIX     | 3   | scPtb-205148 | 0,0375  | 19,721 | 0,507  | 0,0504         |
| BRIX     | 49  | scPth-426196 | 0,0225  | 20,193 | -0,572 | 0,0573         |
| BRIX     | 51  | scPtb-416556 | 0,0408  | 20,197 | -0,496 | 0,0377         |
| BRIX     | 94  | scPth-420641 | 0,0179  | 20,208 | -0,578 | 0,0635         |
| BRIX     | 94  | scPth-173964 | 0,0210  | 20,191 | -0,569 | 0,0593         |
| BRIX     | 98  | scPth-428811 | 0,0275  | 19,680 | 0,537  | 0,0460         |
| BRIX     | 98  | scPtb-428811 | 0,0275  | 19,680 | 0,537  | 0,0564         |
| BRIX     | 105 | scPtb-428310 | 0,0198  | 20,255 | -0,565 | 0,0628         |

As estimativas dos efeitos dos marcadores revelaram que somente scPh-426196, mSSCIR24\_259, scPtb-382621 e mSSCIR67\_103 contribuíram para o incremento na expressão do caráter PESO. Dentre estes, a presença do marcador mSSCIR24\_259 destacou-se por estar associada a um incremento de até 1,6 kg/parcela.

Os marcadores mSSCIR67 e mSSCIR56, também significativamente associados a PESO, de forma positiva e negativa, respectivamente, já haviam sido mapeados por Rossi et al. (2003), embora sem relatos de associação a qualquer caráter quantitativo. Os outros nove marcadores significativamente

associados a PESO demonstraram contribuir negativamente para a média desta variável, estando suas presenças associadas a uma redução da média de até 1,5 kg/parcela.

Resultados semelhantes aos verificados para a variável PESO foram obtidos para a variável COMP, para a qual se observou associação significativa a doze marcadores, os quais explicaram de 3% a 7% da variação genotípica predita. Dentre estes, somente os marcadores scPth-426018, mSSCIR24\_263 e mSSCIR67\_157 contribuíram positivamente para o incremento na expressão do comprimento de colmos. Os demais estiveram associados negativamente ao caráter.

Para as variáveis NCOLM, DIAM e BRIX nenhum marcador SSR apresentou-se significativamente associado. Oito marcadores DarT apresentaram-se significativamente associados a NCOLM, explicando de 4% a 8% de sua variação genotípica predita. Destes, os marcadores scPth-426196, scPtb-430009, scPtb-417627, scPth-383236 e scPtb-166529 contribuíram positivamente para incrementos na média do caráter, destacando-se o último, no grupo de ligação GL88, devido à sua presença contribuir com incremento na média de, aproximadamente, 1,8 colmos/parcela.

Onze marcadores DarT apresentaram-se significativamente associados a DIAM, explicando de 4% a 7% de sua variação genotípica predita. Os marcadores scPtb-254331, scPth-258257, scPtb-382621, scPtb-417627 e scPtb-206567 contribuíram de maneira positiva e, aproximadamente, na mesma proporção, de 0,2 mm, com incrementos na média do caráter.

Quanto à variável BRIX, oito marcadores apresentaram associação significativa a esta variável, explicando entre 4% e 6% de sua variação genotípica predita. Destes, os marcadores scPth-428811, scPtb-428811 e scPtb-205148 contribuíram com incremento positivo de aproximadamente 0,5° brix.

Quanto ao caráter NENT foi possível identificar oito marcadores significativamente associados, os quais explicaram de 4% a 6% da variação genotípica predita. Dentre estes, os marcadores mSSCIR24\_243 e mSSCIR56\_171 contribuíram para uma redução da média do número de entrenós e somente o marcador scPtb-385759 contribuiu com incremento para a média do caráter. Todavia, como foi observado durante as avaliações fenotípicas, genótipos

que apresentavam menor número de entrenós mostravam-se mais vigorosos. Logo, pode-se afirmar que todos os sete marcadores que se apresentaram significativamente associados de forma negativa a este caráter são almejados na obtenção de um ideótipo. São eles: scPtb-417297, scPth-426651, scPth-250985, scPth-417296, mSSCIR24\_243, mSSCIR56\_171 e scPth-382489.

Em alguns grupos de ligação verificou-se, em cada grupo, a ocorrência de um só marcador significativamente associado a diferentes caracteres. Este fato corrobora a observação de que um maior número de entrenós está associado a genótipos menos vigorosos, uma vez que, no GL99 o marcador scPtb-385759 apresenta-se relacionado ao PESO, NCOLM e NENT, contribuindo, ao mesmo tempo, para a redução da média das duas primeiras variáveis e com incrementos na média da última.

Somente o marcador scPth-383236, presente no grupo de ligação GL107, apresentou-se relacionado a PESO e NCOLM concomitantemente, promovendo incrementos na média das duas variáveis. Os marcadores scPth-367708, scPtb-427354, scPth-435948, mSSCIR24\_243, mSSCIR56\_171 e scPth-382489 apresentaram-se relacionados a dois diferentes caracteres (Tabela 4.4) e, para ambos, apresentaram contribuições para a redução de incrementos na média. Logo, não é desejável a presença destes marcadores em trabalhos que visem a seleção de genótipos com melhor desempenho para caracteres relacionados à produção em cana-de-açúcar.

Das sessenta associações significativas identificadas na população proveniente de fecundação cruzada, por meio da análise de marcas simples, 20% (doze) correspondem a marcadores que não participaram de qualquer grupo de ligação no mapa *framework*. Dois destes marcadores estão associados à variável PESO, cinco à COMP, dois à DIAM e três à NENT. Destes, somente os marcadores mSSCIR24\_263, scPtb-417627 e scPtb-206567 contribuíram para incrementos nas médias das variáveis COMP e DIAM. Os demais contribuíram para redução da média dos caracteres a eles relacionados. Logo, a inclusão destes marcadores nas análises que visam à identificação de QTL é de fundamental importância em estudos posteriores, pois permite não somente agregar informações sobre o efeito da presença/ausência destes marcadores, como também comparar seus comportamentos em diferentes populações.

Os resultados referentes à análise de mapeamento por intervalo composto (MIC) permitiram identificar a presença de dois QTL (4 e 5), sendo um relacionado a comprimento dos colmos (COMP) e outro a diâmetro dos colmos (DIAM) (Tabela 4.5). O QTL relacionado a COMP está no grupo de ligação GL6, enquanto aquele relacionado a DIAM, no grupo de ligação GL81. Estes QTL estão indicados nos seus respectivos grupos de ligação no mapa *framework* integrado, conforme ilustrado no Apêndice C.

O QTL 4, presente no grupo de ligação GL6, está associado ao marcador scPtb-427371, o qual explicou 12% da variação genotípica predita para COMP. Observa-se que, pela análise de marcas simples, scPtb-427371 não se apresentou significativamente associado a COMP, mas sim ao caráter peso de colmos (PESO). Além disso, é importante considerar que na análise de marcas simples a associação de scPtb-427371 a PESO explicou somente 4% da variação genotípica predita para este caráter e, ainda, contribuiu para redução de sua média. Da mesma forma, os marcadores scPth-367708 e scPtb-424764, flangeadores de scPtb-427371 no MIC, também apresentaram-se, através da análise de marcas simples, relacionados a outro caráter (DIAM), contribuindo para a redução de sua média.

**Tabela 4.5.** Grupos de ligação (GL) contendo os dois QTL identificados na população proveniente do cruzamento entre RB97327 x RB72454 através da análise de mapeamento por intervalo composto.

| GL | QTL | Posição (cM) | Variável | Marcador associado ao QTL | LOD  | R <sup>2</sup> | Marcadores flangeadores       |
|----|-----|--------------|----------|---------------------------|------|----------------|-------------------------------|
| 6  | 4   | 32           | COMP     | scPtb-427371              | 2,5  | 0,12           | scPth-367708<br>scPtb-424764  |
| 81 | 5   | 11           | DIAM     | scPtb-436552              | 4,65 | 0,17           | scPth- 429526<br>scPtb-396535 |

<sup>1</sup> Posição (cM): Posição do QTL no seu respectivo grupo de ligação;

<sup>2</sup> Variável: variáveis fenotípicas associadas aos QTL. DIAM: diâmetro da parcela em mm; COMP: comprimento do colmo em m;

<sup>3</sup> LOD: Razão entre a hipótese H<sub>1</sub> (presença de QTL) e H<sub>0</sub> (ausência de QTL);

<sup>4</sup> (R<sup>2</sup>): Proporção da variação genotípica predita explicada pela variação do marcador.

Situação semelhante pode ser observada em relação ao QTL 5, presente no grupo de ligação GL81 e associado ao marcador scPtb-436552. Pela análise de marcas simples, verificou-se associação significativa do marcador scPtb-436552 à variável PESO, enquanto na análise de mapeamento por

intervalo composto, a presença deste marcador explicou 17% da variação genotípica predita para a variável DIAM.

Quanto aos marcadores flaqueadores de scPtb-436552, observa-se que somente o marcador scPth-429526 apresentou-se, pela análise de marcas simples, associado a DIAM, sendo que scPtb-396535 esteve significativamente associado a COMP. Além de se observar que através de diferentes metodologias não foi possível associar os mesmos marcadores às mesmas variáveis, deve-se ainda considerar que, pela análise de marcas simples, tanto o marcador scPtb-436552, quanto os marcadores que o flanqueiam, scPth-429526 e scPtb-396535, contribuíram para a redução de incrementos na média de todas as variáveis às quais foi relacionado: PESO, DIAM e COMP, respectivamente.

De acordo com Sills et al. (1995), marcadores *single-dose* analisados em uma progênie  $F_1$  derivada do cruzamento de genótipos com elevada heterozigosidade maximizam o desequilíbrio de ligação, de tal forma que os marcadores encontrados significativamente associados a QTL representam regiões genômicas muito grandes, permitindo que mais de um QTL seja representado por um único marcador. Isto pode provocar atribuição incorreta de efeitos a um único loco. Assim, sabendo-se que o tamanho da população utilizada para obtenção do mapa, a complexidade e o tamanho do genoma da espécie são capazes de influenciar a detecção precisa de QTL (Pastina et al., 2010), sugere-se que na população proveniente do cruzamento RB97327 x RB72454 sejam utilizados maior número de marcadores em uma maior quantidade de genótipos, de forma a se obter maior nível de resolução quanto ao efeito dos QTL 4 e 5 identificados no presente estudo.

#### 4.3.3.3 Considerações gerais sobre os QTL identificados

De maneira geral, observa-se que, embora em ambas as populações os QTL relacionados à produção em cana-de-açúcar e identificados através da análise de marcas simples explicaram pequena fração da variação genotípica predita, entre 3% e 10%, estes dados estão de acordo a literatura. Hoarau et al. (2002) já havia verificado que os QTL identificados em cana-de-açúcar explicavam somente de 3% a 9% da variação fenotípica observada. Piperidis et

al. (2008) também afirmaram que os QTL já identificados em cana-de-açúcar e relacionados ao teor de sacarose explicaram apenas 3% a 8% da variação fenotípica. Para estes autores, vários fatores podem contribuir para a pequena explicação desta variação em cana-de-açúcar como, por exemplo, a forma com que as análises tem sido conduzidas nessa espécie.

Quando associações significativas entre marcadores e QTL são estabelecidas em cana-de-açúcar, as informações que estão sendo consideradas referem-se à de apenas um ou dois alelos, mas pode ocorrer de até oito ou dez alelos contribuírem para a determinação de um dado caráter. Neste caso, como diferentes dosagens dos alelos não são consideradas nas análises de mapeamento em cana-de-açúcar, os QTL que têm sido identificados explicam somente a parte da variação que lhes é devida, podendo ocorrer de, mais alelos que contribuem para a determinação do caráter, não estarem sendo considerados e nem quantificados.

Hoarau et al. (2002) ainda ressaltam que quando a variação explicada pelos marcadores é pequena há evidências de que os QTL possuem pequenos efeitos ou não estão estreitamente ligados aos marcadores. E no caso de poliploides, como vários alelos coexistem para um mesmo loco, para que o efeito de um alelo se destaque é necessário que o seu efeito exceda não apenas o efeito de outro alelo qualquer, mas também o efeito médio dos demais alelos segregantes para o mesmo loco.

Neste contexto, é válido também considerar o complexo padrão de interações epistáticas e dominantes existente entre e dentro dos locos. Desta forma, considera-se que, mesmo que as associações identificadas através da análise de marcas simples tenham explicado uma pequena parte da variação genotípica predita, pode-se afirmar que este estudo forneceu informações relevantes sobre a associação dos marcadores SSR e DArT às variáveis envolvidas na produção de cana-de-açúcar, em dois diferentes tipos de população.

O presente estudo também nos permitiu verificar que os marcadores DArT revelaram-se mais eficientes que os demais marcadores publicados na literatura de cana-de-açúcar, relacionada à obtenção de mapas de ligação e para a detecção de QTL. Isto porque são capazes de gerar grande número de marcas

polimórficas. Contudo, como em cana-de-açúcar diversos locos e alelos de pequeno efeito podem estar envolvidos na determinação dos caracteres quantitativos relacionados à produção, para identificar marcadores que expliquem uma maior proporção da variação genotípica ressalta-se a necessidade de marcadores que sejam capazes de gerar um número de marcas ainda maior que os marcadores DArT. Além disso, é importante também que diferentes dosagens destes marcadores sejam consideradas nas análises.

Quanto aos marcadores SSR, observa-se que alguns QTL associados a esse tipo de marcador e identificados no presente estudo já haviam sido reportados em mapas de ligação publicados na literatura; embora o marcador mSSCIR24 ainda não havia sido relatado em mapas de QTL obtidos para cana-de-açúcar. No presente estudo, mSSCIR24 apresentou-se, através da análise de marcas simples, significativamente associado a diferentes caracteres nas duas populações avaliadas. Na população derivada de autofecundação, este marcador apresentou-se associado a NCOLM e DIAM, de forma positiva e negativa, respectivamente, enquanto na população proveniente de fecundação cruzada, apresentou-se associado de forma positiva e/ou negativa a NENT, COMP e PESO.

Piperidis et al. (2008) ressaltam sobre a importância da identificação de QTL comuns em mais de uma população, sendo este um exercício necessário para a validação dos marcadores. Assim como no trabalho desses autores, no presente estudo houve dificuldades de se verificar, pela análise de marcas simples, as mesmas marcas SSR associadas aos mesmos QTL em ambas as populações, embora os marcadores SSR associados às variáveis avaliadas (mSSCIR11, mSSCIR12, mSSCIR14, mSSCIR24, mSSCIR25, mSSCIR26, mSSCIR35, mSSCIR56 e mSSCIR67), já haviam sido apresentados em mapas de ligação anteriormente publicados para cana-de-açúcar.

Para Pastina et al. (2012), a dificuldade de identificação de QTL comuns em mais de uma população pode estar relacionada ao fato de que, provavelmente, os fatores envolvidos na produção de açúcar estão distribuídos por todo o genoma. Como normalmente as populações utilizadas para construção de mapas são pequenas e apenas um único tipo de segregação é considerado, os mapas obtidos apresentam baixa saturação, implicando em baixa cobertura do



genoma e aumento da dificuldade para identificação de QTL.

De acordo com Pastina et al. (2012), a maioria dos QTL identificados em cana-de-açúcar foi realizada através da análise de marcas simples ou mapeamento por intervalo, sendo raros os QTL obtidos através de mapeamento por intervalo composto. No entanto, a identificação de QTL por análises de marcas simples é muito sujeita à identificação de QTL “fantasmas”, em função da ocorrência de falsos positivos presentes próximos aos QTL alvos, reduzindo muito o poder de detecção dos reais QTL.

Diante disso, o presente estudo permitiu a identificação de QTL através de ambas as metodologias, utilizando dois tipos diferentes de populações e possibilitando, com isso, a identificação de marcadores associados a QTL relacionados às mesmas variáveis através de mais de um tipo de análise. Com isso reduz-se consideravelmente a possibilidade de identificação de falsos positivos.

#### 4.4 CONCLUSÕES

- Em uma população derivada de autofecundação (RB97327), identificou-se através da análise de marcas simples, 63 marcadores (DArT e SSR) significativamente associados às variáveis fenotípicas avaliadas (peso dos colmos, número de colmos, número de entrenós, teor de sólidos solúveis (brix), comprimento e diâmetro dos colmos). Nesta mesma população, através da análise de mapeamento por intervalo composto, três marcadores DArT foram associados a QTL relacionados a brix, diâmetro e comprimento dos colmos.
- Na população proveniente de fecundação cruzada (RB97327 x RB72454) foram identificados, pela análise de marcas simples, 60 marcadores significativamente associados às mesmas variáveis avaliadas na população de autofecundação. Ao se proceder ao mapeamento por intervalo composto, identificou-se dois marcadores DArT associados a QTL relacionados ao comprimento e diâmetro dos colmos.

## 5 CONCLUSÕES GERAIS

Pelos resultados obtidos das análises de mapeamento foi possível obter os primeiros mapas de ligação utilizando marcadores DArT e SSR. Ao se observar que para a obtenção destes houve compartilhamento da maior parte das marcas DArT e SSR (90,2%), mesmo que dois tipos diferentes de população (autofecundação e fecundação cruzada) tenham sido utilizados, garante-se a estabilidade dos mapas de ligação obtidos.

Os marcadores SSR revelam-se importantes em estudos de mapeamento genético, sobretudo por permitirem a identificação dos grupos de homeologia. Já a utilização de genotipagem DArT, por associar baixos custos à rapidez e praticidade, mostrou-se amplamente eficiente no rendimento de marcadores moleculares espalhados por todo o genoma. Assim, estimula-se a sua utilização em estudos de mapeamento e melhoramento de cana-de-açúcar.

Além disso, como se tem conhecimento das sequências de todas as sondas DArT utilizadas no presente estudo, acredita-se que o seu emprego tem muito a contribuir na obtenção de um genoma de referência para cana-de-açúcar. Todavia, dado o grande número de locos monomórficos verificados no estudo, faz-se ainda necessário rever os métodos de redução da complexidade do genoma, etapa fundamental na geração de polimorfismo.

Embora a cobertura alcançada pelos mapas tenha sido baixa, foi possível identificar marcadores DArT e SSR significativamente associados aos caracteres relacionados à produção de açúcar. A utilização das informações fenotípicas associadas às fornecidas pelos marcadores moleculares permitiu um melhor conhecimento do número, da localização e dos efeitos dos locos que contribuíram para a expressão dos caracteres quantitativos, através do mapeamento de QTL.

## 6 REFERÊNCIAS

AITKEN, K. S.; HERMANN, S.; KARNO, K.; BONNETT, G. D.; MCINTYRE, C. L.; JACKSON, P. A. Genetic control of yield related stalk traits in sugarcane. **Theoretical and Applied Genetics**, Berlin, v. 117, n. 7, p. 1191-1203, nov. 2008.

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 and Applied Genetics**, Berlin, v. 110, n. 5, p. 789-801, mar. 2005.

AITKEN, K. S.; JACKSON, P. A.; MCINTYRE, C. L. Construction of a genetic linkage map for *Saccharum officinarum* incorporating both simplex and duplex markers to increase genome coverage. **Genome**, Ottawa, v. 50, n. 8, p. 742-756, ago. 2007.

AITKEN, K. S.; JACKSON, P. A.; MCINTYRE, C. L. Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar x *Saccharum officinarum* population. **Theoretical and Applied Genetics**, Berlin, v. 112, n. 7, p. 1306-1317, mai. 2006.

AKBARI, M.; WENZL, P.; CAIG, V.; CARLING, J.; XIA, L.; YANG, S.; USZYNSKI, G.; MOHLER, V.; LEHMENSIEK, A.; KUCHEL, H.; HAYDEN, M. J.; HOWES, NEIL.; SHARP, P.; VAUGHAN, P.; RATHMELL, B.; HUTTNER, E.; KILIAN, A. Diversity arrays technology (DArT) for high- throughput profiling of the hexaploid wheat genome. **Theoretical and Applied Genetics**, Berlin, v. 113, p. 1409-1420, out. 2006.

ALBINO, J. C.; CRESTE, S.; FIGUEIRA, A. Mapeamento genético da cana-de-açúcar: Uso de marcadores moleculares no mapeamento genético visando o melhoramento da cana-de-açúcar. **Biotecnologia, Ciência e Desenvolvimento**, Brasília, v. 36, n. 9, p. 82-91, jan/jun. 2006.

ALJANABI, S. M.; FORGET, L.; DOOKUN, A. An improved and rapid protocol for the isolation of polysaccharide- and polyphenol-free sugarcane DNA. **Plant Molecular Biology Reporter**, Dordrecht, v. 17, n. 3, p. 1-8, 1999.

AL-JANABI, S. M.; HONEYCUTT, R. J.; MCCLELLAND, M.; SOBRAL, B. W. S. A genetic linkage map of *Saccharum spontaneum* L. 'SES208'. **Genetics**, Baltimore, v. 134, n. 4, p. 1249-1260, ago. 1993.

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

1-14, jan. 2007.

ALLARD, R. W. **Princípios do melhoramento genético de plantas**. São Paulo: Edgard Blucher, 1971. 381 p.

ALIX, K.; PAULET, F.; GLASZMANN, J.C.; D'HONT, A. Inter-Alu like species-specific sequences in the *Saccharum* complex. **Theoretical and Applied Genetics**, Berlin, v. 99, n. 6, p. 962-968, out. 1999.

ALWALA, S.; KIMBENG, C. A.; VEREMIS, J. C.; GRAVOIS, K. A. Identification of molecular markers associated with sugar-related traits in a *Saccharum* interspecific cross. **Euphytica**, Amsterdam, v. 167, n. 1, p. 127-142, mai. 2009.

ALWALA, S.; KIMBENG, C. **Molecular genetic linkage mapping in *Saccharum*: Strategies, resources and achievements**. In: HENRY, R. J.; KOLE, C. (Ed.). Genetics genomics and breeding of sugarcane. Clemson: CRC Press, 2010. cap. 5, p. 69-96.

ALWALA, S.; KIMBENG, C. A.; VEREMIS, J. C.; GRAVOIS, K. A. Linkage mapping and genome analysis in *Saccharum* interspecific cross using AFLP, SRAP and TRAP markers. **Euphytica**, Amsterdam, v. 164, n. 1, p. 37-51, jan. 2008.

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

ASNAGHI, C.; PAULET, F.; KAYE, C.; GRIVET, L.; DEU, M.; GLASZMANN, J. C.; D'HONT, A. Application of synteny across Poaceae to determine the map location of a sugarcane rust resistance gene. **Theoretical and Applied Genetics**, Berlin, v. 101, n. 5-6, p. 962-969, out. 2000.

ASNAGHI, C.; ROQUES, D.; RUFFEL, S.; KAYE, C.; HOARAU, J. Y.; TÉLISMART, H.; GIRARD, J. C.; RABOIN, L. M.; RISTERUCCI, A. M.; GRIVET, L.; D'HONT, A. Targeted mapping of a sugarcane rust resistance gene (*Bru1*) using bulked segregant analysis and AFLP markers. **Theoretical and Applied Genetics**, Berlin, v. 108, n. 4, fev. 2004.

BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J. B.; CARNEIRO, I. F (Ed.). **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE, 2000. cap. 5, p. 63-223.

BERTRAND, C. Y.; MACKILL, D. J. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. **Philosophical Transactions of the Royal Society**, London, v. 363, n. 1491, p. 557-572, fev. 2008.

BESSE, P.; McINTYRE, C. L.; BERDING, N. Characterization of *Erianthus* sect. *Ripidium* and *Saccharum* germoplasm (Andropogoneae - Saccharinae) using RFLP markers. **Euphytica**, Amsterdam, v. 93, n. 3, p. 283-292, fev. 1997.

BHAT, S. R.; GILL, S. S. The implications of 2n egg gametes in nobilization and breeding of sugarcane. **Euphytica**, Amsterdam, v. 34, n. 2, p. 377-384, ago. 1985.

BNDES. **Bioetanol de cana-de-açúcar**: energia para o desenvolvimento sustentável. Rio de Janeiro: BNDES, 2008. 316 p.

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

BRASIL. MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO. **Exportações brasileiras de etanol**. Disponível em: <[http://www.agricultura.gov.br/arq\\_editor/file/Desenvolvimento\\_Sustentavel/Agroenergia/estatisticas/producao/atualizacao\\_20\\_01\\_2011/Producao\\_Mensal\\_Cana\\_Acucar\\_Alcool\\_Brasil.pdf](http://www.agricultura.gov.br/arq_editor/file/Desenvolvimento_Sustentavel/Agroenergia/estatisticas/producao/atualizacao_20_01_2011/Producao_Mensal_Cana_Acucar_Alcool_Brasil.pdf)> Acesso em: 17 mar. 2011.

BREMER, G. Problems in breeding and cytology of sugar cane. **Euphytica**, Amsterdam, v. 10, n. 1, p. 59-78, mar. 1961.

BUETOW, K. H.; CHAKRAVARTI, A. Multipoint gene mapping using seriation I. General methods. **American Journal Human Genetics**, Chicago, v. 41, n. 2, p.180-188, 1987

BUNDOCK, P. C.; ELIOTT, F. G.; ABLETT G.; BENSON, A. D.; CASU, R. E.; AITKEN K. S.; HENRY, R. J. Targeted single nucleotide polymorphism (SNP) discovery in a highly polyploid plant species using 454 sequencing. **Plant Biotechnology Journal**, Malden, v. 7 p. 347–354, dez. 2009.

BUTTERFIELD, M. K.; D'HONT, A. D.; BERDING, N. The sugarcane genome: a synthesis of current understanding, and lessons for breeding and biotechnology. **Proceedings of the South African Sugar Technologists Association**, Durban, v. 75, n. 1, p.1-5, 2001.

CARNEIRO, M. S.; VIEIRA, M. L. C. Mapas genéticos em plantas. **Bragantia**, Campinas, v. 61, n. 2, p.89-100, mai/ago. 2002.

CESNIK, R.; MIOCQUE, J. **Melhoramento da cana-de-açúcar**. Brasília: Embrapa, 2004. 307 p.

COELHO, A. S. G. Considerações gerais sobre a análise de QTL's. In: PINHEIRO, J. B.; CARNEIRO, I. F. (Ed.). **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE, 2000. cap. 1, p.1-36.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento de safra brasileira**: Cana-de-açúcar, 3º Levantamento. Brasília: Conab, 2011. 19 p.

CORDEIRO, G. M.; CASU, R.; McINTYRE, C. L.; MANNERS, J. M.; HENRY, R.J. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. **Plant Science**, Amsterdam, v. 160, n. 6, p.1115-1123, 2001.

CORDEIRO, G. M.; ELIOTT, F.; McINTYRE, C. L.; CASU, R. E.; HENRY, R. J. Characterization of single nucleotide polymorphisms in sugarcane ESTs. **Theoretical Applied Genetic**, Berlin, v. 113, n. 2, p.331–343, 2006.

CORDEIRO, M. G.; PAN, Y. B.; HENRY, R. J. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. **Plant Science**, Amsterdam, v. 165, n. 1, p. 181-189, jul. 2003.

CORDEIRO, G. M.; TAYLOR, G. O.; HENRY, R. J. Characterization of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploidy species. **Plant Science**, Amsterdam, v. 155, n. 2, p. 161-168, 2000.

COVER, C. **Caracterização fenotípica e genotípica de caracteres agrônômicos em uma população de linhagens recombinantes de aveia (*Avena sativa* L.)**. 2010.100 f. Dissertação (Mestrado em Fitotecnia com ênfase em Melhoramento e Biotecnologia Vegetal) – Universidade Federal do Rio Grande do Sul, Porto Alegre. 2010.

DANIELS, J.; ROACH, B. T. Taxonomy and evolution. In: HEIZ, D. J. (Ed.). **Sugarcane improvement through breeding**. Amsterdam: Elsevier, 1987. v. 1, cap. 2, p. 7-84.

DAL-BIANCO, M.; CARNEIRO, M. S.; HOTTA, C. T.; CHAPOLA, R. G.; HOFFMAN, H. P.; GARCIA, A. A. F.; SOUZA, G. M. Sugarcane improvement: how far can we go? **Current Opinion in Biotechnology**, Amsterdam, v. 23, n. 2, p. 1-6, 2011.

Da SILVA, J. A. G; BRESSIANI, J. A. Sucrose synthase molecular marker associated with sugar content in elite sugarcane progeny. **Genetics and Molecular Biology**, São Paulo, v. 28, n. 2, 2005.

Da SILVA, J. A. G.; HONEYCUTT, R. J.; BURNQUIST, W.; AL-JANABI, S.M.; SORRELLS, M. E.; TANKSLEY, S. D.; SOBRAL, B.W. S. *Saccharum spontaneum* L. “SES 208” genetic linkage map combining RFLP- and PCR-based markers. **Molecular Breeding**, Netherlands, v. 1, n. 2, p. 165-179, jun. 1995.

Da SILVA, J. A. G.; SORRELLS, M. E.; BURNQUIST, W. L.; TANKSLEY, S. D. RFLP linkage map and genome analysis of *Saccharum spontaneum*. **Genome**, Ottawa, v. 36, n. 4, p. 782-791, ago. 1993.

DAUGROIS, J.; GRIVET, L.; ROQUES, D.; HOARAU, J.; LOMBARD, H.; GLASZMANN, J.; D'HONT, A. A putative major gene for rust resistance linked with a RFLP marker in sugarcane var 'R570'. **Theoretical and Applied Genetics**, Berlin, v. 92, n. 8, p. 1059-1064, jun. 1996.

D'HONT, A.; GRIVET, L.; FELDMANN, P.; RAO, S.; BERDING, N.; GLASZMANN, J. C. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. **Molecular and General Genetics**, Berlin, v. 250, n. 4, p. 405-413, mar. 1996.

D'HONT, A.; ISON, D.; ALIX, K.; ROUX, C.; GLASZMANN, J. C. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome, Ottawa*, v. 41, n. 2, p. 221-225, 1998.

D'HONT, A.; PAULET, F.; GLASZMANN, J. C. Oligoclonal interspecific origin "North Indian" and "Chinese" sugarcane. *Chromosome Research*, Oxford, v. 10, n. 3, p. 253-262, mar. 2002.

DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento vegetal**. 2000. 293 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas)-Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba. 2000.

DUFOUR, P.; DEU, M.; GRIVET, L.; D'HONT, A.; PAULET, F.; BOUET, A.; LANAUD, C.; GLASZMANN, J. C.; HAMON, P. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploidy. *Theoretical Applied Genetic*, Berlin, v. 94, n. 3-4, p. 409-418, 1997.

EDMÉ, S. J.; GLYNN, N. G.; COMSTOCK, J. C. Genetic segregation of microsatellite markers in *Saccharum officinarum* and *S. spontaneum*. *Heredity*, London, v. 97, n. 5, p. 366-375, 2006.

ELSHIRE, R. J.; GLAUBITZ, J. C.; SUN, Q.; POLAND, J. A.; KAWAMOTO, K.; BUCKLER, E. S.; MITCHELL, S. E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *Plosone*, San Francisco, v. 6, n. 5, 2011.

FEDERER, W. T. Augmented (or Hoonuiaku) designs. *Hawaiian Planter's Records*, Aiea, v. 55, n. 5, p. 191-208, 1956.

FERREIRA, F. M.; BARBOSA, M. H.; CASTRO, R. D.; PETERNELLI, L. A.; CRUZ, C. D. Effects of inbreeding on the selection of sugar cane clones. *Crop Breeding and Applied Biotechnology*, Viçosa, v. 5, n. 4, p. 174-182, 2005.

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

GARCIA, A. A. F.; KIDO, E. A.; MEZA, A. N.; SOUZA, H. M. B.; PINTO, L. R.; PASTINA, M. M.; LEITE, C. S.; SILVA, J. A. G. DA; 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 and Applied Genetics*, Berlin, v. 112, n. 2, p. 298-314, nov. 2006.

GLASZMANN, J. C.; DUFOUR, P.; GRIVET, L.; D'HONT, A.; DEU, M.; PAULET, F.; HAMON, P. Comparative genome analysis between several tropical grasses. *Euphytica*, Amsterdam, v. 96, n. 1, p. 13-21, 1997.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis*

and *Eucalyptus urophylla* using a pseudo-testcross mapping strategy and RAPD markers. **Genetics**, Baltimore, v. 137, n. 4, p. 1121-1137, 1994.

GRIVET L.; ARRUDA, P. Sugarcane genomics: depicting the complex genome of an important tropical plant. **Current Opinion in Plant Biology**, London, v. 5, n. 2, p. 122-127, apr. 2001.

GRIVET, L.; D'HONT, A.; ROQUES, D.; FELDMANN, P.; LANAUD, C. E.; GLASZMANN, J. C. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): Genome Organization in a Highly Polyploid and Aneuploid interespecific Hybrid. **Genetics**, Baltimore, v. 142, n. 3, p. 987-1000, 1996.

GRIVET, L.; GLASZMANN, J. C.; D'HONT, A. Molecular evidences for sugarcane evolution and domestication. In: MOTLEY, T.; ZEREGA, N.; CROSS, H. (Ed). **Darwin's Harvest: New approaches to the origins, evolution, and conservation of crops**. New York: Columbia University Press, 2006. p. 49-66.

GUIMARAES, C. T.; HONEYCUTT, R. J.; SILLS, G. R.; SOBRAL, B. W. S. Genetic maps of *Saccharum officinarum* L. and *Saccharum robustum* Brandes & Jew. ex grassl. **Genetic Molecular Biology**, São Paulo, v. 22, n. 1, mar. 1999.

HALDANE, J. B. S. The combination of linkage values and the calculation of distance between the loci of linkage factors. **Journal of Genetics**, Oxford, v. 8, p. 299-309, 1919.

HALEY, C. S.; KNOTT, S. A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. **Heredity**, London, v. 69, p. 315-324, out. 1992.

HELLER-USZYNSKA, K.; USZYNSKI, G.; HUTTNER, E.; EVERS, M.; CARLIG, J.; CAIG, V.; AITKEN, K.; JACKSON, P.; PIPERIDIS, G.; COX, M.; GILMOUR, R.; D'HONT, A.; BUTTERFIELD, M.; GLASZMANN, J. C.; KILIAN A. Diversity Arrays Technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. **Molecular Breeding**, Netherlands, v. 28, n. 1, p. 37-55, jun. 2011.

HOARAU, J. Y.; GRIVET, L.; OFFMANN, B.; D'HONT, A.; RISTERUCCI, A. M.; ROQUES, D.; GLASZMANN, J. C.; GRIVET, L. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). II. Detection of QTLs for yield components. **Theoretical and Applied Genetics**, Berlin, v. 105, n. 6-7, p. 1027-1037, nov. 2002.

HOARAU, J. Y.; OFFMAN, B.; D'HONT, A.; RISTERUCCI, A. M.; ROQUES, D.; GLASZMANN, J. C.; GRIVET, L. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. **Theoretical and Applied Genetics**, Berlin, v. 103, n. 1, p. 84-97, jan. 2001.

HOGARTH, D. M. The effects of accidental selfing on the analysis of a diallel cross with sugar cane. **Euphytica**, Amsterdam, v. 29, n.1, p. 737-746, 1980.

HOTTA, C. T.; LEMBKE, C. G.; DOMINGUES, D. S.; OCHOA, E. A.; CRUZ, G. M. Q.; MELOTTO-PASSARIN, D. M.; MARCONI, T. G.; SANTOS, M. O.; SOUZA, A.;



MARGARIDO, G. A.; CRIVELLARI, A. C.; SANTOS, W. D.; HOSHINO, A. A.; CARRER, H.; SOUZA, A. P.; GARCIA, A. A. F.; BUCKERIDGE, M. S.; MENOSSE, M.; VANSLUYS, M. A.; SOUZA, G. M. The Biotechnology Roadmap for Sugarcane Improvement. **Tropical Plant Biology**, New York, v. 3, n. 2, p. 75-87, jun. 2010.

HULBERT, S. H.; ILOT, T. W.; EGG, E. J. L.; LINCOLNE, S. E.; LANDER, S.; MICHELMORE, R. W. Genetic analysis of the fungus *Bremia lactucae* using restriction fragment length polymorphisms. **Genetics**, Bethesda, v. 120, n. 4, p. 947-958, dez. 1988.

JACCOUD, D.; PENG, K.; FEINSTEIN, D.; KILIAN, A. Diversity arrays: a solid state technology for sequence information independent genotyping. **Nucleic Acids Research**, Oxford, v. 29, n. 4, jan. 2001.

JANNOO, N.; GRIVET, L.; SEGUIN, M.; PAULET, F.; DOMAINGUE, R.; RAO, P. S.; DOOKUN, A.; D'HONT, A.; GLASZMANN, J. C. Molecular investigation of the genetic base of sugarcane cultivars. **Theoretical and Applied Genetics**, Berlin, v. 99, n.1-2, p. 171-184, 1999.

JANSEN, R. C. A general mixture model for mapping quantitative trait loci by using molecular markers. **Theoretical Applied Genetics**, Berlin, v. 85, p. 252-260, 1992.

JANSEN, R. C. Interval mapping of multiple quantitative trait loci. **Genetics**, Baltimore, v. 135, n. 1, p. 205-211, set. 1993.

JIANG, C.; ZENG, Z. B. Multiple trait analysis of genetic mapping for quantitative trait loci. **Genetics**, Baltimore, v. 140, n. 3, p. 1111-1127, jul. 1995.

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

KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. **Science**, Washington, v. 220, n. 4598, p. 671-680, 1983.

KNOTT, S. A.; HALEY, C. S. Multitrait least squares for quantitative trait loci detection. **Genetics**, Baltimore, v. 156, n. 2, p. 899-911, out. 2000.

KOSAMBI, D. D. The estimation of map distances from recombination values. **Annals of Eugenics**, London, v. 12, n. 1, p. 172-175, jan. 1944.

LANDE, R.; THOMPSON, R. Efficiency of marker assisted selection in the improvement of quantitative traits. **Genetics**, Baltimore, v. 124, n. 3, p. 743-756, mar. 1990.

LANDELL, M. G. A.; PINTO, L. R.; CRESTE, S.; XAVIER, M. A.; ANJOS, I. A.; VASCONCELOS, A. C. M.; BIDÓIA, M. A. P.; SILVA, D. N.; SILVA, M. A. Seleção de novas variedades de cana-de-açúcar e seu manejo de produção. **Encarte Técnico de Informações Agrônomicas**, Piracicaba, v. 1, n. 110, 2005.

LANDER, E. S.; BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Baltimore, v. 121, n. 1, p. 185-199, jan. 1989.

LANDER, E. S.; GREEN, P.; ABRAHAMSON, J.; BARLOW, A.; DALY, M. J.; LINCON, S. E.; NEWBURG, L. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, Orlando, v. 1, n. 1, p. 174-181, 1987.

LIMA, M. L. A. **Análise da similaridade genética entre genótipos de cana-de-açúcar (*Saccharum* spp.) detectada por AFLP e coeficiente de parentesco**. 2000. 87 f. Dissertação (Mestrado em Genética e Biologia Molecular)-Universidade Estadual de Campinas, Campinas, 2000.

LIU, B. H. **Statistical genomics: Linkage, Mapping and QTL Analysis**. Boca Raton: CRC Press, 1998. 611 p.

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

MACE, E. S.; XIA, L.; JORDAN, D. R.; HALLORAN, K.; PARH, D. K.; HUTTNER, E.; WENZL, P.; KILIAN, A. DArT markers: diversity analyses and mapping in *Sorghum bicolor*. **BMC Genomics**, London, v. 9, n. 26, jan. 2008.

MACHADO, F. B. P. O açúcar nos 500 anos do Brasil. **Revista STAB**, Piracicaba, v. 18, n. 5, p. 4-15, 2000.

MALIEPAARD, C.; JANSEN, J.; VAN OOIJEN, J. W. Linkage analysis in a full-sib family of an multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). **Euphytica**, Wageningen, v. 161, n. 1, p. 241-257, 2008.

MATSUOKA, S.; ARIZONO, H.; MASUDA, Y. Variedades de cana: minimizando riscos na adoção. **Revista STAB**, Piracicaba, v. 17, n. 2, p. 18-19, 1999.

MATSUOKA, S.; GARCIA, A. A. F.; ARIZONO, H. Melhoramento da cana-de-açúcar. In: BORÉM, A. (Ed.) **Melhoramento de espécies cultivadas**. Viçosa: Editora UFV, 2005, cap. 7, p. 225-274.

MARGARIDO, G. R. A.; SOUZA, A. P.; GARCIA, A. A. F. OneMap: software for genetic mapping in outcrossing species. **Hereditas**, Lund, v. 144, n. 3, p. 78-79, 2007.

MARGARIDO, G. R. A. **Mapeamento de QTLs em múltiplos caracteres e ambientes em um cruzamento comercial de cana-de-açúcar usando modelos mistos**. 2011. 108 f. Tese (Doutorado em Ciências)—Universidade de São Paulo, Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2011.

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

McINTYRE, C. L.; WHAN, V. A.; CROFT, B.; MAGAREY, R.; SMITH, G. R. Identification and validation of molecular markers associated with *Pachymetra* root rot and brown rust resistance in sugarcane using map- and association-based approaches. **Molecular Breeding**, Netherlands, v. 16, n. 2, p. 151-161, set. 2005.

MICHELMORE, R. W.; PARAN, L.; KESSELI, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. **Proceedings of the National Academy of Sciences**, Washington, v. 88, n. 21, p. 9828-9832, nov. 1991.

MING, R.; DEL MONTE, T. A.; HERNANDEZ, E.; MOORE, P. H.; IRVINE, J. E.; PATERSON, A. H. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploidy genomes. **Genome**, Ottawa, v. 45, n. 5, p. 794-803, out. 2002a.

MING, R.; LIU, S.C.; LIN, Y. R.; DA SILVA, J.; 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**, Baltimore, v. 150, n. 4, p. 1663-1682, dez. 1998.

MING, R.; LIU, S. C.; MOORE, P. H.; IRVINE, J. E.; PATERSON, A. H. Qtl analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. **Genome Research**, New York, v. 11, n. 12, p. 2075-2084, dez. 2001.

MING, R.; WANG, Y. W.; DRAYE, X.; MOORE, P. H.; IRVINE, J. E.; PATERSON, A. H. Molecular dissection of complex traits in autopolyploids: mapping QTL's affecting sugar yield and related traits in sugarcane. **Theoretical and Applied Genetics**, Berlin, v. 105, n. 2-3, p. 332-345, mai. 2002b.

MIOCQUE, J. Origem e história da cana Caiana. **Álcool & Açúcar**, São Paulo, v. 65, n. 1, p. 18-24, jan. 1993.

MISSIAGGIA, A.; GRATTAPAGLIA, D. Plant microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers. **Genetics and Molecular Research**, Ribeirão Preto, v. 5, n. 1, p. 72-78, 2006.

MORAES, M. F.; BASTOS, G. Q.; ANUNCIAÇÃO-FILHO, C. J.; MELO, L. J. T.; REIS, O. V. Avaliação agroindustrial e parâmetros genéticos de progênies de cana-de-açúcar em fase inicial na zona canavieira do litoral norte de Pernambuco. **Ciência e agrotecnologia**, Lavras, v. 34, n. 5, p. 1086-1092, set-out, 2010.

MUDGE, J.; ANDERSEN, W. R.; KEHRER, R. L.; FAIRBANKS, D. J. A RAPD Genetic map of *Saccharum officinarum*. **Crop Science**, Madison, v. 36, n. 5, p. 1362-1366, set. 1996.

NORDBORG, M.; TAVARÉ, S. Linkage disequilibrium: what history has to tell us. **Trends Genetics**, Amsterdam, v. 18, n. 2, p. 83-89, fev. 2002.

OLIVEIRA, K. M. **Desenvolvimento de marcadores moleculares EST-SSRs e mapeamento funcional em cana-de-açúcar (*Saccharum* spp.)**. 2006. 186 f. Tese (Doutorado em Genética e Biologia Molecular) - Universidade Estadual de Campinas, Campinas, 2006.

OLIVEIRA, K. M.; PINTO, L. R.; MARCONI, T. G.; MARGARIDO, G. R. A.; PASTINA, M. M.; TEIXEIRA, L. H. M.; FIGUEIRA, A. V.; ULIAN, E. C.; GARCIA, A. A. F.; SOUZA, A. P. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding**, Netherlands, v. 20, n. 3, p. 189-208, abr. 2007.

OTT, J. **Analysis of human genetic linkage**. Baltimore: Johns Hopkins, 1991. 321 p.

PALHARES, A. C. **Mapeamento genético de marcadores AFLP e de retrotransposons em cana-de-açúcar (*Saccharum* spp.)**. 2010. 98 f. Dissertação (Mestrado em Ciências: Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2010.

PARIDA, S. K.; KALIA, S. K.; KAUL, S.; DALAL, V.; HEMAPRABHA, G.; SELVI, A.; PANDIT, A.; SINGH, A.; GAIKWAD, K.; SHARMA, T. R.; SRIVASTAVA, P. S.; SINGH, N. K.; MOHAPATRA, T. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. **Theoretical and Applied Genetics**, Stuttgart, v. 118, n. 2, p. 327-338, jan. 2009

PASTINA, M.; MALOSETTI, M.; GAZAFFI, R.; MOLLINARI, M.; MARGARIDO, G.; OLIVEIRA, K.; PINTO, L.; SOUZA, A.; VAN EEUWIJK, F.; GARCIA, A. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theoretical and Applied Genetics**, Berlin, v. 124, n. 5, p. 835-849, mar. 2012.

PASTINA, M. M.; PINTO, L. R.; OLIVEIRA, K. M.; SOUZA, A. P.; GARCIA, A. A. F. Molecular mapping of complex traits. In: HENRY, R. J.; KOLE, C. (Ed.). **Genetics genomics and breeding of sugarcane**. Clemson: CRC Press, 2010. cap. 7, p. 117-148.

PINTO, L.; GARCIA, A.; PASTINA, M.; TEIXEIRA, L.; BRESSIANI, J.; ULIAN, E.; BIDOIA, M.; SOUZA, A. Analysis of genomic and functional RFLP derived markers associated with sucrose content, fiber and yield QTLs in a sugarcane (*Saccharum* spp.) commercial cross. **Euphytica**, Netherlands, v. 172, n. 3, p. 313-327, abr. 2010.

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

PIPERIDIS, N.; JACKSON, P. A.; D'HONT, A.; BESSE, P.; HOARAU, J. Y.; COURTOIS, B.; AITKEN, K. S.; MCINTYRE, C. L. Comparative genetics in sugarcane enables structured map enhancement and validation of market-trait associations. **Molecular Breeding**, Netherlands, v. 21, n. 2, p. 233-247, fev. 2008.

PROCANA - JORNAL CANA. **Os impressionantes números do setor sucroenergético (Safrá 2009/10)**. Agosto de 2011. Disponível em: <<http://www.jornalcana.com.br/Conteudo/Conheca%20o%20Setor.asp>>. Acesso em: 12 jan. 2012.

QU, L; HANCOCK, J. F. Detecting and mapping repulsion phase linkage in polyploids with polysomic inheritance. **Theoretical and Applied Genetics**, Berlin, v. 103, n. 1, p. 136-143, jul. 2001.

R Development Core Team (2012). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

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

REFFAY, N.; JACKSON, P. A.; AITKEN, K. S.; HOARAU, J. Y.; D'HONT, A.; BESSE, P.; McINTYRE, C. L. Characterization of genome regions incorporated from an important wild relative into Australian sugarcane. **Molecular Breeding**, Netherlands, v. 15, n. 4, p. 367-381, mai. 2005.

ROACH, B. T. Sugarcane. In: SIMMONDS, N. W. (Ed). **Evolution of crop plants**. London: Longman Group Limited, 1979. 104-108 p.

ROSSI, M.; ARAÚJO, P. G.; PAULET, F.; GARSMEUR, P.; DIAS, V. M.; CHEN, H.; VAN SLUYS, M. A.; D'HONT, A. Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. **Molecular Genetic Genomics**, Berlin, v. 269, n. 3, p. 406-419, jun. 2003.

SANSALONI, C.P. **Desenvolvimento e aplicações de DArT (Diversity Array Technology) e genotipagem por sequenciamento (Genotype-by-Sequencing) para Eucalyptus**. 2011. 121 f. Defesa de qualificação (Doutorado em Biologia Molecular) – Universidade de Brasília, Brasília. 2011.

SCHIFINO-WITTMANN, M. T. Poliploidia e seu impacto na origem e evolução das plantas silvestres e cultivadas. **Revista Brasileira de Agrociência**, Pelotas, v. 10, n. 2, p. 151-157, abr/jun. 2004.

SCHUSTER, I.; CRUZ, C. D. **Estatística genômica aplicada a populações derivadas de cruzamentos controlados**. Viçosa: Editora UFV, 2008. 568 p.

SILLS, G. R.; BRIDGES, W.; AL-JANABI, S. M.; SOBRAL, B. W. S. Genetic analysis of agronomic traits in a cross between sugarcane (*Saccharum officinarum* L.) and its presumed progenitor (*S. robustum* Brandes & Jesw. ex Grassl. **Molecular Breeding**, Netherlands, v. 1, n. 4, p. 355-363, dez. 1995.

SILVA, D. G. **Mapeamento genético de marcadores DArT (Diversity Arrays Technology) em cana-de-açúcar (Saccharum spp.)**. 2012. 76 f. Dissertação

(Mestrado em Genética e Melhoramento de Plantas)–Universidade Federal de Goiás, Goiânia. 2012.

SILVA, M. A.; JERONIMO, E. M; LUCIO, A. D. C. Perfilamento e produtividade de cana-de-açúcar com diferentes alturas de corte e épocas de colheita. **Pesquisa Agropecuária Brasileira**, Brasília, v. 43, n. 8, p. 979-986, ago. 2008.

SILVA, M. B.; MORAIS, A. S. Avaliação energética do bagaço de cana em diferentes níveis de umidade e graus de compactação. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO - A integração de cadeias produtivas com a abordagem da manufatura sustentável, 28., 2008, Rio de Janeiro. **Resumos...** Rio de Janeiro: ABEPRO, 2008. p. 9.

SILVA, P. P. **Divergência genética em genótipos de cana-de-açúcar (*Saccharum* spp.) através de caracteres morfoagronômicos e por marcadores moleculares**. 2006. 114 f. Dissertação (Mestrado em Agronomia: Produção Vegetal)-Centro de Ciências Agrárias, Universidade Federal de Alagoas, Rio Largo, 2006.

TANKSLEY, S. D. Mapping polygenes. **Annual Review Genetics**, Palo Alto, v. 27, p. 205-233, 1993.

TEIXEIRA, L. H. M. **Mapeamento funcional em cana-de-açúcar utilizando ESTs como marcadores moleculares**. 2006. 109 f. Dissertação (Mestrado em Biologia)–Universidade Estadual de Campinas, Campinas, 2006.

THOMPSON, E. A. Crossover counts and likelihood in multipoint linkage analysis. **IMA- Journal of Mathematical Applied Medicine Biology**, Oxford, v. 4, n. 2, p. 93-108, 1987.

VAN OOIJEN, J. W. **MapQTL 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species**. Netherlands: Kyazma BV. 2009.

VAN OOIJEN, J. W.; VOORRIPS, R. E. JoinMap Version 3.0 Software for the calculation of genetics linkage maps. **Plant Research International**, Wageningen, 2001.

VOORRIPS, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, New York, v. 93, n. 1, p.77-78, jan/fev. 2002.

WENZL, P.; CARLING, J.; KUDRNA, D.; JACCOUD, D.; HUTTNER, E.; KLEINHOF, A.; KILIAN, A. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. **Proc Natl Acad Sci USA**, v. 101, n. 26, p. 9915-9920, Jun. 2004.

WEI, X.; JACKSON, P. A.; McINTYRE, C. L.; AITKEN, K. S.; CROFT, B. Associations between DNA markers and resistance to diseases in sugarcane and effects of populations substructure. **Theoretical and Applied Genetics**, New York,

v. 114, n. 1, p. 155-164, 2006.

WEIR, B. **Genetic Data Analysis II**. 2. ed. Sunderland: Sinauer Associates, 1996. 445 p.

WU, K. K.; BURNQUIST, W.; SORRELLS, M. E.; TEW, T. L.; MOORE, P. H.; TANKSLEY, S. D. The detection and estimation of linkage in polyploids using single-dose restriction fragments. **Theoretical Applied Genetic**, Berlin, v. 83, n. 3, p. 294-300, jan. 1992.

WU, R.; MA, C. X.; PAINTER, I.; ZENG, Z. B. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. **Theoretical Population Biology**, New York, v. 61, n. 3, p. 349-363, 2002.

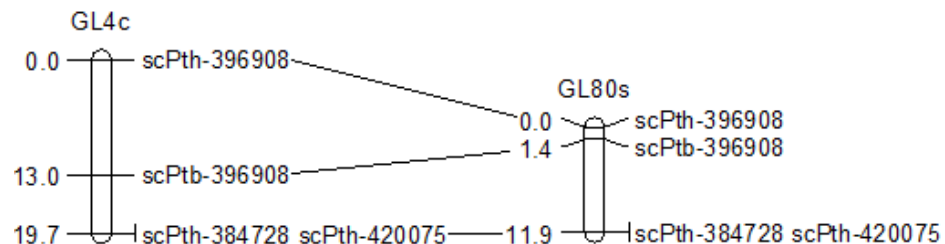
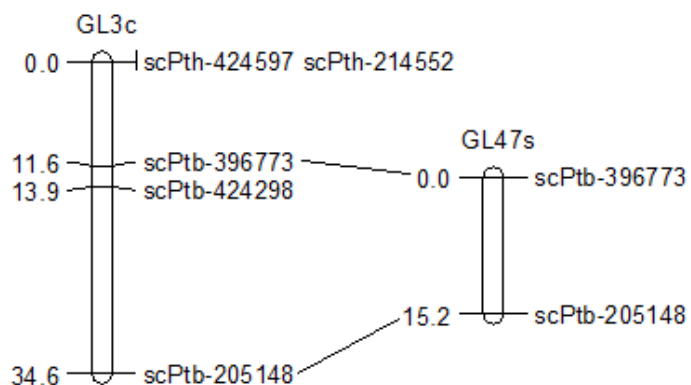
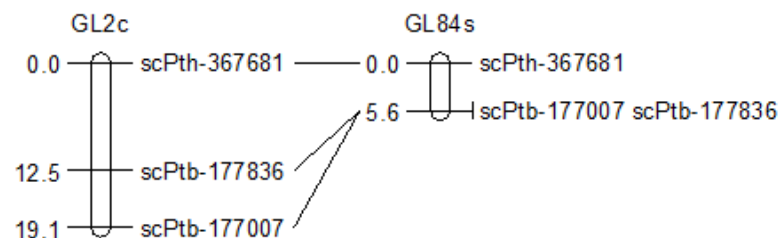
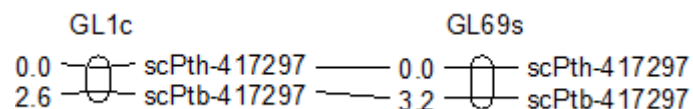
XIE, Y.; MCNALLY, K.; LI, C. Y.; LEUNG, H.; ZHU, Y. Y. A High-throughput Genomic Tool: Diversity Array Technology Complementary for Rice Genotyping. **Journal of Integrative Plant Biology**, v. 48, n. 9, p. 1069-1076, Set. 2006.

XU, Y. **Molecular Plant Breeding**. Oxfordshire: CABI, 2010. 755 p.

ZENG, Z. B. Precision mapping of quantitative trait loci. **Genetics**, Baltimore, v. 136, n. 4, p. 1457-1468, abr. 1994.

ZENG, Z. B. Theoretical basis for separation of multiple linked gene effects in mapping of quantitative trait loci. **Proceedings of the National Academy of Sciences**, Washington, v. 90, n. 23, p. 10972-10976, dez. 1993.

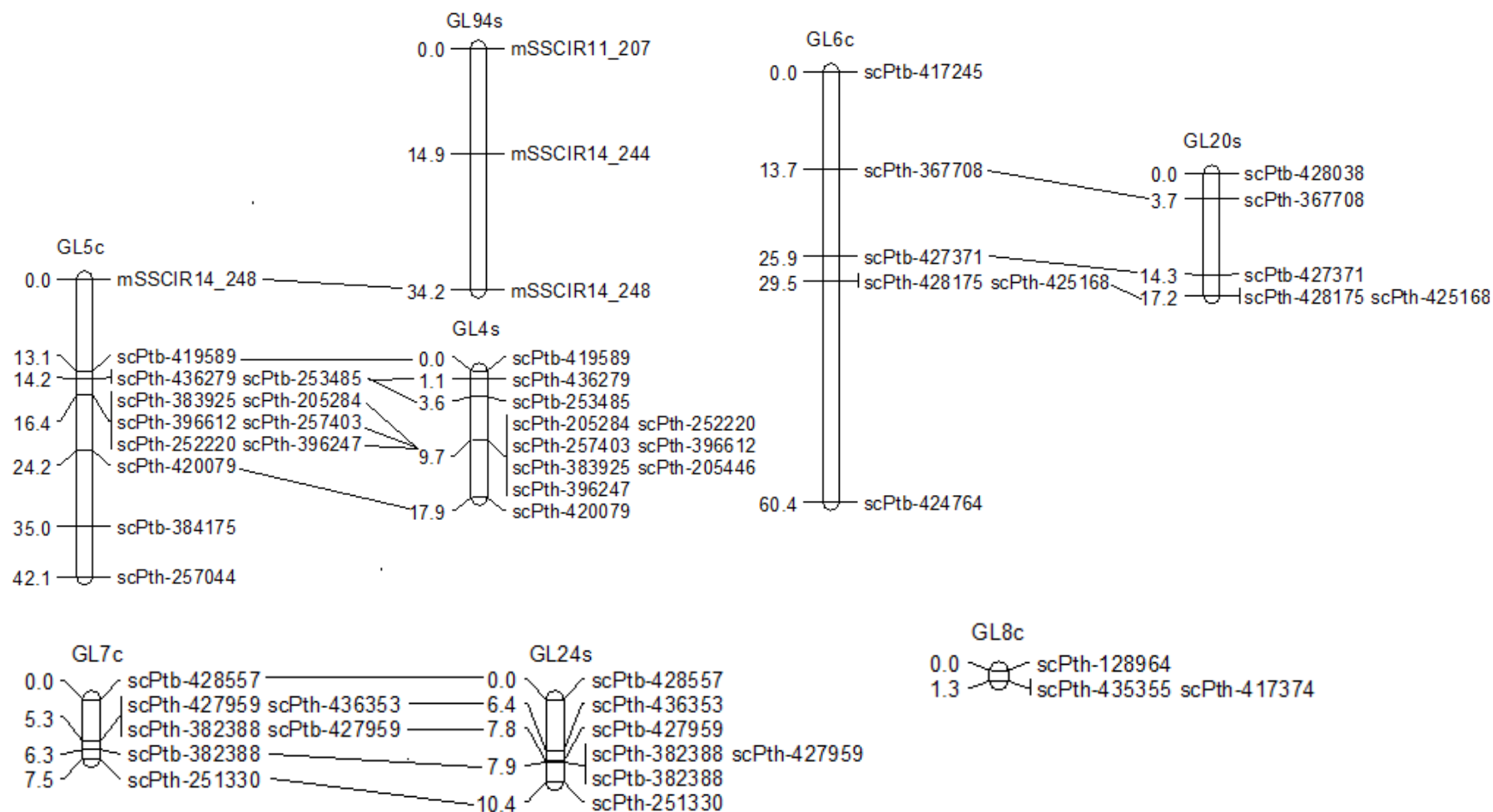
**Apêndice A.** Mapa de ligação integrado desenvolvido a partir da população derivada do cruzamento RB97327 x RB72454 (c, “cross”), e mapa de ligação referente ao genitor RB97327, obtido a partir de uma população derivada de autofecundação (s, “selfing”). À direita está localizada a identificação dos marcadores e à esquerda as distâncias genéticas em cM, estimadas pela função de Kosambi.



Continua...

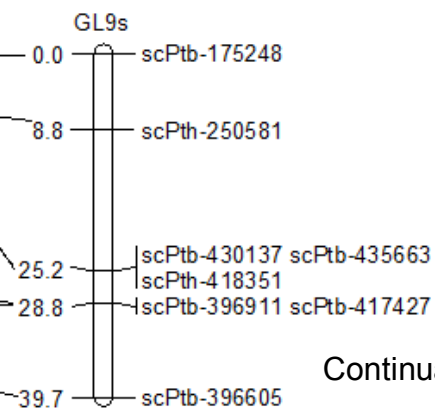
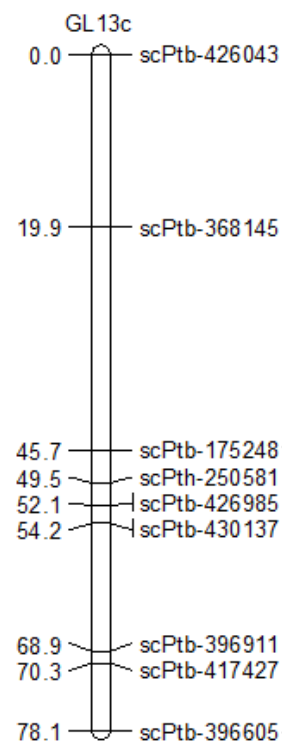
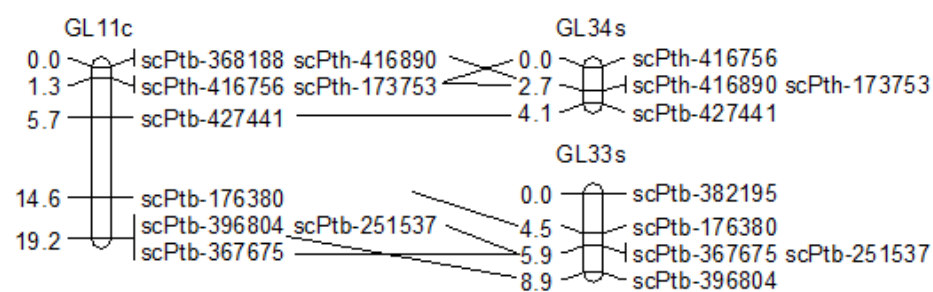
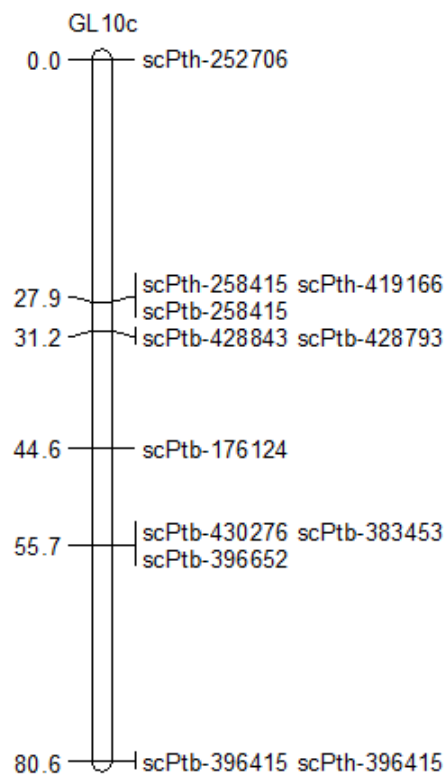
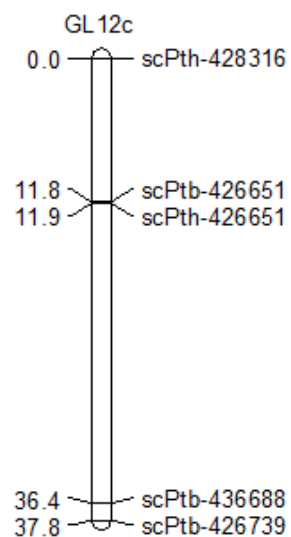
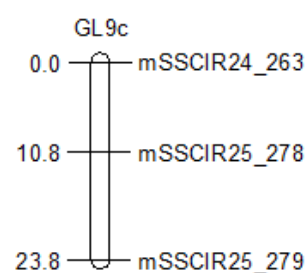


## Apêndice A. Continuação.



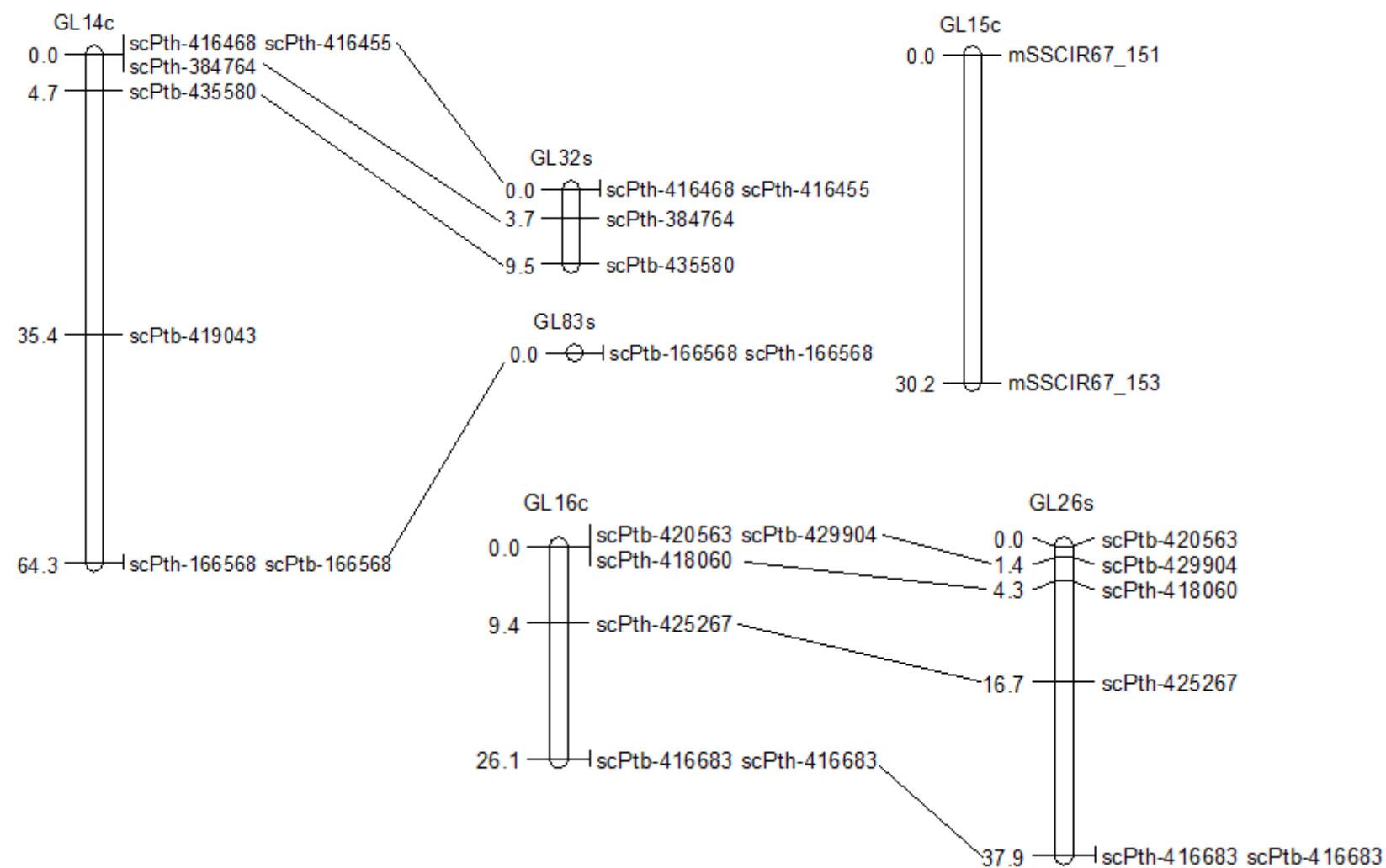
Continua...

## Apêndice A. Continuação.



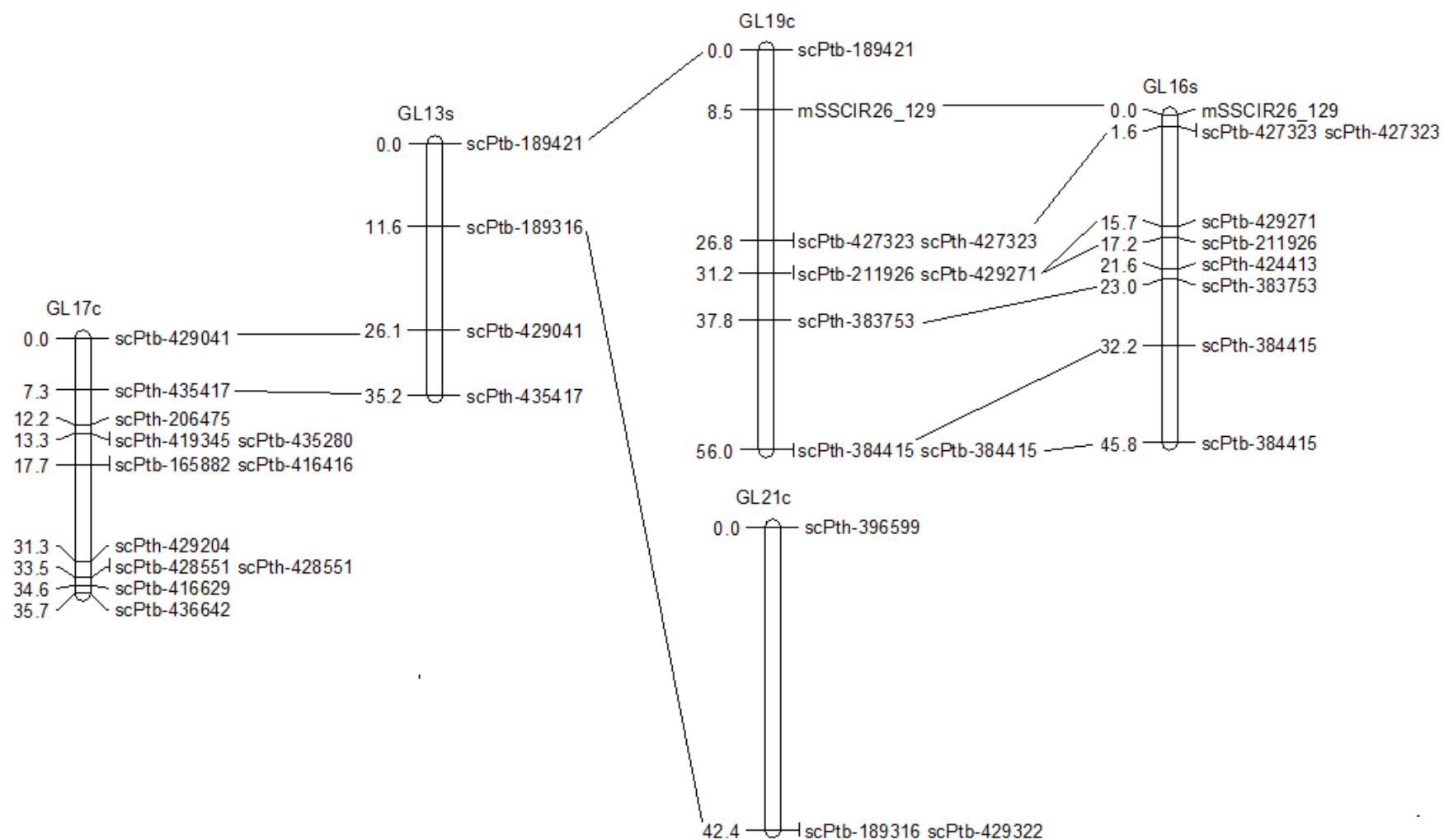
Continua...

## Apêndice A. Continuação.



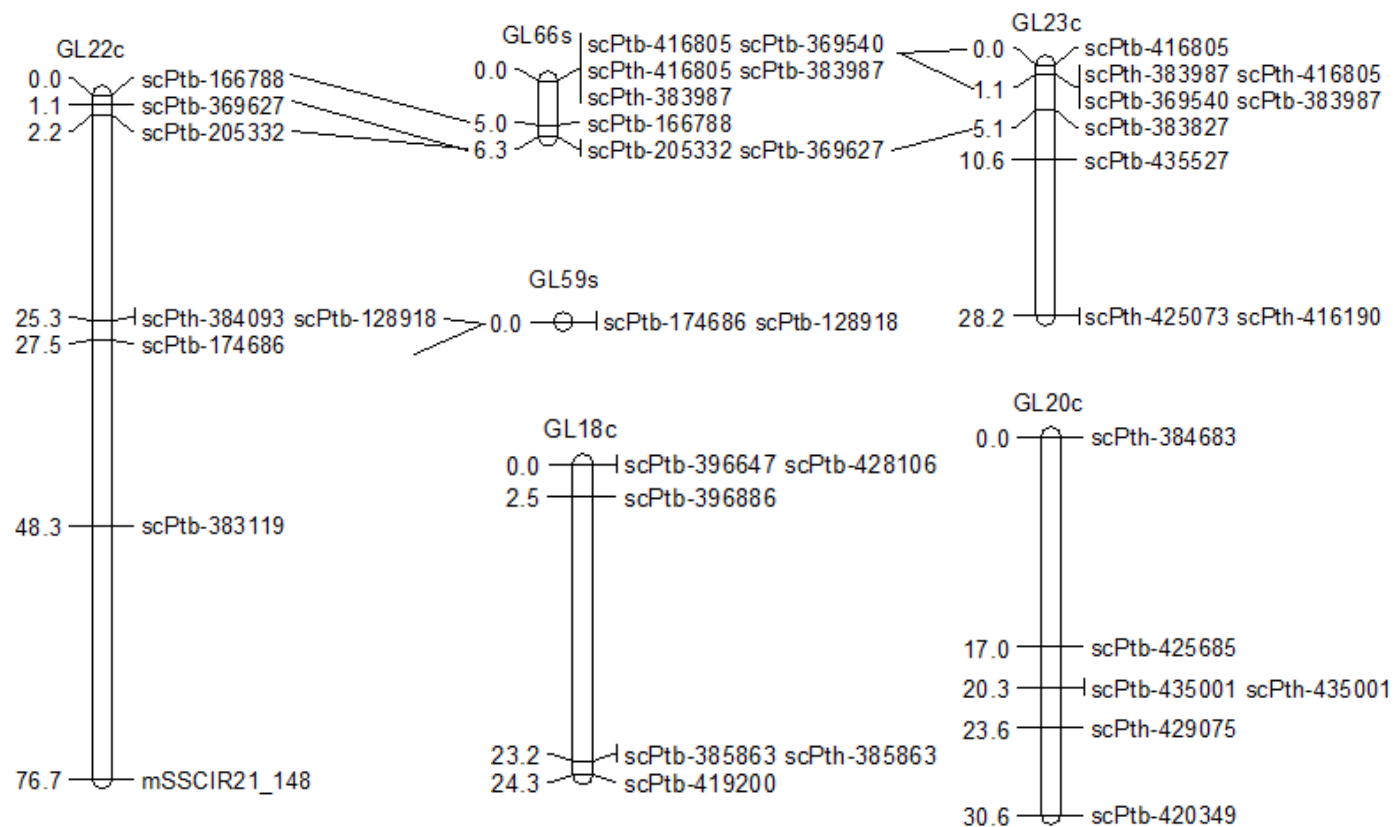
Continua...

## Apêndice A. Continuação.



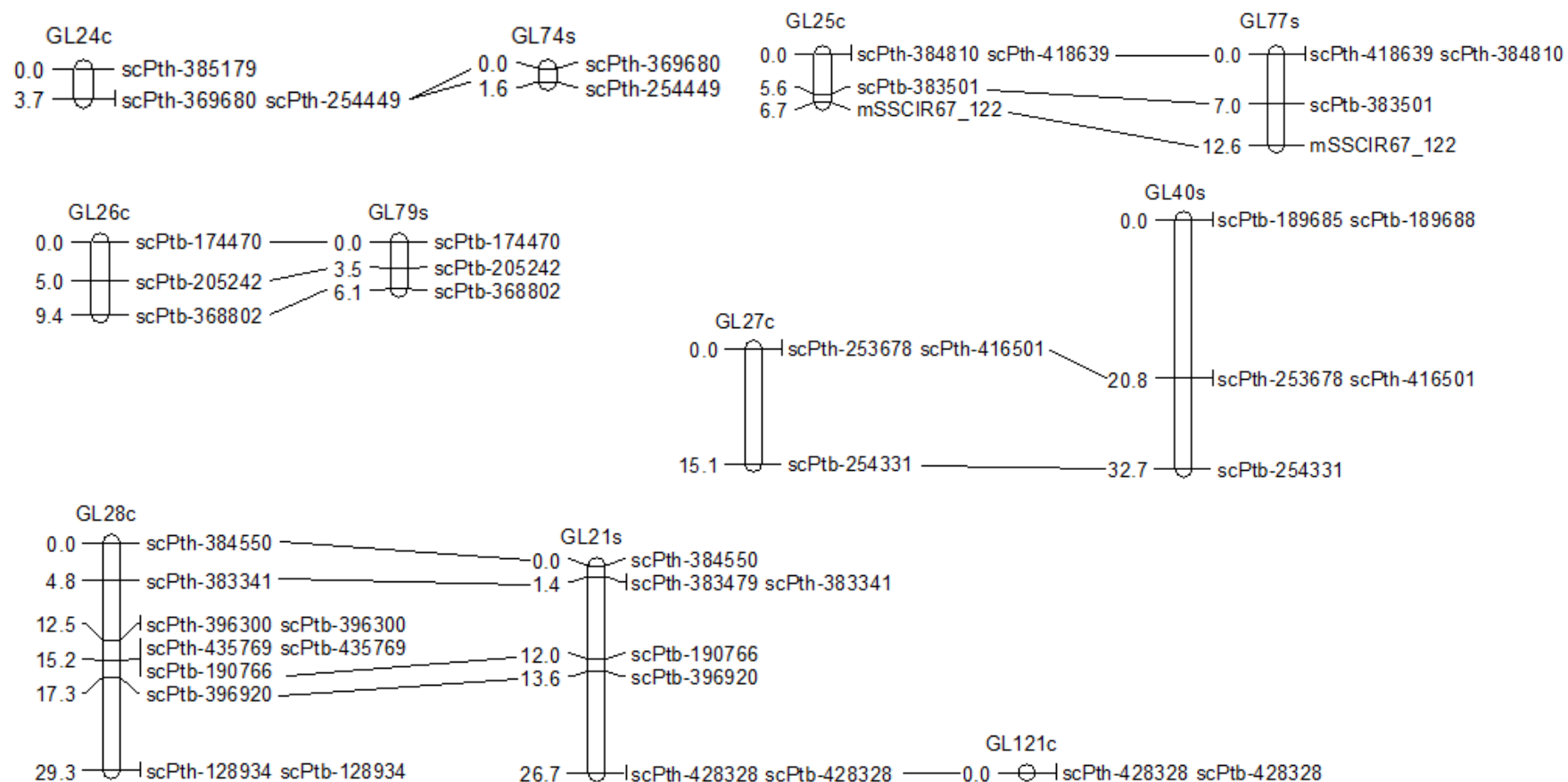
Continua...

## Apêndice A. Continuação.



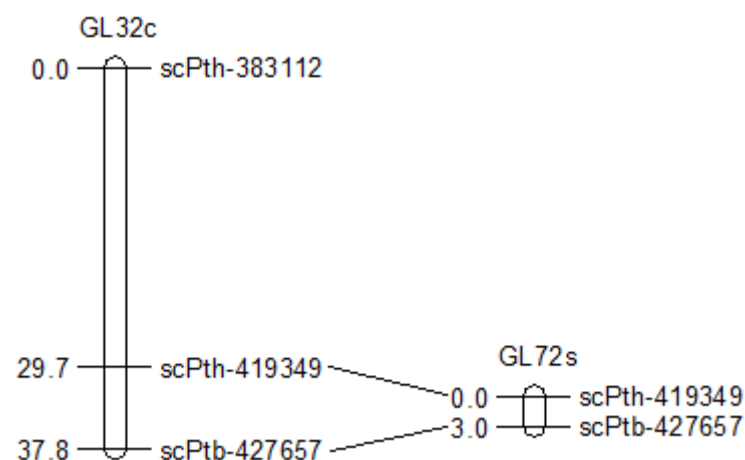
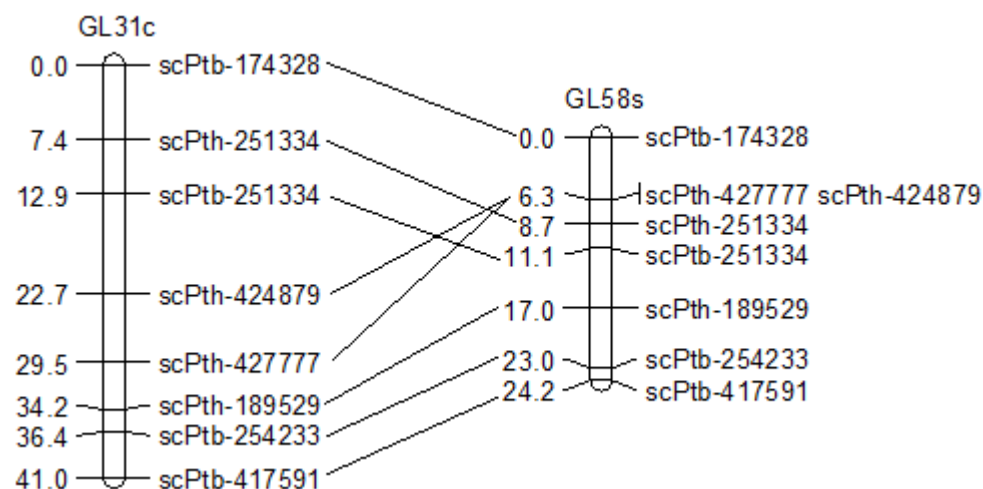
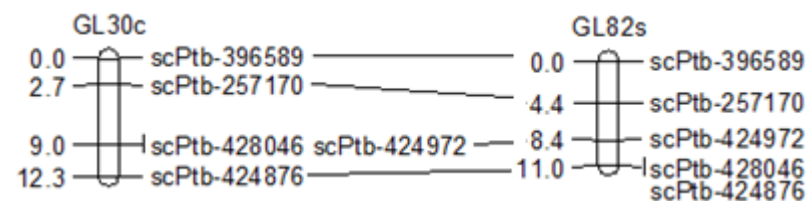
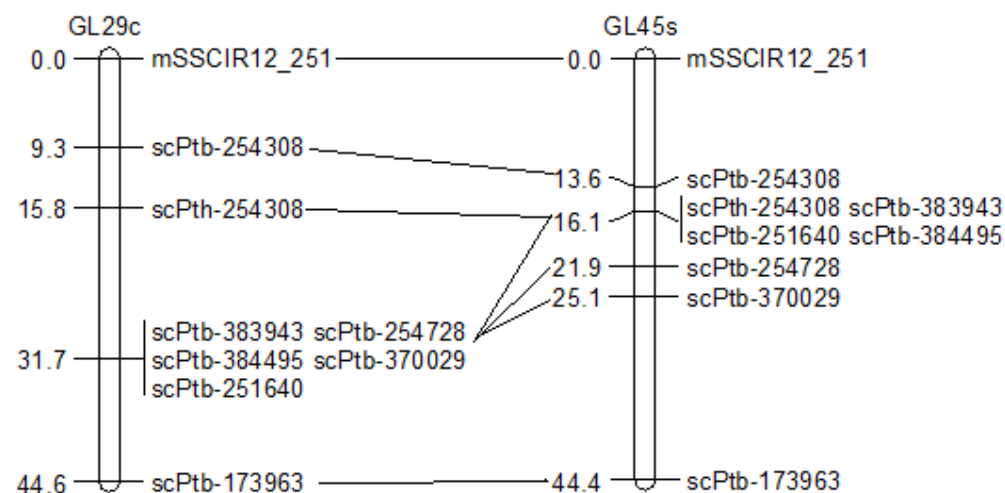
Continua...

## Apêndice A. Continuação.



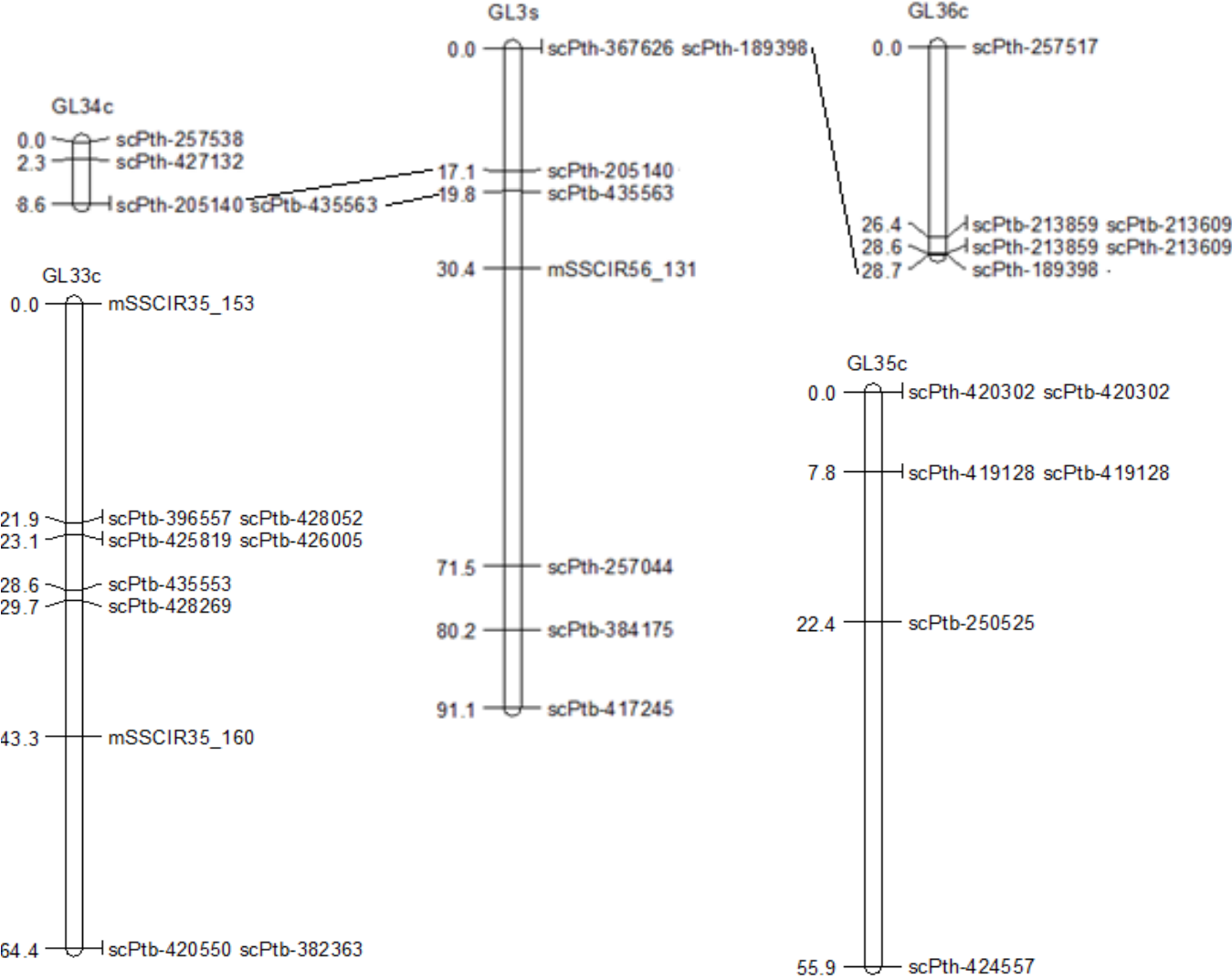
Continua...

## Apêndice A. Continuação.



Continua...

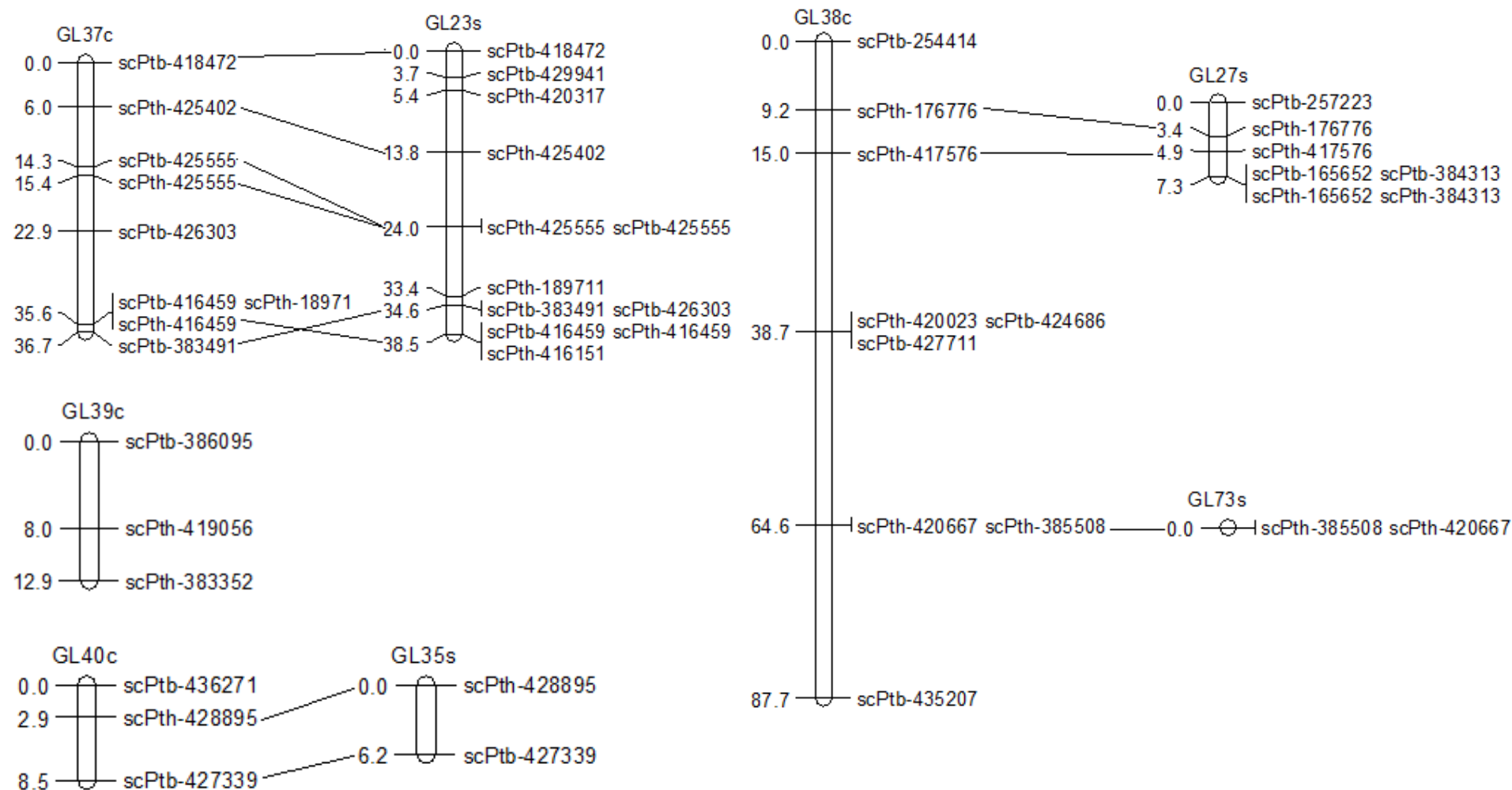
Apêndice A. Continuação.



Continua...

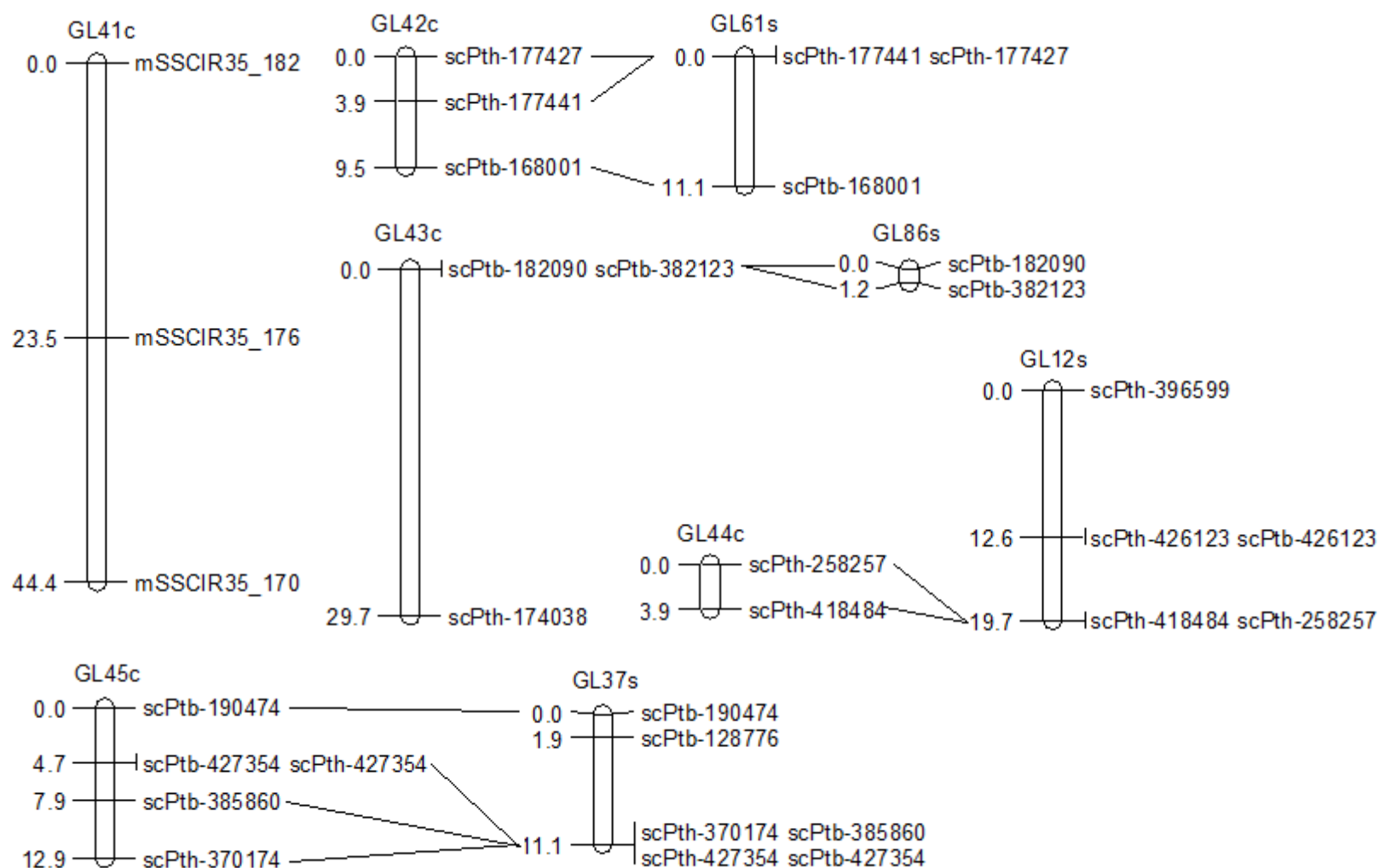


## Apêndice A. Continuação.



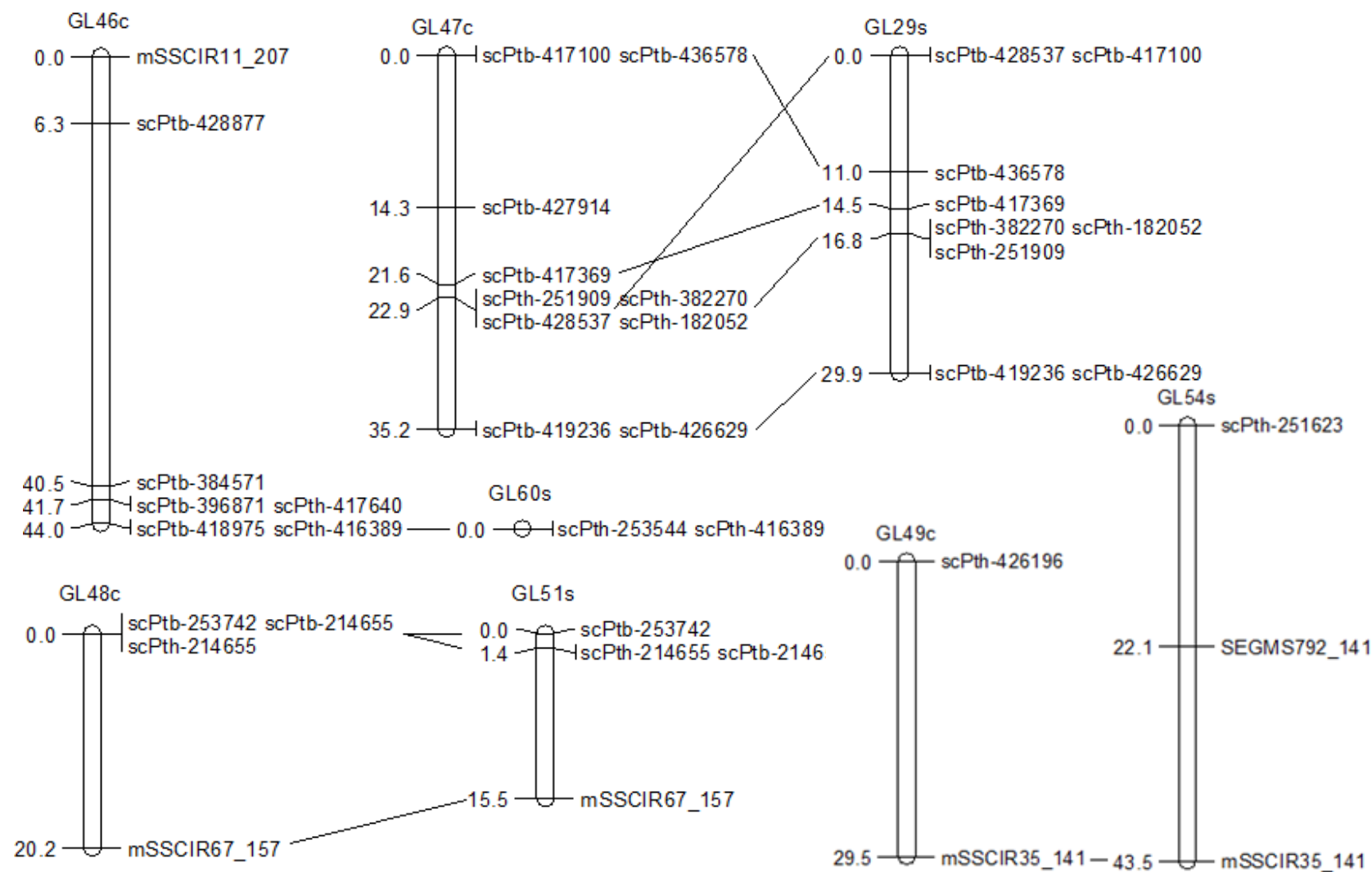
Continua...

## Apêndice A. Continuação.



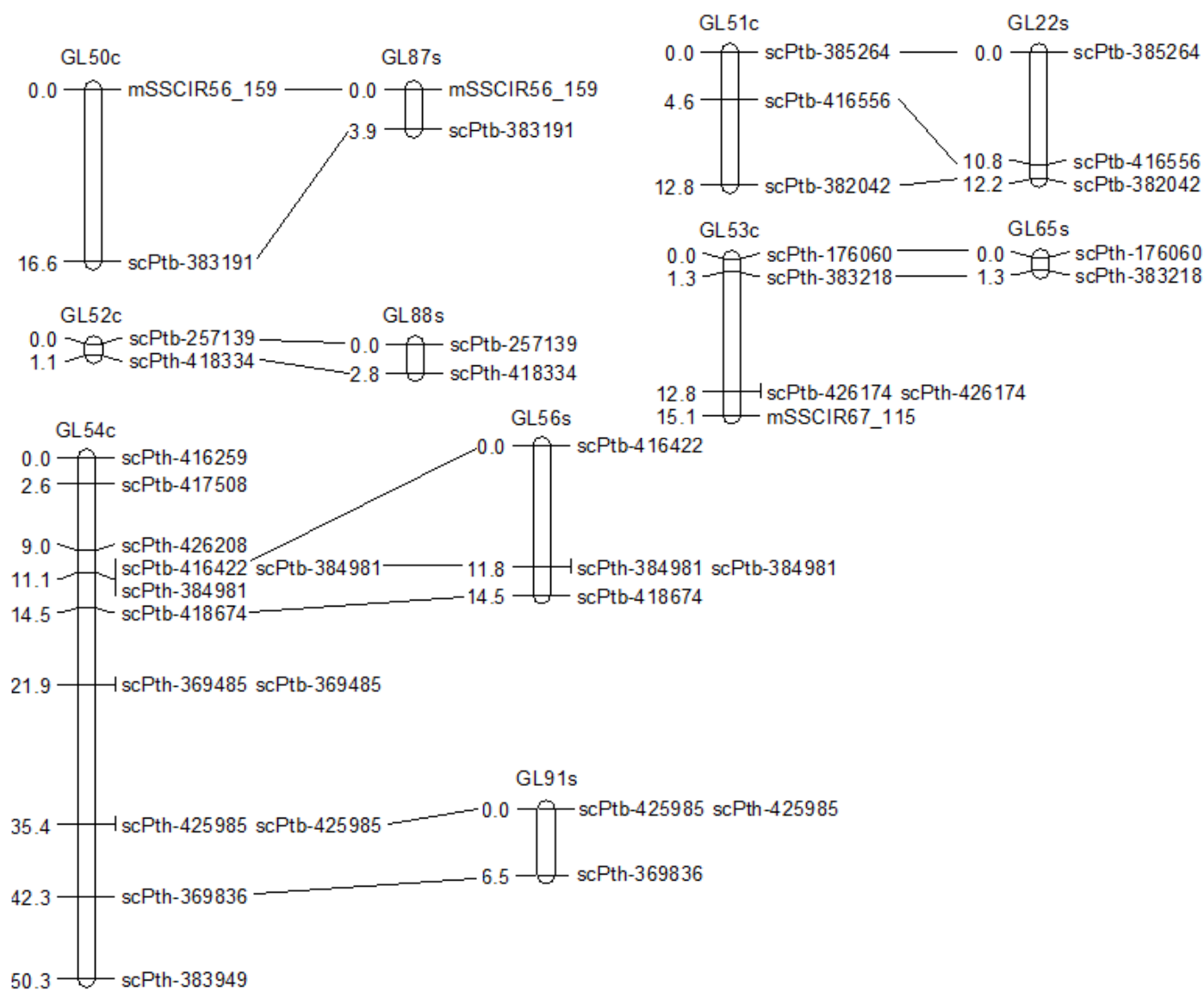
Continua...

## Apêndice A. Continuação.



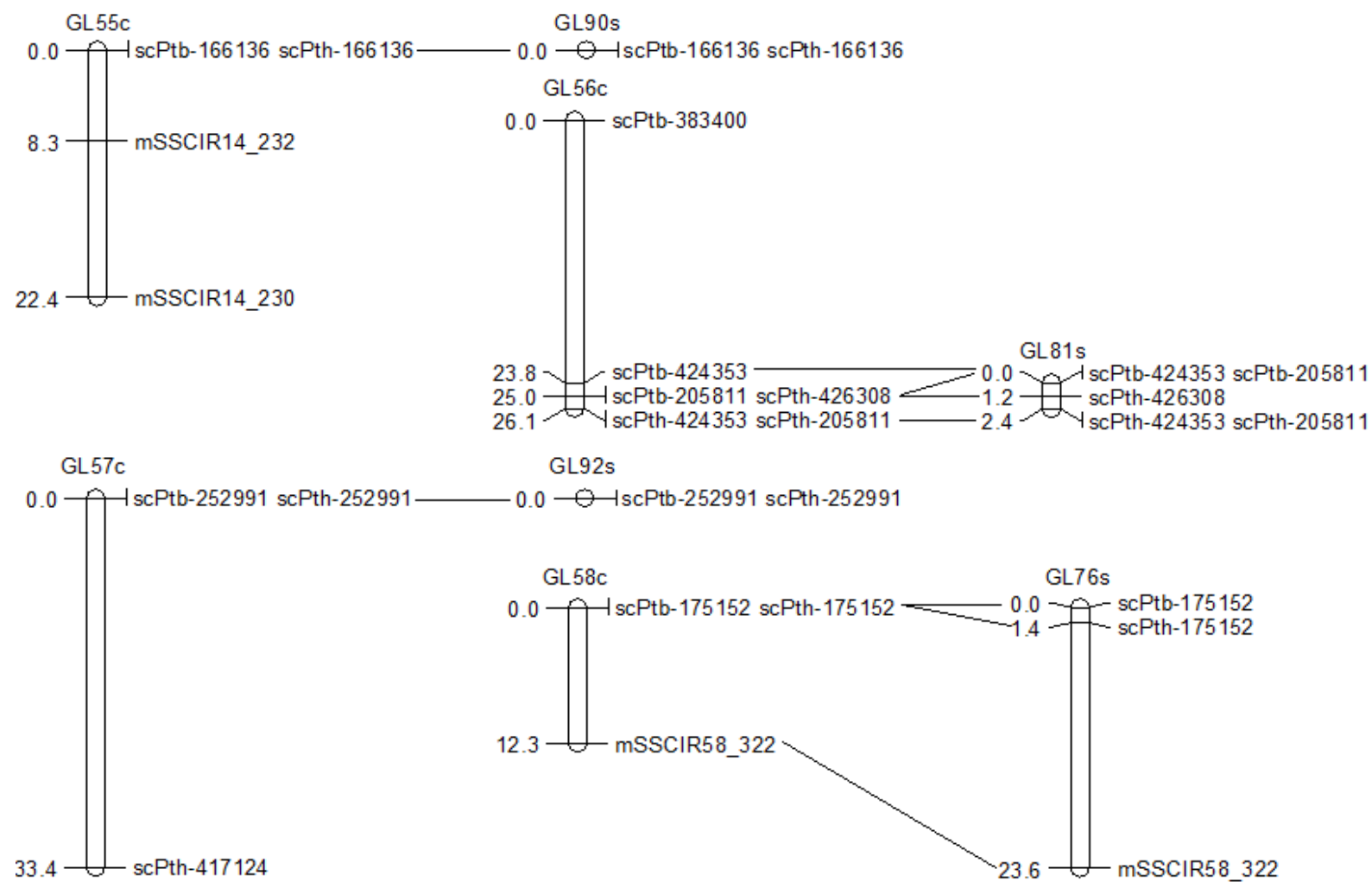
Continua...

## Apêndice A. Continuação.



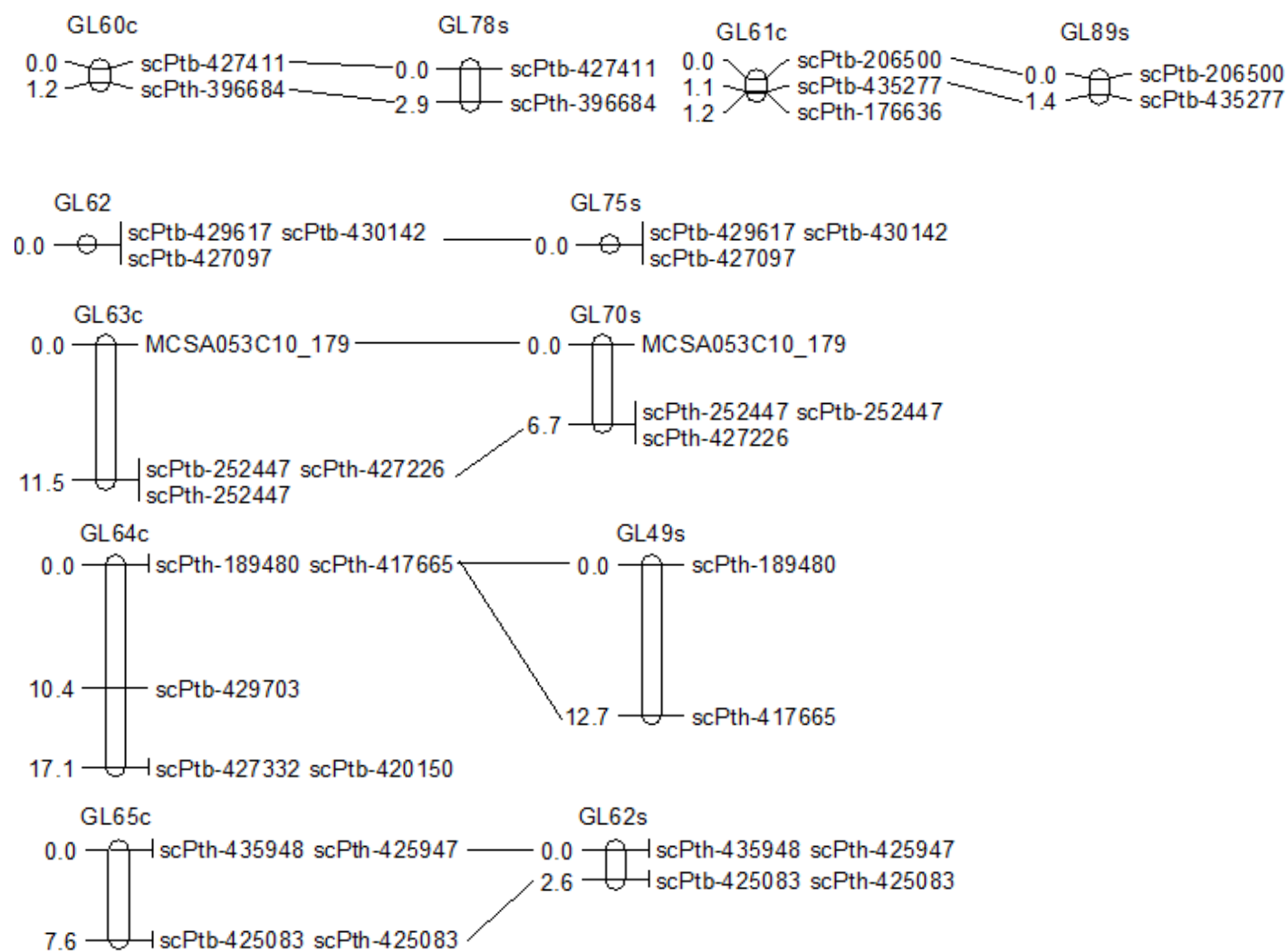
Continua...

## Apêndice A. Continuação.



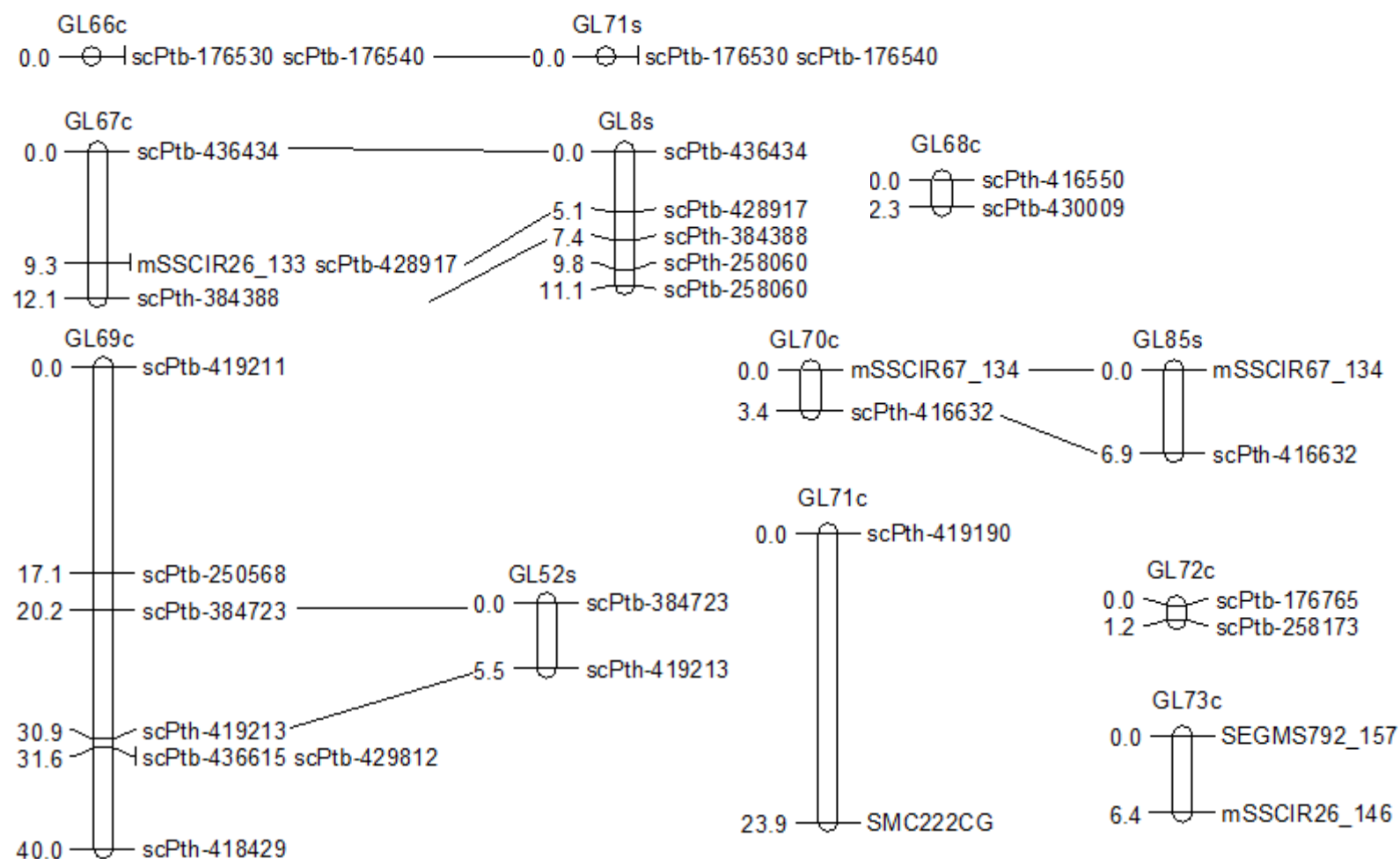
Continua...

## Apêndice A. Continuação.



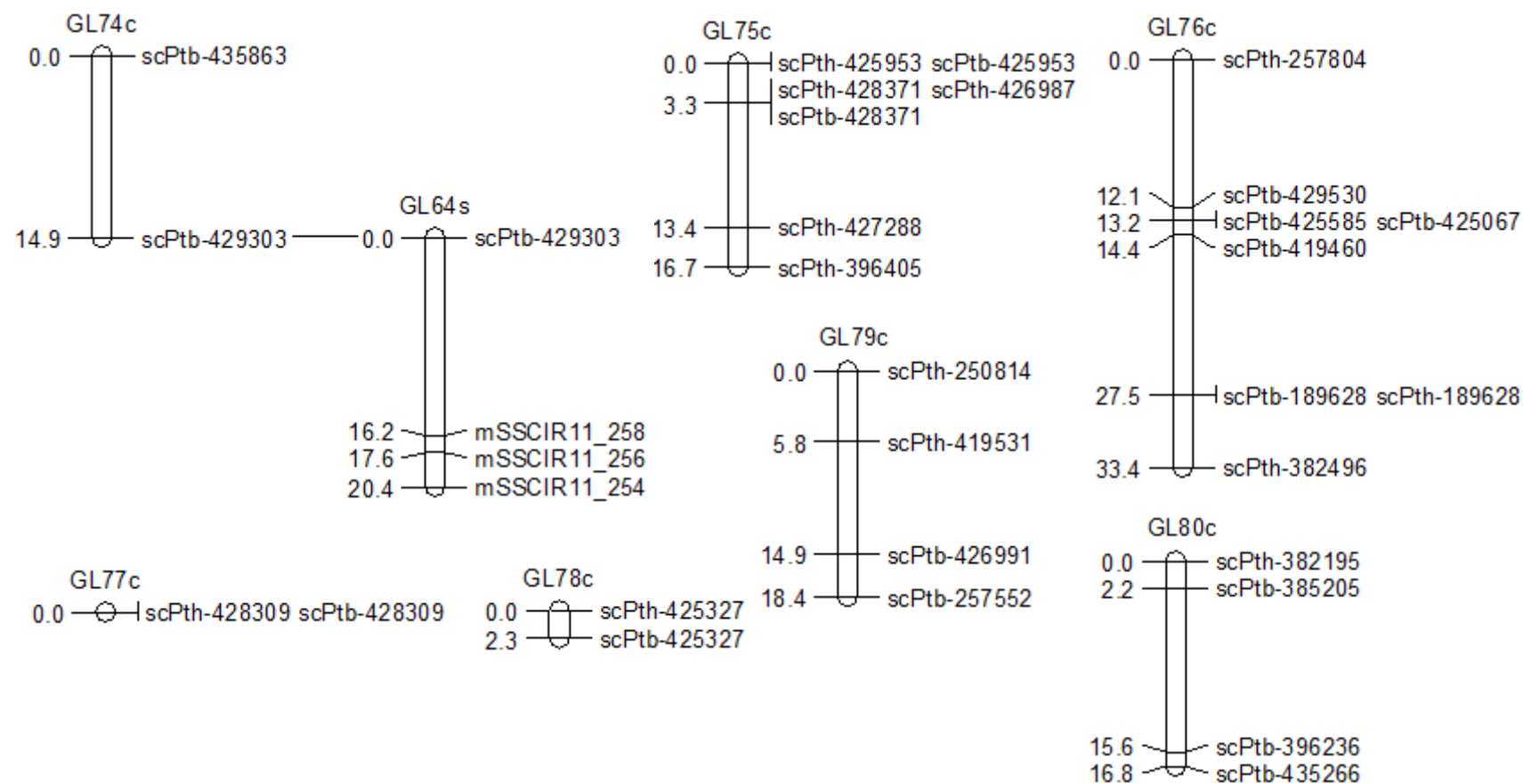
Continua...

## Apêndice A. Continuação.



Continua...

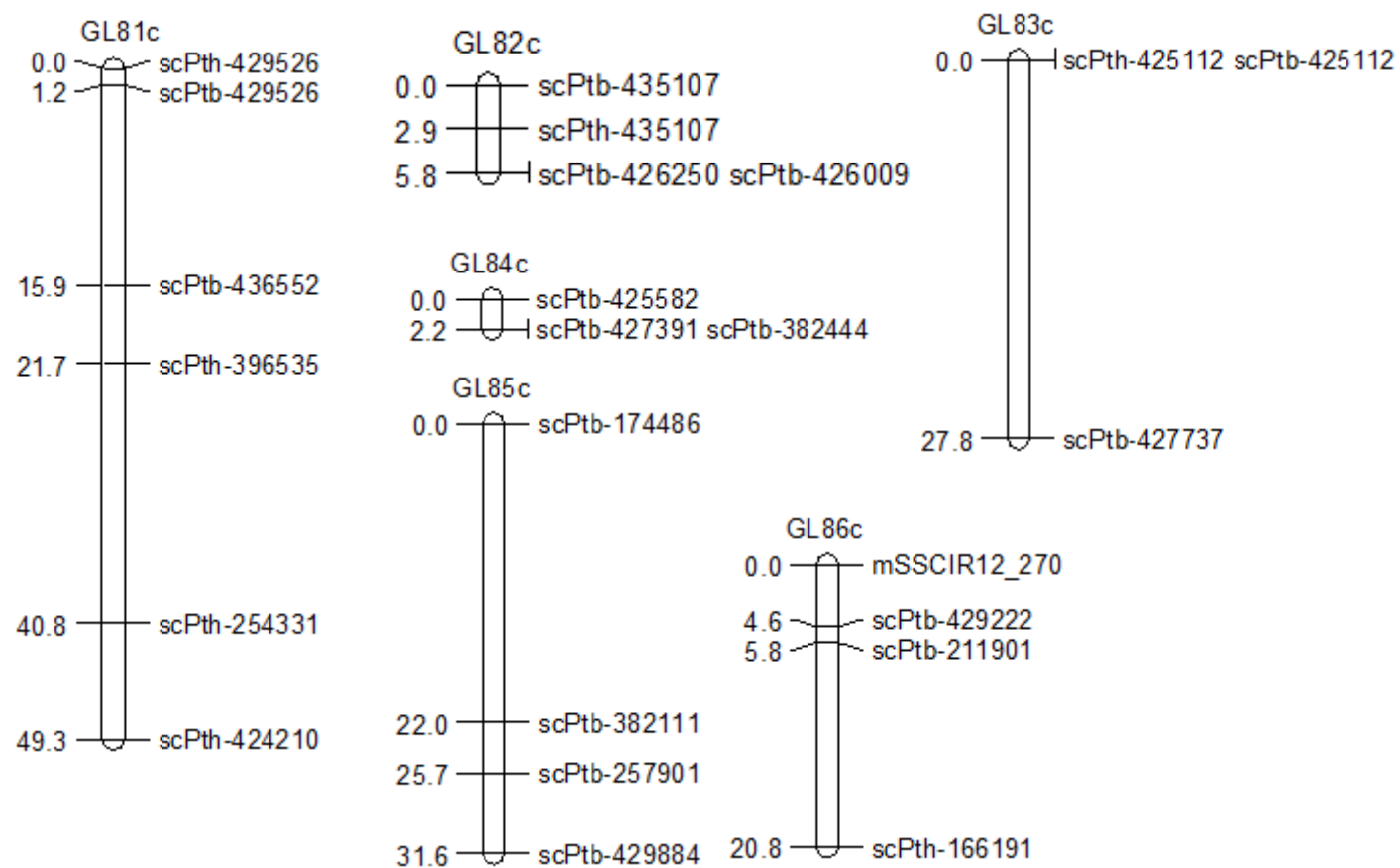
## Apêndice A. Continuação.



Continua...

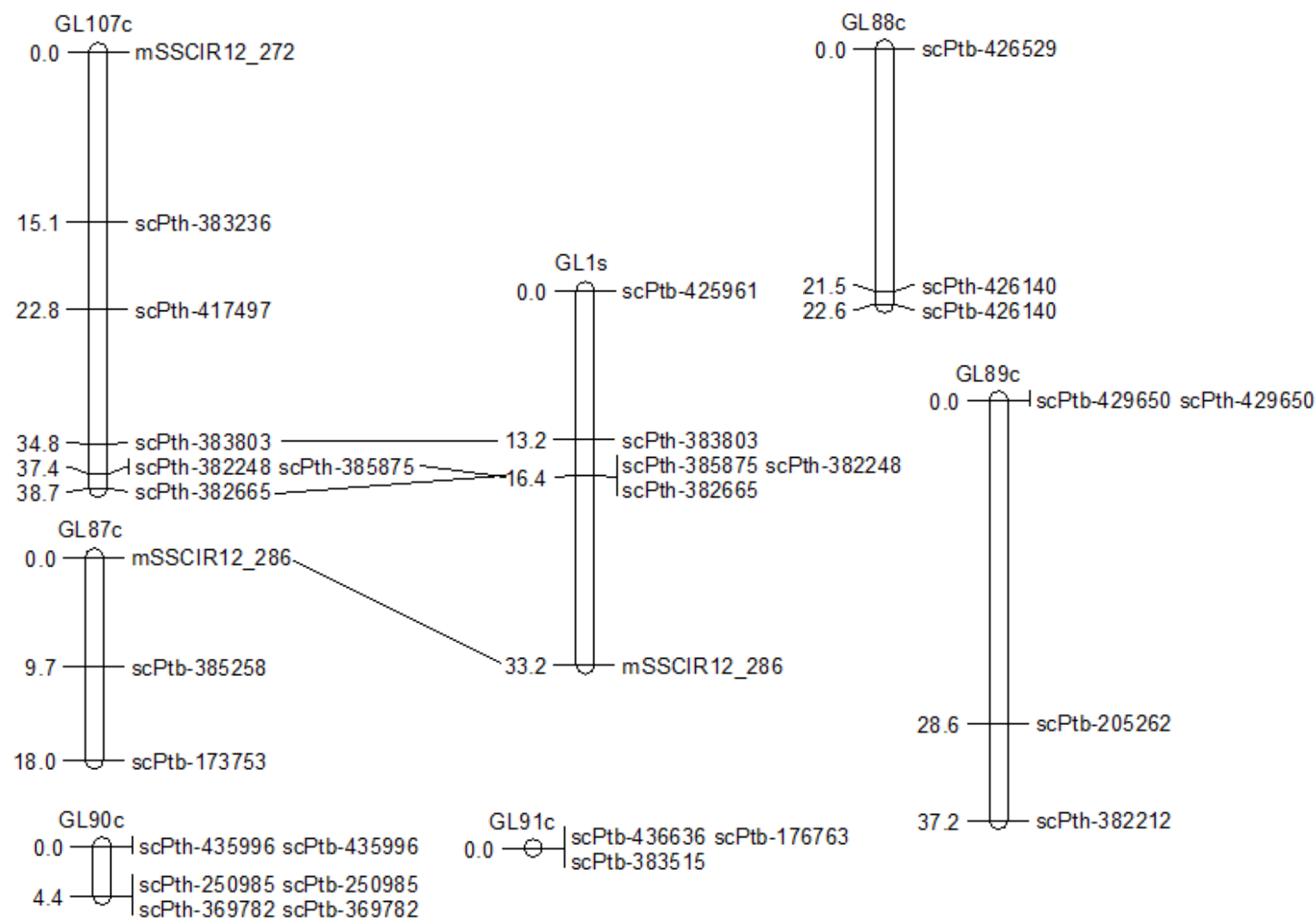


## Apêndice A. Continuação.



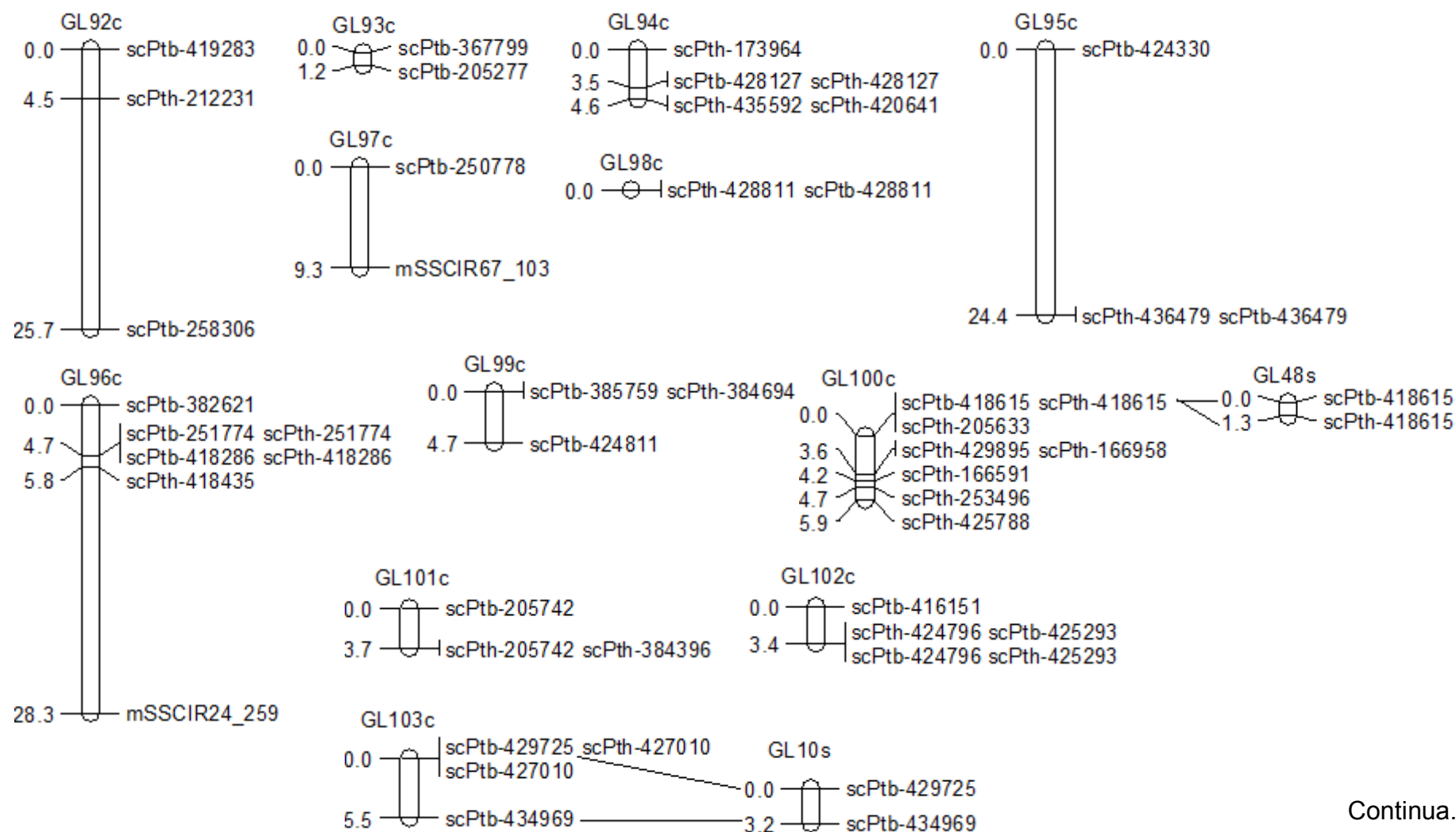
Continua...

## Apêndice A. Continuação.



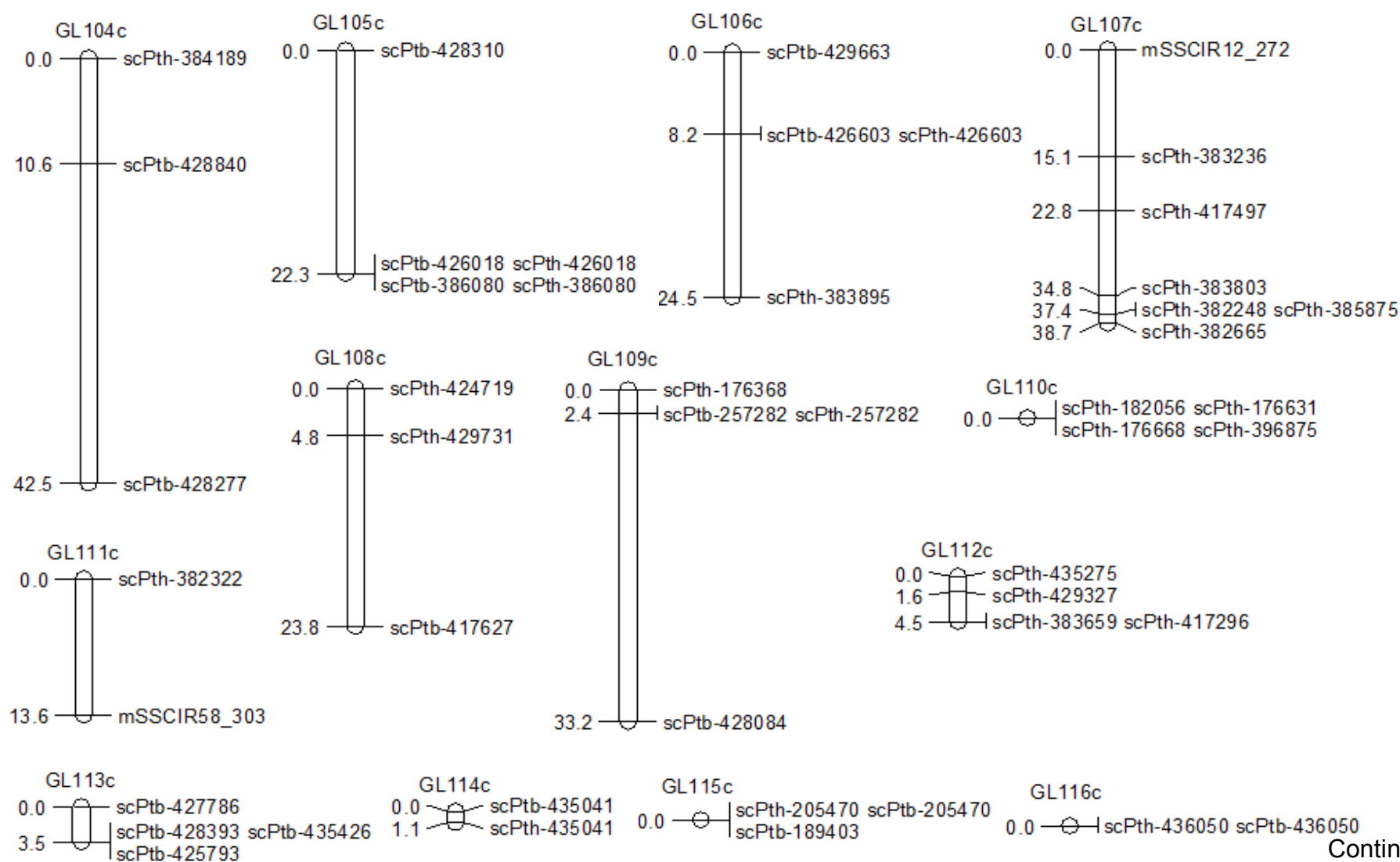
Continua..

## Apêndice A. Continuação.



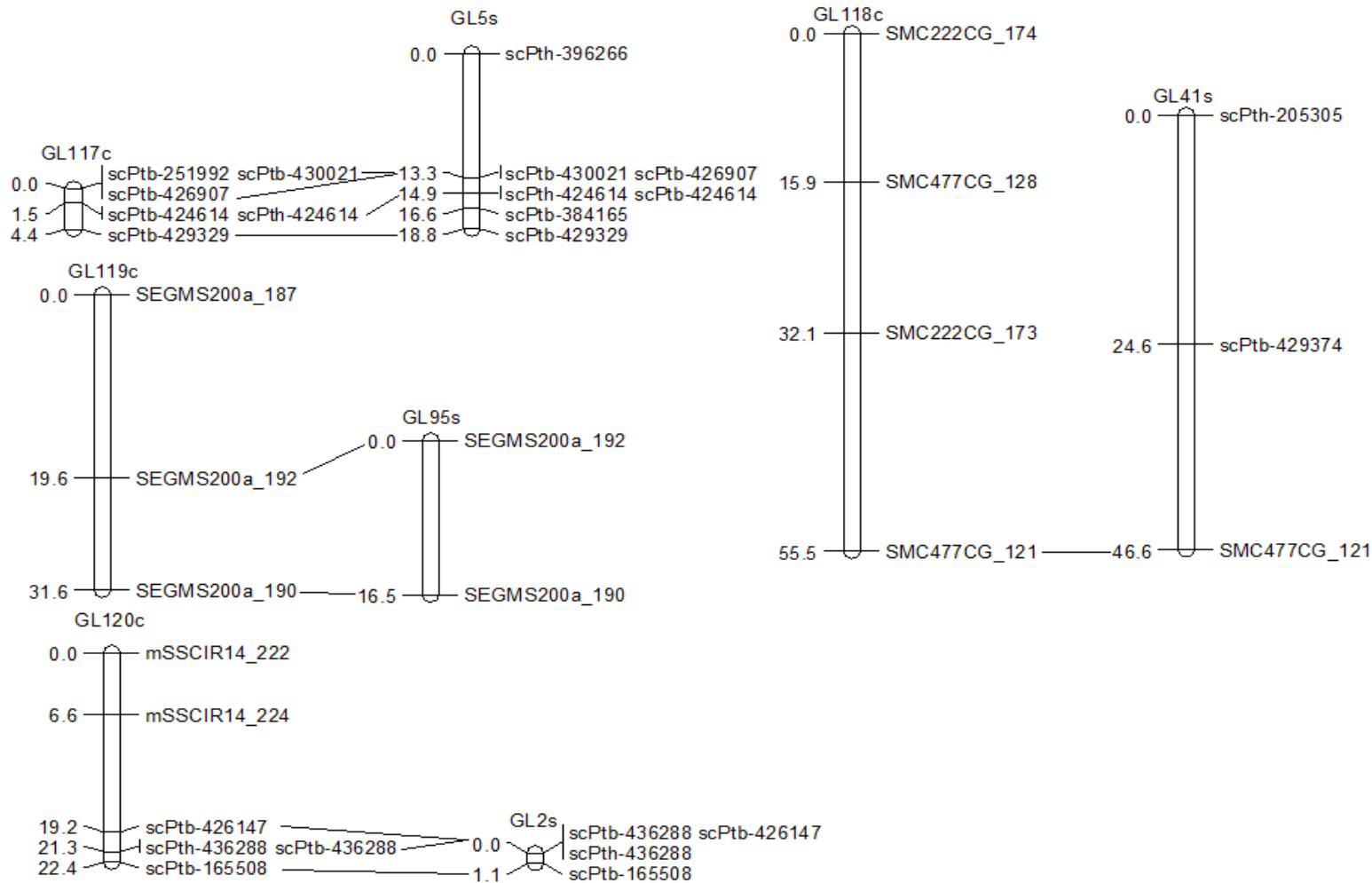
Continua...

## Apêndice A. Continuação.



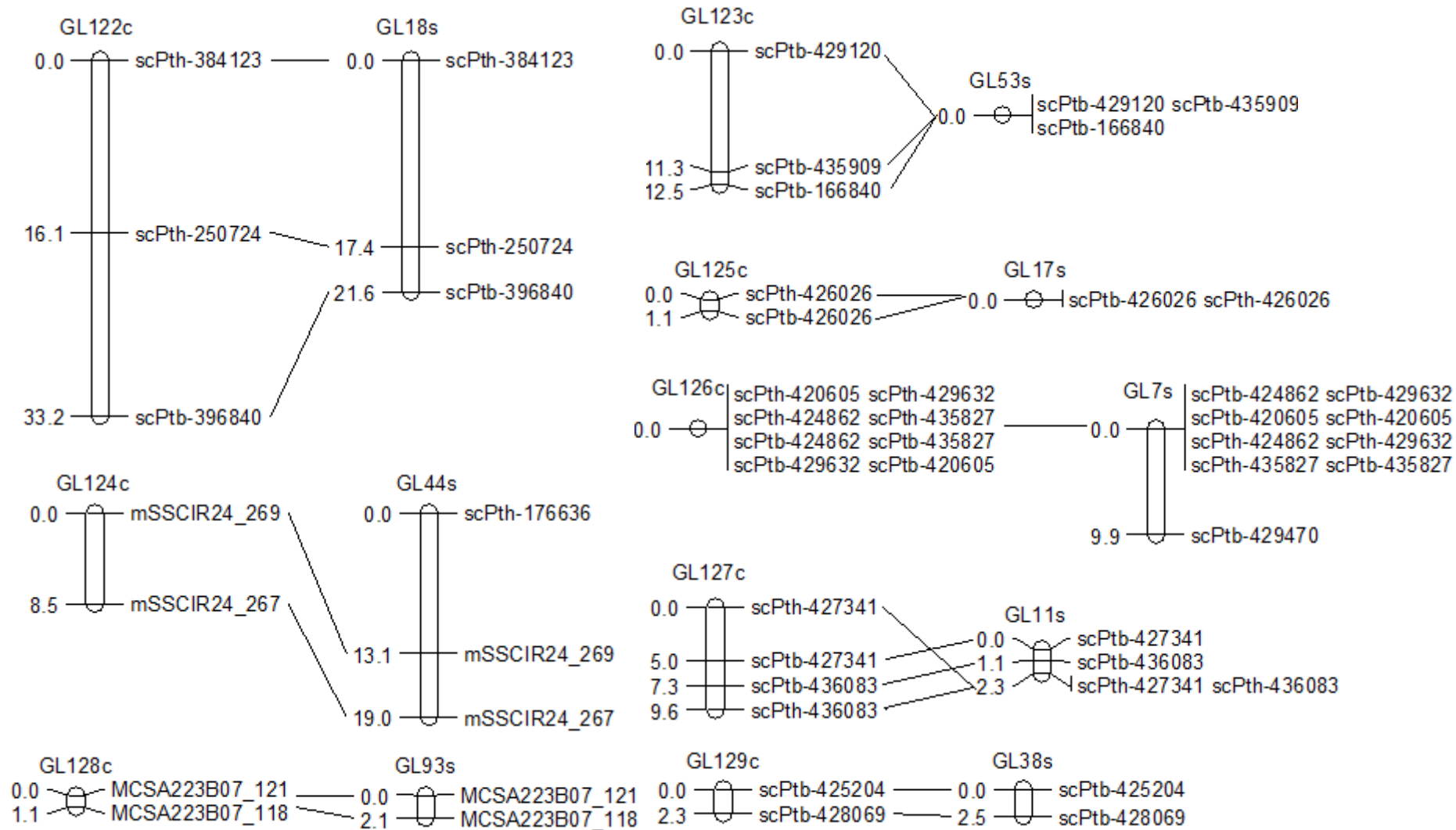
Continua...

**Apêndice A. Continuação.**



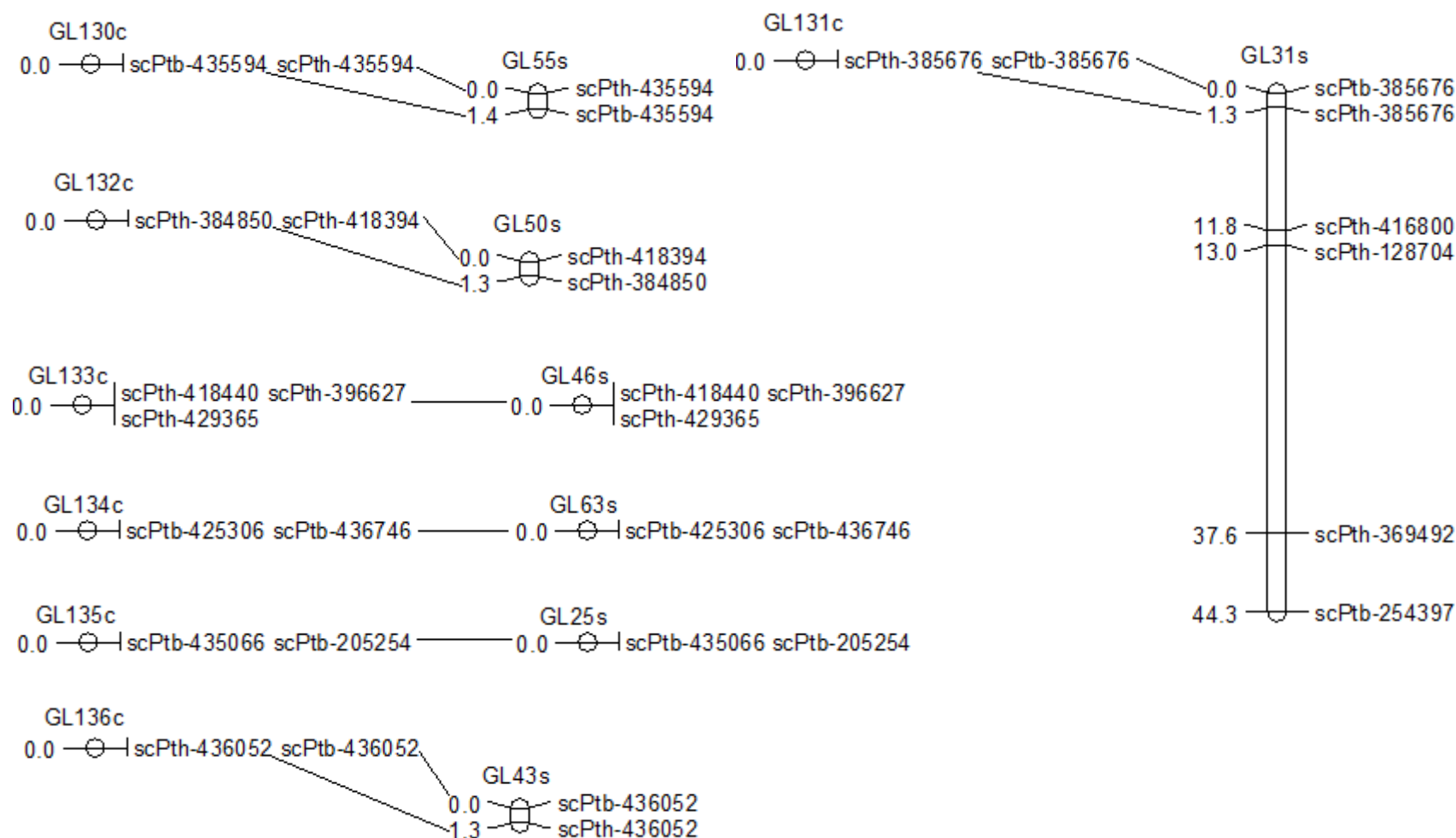
Continua...

## Apêndice A. Continuação.



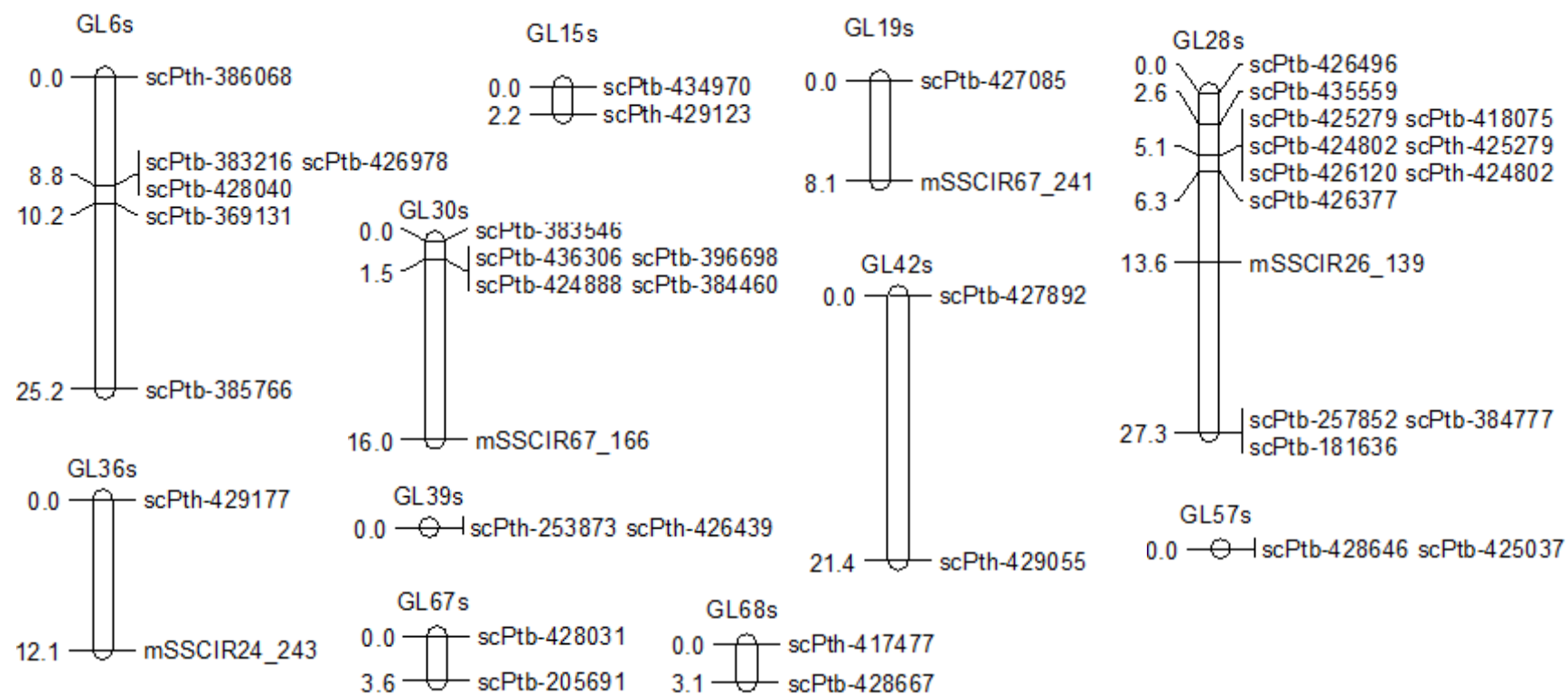
Continua...

## Apêndice A. Continuação.



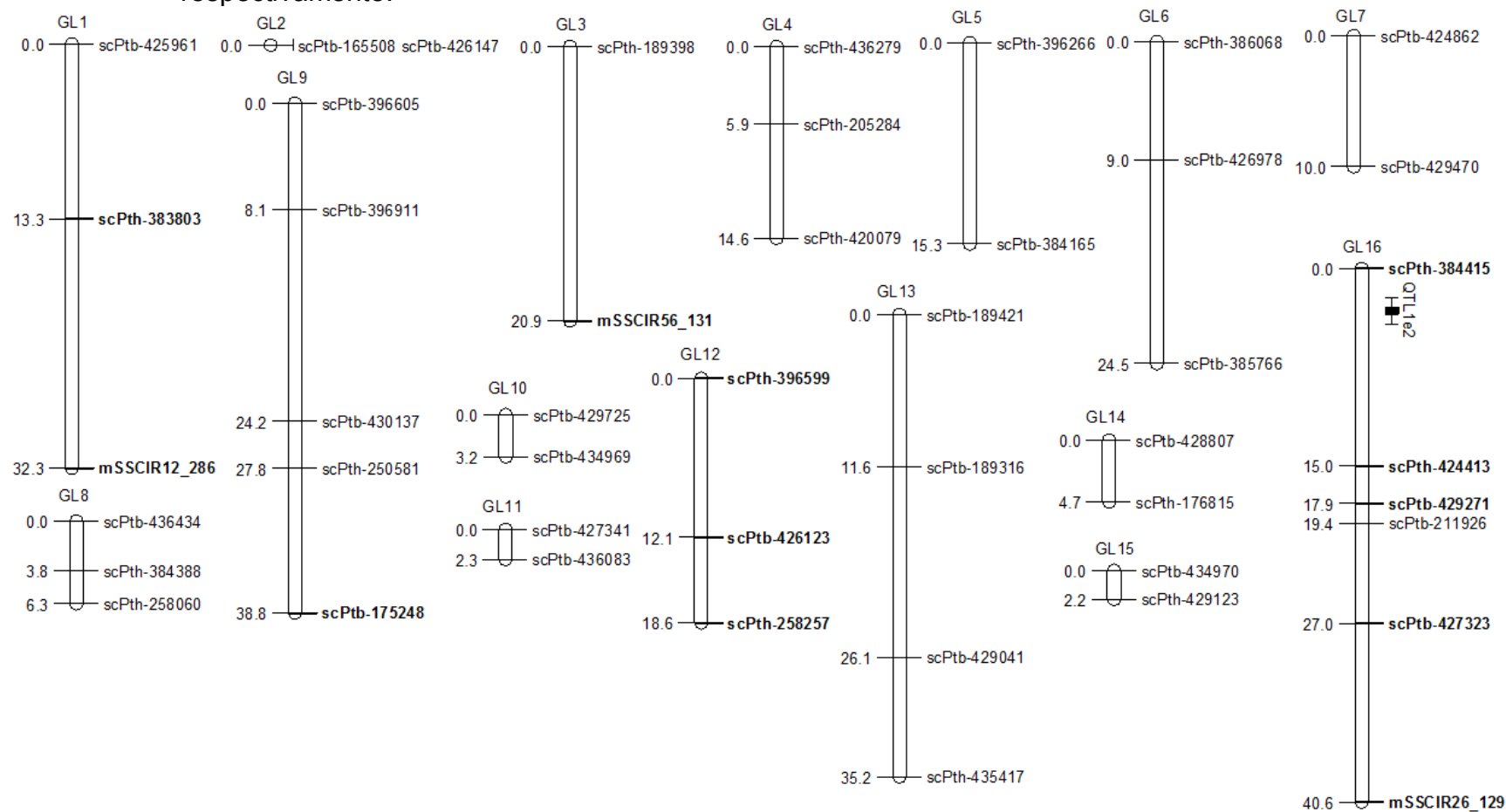
Continua...

## Apêndice A. Continuação.



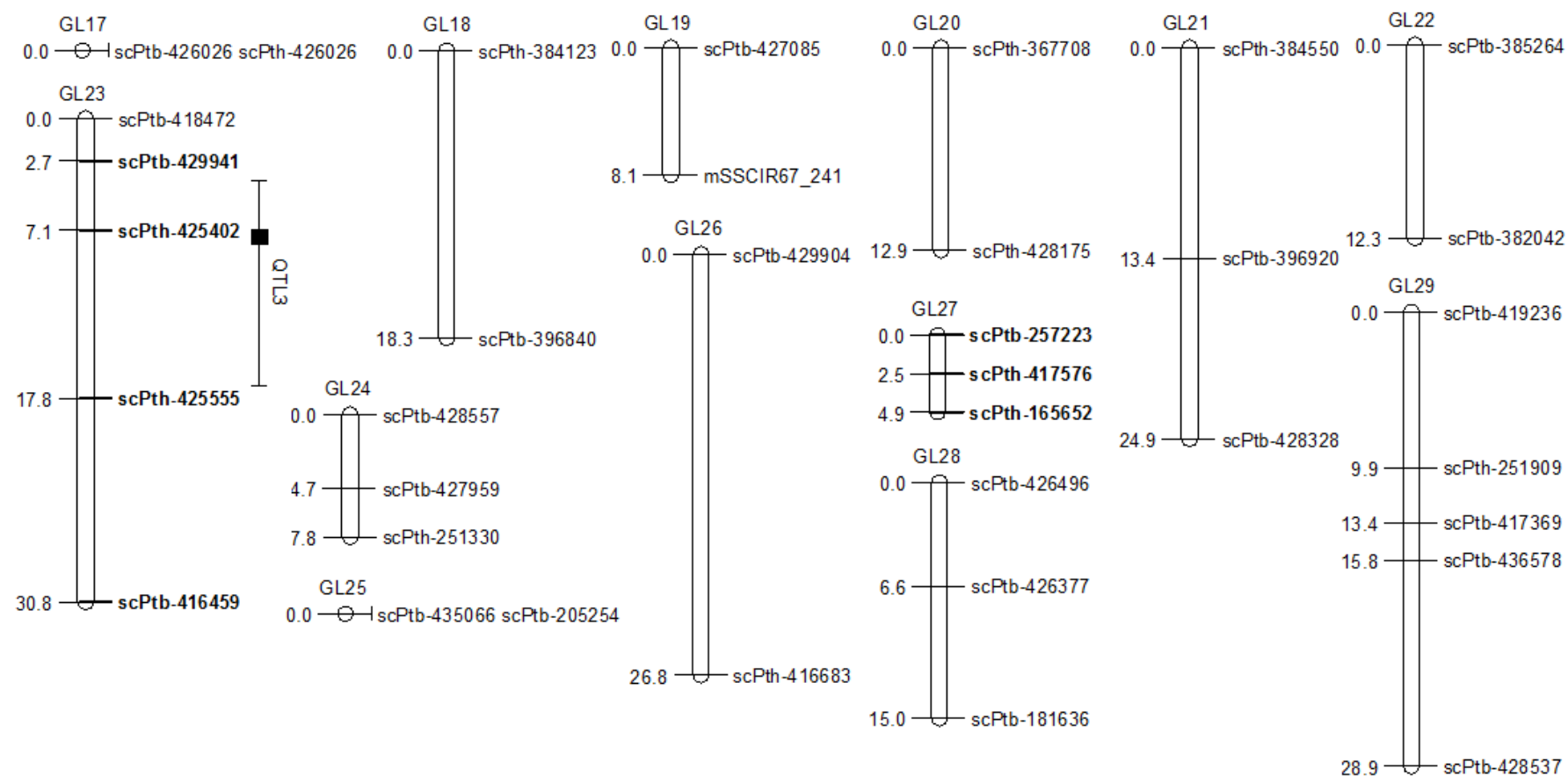


**Apêndice B.** Mapa *framework* obtido a partir da população derivada da autofecundação de RB97327. À direita está localizada a identificação dos marcadores e à esquerda as distâncias genéticas em cM, estimadas pela função de Kosambi. Os marcadores identificados pela análise de marcas simples como significativamente associados aos caracteres quantitativos avaliados no presente estudo estão indicados em **negrito** e os QTL 1, 2 e 3, identificados pela análise de mapeamento por intervalo composto, são destacados nos grupos de ligação GL16 e GL23, respectivamente.



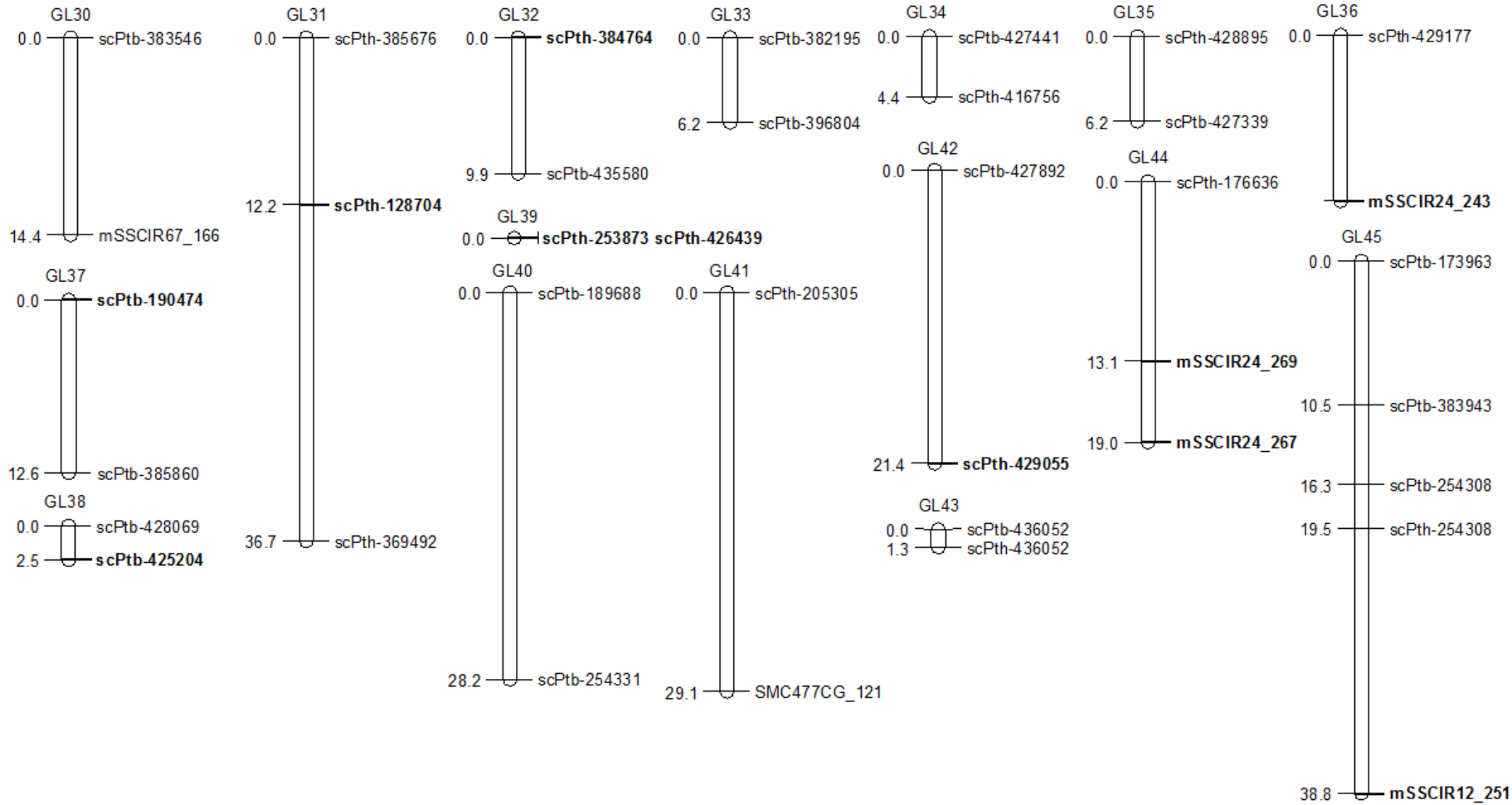
Continua..

## Apêndice B. Continuação.



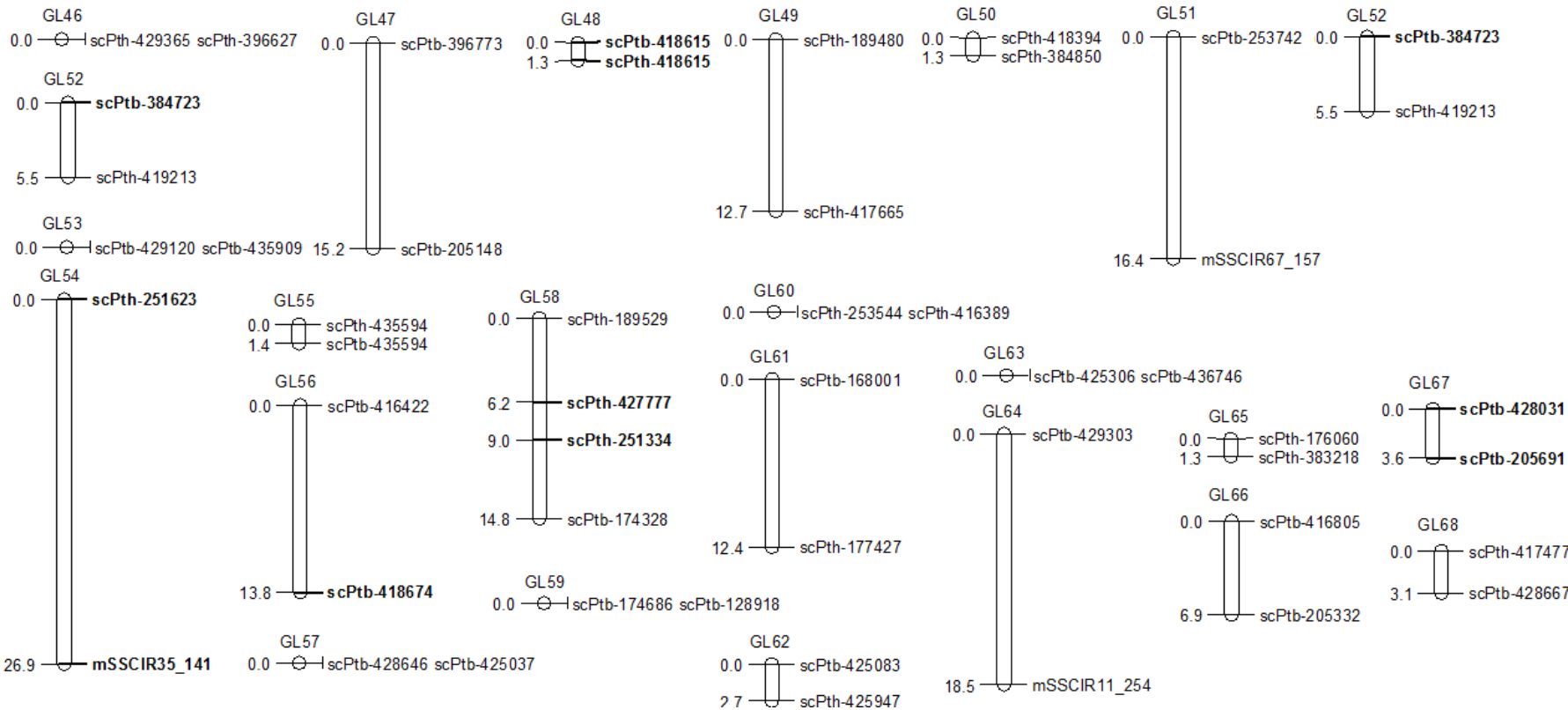
Continua...

Apêndice B. Continuação.



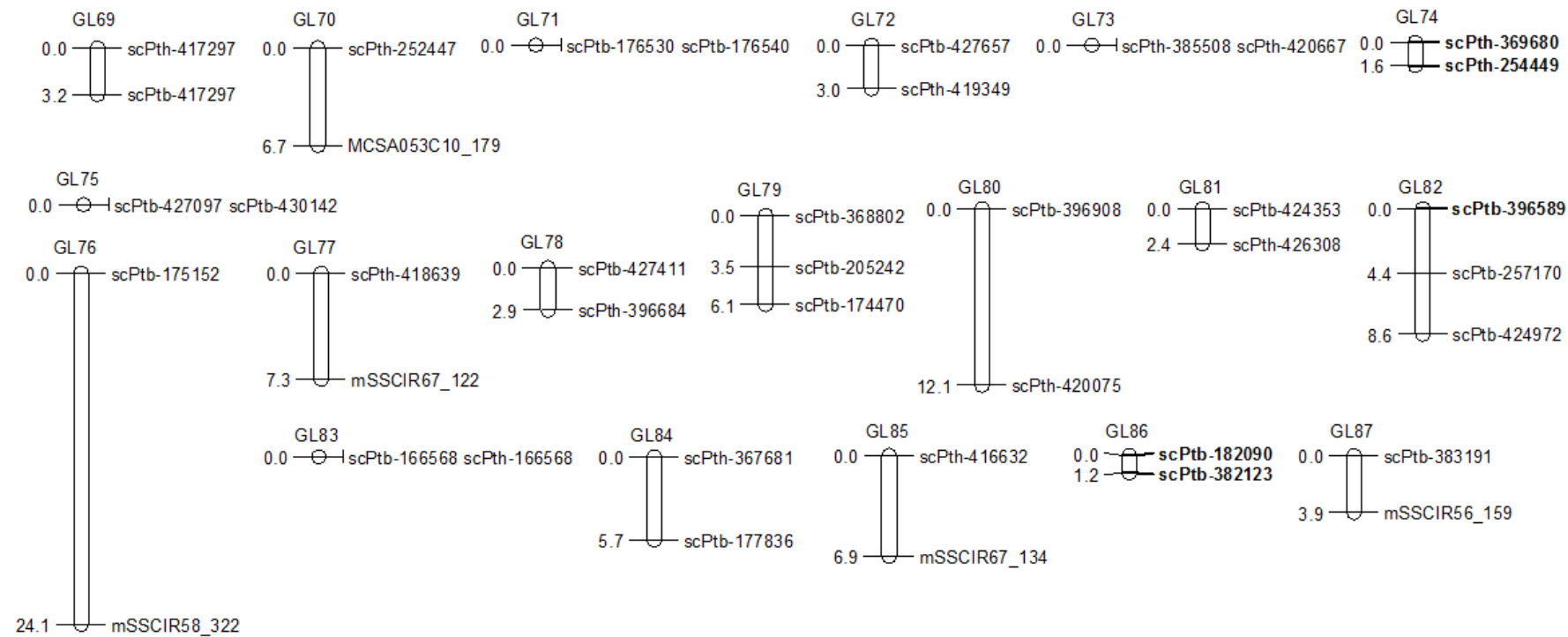
Continua...

# Apêndice B. Continuação.



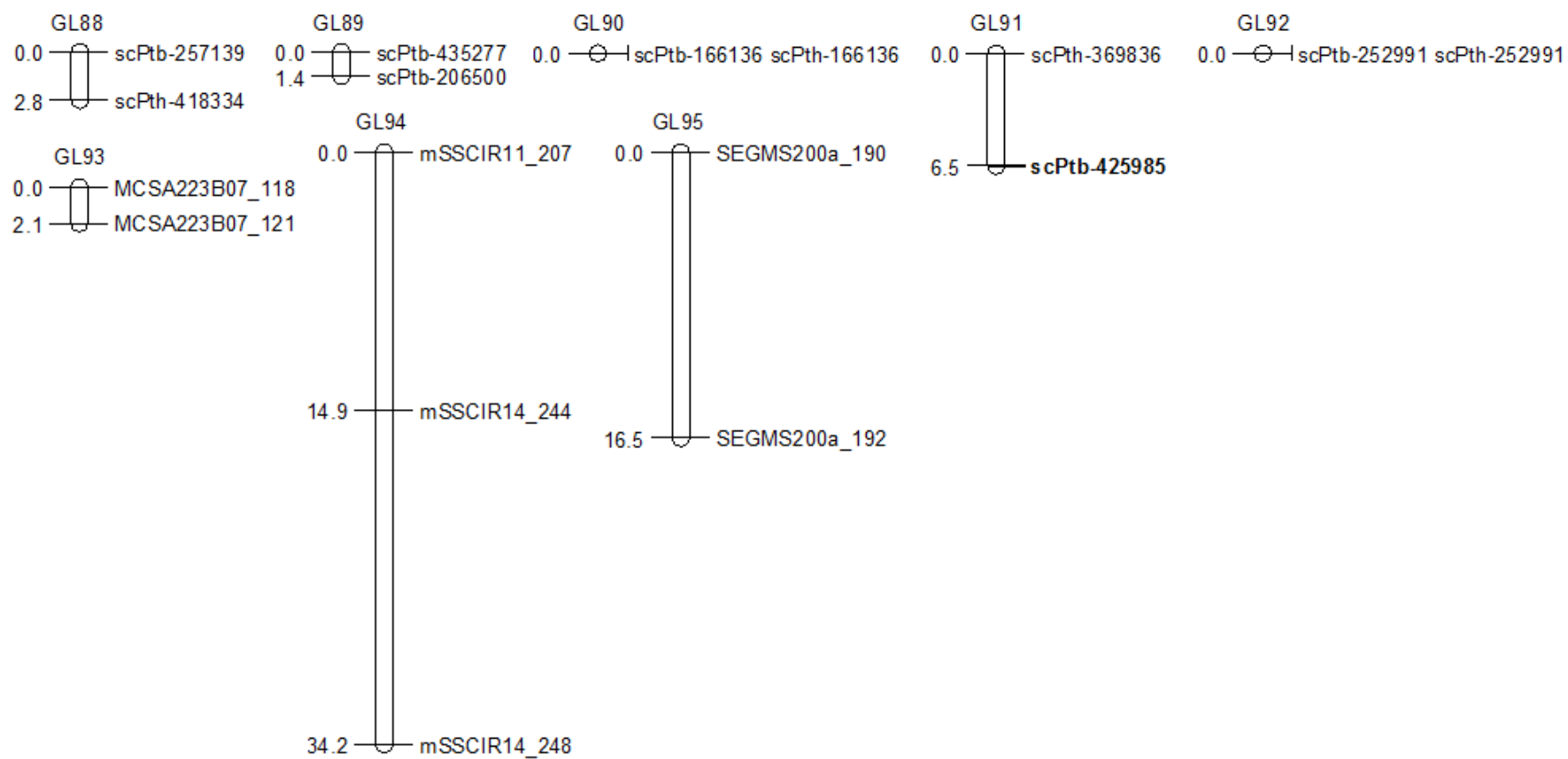
Continua...

**Apêndice B. Continuação.**

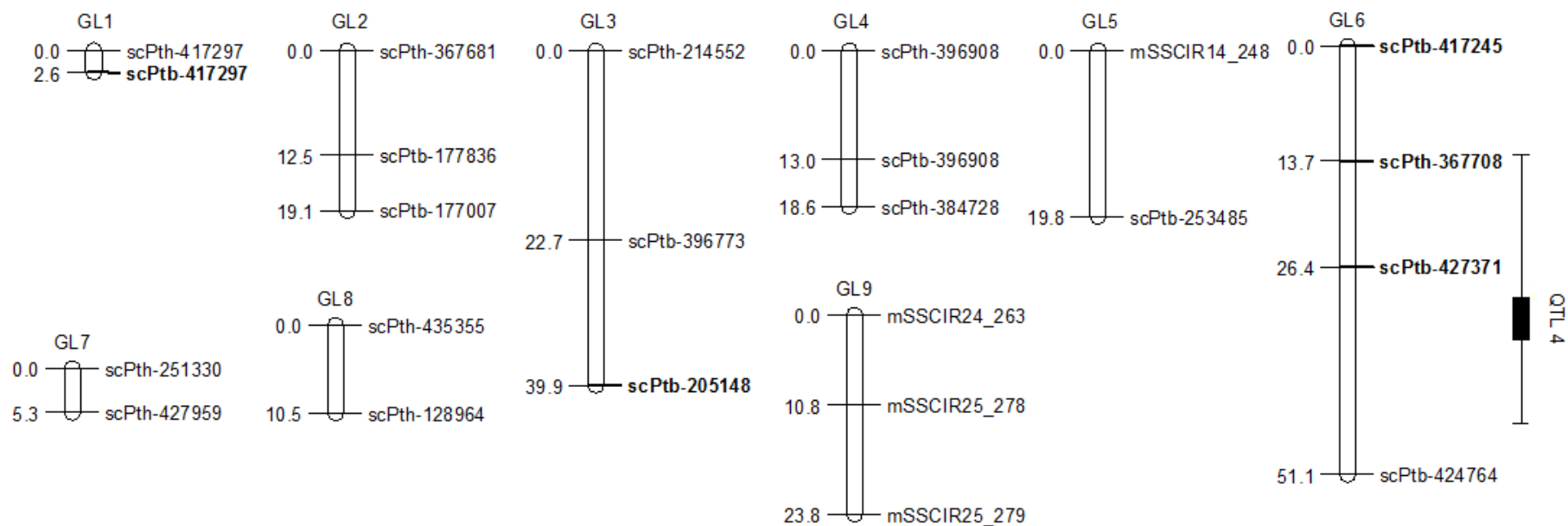


Continua...

## Apêndice B. Continuação.

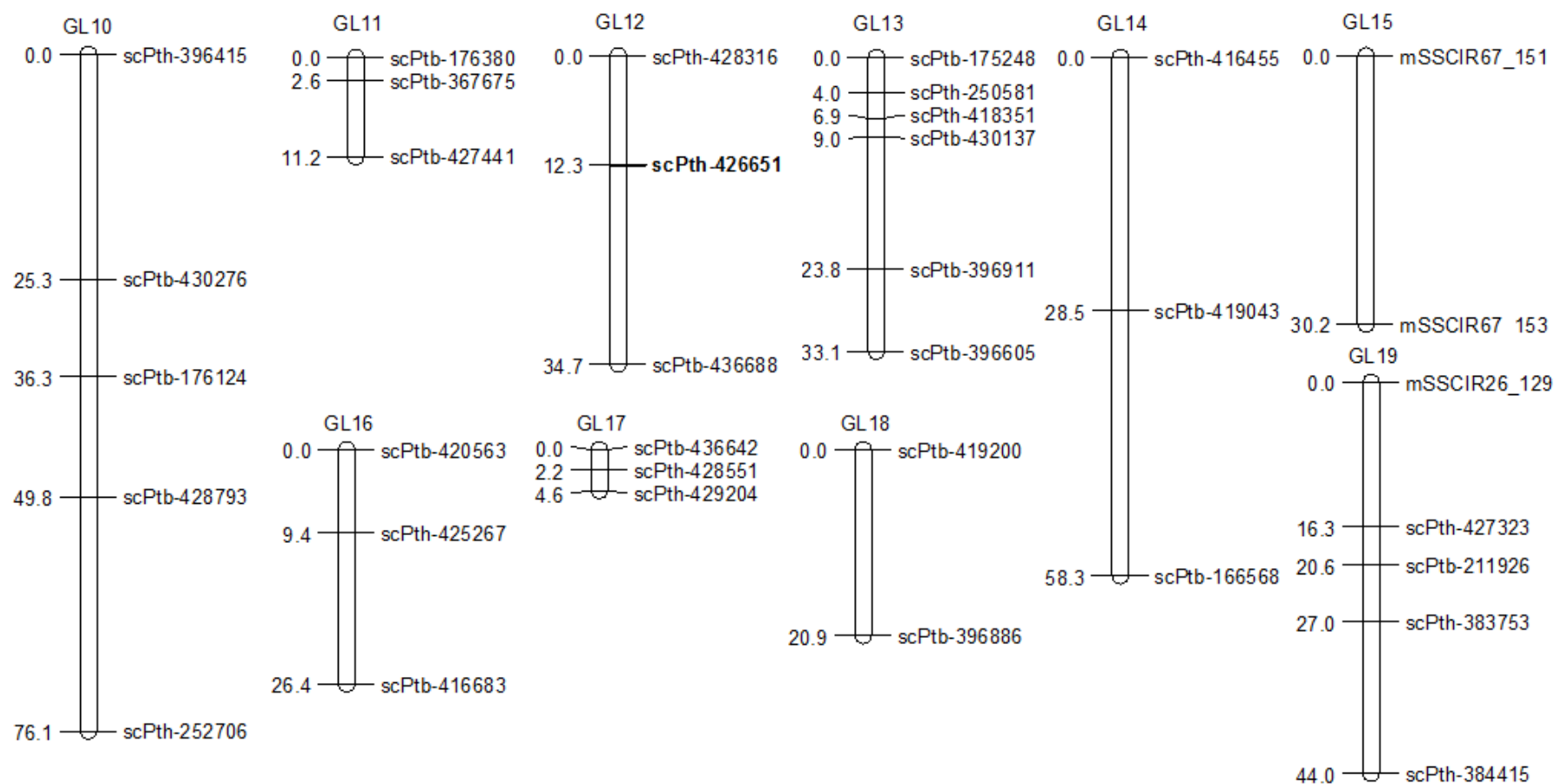


**Apêndice C.** Mapa *framework* obtido a partir da população derivada do cruzamento entre o genótipo-elite da Ridesa RB97327 e a variedade comercial RB72454. À direita está localizada a identificação dos marcadores e à esquerda as distâncias genéticas em cM, estimadas pela função de Kosambi. Os marcadores identificados pela análise de marcas simples como significativamente associados aos caracteres quantitativos avaliados no presente estudo estão indicados em negrito e os QTL 4 e 5, identificados pela análise de mapeamento por intervalo composto, são destacados nos grupos de ligação GL5 e GL81, respectivamente.



Continua...

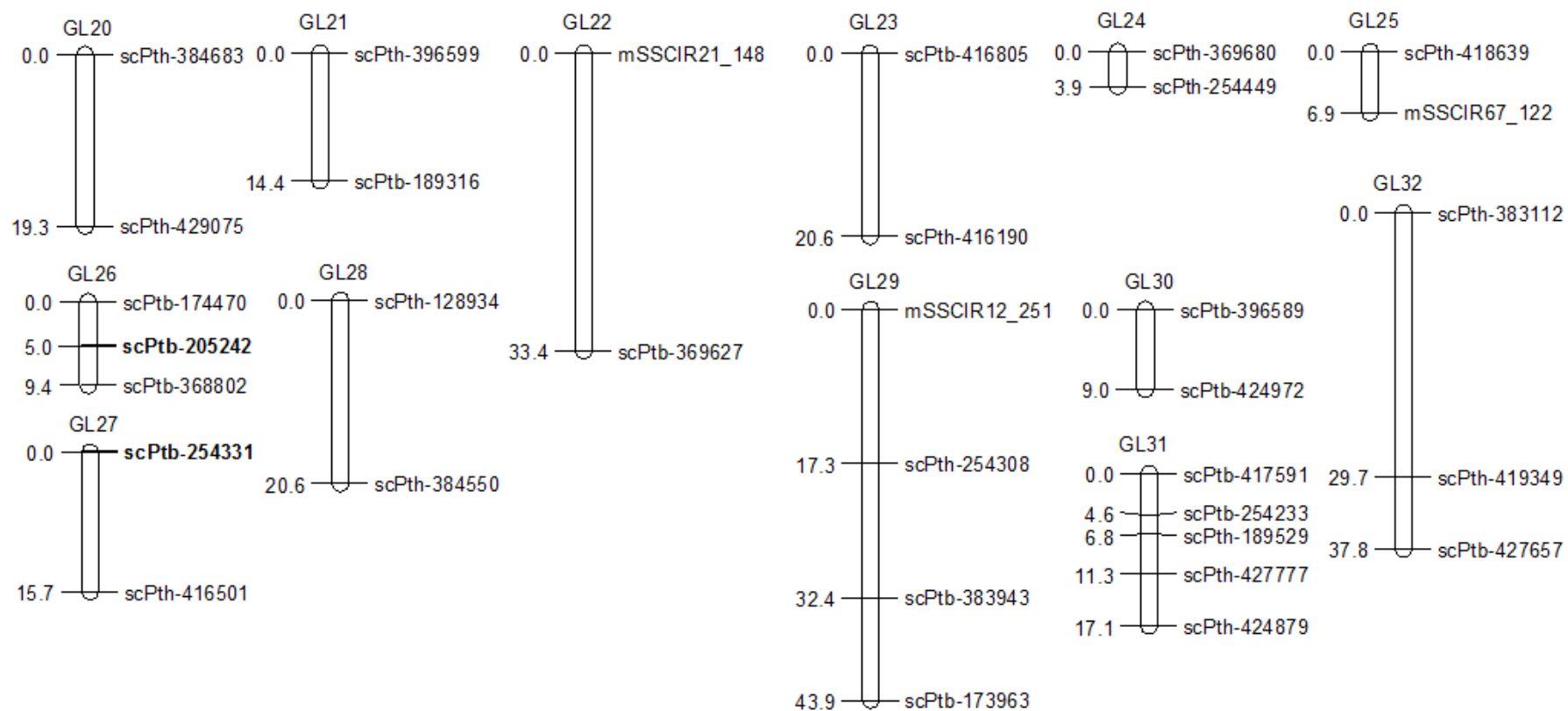
## Apêndice C. Continuação.



Continua...

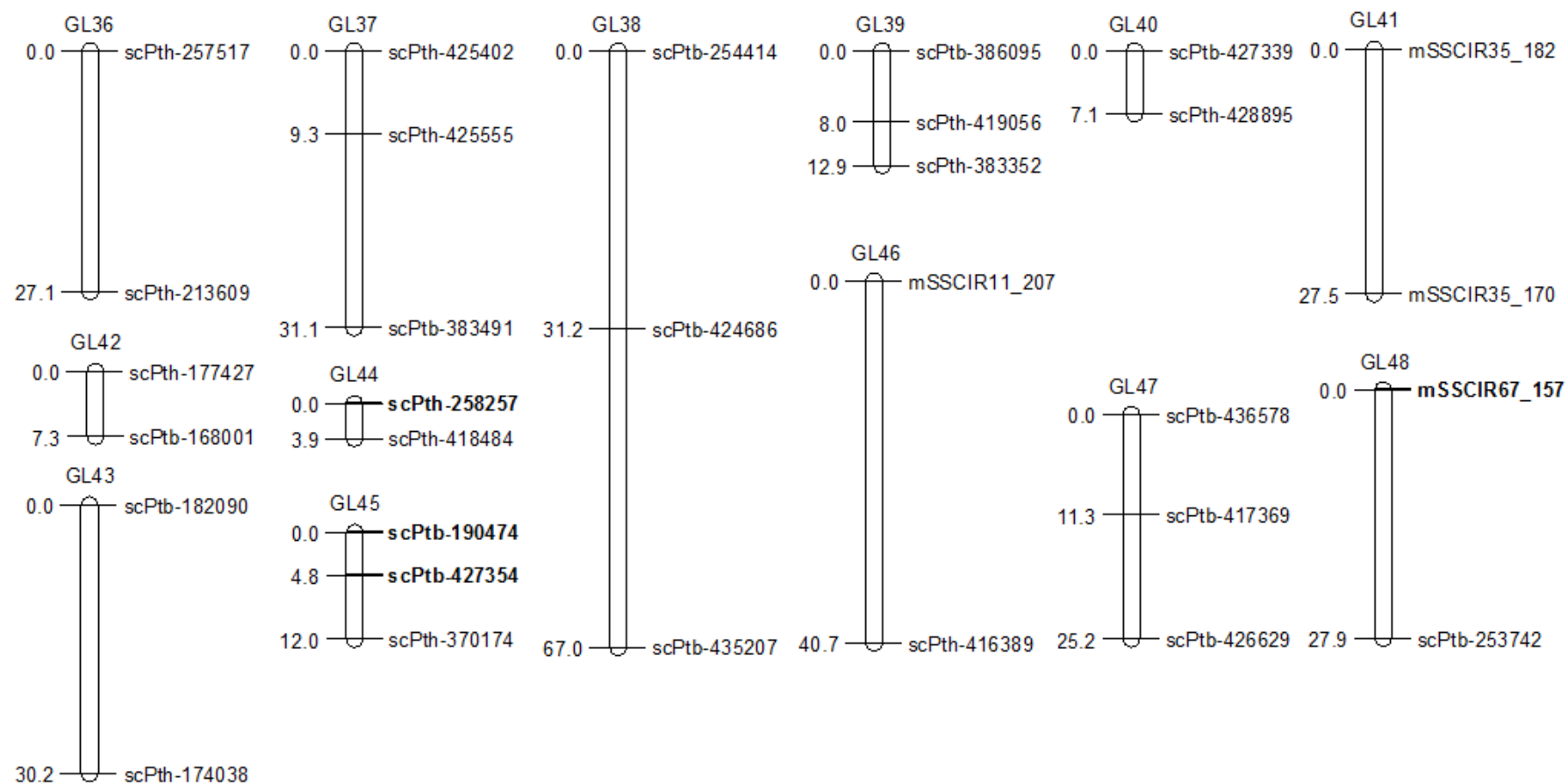


## Apêndice C. Continuação.



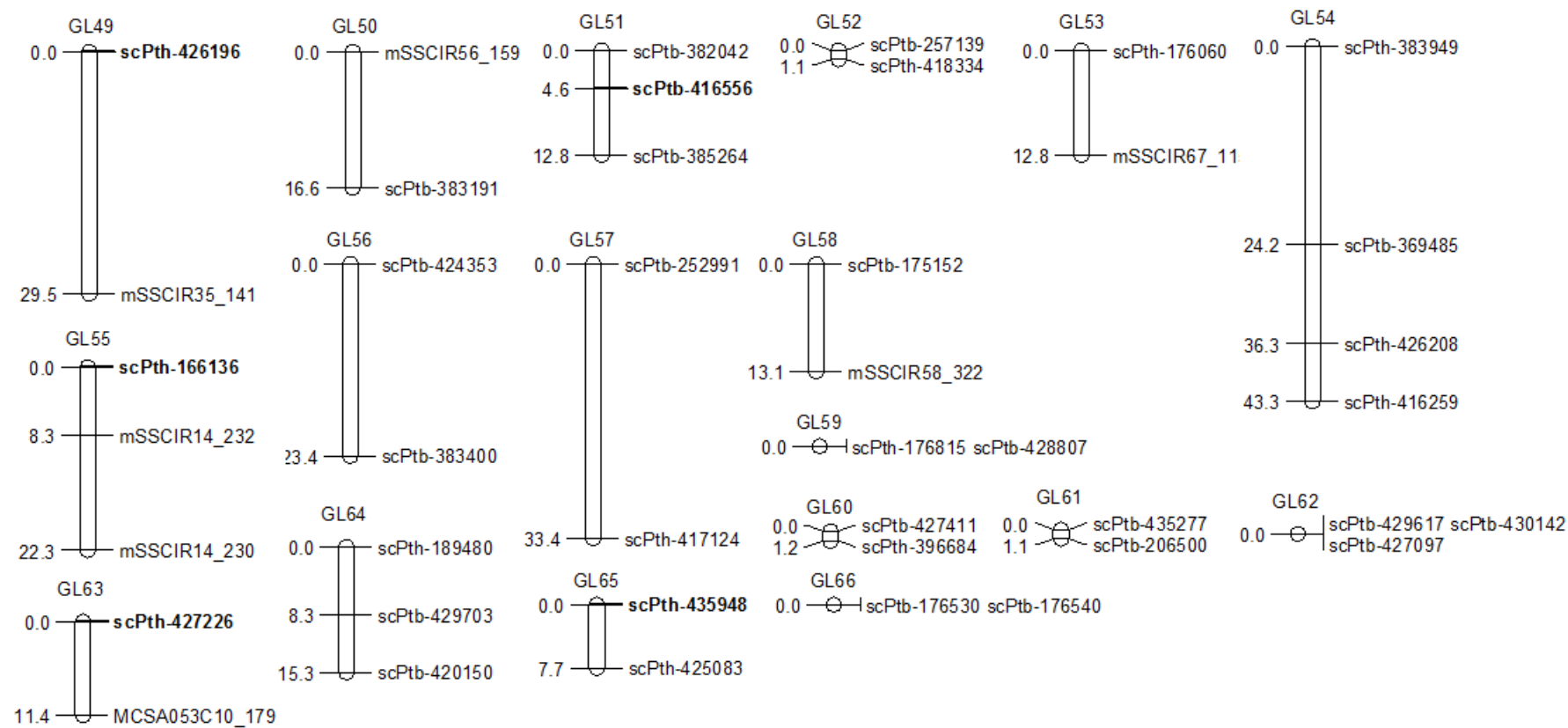
Continua...

## Apêndice C. Continuação.



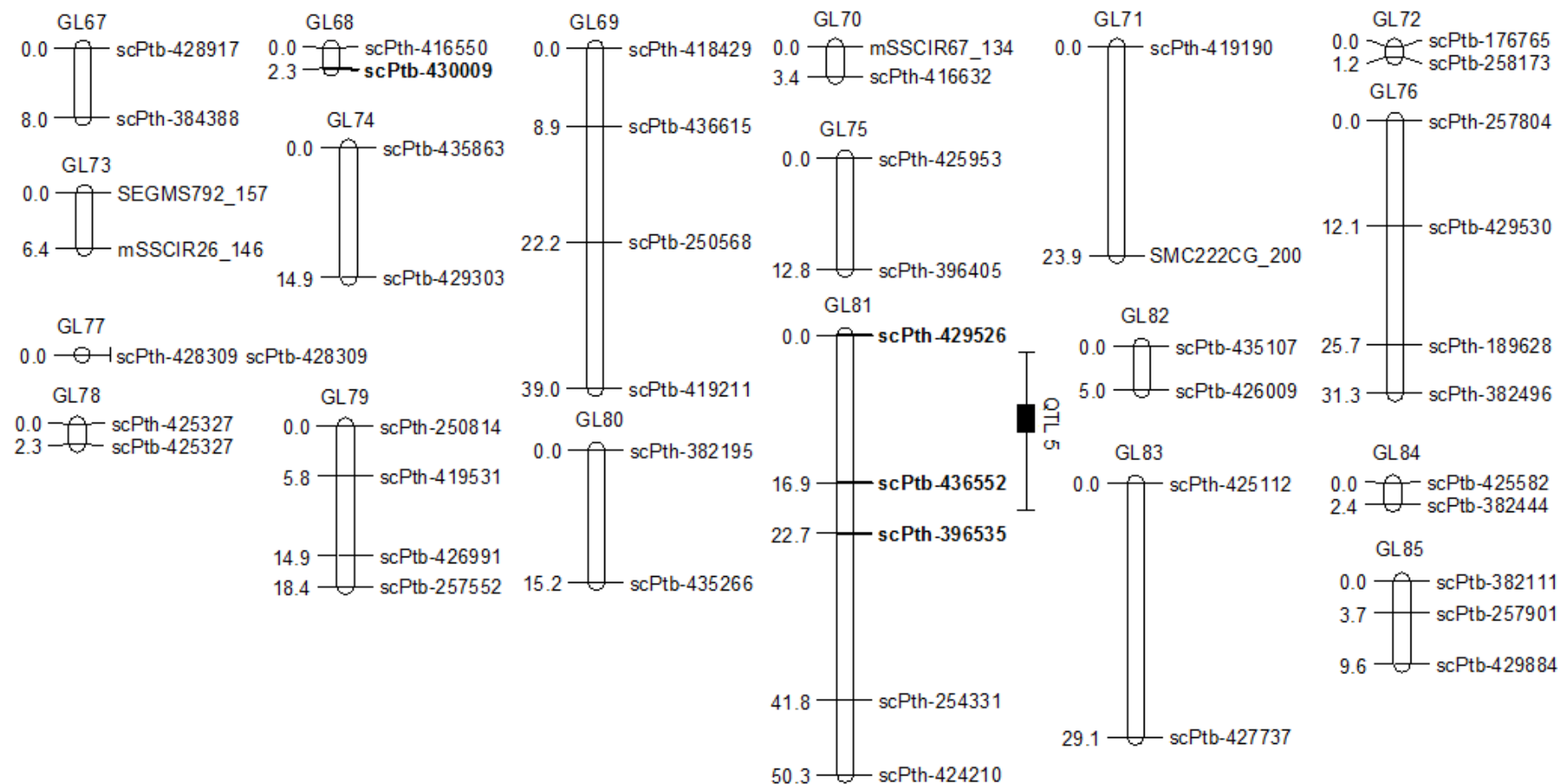
Continua...

## Apêndice C. Continuação.



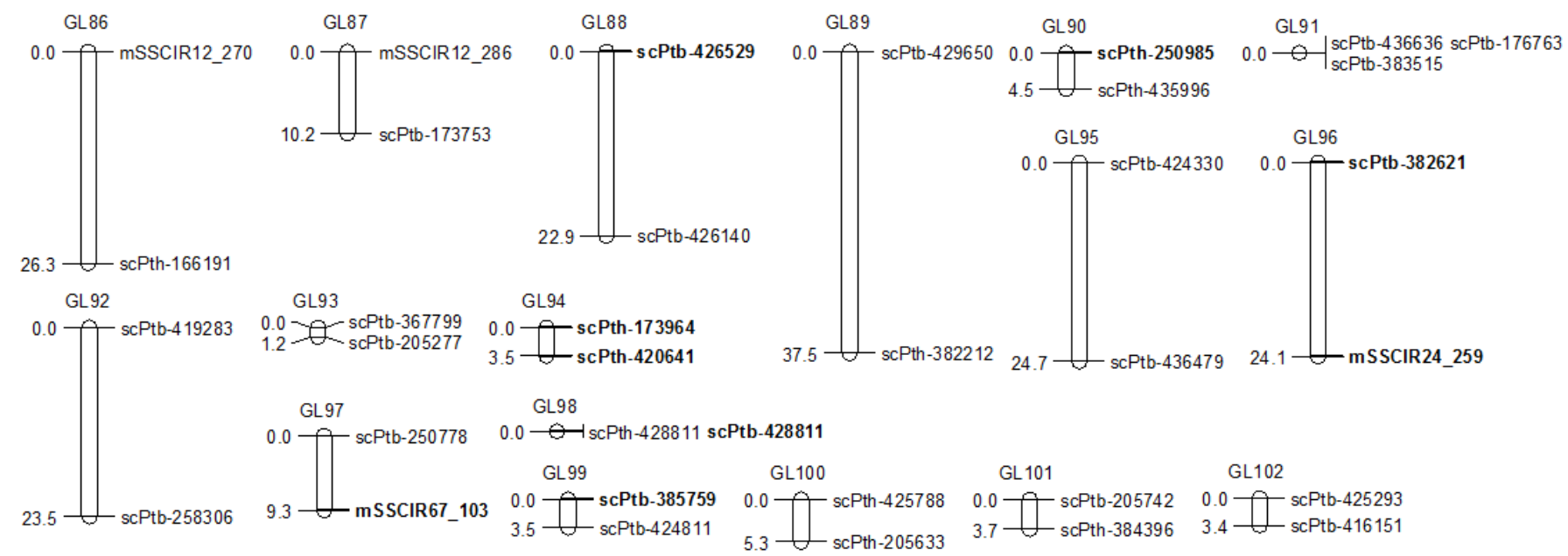
Continua...

## Apêndice C. Continuação.



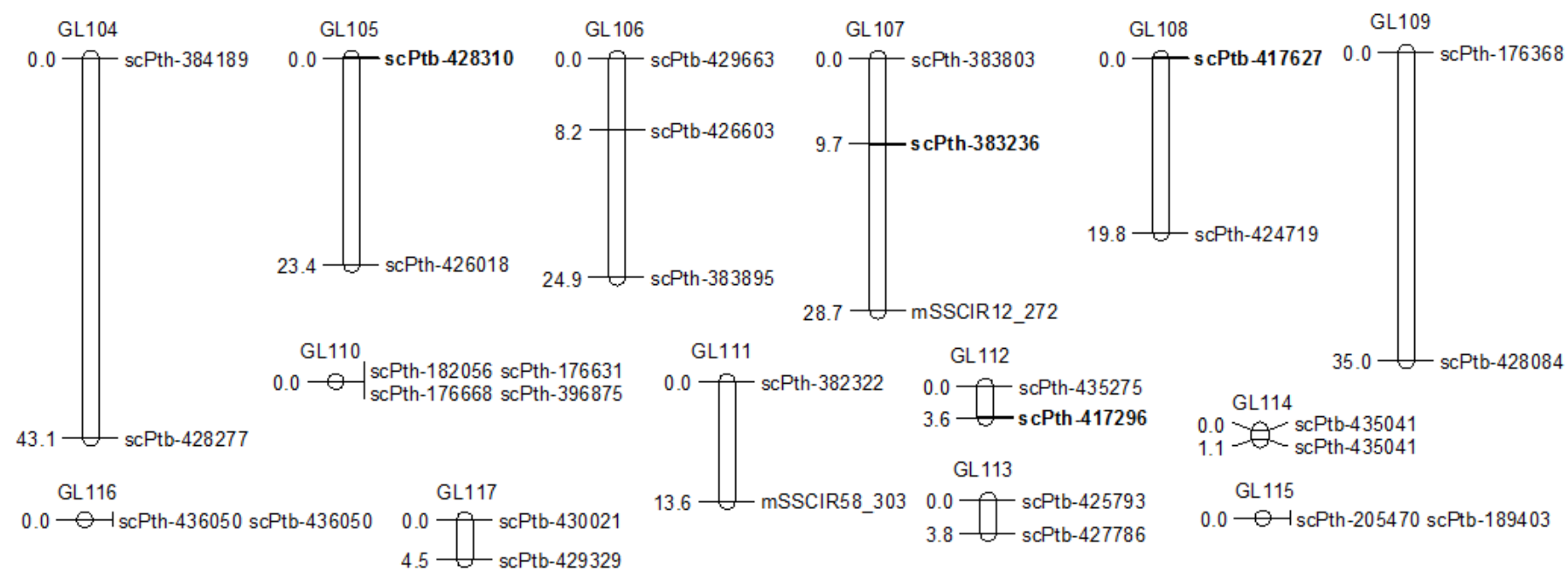
Continua...

Apêndice C. Continuação.



Continua...

## Apêndice C. Continuação.



Continua...

## Apêndice C. Continuação.

