



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
PRODUÇÃO VEGETAL E BIOPROCESSOS ASSOCIADOS

MAPEAMENTO DE QTL'S PARA COMPONENTES ASSOCIADOS A BIOMASSA
EM CANA-DE-AÇÚCAR

MALÚ DEL SOCORRO CARRANZA CRUZ

Araras

2023



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
PRODUÇÃO VEGETAL E BIOPROCESSOS ASSOCIADOS

MAPEAMENTO DE QTL'S PARA COMPONENTES ASSOCIADOS A BIOMASSA
EM CANA-DE-AÇÚCAR

MALÚ DEL SOCORRO CARRANZA CRUZ

ORIENTADOR: PROF. DRA. MONALISA SAMPAIO CARNEIRO
CO-ORIENTADOR: PROF. DR. THIAGO W. ALMEIDA BALSALOBRE

Dissertação apresentada ao Programa de Pós-Graduação em Produção Vegetal e Bioprocessos Associados como requisito parcial à obtenção do título de MESTRE EM PRODUÇÃO VEGETAL E BIOPROCESSOS ASSOCIADOS

Araras

2023

Cruz., Malú Del Socorro Carranza

Mapeamento de QTL's para componentes associados a biomassa em cana-de-açúcar / Malú Del Socorro Carranza Cruz.--2023.
48f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus Araras, Araras

Orientador (a): Monalisa Sampaio Carneiro

Banca Examinadora: Monalisa Sampaio Carneiro,
Rodrigo Gazaffi, André Boldrin Beltrame

Bibliografia

1. Marcadores moleculares. 2. Melhoramento genético vegetal. 3. Biotecnologia Vegetal. I. Cruz., Malú Del Socorro Carranza. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Maria Helena Sachi do Amaral - CRB/8
7083



UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Agrárias
Programa de Pós-Graduação em Produção Vegetal e Bioprocessos Associados

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Malú Del Socorro Carranza Cruz, realizada em 09/08/2023.

Comissão Julgadora:

Profa. Dra. Monalisa Sampaio Carneiro (UFSCar)

Prof. Dr. Rodrigo Gazaffi (UFSCar)

Prof. Dr. André Boldrin Beltrame (EPAGRI)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Produção Vegetal e Bioprocessos Associados.

AGRADECIMENTOS

Primeiramente quero agradecer a Deus pela vida, por me proporcionar tudo que tenho e me permitir estar aqui apesar das adversidades.

À Rosa e Jony, meus pais, por me mostrarem o imenso amor que me dão, pelo apoio incondicional, incentivo para continuar e não vacilar em alcançar meus objetivos e por sempre me motivarem a ser melhor, acima de tudo por me derem a liberdade de voar com minhas próprias asas mesmo que a distância seja longa.

À Katherine e Aylín, minhas irmãs, por estarem comigo durante todo o processo de desenvolvimento deste programa, por serem minha força e a razão pela qual não paro, por deixar nelas o exemplo de superação e de que tudo é possível se você correr riscos.

Ao David, por estar comigo neste processo, pelo apoio constante em todos os meus projetos e por estar sempre comigo, mesmo nos momentos difíceis.

À minha orientadora, Prof. À Dra. Monalisa S. Carneiro, pela generosidade e apoio, bem como pelos conselhos, compreensão, por me dar a oportunidade de trabalhar com ela e, acima de tudo, pela paciência durante todo o mestrado, apesar de todas as adversidades.

Ao meu coorientador, Prof. Dr. Thiago Almeida Balsalobre, por todo apoio, conselhos, grande orientação, paciência e muitos dias de reuniões. Por ser a maior referência pessoal.

Ao meu amigo Joe, por seu apoio e ajuda, por sua orientação no desenvolvimento do mestrado e por se mostrar um amigo fiel.

Ao programa de pós-graduação em Produção Vegetal e Bioprocessos Associados (PPGPVBA), à coordenação pelo auxílio prestado, às secretárias pelo suporte, aos professores por compartilhar todo o conhecimento em sala de aula e graças a isso posso aplicar no trabalho e melhorar minha qualidade profissional.

Agradeço também aos membros da banca de qualificação que me ajudaram a lapidar o trabalho para este momento final.

Obrigado a todos, vida plena e sigam sempre em frente sem medo!

SUMÁRIO

ÍNDICE DE TABELAS	viii
ÍNDICE DE FIGURAS	ix
RESUMO.....	x
ABSTRACT.....	xi
1 INTRODUÇÃO	1
2 OBJETIVOS.....	3
3 REVISÃO DA LITERATURA	4
3.1 Cultura da cana-de-açúcar	4
3.2 A cana-de-açúcar no Perú	5
3.3 Complexidade Genética e Melhoramento da Cana-de-Açúcar	6
3.4 Marcadores moleculares.....	8
3.5 Mapas genéticos.....	11
3.6 <i>Mapeamento de QTL</i>	12
4 MATERIAIS E MÉTODOS	13
4.1 Material vegetal	13
4.2 Marcadores baseados em gbs.....	14
4.3 Estimativa de dosagem alélica	14
4.4 Marcadores baseados em gel.....	15
4.5 Construção do mapa genético de ligação.....	15
4.6 Mapeamento de QTL.....	17
5 RESULTADOS E DISCUSSÃO	18
5.1 Descoberta do polimorfismo de marcadores SNPs	18
5.2 Estimativa da ploidia e dosagem alélica	18
5.3 Genotipagem de marcadores baseados em gel	19
5.4 Mapa genético	19
5.5 Mapeamento QTL.....	19

6 CONCLUSÕES	41
7 LITERATURA CITADA	42

ÍNDICE DE TABELAS

Tabela 1. – Espécies que compõem o “complexo Saccharum” e que deram origem aos híbridos modernos de cana-de-açúcar	6
Tabela 2. Número de marcadores gerados após GBS na população de cana-de-açúcar	18
Tabela 3. Marcadores em dose única baseados em gel e baseados em GBS avaliados na progênie do cruzamento entre as variedades de cana-de-açúcar SP81-3250 e RB925345.	21
Tabela 4. Distribuição dos diferentes tipos de marcadores mapeados de acordo com o tipo de cruzamento.	22
Tabela 5. QTLs mapeados para características SH e SD aplicando um modelo CIM.	35

ÍNDICE DE FIGURAS

- Figura 1.** Teste de segregação mendeliana com correção de múltiplos testes de Bonferroni para os marcadores em dose-única gerados a partir de GBS e bom base em gel.23
- Figura 2.** Mapa genético integrado obtido a partir da genotipagem com marcadores baseados em gel e baseados em GBS da progênie do cruzamento entre as variedades de cana-de-açúcar SP81-3250 e RB925345.....34
- Figura 3.** Mapeamento de intervalo composto (CIM) para diâmetro da planta (SD, mm) e altura da planta da população (SH) da SP81-3250 e RB925345 população F1. À direita, temos um gráfico que indica os limites de LOD obtidos após testes de permutação. As porções dentro do quadrado em vermelho nos grupos de ligação (à esquerda) mostram as posições dos QTLs.....36

MAPEAMENTO DE QTL'S PARA COMPONENTES ASSOCIADOS A BIOMASSA EM CANA-DE-AÇÚCAR

Autor: MALÚ DEL SOCORRO CARRANZA CRUZ

Orientador: Prof. Dr. DRA. MONALISA SAMPAIO CARNEIRO

Co-orientador: Dr. THIAGO W. ALMEIDA BALSALOBRE

RESUMO

A cana-de-açúcar (*Saccharum spp.*) é uma das principais culturas agroindustriais do mundo, cerca de 124 países no mundo produzem açúcar. Na campanha 2021/2022, em nível global, observou-se um aumento de 0,53%, produzindo 181 milhões de toneladas. E, nos últimos 20 anos, no Peru, a cana-de-açúcar apresentou um crescimento visível, sendo 2006 o ano em que o crescimento começou a acelerar. E, em 2021, a produção subiu para 12.309 mil toneladas de cana, das quais 79% foram destinadas à fabricação de açúcar. Sua importância se fundamenta na utilização para a produção de açúcar e etanol. A cana-de-açúcar caracteriza-se por ser autopoliploide com alto nível de ploidia e aneuploidia em seu gigagenoma. O objetivo deste projeto foi a identificação e mapeamento de QTL (caracteres de herança quantitativa) das características agroindustriais da cana-de-açúcar, construindo um mapa genético para identificar regiões genômicas associados à altura e diâmetro das plantas. A população de estudo foi desenvolvida pelo Programa de Melhoramento Genético da Cana-de-açúcar (PMGCA) da Universidade Federal de São Carlos (UFSCar). Ao todo, 238 genótipos, derivados do cruzamento entre SP81-3250 e RB925345, foram genotipados para marcadores SNP a partir de GBS (Genotyping by Sequencing) e avaliados à campo para altura e diâmetros de plantas. A partir do alinhamento das sequências do GBS com os genomas de *Sorghum* e *Sacharum spontaneum* foram identificadas cerca de 152.000 e 250.000 variantes, respectivamente. Deste total, após todas as etapas de filtragem e análise de segregação, foram levados para construção do mapa genético 15.795 marcadores SNP. Em adição, 285 fragmentos de marcadores baseados em gel também foram aptos para construção do mapa genético. Um total de 2016 marcadores fizeram parte do mapa de ligação, composto por 221 grupos de ligação, com tamanho de 5.282 cM e densidade de 3,06 cM. A análise de mapeamento identificou dois QTL (caracteres de herança quantitativa), um para altura e outro para diâmetro plantas. Para altura de plantas o QTL está localizado no grupo de ligação 11, apresentou segregação 1:1 e explicou 6,38% da variação fenotípica observada. Enquanto que o QTL para diâmetro de plantas está localizado no grupo de ligação 53, apresentou segregação 1:2:1 e explicou 4,59% da variação fenotípica observada.

Palavras-chave: *Saccharum spp.*, poliploidia, mapa genético, QTL.

QTL MAPPING FOR COMPONENTS ASSOCIATED WITH BIOMASS IN SUGARCANE

Author: MALÚ DEL SOCORRO CARRANZA CRUZ

Adviser: Prof. Dr. MONALISA SAMPAIO CARNEIRO

Co-adviser: Dr. THIAGO W. ALMEIDA BALSALOBRE

ABSTRACT

Sugarcane (*Saccharum* spp.) is one of the world's main agro-industrial crops, with about 124 countries producing sugar. During the 2021/2022 season, a global increase of 0.53% was observed, producing 181 million tons. Over the past 20 years in Peru, sugarcane has shown visible growth, with 2006 marking the year when growth began to accelerate. In 2021, production reached 12.309 million tons of cane, of which 79% were used for sugar production. Its importance lies in its use for sugar and ethanol production. Sugarcane is characterized as an autoployploid with a high level of ploidy and aneuploidy in its genome. The objective of this project was the identification and mapping of QTLs (quantitative trait loci) for agro-industrial traits in sugarcane, constructing a genetic map to identify genomic regions associated with plant height and diameter. The study population was developed by the Sugarcane Breeding Program at the Federal University of São Carlos (UFSCar). A total of 238 genotypes, derived from the cross between SP81-3250 and RB925345, were genotyped using SNP markers obtained through GBS (Genotyping by Sequencing) and evaluated in the field for plant height and diameter. By aligning the GBS sequences with the genomes of *Sorghum* and *Saccharum spontaneum*, approximately 152,000 and 250,000 variants, respectively, were identified. From this total, after all filtering steps and segregation analysis, 15,795 SNP markers were used to construct the genetic map. In addition, 285 gel-based marker fragments were also suitable for genetic map construction. A total of 2,016 markers were included in the linkage map, which consisted of 221 linkage groups, spanning 5,282 cM with a density of 3.06 cM. The mapping analysis identified two QTLs, one for plant height and another for plant diameter. For plant height, the QTL was located in linkage group 11, exhibited a 1:1 segregation pattern, and explained 6.38% of the observed phenotypic variation. For plant diameter, the QTL was located in linkage group 53, exhibited a 1:2:1 segregation pattern, and explained 4.59% of the observed phenotypic variation.

Keywords: *Saccharum* sp., polyploidy, genetic map, QTL.

1 INTRODUÇÃO

A cultura da cana-de-açúcar pertence à família Poaceae e a tribo Andropogoneae, juntamente com o milho e sorgo (DANIELS; ROACH, 1987-; WELKER *et al.*, 2015). As cultivares modernas de cana-de-açúcar (*Saccharum spp*) são resultantes dos cruzamentos interespecíficos, de uma espécie selvagem (*S. spontaneum*) com uma domesticada (*S. officinarum*), que têm um intervalo de 100 a 130 cromossomos e devido a eventos de hibridização entre diferentes espécies, as cultivares modernas de cana-de-açúcar adquiriram os genomas combinados de seus dois ancestrais poliploides. Como consequência disso, a cultura da cana-de-açúcar apresenta atualmente múltiplas cópias de cromossomos homólogos herdados de ambos os ancestrais o que promove nível variado de ploidia e complexidade genética (DANIELS; ROACH, 1987-; YANG *et al.*, 2018a).

Considerada uma das principais culturas agroindustriais do mundo, a cana-de-açúcar é predominantemente cultivada em países tropicais e subtropicais (CHASTEL, 1994) sendo o Brasil o maior produtor do mundo com 8.598.440 ha em 2019 (FAO, 2019). No Peru, segundo a FAO, a partir de 2018 foram obtidos os melhores rendimentos agrícolas com 121,8 toneladas/ha (MINISTERIO DE AGRICULTURA Y RIEGO, 2020), superando a produtividade brasileira que foi de 73 toneladas/ha. A cana-de-açúcar é utilizada para produção de açúcar, de bioetanol e de bioenergia a partir do resíduo celulósico (COMPANHIA NACIONAL DE ABASTECIMENTO – CONAB, 2023; FRETES; MARTÍNEZ, 2011; YOU *et al.*, 2020) .

Os programas de melhoramento genético têm atuado para liberar novas cultivares que superem os níveis de produtividade das cultivares atuais. Este trabalho pode levar, em média, 15 anos, devido à complexidade genética dos híbridos modernos, o que necessita de alto investimento financeiro (CURSI *et al.*, 2021). Diante deste cenário, novas tecnologias e ferramentas moleculares podem ser utilizadas para contribuir com o avanço desta cultura. Por exemplo, o uso de marcadores moleculares e o posterior mapeamento genético ajudam a encontrar regiões genômicas que controlam as principais características de interesse agrônomo, no entanto melhorar as características agrônomicas e de qualidade da cana-de-açúcar é um desafio devido à sua complexidade genética e natureza poliploide. (BALSALOBRE *et al.*, 2017).

O mapeamento de QTL (caracteres de herança quantitativa) surgiu como uma ferramenta fundamental na pesquisa da cana-de-açúcar, com a finalidade de compreender a base genética de características agrícolas e agroindustriais de grande relevância econômica. Permitindo identificar e localizar regiões específicas do genoma que influenciam características desejáveis como rendimento, teor de açúcar e resistência a doenças, entre outras. Essas regiões estão associadas a genes particulares que contribuem para a variação fenotípica observada nessas características (ANONI *et al.*, 2013; YONEMARU; KASUGA; KAWAHIGASHI, 2022).

Muitos estudos foram realizados e vários mapas genéticos foram gerados em cultivares modernas de cana-de-açúcar, no entanto, nenhum dos mapas atuais oferece um nível de detalhamento completo, o que dificulta a determinação precisa da estrutura genômica das cultivares modernas, principalmente em relação ao número de cópias de cada cromossomo básico (PIPERIDIS; D'HONT, 2020).

O objetivo principal do presente estudo foi mapear QTL's para componentes associados a biomassa em cana-de-açúcar. Para isso, foram empregados métodos de genotipagem e fenotipagem que fornecem dados confiáveis e permitem identificar e localizar os QTL's no mapa genético construído. Os resultados deste estudo têm o potencial de subsidiar o desenvolvimento de estratégias mais eficientes para o melhoramento genético da cana-de-açúcar, possibilitando a obtenção de cultivares superiores. Além disso, contribuem para o avanço no conhecimento sobre a estrutura genética da espécie, particularmente em relação às características associadas a biomassa, fornecendo insights valiosos para futuras pesquisas e aplicações biotecnológicas.

2 OBJETIVOS

Objetivo geral

Construir um mapa genético e identificar regiões genômicas associadas com caracteres de importância econômica em cana-de-açúcar.

Objetivo específico

- Construir um mapa genético de ligação usando marcadores moleculares baseados em sequência e baseados em gel na população do cruzamento comercial SP81-3250 x RB925345.
- Identificar e localizar QTL's para características de interesse econômico em cana-de-açúcar relacionadas com produção de biomassa.

3 REVISÃO DA LITERATURA

3.1 Cultura da cana-de-açúcar

A cana-de-açúcar (*Saccharum* spp.) pertence à família Poaceae e a tribo Andropogoneae, juntamente com o milho e sorgo (CURSI *et al.*, 2022; WELKER *et al.*, 2015). Possui uma origem incerta e existem muitas teorias que datam seu aparecimento há aproximadamente 8.000 anos na Nova Guiné e ilhas vizinhas. (POMPIDOR *et al.*, 2021; RIVERA, 2008). A classificação taxonômica foi feita por Linnaeus, em 1753, como *S. officinarum* (BENDEZÚ, 2014). Desde então, vários autores como Jeswiet e Grassl (1969), reconheceram seis espécies pertencentes ao gênero *Saccharum*: *S. officinarum*, *S. barberi*, *S. sinense*, *S. robustum*, *S. spontaneum* e *S. edule* (D'HONT *et al.*, 1996; DANIELS; ROACH, 1987; SREENIVASAN; AHLOOWALIA; HEINZ, 1987).

De acordo com a origem e distribuição, Daniel e Roach (1987) mencionam que existe uma relação entre os gêneros *Saccharum*, *Erianthus*, *Miscanthus*, *Sclerostachya* e *Narenga*. Esses gêneros possuíam características compartilhadas atribuindo-lhes altos níveis de variação e ploidia. Por causa disso, Mukherjee em 1957 propôs o nome "complexo *Saccharum*" para descrever esse grupo (CURSI *et al.*, 2022; DANIELS; ROACH, 1987; SREENIVASAN; AHLOOWALIA; HEINZ, 1987).

Cultivares comerciais têm um número de cromossomos de 100 a 130, cerca de 70 a 80% dos cromossomos provenientes de *S. officinarum*, cerca de 10-20% de *S. spontaneum* e 5-17% resultantes da recombinação entre ambas as espécies. *S. officinarum* é uma espécie octoploide ($2n = 8x = 80$) e não possui genitores diploides ou tetraploides disponíveis. Em contraste, *S. spontaneum* ($2n = 4x-16x = 32-128$), que exibe notável tolerância ao estresse em cultivares de cana-de-açúcar, apresenta ampla variabilidade no número de cromossomos (YANG *et al.*, 2018a)

Atualmente o Brasil é o maior produtor de cana do mundo com 8.598.440 ha, em 2019, seguido pela Índia (4.400.000 ha), China (1.630.520 ha) e Paquistão (1.029.000 ha) (FAO, 2019). Culturas anuais, como soja, milho, mandioca, algodão, e cana-de-açúcar que é semiperene, representam 80% das áreas cultivadas no Brasil.

A importância da cana-de-açúcar está nos seus diferentes usos. A produção mundial de cana-de-açúcar é obtida principalmente da cana-de-açúcar, cerca de 80%. Em adição, pode ser obtido também o bioetanol, um biocombustível, que é

considerado energia renovável e sustentável (YOU *et al.*, 2020). A cana-de-açúcar também pode ser utilizada para produção de cachaça, e o bagaço pode ser utilizado para produzir etanol de segunda geração e também energia elétrica (CURSI *et al.*, 2021; UMEDA *et al.*, 2021; YOU *et al.*, 2020). No Brasil, o uso do etanol busca substituir uma fatia do mercado de combustíveis com origem no petróleo, como a gasolina; e conseqüentemente reduzir o impacto dos combustíveis fósseis no meio ambiente (CORONEL, 2016). Desde 2006, o bioetanol é declarado pelas autoridades brasileiras como parte do setor estratégico de energia. (MASSOT, 2008).

3.2 A cana-de-açúcar no Perú

O início da indústria açucareira no Peru remonta ao século XV; onde pousou pela primeira vez na cidade de Trujillo. E cresceria gradativamente durante os séculos XVI e XIX, até que os grandes latifúndios se estabilizassem. Nessa época, o Peru era considerado um dos principais produtores de açúcar do mundo, sendo Trujillo e Lambayeque as principais regiões produtoras. Posteriormente, a considerável quantia de 11 milhões de hectares foi concedido a comunidades e cooperativas, que com o tempo faliram, e no início dos anos noventa passariam para o setor privado. Regimes especiais foram criados para o setor agrícola, já que quase meio milhão de peruanos dependem do desenvolvimento deste setor, para melhorar as condições das atividades agrícolas, bem como do capital humano, permitindo o desenvolvimento de toda a indústria agrícola e especialmente da indústria açucareira (PERUCAÑA, 2019).

O Peru a partir de 2018, obteve os melhores rendimentos agrícolas com 121,8 toneladas/ha (MINISTERIO DE AGRICULTURA Y RIEGO, 2020), superando a produtividade do Brasil, que é de aproximadamente 35 toneladas/ha. No Peru, a produção vem crescendo em 1,8% nos últimos anos, sendo as regiões do norte do país as que mais contribuem para o aumento da produção (COMPANHIA NACIONAL DE ABASTECIMIENTO - (CONAB), 2022; CORONEL, 2016). A cana-de-açúcar e sua importância econômica se baseia principalmente na produção do açúcar, por outro lado, também na produção de destilados de caldo de cana, conhecidos como “cañazo” e “primera”, e até mesmo na compostagem (DARGENT, 2017). Além disso, surgiu o interesse pelo uso e produção de combustível (GOÑI; ROJAS, 2016). O conteúdo de bioetanol na mistura de gasolina e etanol, no Peru, é 7,8% do volume. Nos anos de

2016 e 2017, o consumo de bioetanol representou 0,6% do consumo total de combustível, o que gerou uma pequena diversificação no mercado (PACHECO, 2019).

3.3 Complexidade Genética e Melhoramento da Cana-de-Açúcar

Os híbridos interespecíficos de cana-de-açúcar, oriundos do cruzamento entre as espécies do complexo *Saccharum* (Tabela 1), apresentam elevada complexidade genética como, por exemplo, número variado de cromossomos (100 a 130), tamanho do genoma de aproximadamente 10 Gb e presença de aneuploidia (D'HONT, 2005; GRIVET; ARRUDA, 2002). Esta complexidade é derivada desde os primeiros cruzamentos entre a espécie domesticada *S.officinarum* L. ($2n = 80$) e um *S. spontaneum* selvagem, seguido de novos cruzamentos com a espécie domesticada *S. officinarum* ($2n = 40-120$) (GUTIERREZ *et al.*, 2018).

Tabela 1. Espécies que compõem o “complexo Saccharum” e que deram origem aos híbridos modernos de cana-de-açúcar

Espécies	Descrição	Teor de açúcar	Número de cromossomos
<i>S. spontaneum</i> L.	Espécies selvagens	Muito baixo - baixo	$2n=40-128$
<i>S. robustum</i> Brandes e Jeswiet ex Grassl	Espécies selvagens	Muito baixo	$2n=60-200$
<i>S. officinarum</i> L.	Cana nobre	Alto	$2n=80$
<i>S. barberi</i> Jeswiet	Híbrido antigo	Baixo	$2n=111-120$
<i>S. sinense</i> Roxb.	Híbrido antigo	Baixo	$2n=80-124$
<i>S. edule</i> Hassk.	Espécies cultivadas	Baixo	$2n=60-80$ com aneuploidia

Fonte: Cursi *et al.*, 2022 *apud* Daniels e Roach (1987).

ZHANG *et al.*, (2018) analisou o genoma do *S. spontaneum* e comparou haplótipos homólogos, onde encontra milhões de polimorfismos, inserções e deleções, além de variações estruturais, demonstrando alta heterozigosidade em seu

genoma. A partir disso, ele desenvolveu um quadro de análise para estudar as semelhanças entre pares de genes parálogos, mas não conseguiu identificar uma associação clara entre eles na maioria das séries homólogas. No entanto, o estudo revelou eventos de fusão e fissão cromossômica em ancestrais cromossômicos de sorgo que ocorreram antes das rodadas de duplicação no gênero *Saccharum*; da mesma forma, se observa uma redução cromossômica na espécie *Miscanthus* causada pela fusão de um conjunto de cromossomos homólogos. Além disso, ele encontrou inversões cromossômicas em *S. spontaneum* que ocorreram antes das duplicações e em um ancestral comum com o sorgo. Sugerindo que as duas rodadas de duplicação do genoma de *S. spontaneum* são autoploidizações que ocorreram em um curto período de tempo. Por sua vez, ele menciona que essa espécie contribuiu com genes de resistência a doenças em cultivares híbridas modernas de cana-de-açúcar e identificou cerca de 300 sequências que provavelmente codificam para proteínas do sítio de ligação de nucleotídeos, já que são encontradas em quatro cromossomos específicos do *S. spontaneum* (SsChr02, SsChr05, SsChr06 e SsChr07).

Apesar da complexidade genética, o melhoramento genético da cana é concentrado em avaliações fenotípicas, de forma que se busca aumentar rendimento de açúcar ou biomassa por meio da seleção de indivíduos superiores. O ambiente e as características de interesse são a base para o desenvolvimento de estratégias de melhoramento, principalmente para seleção fenotípica (COX; HOGARTH; B, 1994). WELHAM *et al.*, (2010) menciona que nos programas de melhoramento são avaliados clones de cana-de-açúcar durante várias safras e para vários caracteres, pois clones com maior número de alelos favoráveis para características como produtividade, resistência a doenças, tolerância a pragas e estresse abiótico devem ser selecionados.

Segundo Cursi *et al.*, (2021) os programas de melhoramento da cana-de-açúcar possuem três etapas: a) geração da variabilidade genética (cruzamentos); b) seleção (através de experimentos para avaliar as características de interesse) e; c) clonagem (multiplicação dos clones selecionados). No início são realizados experimentos com poucas repetições e com pequenas parcelas, e nas fases finais existem maior número de repetições e locais de experimentação, sendo que todo processo, desde os cruzamentos até obtenção de uma possível nova cultivar pode

levar, em média, de 10 a 15 anos (CURSI *et al.*, 2021). O sucesso de um programa de melhoramento depende do conhecimento e compreensão da diversidade genética encontrada nos bancos de germoplasma, a fim de direcionar os cruzamentos entre genótipos. Entre as características consideradas importantes para os bancos de germoplasma estão a produção em toneladas de cana por hectare, o teor de açúcar, a altura e o diâmetro das plantas. Em adição, essa diversidade genética também pode ser avaliada por meio de marcadores moleculares (ZATTI, 2016).

3.4 Marcadores moleculares

Os marcadores moleculares são ferramentas que tem como base diferentes alterações cromossômicas, aqui são investigadas por polimorfismo entre as sequências dos indivíduos. Tem como características ideais ser codominante, ser reproduzível e, acima de tudo, detectar um maior nível de polimorfismo. Há diversos tipos de marcadores moleculares para análise do genoma da cana-de-açúcar como, por exemplo, AFLP (Amplified Fragment Length Polymorphic), Microsatélites ou SSR (Simple Sequence Repeat), e SNP (Single Nucleotid Polymorphism), DArT (Diversity Array Technology) (AITKEN, 2022; MASUELLI, 1999); no entanto, os principais e mais utilizados são o RFLP que tem como características principais a co-dominância, alta reprodutibilidade, nível médio de polimorfismo, pode ser sequenciado, não requer PCR; por outro lado, requer uma grande quantidade de DNA de boa qualidade, alta abundância do genoma, seus custos são elevados e sua visualização é por meio de radioatividade. Por essas razões, essa técnica não é mais realizada. O marcador RAPD é um marcador dominante, com alta reprodutibilidade, seu nível de polimorfismo é muito alto, requer uma quantidade média de DNA, tem baixo custo de produção; por outro lado, requer DNA de alta qualidade, a abundância do genoma é muito alta, requer a realização de PCR e sua visualização é feita em gel de agarose, e requer pouca quantidade de DNA. O marcador AFLP também é um marcador dominante, com reprodutibilidade intermediária, seu nível de polimorfismo é alto, requer pouca quantidade de DNA, tem alto custo de produção; por outro lado, requer DNA de alta qualidade, a abundância do genoma é muito alta, requer a realização de PCR para a obtenção de uma boa quantidade de genoma, e sua visualização é feita em gel de agarose. Não é mais usado atualmente. O marcador ISSR também é um marcador dominante, com reprodutibilidade média-alta, seu nível de polimorfismo é

alto, não requer DNA de boa qualidade nem requer muita quantidade de DNA, tem alto custo; por outro lado, a abundância do genoma é média, requer a realização de PCR para a obtenção de uma boa quantidade de genoma, e sua visualização é feita em gel de agarose. Continua sendo utilizado nos estudos atuais. O marcador SSR microssatélite é um marcador co-dominante, com alto nível de reprodutibilidade e alto nível de polimorfismo. Requer DNA de alta qualidade, mas não grandes quantidades, apresenta genoma abundante e alto custo de produção, adequado para sequenciamento, requer a realização de PCR e sua visualização é em gel de agarose. Atualmente continua sendo utilizado. Os marcadores SNP são caracterizados por serem co-dominantes, apresentam alta reprodutibilidade, alto nível de polimorfismo, requerem DNA de alta qualidade, mas baixa quantidade. A abundância do genoma é muito alta, seus custos são variáveis, pode-se realizar o sequenciamento e sua visualização é através de SNPs. Os marcadores DArT, marcador dominante, de alta reprodutibilidade, com alto nível de polimorfismo, requerem pouca quantidade de DNA, mas alta qualidade, abundância de genoma muito alta, são muito mais econômicos que os marcadores anteriores, não necessitam de realização de PCR e sua visualização é por meio de microarray. Os retrotransposons são de caráter dominante, com um alto nível de reprodutibilidade e um nível alto de polimorfismo. Requerem DNA de alta qualidade em quantidades baixas, apresentam genoma abundante e custo de produção muito baixo, não são adequados para sequenciamento, requerem a realização de PCR e sua visualização é em gel de agarose. Atualmente continuam sendo utilizados (NADEEM *et al.*, 2018). Também existem os marcadores funcionais, os quais estão distribuídos aleatoriamente por todo o genoma e refletem variações genéticas relevantes no setor agrícola; uma das técnicas para detectar essas variações são os marcadores TRAP ("Target Region Amplified Polymorphism"). Esses marcadores são baseados na técnica de PCR e são especificamente projetados para detectar polimorfismos dentro dos genes. Para sua geração, é utilizado um iniciador sentido projetado a partir de um gene ou sequência de expressão (EST), e um iniciador antisense projetado com um motivo AT ou GC. Dessa forma, os iniciadores se ligam aos íntrons e éxons, respectivamente (PERERA *et al.*, 2020).

São utilizados diversos tipos de marcadores moleculares para identificar características de interesse em várias pesquisas. Esses marcadores podem ser utilizados para estudos de análise da diversidade genética de bancos de germoplasma, seleção assistida por marcadores (SAM), seleção genômica ampla (GWAS), construção de mapas genéticos de ligação e mapeamento de QTLs, dentro outros (MASUELLI, 1999; YOU *et al.*, 2020).

Como resultado do uso de 5 combinações diferentes de iniciadores, Aitken (2006) relatou que foram obtidos 657 marcadores clones de *S. officinarum*, cada um exibindo um padrão único de AFLP (Amplified Fragment Length Polymorphic), permitindo a detecção de um número significativo de marcadores em ambos os grupos de indivíduos. Na coleção de cultivares, houve uma ampla variação no número de marcadores e apenas 6,5% dos marcadores eram monomórficos.

Utilizando um mapa genético de polimorfismo de comprimento de fragmentos de restrição (RFLP) desenvolvido com 77 amostras da cultivar R570, foi encontrado um importante locus de resistência à ferrugem marrom da cana-de-açúcar, conhecido como Bru1 (DAUGROIS *et al.*, 1996). No ano de 2005, foi construído um mapa genético com 192 progênies F1 da cana-de-açúcar contendo 852 marcadores, incluindo RFLPs, SSRs e AFLPs, com o objetivo de identificar três análogos de genes de resistência (RGAs) associados à resistência à ferrugem marrom (MCINTYRE *et al.*, 2005). Outros genes de resistência à ferrugem marrom da cana-de-açúcar também foram identificados, como o Bru2, um QTL para a resistência à doença das manchas amarelas, e o gene Ryl1, para a resistência à doença das manchas amarelas (COSTET *et al.*, 2012). Um estudo mais recente, utilizando a tecnologia DaRT (Diversity Array Technology), foi realizado no México para estimar a diversidade genética de 93 clones de cana-de-açúcar. Essa técnica permitiu a redução eficiente do genoma e forneceu informações sobre a distância genética entre os clones (SILVA-ROJAS; AGUIRRE RAYO; CERVANTES, 2017). Em 2018, foram localizados dois novos QTLs importantes para a resistência à ferrugem marrom da cana-de-açúcar (YANG *et al.*, 2018b). Em 2020, um estudo foi realizado usando marcadores TRAP, que foram utilizados para avaliar a similaridade com uma variedade de cana-de-açúcar não transgênica de diferentes eventos transgênicos, com o objetivo de caracterizar os eventos transgênicos das variedades LCP 85-384, TUCCP 77-42, TUC 95-10 e TUC 03-12 nas primeiras etapas do processo de avaliação (PERERA *et al.*,

2020). O mais recente estudo utilizou SNPs e análise de segregantes em massa para identificar marcadores do vírus da folha amarela e da podridão vermelha usando a estratégia de sequenciamento baseada na análise de segregação em massa (polyBSA-seq) e foram descobertos três marcadores associados à resistência e um marcador relacionado à suscetibilidade à mancha foliar da cana-de-açúcar (WANG *et al.*, 2022a).

3.5 Mapas genéticos

A técnica de mapear ligações genéticas é baseada na frequência de recombinação cromossômica durante a meiose para determinar a ordem dos marcadores genéticos em um cromossomo. Os efeitos da autoploidia no comportamento cromossômico durante diferentes estágios da meiose em clones autotetraplóides, autooctoplóides e autodecaplóides foram examinados usando hibridização fluorescente *in situ* (FISH), verificou-se que os clones autoploidios exibem um padrão estável de acasalamento diploidizado, bem como acasalamento preferencial entre cópias do cromossomo 8 homólogo no clone autotetraploide. Embora não tenham sido encontradas diferenças nas sequências de DNA entre os homólogos do cromossomo 8 e os demais cromossomos, conclui-se que existem mecanismos desconhecidos que influenciam o acasalamento estável de *S. spontaneum* (ZHANG *et al.*, 2023).

Para obter um mapa de ligação preciso, é crucial ter uma população de mapeamento que apresente um número considerável de eventos de recombinação envolvendo diferentes marcadores genéticos. Considerasse uma ferramenta que contribui para o conhecimento do genoma da cana-de-açúcar, estes mapas são como uma representação dos cromossomos da espécie avaliada, sendo que os marcadores moleculares são posicionados nestes cromossomos de acordo com a distância entre eles, medida através de uma fração de recombinação (WU *et al.*, 1992; YANG *et al.*, 2018a).

Em 2014, o método de *pseudotestcross* foi utilizado para criar um mapa genético da variedade Q165 de cana-de-açúcar. Um total de 227 descendentes foram gerados a partir de um cruzamento entre Q165 e IJ76-514. Foram utilizados 2467 marcadores de dose única (SD) para o mapeamento, dos quais 967 foram usados

para construir o mapa. Foi observada distorção na segregação de alguns marcadores SD, e marcadores sesgados foram incluídos no mapa. Os marcadores SD foram agrupados em 160 grupos de ligação, com 2267 marcadores ligados e 200 marcadores não ligados (AITKEN *et al.*, 2014).

Balsalobre em 2017 utilizou marcadores moleculares em dose única (SSR, TRAP e SNPs), que segregam nas proporções 1:1, 3:1 e 1:2:1 para construir um mapa genético em população de cana-de-açúcar derivada de cruzamento bi-parental, ou seja, entre dois genitores conhecidos (SP80-3280 x RB835486). Este foi o primeiro mapa genético a utilizar a abordagem de GBS (genotyping-by-sequencing). Esta técnica permite gerar uma quantidade maior de marcadores que são levados para análise de ligação e necessita de um genoma de referência para descoberta dos polimorfismos. Este mapa permitiu, posteriormente, o mapeamento de QTLs para importantes características agrônômicas da cana-de-açúcar (BALSALOBRE *et al.*, 2017).

Para o ano de 2022, utilizaram a técnica de sequenciamento de fragmentos amplificados de loci específicos (SLAF-seq) para criar um mapa genético integrado de alta densidade com marcadores SNP e SSR adicionais. Este mapa incluiu um total de 3.088 marcadores (3.068 SLAF e 20 SSR) distribuídos em 60 grupos de ligação (LG), abrangendo um total de 6.066,63 cM, com uma distância média de 1,96 cM entre marcadores adjacentes. (GAO *et al.*, 2022).

3.6 Mapeamento de QTL

O mapeamento de QTLs (quantitative trait loci) é uma forma importante de conhecer a base genética de características complexas controladas por vários genes. Para a cana-de-açúcar foram utilizados métodos estatísticos que permitem estimar a fração recombinante e as fases de ligação para diferentes proporções de segregação. Essa abordagem tem sido benéfica para obter mapas de ligação integrados na cana, utilizando marcadores em proporções de 1:1 e 3:1. No entanto, é importante ter em mente que essa estimativa é aproximada devido à natureza poliploide da cana-de-açúcar e por isso que esta análise tem sido realizada com certas limitações, uma vez que depende de genitores contrastantes para a característica de interesse na construção de populações segregantes, uma grande quantidade de marcadores

levados para análise de ligação e uma análise fenotípica bastante assertiva, de preferência realizada em diferentes locais e anos agrícolas (DAL-BIANCO *et al.*, 2012; ISLAM *et al.*, 2021).

Num estudo de mapeamento de QTL em cana-de-açúcar, foi utilizado o mapeamento de intervalo composto para espécies cruzadas. Diferentes níveis de ploidias foram considerados, e uma segregação de 1:2:1 para os SNPs. Além disso, usando marcadores AFLP e SSR, um mapa de ligação genética integrado de 187 indivíduos foi criado (COSTA *et al.*, 2016b).

4 MATERIAIS E MÉTODOS

4.1 Material vegetal

A população de mapeamento foi composta 238 genótipos de irmãos completos oriundos do cruzamento bi-parental entre as cultivares de cana-de-açúcar SP81-3250 (CP70-1547 x SP71-1279) e RB925345 (H59-1966 x ?), que são resistente e suscetível a ferrugem marrom, respectivamente. Esta população foi desenvolvida pelo Programa de Melhoramento Genético da Cana-de-açúcar (PMGCA) da Universidade Federal de São Carlos (UFSCar) conforme descrito em Balsalobre *et al.*, (2016). O DNA genômico total dos genitores e da progênie foram extraídos conforme descrito por Balsalobre *et al.*, (2017).

Avaliação a campo de características fenotípicas

A população de mapeamento foi avaliada em campo durante três anos agrícolas e em dois locais para dez características relacionadas com rendimento e produção de cana-de-açúcar. Os resultados desta avaliação estão sumarizados em Balsalobre *et al.*, 2016. Neste estudo, utilizamos duas das características que foram avaliadas que estão relacionadas com acúmulo e produção de biomassa, altura (SH) e diâmetro de plantas (SD) para buscar QTLs.

4.2 Marcadores baseados em GBS

GBS foi realizado pelo Institute for Genomic Diversity (Cornell University, Ithaca, NY, USA) de acordo com o protocolo descrito por Elshire, (2011) (ELSHIRE *et al.*, 2011). Amostras dos genitores da população foram replicadas três vezes para sequenciamento. Cada indivíduo dentro de uma biblioteca foi parte de uma reação 96-plex. Para fornecer uma alta profundidade de alelo, as bibliotecas foram obtidas pela digestão com a enzima *Pst*I, uma enzima de restrição de seis pares de base parcialmente sensível ao grupamento methyl. Adicionalmente, as bibliotecas 96-plex foram corridas em duas lanes distintas no HiSeq™ 2000 (Illumina® Inc., San Diego, CA, USA).

Para descobrir os polimorfismos, inicialmente foi utilizado o TASSEL-GBS pipeline (GLAUBITZ *et al.*, 2014), o qual foi implementado no software TASSEL (v. 4.3.8) e posteriormente modificado por Pereira *et al.*, (2018) para uso em poliploides. Este pipeline requer um genoma de referência foi utilizado duas referências: (i) o genoma de *Sorghum bicolor* (v. 3.0; ~709 Mb em 10 cromossomos); e (ii) montagem do haploide de *S. spontaneum* AP85-441 (~3.13 Gb em 32 pseudo-cromossomos representando quatro cópias dos oito cromossomos do conjunto básico) (ZHANG *et al.*, 2018). O algoritmo BOWTIE2 (v. 2.3.3) foi utilizado para mapear as tags de 64 pb contra cada referência. A referência exata e a profundidade do alelo alternativo (contagem de reads) foram gravadas em arquivo VCF (variant call format) (PEREIRA; GARCIA; MARGARIDO, 2018).

4.3 Estimativa de dosagem alélica

O GBS gerou dados de contagem de reads alelo-específicos no formato $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ para cada loco bialélico a partir dos indivíduos $i=1, 2, \dots, n$.

A matriz D de cada loco foi analisada no software SuperMASSA (SERANG; MOLLINARI; GARCIA, 2012) utilizando o script VCF2SM (PEREIRA; GARCIA; MARGARIDO, 2018). Em adição foram excluídos loci com menos de 10 reads por amostra, na média.

Todos os níveis de ploidia variando de 2 a 20 foram testados (GARCIA *et al.*, 2013). A ploidia que retornou a maior verossimilhança foi selecionada após o ajuste do modelo de segregação F1 no SuperMASSA. Os dados parentais replicados forneceram restrições adicionais durante a estimativa, e apenas ploídias variando de 6 a 14 foram selecionadas porque eram mais prováveis de aparecer no genoma da cana-de-açúcar e exibiram um maior número de SDMs (GARCIA *et al.*, 2013). O limite de probabilidade individual posterior no SuperMASSA foi ajustado para 0,80, de modo que apenas as chamadas de genótipos com alta certeza nas classes de dosagens foram mantidas. Além disso, foi exigido que pelo menos 75% das amostras atendessem ao limite de 0,80, sendo que amostras com valores abaixo disso foram tratadas como dado perdido. As dosagens atribuídas pelo SuperMASSA foram recodificadas para análise no software R substituindo os respectivos alelos de referência e alternativo por “a” e “b”, conforme notação do software OneMap.

4.4 Marcadores baseados em gel

Um total de 96 marcadores SSR e dois marcadores, o R12H16 e o 9O20-F4-Rsal, ligados ao gene de resistência a ferrugem marrom (*Bru1*) foram genotipados nos 238 irmãos completos e nos dois genitores. Os marcadores SSR foram derivados de sequências genômicas (14) e sequências expressas (EST) (BROWN *et al.*, 1996). As reações de PCR foram realizadas em volume final de 20 µL conforme descrito por (OLIVEIRA *et al.*, 2009). As amplificações foram desnaturadas a 90 °C por 3 min juntamente com 12 µL de buffer (formamida contendo 0,8 mM de EDTA e traços de azul de bromofenol e xileno cianol), resfriado rapidamente em gelo e submetido a eletroforese em géis de poliacrilamida desnaturante a 6% em tampão TBE 1X. As amostras foram carregadas em um sistema de eletroforese vertical duplo (CBS Scientific) e foram corridas a 75 W por 3 h. Um ladder de 10 pb foi usado como padrão de tamanho. As bandas foram visualizadas por coloração com nitrato de prata (CRESTE; TULMANN; FIGUEIRA, 2001).

4.5 Construção do mapa genético de ligação

A análise de ligação foi realizada com single-dose markers (SDMs) utilizando o pacote Onemap (v. 3.0.0) no software R (<http://www.cran.r-project.org>). Os

marcadores foram codificados de acordo com a notação proposta por Wu *et al.*, (2002). Brevemente, os alelos codominantes foram codificados como a e b, enquanto os alelos nulos foram codificados como o e tratados como alelos recessivos. Os padrões de segregação avaliados para marcadores oriundos de GBS foram 'B3.7' (ab x ab), 'D1.10' (ab x aa), 'D2.15' (aa x ab). Enquanto que para os marcadores baseados em gel os padrões de segregação foram 'C.8' (ao x ao), 'D1.13' (ao x oo) e 'D2.18' (oo x ao). 'D1' e 'D2' representam cruzamentos nos quais o loco é heterozigoto (e, portanto, informativo) apenas para SP81-3250 ou para RB925345, respectivamente; espera-se que ambos segreguem em uma proporção de 1:1. 'B3' e 'C' representam cruzamentos nos quais o loco é heterozigoto e simétrico em ambos pais; espera-se que o primeiro segregue em uma proporção de 1:2:1, enquanto o segundo segregará na proporção de 3:1. Os padrões de segregação foram testados pela função TEST_SEGRETATION, o qual realiza um teste qui-quadrado de acordo com a segregação mendeliana com correção de Bonferroni para múltiplo testes.

O mapa de ligação foi construído conforme descrito por Balsalobre *et al.*, (2017). Brevemente, para obter o mapa genético de ligação, primeiro foi realizado teste de dois pontos para identificar os grupos de ligação (LGs). Quaisquer marcadores emparelhados que mostraram um LOD score > 9,0 e uma fração de recombinação < 0,10 foram considerados ligados. Posteriormente, algoritmos de ordenação foram aplicados para cada grupo. Para grupos com menos de seis marcadores, a melhor ordem foi obtida realizando uma busca exaustiva com a função 'compare'. Para aqueles grupos com mais de seis marcadores, foi utilizado o comando 'order.seq', isto é, um conjunto inicial dos cinco marcadores mais informativos (preferencialmente os marcadores dos tipos 'B3' e 'C') foi amostrado para uma busca exaustiva. A melhor ordem foi utilizada como estrutura para a inclusão consecutiva de novos marcadores. Uma vez obtidos esses grupos, a função 'try.seq' foi utilizada para verificar os marcadores considerados desligados conforme o procedimento inicial, sendo possível integrá-los aos grupos pré-ordenados. Nesta etapa, outros marcadores também foram testados: (i) marcadores localizados nos finais dos LGs que estavam a mais de 20 cM de distância do marcador mais próximo; e (ii) marcadores que pertenciam a LGs muito pequenos (com tamanhos menores que 1 cM ou contendo apenas dois loci). Na etapa final, LGs com mais de cinco marcadores foram refinados usando o algoritmo 'ripple' dentro de uma janela deslizante de cinco

marcadores. Os heatmaps dos grupos ordenados foram inspecionados visualmente e correções manuais foram feitas conforme necessário durante todo o processo de criação do mapa. Os LGs foram desenhados no software MapChart (VOORRIPS, 2002).

4.6 Mapeamento de QTL

Para o mapeamento de QTL, foi utilizado o pacote fullsibQTL no software R (R CORE TEAM, 2014). O método CIM (composite interval mapping) foi utilizado para realizar o mapeamento (GAZAFFI *et al.*, 2014). Este método estima três efeitos genéticos: dois aditivos e um para dominância (um para cada genitor) (ZENG, 1994). Para inferir as probabilidades condicionais dos genótipos dos QTLs, probabilidade multiponto foram obtidas utilizando modelo oculto de Markov a cada 1 cM do mapa genético. A estratégia de mapeamento foi realizada em três etapas. Na primeira, foi realizado o mapeamento por intervalo (LANDER; BOTSTEINS, 1989) com a finalidade de selecionar marcadores como cofatores. Os picos com LOD maior que 2 foram incluídos no processo de detecção de QTL. Se o pico não foi coincidente com um marcador, o marcador mais próximo do pico foi considerado como cofator. Na segunda etapa, a busca de QTL foi realizada ao longo do genoma e considerou os cofatores localizados fora do grupo de ligação em análise. Para declarar um QTL, o threshold para cada busca foi obtido a partir de 1.000 permutações com um nível de significância de 0,95 (CHEN; STOREY, 2006). Por fim, os picos acima do threshold de permutação foram totalmente caracterizados, ou seja, a significância de cada efeito genético foi testada juntamente com a fase de ligação entre marcadores e QTLs e o padrão de segregação do QTL. As proporções de variância fenotípica explicada (R^2) para cada QTL detectado foram obtidas para todos os efeitos simultaneamente. Todas as análises foram realizadas no software R (R CORE TEAM, 2014)

5 RESULTADOS E DISCUSSÃO

5.1 Descoberta do polimorfismo de marcadores SNPs

As sequências utilizadas para descoberta de SNPs foram obtidas através da estratégia Genotyping-by-Sequencing (GBS) (ELSHIRE *et al.*, 2011). Do total de 1.834 bilhões de reads com barcode e com boa qualidade, foram obtidos perto de 6.7 milhões de tags a partir do alinhamento com duas diferentes referências. O alinhamento com o genoma de *S. spontaneum* resultou na maior taxa de alinhamento, com 65.98% (4.414.955) das tags alinhadas, enquanto que contra o genoma de *Sorghum* foram alinhadas 33.78% (2.260.432) das tags. Um total de 152.027 e 257.026 variantes foram identificadas com base nos genomas de *Sorghum* e *S. spontaneum*, respectivamente.

5.2 Estimativa da ploidia e dosagem alélica

Os níveis de ploidia avaliados com o software SUPERMASSA variaram de 2 a 20. Uma vez que a ploidia mais provável foi determinada, o software fornece a probabilidade posterior individual de cada indivíduo ser alocado em uma classe de dosagem esperada. Para construção do mapa genético de ligação, foram selecionados os locos que tiveram probabilidades posteriores maiores ou iguais a 0,80, com ploidias variando de 6 a 14. Isso representou 6.608 (4,34%) e 10.216 (3,97%) SNPs com base nas referências de *Sorghum* e *S. spontaneum*, respectivamente (Tabela 2). Desse total de marcadores, 1.029 foram excluídos por não seguirem os padrões de segregação descritos por Wu *et al.* (2002), restando, portanto, 15.795 marcadores SNPs baseados em GBS.

Tabela 2. Número de marcadores gerados após GBS na população de cana-de-açúcar.

Marcadores	Total
Genoma <i>S. spontaneum</i>	10.216
Genoma de <i>Sorghum</i>	6.608

5.3 Genotipagem de marcadores baseados em gel

Um total de 96 SSRs produziram 376 fragmentos polimórficos em dose única, sendo 243 (64,63%) testados para segregação 1:1 e 133 (35,37%) testados para segregação 3:1. Desse total, 285 fragmentos estavam aptos para análise de ligação, sendo 228 (80%) e 57 (20%) originados a partir de SSR gênicos e SSR genômicos, respectivamente (Tabela 3).

5.4 Mapa genético

Um mapa genético integrado foi construído utilizando 238 irmãos completos originados do cruzamento entre as variedades de cana-de-açúcar SP81-3250 e RB925345 (Figura 2). Do total de 15.795 marcadores em dose única avaliados (Tabela 2), 6.497 apresentaram redundância de informação genotípica e foram retirados do processo de análise. Do total de 9.298 marcadores, 7.088 (76%) tiveram distorção da segregação mendeliana (Figura 1). Portanto, 2.498 foram utilizados para análise de ligação (2.213 marcadores baseados em GBS e 285 marcadores SSRs), e 2.016 (80,70%) foram colocados no mapa de ligação (Tabela 3 e Tabela 4). Fazem parte do mapa de ligação 1.907 marcadores baseados em GBS e 109 SSRs. A distribuição dos marcadores mapeados de acordo com os padrões de segregação foram 24 marcadores do tipo C8 (3:1), 995 do tipo D1 ou que foram informativos apenas para SP81-3250 (953 marcadores baseados em GBS e 42 SSRs) e 997 do tipo D2 ou que foram informativos apenas para RB925345 (954 marcadores baseados em GBS e 43 SSRs) (Tabela 2). Os marcadores foram distribuídos ao longo de 221 LGs, com um tamanho cumulativo de 5.282 cM e uma densidade média de marcadores de 3,07 cM. O tamanho dos LGs variou de 1.48 cM (LG 228) a 201 cM (LG 01), com uma média de 28,93 cM; dois LGs mostraram tamanho menores que 2 cM, 54 LGs tiveram tamanho igual ou maior a 2 cM e menores que 10 cM, e os outros 172 LGs exibiram tamanho igual ou maior que 10 cM.

5.5 Mapeamento QTL

Utilizando o modelo CIM para buscar QTLs para altura (SH) e diâmetro (SD) de plantas. Um teste de permutação foi realizado para avaliar a significância estatística

dos QTLs (LANDER; BOTSTEINS, 1989). Valores limiares de LOD para altura de plantas e diâmetros de plantas foram 3,82 e 3,37, respectivamente. Neste estudo, um total de dois QTLs foram identificados. Os valores de R² variaram de 6,38% a 4,59. O QTL com maior LOD Score para o efeito aditivo foi de 0,93 para o genitor SP81-3250, que apresenta QTL associado ao diâmetro da planta (em LG 53) e por outro lado o genitor RB925345, com efeito aditivo de 0,03 associado à altura das plantas (em LG11); entretanto, para a característica SD apresentou valores negativos. O maior LOD Score para o efeito de dominância foi indeterminado para ambos os QTLs. Diferentes padrões de segregação foram exibidos, sendo 1:1 e 1:2:1, correspondendo a SH e SD (diâmetro e altura da planta) respectivamente.

Tabela 3. Marcadores em dose única baseados em gel e baseados em GBS avaliados na progênie do cruzamento entre as variedades de cana-de-açúcar SP81-3250 e RB925345.

Marcadores	Marcadores baseados em gel		Marcadores baseados em GBS	Total
	SSR Genômicos	SSR Gênicos		
Número de SDMs avaliados (baseados em gel e em GBS)	71	305	15,795	16,171
SDMs com segregação 1:1	42	201	9,254	9,497
SDMs com segregação 1:2:1 (baseados em GBS)	-	-	6,541	6,541
Duplo SDMs com segregação 3:1 (baseados em gel)	29	104	-	133
Número de marcadores com distorção de segregação	14	77	13,582	13,673
Total aptos para análise de ligação (1:1, 1:2:1 e 3:1)	57	228	2,213	2,498

Tabela 4. Distribuição dos diferentes tipos de marcadores mapeados de acordo com o tipo de cruzamento.

Tipo de cruzamento	Número de marcadores			
	Marcadores baseados em gel			Total
	SSR Genômico	SSR Gênico	Marcadores baseados em GBS	
D1.10 (<i>ab x aa</i>)	-	-	953	953
D1.13 (<i>ao x oo</i>)	3	39	-	42
D2.15 (<i>aa x ab</i>)	-	-	954	954
D2.18 (<i>oo x ao</i>)	7	36	-	43
B3.7 (<i>ab x ab</i>)	-	-	-	-
C.8 (<i>ao x ao</i>)	3	21	-	24
Total	13	96	1.907	2.016

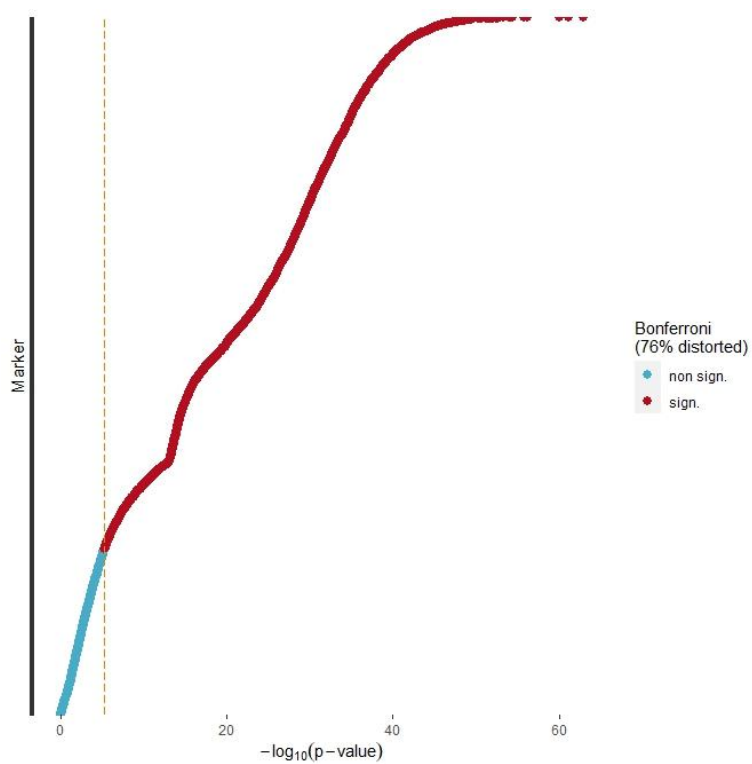
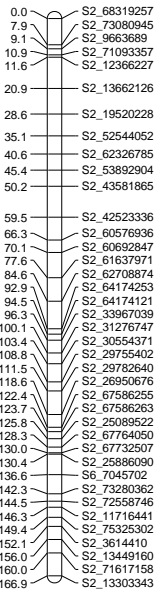
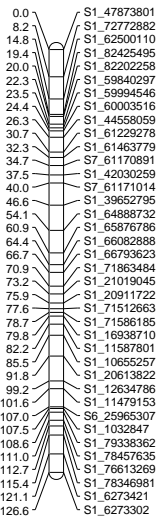


Figura 1. Teste de segregação mendeliana com correção de múltiplos testes de Bonferroni para os marcadores em dose-única.

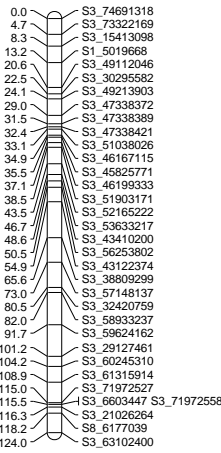
1



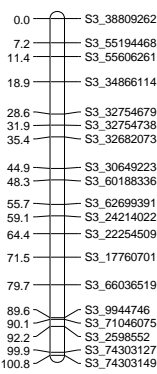
2



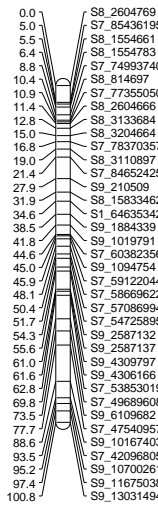
3



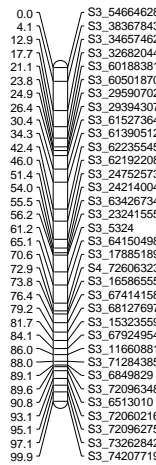
4



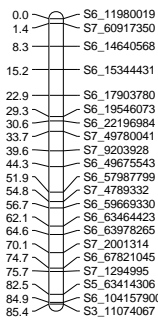
5



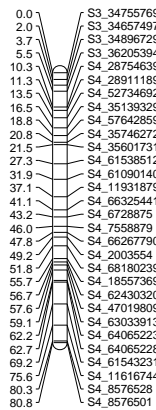
6



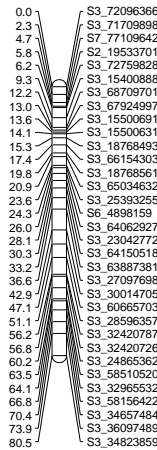
7



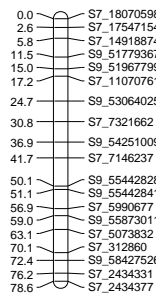
8



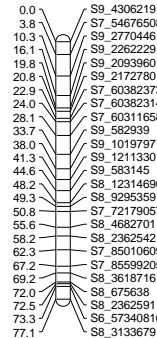
9



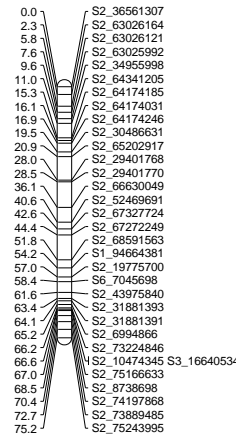
10

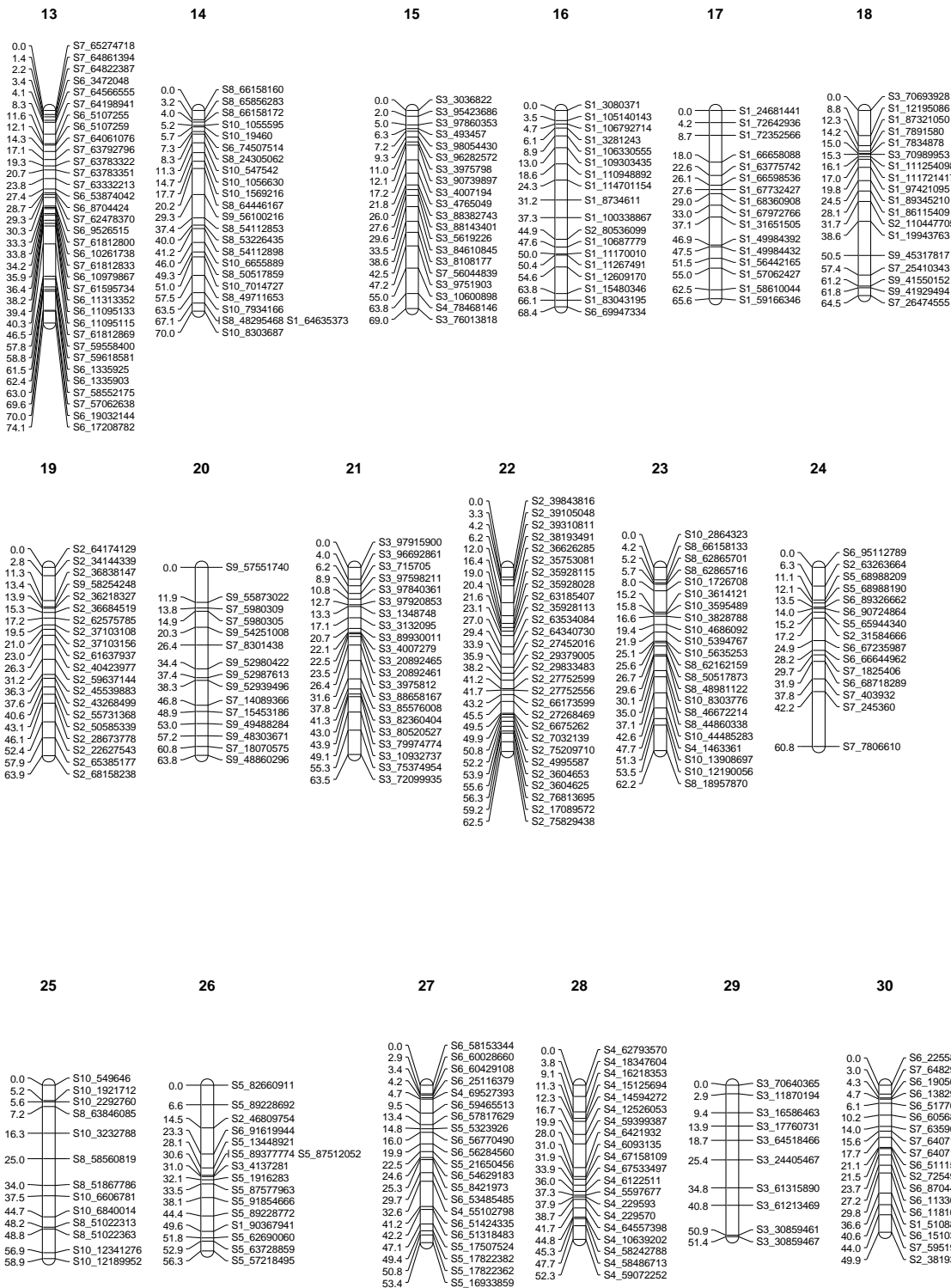


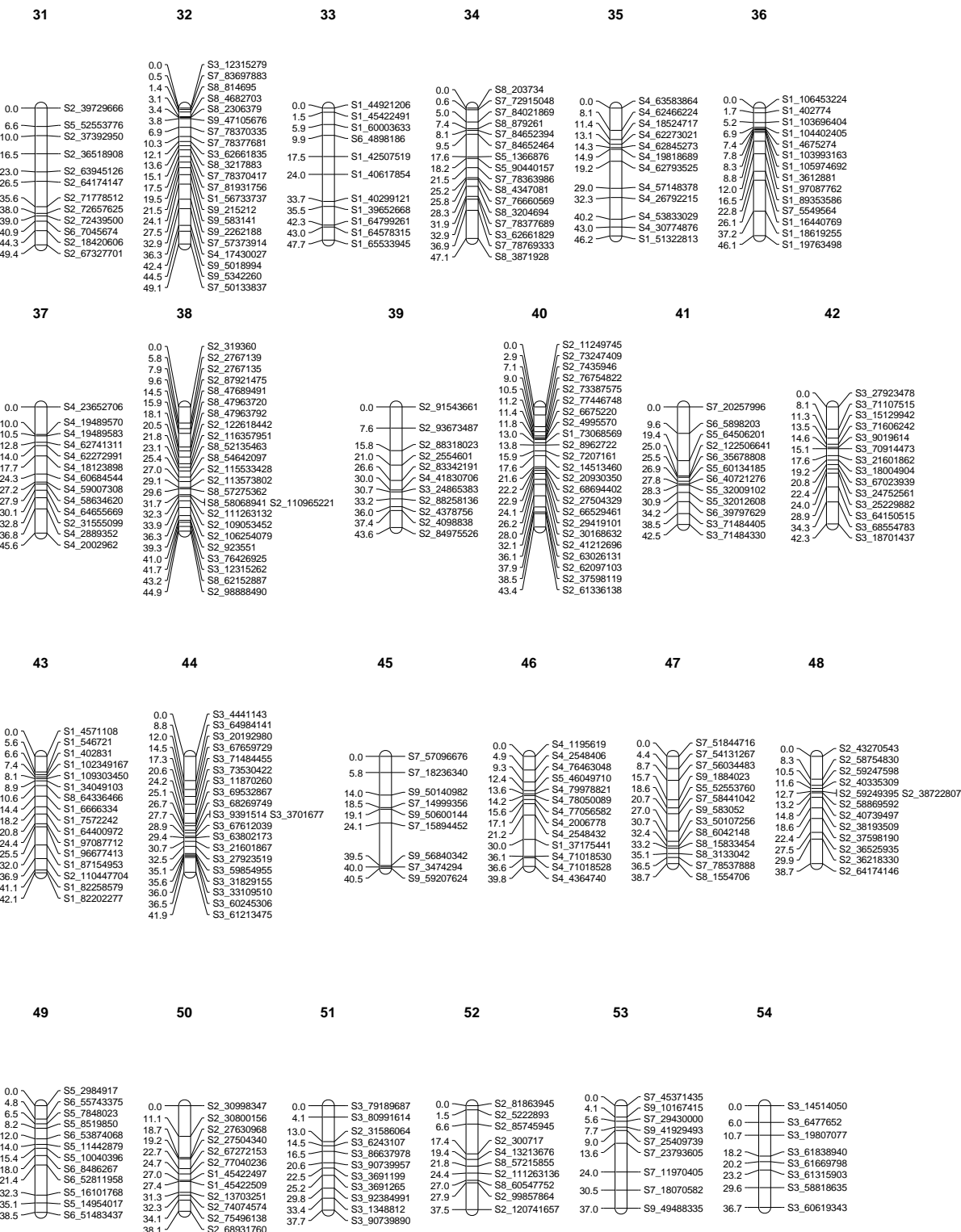
11



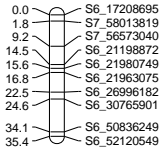
12



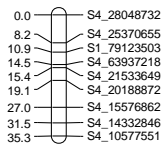




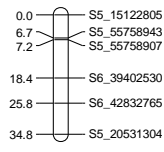
55



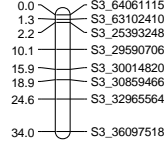
56



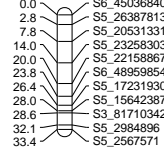
57



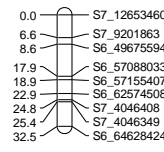
58



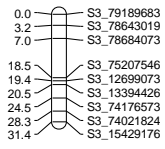
59



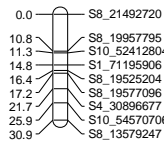
60



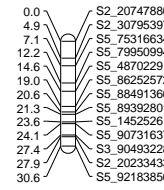
61



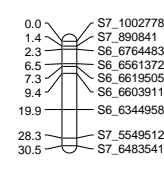
62



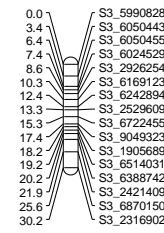
63



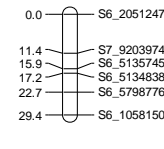
64



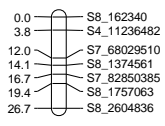
65



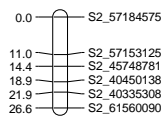
66



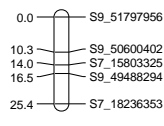
73



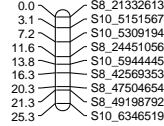
74



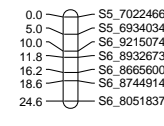
75



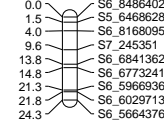
76



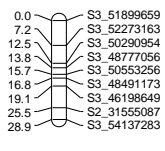
77



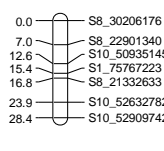
78



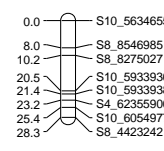
67



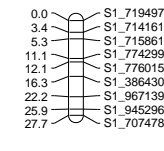
68



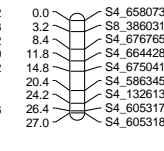
69



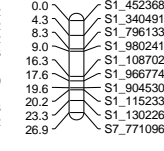
70



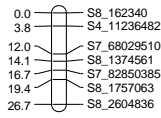
71



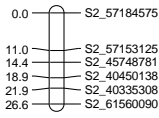
72



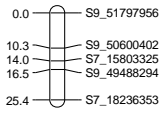
73



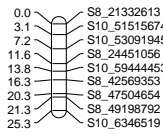
74



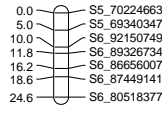
75



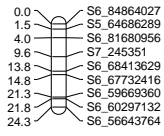
76



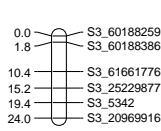
77



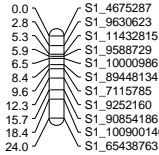
78



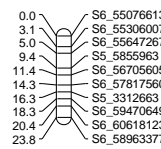
79



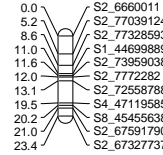
80



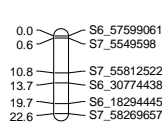
81



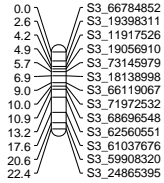
82



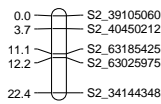
83



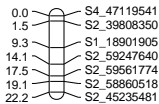
84



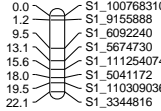
85



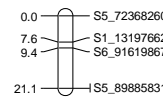
86



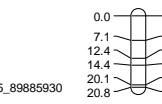
87



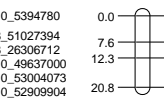
88



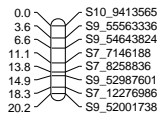
89



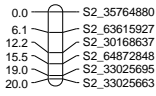
90



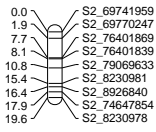
91



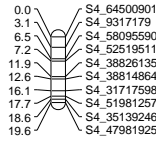
92



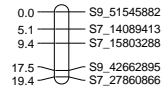
93



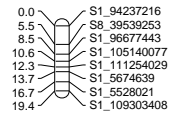
94



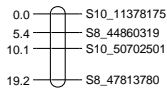
95



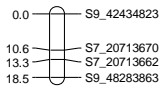
96



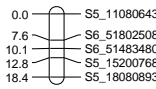
97



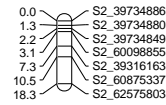
98



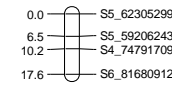
99



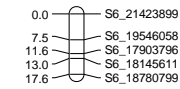
100



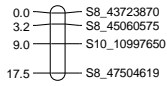
101



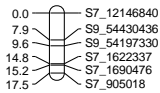
102



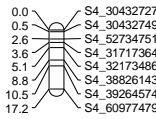
103



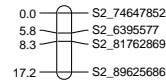
104



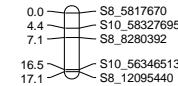
105



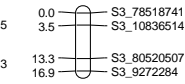
106



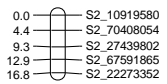
107



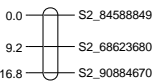
108



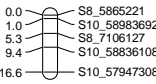
109



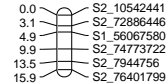
110



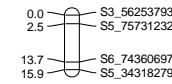
111



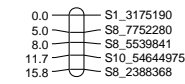
112



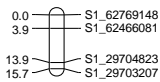
113



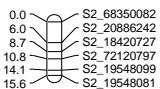
114



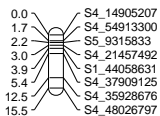
115



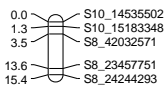
116



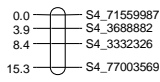
117



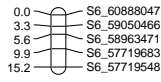
118



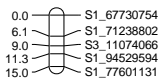
119



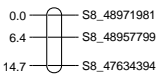
120



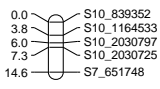
121



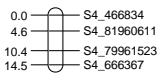
122



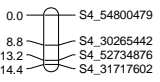
123



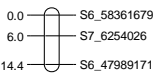
124



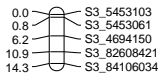
125



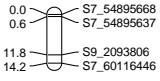
126



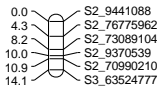
127



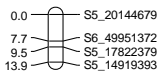
128



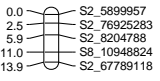
129



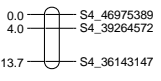
130



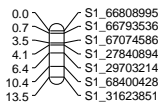
131



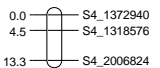
132



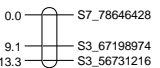
133



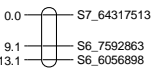
134



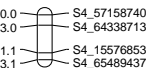
135



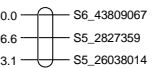
136



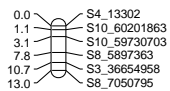
137



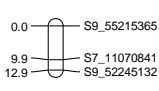
138



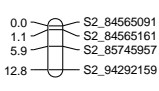
139



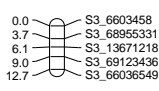
140



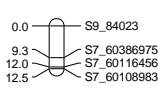
141



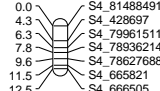
142



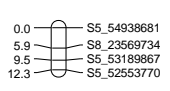
143



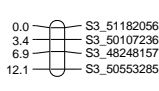
144



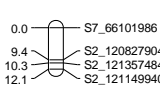
145



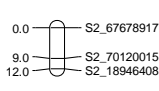
146



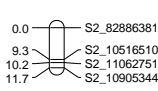
147



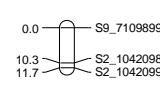
148



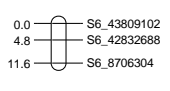
149



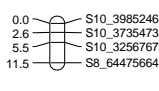
150



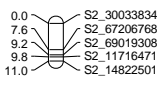
151



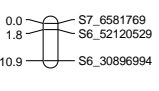
152



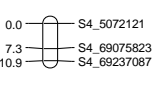
153



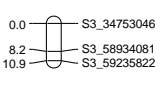
154



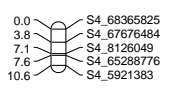
155



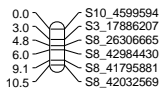
156



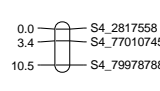
157



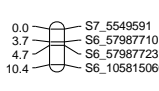
158



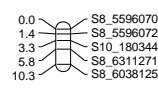
159



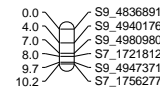
160



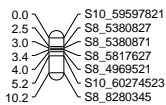
161



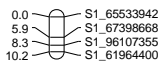
162



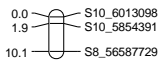
163



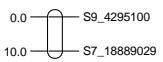
164



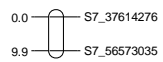
165



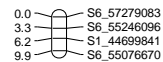
166



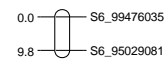
167



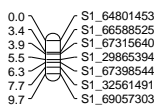
168



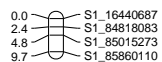
169



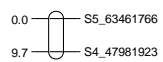
170



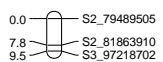
171



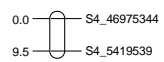
172



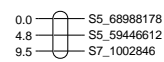
173



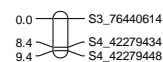
174



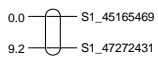
175



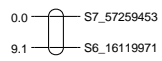
176



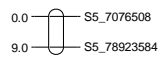
177



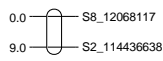
178



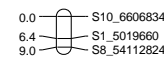
179



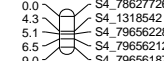
180



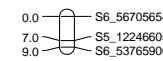
181



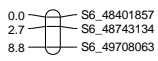
182



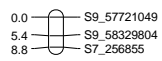
183



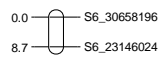
184



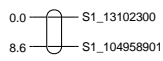
185



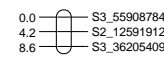
186



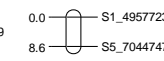
187



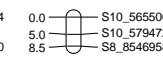
188



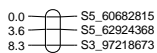
189



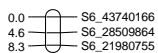
190



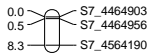
191



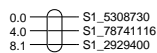
192



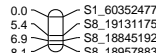
193



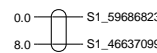
194



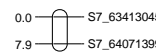
195



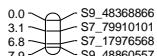
196



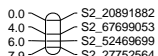
197



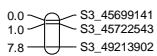
198



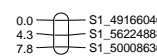
199



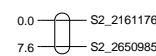
200



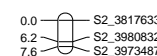
201



202



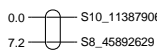
203



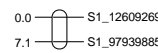
204



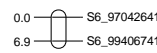
205



206



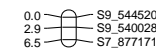
207



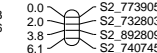
208



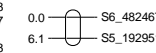
209



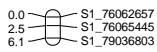
210



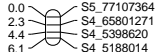
211



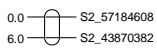
212



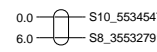
213



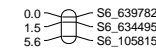
214



215



216



217



218



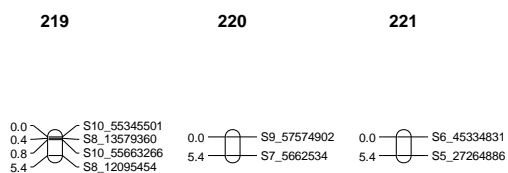


Figura 2. Mapa genético integrado obtido a partir da genotipagem com marcadores baseados em gel e baseados em GBS da progênie do cruzamento entre as variedades de cana-de-açúcar SP81-3250 e RB925345.

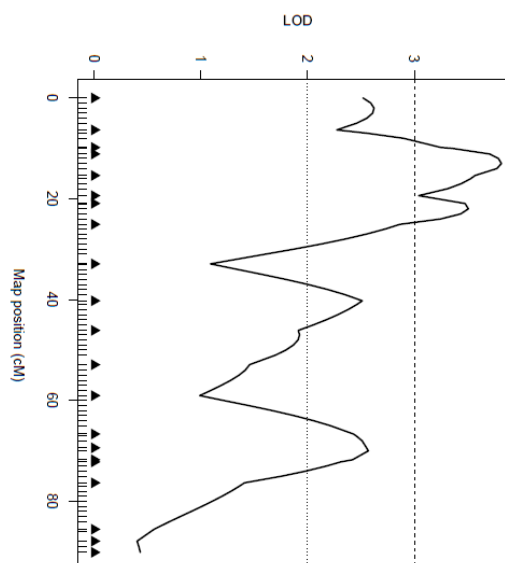
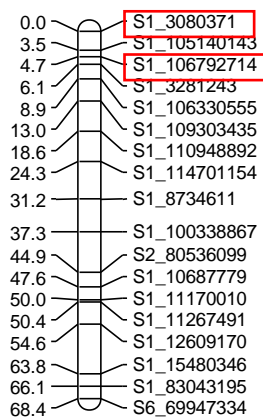
Tabela 5. QTLs mapeados para caracteres altura (SH) e diâmetro (SD) de plantas aplicando o modelo de mapeamento por intervalo composto (CIM).

QTL	Caracter	LG ^a	Posição ^a (cM)	Marcas flanqueadora ^b	Global ^b LOD	R ^{2c}	Efeito aditivo SP81-3250 ^d	LOD ^e	Efeito aditivo RB925345 ^d	LOD ^e	Efeito dominância ^d	LOD ^e	Segregação ^f
1	SH	11	13,00	S1_106792714 - QTL - S1_3080371	3,82	6,38	NA	NA	0,03	3,82	NA	NA	1:1
1	SD	53	17,64	S5_10040396	3,37	4,59	0,93	0,62	-0,25	0,33	NA	NA	1:2:1

^aLG: Grupos de ligação; Posição (cM): posição do QTL no LG; ^bmarcas adjacentes para QTLs e LODs associados; ^cvariação fenotípica explicada; ^defeitos aditivos de pais e efeitos de dominância; ^eLODs dos efeitos aditivos e de dominância; ^festimativa do padrão de segregação do QTL.

Stalk height (SH) - Altura de plantas

LG 11



Stalk diameter (SD) – Diâmetro de plantas

LG 53

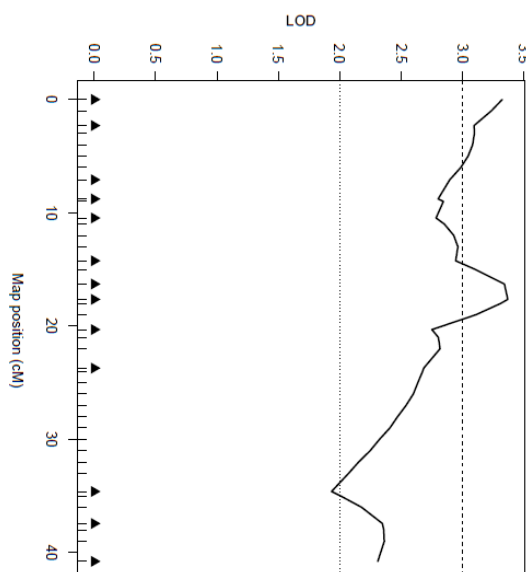
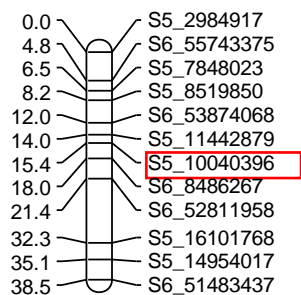


Figura 3. Mapeamento por intervalo composto (CIM) para diâmetro (SD, mm) e altura (SH, cm) de planta população F1 originada do cruzamento entre SP81-3250 e RB925345. À esquerda estão os grupos de ligação (LG) onde os QTLs foram detectados e os retângulos vermelhos indicam os marcadores que flanqueiam os QTLs. À direita, temos os gráficos que indicam os limites de LOD obtidos após testes de permutação; 3,06 e 3,17 para SH e SD, respectivamente.

O genoma da cana-de-açúcar é muito complexo e apresenta algumas limitações para seu estudo; a junção nos híbridos modernos de porções genéticas de diferentes espécies que fazem parte do complexo *Saccharum* (majoritariamente *S. officinarum* e *S. spontaneum*), a ocorrência de aneuploidia, poliploidia e número grande e diferente de cromossomos entre espécies e, conseqüentemente, entre os híbridos modernos. Esta complexidade faz com que os pesquisadores busquem modelos que permitam avanços no entendimento da estrutura genética e genômica da cana-de-açúcar. Assim, pela proximidade genética, o genoma do sorgo tem sido utilizado como referência para a análise comparativa com a cana-de-açúcar (SOUZA *et al.*, 2011; BALSALOBRE *et al.*, 2017). Em adição, os genomas das cultivares de cana-de-açúcar R570 (GARSMEUR *et al.*, 2018) e SP80-3280 (SOUZA *et al.*, 2019), e de um acesso de *S. spontaneum* (ZHANG *et al.*, 2018a) foram disponibilizados recentemente para a comunidade científica, permitindo avanços maiores no entendimento da estrutura genética e genômica da cana-de-açúcar.

Apesar das limitações impostas pela biologia da cana-de-açúcar, alguns trabalhos já foram realizados com o objetivo de associar marcadores moleculares com características fenotípicas (ANONI *et al.*, 2013; BALSALOBRE *et al.*, 2017a; COSTA *et al.*, 2016a; HOARAU *et al.*, 2002; LU *et al.*, 2023). No entanto, o uso de sequenciamento de nova geração para obtenção de marcadores SNPs em cana-de-açúcar é mais recente e as metodologias de análise de dados para descoberta de variantes, classificação em ploidias e dosagens alélicas, e incorporação dessas informações no mapa genético e mapeamento de QTL ainda estão na fronteira do conhecimento (DA SILVA PEREIRA *et al.*, 2020; GAO *et al.*, 2022; MOLLINARI; GARCIA, 2019).

O GBS é um método de genotipagem que ajuda a detectar variações genéticas e interrogar a estrutura do genoma, reduzindo a complexidade do sequenciamento para obter polimorfismos de alta qualidade a baixo custo (CAVANAGH *et al.*, 2013). A cultura da cana-de-açúcar ainda não possui uma referência específica, por isso foram utilizadas 2 pseudo-referências (Tabela 2) para obtenção dos marcadores por este método. Trabalhos anteriores como Balsalobre *et al.*, (2017) e Paterson *et al.*, (2019) permitiram relacionar os marcadores da cana-de-açúcar com o genoma do sorgo (BALSALOBRE *et al.*, 2017; PATERSON *et al.*, 2009).

As sequências utilizadas foram obtidas a partir do mapeamento populacional de dois progenitores das bibliotecas digeridas com a enzima *Pst*I. A enzima *Pst*I (CTGCAG) é uma rara enzima de restrição de corte e isso permite a produção de um alto nível de marcadores confiáveis; além disso, a taxa de linhagem é alta, possivelmente porque essa enzima é sensível à metilação (BALSALOBRE *et al.*, 2017). Além disso, a proporção de loci (SNPs) sobre os marcadores filtrados é maior (POLAND *et al.*, 2012).

Em nosso mapa genético integrado, do total de marcadores avaliados em dose única, 41% foram retirados devido à redundância de informações genotípicas. Dos marcadores totais, 27% foram utilizados para análise de ligação e 80,70% foram colocados no mapa de ligação. O mapa de ligação inclui 1.907 marcadores baseados em GBS e 109 SSRs. A distribuição dos marcadores incluiu 24 marcadores do tipo C8 (3:1), 995 do tipo D1 e 997 do tipo D2, sendo semelhante ao que é mencionado em Gutierrez *et al.* (2018). Essa distribuição de marcadores é maior em comparação com estudos anteriores (GUTIERREZ *et al.*, 2018; OLIVEIRA *et al.*, 2007) o que demonstra um avanço significativo na densidade do mapa e na cobertura do genoma na cana-de-açúcar, distribuídos ao longo de 221 LGs, que foram mais em comparação com os 192 relatados em Oliveira *et al.* (2007), com um tamanho de 5.282 cM e uma densidade de 3,07 cM, que são menores que 6.261,1 cM e 3.658,7 cM para tamanho e densidade, respectivamente (BALSALOBRE *et al.*, 2017a; CARDOSO-SILVA *et al.*, 2014; OLIVEIRA *et al.*, 2007). No entanto, existem evidências de que eles podem ocorrer em tamanhos menores, como mostra em Reffay *et al.* (2005), Aitken *et al.* (2005) e Raboin *et al.* (2006). Por outro lado, para o ano de 2013, Singh *et al.* mapeou marcadores polimórficos em 24 grupos de ligação, com um tamanho de 2606.77 cM (SINGH *et al.*, 2013). O tamanho dos LGs variou de 1,48 cM a 201 cM, com uma média de 28,93 cM. Dois LGs apresentaram tamanhos menores que 2 cM, 54 LGs tiveram tamanhos iguais ou maiores que 2 cM e menores que 10 cM, e os outros 172 LGs exibiram tamanhos iguais ou maiores que 10 cM. E por outro lado, os marcadores SNP, SSR e RFLP identificados foram agrupados em 8 grupos homólogos (HG) e um conjunto adicional de LGs não atribuídos segundo Aiken *et al.* (AITKEN *et al.*, 2014).

Foi utilizado o modelo CIM (Mapeamento de Intervalo Composto) e a análise de permutação para identificar QTLs (GUTIERREZ *et al.*, 2018; LI; YE; WANG, 2007a;

ZENG, 1994) relacionados à altura e ao diâmetro das plantas e foram identificados dois QTLs significativos, um associado ao diâmetro e outro à altura, o que sugere que essas características são influenciadas por diferentes regiões do genoma(WANG *et al.*, 2022b). Além disso, foram observados diferentes padrões de segregação nos QTLs identificados, indicando a existência de variabilidade genética na população estudada. No entanto, os percentuais de variação fenotípica explicada por esses QTLs (MING *et al.*, 2001; PASTINA *et al.*, 2012) foram relativamente baixos, sugerindo a presença de outros fatores contribuindo para essas características. Em outros estudos, utilizaram as mesmas análises, foram identificados sete QTLs relacionados às características de BRIX e POL%C, encontrados na mesma região do genoma, sugerindo uma associação genética entre esses traços (BALSALOBRE *et al.*, 2017a). Em outro estudo, pode-se observar uma correlação negativa entre o diâmetro e a altura da cana com o conteúdo de açúcar (BRIX), indicando que o aumento da biomassa pode estar associado a uma redução no conteúdo de açúcar. Essa observação é relevante para os melhoradores de cana-de-açúcar, que devem encontrar um equilíbrio entre o aumento da biomassa e o conteúdo de açúcar para obter um rendimento ótimo de açúcar (SINGH *et al.*, 2013).

Embora tenham sido encontrados marcadores genéticos relacionados à resistência a doenças na cana-de-açúcar, a implementação da Seleção Assistida por Marcadores (MAS) (CAVALCANTI VASCONCELOS, 2009; LI; YE; WANG, 2007b) em programas de melhoramento genético deste cultivo continua sendo um desafio, pois os traços-chave, como o rendimento, são influenciados por múltiplos loci de características quantitativas, e cada um deles contribui apenas de forma limitada para o efeito fenotípico geral (PASTINA *et al.*, 2012). Até o momento, somente os marcadores PCR para os haplótipos Bru-1 e Bru-2 apresentam um potencial para identificar genes de resistência duradoura à ferrugem marrom no germoplasma da cana-de-açúcar(HASAN *et al.*, 2021) .

São necessários estudos adicionais para validar os marcadores de DNA da cana-de-açúcar e sua aplicação em programas de melhoramento genético e na identificação de germoplasma. O uso de mapas genéticos mais precisos e completos pode ter um impacto significativo na indústria da cana-de-açúcar ao acelerar a obtenção de variedades aprimoradas e sustentáveis que atendam às demandas de uma população em crescimento e a um mercado agrícola em constante evolução.

Nossos resultados são importantes para o melhoramento genético da cana-de-açúcar, pois permitem identificar marcadores genéticos associados a características desejáveis, como maior produtividade e adaptação a diferentes ambientes. Com base nesses resultados, os pesquisadores podem selecionar indivíduos com as características desejadas para cruzamentos futuros, visando o desenvolvimento de cultivares mais produtivas e adaptadas às necessidades agrícolas.

6 CONCLUSÕES

Por meio das análises, foram obtidos resultados que mostram que a estimativa das ploidias e o uso de marcadores gerados em trabalhos anteriores auxiliaram na construção de um mapa genético para o cultivo da cana-de-açúcar.

Conseguiu-se construir um mapa genético integrado utilizando uma população de irmãos completos derivados de duas variedades de cana-de-açúcar. Embora uma quantidade significativa de marcadores tenha sido excluída devido à redundância genotípica, mais de 2.000 marcadores foram mapeados no mapa de ligação, fornecendo informações valiosas sobre a distribuição dos marcadores e sua associação com características de interesse na população estudada.

Os resultados sugerem uma variabilidade genética na segregação mendeliana dos marcadores, o que pode refletir a diversidade genética presente na população parental utilizada para o cruzamento.

7 LITERATURA CITADA

AITKEN, K. S. *et al.* A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array Technology (DArT) markers. [S. l.: s. n.], 2014. Disponível em: <http://www.biomedcentral.com/1471-2164/15/152>. .

AITKEN, K. S. *et al.* AFLP analysis of genetic diversity within *Saccharum officinarum* and comparison with sugarcane cultivars. **Australian Journal of Agricultural Research**, [s. l.], v. 57, n. 11, p. 1167–1184, 2006.

AITKEN, K. S. History and Development of Molecular Markers for Sugarcane Breeding. **Sugar Tech**, [s. l.], v. 24, p. 341–353, 2022. Disponível em: <https://link.springer.com/article/10.1007/s12355-021-01000-7#citeas>. Acesso em: 9 jul. 2023.

ANONI, C. *et al.* QTL Mapping for yield components and quality parameters in sugarcane. Disponível em: <https://www.researchgate.net/publication/268120664%0AQTL>.

ANONI, C. *et al.* QTL mapping for yield components in a sugarcane commercial cross. 2013. Disponível em: www.sbg.org.br. Acesso em: 22 jun. 2023.

BALSALOBRE, T. W. A. *et al.* GBS-based single dosage markers for linkage and QTL mapping allow gene mining for yield-related traits in sugarcane. **BMC Genomics**, [s. l.], v. 18, n. 1, 2017a.

BALSALOBRE, T. W. A. *et al.* GBS-based single dosage markers for linkage and QTL mapping allow gene mining for yield-related traits in sugarcane. **BMC Genomics**, [s. l.], v. 18, n. 1, p. 72, 2017b.

BENDEZÚ, G. “EVALUACIÓN DEL SISTEMA RADICULAR DE CAÑA DE AZÚCAR (*Saccharum officinarum* L.) EN RIEGO POR GOTEIO Y GRAVEDAD-PARAMONGA”. [s. l.], n. 511, p. 128, 2014. Disponível em: <http://repositorio.lamolina.edu.pe/bitstream/handle/UNALM/859/T007011.pdf?sequence=1&isAllowed=y>.

BROWN, S. M. *et al.* Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. **Theoretical and Applied Genetics**, [s. l.], v. 93, n. 1–2, p. 190–198, 1996.

CARDOSO-SILVA, C. B. *et al.* De novo assembly and transcriptome analysis of contrasting sugarcane varieties. **PLoS ONE**, [s. l.], v. 9, n. 2, 2014.

CAVALCANTI VASCONCELOS, J. J. Importância da seleção assistida por marcadores moleculares. [S. l.: s. n.], 2009.

CAVANAGH, C. R. *et al.* Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. **Proceedings of the National Academy of Sciences**, [s. l.], v. 110, n. 20, p. 8057–8062, 2013. Disponível em: <https://pnas.org/doi/full/10.1073/pnas.1217133110>.

CHASTEL, J. M. Le sucre et ses marches. **Agriculture et Developpement**, [s. l.], v. 4, p. 4–11, 1994. Disponível em: https://agritrop.cirad.fr/387346/1/document_387346.pdf.

CHEN, L.; STOREY, J. D. Relaxed significance criteria for linkage analysis. **Genetics**, [s. l.], v. 173, n. 4, p. 2371–2381, 2006.

COMPANHIA NACIONAL DE ABASTECIMENTO - (CONAB). **Acompanhamento da Safra Brasileira - CANA-DE-ACUCAR** Observatório Agrícola. [S. l.: s. n.], 2022. Disponível em: <https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safra-de-cana-de-acucar>. Acesso em: 17 maio 2023.

COMPANHIA NACIONAL DE ABASTECIMENTO – CONAB. **Acompanhamento da safra brasileira de cana-de-açúcar**. Brasília: [s. n.], 2023. Disponível em: <https://www.conab.gov.br/info-agro/safras/cana>. Acesso em: 18 jul. 2023.

CORONEL CORTÉS, J. C. Caña de azúcar: Principales aspectos de la cadena agroproductiva. **Dirección General de Competitividad Agraria**, [s. l.], p. 35, 2016. Disponível em: <https://silo.tips/download/direccion-general-de-competitividad-agraria-caa-de-azucar-principales-aspectos-d>.

COSTA, E. A. *et al.* QTL mapping including codominant SNP markers with ploidy level information in a sugarcane progeny. **Euphytica**, [s. l.], v. 211, n. 1, p. 1–16, 2016a.

COSTA, E. A. *et al.* **QTL mapping including codominant SNP markers with ploidy level information in a sugarcane progeny**. [S. l.]: Springer Netherlands, 2016b.

COSTET, L. *et al.* Haplotype structure around Bru1 reveals a narrow genetic basis for brown rust resistance in modern sugarcane cultivars. **Theoretical and Applied Genetics**, [s. l.], v. 125, n. 5, p. 825–836, 2012.

COX, M. C.; HOGARTH, D. M.; B, H. P. Breeding and selection for high early season sugar content in a sugarcane (*Saccharum* spp. hybrids) improvement program. **Aust. J. Agric. Res.**, [s. l.], v. 45, p. 1569–1575, 1994.

CRESTE, S.; TULMANN NETO, A.; FIGUEIRA, A. Detection of Single Sequence Repeat Polymorphisms in Denaturing Polyacrylamide Sequencing Gels by Silver Staining. **Plant Molecular Biology Reporter**, [s. l.], v. 19, n. 4, p. 299–306, 2001.

CURSI, D. E. *et al.* History and Current Status of Sugarcane Breeding, Germplasm Development and Molecular Genetics in Brazil. **Sugar Tech**, [s. l.], 2021. Disponível em: <https://doi.org/10.1007/s12355-021-00951-1>.

CURSI, D. E. *et al.* Origin, Genetic Diversity, Conservation, and Traditional and Molecular Breeding Approaches in Sugarcane. [S. l.: s. n.], 2022.

DA SILVA PEREIRA, G. *et al.* Multiple QTL mapping in autopolyploids: A random-effect model approach with application in a hexaploid sweetpotato full-sib population. **Genetics**, [s. l.], v. 215, n. 3, p. 579–595, 2020.

DAL-BIANCO, M. *et al.* **Sugarcane improvement: How far can we go?**. [S. l.: s. n.], 2012.

DANIELS, J.; ROACH, B. T. **Taxonomy and evolution**. [S. l.]: Elsevier B.V., 1987-. ISSN 0378519X.v. 11 Disponible em: <http://dx.doi.org/10.1016/B978-0-444-42769-4.50007-2>.

DARGENT, E. **Historia del azúcar y sus derivados en el Perú** Universidad Ricardo Palma. [S. l.: s. n.], 2017. Disponible em: <https://www.urp.edu.pe/pdf/id/20926/n/historia-del-azucar-y-sus-derivados-en-el-peru.pdf>.

DAUGROIS, J. H. *et al.* **A putative major gene for rust resistance linked with a RFLP marker in sugarcane cultivar "R570"** *Theor Appl Genet*. [S. l.]: Springer-Verlag, 1996.

D'HONT, A. *et al.* Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. **Molecular and General Genetics**, [s. l.], v. 250, n. 4, p. 405–413, 1996.

ELSHIRE, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS ONE**, [s. l.], v. 6, n. 5, p. 1–10, 2011.

FAO. Sugar. *Em:* , 2019. **OECD-FAO Agricultural outlook 2019-2028**. [S. l.: s. n.], 2019. p. 154–165.

FRETES, F.; MARTÍNEZ, M. Análisis De La Cadena De Valor En Concepción Y Canindeyú. **USAID/Paraguay**, [s. l.], v. 4, p. 46, 2011.

GAO, Y. *et al.* Quantitative Trait Loci Mapping and Development of KASP Marker Smut Screening Assay Using High-Density Genetic Map and Bulk Segregant RNA Sequencing in Sugarcane (*Saccharum* spp.). **Frontiers in Plant Science**, [s. l.], v. 12, 2022.

GARCIA, A. A. F. *et al.* SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. **Scientific Reports**, [s. l.], v. 3, p. 1–10, 2013.

GARSMEUR, O. *et al.* A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications**, [s. l.], v. 9, n. 1, 2018.

GLAUBITZ, J. C. *et al.* TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. **PLoS ONE**, [s. l.], v. 9, n. 2, 2014.

GOÑI DELIÓN, J. C.; ROJAS DELGADO, M. Manual de combustibles alternativos y tecnología automotriz. led. Lima - Perú: Saxo.com Peru S.A.C, 2016.

GUTIERREZ, A. F. *et al.* Identification of genomic regions controlling leaf scald resistance in sugarcane using a Bi-parental mapping population and selective genotyping by sequencing. **Frontiers in Plant Science**, [s. l.], v. 9, n. June, p. 1–10, 2018.

HASAN, N. *et al.* **Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes**. [S. l.]: Springer Science and Business Media Deutschland GmbH, 2021.

HOARAU, J. Y. *et al.* Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). II. Detection of QTLs for yield components. **Theoretical and Applied Genetics**, [s. l.], v. 105, n. 6–7, p. 1027–1037, 2002.

ISLAM, M. S. *et al.* Identification of quantitative trait loci (QTL) controlling fibre content of sugarcane (*Saccharum* hybrids spp.). **Plant Breeding**, [s. l.], v. 140, n. 2, p. 360–366, 2021.

LANDER, E.; BOTSTEINS, D. Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. **Genetics**, [s. l.], 1989.

LI, H.; YE, G.; WANG, J. A modified algorithm for the improvement of composite interval mapping. **Genetics**, [s. l.], v. 175, n. 1, p. 361–374, 2007a.

LI, H.; YE, G.; WANG, J. A modified algorithm for the improvement of composite interval mapping. **Genetics**, [s. l.], v. 175, n. 1, p. 361–374, 2007b.

LU, G. *et al.* Identification of QTLs and critical genes related to sugarcane mosaic disease resistance. **Frontiers in Plant Science**, [s. l.], v. 14, 2023.

MASUELLI, R. Uso de marcadores moleculares en el mejoramiento genético de especies hortícolas. **Avances en Horticultura**, [s. l.], v. 4(1), n. 5505, p. Edición on-line, 1999.

MCINTYRE, C. L. *et al.* Resistance gene analogues in sugarcane and sorghum and their association with quantitative trait loci for rust resistance. **Genome**, [s. l.], v. 48, n. 3, p. 391–400, 2005.

MING, R. *et al.* QTL analysis in a complex autopolyploid: Genetic control of sugar content in sugarcane. **Genome Research**, [s. l.], v. 11, n. 12, p. 2075–2084, 2001.

MINISTERIO DE AGRICULTURA Y RIEGO. **Reporte trimestral publico de los meses de abril-Junio**. [S. l.: s. n.], 2020. Disponible em: https://cdn.www.gob.pe/uploads/document/file/1026014/Commodities_arroz_abr-jun_2020.pdf. .

MOLLINARI, M.; GARCIA, A. A. F. Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. **G3: Genes, Genomes, Genetics**, [s. l.], v. 9, n. 10, p. 3297–3314, 2019.

NADEEM, M. A. *et al.* DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. [S. l.]: Taylor and Francis Ltd., 2018.

OLIVEIRA, K. M. *et al.* Characterization of new polymorphic functional markers for sugarcane. **Genome**, [s. l.], v. 52, n. 2, p. 191–209, 2009.

OLIVEIRA, K. M. *et al.* Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding**, [s. l.], v. 20, n. 3, p. 189–208, 2007.

PACHECO CANALES, Y. R. Evaluación de la política pública sobre bioetanol como combustible en Perú, 2003-2019. **Semestre Económico**, [s. l.], v. 22, n. 53, p. 127–162, 2019.

PASTINA, M. M. *et al.* A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theoretical and Applied Genetics**, [s. l.], v. 124, n. 5, p. 835–849, 2012.

PATERSON, A. H. *et al.* The Sorghum bicolor genome and the diversification of grasses. **Nature**, [s. l.], v. 457, n. 7229, p. 551–556, 2009.

PEREIRA, G. S.; GARCIA, A. A. F.; MARGARIDO, G. R. A. A fully automated pipeline for quantitative genotype calling from next generation sequencing data in autopolyploids. **BMC Bioinformatics**, [s. l.], v. 19, n. 1, p. 1–10, 2018.

PERERA, F. M. *et al.* Los marcadores moleculares TRAP permiten identificar líneas transgénicas de caña de azúcar (*Saccharum* spp.) genéticamente similares al genotipo sin transformar. **Rev. Ind. y Agríc. de Tucumán**, [s. l.], v. 97, n. 1, p. 1–7, 2020. Disponible em: <https://ri.conicet.gov.ar/handle/11336/147092>. Acceso em: 15 jul. 2023.

PERUCAÑA. Sector agroindustrial del azúcar y derivados en el Perú. [S. l.: s. n.], 2019. Disponible em: https://www.perucana.com/_files/ugd/6ac0b0_b21198ffdf54c16a1e83f2695cbb1.pdf. Acceso em: 20 maio 2023.

PIPERIDIS, N.; D'HONT, A. Sugarcane genome architecture decrypted with chromosome-specific oligo probes. **Plant Journal**, [s. l.], v. 103, n. 6, p. 2039–2051, 2020.

POLAND, J. A. *et al.* Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. **PLoS ONE**, [s. l.], v. 7, n. 2, 2012.

POMPIDOR, N. *et al.* Three founding ancestral genomes involved in the origin of sugarcane. **Annals of Botany**, [s. l.], v. 127, n. 6, p. 827–840, 2021.

R CORE TEAM. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2014. Disponible em: <http://www.R-project.org/>. Acceso em: 2 jun. 2023.

RIVERA, F. El cultivo de la caña de azúcar (*Saccharum officinarum* L) en la región de Cardel, centro de Veracruz. 2008. 55 f. [s. l.], 2008. Disponible em: [http://repositorio.uaaan.mx:8080/xmlui/bitstream/handle/123456789/1417/EL_CULTIVO_DE_LA_CA%20D1A_DE_AZUCAR_\(Saccharum_officinarum_L.\)_EN_LA_REGION_DE_CARDEL,_CENTRO_DE_VERACRUZ.pdf?sequence=1](http://repositorio.uaaan.mx:8080/xmlui/bitstream/handle/123456789/1417/EL_CULTIVO_DE_LA_CA%20D1A_DE_AZUCAR_(Saccharum_officinarum_L.)_EN_LA_REGION_DE_CARDEL,_CENTRO_DE_VERACRUZ.pdf?sequence=1).

SERANG, O.; MOLLINARI, M.; GARCIA, A. A. F. Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. **PLoS ONE**, [s. l.], v. 7, n. 2, p. 1–13, 2012.

SILVA-ROJAS, V. H.; AGUIRRE RAYO, J. M.; CERVANTES, B. Mejoramiento genético de la caña de azúcar mediante el uso de marcadores moleculares en México. [S. l.: s. n.], 2017. Disponible em: <https://www.researchgate.net/publication/321978728>. .

SINGH, R. K. *et al.* Genetic mapping and QTL analysis for sugar yield-related traits in sugarcane. **Euphytica**, [s. l.], v. 191, n. 3, p. 333–353, 2013.

SOUZA, G. M. *et al.* Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. **GigaScience**, [s. l.], v. 8, n. 12, 2019.

SOUZA, G. M. *et al.* The Sugarcane Genome Challenge: Strategies for Sequencing a Highly Complex Genome. **Tropical Plant Biology**, [s. l.], v. 4, n. 3–4, p. 145–156, 2011.

SREENIVASAN, T. V.; AHLOOWALIA, B. S.; HEINZ, D. J. Cytogenetics. **Developments in Crop Science**, [s. l.], v. 11, n. C, p. 211–253, 1987.

UMEDA, M. *et al.* Detection of a major qtl related to smut disease resistance inherited from a Japanese wild sugarcane using GRAS-DI technology. **Breeding Science**, [s. l.], v. 71, n. 3, p. 365–374, 2021.

VOORRIPS, R. E. Mapchart: Software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, [s. l.], v. 93, n. 1, p. 77–78, 2002.

WANG, Z. *et al.* An autoploid-suitable polyBSA-seq strategy for screening candidate genetic markers linked to leaf blight resistance in sugarcane. **Theoretical and Applied Genetics**, [s. l.], v. 135, n. 2, p. 623–636, 2022a.

WANG, Z. *et al.* Isolating QTL controlling sugarcane leaf blight resistance using a two-way pseudo-testcross strategy. **Crop Journal**, [s. l.], n. December, 2022b. Disponible em: <https://doi.org/10.1016/j.cj.2021.11.009>.

WELHAM, S. J. *et al.* A comparison of analysis methods for late-stage variety evaluation trials. **Australian and New Zealand Journal of Statistics**, [s. l.], v. 52, n. 2, p. 125–149, 2010.

WELKER, C. A. D. *et al.* Phylogenetic analysis of *Saccharum* S.L. (Poaceae; Andropogoneae), with emphasis on the circumscription of the South American species. **American Journal of Botany**, [s. l.], v. 102, n. 2, p. 248–263, 2015.

WU, K. K. *et al.* The detection and estimation of linkage in polyploids using single-dose restriction fragments. **Theoretical and Applied Genetics**, [s. l.], v. 83, n. 3, p. 294–300, 1992.

YANG, X. *et al.* Sugarcane genome sequencing and genetic mapping. [s. l.], n. March, p. 3–34, 2018a.

YANG, X. *et al.* Sugarcane genome sequencing and genetic mapping. *Em*: [S. l.: s. n.], 2018b. p. 3–34.

YONEMARU, J.; KASUGA, S.; KAWAHIGASHI, H. QTL analysis of regrowth ability inbmrsorghum (*Sorghumbicolor*[L.] Moench) x sudangrass (*S. bicolor* subsp.drummondii)populations. **Grassland Science**, [s. l.], v. 68, n. 4, p. 310–318, 2022. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/grs.12365>.

YOU, Q. *et al.* Identifying genomic regions controlling ratoon stunting disease resistance in sugarcane (*Saccharum* spp.) clonal F1 population. **Crop Journal**, [s. l.], n. xxxx, 2020. Disponível em: <https://doi.org/10.1016/j.cj.2020.10.010>.

ZATTI BARRETO, F. Caracterização fenotípica e molecular do painel brasileiro de genótipos de cana-de-açúcar. 2016. 91 f. - Universidade Federal de São Carlos, 2016, [s. l.], 2016. Disponível em: <https://repositorio.ufscar.br/bitstream/handle/ufscar/10901/DissFZB.pdf?sequence=1&isAllowed=y>.

ZATTI BARRETO, F. *et al.* Validação de marcadores moleculares associados à resistência à ferrugem marrom em cana-de-açúcar. **Summa Phytopathol.**, [s. l.], v. 43, n. 1, p. 36–40, 2017.

ZENG, Z.-B. Precision Mapping of Quantitative Trait Loci. **Genetics**, North Carolina, p. 1457–1468, 1994. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1205924/>. Acesso em: 21 jun. 2023.

ZHANG, J. *et al.* Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. **Nature Genetics**, [s. l.], v. 50, n. 11, p. 1565–1573, 2018a.

ZHANG, J. *et al.* Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. **Nature Genetics**, [s. l.], v. 50, n. 11, p. 1565–1573, 2018b. Disponível em: <http://www.nature.com/articles/s41588-018-0237-2>.

ZHANG, X. *et al.* Characterization of meiotic chromosome behavior in the autopolyploid *Saccharum spontaneum* reveals preferential chromosome pairing without distinct DNA sequence variation. **Crop Journal**, [s. l.], 2023.