

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

**MAPEAMENTO DE MICROSSATÉLITES FUNCIONAIS
DE SORGO (EST-SSR) EM CANA-DE-AÇÚCAR**

Silvan Gomes de Brito
Engenheiro Agrônomo

2016

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**MAPEAMENTO DE MICROSSATÉLITES FUNCIONAIS DE
SORGO (EST-SSR) EM CANA-DE-AÇÚCAR**

Silvan Gomes de Brito

Orientadora: Profa. Dra. Luciana Rossini Pinto

Co-orientador: Prof. Dr. Renato Vicentini dos Santos

**Tese apresentada à Faculdade de Ciências
Agrárias e Veterinárias – Unesp, Câmpus de
Jaboticabal, como parte das exigências para
a obtenção do título de Doutor em
Agronomia (Genética e Melhoramento
Genético de Plantas)**

2016

Brito, Silvan Gomes
M499a Mapeamento de microssatélites funcionais de sorgo (EST-SSR) em cana-de-açúcar. / Silvan Gomes de Brito. – – Jaboticabal, 2016
xix, 110 p. ; 28 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2016
Orientador: Luciana Rossini Pinto
Banca examinadora: Dilermando Perecin, Fernanda Raquel Camilo dos Santos, Gustavo Vitti Moro, Michael dos Santos Brito

Bibliografia

1. Colinearidade. 2. Genes ortólogos. 3. Microssatélites. 4. Mapeamento comparativo. 5. Marcador molecular. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 619:616.728.3:636.92

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: MAPEAMENTO DE MICROSSATÉLITES FUNCIONAIS DE SORGO (EST-SSR)
EM CANA-DE-AÇÚCAR

AUTOR: SILVAN GOMES DE BRITO

ORIENTADORA: LUCIANA ROSSINI PINTO

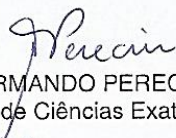
CO-ORIENTADOR: RENATO VICENTINI DOS SANTOS

Aprovado como parte das exigências para obtenção do Título de Doutor em AGRONOMIA
(GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:



Profa. Dra. LUCIANA ROSSINI PINTO

Centro Avançado de Pesquisa Tecnológica do Agronegócio de Cana / IAC / Instituto Agrônomo de Campinas -
Ribeirão Preto, SP



Prof. Dr. DILERMANDO PERECIN

Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal



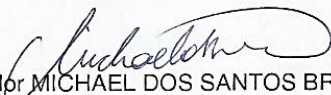
Prof. Dr. GUSTAVO VITTI MÔRO

Departamento de Produção Vegetal / FCAV / UNESP - Jaboticabal



Pós-doutoranda FERNANDA RAQUEL CAMILO DOS SANTOS

Instituto Agrônomo de Campinas / Centro de Cana de Açúcar / Ribeirão Preto, SP



Pesquisador MICHAEL DOS SANTOS BRITO

Instituto Agrônomo de Campinas / Centro de Cana de Açúcar / Ribeirão Preto, SP

Jaboticabal, 01 de julho de 2016

DADOS CURRICULARES DO AUTOR

SILVAN GOMES DE BRITO – Nasceu em 05 de maio de 1984, no município de Santa Rita- Paraíba. Em dezembro de 2005, concluiu o curso Técnico em agropecuária, no Centro de Formação Tecnólogos, Universidade Federal da Paraíba, Bananeiras-PB. Ingressou no curso de Engenharia Agrônômica (Bacharel), no ano de 2004, no Centro de Ciências Agrárias, Universidade Federal da Paraíba, Areia-PB, onde na mesma, obteve o título de Engenheiro Agrônomo no ano de 2009. Em julho de 2012, obteve o título de Mestre em Agronomia, pela Universidade Federal Rural de Pernambuco, Recife-PE. Em agosto de 2012, ingressou no programa de Pós-graduação em Agronomia (Genética e Melhoramento de Plantas), pela Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, São Paulo. Realizou estágio curricular para desenvolvimento da tese, no Centro Avançado da Pesquisa Tecnológica do Agronegócio de Cana, Instituto Agrônomo de Campinas-IAC/RP.

Tua caminhada ainda não terminou... A realidade te acolhe dizendo que pela frente o horizonte da vida necessita de tuas palavras e do teu silêncio.

Se amanhã sentires saudades, lembra-te da fantasia e sonha com tua próxima vitória. Vitória que todas as armas do mundo jamais conseguirão obter, porque é uma vitória que surge da paz e não do ressentimento.

É certo que irás encontrar situações tempestuosas novamente, mas haverá de ver sempre o lado bom da chuva que cai e não a faceta do raio que destrói.

Tu és jovem. Atender a quem te chama é belo, lutar por quem te rejeita é quase chegar à perfeição.

A juventude precisa de sonhos e se nutrir de lembranças, assim como o leito dos rios precisa da água que rola e o coração necessita de afeto.

Não faças do amanhã o sinônimo de nunca, nem o ontem te seja o mesmo que nunca mais.

Teus passos ficaram. Olhes para trás... mas vá em frente pois há muitos que precisam que chegues para poderem seguir-te.

Charles Chaplin

A Deus

Ofereço

Aos meus pais, Sebastião Gomes e Maria da Penha, a minha irmã, Silvana e aos amores de minha vida (minhas filhas), Janice, Samantha Gabrielle e Lavínia Fernanda.

Dedico

Agradecimentos

A Deus em primeiro lugar, por ter me dado a oportunidade de viver e a capacidade para enfrentar e vencer as dificuldades proporcionadas ao longo da vida. À minha família, pelo apoio e confiança no meu trabalho. Aos meus queridos pais por terem proporcionado toda condição para concluir este meu objetivo.

À minha irmã, pela força e carinho. Ao meu cunhado Eduardo, pelo apoio e grande amizade. À Paula Fernanda e seus pais, José Gomes e Anunciada Dutra e Berenice Eugênio, por acreditar em mim, dando crédito a minha escolha profissional e a minha inserção em sua família.

A minha orientadora Prof^{fa}. Dra. Luciana Rossini Pinto, pelo convívio, paciência, amizade e ensinamentos transmitidos, contribuindo para minha formação profissional e acadêmica ao longo desses anos.

Ao meu Co-orientador. Dr. Renato Vicentini dos Santos, pelas valiosas contribuições que melhoraram a qualidade deste trabalho.

Aos Professores do Programa de Pós-Graduação em Agronomia (Genética e Melhoramento Genético de plantas) da UNESP/FCAV, pelos ensinamentos e dedicação.

Ao professor Dr. Dilermando Perecin, pelos ensinamentos, paciência e a valiosa contribuição na análise de dados.

A Dra. Karen Toledo, por doar as sementes dos híbridos de sorgo da empresa Dow AgroScienses utilizadas nesse estudo.

Aos funcionários da seção técnica do Programa de Pós-Graduação da UNESP/FCAV, pela paciência e constantes ajudas fornecidas.

Aos amigos Alisson Esdras, Alisson Jales, Guilherme, Gustavo Hugo, Hudson, Ivanildo, Jacqueline Ferraça, Lucas (o paulista), Marília, Miquéias.

Aos companheiros GOU (Grupo de Oração Universitária), Andréa, Daniel, Denise, Liziane, Natalia Águila e Roberto pelos momentos de oração e amizade construída ao longo desses anos.

Aos amigos do laboratório de Biotecnologia do Centro de Cana-IAC/RP, Alexandre, Carol Neuber, Cibele, Fernanda Raquel, Isadora (Izinha), Jéssica, Juliana,

Larissa, Maria Natália, Maria Leticia, Marcel, Michael, Rafael e em especial ao Maycon (Maycão) e a Thaís, pela ajuda e por compartilharem seus conhecimentos.

Aos funcionários do Centro de Cana-IAC/RP, pela paciência e constantes ajudas fornecidas, em especial a D. Marta, pelo carinho, atenção e amizade.

Ao apoio institucional e financeiro da Universidade Estadual Paulista e Centro Avançado da Pesquisa Tecnológica do Agronegócio de Cana/Apta-IAC/RP. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pela concessão à bolsa.

Enfim, a todos e aqueles que não pude citar, mas que contribuíram para a conclusão desta minha etapa profissional.

Muito obrigado!

SUMÁRIO

RESUMO.....	i
ABSTRACT.....	ii
CAPÍTULO I.....	1
1 INTRODUÇÃO GERAL.....	2
2 REVISÃO DE LITERATURA.....	4
2.1 Aspectos gerais da cultura da cana-de-açúcar.....	4
2.2 Marcadores microssatélites derivados das etiquetas de sequências expressas (EST-SSRs) e transferibilidade entre gêneros correlatos.....	6
2.3 Mapeamento genético.....	9
2.4 Construção de mapa genético.....	11
2.5 Mapeamento comparativo.....	11
2.6 Mapeamento de QTL em cana-de-açúcar.....	13
2.7 Desenvolvimento de marcadores microssatélites a partir de sequências ortólogas conservadas (COS) entre cana e sorgo.....	15
2.8 Importância do sorgo para o mapeamento genético da cana-de-açúcar.....	18
3 REFERÊNCIAS.....	21
CAPÍTULO II.....	34
DESENVOLVIMENTO E CARACTERIZAÇÃO DE MARCADORES COS-EST-SSRS ENTRE CANA E SORGO.....	34
RESUMO.....	35
ABSTRACT.....	36
1 INTRODUÇÃO.....	37
2 MATERIAL E MÉTODOS.....	39
2.1 Material vegetal.....	39
2.2 Obtenção de sequências de etiquetas expressas ortólogas conservadas entre cana e sorgo (COS-EST) contendo microssatélites.....	40
2.3 Desenho dos pares de <i>primers</i> (COS-EST-SSRs).....	40
2.4 Anotação funcional dos COS-EST-SSR.....	41
2.5 Extração do DNA genômico.....	41
2.6 Amplificação dos COS-EST-SSR.....	42

2.7	Eletroforese em gel de poliacrilamida desnaturante	42
2.8	Análise de polimorfismo dos COS-EST-SSR	46
3	RESULTADOS	46
3.1	Identificação dos COS-EST-SSRs	46
3.2	Anotação funcional dos COS-EST-SSR	48
3.3	Análise de polimorfismo dos COS-EST-SSRs em cana e sorgo	48
4	DISCUSSÃO	53
5	CONCLUSÃO	56
6	REFERÊNCIAS	57
	CAPÍTULO III	62
	MAPEAMENTO DE MARCADORES MOLECULARES DE SORGO EM CANA-DE-AÇÚCAR	62
	RESUMO	63
	ABSTRACT	64
1	INTRODUÇÃO	65
2	MATERIAL E MÉTODOS	66
2.1	Material vegetal	66
2.2	Genotipagem da população de mapeamento	67
2.2.1	Extração e quantificação do DNA genômico	67
2.2.2	Marcadores microssatélites genômicos (gSSRs) e derivados de sequências expressas (EST-SSRs) de sorgo	67
2.2.3	Amplificação dos COS-EST-SSR	67
2.3	Eletroforese em gel de poliacrilamida desnaturante	68
2.4	Fenotipagem da população de mapeamento	68
2.4.1	Delineamento experimental	68
2.4.2	Análise fenotípica	68
2.5	Análise segregação de marcadores em dose única	71
2.6	Construção do mapa genético	72
2.7	Análise de associação	72
3	RESULTADOS	73
3.1	Marcadores em dose única gSSRs, EST-SSRs de sorgo e COS-EST-SSRs	73
3.2	Mapa de ligação	75

3.3	Comparação entre os grupos de ligação de cana-de-açúcar e o mapa consenso de sorgo	75
3.4	Associação dos marcadores de sorgo com os caracteres fenotípicos	79
3.4.1	Pol%Cana	81
3.4.2	Fibra%.....	82
4	DISCUSSÃO	83
5	CONCLUSÃO.....	87
6	REFERÊNCIAS	88

Mapeamento genético de microssatélites funcionais de sorgo (EST-SSRs) em cana-de-açúcar

RESUMO - O genoma do sorgo tem sido utilizado como referência em estudos de genômica funcional e comparativa em cana-de-açúcar. A utilização do genoma do sorgo, em estudos de mapeamento comparativo, tem contribuído para auxiliar a detecção de QTL (locus de caracteres quantitativos) em cana-de-açúcar. Marcadores microssatélites de sorgo podem ser empregados no mapeamento comparativo com a cana-de-açúcar, assim como, o desenvolvimento de marcadores microssatélites a partir de sequências ortólogas conservadas (COS - *Conserved Orthologous Sequences*) podem servir de “marcadores âncoras” entre estas duas espécies. No presente trabalho, marcadores microssatélites de sorgo foram incorporados em um mapa prévio de ligação derivado do cruzamento entre o clone IACSP95-3018 e a cultivar IACSP93-3046, ambos do Programa de Melhoramento Genético de Cana-de-Açúcar do Instituto Agronômico de Campinas (IAC). Paralelamente, um conjunto de marcadores microssatélites obtidos a partir da identificação de sequências expressas ortólogas conservadas (COS) entre cana-de-açúcar e sorgo foram desenvolvidos e avaliados quanto ao seu potencial de polimorfismo e mapeamento em cana-de-açúcar. Os marcadores obtidos também foram utilizados para identificar associações putativas aos parâmetros de qualidade (Fibra% e Pol%Cana). Um total de 41 pares de microssatélites COS-EST-SSRs foi desenhado, dos quais 34 produziram amplicons em um grupo de 24 genótipos de *Saccharum* e *Sorghum*. O número de alelos variou de 2 a 15 com valor médio de PIC (conteúdo de informação de polimorfismo) de 0,71. Os 29 pares de *primers* de microssatélites de sorgo, juntamente com os 41 pares de *primers* COS-EST-SSRs genotipados na população de mapeamento produziram um total de 120 marcadores polimórficos, dos quais 92 segregaram em dose única. Destes 92, 52 (56,5%) foram incorporados ao mapa prévio originando 138 grupos de ligação agrupados em 12 grupos de homologia. A maioria das associações entre marcadores e o caractere Pol%Cana detectadas por meio da análise de marcas simples foram de efeito positivo contribuindo para o aumento deste caractere sendo apenas duas com contribuição negativa. De forma geral, os marcadores derivados de sorgo foram promissores no mapeamento da população aqui apresentada.

Palavra-chave: colinearidade, genes ortólogos, mapeamento comparativo, marcador molecular, microssatélites

Genetic mapping of sorghum functional microsatellites (EST-SSRs) markers in sugarcane

ABSTRACT- Sorghum genome has been used as a reference in functional and comparative genomics studies of sugarcane. The sorghum genome has aided to colocalize QTL (quantitative trait locus) in sugarcane through comparative mapping approach. Sorghum microsatellite markers, as well as, the development of microsatellite markers derived from orthologous conserved sequences (COS) can serve as "anchors markers" in comparative mapping between these two species. In the present study, sorghum microsatellite markers were included in a previous map derived from a bi-parental cross between a clone (IACSP95-3018) and a sugarcane cultivar (IACSP93-3046) from the Instituto Agronômico de Campinas (IAC) sugarcane breeding program. A set of microsatellite markers derived from conserved orthologous (COS-EST-SSRs) sequences between sugarcane and sorghum was developed and evaluated for their polymorphism and mapping potential in sugarcane. These markers were also used to identify putative associations for sugarcane quality parameters (fiber% and Pol% Cane). A total of 41 COS-EST-SSRs primer pairs was designed, of which 34 produced amplicons in a group of 24 genotypes of *Saccharum* and *Sorghum*. The number of alleles ranged from 2 to 15 with an PIC (polymorphism information content) average value of 0.71. The 29 sorghum SSR primer pairs along with the 41 COS-EST-SSR primer pairs was genotyped in the mapping population and produced a total of 120 polymorphic markers, of which 92 were single dose markers. Of these, 52 (56.5%) were added to the previous map generating 138 linkage groups grouped into 12 homology groups. The majority of the single marker-trait associations detected for Pol% Cane showed positive effects contributing to increase this trait (only two showed negative contribution to Pol% Cane). In general, the sorghum derived markers were promising in the genetic mapping of the bi-parental population evaluated.

Keyword: collinearity, orthologous genes, comparative mapping, molecular markers, microsatellites

LISTA DE TABELAS

CAPÍTULO II - IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MARCADORES COS-EST-SSRS ENTRE CANA E SORGO

Tabela 1 -	Genótipos utilizados na avaliação dos pares de <i>primers</i> flanqueadores de microssatélites derivados de sequências de etiquetas expressas ortólogas conservadas (COS-EST-SSRs) entre cana e sorgo.....	39
Tabela 2 -	Identificação das sequências ortólogas conservadas na base de dados SUCEST, iniciadores COS-EST-SSRs com suas respectivas sequências (forward/reverse), temperatura de anelamento (T _m), tamanho do produto esperado (TPes) e similaridade esperada.....	43
Tabela 3 -	Classes de motivos de COS-EST-SSR e respectiva porcentagem de amplificação nos gêneros de <i>Saccharum</i> e <i>Sorghum</i>	50
Tabela 4 -	Identificação dos COS-EST-SSRs, número de alelos (NA), variação do tamanho (pb), conteúdo de informação polimórfica (PIC) e heterozigosidade esperada (He) dos marcadores COS-EST-SSRs avaliados em 24 genótipos de cana e sorgo.....	51

CAPÍTULO III - MAPEAMENTO E ASSOCIAÇÃO DE MARCADORES COS-EST-SSRs, gSSR e EST-SSR DE SORGO EM CANA-DE-AÇÚCAR

Tabela 1 -	Relação dos pares de <i>primers</i> sequências forward e reverse, temperatura de anelamento (T _a) e tamanho do produto amplificado para os gSSRs e EST-SSRs avaliados.....	69
Tabela 2 -	Número de marcadores em dose única obtidos por EST-SSR e gSSR de sorgo e COS-EST-SSRs em cruzamento bi-parental.....	74
Tabela 3 -	Análise descritiva dos dados fenotípicos da população de mapeamento derivada de um cruzamento bi-parental entre o clone IACSP95-3018 e a cultivar IACSP93-3046.....	79
Tabela 4 -	Efeito de marcadores COS-EST-SSRs, gSSR e EST-SSRs associados a variável Pol%Cana avaliados na população derivada do cruzamento bi-parental.....	81

- Tabela 5** - Percentagem da variação fenotípica explicada pelas associações ($R^2\%$) parciais e múltiplas (R^2 total) para o carácter Pol%Cana para marcadores do tipo COS-EST-SSRs, gSSR e EST-SSRs via procedimento *stepwise* do SAS..... 82
- Tabela 6** - Efeito de associações ($R^2\%$) parciais e múltiplas (R^2 total) para o carácter Fibras para marcadores do tipo COS-EST-SSRs, gSSR e EST-SSRs via procedimento *Stepwise* do SAS..... 83

LISTA DE FIGURAS

CAPÍTULO I – INTRODUÇÃO GERAL

- Figura 1 -** Duplicação do gene. A e A' são ortólogos porque são relacionados por ascendência. Adaptada de Neale e Krutovsky (2005)..... 16

CAPÍTULO II - IDENTIFICAÇÃO E CARACTERIZAÇÃO DE MARCADORES COS-EST-SSRS ENTRE CANA E SORGO

- Figura 1 -** Identificação de sequências EST-SSR ortólogos conservados (COS-EST-SSR) entre os genomas de cana e sorgo. **Fonte:** Adaptado de Vicentini et al. (2012)..... 47
- Figura 2 -** Alinhamento realizado utilizando a ferramenta Clustal W: sequências ortólogas de cana em sorgo (SCAGLB1070D03.g-Sb04g000230.1, SCCCLR1069H09.g-Sb02g041350.1, SCCCL3005C02.g-Sb01g030835.1, SCCRZ2C04A03.g-Sb01g019290.1 e SCJLFL4187G07.g-Sb01g011810.1, respectivamente). Alinhamento utilizado na obtenção de *primers* COS-EST-SSRs: (A) alinhamento perfeito na região complementar ao primer forward (retângulo verde), na região do COS-EST-SSR (retângulo vermelho) e na região complementar ao primer reverse (retângulo azul). Alinhamentos não utilizados na obtenção de *primers* COS-EST-SSRs: (B) *Gaps* e incompatibilidade de nucleotídeos na região complementar ao *primer* (retângulo verde); região do COS-EST-SSR (retângulo vermelho); (C) COS-EST-SSR localizados no início de uma sequência alinhada; (D) Alinhamento com tamanho inferior a 150 pb e (E) COS-EST-SSR localizados no final de uma sequência alinhada..... 49
- Figura 3 -** Análise eletroforética dos produtos amplificados com *primers* COS-EST-SRR separados em gel de poliacrilamida desnaturante 6% e visualizados por meio de coloração com prata. M: marcador 10-pb; híbridos de sorgo (1: 1G100; 2: 1G220; 3: 1G233; 4: 1G244; 5: 1G282; 6: SS318; 7: 50A40); cultivares comerciais de cana (8: IACSP95-5000; 9: RB867517; 10: SP801842; 11: SP803250; 12: RB835054; 13: RB855453); clones de *S. officinarum* (14: Ajax; 15: Badila; 16: Badila de java; 17: Bourbon sur; 18: Creola; 19: Sabura); clones de *S. spontaneum* (20: Dacca; 21: Djantoerll; 22: Formosa; 23: Gehrabon e 24: Holes1)..... 50

- Figura 4 -** Dendrograma (UPGMA) baseado na dissimilaridade estimada pelo complemento do coeficiente de similaridade de Jaccard para 24 genótipos..... 52

CAPÍTULO III - MAPEAMENTO E ASSOCIAÇÃO DE MARCADORES COS-EST-SSRs, gSSR e EST-SSR DE SORGO EM CANA-DE-AÇÚCAR

- Figura 1 -** Padrão de amplificação obtido para o locus COS-EST-SSR2, por eletroforese em gel de poliacrilamida. M-marcador de peso molecular (10pb); LS-linhagem de sorgo; P1-genitor feminino (IACSP95-3018); P2-genitor masculino (IACSP93-3046) e 54 indivíduos da população de mapeamento..... 74
- Figura 2 -** Mapa de ligação de cana-de-açúcar integrado para a população derivada do cruzamento bi-parental entre os genitores IACSP95-3018 e IACSP93-3046. Os números superiores referem-se ao grupo de ligação (GL). Os números à esquerda são as distâncias em centiMorgans (cM) entre os marcadores. Os nomes dos marcadores estão à direita, seguidos pela sigla D1 (presente no IACSP95-3018), D2 (presente no IACSP93-3046) e C (presente em ambos os genitores). Os GL foram reagrupados com base na presença de marcadores derivados de um mesmo loco de microssatélite (loco definido pelo par de *primers* que flanqueia o microssatélite) formando os grupos de homologia (GH)..... 76
- Figura 3 -** Alinhamento dos grupos de ligação obtido a partir do cruzamento entre IACSP95-3018 e IACSP93-3046 com o mapa consenso do sorgo (http://sorgblast3.tamu.edu/linkage_groups.htm), MACE et al. (2009). A identificação do locus da cana e do sorgo encontra-se a direita do grupo de ligação de cada espécie. As linhas na diagonal entre os grupos de ligação da cana e do sorgo indicam a posição correspondente do locus..... 80

CAPÍTULO I

INTRODUÇÃO GERAL

1 INTRODUÇÃO GERAL

Dentre os diferentes tipos de marcadores moleculares disponíveis para a análise em plantas, marcadores do tipo microssatélites tem-se mostrado úteis em diversas aplicações, tanto na genética quanto no melhoramento, principalmente devido a sua reprodutibilidade e abundância no genoma. Estes marcadores representam sequências repetitivas compostas por motivos (1 a 6 pares de bases) repetidos lado a lado, e cujas sequências que o flanqueiam apresentam-se conservadas a ponto de permitir o desenho de um par de iniciadores de síntese específica (*primers*) para a sua amplificação via reação em cadeia da polimerase (*Polimerase Chain Reaction*).

Os Projetos de Sequenciamento das Etiquetas de Sequências Expressas (ESTs) de culturas de importância, tais como a cana-de-açúcar (<http://sucest.lad.dcc.unicamp.br/en/>) e o sorgo (<http://www.phytozome.net>) têm permitido o desenvolvimento de microssatélites de forma rápida e econômica pela simples mineração de sequências repetitivas nos bancos de dados. Tais microssatélites por derivarem de ESTs são chamados de microssatélites funcionais (EST-SSRs) e são considerados de grande utilidade para acessar a diversidade a nível de genes expressos em coleções de germoplasma e também no mapeamento genético, visto que o polimorfismo gerado permite o mapeamento direto de genes expressos, os quais podem proporcionar a identificação de marcadores funcionais associados a características de interesse.

Os microssatélites derivados de sequências expressas (EST-SSRs) apresentam uma maior transferibilidade entre espécies correlacionadas podendo ser utilizados como marcadores âncoras no mapeamento comparativo (CABRERA et al., 2009).

Os bancos de dados de sequências expressas também têm permitido o desenvolvimento de marcadores microssatélites a partir de sequências ortólogas conservadas (*Conserved Orthologous Sequences-COS*) entre espécies correlacionadas. Os COS são definidos como uma coleção de genes de cópia única que são ortólogos entre espécies (JEONG et al., 2014). Genes de cópias únicas possibilitam o desenvolvimento de marcadores universais (âncoras) ou os chamados

Conserved Orthologous Markers. De acordo com esses autores, esses marcadores provenientes de sequências ortólogas conservadas podem ser utilizados em estudos de filogenia e mapeamento comparativo.

A cana-de-açúcar pertence à tribo *Andropogoneae*, família das gramíneas, cuja respectiva tribo, inclui também as gramíneas tropicais e subtropicais, tais como os cereais do gênero *Sorghum* e *Zea*. As cultivares atuais são híbridos interespecíficos complexos de elevado nível de ploidia ($2n = 100$ a 120 cromossomos) apresentando, em alguns casos, aneuploidia com segregação irregular (D'HONT et al., 1996; PIPERIDIS et al., 2010; SOBHAKUMARI, 2013). A complexidade e o conteúdo cromossômico variável do genoma da cana são as maiores preocupações no procedimento de sequenciamento do genoma (DILLON et al., 2007).

A complexidade genômica da cana-de-açúcar dificulta os trabalhos de mapeamento genético e a detecção de características de interesse agrônomo aos programas de melhoramento, especialmente quando os mapas desenvolvidos são construídos utilizando-se cultivares comerciais como genitores. Segundo Le Cunff et al. (2008) a otimização de sistemas de marcadores moleculares contribui para aumentar a eficiência no desenvolvimento de mapas genéticos de cana-de-açúcar e sua utilização no mapeamento de QTL, o que pode proporcionar avanços em estudos genéticos, genômicos e aplicações nos programas de melhoramento genético (JAMES et al., 2011).

Por outro lado, o sorgo é uma espécie diplóide ($2n=20$), com um genoma relativamente pequeno e com uma vasta adaptação a ambientes severos e tolerância às condições de estresse bióticos e abióticos, o que contribui para que seja apontada como uma importante cultura modelo (SINGH et al., 2013b). A utilização do genoma pequeno do sorgo como referência em estudos de mapeamento comparativo tem contribuído para acelerar a análise do genoma da cana-de-açúcar, principalmente em relação à detecção de QTL relacionados à resistência a doenças (MCINTYRE et al., 2005b) e parâmetros agroindustriais (JORDAN et al., 2004) obtidos através de marcadores provenientes de sondas heterólogas em ensaios de RFLP (*Restriction fragment lenght polymorphism*).

Nesse contexto, esse estudo apresenta o desenvolvimento de marcadores microsatélites ortólogos conservados (COS-EST-SSRs) entre cana e sorgo bem

como o mapeamento destes COS-EST-SSRs e de EST-SSRs de sorgo em uma população de mapeamento genético de cana-de-açúcar do Programa Cana IAC. Para o nosso conhecimento, até o presente momento, não foram desenvolvidos marcadores microssatélites ortólogos conservados derivados de sequências expressas (COS-EST-SSRs) entre cana-de-açúcar e sorgo.

2 REVISÃO DE LITERATURA

2.1 Aspectos gerais da cultura da cana-de-açúcar

A cana-de-açúcar (*Saccharum* spp.) apresenta um papel importante na economia mundial. Seu cultivo ocorre principalmente em regiões de clima tropical e subtropical, sendo a principal fonte de produção de sacarose e biocombustíveis renováveis. O Brasil ocupa o primeiro lugar na produção mundial de açúcar e etanol, com uma área de colhida estimada em 9073,7 mil ha (Conab, 2016) e vem conquistando, cada vez mais, o mercado externo com o uso do biocombustível como alternativa energética. De acordo com o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o Brasil é responsável por mais da metade do açúcar comercializado no mundo, podendo alcançar taxa média de aumento da produção de 3,25%, até 2018/19, com uma safra estimada de 47,34 milhões de toneladas de cana.

A cana-de-açúcar é uma gramínea pertencente ao gênero *Saccharum*, família Poaceae. Este gênero pertence à tribo Andropogoneae, o qual inclui também o gênero *Sorghum* (DANIELS; ROACH, 1987). Dentre as espécies que compõe o gênero *Saccharum* seis estão diretamente relacionadas com a cana-de-açúcar. *Saccharum officinarum* L. é conhecida como cana nobre por ser rica em sacarose, apresenta colmos mais grossos (acima de 3,5 cm de diâmetro) e macios devido ao menor teor de fibra, enquanto *Saccharum spontaneum*, conhecida como a espécie selvagem apresenta colmos finos, é rica em fibra, tem alto vigor e adaptabilidade a diferentes condições ambientais. Além das espécies *S. officinarum* e *S. spontaneum*, o gênero consta com mais 4 espécies: *S. edule*, *S. barberi*, *S. sinensis* e *S. robustum* (SOBHAKUMARI, 2013).

As cultivares modernas de cana-de-açúcar são resultantes de vários ciclos de cruzamentos e retrocruzamentos, iniciados no final do século XIX, envolvendo predominantemente, as espécies *S. officinarum* L. ($2n=80$) e *S. spontaneum* ($2n=40$ a 120) (SINGH et al., 2013a). Os sucessivos retrocruzamentos dos híbridos resultantes com a espécie *S. officinarum* L., utilizada como genitor feminino, realizados para se atingir níveis desejáveis de sacarose, fez com que as cultivares atuais apresentassem um alto grau de poliploidia e frequente aneuploidia (HOARAU et al., 2001). Segundo Piperidis et al. (2010), a hibridação interespecífica proporcionou maior resistência a patógenos, além de aumentar o vigor e a adaptabilidade das cultivares atuais. As cultivares atuais apresentam cerca de 70-80% dos cromossomos de *S. officinarum*, 10-20% de *S. spontaneum* e o restante de recombinações entre estas duas espécies (D'HONT et al., 1996). Além disso, os genótipos cultivados de cana-de-açúcar apresentam um elevado nível de heterozigosidade devido ao alto grau de ploidia. Sendo assim, as cultivares modernas de cana-de-açúcar tem uma estrutura genômica complexa, composta de 10-13 cópias de cromossomos homeólogos (MOYLE; BIRCH, 2013).

Considerando sua importância mundial, da década de 70 até os dias atuais, a cana-de-açúcar tem se destacado como uma alternativa de energia renovável. Além da produção de etanol, a partir do caldo de cana ou do melaço, o bagaço de cana também possui potencial para ser usado diretamente na produção de etanol (BARNABAS et al., 2015).

Apesar dos avanços na biotecnologia, os programas de melhoramento da cana-de-açúcar tem buscado ferramentas que auxiliem o melhorista na condução de metodologias que ajudem a identificar genes inéditos e de interesse econômico. Desde o final da década de 80, os melhoristas e geneticistas têm explorado o uso de várias ferramentas para facilitar o processo de obtenção de novas cultivares de cana, entre elas o uso de marcadores moleculares baseados em DNA.

Diversos tipos de marcadores têm sido utilizados em cana-de-açúcar, entre esses podemos citar AFLP (*Amplified Fragment Length Polymorphism*), TRAP (*Targed Region Amplification Polymorphism*), SNPs (*Single Nucleotide Polymorphism*) e SSRs (*Simple Sequence Repeats*) (Pan, 2006). Os marcadores microsatélites (SSRs) são úteis para a avaliação da diversidade genética e

identificação varietal, por serem abundantes no genoma, apresentarem herança co-dominante e alta reprodutibilidade (CORDEIRO et al., 2000; CORDEIRO et al., 2001).

2.2 Marcadores microssatélites derivados das etiquetas de sequências expressas (EST-SSRs) e transferibilidade entre gêneros correlatos

O uso de marcadores moleculares no processo de seleção pode ser de imenso valor no aumento da eficiência da seleção e na redução do tempo de obtenção de cultivares de cana (SINGH et al., 2013a) em programas de melhoramento.

Uma grande variedade de marcadores moleculares como, RFLP (PINTO et al., 2010), AFLP (LIU et al., 2016), SNP (MCINTYRE et al., 2015), SSR (BAER; LALUSIN, 2013) têm sido desenvolvidos e utilizados em estudos de análises genéticas de cana-de-açúcar. Entre as diferentes classes de marcadores moleculares, os marcadores microssatélites têm uma variedade de aplicações no melhoramento de plantas devido sua reprodutibilidade, natureza multi-alélica, heranças co-dominante e abundância no genoma (SINGH et al., 2013a).

Dependendo da fonte de dados, dois tipos de marcadores microssatélites são conhecidos, os microssatélites genômicos (gSSRs) e os microssatélites derivados de etiquetas de sequências expressas (EST-SSRs). Os gSSRs são os mais abundantes devido estarem localizados em regiões não codante do genoma (ADAL et al., 2015). Por outro lado, os EST-SSRs são repetições localizadas em regiões expressas do genoma e podem ser obtidos a partir da mineração em bases de dados de ESTs.

Geralmente, gSSRs não apresentam nenhuma função gênica, nem ligação íntima com a transcrição de regiões, enquanto que os EST-SSRs são potencialmente localizados dentro de genes que apresentam pelo menos uma função putativa conhecida e que podem controlar algumas características agronômicas. Esses marcadores tendem a ser mais prontamente transferíveis entre espécies ou gêneros aparentados, uma vez que as sequências de codificação são mais conservadas do que as sequências não-codificantes (BOCCACCI et al., 2015).

Apesar do número de repetições e do comprimento total, os EST-SSRs são relativamente pequenos em comparação com gSSR (GAO et al., 2004).

Os projetos de sequenciamento de ESTs tem gerado uma vasta quantidade de informação sobre sequência que podem ser utilizadas para a identificação de genes, estudos genéticos funcionais e desenvolvimento de marcadores (PASHLEY et al., 2006). Os EST-SSRs podem ser rapidamente desenvolvidos a baixo custo a partir da análise de dados de EST *in silico*, e devido à sua presença em regiões expressas, eles podem levar ao desenvolvimento de mapas genéticos que podem aumentar a eficiência da seleção assistida por marcadores (SAM) através da utilização de genes candidatos. Zhang et al. (2005) a partir de uma análise *in silico* em trigo, transferiram 73,2% de EST-SSRs de trigo para outras espécies diploides relacionadas. O potencial destes marcadores é devido à sua associação com regiões codantes do genoma o que pode favorecer a detecção de características agronomicamente importantes (GADALETA et al., 2009). Além disso, se um marcador EST está geneticamente associado com a característica de interesse, é possível que o gene mapeado afete diretamente a característica. Segundo Cato et al. (2001), quando os ESTs apresentam homologia a genes candidatos ou que estejam diferencialmente expressos no tecido de interesse, podem ser utilizados no mapeamento genético.

Depois de identificar sequências de microssatélites, iniciadores específicos (*primers*) que flanqueiam estas repetições podem ser desenhados e a região, contendo o microssatélite, amplificada. Esta estratégia elimina a necessidade de sequenciamento do DNA, bem como todas as etapas iniciais da estratégia tradicional, que tornam o processo de desenvolvimento de microssatélite uma tarefa trabalhosa. Sabe-se que, ao longo da evolução, as sequências da maioria dos genes permaneceram altamente conservadas, de modo que os marcadores pertencentes às regiões transcritas do genoma são facilmente transferidos entre espécies relacionadas e podem ser utilizados na construção de mapas genéticos comparativos, facilitando o estudo da sintenia e colinearidade entre genomas relacionados (GUPTA et al., 2003). A sintenia é um termo utilizado para descrever genes localizados em um mesmo cromossomo, enquanto colinearidade é a conservação da ordem desses genes em um mesmo cromossomo.

Para a cana-de-açúcar, o estudo de sintenia fornece uma ferramenta poderosa para compreensão e manipulação da biologia da cana (SINGH et al., 2013a).

A transferibilidade de marcadores microssatélites depende do grau de relacionamento genético entre as espécies avaliadas, o qual inclui as diferenças na sequência do DNA, tamanho do genoma e taxa de evolução.

Diversos trabalhos têm demonstrado a transferibilidade de microssatélites EST-SSRs para espécies ou gêneros relacionados (PANDIAN et al., 2000; HERNÁNDEZ et al., 2001; VARSHNEY et al., 2005; DAI et al., 2013; SINGH et al., 2013a). Estes marcadores apresentam algumas vantagens em relação a outros marcadores moleculares: a) são amplamente transferíveis entre espécies, e até mesmo entre plantas de gêneros diferentes (CHAGNE et al., 2004; FRASER et al., 2005); b) por estarem localizados em genes transcritos, a identificação de diferentes *loci* de EST-SSRs em análises do genoma pode identificar diretamente genes candidatos para seleção (VASEMAGI, 2005; KANE; RIESEBERG, 2007); c) os EST-SSRs muitas vezes exibem níveis reduzidos de polimorfismo em comparação com SSRs genômicos (EUJAYL et al., 2002; CHABANE et al., 2005).

Na família das Poaceae, a transferibilidade de marcadores microssatélites tem sido estimada ao redor de 57% (GUPTA et al., 2003; SAHA et al., 2004; WANG et al., 2005). Hernández et al. (2001) observaram uma taxa de transferibilidade de 75% de *loci* EST-SSRs de milho para *Miscanthus*. Singh et al. (2013a), desenvolveram marcadores EST-SSRs e relataram que estes foram altamente transmissíveis dentro de um grupo do complexo *Saccharum* (84,1% em *S. barberi* a 93,4% em *S. officinarum*), em outros gêneros relacionados (80,6% em *Narenga* para 87,7% em *Erianthus*) e em cereais (76,7% em *Pennisetum* a 80,6% em Sorgo). Singh et al. (2013) desenvolveram 227 EST-SSRs de cana-de-açúcar e encontraram percentagens de transferibilidade, respectivamente, na ordem de 80,6%, 79,3%, 78,4%, 76,7% para sorgo, milho, arroz e milheto.

Relações filogenéticas entre diferentes espécies do gênero *Saccharum* e outras tribos aparentadas da família Poaceae foram avaliadas através da transferibilidade de *primers* de microssatélites de cana. Foi revelado 100% de transferência em todas as variedades de cana e 74,07% em plantas de sorgo, utilizando 27 *primers* de microssatélites de cana (SINGH et al., 2011).

Parmessur et al. (2011) utilizaram marcadores EST-SSR de sorgo para mapear fisicamente, *in sílico*, o genoma do sorgo e foi encontrado sintenia e colinearidade em

diversos cromossomos do sorgo, além de similaridade entre os genomas de sorgo e cana revelando transferência do marcador da ordem de 95% entre as duas espécies.

Dillon et al. (2007) relataram altas taxas de transferência dos marcadores SSRs derivados de *S. bicolor* para espécies selvagens e seu alto nível de diversidade, sugere que estes SSRs são um recurso eficiente e altamente informativos para espécies de sorgo não domesticadas.

As condições de amplificação, tais como a quantidade de DNA e temperatura de anelamento, também influenciam na transferibilidade, sendo que este último constitui um dos parâmetros mais críticos, além do tamanho do produto amplificado apresentar ou não o mesmo tamanho do produto amplificado pelo *primer* desenhado na espécie original (WANG et al., 2005).

O mapeamento de 9 *loci* EST-SSRs de cevada em uma população de centeio mostrou que 5 *loci* foram mapeados em braços de cromossomos homólogos enquanto quatro *loci* foram mapeados em braços de cromossomos não homólogos em translocações conhecidas entre os genomas de cevada e centeio, os quais representam blocos sintênicos de ligação (KORZUN et al., 2001). Yu et al. (2004), utilizaram com sucesso *loci* de EST-SSRs de arroz no mapeamento da população de referência de trigo derivada do cruzamento entre W7984 x Opata 85, e marcadores EST-SSRs de trigo no mapeamento de 46 linhagens duplo-haplóides derivadas do cruzamento entre IR64 x Azucena, bem como marcadores COS-ESTs desenvolvidos com base em ambas as espécies. Eles afirmam que os marcadores âncoras identificados pelos EST-SSRs suportaram os padrões prévios de conservação entre os genomas de trigo e arroz.

2.3 Mapeamento genético

Os mapas genéticos baseados em marcadores de DNA, constituem uma ferramenta eficiente para muitas aplicações no melhoramento de plantas. É apresentado como uma ferramenta valiosa para identificar fatores genéticos que são quantitativamente herdados ou expressos apenas em condições particulares. Os marcadores genéticos ligados a características agronomicamente importantes podem

ser utilizados na seleção assistida por marcadores nas fases iniciais de um programa de melhoramento para acelerar a criação de cultivares desejáveis (KIM et al., 2012). Além disso, possibilitam a cobertura e análise completa de genomas, a decomposição de características genéticas complexas nos seus componentes mendelianos, a localização de regiões genômicas que controlam caracteres de importância, a quantificação do efeito dessas regiões na característica estudada e a canalização de toda essa informação para uso em programas de melhoramento (FALEIRO et al., 2003).

A construção de mapas genéticos de ligação, utilizando marcadores moleculares, é um passo necessário para entender a organização genômica e a relação entre genes e fenótipo. Para isto, é preciso encontrar a ordem e o espaçamento correto entre marcas no mapa de uma população. Um mapa genético é uma representação linear ou circular, de modo que marcadores estão ordenados e distanciados ao longo do cromossomo de um organismo (CRANE; CRANE, 2005).

Vários mapas de ligação têm sido construídos para cana-de-açúcar (HOARAU et al., 2001; MING et al., 2002b; AITKEN et al., 2005a; GARCIA et al., 2006; OLIVEIRA et al., 2007; PIPERIDIS et al., 2008; ANDRU et al., 2011; PALHARES et al., 2012; SINGH et al., 2013b; AITKEN et al., 2014a), porém a maioria deles com um pequeno número de marcadores, o que dificulta a cobertura total do genoma. Contudo, a cana-de-açúcar apresenta um genoma grande (10 Gb) (LE CUNFF et al., 2008) e um mapa genético com alta densidade é uma ferramenta valiosa para compreender a genética e organização genômica desta cultura (AITKEN et al., 2014a). Outra particularidade que merece ser salientada é a poliploidia, que tem comprometido o desenvolvimento de mapas genéticos saturados, requerendo mais trabalho do que para espécies diploides (SINGH et al., 2013a). Devido à sua natureza poliploide, a cana-de-açúcar apresenta vários alelos segregando em um mesmo loco, os quais podem estar em dose única ou múltipla. Dispondo destas informações, o estudo de melhoramento genético pode ser otimizado, tanto na eficiência da seleção dos melhores genótipos em uma população com variabilidade, quanto na velocidade que os ganhos genéticos serão obtidos. Isto se deve a possibilidade de seleção com base nos marcadores moleculares, excluindo-se o efeito ambiental sobre a expressão dos diferentes caracteres de interesse e dispensando etapas de avaliação e seleção.

2.4 Construção de mapa genético

Um mapa genético é um diagrama onde são representados os genes com suas respectivas posições no cromossomo, ou seja, a distribuição sequencial dos genes ao longo dos cromossomos (NOMELINI et al., 2009). A partir desta ferramenta, é possível avançar em estudos que permitem compreender a ordem e espaçamento dos marcadores de genomas ainda desconhecidos, bem como comparar com de outras espécies, por meio de mapeamento comparativo e sequências do genoma de outras espécies vegetais.

Na maioria das situações, não sabemos como se comportam os genes e onde estão localizados no genoma. Com o rápido desenvolvimento da tecnologia molecular, grandes quantidades de dados moleculares estão agora disponíveis, fornecendo possibilidades para estimar os efeitos, bem como localização de *loci* de característica quantitativa (LUO; XU, 2003).

A influência do mapa genético não se limita a espécie de plantas cujos genomas não foram ainda sequenciados podendo ser estimado a partir de um conjunto de dados derivado de uma população de mapeamento, que possui certas características que devem ser levados em conta no processo de construção do mapa, como o tipo da população, onde as mais utilizadas são retrocruzamento (RC), F2, duplo-haplóide (DH) e linhagens endogâmicas recombinantes (RILs) (CHEEMA; DICKS, 2009).

A construção de um mapa genético compreende quatro etapas a seguir: a) identificação de marcadores genéticos; b) desenvolvimento de uma população segregante; c) análise da herança dos marcadores genéticos e d) estabelecimento da ordem dos marcadores genéticos e da distâncias entre eles (VIEIRA et al., 2006).

2.5 Mapeamento comparativo

A comparação de mapas genéticos e sequências EST-SSRs são as principais ferramentas utilizadas para entender o genoma da cana, além de permitir a identificação de características de interesse agrônomo. Segundo Yadav et al. (2014) genes candidatos com funções preditas podem ser utilizados como recurso para gerar marcadores moleculares novos e informativos.

O mapeamento genético comparativo entre cereais tem mostrado que não apenas os genes, como também a sua ordem ao longo (AITKEN, 2007; ALWALA et al., 2009) dos cromossomos manteve-se amplamente conservados ao longo da história evolutiva das gramíneas, a ponto do genoma das gramíneas representarem um único sistema genético (GALE; DEVOS, 1998).

Um mapa de consenso das espécies cana-de-açúcar e sorgo revelou um elevado grau de conservação entre os dois genomas, que pode ser fundamental na tradução da informação genética para o desenvolvimento de um mapa genético detalhado (MING et al., 2002b).

De acordo com Dufour et al. (1997) a análise comparativa entre os genomas de sorgo e cana-de-açúcar demonstrou que os cromossomos dessas espécies apresentam forte colinearidade e são altamente sintênicos, levando-os a sugerir que a ordem das sequências de DNA presentes em seus cromossomos tenha sido conservada. Guimarães (1999) conseguiu identificar uma elevada colinearidade em um estudo de mapeamento comparativo envolvendo o genoma do sorgo e espécies de *Saccharum* e relatou que o sorgo pode ser considerado a espécie diploide filogeneticamente mais próxima da cana.

O mapeamento comparativo em cana-de-açúcar, utilizando sondas heterólogas de milho e sorgo, mostrou que, em relação à organização dos cromossomos, os genomas da cana-de-açúcar e do sorgo estão muito mais relacionados entre si quando comparado com o do milho (GRIVET et al., 1994; DUFOUR et al., 1997). Essa proximidade tem permitido, através de sondas heterólogas de sorgo, ancorar os mapas genéticos de cana-de-açúcar nos mapas de sorgo (GUIMARÃES et al., 1997; MING et al., 1998) proporcionando a utilização do genoma do sorgo na identificação de genes de interesse em cana-de-açúcar. Sendo assim, o sorgo parece ser uma excelente cultura modelo para a cana, devido à colinearidade entre seus cromossomos (GRIVET; ARRUDA, 2001).

Wang et al. (2010) utilizando genes ortólogos como âncoras, compararam sequências de cana com o genoma do sorgo e afirmaram que a maioria das regiões alinhadas compartilham uma média de 95,2% de identidade de sequências, revelando que o genoma da cana exibiu um alto grau de colinearidade com o genoma do sorgo.

De acordo com esses autores, o genoma do sorgo fornece uma oportunidade excepcional para desvendar o genoma complexo da cana-de-açúcar.

Asnaghi et al. (2000) utilizaram da sintenia entre cana-de-açúcar e outras gramíneas como arroz, milho e sorgo para saturar a região ao redor de um gene de resistência a ferrugem, o qual mostrou associação à sonda de cana-de-açúcar CDSR29. McIntyre et al. (2005a) ressaltaram a importância do mapeamento simultâneo em cana-de-açúcar e sorgo visto que três sondas de genes análogos de resistência associadas à ferrugem na população derivada do cruzamento Q117 x 74C42 em cana-de-açúcar mapearam em sorgo no mesmo grupo de ligação. A localização de características agroindustriais em cana-de-açúcar, como perfilhamento e brotação da soqueira, também puderam ser determinadas pela utilização de sondas heterólogas de sorgo associadas à estas mesmas características (JORDAN et al., 2004).

2.6 Mapeamento de QTL em cana-de-açúcar

Na última década, a cana-de-açúcar tem se destacado na produção de energia renovável e despertado o interesse e curiosidade da comunidade científica em solucionar questões que favoreçam a obtenção de genótipos superiores. No entanto, seus esforços não têm sido suficientes para superar as dificuldades, relacionadas principalmente às interações genótipo-ambiente.

Como uma alternativa viável, os marcadores moleculares, vem sendo utilizado com sucesso no melhoramento de plantas, auxiliando na seleção indireta de características de interesse agrônomo. No melhoramento genético de plantas, tem-se buscado, cada vez mais, a seleção assistida por marcadores moleculares, visando obter maior eficiência na transferência de fatores genéticos. Os marcadores moleculares são ferramentas úteis para detectar variações no genoma e estimar parâmetros usados na seleção indireta, aumentando o poder da análise genética das plantas (CAIXETA et al. 2006)

Numa população de mapeamento bi-parental, um marcador em dose única (MDU) tem apenas uma única cópia de um alelo de um pai segregando em 1:1

(presença:ausência) ou uma única cópia do mesmo alelo em ambos os pais segregando 3:1 (presença:ausência) (SINGH et al., 2013b). Segundo Wu et al. (2002) os *loci* que apresentam padrões de segregação 1:1 ou 3:1 são mais informativos na construção de mapa genético do que aqueles com padrões de segregação mais complexos.

A descoberta e o mapeamento de genes responsáveis pelo controle de características qualitativas ou quantitativas através do emprego de marcadores moleculares, tem complementado mapas de ligação. Os mapas genéticos baseados em marcadores moleculares possibilitam a localização de genes nos cromossomos, a contagem aproximada dos principais *loci* envolvidos no controle da característica e a estimativa da magnitude do efeito destes *loci* no fenótipo de interesse (FERREIRA; GRATTAPAGLIA 2006).

Ming et al. (2002c) avaliando uma população segregante derivada do cruzamento entre as espécies *S. officinarum* e *S. spontaneum*, detectaram QTL associados significativamente aos caracteres peso, número de colmos, brix, altura de planta e florescimento.

A identificação de marcadores associados a características relacionadas a parâmetros tecnológicos, tais como brix e pol, por exemplo, pode contribuir para melhorar o teor de sacarose na cana-de-açúcar. De modo geral, há poucos trabalhos sobre QTL que controlam o conteúdo de sacarose, características relacionadas a produção de açúcar e componentes de produção em cana. LIU et al. (2016) identificaram 24 QTL associados ao teor de sacarose em uma população de 237 indivíduos. Já Diola et al. (2014) relataram 99 *loci* associados a caracteres agrônômicos, sendo que 21 estiveram associados com brix, 22 com peso seco, 12 com perfilhamento, 16 com a angulação pronunciada da touceira, 12 com amadurecendo precoce e 16 associações com o caractere peso fresco. McIntyre et al. (2015) analisando QTL para parâmetros tecnológicos, como teor de açúcar e fibra, identificaram associações de QTL com efeito pequeno, porém a maioria dos caracteres apresentaram herança quantitativa em cana-de-açúcar. Pinto et al. (2010) afirmam que progênies derivadas a partir de clones elite têm a vantagem de mapear QTL associados a alelos favoráveis para as características de interesse, uma vez que estas progênies foram submetidas a vários ciclos de seleção.

Apesar de vários QTL terem sido identificados para diversas características agronômicas de importância em cana-de-açúcar, a grande maioria deles explicam apenas uma pequena parte da variação observada para as características quantitativas avaliadas (MING et al., 2002a; MING et al., 2002c). Os resultados obtidos a partir da análise de marcadores moleculares associados a características de interesse são de grande importância para os melhoristas, uma vez que permitem a identificação de *loci* que controlam a expressão de genes. Essa estratégia tem sido utilizada com sucesso para resolver ou diminuir problemas relacionados a interação do ambiente, por exemplo, na seleção de fenótipos superiores em programas de melhoramento.

2.7 Desenvolvimento de marcadores microssatélites a partir de sequências ortólogas conservadas (COS) entre cana e sorgo

Os marcadores moleculares desenvolvidos a partir de técnicas baseadas no uso da PCR podem ser úteis na identificação de genes responsáveis pela expressão de uma dada característica de interesse para o melhorista. Recentemente, ferramentas de bioinformática tem sido utilizada para explorar bases de dados que comportam sequências de ESTs, além de facilitar o desenvolvimento de marcadores SSRs. A identificação de genes ortólogos conservados de cópia única visando o desenvolvimento de marcadores universais foi relatado em algumas famílias de plantas cultivadas, como Rosaceae (CABRERA et al., 2009), Malvaceae (KUHN et al., 2012), Poaceae (LIU et al., 2013; TERRACCIANO et al., 2013), Zamiaceae (SALAS-LEIVA et al., 2014), Fabaceae (CHAPMAN, 2015), Solanaceae (STANIASZEK et al., 2014; VILLANO et al., 2016).

Os genes são tidos como homólogos quando compartilham uma mesma origem evolucionária. Genes homólogos, em diferentes espécies, são definidos como ortólogos quando os mesmos descendem de um único gene em relação ao último ancestral comum (FULTON et al., 2002; FORSLUND et al., 2011). Os *loci* ortólogos são aqueles que estão relacionados evolutivamente e que divergiram de seu ancestral comum mais recente por um processo de especiação (Figura 1) (DEWEY, 2012).

Desta forma, uma coleção de genes que são conservados em sequência e número de cópias ao longo da evolução tem sido designada como um conjunto de ortólogos conservados (FORSLUND et al., 2011). De acordo com esses autores, esses marcadores provenientes de sequências ortólogas conservadas podem ser utilizados em estudos de filogenia e mapeamento comparativo.

Desta forma, uma coleção de genes que são conservados em sequência e número de cópias ao longo da evolução tem sido designada como um conjunto de ortólogos conservados (FORSLUND et al., 2011). De acordo com esses autores, esses marcadores provenientes de sequências ortólogas conservadas podem ser utilizados em estudos de filogenia e mapeamento comparativo.

Estudos comparativos têm revelado grande colinearidade entre os genomas, demonstrando que um número reduzido de rearranjos cromossômicos distingue espécies e gêneros correlacionados.

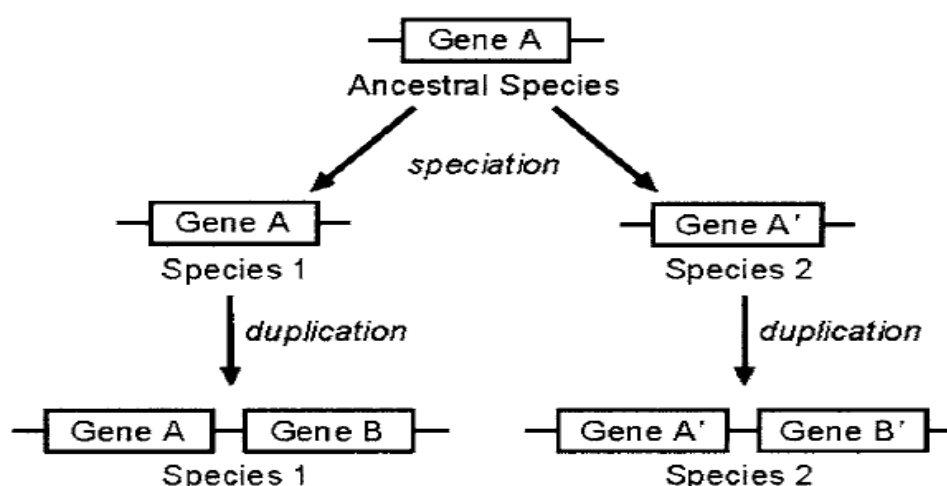


Figura 1 – Duplicação do gene. A e A' são ortólogos porque são relacionados por ascendência. Adaptada de Neale e Krutovsky (2005).

A identificação de marcadores a partir de genes conservados é grande importância em estudos de mapeamento comparativo uma vez que os mesmos permitem determinar, sem ambiguidade, o grau de sintenia. Comparando os genomas do sorgo e cana-de-açúcar, McIntyre et al. (2015) identificaram e mapearam

marcadores em dose única, relatando a existência de regiões sintênicas com uma região do cromossomo 9 do sorgo.

A estratégia de comparação de sequências genômicas foi inicialmente proposta por Fulton et al. (2002) que compararam sequências genômicas de *Arabidopsis* com ESTs de tomate. De acordo com esses autores, a identificação de genes conservados ou conjunto de marcadores ortólogos conservados (COS), como sugerido por eles, são úteis no estudo de mapeamento comparativo entre genomas diferentes, espécies de mesma família ou ainda estudos filogenéticos. No entanto, o desenvolvimento de marcadores COS, depende da disponibilidade de genes ortólogos conservados válidos (ROKAS et al., 2003; BLANCHETTE et al., 2004). Para Cabrera et al. (2009) os genes que são altamente conservados e estão presente em cópia única nos genomas, podem ser utilizados como marcadores em estudos de evolução ou análise comparativa do genoma.

É possível identificar conjuntos de genes ortólogos em espécies relacionadas, como por exemplo, as gramíneas, sorgo e arroz, que evoluíram de um mesmo ancestral comum (HACKAUF et al., 2012) a partir da análise comparativa de sequências genômicas depositadas em base de dados de ESTs (FULTON et al. 2002). As sequências de genes já conhecidas permanecem altamente conservadas durante a evolução. Assim, os marcadores pertencentes à região transcrita do genoma são transferíveis entre espécies relacionadas e podem ser usados para construir mapas genéticos comparativos (GUPTA et al., 2003).

O desenvolvimento de marcadores COS em cana-de-açúcar representa um grande desafio. Estes marcadores têm sido intensamente estudados em algumas plantas da família Poaceae, como trigo (QURAISHI et al., 2009), centeio (HACKAUF et al., 2012) e sorgo (JAYASHREE et al., 2006), mas pouco estudado em cana. No entanto, uma vez desenvolvido, representará uma poderosa ferramenta para acelerar a clonagem de QTL, bem como otimizar análises de diversidade genética em programas melhoramento genético. Sendo assim, fica claro a importância dos marcadores COS em estudos genômicos da cana-de-açúcar como uma grande ferramenta para o desenho de *primers* universais. Os marcadores COS podem ser utilizados no mapeamento comparativo para identificar genes ortólogos conservados em qualquer espécie de planta, permitindo o entendimento de processos evolutivos

do genoma de espécies que ainda não possuem tais informações a partir de análises comparativas com genomas totalmente sequenciados (WU et al., 2006), uma vez que esses marcadores apresentam uma alta transferibilidade entre espécies relacionadas de trigo (MOLNAR et al., 2015).

2.8 Importância do sorgo para o mapeamento genético da cana-de-açúcar

O sorgo é cultivado em uma grande diversidade de ambientes, incluindo áreas sujeitas a climas muito secos e ou muito quentes. É um cereal com grande expressão no cenário brasileiro e mundial, com potencial para produção de biocombustíveis.

Assim como a cana-de-açúcar, o sorgo pertence à família Poaceae (AITKEN et al., 2014b) e além disso, compartilharam do mesmo ancestral comum em pelo menos 10 milhões de anos atrás (JANNOO et al., 2007). É uma espécie diplóide ($n=10$) predominantemente autógama, com facilidade para cruzamentos e autofecundação, favorecendo tanto a obtenção de linhagens como o desenvolvimento de híbridos, além de apresentar fácil adaptação a regiões de clima seco. Mesmo com um genoma relativamente pequeno, aproximadamente 730Mb (PATERSON et al., 2009), apresenta uma grande diversidade genética e é menos complexo do que o genoma de outras culturas, como por exemplo, as espécies de cana-de-açúcar, que apresentam genoma monoplóide, com um tamanho estimado em aproximadamente 760-930Mb (D'HONT; GLASZMANN, 2001; WANG et al., 2010).

Nos últimos anos, o sorgo tem sido explorado como uma cultura chave para análise comparativa entre genomas de gramíneas. Análises genéticas têm mostrado que o sorgo e a cana apresentam uma estrutura genômica similar e por isso pode ser utilizada para explorar a base genética da cana-de-açúcar. O genoma do sorgo já se encontra totalmente sequenciado (PATERSON et al., 2009) e apresenta um tamanho menor do que o genoma da cana-de-açúcar, podendo, portanto, ser utilizado como genoma de referência para estudos genômicos da cana-de-açúcar. O uso do genoma pequeno do sorgo, como referência, tem contribuído para acelerar a análise do genoma da cana-de-açúcar (MCINTYRE et al., 2005b).

Recentemente, em um estudo genético comparativo entre espécies da família Poaceae, Aintken et al. (2014a) alinharam grupos homólogos de cana com grupos de

ligação de sorgo utilizando marcadores EST-SSRs e identificaram um alto nível de sintonia (90%) entre os dois genomas, relatando que o sorgo está intimamente relacionado com a cana-de-açúcar e pode ser considerado como uma espécie diplóide modelo (SINGH et al., 2013a) em estudos de mapeamento genético da cana-de-açúcar. Esta estratégia pode fornecer o conhecimento sobre a composição e estrutura genômica da cana-de-açúcar, que apresenta um genoma complexo entre as gramíneas. McIntyre et al. (2005) comparando sequências de genes de resistência entre cana e sorgo, relataram que o mapeamento de sequências EST-SSRs, fornece um recurso importante para futuros estudos de resistência a pragas e doenças nestas espécies e pode facilitar a clonagem de genes de resistência.

A cana (*Saccharum* L. spp.) é uma gramínea que contribui com 60% do açúcar bruto produzido no mundo. Apresenta uma importância significativa no que diz respeito à produção de biocombustíveis. Por isso, esforços consideráveis estão sendo feitos com o intuito de melhorar a produção de biomassa de cana e produção de açúcar.

Devido à sua complexidade, o genoma da cana ainda não foi completamente sequenciado. Isto se deve principalmente à dificuldade de atribuir genes contendo fragmentos de um cromossomo homólogo ou homeólogo específico. No entanto, a sequência do genoma de referência poderia ser montada a partir de fragmentos de cromossomos homólogos e homeólogos diferentes e, mesmo que essa sequência de referência seja quimérica, pode ser utilizada para a análise comparativa de genomas com parentes próximos, como o sorgo (ABROUK et al., 2010).

A complexidade do genoma autopoliploide das cultivares modernas de cana-de-açúcar tem limitado a aplicação de ferramentas genômicas e o avanço dos programas de melhoramento de cana-de-açúcar na obtenção de novas cultivares.

Entre as gramíneas que apresentam o genoma completamente sequenciado e anotado, o genoma do sorgo apresenta maior relação de parentesco com o genoma da cana-de-açúcar, podendo ser utilizado como genoma de referência para análise comparativa (DE SETTA et al., 2014).

Em cana-de-açúcar, o uso do genoma pequeno do sorgo, como referência, tem contribuído para acelerar a análise do genoma da cana (MCINTYRE et al., 2005a).

Diante do exposto, observa-se que à compreensão do genoma da cana-de-açúcar é dependente do avanço da biotecnologia, que tem contribuído para acelerar a identificação, seleção e utilização de genes responsáveis pela expressão de características de importância agronômica. A biologia molecular permite o desenvolvimento de marcadores COS-EST-SSR que podem ser utilizados para estudar a variabilidade genética entre diferentes genótipos ou ainda mapear regiões associadas a características de interesse.

3 REFERÊNCIAS

ABROUK, M.; MURAT, F.; PONT, C.; MESSING, J.; JACKSON, S.; FARAUT, T.; TANNIER, E.; PLOMION, C.; COOKE, R.; FEUILLET, C.; SALSE, J. Palaeogenomics of plants: syteny based modelling of extinct ancestors. **Trends in Plant Science**, v.15, p. 479-487, 2010.

ADAL, A. M.; DEMISSIE, Z. A.; MAHMOUD, S. S. Identification, validation and cross-species transferability of novel *Lavandula* EST-SSRs. **Planta**, v. 241, n. 4, p. 987-1004, 2015.

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. **Theor Appl Genet**, v. 110, n. 5, p. 789-801, 2005a.

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**, v. 50, n. 8, p. 742-56, 2007.

AITKEN, K. S.; MCNEIL, M. D.; BERKMAN, P. J.; HERMANN, S.; KILIAN, A.; BUNDOCK, P. C.; LI, J. C. Comparative mapping in the Poaceae family reveals translocations in the complex polyploid genome of sugarcane. **Bmc Plant Biology**, v. 14, 2014a.

AITKEN, K. S.; MCNEIL, M. D.; HERMANN, S.; BUNDOCK, P. C.; KILIAN, A.; HELLER-USZYNSKA, K.; HENRY, R. J.; LI, J. C. A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array Technology (DART) markers. **Bmc Genomics**, v. 15, 2014b.

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**, v. 167, n. 1, p. 127-142, 2009.

ANDRU, S.; PAN, Y. B.; THONGTHAWEE, S.; BURNER, D. M.; KIMBENG, C. A. Genetic analysis of the sugarcane (*Saccharum* spp.) cultivar 'LCP 85-384'. I. Linkage mapping using AFLP, SSR, and TRAP markers. **Theoretical and Applied Genetics**, 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 sugarcane rust resistance gene. **Theoretical and Applied Genetics**, v.101, p.962-969, 2000.

BAER, O. T.; LALUSIN, A. G. Molecular Markers Associated to Downy Mildew [*Peronosclerospora philippinensis* (W. Weston) CG Shaw] Resistance in Sugarcane (*Saccharum officinarum* L.) Hybrids (CP 57-604 X PHIL 84-77). **Philippine Journal of Crop Science**, v. 38, n. 3, p. 37-45, 2013.

BARNABAS, L.; RAMADASS, A.; AMALRAJ, R. S.; PALANIYANDI, M.; RASAPPA, V. Sugarcane proteomics: An update on current status, challenges, and future prospects. **Proteomics**, v. 15, n. 10, p. 1658-1670, 2015.

BLANCHETTE, M. KENT, W. J.; RIEMER, C.; ELNITSKI, L.; SMIT, A. F. A.; ROSKIN, R. B.; ROSENBLOOM, K.; CLAWSON, H.; GREEN, E. D.; HAUSSLER, D.; MILLER, W. Aligning multiple genomic sequences with the threaded blockset aligner. **Genome research**, v. 14, n. 4, p. 708-715, 2004.

BOCCACCI, P.; BELTRAMO, C.; PRANDO, M. A. S.; LEMBO, A.; SARTOR, C.; MEHLENBACHER, S. A.; BOTTA, R.; MARINONI, D. T. In silico mining, characterization and cross-species transferability of EST-SSR markers for European hazelnut (*Corylus avellana* L.). **Molecular Breeding**, v. 35, n. 1, 2015.

CABRERA, A.; KOZIK, A.; HOWAD, W.; ARUS, P.; IEZZONI, A. F.; VAN DER KNAAP, E. Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. **BMC Genomics**, v. 10, p. 562, 2009.

CAIXETA, E. T.; OLIVEIRA, A. C. B.; BRITO, G. G.; SAKIYAMA, N. S. Tipos de marcadores moleculares. In: BORÉM, A.; CAIXETA, E. T. **Marcadores moleculares**. Viçosa: UFV, 2006. p.9-78.

CATO, S. A.; GARDNER, R. C.; KENT, J.; RICHARDSON, T. E. A rapid PCR-based method for genetically mapping ESTs. **Theoretical and Applied Genetics**, v. 102, n. 2-3, p. 296-306, 2001.

CHABANE, K.; ABLETT, G. A.; CORDEIRO, G. M.; VALKOUN, J.; HENRY, R. J. EST versus Genomic Derived Microsatellite Markers for Genotyping Wild and Cultivated Barley. **Genetic Resources and Crop Evolution**, v. 52, n. 7, p. 903-909, 2005.

CHAGNE, D.; CHAUMEIL, P.; RAMBOER, A.; COLLADA, C.; GUEVARA, A.; CERVERA, M. T.; VENDRAMIN, G. G.; GARCIA, V.; FRIGERIO, J. M. M.; ECHT, C.; RICHARDSON, T.; PLOMION, C. Cross-species transferability and mapping of genomic and cDNA SSRs in pines. **Theoretical and Applied Genetics**, v. 109, n. 6, p. 1204-1214, 2004.

CHAPMAN, M. A. Transcriptome Sequencing and Marker Development for Four Underutilized Legumes. **Applications in Plant Sciences**, v. 3, n. 2, 2015.

CHEEMA, J.; DICKS, J. Computational approaches and software tools for genetic linkage map estimation in plants. **Brief Bioinform**, v. 10, n. 6, p. 595-608, 2009.

CONAB-Companhia Nacional de Abastecimento. Acompanhamento da Safra Brasileira Cana-de-açúcar: Primeiro levantamento/abril 2016 - Brasília, v.3, n.7, p. 1-158, abril. 2016. Disponível em:<
http://www.conab.gov.br/OlalaCMS/uploads/arquivos/16_04_18_14_27_15_boletim_cana_portugues_-_1o_lev_-_16.pdf.

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**, v. 160, n. 6, p. 1115-1123, 2001.

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

CRANE, C. F.; CRANE, Y. M. A nearest-neighboring-end algorithm for genetic mapping. **Bioinformatics**, v. 21, n. 8, p. 1579-91, 2005.

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

DE SETTA, N. et al. Building the sugarcane genome for biotechnology and identifying evolutionary trends. **BMC genomics**, v. 15, n. 1, p. 540, 2014.

DEWEY, C. N. Whole-genome alignment. In: ANISIMOVA, M. **Evolutionary Genomics: Statistical and Computational Methods**. New York: Humana Press, 2012. cap. 8, p. 237-258.

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 & General Genetics**, v. 250, n. 4, p. 405-413, 1996.

D'HONT, A.; GLASZMANN, J. C. Sugarcane genome analysis with molecular markers: A first decade of research. **International Society of Sugar Cane Technologists, Vol II, Proceedings**, p. 556-559, 2001.

DAI, L. J.; WANG, B.; ZHAO, H.; PENG, J. H. Transferability of Genomic Simple Sequence Repeat and Expressed Sequence Tag-Simple Sequence Repeat Markers from *Sorghum* to *Miscanthus sinensis*, a Potential Biomass Crop. **Crop Science**, v. 53, n. 3, p. 977-986, 2013.

DILLON, S. L.; SHAPTER, F. M.; HENRY, R. J.; CORDEIRO, G.; IZQUIERDO, L.; LEE, L. S. Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (*Andropogoneae*). **Ann Bot**, v. 100, n. 5, p. 975-89, 2007.

DIOLA, V.; BARBOSA, M. H. P.; VEIGA, C. F. M.; FERNANDES, E. C. Molecular Markers EST-SSRs for Genotype-Phenotype Association in Sugarcane. **Sugar Tech**, v. 16, n. 3, p. 241-249, 2014.

DUFOUR, P. et al. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. **Theoretical and Applied Genetics**, v. 94, n. 3-4, p. 409-418, 1997.

EUJAYL, I. SORRELLS, M.E.; BAUM, M.; WOLTERS, P.; POWEL, W. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. **Theoretical and Applied Genetics**, v. 104, n. 2-3, p. 399-407, 2002.

FALEIRO, F. G.; SCHUSTER, I.; RAGAGNIN, V. A.; CRUZ, C. D.; CORRÊA, R. X.; MOREIRA, M. A.; BARROS, E. G. Caracterização de linhagens endogâmicas recombinantes e mapeamento de locos de características quantitativas associados a ciclo e produtividade do feijoeiro-comum. **Pesquisa Agropecuária Brasileira**, Brasília, v.38, p.1387-1397, 2003.

FERREIRA, M. E.; GRATTAPAGLIA, D. Genética de associação em plantas. In: BORÉM, A.; CAIXETA, E. T. **Marcadores moleculares**. Viçosa: UFV, 2006. p.273-306.

FORSLUND, K.; PEKKARI, I.; SONNHAMMER, E. L. L. Domain architecture conservation in orthologs. **BMC bioinformatics**, v. 12, n. 1, p. 1, 2011.

FRASER, L. G.; MCNEILAGE, M. A.; TSANG, G. K.; HARVEY, C. F.; DE SILVA, H. N. Cross-species amplification of microsatellite loci within the dioecious, polyploid genus *Actinidia* (*Actinidiaceae*). **Theor Appl Genet**, v. 112, n. 1, p. 149-57, 2005.

FULTON, T. M.; VAN DER HOEVEN, R.; EANNETTA, N. T.; TANKSLEY, S. D. Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. **Plant Cell**, v. 14, n. 7, p. 1457-67, 2002.

GADALETA, A.; GIANCASPRO, A.; GIOVE, S. L.; ZACHEO, S.; MANGINI, G.; SIMEONE, R.; SIGNORILE, A.; BLANCO, A. Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat. **Theor Appl Genet**, v. 118, n. 5, p. 1015-25, 2009.

GALE, M. D.; DEVOS, K. M. Comparative genetics in the grasses. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 5, p. 1971-1974, 1998.

GAO, L. F.; JING, R. L.; HUO, N. X.; LI, Y.; LI, X. P.; ZHOU, R. H.; CHANG, X. P.; TANG, J. F.; MA, Z. Y.; JIA, J. Z. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. **Theor Appl Genet**, v. 108, n. 7, p. 1392-1400, 2004.

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

GRIVET, L.; D'HONT, A.; HAMON, P.; ROQUES, D.; GLASZMANN, J. C. Comparative genome mapping of sugar cane with other species within the Andropogoneae tribe. **Heredity**, v. 73, n. 5, p. 500-508, 1994.

GRIVET, L.; ARRUDA, P. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Curr Opin Plant Biol**, v.5, p.122–127, 2001.

GUIMARÃES, C. T. **Mapeamento comparative e detecção de QTLs em cana-de-açúcar utilizando marcadores moleculares.** 1999. 71. Tese (Doutorado) – Universidade Federal de Viçosa, Viçosa.

GUIMARÃES, C. T.; GAVIN, R. S.; BRUNO, W. S. S. Comparative mapping of Andropogoneae: *Saccharum* L. (sugarcane) and its relation to sorghum and maize. **Proceedings of the National Academy of Sciences.** v.94, n.26, p. 14261-14266, 1997.

GUPTA, P. K.; RUSTGI, S.; SHARMA, S.; SINGH, R.; KUMAR, N.; BALYAN, H. S. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. **Mol Genet Genomics**, v. 270, n. 4, p. 315-23, 2003.

HACKAUF, B.; KORZUN, V.; WORTMANN, H.; WILDE, P.; WEHLING, P. Development of conserved ortholog set markers linked to the restorer gene Rfp1 in rye. **Molecular Breeding**, v. 30, n. 3, p. 1507-1518, 2012.

HERNÁNDEZ, P.; DORADO, G.; LAURIE, D.A.; MARTIN, A.; SNAPE, J.W. Microsatellites and RFLP probes from maize are efficient sources of molecular markers for the biomass energy crop *Miscanthus*. **Theoretical and Applied Genetics**, v. 102, n. 4, p. 616-622, 2001.

HOARAU, J. Y.; OFFMANN, 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**, v. 103, n. 1, p. 84-97, 2001.

JAMES, B. T.; CHEN, C.; RUDOLPH, A.; SWAMINATHAN, K.; MURRAY, J. E.; NA, J.-K.; SPENCE, A. K.; SMITH, B.; HUDSON, M. E.; MOOSE, S. P.; MING, R. Development of microsatellite markers in autopolyploid sugarcane and comparative analysis of conserved microsatellites in sorghum and sugarcane. **Molecular Breeding**, v. 30, n. 2, p. 661-669, 2011.

JANNOO, N.; GRIVET, L.; CHANTRET, N.; GARSMEUR, O.; GLASZMANN, J. C.; ARRUDA, P.; D'HONT, A. Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome. **Plant J**, v. 50, n. 4, p. 574-85, 2007.

JAYASHREEA, B.; PUNNAB, R.; PRASADA, P.; BANTTEB, K.; C. HASHB, C. T.; CHANDRAA, S.; HOISINGTONB, D. A.; VARSHNEYB, R. K. A database of simple sequence repeats from cereal and legume expressed sequence tags mined *in silico*: survey and evaluation. **In Silico Biology**, v.6, p.607-620, 2006.

JEONG, Y. M.; CHUNG, W. H.; CHUNG, H.; KIM, N.; PARK, B. S.; LIM, K. B.; Yu. H.; MUN, J. H. Comparative analysis of the radish genome based on a conserved ortholog set (COS) of Brassica. **TAG Theoretical and applied genetics Theoretische und angewandte Genetik**, v.127, p.1975-1989, 2014.

JORDAN, D. R.; CASU, R. E.; BESSE, P.; CARROLL, B. C.; BERDING, N.; MCINTYRE, C. L. Markers associated with stalk number and suckering in sugarcane colocate with tillering and rhizomatousness QTLs in sorghum. **Genome**, v. 47, n. 5, p. 988-993, 2004.

KANE, N. C.; RIESEBERG, L. H. Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus*. **Genetics**, v. 175, n. 4, p. 1823-34, 2007.

KIM, C.; ZHANG, D.; AUCLAND, S. A.; RAINVILLE, L. K.; JAKOB, K.; KRONMILLER, B.; SACKS, E. J.; DEUTER, M.; PATERSON, A. H. SSR-based genetic maps of *Miscanthus sinensis* and *M. sacchariflorus*, and their comparison to sorghum. **Theor Appl Genet**, v. 124, n. 7, p. 1325-38, 2012.

KORZUN, V.; MALYSHEV, S.; VOYLOKOV, A. V.; BORNER, A. A genetic map of rye (*Secale cereale* L.) combining RFLP, isozyme, protein, microsatellite and gene loci. **theoretical and Applied Genetics**, v. 102, n. 5, p. 709-717, 2001.

KUHN, D. N.; LIVINGSTONE, D.; MAIN, D.; ZHENG, P.; SASKI, C.; FELTUS, F. A.; MOCKAITIS, K.; FARMER, A. D.; MAY, G. D.; SCHNELL, R. J.; MOTAMAYOR, J. C. Identification and mapping of conserved ortholog set (COS) II sequences of cacao and their conversion to SNP markers for marker-assisted selection in *Theobroma cacao* and comparative genomics studies. **Tree Genetics & Genomes**, v. 8, n. 1, p. 97-111, 2012.

LE CUNFF, L.; GARSMEUR, O.; RABOIN, L. M.; PAUQUET, J.; TELISMART, H.; SELVI, A.; GRIVET, L.; PHILIPPE, R.; BEGUM, D.; DEU, M.; COSTET, L.; WING, R.; GLASZMANN, J. C.; D'HONT, A. Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance gene (Bru1) in highly polyploid sugarcane (2n approximately 12x approximately 115). **Genetics**, v. 180, n. 1, p. 649-60, 2008.

LIU, H.; GUO, X.; WU, J.; CHEN, G. B.; YING, Y. Development of universal genetic markers based on single-copy orthologous (COSII) genes in Poaceae. **Plant Cell Rep**, v. 32, n. 3, p. 379-88, 2013.

LIU, P.; CHANDRA, A.; QUE, Y.; CHEN, P.-H.; GRISHAM, M. P.; WHITE, W. H.; DALLEY, C. D.; TEW, T. L.; PAN, Y.-B. Identification of quantitative trait loci controlling sucrose content based on an enriched genetic linkage map of sugarcane (*Saccharum* spp. hybrids) cultivar 'LCP 85-384'. **Euphytica**, v. 207, n. 3, p. 527-549, 2016.

LUO, L.; XU, S. Mapping viability loci using molecular markers. **Heredity (Edinb)**, v. 90, n. 6, p. 459-67, 2003.

MCINTYRE, C. L.; CASU, R. E.; DRENTH, J.; KNIGHT, D.; WHAN, V. A.; CROFT, B. J.; JORDAN, D. R.; MANNERS, J. M. Resistance gene analogues in sugarcane and sorghum and their association with quantitative trait loci for rust resistance. **Genome**, v. 48, n. 3, p. 391-400, 2005a.

MCINTYRE, C. L.; GOODE, M. L.; CORDEIRO, G.; BUNDOCK, P.; ELIOTT, F.; HENRY, R. J.; CASU, R. E.; BONNETT, G. D.; AITKEN, K. S. Characterization of alleles of the sucrose phosphate synthase gene family in sugarcane and their association with sugar-related traits. **Molecular Breeding**, v. 35, n. 3, 2015.

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**, v. 16, n. 2, p. 151-161, 2005b.

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

MING, R.; LIU, S. C.; BOWERS, J. E.; MOORE, P. H.; IRVINE, J. E.; PATERSON, A. H. Construction of a *Saccharum* consensus genetic map from two interspecific crosses. **Crop Science**, v. 42, n. 2, p. 570-583, 2002b.

MING; S.-C. LIU; Y.-R. LIN; J. DA SILVA; WENSLAFF, T. F.; WU, K. K.; WILSON, W.; BRAGA, D.; IRVINE, J. E.; A. H. PATERSON; P. H. MOORE; W. BURNQUIST; A. VAN DEYNZE; M. E. SORRELLS. Detailed Alignment of *Saccharum* and *Sorghum* Chromosomes: Comparative Organization of Closely Related Diploid and Polyploid Genomes. 1998.

MING, R.; WANG, W.; DRAYE, X.; MOORE, H.; IRVINE, E.; PATERSON, H. Molecular dissection of complex traits in autopolyploids: mapping QTLs affecting sugar yield and related traits in sugarcane. **Theor Appl Genet**, v. 105, n. 2-3, p. 332-345, 2002c.

MOLNAR, I.; VRANA, J.; FARKAS, A.; KUBALAKOVA, M.; CSEH, A.; MOLNAR-LANG, M.; DOLEZEL, J. Flow sorting of C-genome chromosomes from wild relatives of wheat *Aegilops markgrafii*, *Ae. triuncialis* and *Ae. cylindrica*, and their molecular organization. **Ann Bot**, v. 116, n. 2, p. 189-200, 2015.

MOYLE, R. L.; BIRCH, R. G. Diversity of sequences and expression patterns among alleles of a sugarcane loading stem gene. **Theor Appl Genet**, v. 126, n. 7, p. 1775-82, 2013.

NEALE, D. B.; KRUTOVSKY, K. V. Comparative Genetic Mapping in Trees: The Group of Conifers. In: NAGATA, T.; LÖRZ, H.; WIDHOLM, J. M. **Biotechnology in Agriculture and Forestry 55: Molecular Marker Systems in Plant Breeding and Crop Improvement**. Germany: Springer, 2005. p. 267-277.

NOMELINI, Q. S. S.; SILVA, H. D.; FARIA, T. C. EFICIÊNCIA DOS MÉTODOS DE OTIMIZAÇÃO SIMULATED ANNEALING, DELINEAÇÃO RÁPIDA EM CADEIA E RAMOS E CONEXÕES PARA CONSTRUÇÃO DE MAPAS GENÉTICOS. **Ciênc. agrotec.,(Impr.)**, v. 33, n. 6, p. 1534-1537, 2009.

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**, v. 20, n. 3, p. 189-208, 2007.

PALHARES, A. C.; RODRIGUES-MORAIS, T. B.; VAN SLUYS, M. A.; DOMINGUES, D. S.; MACCHERONI, W.; JORDAO, H.; SOUZA, A. P.; MARCONI, T. G.; MOLLINARI, M.; GAZAFFI, R.; GARCIA, A. A. F.; VIEIRA, M. L. C. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. **Bmc Genetics**, v. 13, 2012.

PAN, Yong-Bao. Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. **Sugar Tech**, v. 8, n. 4, p. 246-256, 2006.

PANDIAN, Anand; FORD, Rebecca; TAYLOR, Paul WJ. Transferability of sequence tagged microsatellite site (STMS) primers across four major pulses. **Plant Molecular Biology Reporter**, v. 18, n. 4, p. 395-395, 2000.

PARMESSUR, Y.; RAMU, P.; HASH, T.; DOOKUN-SAUTALLY, A. Application of sorghum est-ssr to the linkage mapping of Sugarcane variety m 134/175. In: 10th GERMPLASM & BREEDING: 7th MOLECULAR BIOLOGY WORKSHOP. 2011, Maceió. **Abstract...** Maceió: STAB- EAST SECTION, 2011, p.56, ref. MO9.

PASHLEY, C. H.; ELLIS, J. R.; MCCAULEY, D. E.; BURKE, J. M. EST databases as a source for molecular markers: Lessons from Helianthus. **Journal of Heredity**, v. 97, n. 4, p. 381-388, 2006.

PATERSON, A. H.; BOWERS, J. E.; BRUGGMANN, R.; DUBCHAK, I.; GRIMWOOD, J.; GUNDLACH, H.; HABERER, G.; HELLSTEN, U.; MITROS, T.; POLIAKOV, A.; SCHMUTZ, J.; SPANNAGL, M.; TANG, H.; WANG, X.; WICKER, T.; BHARTI, A. K.; CHAPMAN, J.; FELTUS, F. A.; GOWIK, U.; GRIGORIEV, I. V.; LYONS, E.; MAHER, C. A.; MARTIS, M.; NARECHANIA, A.; OTILLAR, R. P.; PENNING, B. W.; SALAMOV, A. A.; WANG, Y.; ZHANG, L.; CARPITA, N. C.; FREELING, M.; GINGLE, A. R.; HASH, C. T.; KELLER, B.; KLEIN, P.; KRESOVICH, S.; MCCANN, M. C.; MING, R.; PETERSON, D. G.; MEHBOOB UR, R.; WARE, D.; WESTHOFF, P.; MAYER, K. F.; MESSING, J.; ROKHSAR, D. S. The Sorghum bicolor genome and the diversification of grasses. **Nature**, v. 457, n. 7229, p. 551-6, 2009.

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

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

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

QURASHI, U. M.; ABROUK, M.; BOLOT, S.; PONT, C.; THROUDE, M.; GUILHOT, N.; CONFOLENT, C.; BORTOLINI, F.; PRAUD, S.; MURIGNEUX, A.; CHARMET, G.; SALSE, J. Genomics in cereals: from genome-wide conserved orthologous set (COS) sequences to candidate genes for trait dissection. **Functional e Integrative Genomics**, v.9, p.473-484, 2009.

ROKAS, A.; KING, N.; FINNERTY, J.; CARROLL, S. B. Conflicting phylogenetic signals at the base of the metazoan tree. **Evolution & development**, v. 5, n. 4, p. 346-359, 2003.

SAHA, M. C.; MIAN, M. A. R.; EUJAYL, I. ZWONITZER, J. C.; WANG, L.; MAY, G. D. Tall fescue EST-SSR markers with transferability across several grass species. **Theoretical and Applied Genetics**, v. 109, n. 4, p. 783-791, 2004.

SALAS-LEIVA, D. E.; MEEROW, A. W.; FRANCISCO-ORTEGA, J.; CALONJE, M.; GRIFFITH, M. P.; STEVENSON, D. W.; NAKAMURA, K. Conserved genetic regions across angiosperms as tools to develop single-copy nuclear markers in gymnosperms: an example using cycads. **Mol Ecol Resour**, v. 14, n. 4, p. 831-45, 2014.

SINGH, R. K.; JENA, S. N.; KHAN, S.; YADAV, S.; BANARJEE, N.; RAGHUVANSHI, S.; BHARDWAJ, V.; DATTAMAJUMDER, S. K.; KAPUR, R.; SOLOMON, S.; SWAPNA, M.; SRIVASTAVA, S.; TYAGI, A. K. Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane. **Gene**, v. 524, n. 2, p. 309-329, 2013a.

SINGH, R. K.; SINGH, R. B.; SINGH, S. P.; SHARMA, M. L. Identification of sugarcane microsatellites associated to sugar content in sugarcane and transferability to other cereal genomes. **Euphytica**, v. 182, n. 3, p. 335-354, 2011.

SINGH, R. K.; SINGH, S. P.; TIWARI, D. K.; SRIVASTAVA, S.; SINGH, S. B.; SHARMA, M. L.; SINGH, R.; MOHAPATRA, T.; SINGH, N. K. Genetic mapping and QTL analysis for sugar yield-related traits in sugarcane. **Euphytica**, v. 191, n. 3, p. 333-353, 2013b.

SOBHAKUMARI, V. P. New determinations of somatic chromosome number in cultivated and wild species of *Saccharum*. **Caryologia**, v. 66, n. 3, p. 268-274, 2013.

STANIASZEK, M.; SZCZETCHURA, W.; MARCZEWSKI, W. Identification of a New Molecular Marker C2-25 Linked to the *Fusarium oxysporum* f.sp *radicis- lycopersici* Resistance Frl Gene in Tomato. **Czech Journal of Genetics and Plant Breeding**, v. 50, n. 4, p. 285-287, 2014.

TERRACCIANO, I.; MACCAFERRI, M.; BASSI, F.; MANTOVANI, P.; SANGUINETI, M. C.; SALVI, S.; SIMKOVA, H.; DOLEZEL, J.; MASSI, A.; AMMAR, K.; KOLMER, J.; TUBEROSA, R. Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of Lr14a in durum wheat (*Triticum durum* Desf.). **Theoretical and Applied Genetics**, v. 126, n. 4, p. 1077-1101, 2013.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v. 23, n. 1, p. 48-55, 2005.

VASEMAGI, A. Expressed Sequence Tag-Linked Microsatellites as a Source of Gene-Associated Polymorphisms for Detecting Signatures of Divergent Selection in Atlantic Salmon (*Salmo salar* L.). **Molecular Biology and Evolution**, v. 22, n. 4, p. 1067-1076, 2005.

VIEIRA, E. A.; NODARI, R.O.; CARVALHO, F. I. F.; FIALHO, J. F. **Mapeamento genético de caracteres quantitativos e sua interação com o ambiente**. Planaltina-DF: Embrapa Cerrados, 2006. 26.p. (Embrapa Cerrados. Documentos, 170). Disponível em: < <http://www.infoteca.cnptia.embrapa.br/handle/doc/570277>>. Acesso em: 20 mai. 2015.

VILLANO, C.; MIRAGLIA, V.; IORIZZO, M.; AVERSANO, R.; CARPUTO, D. Combined Use of Molecular Markers and High-Resolution Melting (HRM) to Assess Chromosome Dosage in Potato Hybrids. **Journal of Heredity**, v. 107, n. 2, p. 187-192, 2016.

WANG, J.; ROE, B.; MACMIL, S.; YU, Q.; MURRAY, J. E.; TANG, H.; CHEN, C.; NAJAR, F.; WILEY, G.; BOWERS, J.; VAN SLUYS, M. A.; ROKHSAR, D. S.;

HUDSON, M. E.; MOOSE, S. P.; PATERSON, A. H.; MING, R. Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. **BMC Genomics**, v. 11, p. 261, 2010.

WANG, M. L.; WANG, M. L.; BARKLEY, N. A.; YU, J. K.; DEAN, R. E.; NEWMAN, M. L.; SORRELLS, M. E.; PEDERSON, G. A. Transfer of simple sequence repeat (SSR) markers from major cereal crops to minor grass species for germplasm characterization and evaluation. **Plant Genetic Resources: characterization and utilization**, v. 3, n. 1, p. 45-57, 2005.

WU, R.; MA, C. X.; PAINTER, I.; ZENG, Z. B. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. **Theor Popul Biol**, v. 61, n. 3, p. 349-63, 2002.

WU, F.; MUELLER, L. A.; CROUZILLAT, D.; PÉTIARD, V.; TANKSLEY, S. D. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade. **Genetics**, v. 174, n. 3, p. 1407-1420, 2006.

YADAV, S.; GAUR, V. S.; JAISWAL, J. P.; KUMAR, A. Simple sequence repeat (SSR) analysis in relation to calcium transport and signaling genes reveals transferability among grasses and a conserved behavior within finger millet genotypes. **Plant Systematics and Evolution**, v. 300, n. 6, p. 1561-1568, 2014.

YU, J. K.; LA ROTA, M.; KANTETY, R. V.; SORRELLS, M. E. EST derived SSR markers for comparative mapping in wheat and rice. **Mol Genet Genomics**, v. 271, n. 6, p. 742-51, Jul,2004.

ZHANG, L. Y.; BERNARD, M.; LEROY, P.; FEUILLET, C.; SOURDILLE, P. High transferability of bread wheat EST-derived SSRs to other cereals. **Theor Appl Genet**, v. 111, n. 4, p. 677-87, 2005.

CAPÍTULO II

DESENVOLVIMENTO E CARACTERIZAÇÃO DE MARCADORES COS-EST-SSRS ENTRE CANA E SORGO

Desenvolvimento e caracterização de marcadores COS-EST-SSRs entre cana-de-açúcar e sorgo

RESUMO - Nos últimos anos, os bancos de dados de ESTs (Etiquetas de Sequências Expressas) tem sido amplamente explorados na busca de sequências que possam ser empregadas no desenvolvimento de marcadores moleculares. Através da análise comparativa entre bancos de dados de ESTs de diferentes espécies é possível o desenvolvimento de marcadores derivados de sequências ortólogas conservadas (*COS-Conserved Ortholog Sequence*). O objetivo do presente estudo foi identificar sequências ortólogas conservadas entre os genomas de cana e sorgo para desenvolver um grupo de pares de *primers* flanqueadores de microssatélites (COS-EST-SSRs) e testar sua eficiência em genótipos de ambas espécies. A partir da utilização de 51.247 ESTs ortólogas conservadas entre cana e sorgo foram obtidos 320 *clusters* de cana contendo microssatélites (COS-EST-SSR) dos quais, 58% foram constituídos por motivos trinucleotídeos e o restante de motivos de dinucleotídeos. Sessenta e oito por cento das sequências que continham regiões ortólogas conservadas mostrou similaridade com proteínas conhecidas em espécies relacionadas. Dos 320 clusters de cana foram desenhados 41 pares de *primers* COS-EST-SSRs que foram avaliados quanto ao polimorfismo em 24 genótipos de *Saccharum* e *Sorghum*. Foi verificada uma maior porcentagem de polimorfismo entre os genótipos de cana (79%) do que nos híbridos de sorgo (18%). De acordo com o valor médio do PIC (0,71), os COS-EST-SSRs foram capazes de capturar a variabilidade genética entre os genótipos estudados o que reforça a utilidade destes marcadores em detectar diferenças entre e dentro de indivíduos das duas espécies avaliadas. Os marcadores COS-EST-SSRs foram eficientes no estudo de divergência genética e podem ser utilizados em estudos de filogenia, na avaliação da variabilidade genética e no mapeamento comparativo.

Palavras chave: ortologia, sequências conservadas, conteúdo de informação de polimorfismo, marcadores microssatélites, Poaceae

Development and characterization of COS-EST-SSRs markers between sugarcane and sorghum

ABSTRACT - In recent years, ESTs (Expressed Sequence Tags) data banks has been widely explored for molecular markers development. The COS markers (*Conserved Ortholog Sequence*) can be developed through comparative analysis of ESTs databases. The objective of this study was to identify orthologous sequences between sugarcane and sorghum to develop a set of COS-EST-SSRs markers and test their efficiency among genotypes of both species. From 51,247 ESTs, 320 orthologous sugarcane clusters conserved between sugarcane and sorghum were obtained containing microsatellites (COS-EST-SSR). Of these, 58% consisted of trinucleotide motifs and the remaining ones of dinucleotides. Sixty eight percent of the COS showed similarity to known proteins in related species. Of the 320 COS-EST-SSRs, 41 COS-EST-SSRs *primer* pairs were designed and evaluated for polymorphism in 24 genotypes of *Saccharum* and *Sorghum*. The percentage of polymorphism between sugarcane genotypes (79%) was higher than that observed in sorghum hybrids (18%). According to the PIC average value (0.71) the COS-EST-SSRs were able to capture the genetic variability among genotypes reinforcing their usefulness in detecting differences among and within these two species. COS-EST-SSRs markers were efficient to access genetic divergence and can be used in phylogenetic and comparative mapping studies.

Keywords: orthology, conserved sequences, polymorphism information content, microsatellite markers, Poaceae

1 INTRODUÇÃO

Vários tipos de marcadores baseado em DNA foram desenvolvidos na última década. Os marcadores moleculares têm sido aplicados no melhoramento de plantas e estudos genéticos, tais como mapeamento de genes e análises de *loci* com caracteres quantitativos (*Quantitative Trait Loci*-QTL) (MA et al., 2012; ZHANG et al., 2015), análise genética e estrutura de populações (SINGH et al., 2013), mapeamento comparativo (AITKEN et al., 2014) e seleção assistida por marcadores (SAM) (ZHANG et al., 2014).

Devido à sua herança codominante, abundância em genomas e alta reprodutibilidade, os marcadores microssatélites têm sido desenvolvidos e utilizados em estudos genéticos moleculares para muitas espécies (WANG et al., 2013). Os marcadores do tipo microssatélites são sequências repetitivas compostas por 1 a 6 nucleotídeos (motivo) repetidos em tandem.

Recentemente, devido à disponibilidade de uma grande quantidade de dados de domínio público para as etiquetas de sequências expressas (ESTs), a ênfase passou de SSR (*Single Sequence Repeat*) genômicos para EST-SSR, que pertencem a regiões transcritas do genoma e podem ter um papel fundamental na expressão ou função do gene (KAYESH et al., 2014).

A análise comparativa de sequências de ESTs entre genomas de plantas depositadas em bancos de dados pode permitir o desenvolvimento de marcadores a partir de sequências ortólogas conservadas (COS- *Conserved Ortholog Sequences*) (FULTON et al., 2002). Os COS são definidos como uma coleção de genes de cópia única que são conservados ou ortólogos entre espécies (JEONG et al., 2014). Genes de cópias únicas possibilitam o desenvolvimento de marcadores universais (âncoras). Isso implica dizer que uma espécie que não apresente seu genoma sequenciado pode utilizar marcadores desenvolvidos a partir de sequências ortólogas conservadas oriundas de espécies relacionadas (LIU et al., 2013). Os genes ortólogos são originados de um único gene do último ancestral comum entre as espécies (REMM et al., 2013). Os marcadores COS têm sido utilizados em diversos estudos envolvendo famílias de plantas como Pinaceae (LIEWLAKSANEEYANAWIN et al., 2009), Malvaceae (KUHN et al., 2012), Rosaceae (BUSHAKRA et al., 2012), Poaceae (LIU

et al., 2013), Solanaceae (LINDQVIST-KREUZE et al., 2013) e Brassicaceae (JEONG et al., 2014). Segundo Fulton et al. (2002) marcadores COS podem ser úteis em estudos de mapeamento comparativo entre plantas de mesma família.

A cana pertence à família Poaceae e entre as culturas de importância econômica dessa família, o sorgo surge como planta modelo, uma vez que seu genoma encontra-se totalmente sequenciado (PATERSON et al., 2009) e também por apresentar um genoma relativamente pequeno (~730 Mb) e com menor nível de complexidade do que o milho e a cana-de-açúcar, por exemplo (RAMU et al., 2009). Jannoo et al. (2007) demonstraram que sequências de sorgo e cana-de-açúcar apresentaram 88,9% de identidade com espécies de *Saccharum officinarum* e *S. spontaneum*, revelando alta conservação da estrutura do gene e similaridade de nucleotídeos. Estima-se que, a divergência entre cana e sorgo data de 7,7 milhões anos (KIM et al., 2009; WANG et al., 2010), mas apesar desse longo período, ambas as espécies tiveram seus genes conservados no processo de evolução. Sendo conservados na evolução, os genes ortólogos podem ser utilizados no mapeamento comparativo entre genomas altamente divergentes, estudo de sintonia e construção de mapas genéticos e análises de diversidade genética entre espécies (LIU et al., 2013).

A cana-de-açúcar (*Saccharum* spp) é uma gramínea tropical de grande importância econômica, responsável por 70% da produção de açúcar no mundo e nos últimos anos, tem recebido uma maior atenção por causa do potencial de seus derivados, etanol e bagaço, como fontes importantes de biocombustíveis renováveis (PIPERIDIS et al., 2010). Por ser uma cultura de destaque na economia do país, é preciso desenvolver ferramentas que auxiliem os programas de melhoramento na obtenção de cultivares mais produtivas e tolerantes a fatores bióticos e abióticos. Até o presente momento, poucos estudos foram realizados para cana-de-açúcar envolvendo marcadores moleculares desenvolvidos a partir de sequências ortólogas conservadas. Sendo assim, torna-se necessário desenvolver e validar marcadores baseados em sequências ortólogas para o genoma da cana que poderão ser úteis na identificação de QTL, análises de diversidade genética e no mapeamento comparativo envolvendo outras espécies da família Poaceae. Desta forma, o objetivo do presente estudo foi identificar sequências ortólogas conservadas entre os genomas de cana-

de-açúcar e sorgo com a finalidade de desenvolver um conjunto de pares de *primers* capazes de gerar marcadores COS-EST-SSRs e testar sua eficiência em genótipos de ambas as espécies.

2 MATERIAL E MÉTODOS

2.1 Material vegetal

Para a avaliação do polimorfismo dos pares de *primers* flanqueadores de microssatélites derivados de sequências de etiquetas expressas ortólogas conservadas (COS-EST-SSRs) entre cana e sorgo foi utilizado um total de 24 genótipos: seis pertencentes a espécie *S. officinarum*, cinco da espécie *S. spontaneum*, seis cultivares comerciais, além de sete híbridos comerciais de sorgo (Tabela 1).

Tabela 1 - Genótipos utilizados na avaliação dos pares de *primers* flanqueadores de microssatélites derivados de sequências de etiquetas expressas ortólogas conservadas (COS-EST-SSRs) entre cana e sorgo.

<i>S. officinarum</i>	<i>S. spontaneum</i>	Cultivares comerciais	Híbridos de sorgo ^a
Ajax	Dacca	IACSP955000	1G100
Badila	Djantoerll	RB867515	1G220
Badila de Java	Formosa	SP801842	1G233
Bourbon Sur	Gehrabon	SP803250	1G244
Creoula	Holes 1	RB835054	1G282
Sabura	-	RB855453	SS318
-	-	-	50A40

^a Híbridos de sorgo cedidos pela empresa de sementes Dow Agrosciences.

2.2 Obtenção de sequências de etiquetas expressas ortólogas conservadas entre cana e sorgo (COS-EST) contendo microssatélites

A base de dados de Genômica Funcional de Cana-de-açúcar do projeto SUCEST conta com, aproximadamente, 238.000 ESTs organizados em 43.141 clusters. Vicentini et al. (2012) juntaram estes clusters a um conjunto de 8.106 sequências de ESTs, não geradas pelo projeto SUCEST, da base de dados Sugarcane Gene Index – SoGI, o que resultou em um conjunto de dados contendo 51.247 clusters, que passou a integrar as Sequências Montadas de Cana – SASs (Figura 1), denominadas pelos autores de agrupamentos consolidados de cana-de-açúcar ou CSCs (*Consolidated Sugarcane Clusters*).

As etiquetas de sequências expressas ortólogas conservadas (COS-EST) entre cana e sorgo foram obtidas com base nessas 51.247 sequências não redundantes identificadas por Vicentini et al. (2012) a partir da integração entre as sequências dos bancos de dados do projeto SUCEST (http://sucest-fun.org/cgi-bin/cane_regnet/sucamet/search_transcript.cgi) e da base de dados - SoGI. A sequência completa do genoma do sorgo foi obtida do Phytozome (<http://www.phytozome.net/>). Essas 51.247 sequências foram submetidas a um processo de alinhamento utilizando a ferramenta Blastx ($<1e^{-5}$) contra o proteoma do sorgo (DOE JGI's *Sorghum bicolor* version 1.4 (<http://genome.jgi-psf.org/Sorbi1/Sorbi1.home.html>)). Em seguida, foi realizada uma análise de filogenia que permitiu identificar 19.315 CSCs ortólogos (Vicentini et al. 2012), os quais foram utilizados no presente trabalho para a busca de microssatélites (SSRs) (Figura 1).

Após a identificação das sequências ortólogas conservadas entre cana-de-açúcar e sorgo (COS-EST) o programa SSRFINDER (<http://www.fresnostate.edu/ssrfinder/>) foi utilizado para localizar microssatélites. Buscou-se por microssatélites do tipo: di, tri e tetranucleotídeos, repetidos em tandem, com repetições ≥ 5 .

2.3 Desenho dos pares de *primers* (COS-EST-SSRs)

Os pares de sequências ortólogas foram alinhados pelo programa Clustal W (<http://www.genome.jp/tools/clustalw/>) e aqueles com o melhor E-value ($<1e^{-5}$) foram

selecionados para o desenho de *primers*. Este procedimento permitiu identificar os pares de sequências ortólogas com sobreposição adequada das respectivas sequências entre cana e sorgo (sequência de cana posicionada totalmente depois do início e antes do final da sequência de sorgo) para que fosse possível o desenho dos pares de *primers* COS-EST-SSR.

Os oligos de síntese específica (*primers*) flanqueadores das regiões repetitivas foram desenhados utilizando o programa Primer 3 (<http://primer3.sourceforge.net/>). Os parâmetros utilizados para desenvolvimento dos *primers* foram: sequências com o tamanho do produto da amplificação da PCR variando de 150 a 300 pares de bases; *primers* com tamanho de 18 a 22 bases; conteúdo de guanina e citosina variando de 40 a 60% de bases e temperatura de anelamento variando de 55 a 62 °C. Foram desenhados 41 pares de *primers* COS-EST-SSRs cujas sequências estão apresentadas na Tabela 2.

2.4 Anotação funcional dos COS-EST-SSR

As sequências referentes aos COS-EST-SSR foram analisadas na base de dados de proteínas e anotação UniProt (Universal Protein Resource) (<http://www.uniprot.org/>). Esta ferramenta permitiu verificar a existência de similaridade entre as sequências COS-EST-SSR com as proteínas de diversos organismos, quando considerado E-value $\leq 1e^{-5}$.

2.5 Extração do DNA genômico

O DNA genômico total foi obtido de 300 mg de tecido foliar jovem liofilizado e moído utilizando o procedimento de extração com base no método de Aljanabi et al. (1999). O DNA de cada indivíduo foi quantificado na presença de um padrão de DNA do fago lambda de quantidades conhecidas em gel de agarose 0,8% (p/v) corado com brometo de etídio.

2.6 Amplificação dos COS-EST-SSR

Para definição da melhor temperatura de anelamento, foram utilizados três genótipos: uma linhagem de sorgo (ATX 3381) e dois genitores (IACSP953018 e IACSP933046) de uma população de mapeamento genético de cana-de-açúcar do Programa de Melhoramento Genético de Cana-de-açúcar do Instituto Agronômico.

As reações de PCR foram efetuadas em um volume final de reação de 18 µL, contendo 40 ng do DNA molde, 0.2µM de cada par de primer (forward e reverse), 100 µM de cada dNTP, 2.0 mM MgCl₂, 10 mM Tris-HCl, 50 mM KCl, e 0.5 unidades de Taq DNA polimerase.

A temperatura de anelamento de cada par de primer foi identificada usando gradiente de temperatura variando de 50 a 61 °C em um termociclador gradiente (My-Cicler, Bio-rad).

Após definida a melhor temperatura de anelamento para cada par de primer COS-EST-SSR, as seguintes condições de amplificação foram adotadas: desnaturação inicial à 94 °C por 3 minutos, seguida de 35 ciclos, cada ciclo contendo uma etapa de desnaturação à 94 °C por 1 minuto, temperatura de anelamento específica para cada par de primer (forward/reverse) por 1 minuto, extensão à 72 °C por 1 minuto; e um ciclo final à 72 °C por 2 minutos utilizando o termociclador Veriti (Applied Biosystems).

Aos produtos da amplificação da reação de PCR foram adicionados 7 µL de tampão formamida (98% formamida; 10 mM EDTA pH 8,0; 0.025% azul de bromofenol; 0.025% xileno cianol) seguido de desnaturação à 95 °C por 5 minutos.

2.7 Eletroforese em gel de poliacrilamida desnaturante

Os fragmentos foram separados por eletroforese em gel de poliacrilamida desnaturante a 6% w/v, utilizando ladder de 10 bp como marcador de peso molecular, a uma potência constante de 55 Watts. Os produtos da amplificação foram corados com nitrato de prata de acordo com o protocolo estabelecido por Creste et al. (2001) para visualização das marcas (alelos).

Tabela 2 - Identificação das sequências ortólogas conservadas na base de dados SUCEST, iniciadores COS-EST-SSRs com suas respectivas sequências (forward/reverse), temperatura de anelamento (T_m), tamanho do produto esperado (TP_{es}) e similaridade esperada.

ID SUCEST	Primer	Motivo	Forward (5'→3')	Reverse (3'→5')	T _m (°C)	TP _{es} (pb)	Similaridade dos COS-EST-SSR
SCUTLR2030G02.g	COS-EST-SSR 1	(AAG) ₅	AGATCTCCAACGCCGACTC	TGATGAATCCGATGAAGACG	59,60	237	Endoglucanase 10
SCVPCL6064B03.g	COS-EST-SSR 2	(GC) ₅	CATGATCCGCTGGGACAA	GGAACCTGGACACGAGGAT	59,91	179	Trehalose-phosphatase family protein, expressed
SCAGLR2033D02.g	COS-EST-SSR 3	(GC) ₅	GATGTAATCGCGGAGCATCT	CTGCTCAGCAACAGGAAGGT	60,59	225	TPR Domain containing protein, expressed
SCEZLB1006H04.g	COS-EST-SSR 4	(GC) ₅	ATCTGCAGCAGGTGCGAGT	AGGACTGCTTCGTGCTCAA	59,72	154	Lipase, putative, expressed
SCAGLR1064B05.g	COS-EST-SSR 5	(CG) ₅	TACCTCCAGCACGTTTCATCA	AGGCTGACGTTTCGACATGA	60,42	174	Cytochrome P450 family protein, expressed
SCEZLB1012C07.g	COS-EST-SSR 6	(CG) ₅	AGGAGCAGGTTCTCGAGCTT	ACCCGCTCCAAGATCTACCT	60,10	188	GRMZM2G412601; putative CBL-interacting protein kinase family protein
SCCCLR1078G04.g	COS-EST-SSR 7	(CAG) ₆	CCCTGAGCCTTATTGGAACA	GTTGCACTTGATCTGAAGG	59,45	156	Os07g0571100 protein
SCEZLR1009C07.g	COS-EST-SSR 8	(CAG) ₅	CACAGCAATTTTCAGCAGCTC	CCAAGGACCACATTCCTGTT	59,82	151	Os01g0607600 protein
SCSFST1064G02.g	COS-EST-SSR 9	(TGC) ₅	CACCTTCCCCTGACTATCCA	CTGGCTTCAGCAAATTCAT	60,21	202	Auxilin-like protein
SCAGLB1070D03.g	COS-EST-SSR 10	(CGC) ₅	TGTCATGGTTGAGTTCTACGC	AGAAGAGGATGGTGGGGAAG	60,45	191	Protein disulfide isomerase
SCACLR2007D01.g	COS-EST-SSR 11	(GGT) ₅	GCGAAAGCACGAGATCCAC	GAAGATCTCCGGGGAGGAG	61,09	175	hydrolase, acting on ester bonds
SCAGST3137D03.g	COS-EST-SSR 16	(GAC) ₅	TTTTGGGCAAAGAAGCTGTT	GAGGAGGTGCTCCTTCAATTC	60,20	186	SH3 domain-containing protein 3, putative, expressed
SCCCLR1024F03.g	COS-EST-SSR 20	(CCG) ₅	GATGATGCTGATCGACTGCT	CGCGAGTACCGATACGAGAT	60,26	228	Os10g0565100 protein
SCCCRT1004D09.g	COS-EST-SSR 23	(CG) ₇	GTCCCGTAGTCGTTGGAGTT	ACTTCAACCTCACGCTGCTC	60,60	244	Pyruvate decarboxylase isozyme 2
SCEZLB1006A05.g	COS-EST-SSR 26	(GGC) ₅	ACCCTCGACGACATCATCA	GCAGGGTACATGTCGTGCT	59,72	218	RNA and export factor-binding protein, putative, expressed
SCBGLR1082B07.g	COS-EST-SSR 29	(CCG) ₅	CTCGTGCTCCTCCTCATCCT	GATCTCCTGGATGTCCTTGG	59,46	159	Protein kinase domain containing protein, expressed

Tabela 2 - Cont.

ID SUCEST ^a	Primer	Motivo	Forward (5'→3')	Reverse (3'→5')	T _m (°C)	TP _{es} (pb)	Similaridade dos COS-EST-SSR
SCEZRZ1014H03.g	COS-EST-SSR 32	(CA) ₅	CATGGACCAGAAAGGGAAGA	GGCAGCCAATAGACAACACA	59,72	233	Putative ubiquitin conjugating enzyme 7 interacting protein 4
SCJFRZ2027B09.g	COS-EST-SSR 33	(CT) ₅	GGCAACCATGCATACAACCTG	GTGTAGCTGCTCGCCATGA	61,16	254	Squamosa promoter-binding-like protein 17
SCCCCL4004C12.g	COS-EST-SSR 36	(GA) ₅	CAAGCTGTGGATGGACTGAA	ATTCCTTTTGCAGGAGAGCA	59,96	207	Chromosome chr6 scaffold_25, whole genome shotgun sequence
SCJFLR1013E09.g	COS-EST-SSR 37	(GA) ₅	TGCACCAAGGAAGTGATTGA	CGGTTTGGATGGTATCCTGT	59,67	229	Mitochondrial inner membrane translocating protein-like protein
SCCCLR1022A08.g	COS-EST-SSR 40	(AGA) ₅	TGTTTCTGTTCCTGCGAAGA	GCAGCATCACTTCCTTCCTT	59,43	212	Putative kinetochore protein
SCCCCL5002C02.g	COS-EST-SSR 42	(CG) ₅	CTACGGTGCGGTCAAGAAGT	ACTGGGCATGGAGGTGATGT	62,76	196	isopenicillin N epimerase; zcef1; PCO061342; csu188; csu188(gfu); p-csu188; PCO061342_ov
SCCCLR1C02D02.g	COS-EST-SSR 45	(TG) ₅	GCGGAACCATATAAGGGACA	GAGTGGAGTGCCTGTGATGA	59,83	223	Probable chromatin-remodeling complex ATPase chain
SCVPCL6038H06.g	COS-EST-SSR 47	(CAG) ₅	CTTGCCATAGAGGGGAAGAA	CTCCACCCATGGAAGATGG	61,30	247	Heterogeneous nuclear ribonucleoprotein A2/B1-like
SCEZFL5091D09.g	COS-EST-SSR 48	(CG) ₅	GCCGTGATGTCTTCAAGAGG	GGAGGAAGTTGATGGTGAGG	59,51	186	Beta adaptin-like
SCEQLR1007F12.g	COS-EST-SSR 101	(AG) ₅	TCCTCATCAGCATCATCGTC	CCACAGGTGTGAGGAGGTG	60,14	228	upstream in-frame stop codon
SCVPLR1049E07.g	COS-EST-SSR 102	(TG) ₅	TTCCAGTTGCGGTGAGTGTA	CAAGAAGGCCACCAACAGTC	60,69	180	Putative uncharacterized protein - Sb01g009520
SCACLR1057A01.g	COS-EST-SSR 103	(CT) ₅	CAGGACGCTCTTTTTGAGGA	TCAGCTTTTCTTGGGTTCAGA	59,98	209	Putative uncharacterized protein - Sb01g010130

Tabela 2 - Cont.

ID SUCEST ^a	Primer	Motivo	Forward (5'→3')	Reverse (3'→5')	T _m (°C)	TP _{es} (pb)	Similaridade dos COS-EST-SSR
SCSFRT2071G03.g	COS-EST-SSR 104	(GAC) ₆	GAGGTTGACACCGTTATTGAC A	CGCCTCAAAGCTTTTGTACC	59,88	161	Putative uncharacterized protein - Sb01g015540
SCAGLR2026A05.g	COS-EST-SSR 105	(CAG) ₉	CTCCATGGACTTCTCCAAGA	GCTCAGGAAGTTCTTCCTTCC	59,45	223	Putative uncharacterized protein - Sb01g045810
SCCCLB1021A03.g	COS-EST-SSR 106	(GA) ₅	TTGGACAATGAGAGGACTGCT	CTCTTCAGGGCCATCTTCAG	59,94	248	Putative uncharacterized protein - Sb01g028010
SCCCLB1004A12.g	COS-EST-SSR 107	(TGA) ₅	CCTGCTAGCAAAGGTGCATA	CCATTCGCTCATCACTGCTA	59,97	160	Não mostrou homologia com nenhum gene de sorgo, no entanto, para milho e arroz Putative uncharacterized protein
SCEQRT2027A08.g	COS-EST-SSR 108	(CG) ₇	ACTTCAACCTCACGCTGCTC	GGTTGGTCCCGTAGTCGTT	59,84	248	Putative uncharacterized protein - Sb01g010130
SCRLLR1038B09.g	COS-EST-SSR 109	(CAG) ₅	GCATGTGTTGAGGTTGATGC	CCAGCTGTGCATCCTGATAA	59,82	207	Pyruvate decarboxylase isozyme 2
SCVPLB1018A01.g	COS-EST-SSR 110	(TGA) ₅	CATTTACAGGGGCTGGAAGA	ACAAGATGCCCACATGGAAT	60,20	198	Putative uncharacterized protein - Sb01g045810
SCAGST3141D06.g	COS-EST-SSR 111	(GC) ₅	AGAAGAAGCGGGTGGTCAT	CGGGAAGTCGGAGTACTGGT	61,46	188	Putative uncharacterized protein - Sb01g028010
SCSGST3120A05.g	COS-EST-SSR 112	(GC) ₅	TCCTCAGTATCTGCGCTTCA	AACCAGTACAACACGCTCCA	59,21	178	DnaK protein, expressed C3HC4 zinc finger
SCJFST1010F10.g	COS-EST-SSR 113	(GC) ₅	CTTCTCACCACCCTGCAGAC	GAGCCTCGTGCAGCTTCTT	60,84	151	containing protein putative, expressed
SCCCLR1C03A10.g	COS-EST-SSR 114	(GTC) ₅	GGGAAGCTCTGGAACCTCTG	CGACTTAGCCACTGACGAGA	59,19	195	Putative uncharacterized protein - Sb02g036280
SCJFRT2057C03.g	COS-EST-SSR 115	(CAG) ₆	TCAGCAAGAAGCAGCAGAAG	TTTGAAGTGGATCTCCTCGAC	57,78	151	Putative uncharacterized protein e Os02g0686400 protein, respectivamente
SCQGRT1038G12.g	COS-EST-SSR 116	(CG) ₅	GTGGCCGCAAGAACCAGTA	GAAGTTGACCTTGGCTTTGG	59,17	185	Putative uncharacterized protein - Sb03g012890

2.8 Análise de polimorfismo dos COS-EST-SSR

As marcas foram genotipadas com base na sua presença (1) ou ausência (0). Os fragmentos obtidos pela amplificação de cada par de *primers* foram considerados como alelos. O conteúdo de informação de polimorfismo (PIC) de cada par de primer COS-EST-SSR, foi obtido com base na expressão:

$PIC = 1 - \sum p_{ij}^2$, em que p_{ij} é a frequência do j-ésimo alelo (marca) para o i-ésimo locus COS-EST-SSR.

O tamanho dos alelos foi determinado por comparação com um marcador de peso molecular padrão (10-pb “ladder” – Invitrogen®).

A dissimilaridade entre os 24 genótipos utilizados para avaliação do polimorfismo dos COS-EST-SSRs foi estimada pelo complemento do coeficiente de similaridade de Jaccard e visualizadas pela construção de um dendrograma pelo método de agrupamento UPGMA (*Unweighted Pair Group Method with Arithmetic Averages*), ambos gerados pelo programa NTSYS-pc versão 2.0 (ROHLF 1992).

3 RESULTADOS

3.1 Identificação dos COS-EST-SSRs

Do total de 19.315 CSCs ortólogos entre cana e sorgo previamente identificados por Vicentini et al. (2012), 3.763 (19,48%) e 5.209 (27%) CSCs, respectivamente de cana e sorgo, apresentaram microsatélites (SSRs). Estas sequências de cana e sorgo contendo SSRs foram comparadas entre si, o que permitiu eliminar SSRs presentes apenas em cana ou em sorgo. Este processo resultou em 998 sequências ortólogas contendo SSRs que foram consideradas como pares de ortólogos em cana e sorgo. A triagem destes pares de ortólogos permitiu obter 365 pares de sequências ortólogas de cana e sorgo contendo o mesmo SSR (Figura 1).

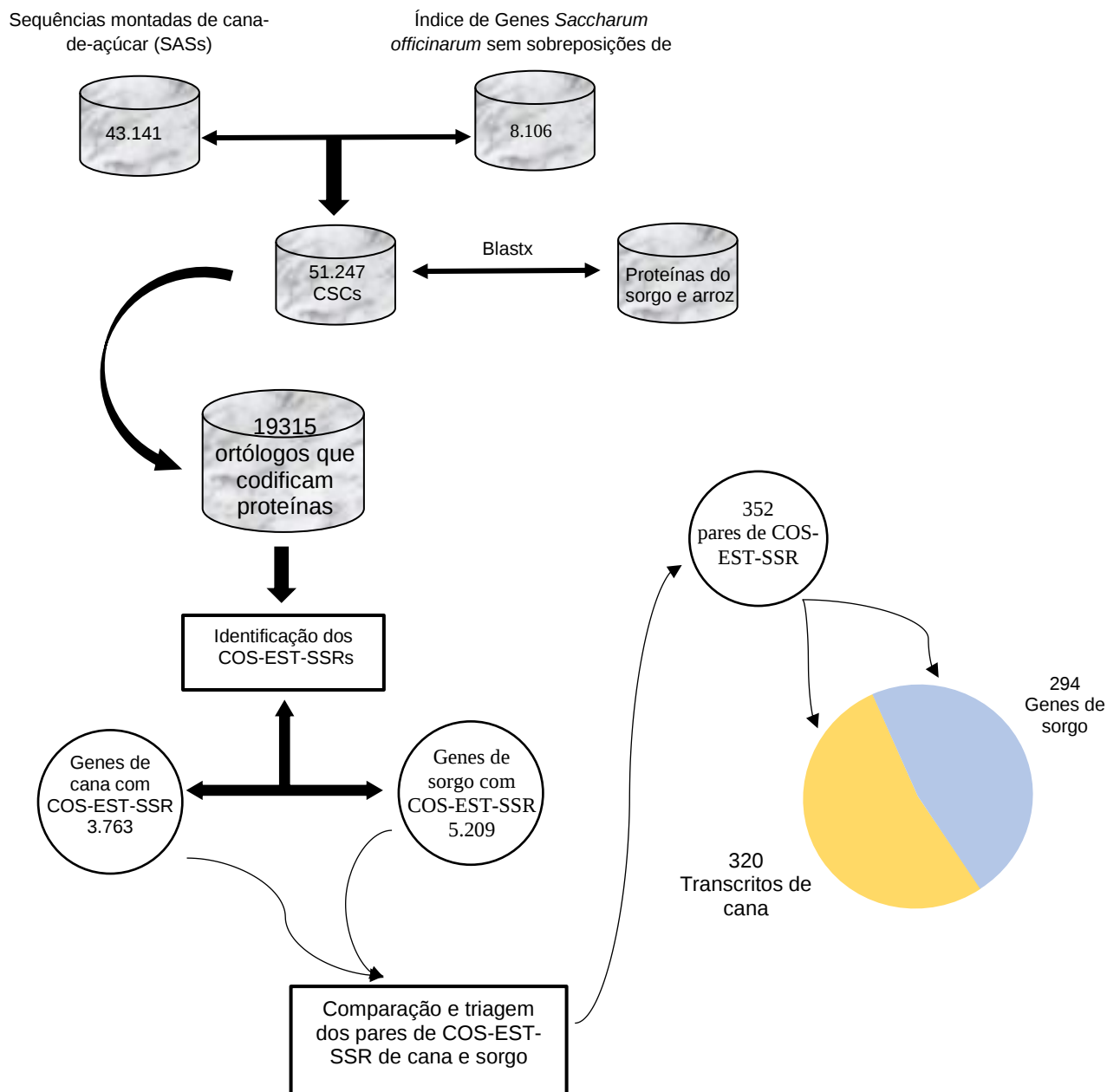


Figura 1 - Identificação de sequências EST-SSR ortólogos conservados (COS-EST-SSR) entre os genomas de cana e sorgo. **Fonte:** Adaptado de Vicentini et al. (2012).

O alinhamento dos pares de sequências ortólogas pelo software Clustal W permitiu a seleção de 352 pares de COS-EST-SSR entre cana e sorgo, os quais corresponderam a 320 clusters de cana alinhados com 294 genes de sorgo. Os 320 clusters de cana (COS-EST-SSRs) deram origem a um banco de dados para o desenho dos pares de *primers* COS-EST-SSRs.

Entre os 320 transcritos de cana (COS-EST-SSRs), 135 (42%) e 185 (58%) foram compostos de motivos de di e trinucleotídeos, respectivamente, enquanto não foram encontrados motivos de tetranucleotídeos repetidos acima de cinco vezes (Tabela 3).

3.2 Anotação funcional dos COS-EST-SSR

Dos 320 CSC, 151 (47%) apresentaram similaridade com diversos tipos de proteínas de várias espécies, porém apenas 41 sequências COS-EST-SSRs foram selecionadas para desenho de *primers*. As 41 sequências COS-EST-SSRs apresentaram similaridade a diversos tipos de proteínas (*endoglucanase 10*, *trehalose 6-phosphate phosphatase*, *non-specific serine/threonine protein kinase*, *TPR domain containing protein*, *lipase putative expressed*, *cytochrome P450 family protein*, *disulfide isomerase*, *pyruvate decarboxylase isozyme 2*, etc.) relacionadas, principalmente, com as espécies de sorgo, arroz e milho. Dos 41 COS-EST-SSRs desenhados, 28 (68%) mostraram similaridade com proteínas já caracterizadas (COS-EST-SSR 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20, 23, 26, 29, 32, 33, 36, 37, 40, 42, 45, 47, 48, 101, 109, 112 e 113). As 201 sequências remanescentes (53%) não apresentaram homologia com a espécie *Sorghum bicolor* (L.) Moench.

As sequências remanescentes tiveram similaridade com proteínas desconhecidas quando analisadas na base de dados UniProt (Tabela 2). Para efeito de resultados, a sequência COS-EST-SSR 6 não foi utilizada em análises subsequentes por não apresentar produto amplificado durante a eletroforese.

3.3 Análise de polimorfismo dos COS-EST-SSRs em cana e sorgo

Para o desenho de pares de *primers*, buscou-se selecionar os COS-EST-SSRs não redundantes quanto à similaridade de proteína, com ausência de gaps ou incompatibilidade de nucleotídeos na sequência ou em regiões muito próximas, de sequências com tamanho inferior a 150 pb e com microssatélites não posicionados no início ou no final da sequência (Figura 2).

Dos 320 COS-EST-SSRs, 41 foram selecionados para o desenho de *primers* (Tabela 2) com a finalidade de avaliar o potencial dos mesmos em gerar polimorfismo tanto nos genótipos de cana quanto nos de sorgo.

Dos 41 pares de *primers* avaliados quanto ao polimorfismo, 34 (~83%) produziram perfis desejáveis e puderam ser avaliados de acordo com o produto produzido (Figura 3), enquanto os outros 7 (17%) não amplificaram.

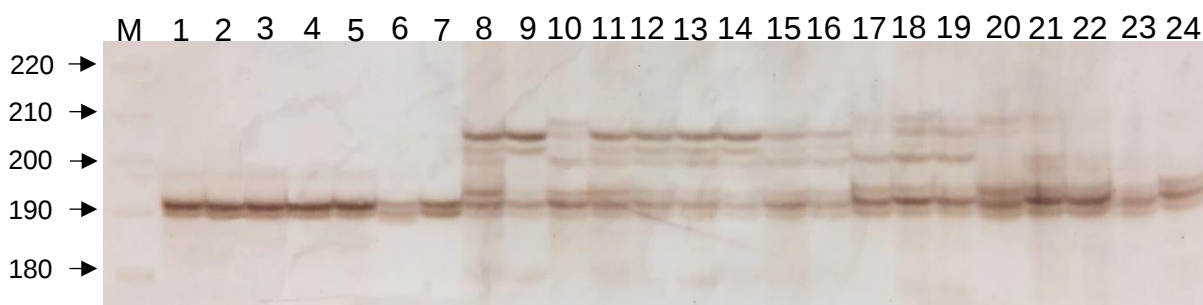


Figura 3 - Análise eletroforética dos produtos amplificados com *primers* COS-EST-SSR separados em gel de poliacrilamida desanaturante 6% e visualizados por meio de coloração com prata. M: marcador 10-pb; híbridos de sorgo (1: 1G100; 2: 1G220; 3: 1G233; 4: 1G244; 5: 1G282; 6: SS318; 7: 50A40); cultivares comerciais de cana (8: IACSP955000; 9: RB867517; 10: SP801842; 11: SP803250; 12: RB835054; 13: RB855453); clones de *S. officinarum* (14: Ajax; 15: Badila; 16: Badila de java; 17: Bourbon sur; 18: Creola; 19: Sabura); clones de *S. spontaneum* (20: Dacca; 21: Djantoerli; 22: Formosa; 23: Gehrabon e 24: Holes1).

Na tabela 3 são apresentados os 29 *primers* COS-EST-SSR que foram polimórficos nos 24 genótipos avaliados. Destes, 27 (79%) foram polimórficos nas cultivares de cana e acessos selvagens, enquanto nos híbridos de sorgo, apenas 6 (~18%) pares de *primers* mostraram-se polimórficos.

Tabela 3 - Classes de motivos de COS-EST-SSR e respectiva porcentagem de amplificação nos gêneros de *Saccharum* e *Sorghum*

Classes de repetições COS-EST-SSR			Pares de <i>primers</i> desenhados	Cana	Sorgo	Não amplificaram cana/sorgo
Di	Tri	Tetra		Polimórfico	Polimórfico	
135 (42%)	185 (58%)	0	41	27 (79%)	6 (18%)	7 (17%)

O número e o tamanho dos alelos variaram de 3 (COS-EST-SSR 4, COS-EST-SSR 8, COS-EST-SSR 16, COS-EST-SSR 32, COS-EST-SSR 42 e COS-EST-SSR 110) a 15 (COS-EST-SSR 109) e de 68 (COS-EST-SSR 4) a 319 (COS-EST-SSR 103) pb, respectivamente (Tabela 4).

Tabela 4 - Identificação dos COS-EST-SSRs, número de alelos (NA), variação do tamanho (pb), conteúdo de informação polimórfica (PIC) e heterozigosidade esperada (He) dos marcadores COS-EST-SSRs avaliados em 24 genótipos de cana e sorgo.

COS-EST-SSRs	NA	Amplitude (pb)	PIC	He
COS-EST-SSR 1	6	237 – 251	0,73	0,77
COS-EST-SSR 2	8	186 – 199	0,81	0,83
COS-EST-SSR 4	3	68 – 75	0,37	0,42
COS-EST-SSR 5	4	160 – 178	0,52	0,60
COS-EST-SSR 7	10	146 – 171	0,81	0,83
COS-EST-SSR 8	3	151 – 153	0,40	0,52
COS-EST-SSR 9	4	164 – 179	0,69	0,74
COS-EST-SSR 10	5	178 – 202	0,72	0,76
COS-EST-SSR 11	6	81 – 94	0,67	0,71
COS-EST-SSR 16	3	335 – 347	0,54	0,61
COS-EST-SSR 20	7	106 – 120	0,75	0,79
COS-EST-SSR 23	8	109 – 121	0,73	0,77
COS-EST-SSR 26	8	282 – 297	0,81	0,83
COS-EST-SSR 29	11	148 – 183	0,87	0,88
COS-EST-SSR 32	3	168 – 172	0,54	0,61
COS-EST-SSR 33	5	172 – 183	0,74	0,78
COS-EST-SSR 36	5	247 – 269	0,61	0,67
COS-EST-SSR 40	4	196 – 229	0,67	0,72
COS-EST-SSR 42	3	92 – 97	0,38	0,47
COS-EST-SSR 45	4	216 – 222	0,60	0,66
COS-EST-SSR 47	12	157 – 177	0,84	0,86
COS-EST-SSR 48	4	186 – 199	0,63	0,67
COS-EST-SSR 101	5	123 – 140	0,62	0,67
COS-EST-SSR 103	5	293 – 319	0,64	0,70
COS-EST-SSR 104	7	270 – 311	0,76	0,80
COS-EST-SSR 106	8	209 – 262	0,83	0,84
COS-EST-SSR 109	15	164 – 212	0,87	0,88
COS-EST-SSR 110	3	195 – 199	0,59	0,67
COS-EST-SSR 115	5	137 – 155	0,61	0,67
Média	6		0,67	0,71

Dos 29 pares de *primers* que foram polimórficos, 14 (48%) produziram o tamanho do produto amplificado de acordo com o esperado, tanto nos híbridos de sorgo quanto nos genótipos de cana. Vinte e um foram polimórficos em todos os genótipos de cana avaliados. Foi observado um número maior de alelos nestes

genótipos do que nos híbridos de sorgo. Foram detectados 174 alelos nos 29 primers COS-EST-SSR polimórficos, com um número médio de 6 alelos por locus. O valor do PIC calculado para estimar a informatividade de cada par de *primers* COS-EST-SSR, tanto em sorgo como em cana, variou de 0,37 (COS-EST-SSR 4) a 0,87 (COS-EST-SSR 29 e COS-EST-SSR 109), com um valor médio de 0,67 (Tabela 4).

A análise de agrupamento realizada com os marcadores COS-EST-SSRs separou os genótipos estudados em dois grupos em conformidade com o respectivo gênero. O grupo I agrupou os híbridos de sorgo, com uma dissimilaridade média entre eles de 0,138 enquanto o grupo II, formado pelos genótipos do gênero *Saccharum*, apresentou uma dissimilaridade média de 0,395 (Figura 4).

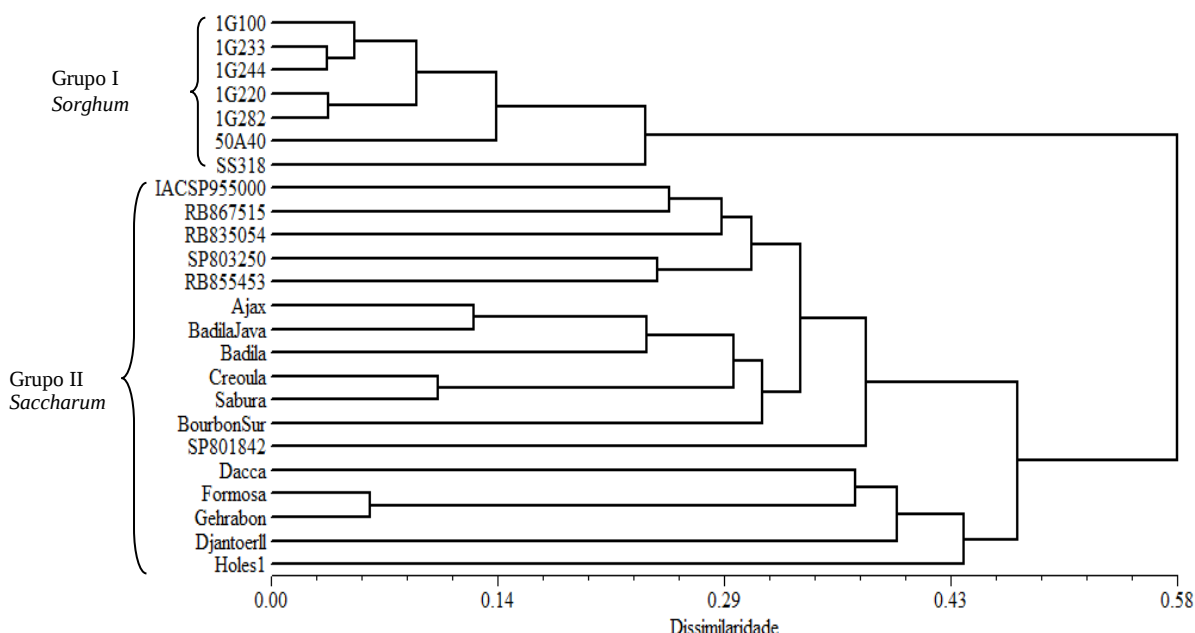


Figura 4 - Dendrograma (UPGMA - baseado na dissimilaridade estimada pelo complemento do coeficiente de similaridade de Jaccard para 24 genótipos.

No grupo II, os acessos de *S. spontaneum* formaram um sub-grupo distante dos acessos de *S. officinarum*, enquanto as cultivares, com exceção da SP801842, foram agrupadas em um mesmo sub-grupo próximo aos acessos de *S. officinarum*. A dissimilaridade média entre o grupo formado pelos híbridos de sorgo e o grupo formado pelos acessos de *S. officinarum* foi ligeiramente menor (0,578) que a encontrada (0,606) entre os híbridos de sorgo e os acessos de *S. spontaneum*. Considerando todos os genótipos analisados, a menor dissimilaridade (0,035) foi

observada entre os híbridos 1G244 e IG233 enquanto a maior (0,707) entre o acesso de Formosa (*S. officinarum*) e o híbrido de sorgo SS318. A dissimilaridade média entre todos os genótipos foi de 0,455.

4 DISCUSSÃO

Nos últimos anos, as ferramentas de bioinformática têm auxiliado significativamente na análise e compreensão de dados armazenados em base de dados. Com a ajuda desse recurso tornou-se possível identificar genes ortólogos a partir de regiões conservadas entre genomas e desenvolver marcadores derivados de sequências conservadas. Vários estudos têm identificado e analisado marcadores construídos a partir de sequências conservadas no genoma em várias espécies cultivadas de plantas (FULTON et al., 2002; CASTELBLANCO; FREGENE 2006; LI et al., 2008; CABRERA et al., 2009; WEI et al., 2012).

No presente estudo, os bancos de dados de sequências expressas de cana e sorgo foram explorados e um total de 352 pares de COS-EST-SSRs foram identificados a partir da utilização de 51.247 etiquetas de sequências expressas ortólogas conservadas (COS-EST) não redundantes entre cana e sorgo identificadas por Vicentini et al. (2012). Os marcadores COS-EST-SSR identificados neste estudo, podem ser considerados uma forte ferramenta para estudos de genética funcional e mapeamento comparativo, uma vez que se encontram conservados no genoma, bem como servir de marcadores âncoras para estudos de sintenia e filogenia entre as duas espécies ou gêneros relacionados da família Poaceae. De fato, Liu et al. (2013) estudando genomas sequenciados de cinco espécies de Poaceae (*Oryza sativa* Nipponbare/japônica; *Sorghum bicolor*; *Setaria italica*; *Brachypodium distachyon*; e *Zea mays*), identificaram 2.681 genes COS nos 12 cromossomos do arroz. Um estudo realizado por Jannoo et al. (2007) revelou alta colinearidade e similaridade entre sequências de sorgo e cana quando alinhadas. Embora as sequências de EST-SSR utilizadas neste trabalho tenham sido buscadas nos bancos de dados SoGI e SUSCEST, no resultado final foram identificadas sequências encontradas apenas no banco de dados do SUCEST. De certo modo este resultado era esperado, uma vez

que o projeto SUCEST abriga um maior número de ESTs de cana. Este resultado ocorreu naturalmente sem critérios que beneficiassem o banco do SUCEST.

Os resultados encontrados para trinucleotídeos estão de acordo com aqueles encontrados por James et al. (2011), que identificaram uma maior frequência de repetições de trinucleotídeos em sequências de EST de cana. Bushman et al. (2011) identificaram 85 e 6% de repetições tri e dinucleotídeos, respectivamente, em sequências de EST derivadas de unigenes do gênero *Dactylis* (família Poaceae). Essas diferenças podem ser atribuídas aos critérios pré-estabelecidos pelos autores no processo de identificação dos EST-SSRs, uma vez que não há uma definição de quantas repetições seriam necessárias para que a sequência seja considerada microsatélite. Singh et al. (2013), estudando diversidade genética em cana-de-açúcar utilizando marcadores EST-SSRs, identificaram uma porcentagem maior de repetições trinucleotídeos do que dinucleotídeos. De acordo com Bushman et al. (2011), as repetições de trinucleotídeos são ricas em regiões que codificam proteínas.

A similaridade dos marcadores COS-EST-SSRs foi investigada e nosso resultado foi superior aqueles relatados por Singh et al. (2013), em que 30% das sequências de EST-SSRs de cana apresentaram similaridade a proteínas conhecidas, enquanto que 70% resultaram em proteínas hipotéticas. Os resultados obtidos aqui com a análise de similaridade, demonstraram que 57% das sequências que continham microsatélites do tipo trinculeotídeo, e que estas estão associadas a genes que codificam proteínas com funções diferentes, enquanto que as remanescentes, a alguma proteína hipotética.

A partir da utilização dos marcadores COS-EST-SSRs, será possível atribuir informação funcional a genes com funções desconhecidas para espécies de gêneros relacionados ou ainda de mesma família, como é o caso da cana-de-açúcar e o do sorgo, podendo estes serem utilizados como marcadores universais em estudos de mapeamento genético. É possível que esse número seja maior, uma vez que foram considerados apenas os 41 pares de *primers* COS-EST-SSRs que foram sintetizados para uso neste estudo.

Na avaliação dos 41 *primers* COS-EST-SSRs, 34 puderam amplificar com sucesso bandas nos genótipos de cana e sorgo. No entanto, foi verificado uma maior porcentagem de polimorfismo entre os genótipos de cana (79%) do que nos híbridos

de sorgo (18%). Estes resultados podem ser atribuídos ao fato de que a cana é uma espécie altamente poliploide e aos marcadores COS-EST-SSRs serem derivados de EST-SSRs que apresentam uma natureza multialélica.

Marcadores COS-EST-SSRs tem apresentado bons resultados de polimorfismo quando avaliado entre espécies com gêneros relacionados (FULTON et al., 2002; LI et al., 2008; CABRERA et al., 2009; LIEWLAKSANEEYANAWIN et al. 2009; LIU et al., 2013; JEONG et al., 2014). No nosso trabalho, o número de alelos por locus variou de 3 a 15 nos genótipos de cana e sorgo, com uma média de 6 alelos. Estes resultados estão de acordo com os obtidos por Pinto et al. (2004) e Singh et al. (2013) que trabalharam com microssatélites derivados de sequências expressas. O valor médio do PIC (0,67) dos COS-EST-SSRs desenvolvidos no presente estudo, foi inferior aqueles obtidos em estudos anteriores envolvendo marcadores EST-SSRs em cana-de-açúcar (KITIPAT 2012; SINGH et al., 2013), porém semelhante ao obtido por Pinto et al. (2004).

Estes resultados sugerem que os marcadores COS-EST-SSRs avaliados aqui, originados de sequências conservadas, apresentam uma alta variabilidade entre os genótipos de cana. Isto pode ser explicado pela elevada complexidade encontrada no genoma da cana devido ao alto nível de ploidia com diferentes números de cromossomos no gênero *Saccharum*. Além disso, reforça a utilidade dos marcadores COS-EST-SSRs em detectar diferenças entre indivíduos de mesma espécie ou gênero relacionado. O presente estudo avaliou a eficiência dos marcadores COS-EST-SSRs derivados de sequências ortólogas conservadas baseada em análise de sequência e amplificação via PCR. Nossos resultados mostraram que os marcadores COS-EST-SSRs são úteis para tais propósitos.

Os marcadores COS-EST-SSRs foram eficientes no estudo de divergência genética. Pela análise de dissimilaridade, observa-se uma nítida separação entre os grupos formados pelos genótipos do gênero *Saccharum* (Figura 4). Embora as variedades de cana estudadas no presente trabalho sejam derivadas de sucessivas hibridações entre as espécies *S. officinarum* e *S. spontaneum*, os marcadores COS-EST-SSRs puderam agrupar as variedades de cana em um subgrupo diferente dos grupos das espécies *S. officinarum* e *S. spontaneum*. Liu et al. (2013) avaliando marcadores COS em espécies de bambu, construíram uma árvore filogenética e

encontraram resultados semelhantes à classificação morfológica. De acordo com Li et al. (2008), marcadores COS podem ser utilizados na reconstrução filogenética de espécies relacionadas. Já Enciso-Rodriguez et al. (2010) estudando marcadores COS, observaram uma pronunciada variabilidade entre acessos de lulo (*Solanum quitoense* Lam.). Para esses autores, os marcadores COSII são adequados para análises de diversidade.

5 CONCLUSÃO

Podemos concluir que a partir de sequências ortólogas conservadas entre os genomas de cana-de-açúcar e sorgo, foi possível desenvolver, caracterizar e testar a eficiência de marcadores COS-EST-SSRs. Estes marcadores podem ser utilizados em estudos de filogenia entre espécies relacionadas, na avaliação da variabilidade genética e no mapeamento comparativo.

6 REFERÊNCIAS

AITKEN, K. S.; MCNEIL, M. D.; BERKMAN, P. J.; HERMANN, S.; KILIAN, A.; BUNDOCK, P. C.; LI, J. C. Comparative mapping in the Poaceae family reveals translocations in the complex polyploid genome of sugarcane **Bmc Plant Biol**, v.14, p.1-15, 2014.

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**, v. 17, n. 3, p. 281-281, 1999.

BUSHAKRA, J. M.; SARGENT, D. J.; CABRERA, A.; CROWHURST, R.; GIRONA, E. L.; VELASCO, R.; SYMONDS, V. V.; KNAAP, E. V. D.; TROGGIO, M.; GARDINER, S. E.; CHAGNÉ, D. Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between *Malus* and *Fragaria*. **Tree Genet Genomes**, v.8, p.643-658, 2012.

BUSHMAN, B. S.; LARSON, S. R.; TUNA, M.; WEST, M. S.; HERNANDEZ, A. G.; VULLAGANTI, D.; GONG, G.; ROBINS, J. G.; JENSEN, K. B.; THIMMAPURAM, J. Orchardgrass (*Dactylis glomerata* L.) EST and SSR marker development, annotation, and transferability. **Theoretical and applied genetics Theoretische und angewandte Genetik**, v.123, p.119-129, 2011.

CABRERA, A.; KOZIK, A.; HOWAD, W.; ARUS, P.; IEZZONI, A. F.; VAN DER KNAAP, E. Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. **BMC genomics**, v.10, p.562, 2009.

CASTELBLANCO, W.; FREGENE, M. SSCP-SNP-based conserved ortholog set (COS) markers for comparative genomics in cassava (*Manihot esculenta* Crantz). **Plant Mol Biol Rep**, v.24, p.229-236, 2006.

CRESTE, S.; NETO, A. T.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining **Plant Mol Biol Rep**, v.19, p.299-306, 2001.

ENCISO-RODRIGUEZ, F.; MARTINEZ, R.; LOBO, M.; BARRERO, L. S. Genetic variation in the Solanaceae fruit bearing species lulo and tree tomato revealed by Conserved Ortholog (COSII) markers. **Genet Mol Biol**, v.33, p.271-274, 2010.

FULTON, T. M.; VAN DER HOEVEN, R.; EANNETTA, N. T.; TANKSLEY, S. D. Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. **Plant Cell**, v.14, p.1457-1467, 2002.

JAMES, B. T.; CHEN, C.; RUDOLPH, A.; SWAMINATHAN, K.; MURRAY, J. E.; NA, J. K.; SPENCE, A. K.; SMITH, B.; HUDSON, M. E.; MOOSE, S. P.; MING, R. Development of microsatellite markers in autopolyploid sugarcane and comparative analysis of conserved microsatellites in sorghum and sugarcane. **Molecular Breeding**, v.30, p.661-669, 2011.

JANNOO, N.; GRIVET, L.; CHANTRET, N.; GARSMEUR, O.; GLASZMANN, J. C.; ARRUDA, P.; D'HONT, A. Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome. **The Plant journal : for cell and molecular biology**, v.50, p.574-585, 2007.

JEONG, Y. M.; CHUNG, W. H.; CHUNG, H.; KIM, N.; PARK, B. S.; LIM, K. B.; Yu. H.; MUN, J. H. Comparative analysis of the radish genome based on a conserved ortholog set (COS) of Brassica. **TAG Theoretical and applied genetics Theoretische und angewandte Genetik**, v.127, p.1975-1989, 2014.

KAYESH, E.; BILKISH, N.; LIU, G. S.; CHEN, W.; LENG, X. P.; FANG, J. G. Characterization of EST-derived and non-EST simple sequence repeats in an F-1 hybrid population of *Vitis vinifera* L. **Genet Mol Res**, v.13, p.2220-2230, 2014.

KIM, C.; TANG, H.; PATERSON, A. H. Duplication and Divergence of Grass Genomes: Integrating the Chloridoids. **Tropical Plant Biology**, v.2, p.51-62, 2009.

KITTIPAT, U. Novel expressed sequence tag- simple sequence repeats (EST-SSR) markers characterized by new bioinformatic criteria reveal high genetic similarity in sugarcane (*Saccharum* spp.) breeding lines **African Journal of Biotechnology**, v.11, p.1337-1363, 2012.

KUHN, D. N.; LIVINGSTONE III, D.; MAIN, D.; ZHENG, P.; SASKI, C.; FELTUS, F. A.; MOCKAITIS, K.; FARMER, A. D.; MAY, G. D.; SCHNELL, R. J.; MOTAMAYOR, J. C. Identification and mapping of conserved ortholog set (COS) II sequences of cacao and their conversion to SNP markers for marker-assisted selection in *Theobroma cacao* and comparative genomics studies. **Tree Genet Genomes**, v.8, p.97-111, 2012.

LI, M.; WUNDER, J.; BISSOLI, G.; SCARPONI, E.; GAZZANI, S.; BARBARO, E.; SAEDLER, H.; VAROTTO, C. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species **Cladistics**, v.24, p.727-745, 2008.

LIEWLAKSANEEYANAWIN, C.; ZHUANG, J.; TANG, M.; FARZANEH, N.; LUENG, G.; CULLIS, C.; FINDLAY, S.; RITLAND, C. E.; BOHLMANN, J.; RITLAND, K. Identification of COS markers in the Pinaceae **Tree Genet Genomes**, v.5, p.247-255, 2009.

LINDQVIST-KREUZE, H.; CHO, K.; PORTAL, L.; RODRÍGUEZ, F.; SIMON, R.; MUELLER, L. A.; SPOONER, D. M.; BONIERBALE, M. Linking the potato genome to the conserved ortholog set (COS) markers **Bmc Genet**, v.14, p.1-12, 2013.

LIU, H.; GUO, X.; WU, J.; CHEN, G. B.; YING, Y. Development of universal genetic markers based on single-copy orthologous (COSII) genes in Poaceae **Plant cell reports**, 32:379-388, 2013.

MA, X. F.; JENSEN, E.; ALEXANDROV, N.; TROUKHAN, M.; ZHANG, L.; THOMAS-JONES, S.; FARRAR, K.; CLIFTON-BROWN, J.; DONNISON, I.; SWALLER, T.; FLAVELL, R. High Resolution Genetic Mapping by Genome Sequencing Reveals Genome Duplication and Tetraploid Genetic Structure of the Diploid *Miscanthus sinensis* **Plos One**, v.7, p.e33821, 2012.

PATERSON, A. H.; BOWERS, J. E.; BRUGGMANN, R.; DUBCHAK, I.; GRIMWOOD, J.; GUNDLACH, H.; HABERER, G.; HELLSTEN, U.; MITROS, T.; POLIAKOV, A.; SCHMUTZ, J.; SPANNAGL, M.; TANG, H.; WANG, X.; WICKER, T.; BHARTI, A. K.; CHAPMAN, J.; FELTUS, F. A.; GOWIK, U.; GRIGORIEV, I. V.; LYONS, E.; MAHER, C. A.; MARTIS, M.; NARECHANIA, A.; OTILLAR, R. P.; PENNING, B. W.; SALAMOV, A. A.; WANG, Y.; ZHANG, L.; CARPITA, N. C.; FREELING, M.; GINGLE, A. R.; HASH, C. T.; KELLER, B.; KLEIN, P.; KRESOVICH, S.; MCCANN, M. C.; MING, R.; PETERSON, D. G.; MEHBOOB UR, R.; WARE, D.; WESTHOFF, P.; MAYER, K. F.; MESSING, J.; ROKHSAR, D. S. The *Sorghum bicolor* genome and the diversification of grasses. **Nature**, v. 457, n. 7229, p. 551-6, 2009.

PINTO, L. R.; OLIVEIRA, K. M.; ULIAN, E. C.; GARCIA, A. A.; DE SOUZA, A. P. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome / National Research Council Canada = Genome / Conseil national de recherches Canada**, v.47, p.795-804, 2004.

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

RAMU, P.; KASSAHUN, B.; SENTHILVEL, S.; KUMAR, C. A.; JAYASHREE, B.; FOLKERTSMA, R. T.; REDDY, L. A.; KURUVINASHETTI, M. S.; HAUSSMANN, B. I. G.; HASH, C. T. Exploiting rice-sorghum synteny for targeted development of EST-SSRs to enrich the sorghum genetic linkage map. **TAG Theoretical and applied genetics Theoretische und angewandte Genetik**, v.119, p.1193-1204, 2009.

REMM, M.; STORM, C. E.; SONNHAMMER, E. L. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. **J Mol Biol**, v.314, p.1041-1052, 2001.

ROHLF, F. J. NTSYSpc: numerical taxonomy and multivariate system. Applied Biostatistics Inc., New York, 1992.

SINGH, R. K. et al. Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane **Gene**, v.524, p.309-329, 2013.

VICENTINI, R.; DEL BEM, L. E. V.; VAN SLUYS, M. A.; NOGUEIRA, F. T. S.; VINCENTZ, M. Gene Content Analysis of Sugarcane Public ESTs Reveals Thousands of Missing Coding-Genes and an Unexpected Pool of Grasses Conserved ncRNAs **Tropical Plant Biology**, v.5, p.199-205, 2012.

WANG, J. et al. Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes **BMC genomics**, v.11, p.1-17, 2010.

WANG, Z.; YAN HW, F. U. X. N.; LI, X. H, GAO, H. W. Development of simple sequence repeat markers and diversity analysis in alfalfa (*Medicago sativa* L.). **Mol Biol Rep**, v.40, p.3291-3298, 2013.

WEI, J. L.; HU, X. R.; YANG, J. J.; YANG, W. C. Identification of Single-Copy Orthologous Genes between *Physalis* and *Solanum lycopersicum* and Analysis of Genetic Diversity in *Physalis* Using Molecular Markers. **Plos One**, v.7, n. 11, p. e50164, 2012

ZHANG, D.; KONG, W.; ROBERTSON, J.; GOFF, V. H.; EPPS, E.; KERR, A.; MILLS, G.; CROMWELL, J.; LUGIN, Y.; PHILLIPS, C.; PATERSON, A. H. Genetic analysis of inflorescence and plant height components in sorghum (Panicoidae) and comparative genetics with rice (Oryzoidae) **Bmc Plant Biol**, v.15, p.1-15, 2015.

ZHANG, J. F.; FANG, H.; ZHOU, H. P.; SANOGO, S. M. A. Z. Y. Genetics, Breeding, and Marker-Assisted Selection for Verticillium Wilt Resistance in Cotton. **Crop Sci**, v.54, p.1289-1303, 2014.

CAPÍTULO III

MAPEAMENTO DE MARCADORES MOLECULARES DE SORGO EM CANA-DE-AÇÚCAR

Mapeamento de marcadores moleculares de Sorgo em Cana-de-Açúcar

RESUMO - A partir dos mapas de ligação é possível conhecer a organização do genoma e identificar os marcadores que estão ligados a características que são agronomicamente importantes. O presente trabalho teve por objetivo incorporar em um mapa prévio de ligação de uma população de mapeamento resultante do cruzamento bi-parental envolvendo o clone IACSP95-3018 e a cultivar IACSP93-3046 marcadores moleculares do tipo gSSRs e EST-SSRs de sorgo como marcadores microssatélites derivados de sequências ortólogas conservadas (COS-EST-SSRs) entre estas duas espécies. O experimento foi conduzido em campo, no período de 2012 a 2014, no Centro de Cana-IAC, em Ribeirão Preto-SP. Os 52 marcadores moleculares derivados de sorgo segregando em dose única foram incorporados no mapa prévio de ligação do cruzamento bi-parental. O mapa de ligação foi composto por 603 marcadores distribuídos em 138 grupos de ligação que deram origem a doze grupos de homologia, com uma extensão total de 7371 cM e uma distância média entre marcas de 12,22 cM, com distribuição irregular ao longo do grupo de ligação. A maioria das associações entre marcadores e o caractere Pol%Cana detectadas por meio da análise de marcas simples foram de efeito positivo contribuindo para o aumento deste caractere sendo apenas duas com contribuição negativa. De forma geral, os marcadores derivados de sorgo foram promissores no mapeamento da população aqui apresentada.

Palavra-chave: transferibilidade, microssatélites, Sequências Ortólogas Conservadas

Genetic mapping of Sorghum derived markers in Sugarcane

ABSTRACT - Genetic linkage maps allow the identification of markers linked to agronomical important traits. In the present study, sorghum microsatellite markers (SSRs) and Conserved Orthologous SSRs (COS-EST-SSRs) derived markers between sugarcane and sorghum were included in a previous map derived from a bi-parental cross between a clone (IACSP95-3018) and a sugarcane cultivar (IACSP93-3046) from the Instituto Agronômico de Campinas (IAC). The experiment was conducted in the field, in the period 2012-2014 at Ribeirão Preto, SP. The 52 single dose molecular markers derived from sorghum were incorporated in the previous bi-parental map. The final linkage map consisted of 603 markers distributed into 138 linkage groups and twelve homology groups. The total map length had 7371cM and an average distance between markers of 12.22 cM with irregular marker distribution across the linkage groups. The majority of the single marker-trait associations detected for Pol% Cane showed positive effects contributing to increase this trait (only two showed negative contribution to Pol% Cane). In general, the sorghum derived markers were promising in the genetic mapping of the bi-parental population evaluated.

Palavra-chave: transferability, microsatellites, Conserved Orthologous Sequences

1 INTRODUÇÃO

A cana-de-açúcar é um membro da família Poaceae, que pertence à tribo *Andropogoneae*, cuja respectiva tribo, inclui também as gramíneas tropicais e subtropicais, como os cereais do gênero *Sorghum* e *Zea* (HENRY, 2010).

Os gêneros *Saccharum* e *Sorghum* compartilharam do mesmo ancestral comum em pelo menos 10 milhões de anos atrás (JANNOO et al., 2007), sendo que o sorgo, é uma espécie diploide ($2n=20$), com um genoma relativamente pequeno, aproximadamente 730Mb (PATERSON et al., 2009) e menos complexo do que o genoma de outras culturas, como por exemplo, o da cana-de-açúcar, que apresenta genoma monoplóide, com um tamanho estimado em aproximadamente 760-930Mb (D'HONT; GLASZMANN, 2001; WANG et al., 2010). Essa íntima relação entre os genomas do sorgo e da cana-de-açúcar pode ser identificada através de comparações entre sequências de genes das duas espécies.

A utilização do genoma do sorgo, como referência, em estudos de mapeamento comparativo tem contribuído para acelerar a análise do genoma da cana-de-açúcar, principalmente, em relação à detecção de QTL (locus de características quantitativas) relacionados à resistência a doenças (MCINTYRE et al., 2005) e parâmetros agroindustriais (Brix, Pol%Cana e Fibra%) (JORDAN et al., 2004) obtidos através de marcadores provenientes de sondas heterólogas em ensaios de RFLP (*Restriction fragment lenght polymorphism*).

Comparando os genomas do sorgo e cana-de-açúcar, McIntyre et al. (2015) identificaram e mapearam marcadores em dose única, relatando a existência de regiões sintênicas com uma região do cromossomo 9 do sorgo. Aintken et al. (2014a) em um estudo de mapeamento comparativo entre espécies da família Poaceae, alinhou grupos homólogos de cana com grupos de ligação de sorgo utilizando marcadores EST-SSRs e identificaram um alto nível de sintonia (90%) entre os dois genomas.

A alta taxa de transferibilidade de marcadores microssatélites entre cana e sorgo (SINGH et al., 2011) abrem a oportunidade da utilização destes marcadores na realização de estudos de mapeamento comparativo, da mesma forma que marcadores do tipo COS-EST-SSRs entre estas duas espécies podem servir de

âncoras para a comparação de grupos de ligação em ambas as espécies. Atualmente, um grande número de marcadores moleculares, como AFLP (CHEN et al., 2015), SSRs (LU et al., 2015), SNP (MCINTYRE et al., 2015), e ISSR (FAHMY et al., 2016), tem sido utilizado em estudo genéticos envolvendo a cana-de-açúcar. No entanto, limitações como o tamanho do genoma da cana, o tempo para desenvolvimento desses marcadores e ferramentas de sequenciamento de DNA tem tornado o processo lento e dispendioso. O desenvolvimento de marcadores EST-SSRs pode ser feito a partir da mineração de sequências de SSRs em bancos de dados de uma forma simples, rápida e econômica. Uma vez que estes marcadores já se encontram desenvolvidos para uma determinada espécie, podem ser utilizados em estudos de transferibilidade ou mapeamento comparativo envolvendo espécies aparentadas, como é o caso da cana-de-açúcar e o do sorgo.

Neste contexto, este trabalho teve por objetivo incorporar marcadores derivados de sorgo, tais como EST-SSRs, gSSRs e COS-EST-SSRs em um mapa de ligação oriundo de uma população de mapeamento genético de cana-de-açúcar do Programa Cana IAC, bem como buscar possíveis associações entre estes marcadores e caracteres de qualidade (Pol%Cana e Fibr%).

2 MATERIAL E MÉTODOS

2.1 Material vegetal

Os marcadores EST-SSRs, gSSRs de sorgo e COS-EST-SSRs foram incorporados em um mapa prévio de ligação derivado de uma população de 175 indivíduos oriunda do cruzamento bi-parental envolvendo o clone elite IACSP95-3018 como genitor feminino e a cultivar IACSP93-3046 como genitor masculino, ambos pertencentes ao Programa de Melhoramento de Cana-de-Açúcar do Centro Avançado da Pesquisa Tecnológica do Agronegócio de Cana - Apta, Instituto Agrônomo de Campinas (IAC). Esta população vem sendo explorada para o mapeamento de diversos caracteres agronômicos (atributos biométricos e de qualidade) bem como para resistência a ferrugem marrom (MANCINI et al. 2010, SANTOS et al., 2015).

2.2 Genotipagem da população de mapeamento

2.2.1 Extração e quantificação do DNA genômico

O DNA genômico total foi obtido de 300mg de tecido foliar utilizando o procedimento de extração com base no método do CTAB (ALJANABI et al., 1999). O DNA dos genitores e da progênie foi quantificado na presença de um padrão de DNA do fago lambda de quantidades conhecidas em gel de agarose 0,8% (p/v) corado com brometo de etídio.

2.2.2 Marcadores microsatélites genômicos (gSSRs) e derivados de sequências expressas (EST-SSRs) de sorgo

Um conjunto de setenta pares de *primers* de microsatélites (SSRs) foram utilizados na população de mapeamento, sendo quarenta e um COS-EST-SSRs (descritos no capítulo 2), vinte microsatélites genômicos (gSSRs) de sorgo desenvolvidos pela Universidade do Texas (TAMU) (<http://sorgblast3.tamu.edu>), nove EST-SSRs de sorgo (CF756750, CN131982, CF488561, CF488525, CNL162, CNL170, CNL171, CNL182 e CNL183) obtidos por WANG et al. (2005) (Tabela 1).

2.2.3 Amplificação dos COS-EST-SSR

Os marcadores COS-EST-SSRs foram obtidos conforme descrito no Capítulo 2. Para os COS-EST-SSRs, as reações de PCR foram efetuadas em um volume final de reação de 18 µL, contendo 40 ng do DNA molde, 0,2µM de cada par de primer (forward e reverse), 100 µM de cada dNTP, 2,0 mM MgCl₂, 10 mM Tris-HCl, 50 mM KCl, e 0,5 unidades de Taq DNA polimerase. As seguintes condições de amplificação foram: desnaturação inicial à 94 °C por 3 minutos, seguida de 35 ciclos, cada ciclo contendo uma etapa de desnaturação à 94 °C por 1 minuto, temperatura de anelamento específica para cada par de primer (forward/reverse) por 1 minuto, extensão à 72 °C por 1 minuto; e um ciclo final à 72 °C por 2 minutos utilizando o termociclador Veriti (Applied Biosystems).

Aos produtos da amplificação da reação de PCR foram adicionados 7 µL de tampão formamida (98% formamida; 10 mM EDTA pH 8,0; 0.025% azul de bromofenol; 0,025% xileno cianol) seguido de desnaturação à 95 °C por 5 minutos.

2.3 Eletroforese em gel de poliacrilamida desnaturante

Os fragmentos foram separados por eletroforese em gel de poliacrilamida desnaturante a 6% w/v, utilizando padrão de peso molecular de 10 bp como marcador de peso molecular, a uma potência constante de 55 Watts. Os produtos da amplificação foram corados com nitrato de prata de acordo com o protocolo estabelecido por Creste et al. (2001) para visualização das marcas (alelos).

2.4 Fenotipagem da população de mapeamento

2.4.1 Delineamento experimental

O experimento foi conduzido em campo no período de 2012 a 2014, no Centro de Cana-IAC, em Ribeirão Preto-SP. O ensaio consistiu em um experimento em blocos casualizados, com 3 repetições, com parcelas representadas por um sulco de 2 metros de comprimento espaçadas em 1,5 metro entre linhas. A análise tecnológica foi conduzida a partir de 10 colmos coletados de cada parcela em cana soca, em julho de 2012, 2013 e 2014, totalizando 3 cortes (anos). As análises foram realizadas no Laboratório de Tecnologia do Centro de Cana adotando os métodos de análise tecnológica descritos no guia da Consecana (2006) para avaliação de Pol%Cana e Fibra%.

2.4.2 Análise fenotípica

Os dados fenotípicos dos 3 anos foram analisados utilizando o programa SAS (SAS 2008). Os contrastes ortogonais foram utilizados para obtenção das médias fenotípicas utilizando o procedimento PROC GLM adotando-se 0,05 como nível crítico de probabilidade.

Tabela 1- Relação dos pares de *primers* sequências forward e reverse, temperatura de anelamento (Ta) e tamanho de microssatélites (gSSRs e EST-SSRs) avaliados.

SSRs de sorgo	Sequência Forward	Sequência Reverse	Ta (°C)	Tam (pb)
BE355584 (AW15)	TCATGTACAGAGGCGACACG	AGGTCGCAACAGACACCTTC	55	190
BE125697 (AW2)	CGCTGCCAAAATCTAAGCTC	CACGGTGGTCACATCAGAAG	55	220
CF488525	GAAGAACCTCGTGAGGGCGT	CTTCTTGGAGCACTCCCGCA	-	166
CF488561	GCCATAACCAAGAGCTGCGTA	CGGTATTGGTTCTCGACGCAT	-	300
CF756750	TCTGCCCATGTTAACCGGCC	AGACGGGTACAATTGGCCGG	-	236
CN131982	GGAATGAGTAACTCGTCAGA	CCTTACTCATTGAGCAGTCT	-	302
CNL162	TTGGAGGGAAGAAGACAAGC	AGCATGAGACCTGGAAGCAG	-	257
CNL170	GAGGGCGTACAGGAAGAACA	CCGAGAAGGACTTGGTGAAG	-	206
CNL171	TCGAGTGCCTCAAGGAGTG	CAACAACGAAGAAACGACGA	-	187
CNL182	CTCCACAAACCACCACCAC	CCGGCCTCGTTGTAGTAGAC	-	260
CNL183	AGGGAAGGAACCAAGACCAT	CTCAGTCACCTCCACCGAAG	-	372
Xtxp14	GTAATAGTCATGACCGAGG	TAATAGACGAGTGAAAGCCC	50	160
Xtxp21	GAGCTGCCATAGATTTGGTCG	ACCTCGTCCCACCTTTGTTG	50	170
Xtxp24	CCATTGAGCTTCTGCTATCTC	TTCTAAGCCCACCGAAGTTG	50	145
Xtxp30	AAAAAGGACGCGCAGCTG	CTGGTCTCCACCATCCGTAG	50	273
Xtxp31	TGCGAGGCTGCCCTACTAG	TGGACGTACCTATTGGTG	50	222
Xtxp33	GAGCTACACAGGGTTCAAC	CCTAGCTATTCCTTGGTTG	50	160

Tabela 1 – Continuação

SSRs de sorgo	Sequência Forward	Sequência Reverse	Ta (°C)	Tam (pb)
Xtxp38	ACAAACCGCGACGAAGTAAC	ACAAGGCAAAGCCACAAAGC	50	170
Xtxp58	CAAAGTGCCCGGTTAAGACCT	TTCCCTTGCTGTTGCTTGTG	50	210
Xtxp60	GCTAGCTGACGCACGTCTCTG	TGCAACCGAGCGGTGACTA	55	224
Xtxp72	TTATGGAAGCAAAATGAC	CGAATCCTAATTGAGGTAAGC	50	180
Xtxp75	CGATGCCTCGAAAAAAAAACG	CCGATCAGAGCGTGGCAGG	50	172
Xtxp201	GCGTTTATGGAAGCAAAAT	CTCATAAGGCAGGACCAAC	55	230
Xtxp231	GGAAATCCAGGATAGGGT	AGGCAAAGGGTCATCA	50	188
Xtxp248	GGGTGTCCAATGTTGTCTGC	GGCCGTTACTGTCCCTTACTCA	50	140
Xtxp299	CTCTCCCCCTTTGTCATCCATC	TCTTGCCCCACCAGGACTTCTC	50	160
Xtxp304	ACATAAAAGCCCCTCTTC	CTTTCACACCCTTTATTCA	50	306
Xtxp316	CCAGCTTCACTTACGAGGAGATG	ATGCCCGTTTTCTAATTCTTCTACT	50	180
Xtxp343	CGATTGGACATAAGTGTTTC	TATAAACATCAGCAGAGGTG	50	100

A metodologia de modelos mistos foi utilizada para obter a predição dos valores genotípicos em relação à média geral, os BLUPs (*Best Linear Unbiased Predictors*), segundo o modelo:

$$Y = X\beta + Z\alpha + e$$

em que Y é o vetor das observações fenotípicas; X , a matriz de incidência dos efeitos fixos β ; Z , a matriz de incidência dos efeitos aleatórios α (genéticos ou ambientais); β , o vetor de efeitos fixos a serem estimados; α , o vetor de efeitos aleatórios ou valores genotípicos a serem preditos; e e , o vetor de erros aleatórios associados a cada observação. Para essa análise foi utilizado o procedimento PROC MIXED do programa SAS (SAS, 2008), em que o efeito de bloco e de ano foi considerado fixo e o de genótipo, considerado aleatório.

2.5 Análise segregação de marcadores em dose única

Os genitores e a população de mapeamento foram genotipados com base na presença (1) e ausência (0) da marca. A cada marca foi aplicado o teste estatístico qui-quadrado adotando-se nível de significância ($\alpha=0,05$) para avaliar a distorção de segregação em relação às proporções 1:1 e 3:1, as quais referem, respectivamente, marcador em dose simples em apenas um dos genitores e marcador em dose simples em ambos os genitores. A correção de Bonferroni foi aplicada para controlar o erro Tipo I para múltiplos testes (PROVINCE, 1999).

Os marcadores derivados dos pares de *primers* provenientes de sorgo foram designados segundo a seguinte nomenclatura: nome do par de *primer* seguido por um ponto e um número denotando o alelo (marcador) amplificado e uma letra referente a origem da marca em relação aos seus genitores (“D1” loco marcador presente no genitor feminino IACSP95-3018; “D2” loco marcador presente no genitor masculino IACSP93-3046 e “C” loco marcador presente em ambos genitores). Os marcadores obtidos de pares de *primers* provenientes de cana-de-açúcar foram designados da seguinte forma: a) marcadores microssatélites derivados de sequências expressas (SCA, SCB e SCC e CV) b) marcadores microssatélites genômicos (SMC e CIR) c) marcadores AFLP designados por letras.

2.6 Construção do mapa genético

A construção do mapa genético foi realizada com o auxílio do software OneMap (MARGARIDO et al., 2007). Aos marcadores EST-SSRs de sorgo e COS-EST-SSRs foram adicionados 898 marcadores em dose única (MDUs) derivados das técnicas de AFLP (*Amplified Length Polymorphism*), TRAP (*Target Region Amplification Polymorphism*), microssatélites genômicos e funcionais de cana-de-açúcar. Estes marcadores foram previamente genotipados na população de mapeamento pelo grupo de pesquisa da Dr^a Luciana Rossini Pinto, do Centro de Cana do IAC (MANCINI, 2010; SANTOS, 2013).

A frequência de recombinação foi estimada pelo método da máxima verossimilhança, em que foi considerado um LOD score mínimo de 3 e uma fração de recombinação $\leq 0,45$ como parâmetros para definição dos grupos de ligação (GL). A função de mapeamento Kosambi (1944) foi utilizada para converter as frações de recombinação em distâncias genéticas (cM). Em seguida, os grupos de ligação foram reunidos em grupos de homologia (GH). Os grupos de homologia putativos foram definidos com base na presença de marcadores microssatélites. Desta forma, os grupos de ligação contendo marcadores derivados do mesmo par de *primers* de SSRs foram agrupados no mesmo grupo de homologia.

Os grupos de ligação obtidos para cana-de-açúcar foram comparados com grupos de ligação de *Sorghum bicolor* obtidos em um mapa consenso publicado por Mace et al. (2009). Para comparação foram considerados apenas os marcadores SSRs.

2.7 Análise de associação

A análise de associação para os marcadores COS-EST-SSRs, gSSRs e EST-SSRs de sorgo foi realizada utilizando-se as médias fenotípicas ajustadas para os efeitos de blocos e de anos, considerando os atributos avaliados ao longo dos 3 cortes (2012, 2013 e 2014).

A metodologia utilizada para inferir a associação ou não dos marcadores com as medidas fenotípicas (Pol%Cana e Fibr%) foi à análise de regressão simples e regressão múltipla *stepwise*.

A análise de regressão simples seguiu o seguinte modelo:

$$Y_i = \beta_0 + \beta_1 X_i + e_i,$$

em que Y_i é a média ajustada (valor do fenótipo) do i -ésimo indivíduo; β_0 é a constante de regressão; β_1 é o coeficiente de regressão (efeito do marcador); X_i é o código do marcador (0 ou 1) e e_i é o erro aleatório na i -ésima observação, enquanto o modelo empregado para regressão múltipla foi

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni} + e_i$$

em que Y_i representa a média ajustada do i -ésimo indivíduo, referente às observações fenotípicas; β_0 , o intercepto da equação; X_{1i} a X_{ni} corresponde aos valores (0 ou 1) dos n marcadores moleculares; β_1 a β_n são coeficientes de regressão (efeitos simultâneos dos marcadores), que especificam a relação entre Y e os valores X_{1i} a X_{ni} ; e_i é o vetor de efeitos residuais. Para esta análise utilizou-se também o procedimento *stepwise*, a partir da marca mais relacionada (maior R^2 individual), cujas marcas adicionais foram selecionadas de modo a maximizar o R^2 conjunto. Nesse processo, marcas muito dependentes (ligadas) as marcas já selecionadas e que não possuem “p-value” adicional mínimo (15%) foram automaticamente excluídas, de modo que o processo escolheu o menor conjunto de marcas para R^2 múltiplo alcançado. Na escolha, também foi considerado o valor de $C(p)$ de Mallows, próximo ao número de marcas no modelo.

3 RESULTADOS

3.1 Marcadores em dose única gSSRs, EST-SSRs de sorgo e COS-EST-SSRs

Dos 70 pares de *primers* avaliados (COS-EST-SSRs, EST-SSRs e gSSR) 32 (46%) produziram perfis desejáveis de amplificação (Figura 1), enquanto os outros 38 (54%) foram monomórficos ou não produziram padrões claros de amplificação. Dos 32 pares de *primers* polimórficos foram obtidos 120 marcadores, dos quais 63 (52,5%)

segregaram 1:1 e 29 (24%) 3:1, totalizando 92 (77%) marcadores em dose única a serem utilizados para as análises de mapeamento (Tabela 2).

M LS P1 P2

Indivíduos da população de mapeamento

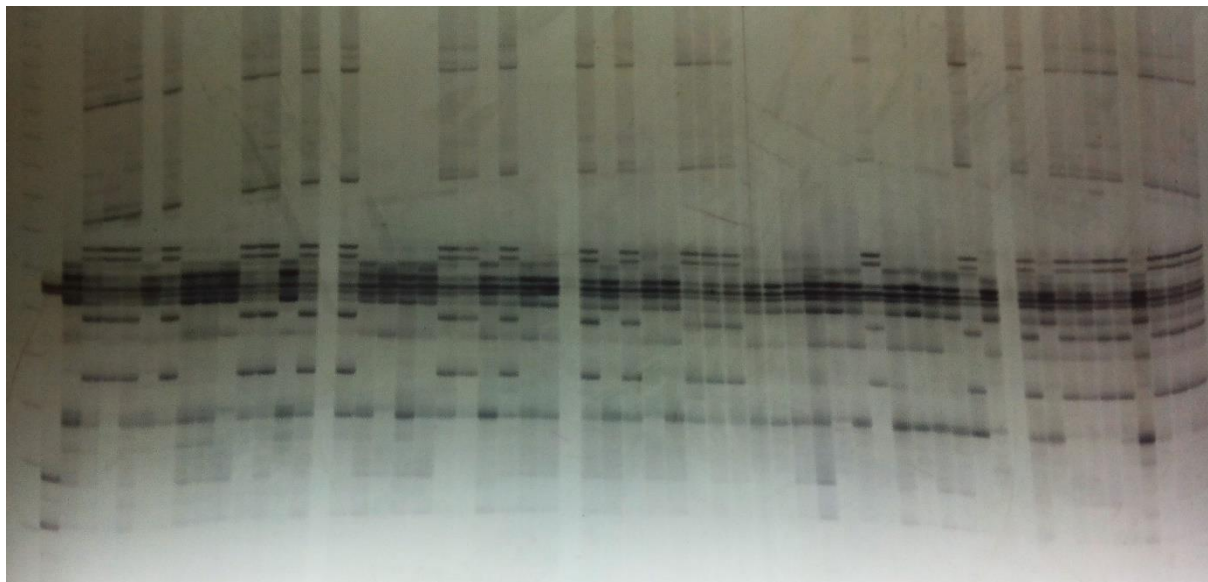


Figura 1 – Padrão de amplificação obtido para o locus COS-EST-SSR2, por eletroforese em gel de poliacrilamida. M-marcador de peso molecular (10pb); LS-linhagem de sorgo; P1-genitor feminino (IACSP95-3018); P2-genitor masculino (IACSP93-3046) e 54 indivíduos da população de mapeamento.

Tabela 2 - Número de marcadores em dose única obtidos por EST-SSR e gSSR de sorgo e COS-EST-SSRs em cruzamento bi-parental.

	EST-SSR	gSSR	COS-EST-SSRs	Total
^a N° marcadores em dose única (1:1)	14	17	32	63
^b N° marcadores em dose única (3:1)	8	9	12	29
Total (1:1 e 3:1)	22	26	44	92

^a Marcador presente em um genitor e ausente no outro segregando na proporção 1:1 na população bi-parental. ^b Marcador presente em ambos os genitores segregando na proporção 3:1 na população bi-parental.

3.2 Mapa de ligação

O mapa de ligação foi construído com 1050 (66,3%) marcadores em dose única (1:1 ou 3:1). Das 1050 marcas analisadas, 603 (57,4%) foram agrupadas em 138 grupos de ligação, enquanto 447 (42,6%) permaneceram não ligadas (Figura 2).

O número de marcadores por grupo de ligação variou de 2 (GL 4, 8, 13, 14, 15, 27, 30, 31, 36, 37, 39, 44, 55, 57, 59, 63, 64, 65, 69, 73, 74, 75, 77, 78, 79, 82, 85, 86, 87, 88, 89, 90, 91, 94, 95, 99, 100, 101, 104, 114, 115, 120, 127, 129, 121, 122, 126, 128, 131, 132, 133, 134, 135, 136, 137) a 33 (GL 1) enquanto o comprimento dos grupos de ligação variou de 0,7 cM (GL 90) a 512,1 cM (GL 1) com uma distância média entre marcadores de 12,22 cM com um comprimento total de 7.371 cM.

Do total de 138 grupos de ligação, 90 (65,3%) puderam ser organizados em 12 grupos de homologia, sendo que os 48 (34,7%) grupos de ligação restantes permaneceram como grupos independentes. O maior grupo de homologia (GH-I) foi constituído por 51 grupos de ligação enquanto os menores (GH-VI, GH-VII, GH-IX, GH-X e GH-XI) foram compostos por apenas 2 grupos de ligação cada (Figura 2).

Dos 92 marcadores derivados de sorgo (gSSRs, EST-SSRs e COS-EST-SSRs), 52 (56,5%) foram ligados aos 138 grupos de ligação, sendo que 40 marcadores foram agrupados dentro dos 12 grupos de homologia.

3.3 Comparação entre os grupos de ligação de cana-de-açúcar e o mapa consenso de sorgo

A comparação entre o mapa de ligação obtido a partir do cruzamento entre IACSP95-3018 e IACSP93-3046 com o mapa consenso do sorgo (http://sorgblast3.tamu.edu/linkage_groups.htm) (MACE et al., 2009) permitiu estabelecer prováveis correspondências entre os grupos de ligação destas duas espécies.

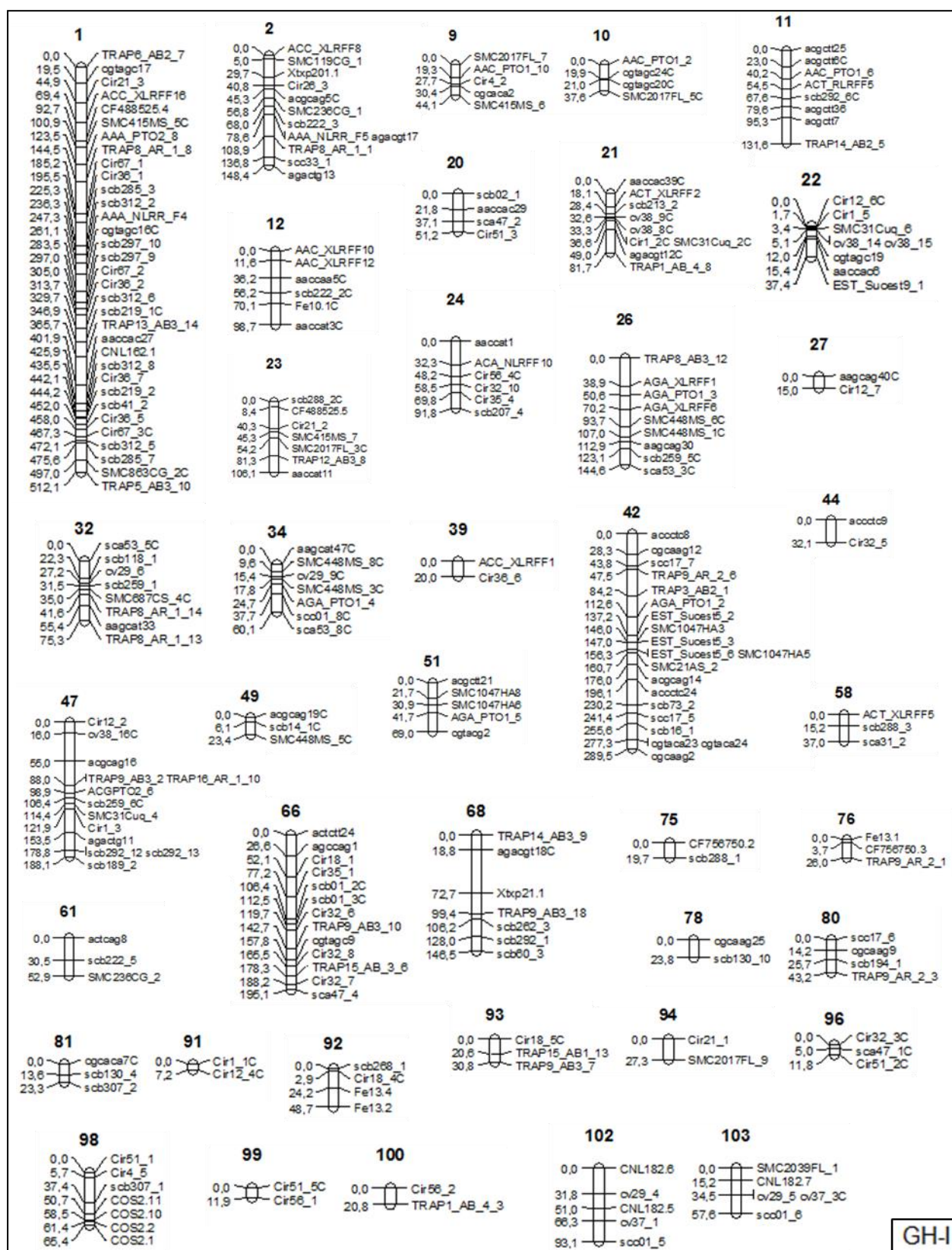


Figura 2 – Os números superiores referem-se ao grupo de ligação (GL). Os números à esquerda são as distâncias em centiMorgans (cM) entre os marcadores. Os nomes dos marcadores estão à direita, seguidos pela sigla D1 (presente no IACSP95-3018), D2 (presente no IACSP93-3046) e C (presente em ambos os genitores). Os GL foram reagrupados com base na presença de marcadores derivados de um mesmo loco de microssatélite (loco definido pelo par de *primers* que flanqueia o microssatélite) formando os grupos de homologia (GH)

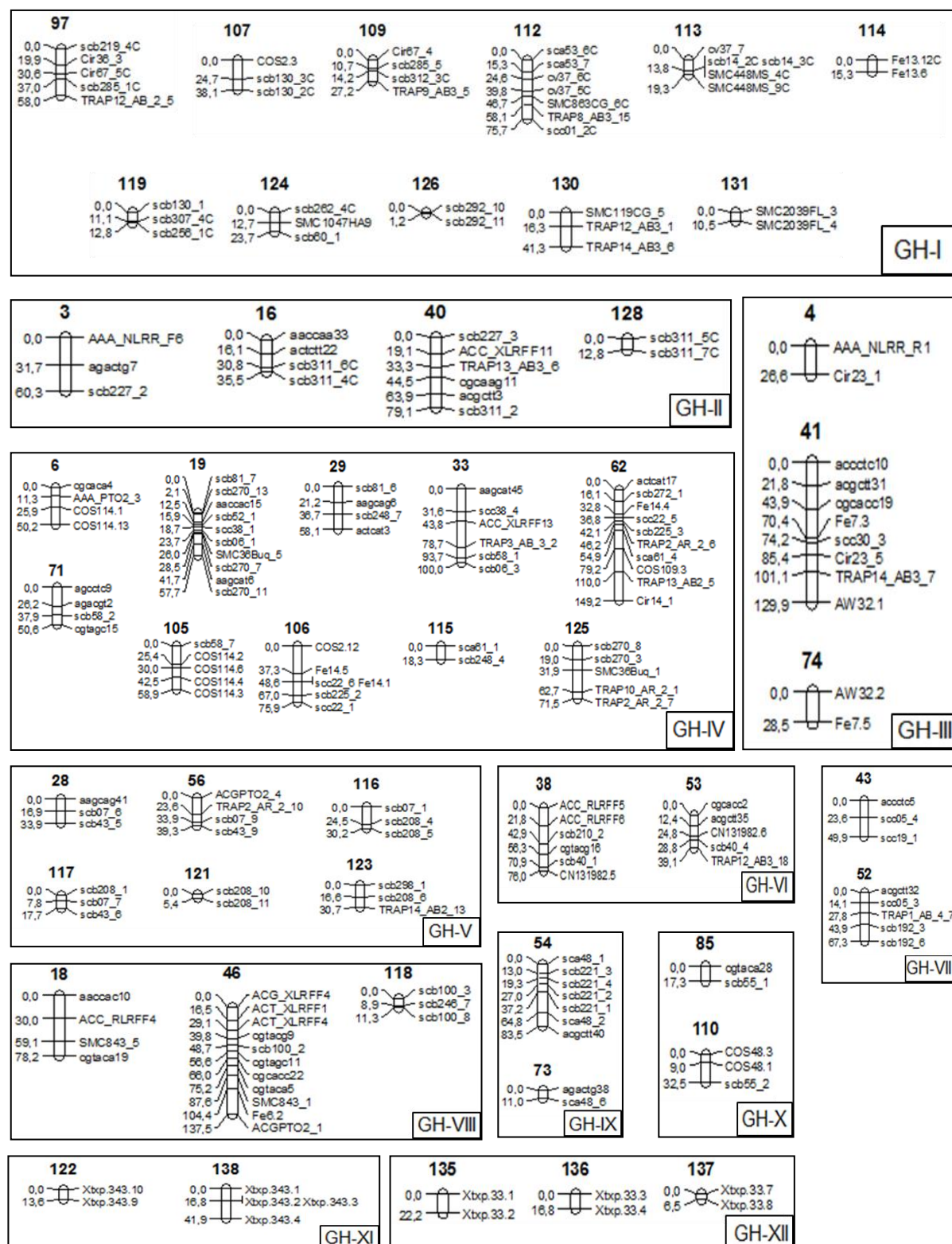


Figura 2 – Continuação

Grupos de ligação Independentes

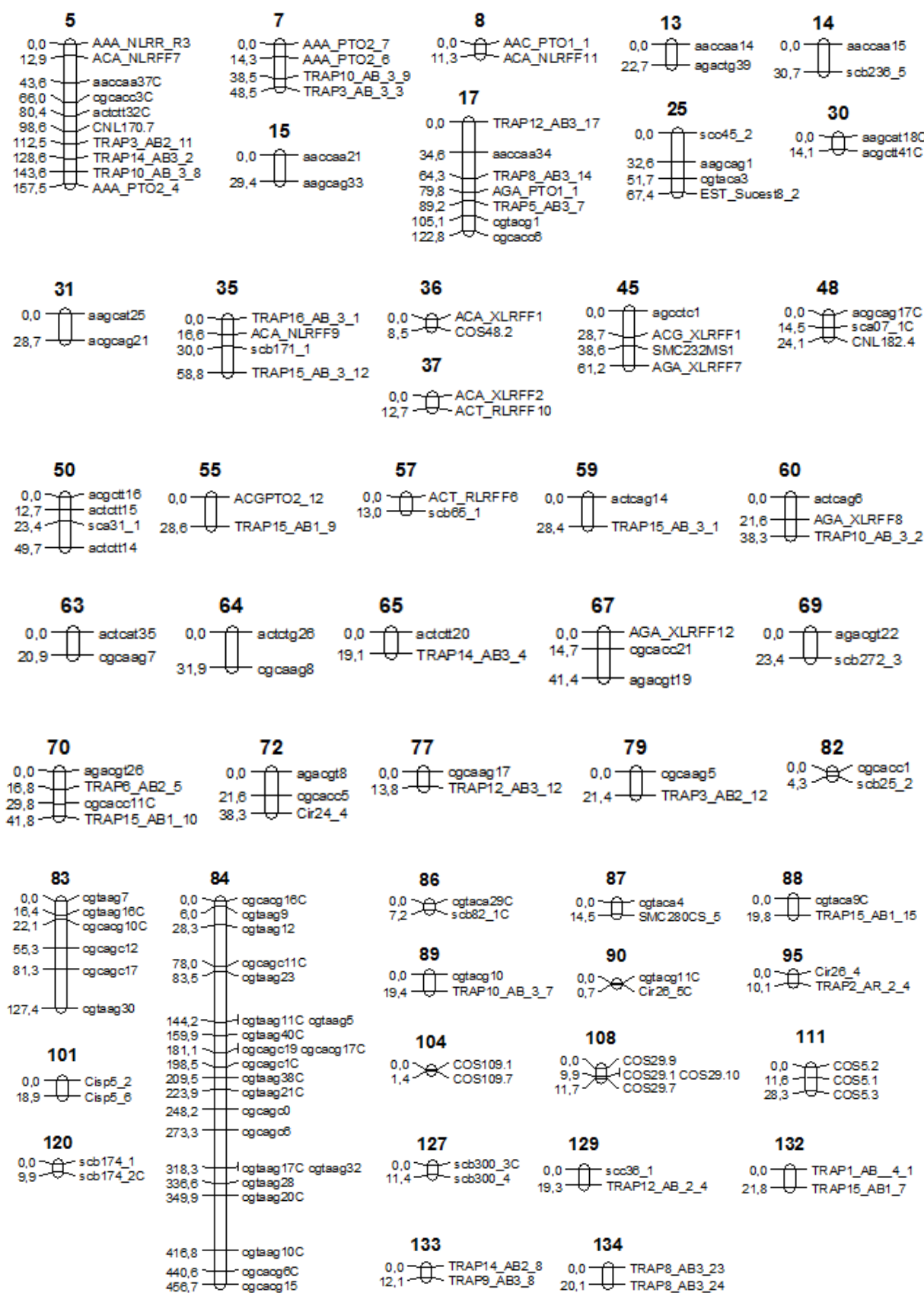


Figura 2 – Continuação

Dos vinte pares de *primers* de gSSRs de sorgo, apenas 4 (Xtxp21, Xtxtp33, Xtxp201 e Xtxp343) geraram marcadores que foram mapeados em 6 GL (GL68, GL135, GL136, GL137, GL2 e GL138) do mapa da cana-de-açúcar e que permitiram inferir uma provável sintenia com os grupos de ligação (2, 3 e 4) presentes no mapa consenso de sorgo (http://sorgblast3.tamu.edu/linkage_groups.htm) (Mace et al. 2009).

Desta forma, o loco Xtxp201 sugere uma correspondência entre o GL 2 do mapa da cana-de-açúcar com o GL 2 do sorgo. O loco Xtxp33 sugere uma correspondência entre os GL135, 136 e 137 de cana com o GL3 do sorgo. É interessante notar que os GL (135, 136 e 137) do mapa da cana pertencem ao mesmo grupo de homologia (GHXII). O loco Xtxp343 sugere uma correspondência entre os GL122 e 138 (ambos do GHXI) do mapa da cana-de-açúcar com o GL 4 do mapa do sorgo. Ainda este mesmo grupo de ligação do sorgo mostrou correspondência ao GL68 do mapa da cana em função do loco Xtxp21 (Figura 3).

3.4 Associação dos marcadores de sorgo com os caracteres fenotípicos

Foram avaliadas associações entre os marcadores e os valores fenotípicos referentes à Pol%cana e Fibra% da população de progênies. Os dados fenotípicos estão representados por médias ajustadas (três cortes e dois blocos por ano-corte). De modo geral, as médias mostraram relativa uniformidade (Tabela 3).

Tabela 3 – Análise descritiva dos dados fenotípicos da população de mapeamento derivada de um cruzamento bi-parental entre o clone IACSP95-3018 e a cultivar IACSP93-3046.

Variáveis	N	Média	Desvio Padrão	Mínimo	Máximo
Pol%Cana	175	13,89	0,71	11,44	15,49
Fibra%	175	13,16	0,68	11,87	17,79

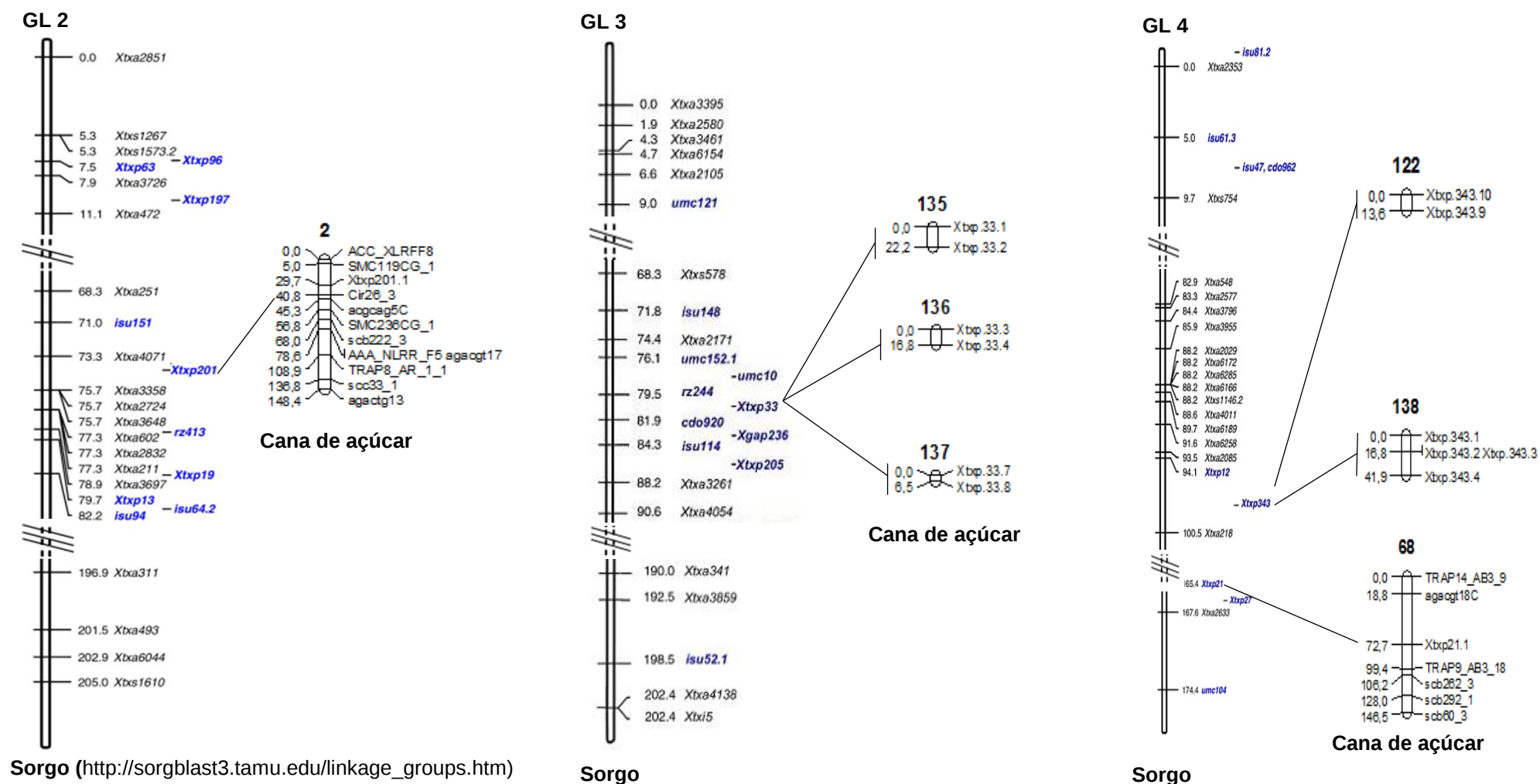


Figura 3 - Alinhamento dos grupos de ligação obtido a partir do cruzamento entre IACSP95-2018 e IACSP93-3046 com o mapa consenso do sorgo (http://sorgblast3.tamu.edu/linkage_groups.htm), MACE et al. (2009). A identificação do locus da cana e do sorgo encontra-se a direita do grupo de ligação de cada espécie. As linhas na diagonal entre os grupos de ligação da cana e do sorgo indicam a posição correspondente do locus.

Dos 92 marcadores, COS29_11 e COS29_13 mostraram-se iguais no conjunto de genótipos avaliados e a primeira foi eliminada para os estudos de associação, restando 91 marcadores.

A regressão múltipla com todos os 91 marcadores mostrou $R^2 = 62,72\%$ ($R^2_{aj}=16,25\%$) para Pol%Cana e $R^2=54,48\%$ ($R^2_{aj}=0,02\%$) para Fibra%. Tais valores são relativamente baixos, especialmente o R^2 ajustado da Fibra%, sugerindo ausência de associação para este caractere.

3.4.1 Pol%Cana

A análise de marcas simples para associação com Pol%Cana (Tabela 4) mostrou 13 marcadores com $p < 0,05$, sendo dois (Xtxp33_4, COS26_2) com efeito negativo, isto é, contribuem para diminuição do caractere e 11 com efeito positivo (aumento de Pol%Cana quando o marcador está presente). Nenhum dos marcadores atingiu *p-values* da ordem exponencial de (-5) , o que indicaria provável associação com QTL forte.

Tabela 4 – Efeito de marcadores COS-EST-SSRs, gSSR e EST-SSRs associados a variável Pol%Cana avaliados na população derivada do cruzamento bi-parental.

Marcador	Efeito do marcador no fenótipo	Pr > t
Xtxp21_1	0,2354	0,0298
Xtxp33_1	0,2547	0,0188
Xtxp343_2	0,2287	0,0348
Xtxp33_4	-0,2701	0,0352
XtxpP33_7	0,2900	0,0236
COS26_2	-0,1367	0,0090
COS29_7	0,2862	0,0231
COS29_10	0,3018	0,0156
COS29_13	0,2438	0,0249
COS47_1	0,2258	0,0372
CF56750	0,2141	0,0471
CNL182_5	0,2262	0,0363
CNL182_6	0,2836	0,0095

A regressão *stepwise* referente à associação entre marcadores e Pol%Cana (Tabela 5) mostrou que até 14 marcas podem ser incluídas para $p < 0,05$, com R^2 de 36,64%. Entretanto, o modelo com seis marcas foi sugerido em virtude da escolha de

um modelo com estatística C(p) (estatística Cp de Mallows) com valor próximo do número de parâmetros. Assim, foi obtida a seguinte equação (regressão múltipla) com os seis marcadores: $\text{Pol\%Cana} = 13,97192 - (0,82248 * \text{COS26_2}) + (0,2174 * \text{COS29_11}) + (0,39195 * \text{Xtxp33_7}) + (0,42945 * \text{COS26_3}) + (0,23363 * \text{Xtxp21_1}) - (0,33950 * \text{Xtxp58_4})$.

Tabela 5 – Percentagem da variação fenotípica explicada pelas associações ($R^2\%$) parciais e múltiplas (R^2 total) para o caráter Pol%Cana para marcadores do tipo COS-EST-SSRs, gSSR e EST-SSRs via procedimento *stepwise* do SAS.

Marcador	R^2 parciais	R^2 total	C(p)	F Value	Pr>F
COS26_2	0,0429	0,0429	29,5149	7,31	0,0076
COS29_11	0,0460	0,0890	22,3496	8,19	0,0048
Xtxp33_7	0,0365	0,1255	17,0851	6,72	0,0104
COS26_3	0,0282	0,1537	13,4665	5,34	0,0222
Xtxp21_1	0,0255	0,1792	10,3810	4,95	0,0275
Xtxp58_4	0,0272	0,2064	6,9636	5,42	0,0212
COS47_1	0,0244	0,2309	4,0971	4,99	0,0269
CNL182_5	0,0217	0,2526	1,7683	4,54	0,0347
COS29_7	0,0201	0,2728	-0,2403	4,29	0,0399
CNL183_4	0,0190	0,2917	-2,0158	4,12	0,0440
COS5_3	0,0195	0,3113	-3,8988	4,33	0,0390
Xtxp33_1	0,0186	0,3299	-5,6052	4,22	0,0416
Xtxp343_2	0,0203	0,3502	-7,6537	4,73	0,0313
Xtxp58_2	0,0162	0,3664	-8,8690	3,82	0,0524

3.4.2 Fibra%

A regressão *Stepwise* de associação entre marcadores e Fibra% (Tabela 6) mostrou que até quatro marcadores podem ser incluídos para $p < 0.05$, com R^2 de 14,15% (valor baixo). Entretanto, como não é possível escolher um modelo com estatística C(p) com valor próximo do número de parâmetros (todos C(p) são negativos) não foi possível fazer previsão de Fibra%, com as marcas analisadas.

Tabela 6 – Efeito de associações ($R^2\%$) parciais e múltiplas (R^2 total) para o caráter Fibr% para marcadores do tipo COS-EST-SSRs, gSSR e EST-SSRs via procedimento *stepwise* do SAS.

Marcador	R^2 parciais	R^2 total	C(p)	F Value	Pr>F
CNL183_4	0,0371	0,0371	-6,58	6,28	0,0132
COS2_9	0,0398	0,0769	-10,95	6,98	0,0091
CNL162_1	0,0452	0,1221	-16,21	8,29	0,0045
COS29_12	0,0234	0,1415	-17,96	4,38	0,0380

4 DISCUSSÃO

O mapa de ligação apresentou um comprimento total de 7.371 cM e uma densidade média de 12,22cM por marca, que foi 3,92 cM menor do que a encontrada por Liu et al. (2016), porém o comprimento total foi maior do que aquele identificado por Andru et al. (2011) e Singh et al. (2013b) que relataram uma cobertura do genoma da cana-de-açúcar de 5.387 cM e de 2606,77 cM, respectivamente. O número de marcas utilizado na construção do mapa de ligação foi suficiente para obter um número de GL (138) semelhante ao número de cromossomos existente nas cultivares modernas de cana-de-açúcar. No entanto, há a necessidade de saturar ainda mais o mapa de ligação uma vez que 55 grupos de ligação (39%) apresentaram apenas duas marcas por grupo. Esse resultado foi superior aquele obtido por Mancini (2010) e Santos (2013) que relataram 72 e 113 GL, respectivamente, nesta mesma população de mapeamento.

A saturação do mapa de ligação avaliado foi inferior àquela obtida em outros mapas genéticos construídos para a cultura da cana-de-açúcar (AITKEN et al., 2014b; LIU et al., 2016), porém superior aquela relatada por PALHARES et al. (2012). De fato, o número elevado de marcadores não ligados, aliado ao tamanho pequeno de alguns dos grupos de ligação e ao reduzido número de marcas por grupo indica que o mapa ainda não está saturado. Embora o mapa gerado nesse estudo não apresente um número de marcas suficiente para cobrir todo genoma da cana, poderá ser utilizado em estudos que visem a identificação de regiões do genoma da cana-de-açúcar responsáveis pelo controle genético de alguma característica de interesse. No entanto, é preciso aumentar a densidade do mapa genético com mais marcadores

para um melhor entendimento da organização genômica da cana-de-açúcar. Desse modo, os marcadores COS-EST-SSRs aqui desenvolvidos também poderão ser utilizados para integrar outros mapas genéticos desenvolvidos para a cana-de-açúcar.

Os grupos de ligação que apresentaram poucas marcas podem representar porções de outros grupos de ligação. Segundo Smith et al. (2005) pequenos grupos de ligação podem representar porções periféricas de cromossomos que são representados por grupos maiores. Isso pode ser devido ao fato da quantidade de marcadores utilizados nesse estudo não ter sido suficiente para cobrir todo o genoma da cana-de-açúcar. Sendo assim, esse resultado indica que mais marcadores são necessários para unir estes grupos de ligação formando grupos maiores e mais saturados.

De modo geral, os microssatélites de sorgo mostraram-se eficientes para serem utilizados no mapeamento em cana-de-açúcar, visto que 77% (92 marcadores) segregaram em dose única e destes 56,5% (52) foram incorporados no mapa prévio de ligação do cruzamento bi-parental (IACSP95-3018 x IACSP93-3046).

Embora uma taxa de 80% de transferibilidade de EST-SSRs entre genótipos de cana e sorgo foi encontrada por Singh et al. (2013a), vale ressaltar que a transferibilidade de 46% obtida em nosso trabalho se refere a apenas dois genótipos, ou seja, os genitores de uma população bi-parental, em que ambos representam genótipos melhorados, que apesar de contrastantes apresentam um certo grau de parentesco.

Seis (COS-EST-SSR2, COS-EST-SSR5, COS-EST-SSR29, COS-EST-SSR48, COS-EST-SSR109, COS-EST-SSR114) dos 23 pares de *primers* COS-EST-SSRs, ou seja, 26% foram genotipados na população de mapeamento. No total, estes 6 pares de *primers* COS-EST-SSRs produziram 65 marcadores segregantes, dos quais, 60% (39) encontram-se em dose única. Destes, 25 (64%) foram incorporados no mapa ao longo de 9 grupos de ligação e 3 grupos de homologia (GH-I, GH-V e GH-X). Tal resultado pode ser considerado satisfatório, já que se trata de marcadores desenvolvidos a partir de regiões ortólogas conservadas.

Ainda, podemos destacar os COS-EST-SSR2 e COS-EST-SSR5, que amplificam regiões com similaridade as proteínas da família trealose-fosfatase e citocromo P450. Os genes da família trealose-fosfatase estão envolvidos em

processos de estresse abióticos, como por exemplo, tolerância a seca e ao frio (NUCCIO et al., 2015; XIE et al., 2015), enquanto os genes da família citocromo P450 estão envolvidos no mecanismo de defesa de algumas plantas da família Poaceae, como sorgo (DU et al., 2010) e milho (IRMISCH et al., 2015). Sekhwal et al. (2013) realizaram a predição funcional de proteínas anotadas como proteínas putativas descaracterizadas no genoma do sorgo.

Os COS-EST-SSRs estudados aqui, por exibirem similaridade com genes candidatos podem auxiliar a caracterização funcional de proteínas e posteriormente serem utilizados na seleção assistida para caracteres relacionados ao estresse abiótico ou mecanismos de defesa da planta. Além disso, este estudo reforça a importância e utilidade dos marcadores COS-EST-SSRs em identificar loci associados a características de interesse agrônomo a fim de acelerar programas de melhoramento da cana-de-açúcar.

Dos dez pares de *primers* de gSSR de sorgo que geraram polimorfismo na população de mapeamento, apenas 4 (40%), (Xtxp21, Xtxp33, Xtxp201 e Xtxp343) tiveram os seus produtos de amplificação mapeados em 7 grupos de ligação (68, 135, 136, 137, 2, 122 e 138) o que comprometeu a comparação entre os grupos de ligação do mapa do cruzamento bi-parental com o do mapa consenso do sorgo. Apesar da quantidade de marcadores utilizados não ter sido suficiente para identificar a alta colinearidade existente entre os genomas de cana e sorgo, esta análise comparativa preliminar permitiu identificar possíveis regiões homólogas entre os genomas de cana e sorgo. Ainda que os resultados sejam preliminares devido ao reduzido número de marcadores de sorgo que foram mapeados em cana é possível que aqueles que foram mapeados sinalizem as regiões cromossômicas da cana com maior sintonia com o sorgo. Segundo Aitken et al. (2014a) alguns cromossomos da cana-de-açúcar mostram maiores níveis de sintonia com sorgo do que outros.

A maioria das associações entre marcadores e o caractere Pol%Cana detectadas por meio da análise de marcas simples foram de efeito positivo, ou seja, contribuem para o aumento deste caractere. Entretanto, dois marcadores Xtxp33_4 e COS26_2 contribuíram negativamente para incrementos das médias. Desta forma, incrementos na média desta variável poderiam ser atingidos com a ausência destes marcadores em genótipos selecionados em programas de melhoramento.

Por meio da análise de regressão *stepwise* (com $p < 0,05$) 14 marcadores mostraram-se associadas com o caráter Pol%Cana, os quais explicaram, conjuntamente, 36,64% da variação fenotípica. Esse resultado foi superior aquele relatado por Mancini (2010) e Liu et al. (2016) para este mesmo caractere. Dos 14 marcadores, sete (Xtxp21_1, CNL182_5, COS5_3, COS29_7, Xtxp33_1, Xtxp33_7 e Xtxp343_2) encontram-se localizados em 7 GL (68, 102, 108, 111, 135, 137 e 138, respectivamente). Entre esses marcadores merece destaque o COS5_3, por ser derivado do par de *primer* COS_EST-SSR5 desenhado a partir de uma sequência com similaridade a proteína da citocromo P450. Yu et al. (2015) verificaram que o gene do citocromo P450 é um dos alvos principais de miRNAs (micro RNAs) diferencialmente expressos durante o acúmulo de açúcar no colmo de sorgo sacarino. Entretanto, segundo a estatística de Mallows o modelo com seis marcadores se mostrou mais adequado. Pela equação de Mallows, o melhor valor predito esperado seria alcançado com ausências (valor zero de COS26_2 e Xtxp58_4) e presenças (valor um de COS29_11, Xtxp33_7, COS26_3 e Xtxp21_1), o que poderia ser utilizado como “tentativa” de seleção assistida. Portanto, o Pol%Cana predito seria de 15,24 ($13,97 + 1,27$), ou seja, ganho esperado de 1,27 sobre a média do Pol%Cana obtido na média das progênies.

Para o caractere Fibras os efeitos foram pequenos a ponto de ter de fato um marcador com potencial de associação a esse caractere. A baixa associação encontrada pode ser atribuída entre outros fatores pela relativa uniformidade dos dados fenotípicos apresentados pelas progênies da população estudada. Resultados semelhantes têm sido relatados em estudos de associação envolvendo os caracteres Pol%Cana e Fibras com efeitos pequenos. Pinto et al. (2010) identificaram 15 e 34 associações de efeito pequeno para Fibras e Pol%Cana, respectivamente. Já Alwala et al. (2009), observaram 6 associações para o caráter Pol%Cana e uma variação fenotípica total de 17,2%. Sabe-se que grande parte das características de interesse agrônomo é controlada por vários genes. Além disso, as cultivares modernas de cana agregam uma grande quantidade de alelos favoráveis, o que contribui para diminuição da variabilidade genética. Banerjee et al. (2015) relataram efeitos menores na variação fenotípica e atribuíram esse resultado a baixa diversidade alélica da população.

5 CONCLUSÃO

A quantidade de grupos de ligação com baixo número de marcadores reforça a necessidade de uma maior cobertura do mapa de ligação do cruzamento bi-parental aqui apresentado.

Os marcadores ortólogos conservados (COS-EST-SSRs) entre cana-de-açúcar e sorgo aqui desenvolvidos podem ser utilizados no mapeamento genético da cana-de-açúcar visando estudos de mapeamento comparativo com o sorgo.

Os marcadores COS-EST-SSRs poderão ser integrados a mapas genéticos de referência, contribuindo para aumentar a quantidade de informações disponíveis para cana-de-açúcar e outras espécies relacionadas.

6 REFERÊNCIAS

AITKEN, K. S.; MCNEIL, M. D.; BERKMAN, P. J.; HERMANN, S.; KILIAN, A.; BUNDOCK, P. C.; LI, J. C. Comparative mapping in the Poaceae family reveals translocations in the complex polyploid genome of sugarcane. **Bmc Plant Biology**, v. 14, 2014a.

AITKEN, K. S.; MCNEIL, M. D.; HERMANN, S.; BUNDOCK, P. C.; KILIAN, A.; HELLER-USZYNSKA, K.; HENRY, R. J.; LI, J. C. A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array Technology (DART) markers. **Bmc Genomics**, v. 15, 2014b.

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**, v. 17, n. 3, p. 281-281, 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**, v. 167, n. 1, p. 127-142, 2009.

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

BANERJEE, N.; SIRAREE, A.; YADAV, S.; KUMAR, S.; SINGH, J.; KUMAR, S.; PANDEY, D. K.; SINGH, R. K. Marker-trait association study for sucrose and yield contributing traits in sugarcane (*Saccharum* spp. hybrid). **Euphytica**, v. 205, n. 1, p. 185-201, 2015.

CHEN, J. W.; LAO, F. Y.; CHEN, X. W.; DENG, H. H.; LIU, R.; HE, H. Y.; FU, C.; CHEN, Y. S.; LIU, F. Y.; LI, Q. W.; JACKSON, P. DNA Marker Transmission and Linkage Analysis in Populations Derived from a Sugarcane (*Saccharum* spp.) x *Erianthus arundinaceus* Hybrid. **PLoS ONE**, v. 10, n. 6, 2015.

CONSELHO DOS PRODUTORES DE CANA DE AÇÚCAR, AÇÚCAR E ÁLCOOL DO ESTADO DE SÃO PAULO - CONSECANA. **Manual de instruções**. 5.ed. Piracicaba, 2006. 112p.

CRESTE, S.; NETO, A. T.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, v. 19, n. 4, p. 299-306, 2001.

D'HONT, A.; GLASZMANN, J. C. Sugarcane genome analysis with molecular markers: A first decade of research. **International Society of Sugar Cane Technologists, Vol II, Proceedings**, p. 556-559, 2001.

DU, Y.; CHU, H.; WANG, M.; CHU, I. K.; LO, C. Identification of flavone phytoalexins and a pathogen-inducible flavone synthase II gene (SbFNSII) in sorghum. **J Exp Bot**, v. 61, n. 4, p. 983-94, 2010.

FAHMY, E. M.; EL-GAWAD, N. M. A.; EL-GEDDAWY, I. H.; SALEH, O. M.; EL-AZAB, N. M. Development of RAPD and ISSR markers for drought tolerance in sugarcane (*Saccharum officinarum* L.). **Egyptian Journal of Genetics And Cytology**, v. 37, n. 1, 2016.

HENRY, R. J. Basic Information on the Sugarcane Plant. In: HENRY, R. J.; KOLE, C. **GENETICS, GENOMICS AND BREEDING OF SUGARCANE**. New York: Science Publishers, 2010. cap. 1, p. 1-7.

IRMISCH, S.; ZELTNER, P.; HANDRICK, V.; GERSHENZON, J.; KOLLNER, T. G. The maize cytochrome P450 CYP79A61 produces phenylacetaldoxime and indole-3-acetaldoxime in heterologous systems and might contribute to plant defense and auxin formation. **BMC Plant Biol**, v. 15, p. 128, 2015.

JANNOO, N.; GRIVET, L.; CHANTRET, N.; GARSMEUR, O.; GLASZMANN, J. C.; ARRUDA, P.; D'HONT, A. Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome. **Plant J**, v. 50, n. 4, p. 574-85, 2007.

JORDAN, D. R.; CASU, R. E.; BESSE, P.; CARROLL, B. C.; BERDING, N.; MCINTYRE, C. L. Markers associated with stalk number and suckering in sugarcane colocate with tillering and rhizomatousness QTLs in sorghum. **Genome**, v. 47, n. 5, p. 988-993, 2004.

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

LIU, P.; CHANDRA, A.; QUE, Y.; CHEN, P.-H.; GRISHAM, M. P.; WHITE, W. H.; DALLEY, C. D.; TEW, T. L.; PAN, Y.-B. Identification of quantitative trait loci controlling sucrose content based on an enriched genetic linkage map of sugarcane (*Saccharum* spp. hybrids) cultivar 'LCP 85-384'. **Euphytica**, v. 207, n. 3, p. 527-549, 2016.

LU, X.; ZHOU, H.; PAN, Y. B.; CHEN, C. Y.; ZHU, J. R.; CHEN, P. H.; LI, Y. R.; CAI, Q.; CHEN, R. K. Segregation analysis of microsatellite (SSR) markers in sugarcane polyploids. **Genet. Mol. Res.**, v. 14, p. 18384-18395, 2015.

MACE, E. S.; RAMI, J. F.; BOUCHET, S.; KLEIN, P. E.; KLEIN, R. R.; KILIAN, A.; WENZL, P.; XIA, L.; HALLORAN, K.; JORDAN, D. R. A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DART) markers. **BMC Plant Biology**, v. 9, n. 1, p. 1, 2009.

MANCINI, M. C.; LEITE, D. C.; PERECIN, D.; BIDÓIA, M. A. P.; XAVIER, M. A.; LANDELL, M. G. A.; PINTO, L. R. Characterization of the genetic variability of a sugarcane commercial cross through yield components and quality parameters. **Sugar Tech**, v. 14, n. 2, p. 119-125, 2012.

MANCINI, M. C. **MAPEAMENTO GENÉTICO E IDENTIFICAÇÃO DE MARCADORES AFLP, MICROSSATÉLITES GENÔMICOS E FUNCIONAIS ASSOCIADOS A PARÂMETROS AGROINDUSTRIAIS EM CANA-DE-AÇÚCAR**. 2010. 113p. Dissertação (Mestrado) - UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO", Jaboticabal.

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

MCINTYRE, C. L.; CASU, R. E.; DRENTH, J.; KNIGHT, D.; WHAN, V. A.; CROFT, B. J.; JORDAN, D. R.; MANNERS, J. M. Resistance gene analogues in sugarcane and sorghum and their association with quantitative trait loci for rust resistance. **Genome**, v. 48, n. 3, p. 391-400, 2005.

MCINTYRE, C. L.; GOODE, M. L.; CORDEIRO, G.; BUNDOCK, P.; ELIOTT, F.; HENRY, R. J.; CASU, R. E.; BONNETT, G. D.; AITKEN, K. S. Characterisation of alleles of the sucrose phosphate synthase gene family in sugarcane and their association with sugar-related traits. **Molecular Breeding**, v. 35, n. 3, 2015.

NUCCIO, M. L.; WU, J.; MOWERS, R.; ZHOU, H. P.; MEGHJI, M.; PRIMAVESI, L. F.; PAUL, M. J.; CHEN, X.; GAO, Y.; HAQUE, E.; BASU, S. S.; LAGRIMINI, L. M. Expression of trehalose-6-phosphate phosphatase in maize ears improves yield in well-watered and drought conditions. **Nat Biotechnol**, v. 33, n. 8, p. 862-9, 2015.

PALHARES, A. C.; RODRIGUES-MORAIS, T. B.; VAN SLUYS, M. A.; DOMINGUES, D. S.; MACCHERONI, W.; JORDAO, H.; SOUZA, A. P.; MARCONI, T. G.; MOLLINARI, M.; GAZAFFI, R.; GARCIA, A. A. F.; VIEIRA, M. L. C. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. **Bmc Genetics**, v. 13, 2012.

PATERSON, A. H.; BOWERS, J. E.; BRUGGMANN, R.; DUBCHAK, I.; GRIMWOOD, J.; GUNDLACH, H.; HABERER, G.; HELLSTEN, U.; MITROS, T.; POLIAKOV, A.; SCHMUTZ, J.; SPANNAGL, M.; TANG, H.; WANG, X.; WICKER, T.; BHARTI, A. K.; CHAPMAN, J.; FELTUS, F. A.; GOWIK, U.; GRIGORIEV, I. V.; LYONS, E.; MAHER, C. A.; MARTIS, M.; NARECHANIA, A.; OTILLAR, R. P.; PENNING, B. W.; SALAMOV, A. A.; WANG, Y.; ZHANG, L.; CARPITA, N. C.; FREELING, M.; GINGLE, A. R.; HASH, C. T.; KELLER, B.; KLEIN, P.; KRESOVICH, S.; MCCANN, M. C.; MING, R.; PETERSON, D. G.; MEHBOOB UR, R.; WARE, D.; WESTHOFF, P.; MAYER, K. F.; MESSING, J.; ROKHSAR, D. S. The *Sorghum bicolor* genome and the diversification of grasses. **Nature**, v. 457, n. 7229, p. 551-6, 2009.

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

PROVINCE, M. A. Sequential methods of analysis for genome scan. In: RAO, D. C.; PROVINCE, M. A. **Dissection of complex traits**. San Diego: Academic, p.578-583, 1999.

REMINGTON, D.L.; THORNSBERRY, J.M.; MATSUOKA, Y.; WILSON, L.M.; WHITT, S.R.; DOEBLEY, J.; KRESOVICH, S.; GOODMAN, M.M.; BUCKLER, E.S. Structure of linkage disequilibrium and phenotypic associations in the maize genome. **Proceedings of the National Academy of Science**, Washington, v. 98, n. 20, p. 11479-11484, 2001.

SANTOS, F. R. C.; PINTO, L. R.; CARLINI-GARCIA, L. A.; GAZAFFI, R.; MANCINI, M. C.; GONCALVES, B. S.; MEDEIROS, C. N. F.; PERECIN, D.; GARCIA, A. A. F.; SOUZA, A. P.; ZUCCHI, M. I. Marker-trait association and epistasis for brown rust resistance in sugarcane. **Euphytica**, v. 203, n. 3, p. 533-547, 2015.

SANTOS, F. R. **IDENTIFICAÇÃO DE MARCADORES MOLECULARES ASSOCIADOS À RESISTÊNCIA À FERRUGEM MARROM (*Puccinia melanocephala*) EM CANA-DE-AÇÚCAR**. 2013. 110p. Tese (Doutorado) – Instituto Agrônomo de Campinas, Campinas.

SAS, SAS; GUIDE, STAT User's. SAS Institute Inc. **Cary, NC, USA**, 2008.

SEKHWAL, M. K.; SHARMA, V.; SARIN, R. Identification of MFS proteins in sorghum using semantic similarity. **Theory Biosci**, v. 132, n. 2, p. 105-13, 2013.

SINGH, R. K.; JENA, S. N.; KHAN, S.; YADAV, S.; BANARJEE, N.; RAGHUVANSHI, S.; BHARDWAJ, V.; DATTAMAJUMDER, S. K.; KAPUR, R.; SOLOMON, S.; SWAPNA, M.; SRIVASTAVA, S.; TYAGI, A. K. Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane. **Gene**, v. 524, n. 2, p. 309-329, 2013a.

SINGH, R. K.; SINGH, R. B.; SINGH, S. P.; SHARMA, M. L. Identification of sugarcane microsatellites associated to sugar content in sugarcane and transferability to other cereal genomes. **Euphytica**, v. 182, n. 3, p. 335-354, 2011.

SINGH, R. K.; SINGH, S. P.; TIWARI, D. K.; SRIVASTAVA, S.; SINGH, S. B.; SHARMA, M. L.; SINGH, R.; MOHAPATRA, T.; SINGH, N. K. Genetic mapping and QTL analysis for sugar yield-related traits in sugarcane. **Euphytica**, v. 191, n. 3, p. 333-353, 2013b.

SMITH, J. J.; KUMP, D. K.; WALKER, J. A.; PARICHY, D. M.; VOSS, S. R. A comprehensive expressed sequence tag linkage map for tiger salamander and Mexican axolotl: enabling gene mapping and comparative genomics in Ambystoma. **Genetics**, v. 171, n. 3, p. 1161-71, 2005.

WANG, J.; ROE, B.; MACMIL, S.; YU, Q.; MURRAY, J. E.; TANG, H.; CHEN, C.; NAJAR, F.; WILEY, G.; BOWERS, J.; VAN SLUYS, M. A.; ROKHSAR, D. S.; HUDSON, M. E.; MOOSE, S. P.; PATERSON, A. H.; MING, R. Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. **BMC Genomics**, v. 11, p. 261, 2010.

WANG, M. L.; WANG, M. L.; BARKLEY, N. A.; YU, J. K.; DEAN, R. E.; NEWMAN, M. L.; SORRELLS, M. E.; PEDERSON, G. A. Transfer of simple sequence repeat (SSR) markers from major cereal crops to minor grass species for germplasm characterization and evaluation. **Plant Genetic Resources: characterization and utilization**, v. 3, n. 1, p. 45-57, 2005.

XIE, D. W.; WANG, X. N.; FU, L. S.; SUN, J.; ZHENG, W.; LI, Z. F. Identification of the trehalose-6-phosphate synthase gene family in winter wheat and expression analysis under conditions of freezing stress. **Journal of genetics**, v. 94, n. 1, p. 55-65, 2015.

YU, H.; CONG, L.; ZHU, Z.; WANG, C.; ZOU, J.; TAO, C.; SHI, Z.; LU, X. Identification of differentially expressed microRNA in the stems and leaves during sugar accumulation in sweet sorghum. **Gene**, v. 571, n. 2, p. 221-230, 2015.