



**ESTIMAÇÃO DO COEFICIENTE DE PARENTESCO ENTRE  
ACESSOS DE CANA DO BANCO DE GERMOPLASMA DO IAC  
COM APLICAÇÕES NO MELHORAMENTO GENÉTICO DA  
CULTURA**

**ANDRÉA MORETI CARVALHO VALESAN**

**SECRETARIA DE AGRICULTURA E ABASTECIMENTO DO ESTADO DE  
SÃO PAULO  
AGÊNCIA PAULISTA DE TECNOLOGIA DOS AGRONEGÓCIOS  
INSTITUTO BIOLÓGICO  
PROGRAMA DE PÓS-GRADUAÇÃO EM SANIDADE, SEGURANÇA ALIMENTAR  
E AMBIENTAL NO AGRONEGÓCIO**

**ESTIMAÇÃO DO COEFICIENTE DE PARENTESCO ENTRE ACESSOS DE CANA  
DO BANCO DE GERMOPLASMA DO IAC COM APLICAÇÕES NO  
MELHORAMENTO GENÉTICO DA CULTURA**

**ANDRÉA MORETI CARVALHO VALESAN**

Dissertação apresentada para a obtenção do título de Mestre em Sanidade, Segurança Alimentar e Ambiental no Agronegócio. Área de concentração: Segurança Alimentar e Sanidade no Agroecossistema.

SÃO PAULO  
2022

**ESTIMAÇÃO DO COEFICIENTE DE PARENTESCO ENTRE ACESSOS DE CANA  
DO BANCO DE GERMOPLASMA DO IAC COM APLICAÇÕES NO  
MELHORAMENTO GENÉTICO DA CULTURA**

**ANDRÉA MORETI CARVALHO VALESAN**

Dissertação apresentada para a obtenção do título de  
Mestre em Sanidade, Segurança Alimentar e  
Ambiental no Agronegócio.

Área de concentração: Segurança Alimentar e  
Sanidade no Agroecossistema.

Orientadora: Dr<sup>a</sup>. Luciana Aparecida Carlini Garcia

SÃO PAULO  
2022

Eu **Andréa Moreti Carvalho Valesan**, autorizo o Instituto Biológico (IB-APTA), da Secretaria de Agricultura e Abastecimento do Estado de São Paulo, a disponibilizar gratuitamente e sem ressarcimento dos direitos autorais, o presente trabalho acadêmico de minha autoria, no portal, biblioteca digital, catálogo eletrônico ou qualquer outra plataforma eletrônica do IB para fins de leitura, estudo, pesquisa e/ou impressão pela Internet desde que citada a fonte.

Assinatura: Andréa M.C. Valesan Data 01/07/22

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**Secretaria de Agricultura e Abastecimento do Estado de São Paulo**  
**Núcleo de Informação e Documentação – IB**

---

Valesan, Andréa Moreti Carvalho.

Estimação do coeficiente de parentesco entre acessos de cana do banco de germoplasma do IAC com aplicações no melhoramento genético da cultura. / Andréa Moreti Carvalho Valesan. - São Paulo, 2022.

41 p.

doi: 10.31368/PGSSAAA.2022D.AV004

Dissertação (Mestrado). Instituto Biológico (São Paulo). Programa de Pós-Graduação.

Área de concentração: Segurança Alimentar e Sanidade no Agroecossistema.  
Linha de pesquisa: Biodiversidade: caracterização, interações, interações ecológicas em agroecossistemas.

Orientador: Luciana Aparecida Carlini Garcia.

Versão do título para o inglês: Estimation of the coefficient of relationship among sugarcane accessions of the among sugarcane accessions of the IAC's germplasm bank, applied to sugarcane plant breeding.

1. Saccharum 2. Poliploidia 3. Genealogia I. Valesani, Andréa II. Garcia, Luciana Aparecida Carlini III. Instituto Biológico (São Paulo) IV. Título.

IB/Bib/2022/004

---

## FOLHA DE APROVAÇÃO

Nome: Andréa Moreti Carvalho Valesan

Título: Estimação do coeficiente de parentesco entre acessos de cana do banco de germoplasma do IAC com aplicações no melhoramento genético da cultura.

Dissertação apresentada ao Programa de Pós-graduação em Sanidade, Segurança Alimentar e Ambiental no Agronegócio do Instituto Biológico, Agência Paulista de Tecnologia dos Agronegócios, Secretaria de Agricultura e Abastecimento do Estado de São Paulo, para a obtenção do título de Mestre em Sanidade, Segurança Alimentar e Ambiental no Agronegócio.

Aprovado em: 01/07/22

### Banca Examinadora:

Prof. Dra. Luciana Aparecida Carlini Garcia

Instituição: URPD de Piracicaba/APTA  
Regional/APTA/SAASP

Julgamento: APROVADA

Assinatura: 

Prof. Dra. Luciana Rossini Pinto

Instituição: Centro de cana/IAC/APTA

Julgamento: APROVADA

Assinatura: 

Prof. Dr. Rodrigo Rampazo Amadeu

Instituição: Bayer Crop Science - USA

Julgamento: APROVADA

Assinatura: 

## **AGRADECIMENTOS**

Primeiramente a Deus pela oportunidade de realizar este caminho e por me fortalecer nos momentos difíceis desta trajetória.

À minha orientadora, Professora Dra. Luciana Aparecida Carlini Garcia, pelos ensinamentos transmitidos, por toda paciência e dedicação ao me orientar.

Aos professores do Programa de Pós-Graduação em Sanidade, Segurança Alimentar e Ambiental no Agronegócio por todos os ensinamentos.

Aos membros da banca de Qualificação e Defesa de Mestrado pelas sugestões e correções para melhoria do trabalho.

Aos meus pais Albino Valesan e Maria M. C. Valesan por sempre me apoiarem e sempre me incentivarem na realização dos meus sonhos.

Ao meu marido Bruno Magalhães da Silva e minha filha Laura M. C. Valesan por todo apoio, carinho, incentivo e paciência nos momentos em que não pude dar atenção a vocês para poder me dedicar ao trabalho.

À minha tia Alcione Carvalho que sempre me auxiliou com carinho antes de cada apresentação do trabalho.

Aos meus colegas do mestrado por sempre estarem dispostos a ajudar e por compartilharem conhecimento que foi muito importante na realização das disciplinas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

## RESUMO

VALESAN, A.M.C. **Estimação do coeficiente de parentesco entre acessos de cana do banco de germoplasma do IAC com aplicações no melhoramento genético da cultura.** 2022. 41f. Dissertação. (Mestrado em Sanidade, Segurança Alimentar e Ambiental no Agronegócio) – Instituto Biológico, Agência Paulista de Tecnologia dos Agronegócios, Secretaria de Agricultura e Abastecimento do Estado de São Paulo, São Paulo, 2022.

O Brasil é o maior produtor mundial de cana-de-açúcar (*Saccharum spp*), cultura que tem grande importância na produção de açúcar e etanol. As variedades modernas de cana são organismos poliploides complexos, sendo *Saccharum officinarum* a espécie que mais contribui na constituição de seus genomas. A caracterização dos acessos dos bancos de germoplasma é essencial para que os programas de melhoramento sejam conduzidos de modo mais eficiente. O presente trabalho teve como objetivos: estimar os coeficientes de parentesco entre 83 acessos de um painel de cana do IAC, a partir de dados de genealogia, com uso do método adequado para organismos poliploides complexos; estimar os coeficientes de parentesco pelo método estendido de VanRaden e os índices de similaridade de Jaccard dos mesmos acessos, a partir de dados de marcadores moleculares (AFLPs e/ou SSRs); comparar os resultados obtidos com esses métodos para verificar o nível de correspondência entre eles nesse painel de genótipos de cana. Os resultados obtidos indicam ser mais vantajoso utilizar o cálculo do parentesco molecular de acordo método estendido de VanRaden, considerando os dois tipos de marcadores concomitantemente, uma vez que as estimativas de parentesco obtidas desse modo são as que mais se assemelham às estimativas obtidas a partir de dados genealógicos. Os resultados desse estudo são relevantes, uma vez que informações sobre as relações de parentesco são importantes para o estabelecimento de estratégias de hibridações em programas de melhoramento.

**Palavras-chave:** *Saccharum*. Poliploidia. Genealogia.

## ABSTRACT

VALESAN, A.M.C. **Estimation of the coefficient of relationship among sugarcane accessions of the IAC's germplasm bank, applied to sugarcane plant breeding.** 2022. 41f. Dissertação. (Mestrado em Sanidade, Segurança Alimentar e Ambiental no Agronegócio) – Instituto Biológico, Agência Paulista de Tecnologia dos Agronegócios, Secretaria de Agricultura e Abastecimento do Estado de São Paulo, São Paulo, 2022.

Brazil is the largest producer of sugarcane (*Saccharum spp*), an important crop from which sugar and ethanol are obtained. The modern sugarcane varieties are complex polyploid organisms and *Saccharum officinarum* is the species that most contributed to the composition of their genome. The description of the germplasm bank accessions is fundamental for efficiently conducting the breeding programs. The objectives of this research were: to estimate the coefficient of relationship among 83 accessions of a sugarcane panel from IAC, based on their genealogical information and on the use of a method proposed for autopolyploid species; to estimate the coefficient of relationship according to extended VanRaden method and the Jaccard similarity coefficient among these accessions based on molecular markers (AFLPs and/or SSRs); to evaluate the resemblance of the results produced by these methods applied to this sugarcane panel. According to the results, the relationship coefficients estimates produced by the extended VanRaden method, considering both molecular markers (AFLPs and SSRs), are the most similar to the relationship coefficient estimates obtained based on the genealogy. The results generated in this study are relevant since information about the relationship among genotypes are important to establish breeding strategies in breeding programs.

**Key words:** *Saccharum*. Polyploidy. Genealogy.

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO.....</b>	<b>10</b>
<b>2.</b>	<b>OBJETIVOS .....</b>	<b>12</b>
2.1	OBJETIVOS GERAIS.....	12
2.2	PERSPECTIVAS FUTURAS.....	12
<b>3.</b>	<b>REVISÃO DE LITERATURA .....</b>	<b>14</b>
3.1	PRODUÇÃO DE CANA-DE-AÇÚCAR.....	14
3.2	ORIGEM DA CANA-DE-AÇÚCAR.....	15
3.3	MELHORAMENTO GENÉTICO DA CANA-DE-AÇÚCAR.....	16
<b>4.</b>	<b>MATERIAL E MÉTODOS .....</b>	<b>22</b>
<b>5.</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>26</b>
<b>6.</b>	<b>CONCLUSÕES .....</b>	<b>34</b>
<b>7.</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>35</b>

## 1. INTRODUÇÃO

A cana-de-açúcar (*Saccharum* spp.) é uma gramínea tropical do tipo C4, com grande capacidade de acumular sacarose. Os principais produtos obtidos a partir dela são o açúcar e o etanol (NEVES; CONEJERO, 2010a). Originária da Nova Guiné foi difundida pelo mundo por meio da migração humana. Foi introduzida no Brasil em 1520 e se tornou um dos cultivos mais importantes do país (NOVA CANA, 2021).

Atualmente, as variedades de cana são organismos interespecíficos autopoliploides complexos com predomínio do genoma da espécie *S. officinarum*, muito usada em programas de melhoramento por apresentar elevada concentração de sacarose e baixo teor de fibra. Outras espécies, em menores proporções, também contribuem na hibridação, principalmente, *S. spontaneum*, que por apresentar maior quantidade de fibras e contribuir para que as variedades sejam mais vigorosas e resistentes a estresses bióticos e abióticos (BREMER, 1961; SCARPARI; BEAUCLAIR, 2008).

Embora a base genética da cana seja estreita, isso vem sendo, de certa forma, contornado devido ao elevado nível de ploidia das cultivares. Além disso, a introgressão de alelos de outras espécies do complexo *Saccharum* vem aumentando a variabilidade genética da cultura e reduzindo a vulnerabilidade da mesma a situações adversas (AITKEN; MCNEIL, 2010).

De modo geral, os programas de melhoramento genético da cana têm como objetivo obter variedades com bom perfilhamento, alta produtividade, baixa porcentagem de fibra, resistentes a pragas e doenças, entre outras características. Caracterizar os acessos dos bancos de germoplasma da cultura, conhecer a estrutura genética desses bancos e o parentesco entre os acessos é de suma importância para os programas de melhoramento (SCARPARI; BEAUCLAIR, 2008). Nesse contexto, obter as estimativas do coeficiente de parentesco entre os genótipos, com base em informações genealógicas ou de marcadores moleculares, dependendo do tipo dos dados que se têm disponíveis, é de extrema valia para o estabelecimento de estratégias do programa de melhoramento, especialmente no que diz respeito à fase de hibridações.

Há relativamente poucos estudos em que se estimam os coeficientes de parentesco de cana-de-açúcar utilizando métodos apropriados para espécies autopoliploides complexas como no caso da referida cultura. É nesse cenário que o presente estudo se insere, tendo por objetivos estimar, entre os pares de genótipos avaliados, os coeficientes de parentesco a partir de dados genealógicos e moleculares, bem como os coeficientes de

similaridade Jaccard também com base em dados de moleculares. Os resultados obtidos com diferentes métodos serão comparados e as informações geradas serão disponibilizadas para uso em programas de melhoramento de cana.

## 2. OBJETIVOS

### 2.1 OBJETIVOS GERAIS

a) Estimar os coeficientes de parentesco entre 83 genótipos de cana do banco ativo de germoplasma (BAG) do Centro de Cana do IAC, a partir de dados de genealogia desses acessos e com uso do método proposto por Kerr et al., (2012), indicado para os casos de espécies autoploidoides complexas e que apresentem citótipos distintos;

b) Estimar os coeficientes de parentesco entre esses mesmos genótipos, a partir de dados de marcadores moleculares, com uso do método de VanRaden (2008) estendido por Ashraf et al. (2016). As estimativas foram obtidas com base em marcadores AFLPs e SSRs conjuntamente e separadamente;

c) Estimar os coeficientes de similaridade de Jaccard entre pares desses genótipos, a partir de dados de marcadores moleculares, considerando marcadores AFLPs e SSRs concomitante e isoladamente;

d) Comparar os resultados obtidos com os diferentes métodos, verificando o nível de correspondência entre eles e se, dentre os métodos empregados, se algum se mostrou mais promissor na avaliação das relações genéticas em cana.

### 2.2 PERSPECTIVAS FUTURAS

a) Disponibilizar as informações resultantes desse trabalho para obtenção de estimativas do valor genético dos genótipos avaliados, no que se refere a diversos caracteres de interesse (produtividade, resistências a doenças e pragas, tolerância a estresses abióticos, etc.), que poderão ser avaliados fenotipicamente;

b) Usar as informações do item anterior em estudos de mapeamento associativo, seleção genômica, entre outros, com a finalidade de tornar o processo de melhoramento genético e obtenção de variedades mais eficientes e dinâmicas;

c) Por meio das futuras hibridações a serem programadas, considerando as informações obtidas nessa pesquisa, e

d) Utilizar as informações geradas nesta pesquisa no direcionamento de hibridações em cana, com o objetivo de: ampliar a base genética da cultura, gerando variabilidade genética a ser explorada pelos programas de melhoramento; obter, de maneira eficaz, novas cultivares de cana, com características que atendam às demandas do setor

canavieiro e que levem, possivelmente, a reduções no custo de produção e no impacto ambiental.

### 3. REVISÃO DE LITERATURA

#### 3.1 PRODUÇÃO DE CANA-DE-AÇÚCAR

O Brasil se destaca mundialmente na produção de cana-de-açúcar e, conseqüentemente, na produção de açúcar e o etanol, sendo este uma importante alternativa aos combustíveis fósseis. O consumo de álcool tem sido superior ao de gasolina no país (MAPA, 2020) e essa substituição, além de economicamente viável, é vantajosa, pois os biocombustíveis são produzidos a partir de matéria-prima renovável e são menos poluentes que os combustíveis fósseis.

Na safra de 2022/23, estima-se que será colhido 596,1 milhões de toneladas de cana, o que representa um aumento de 1,9% em relação à produção da safra anterior. A partir da cana-de-açúcar colhida, deverão ser produzidas cerca de 40.282,3 mil toneladas de açúcar além de 24,81 bilhões de litros de etanol total, configurando aumentos de 14,9% e 7,4%, respectivamente, em relação à safra 2020/21 (CONAB, 2022).

Como mencionado por Oriolo (2016), os programas de melhoramento genético da cana-de-açúcar, tanto no Brasil como em todo mundo, buscam obter genótipos altamente produtivos, com elevado teor de sacarose e baixo teor de fibras, com resistência a estresses bióticos e abióticos. Buscando agregar as características de interesse, as cultivares atuais são híbridos interespecíficos entre *S. officinarum*, espécie que apresenta elevada produtividade e baixo teor de fibra, e outras espécies, como *S. spontaneum*, que se caracteriza por apresentar alto teor de fibra, vigor e resistência/tolerância a vários estresses bióticos e abióticos (SCARPARI; BEAUCLAIR, 2008). Também buscam-se cultivares que apresentem perfil voltado à produção de etanol celulósico, que são caracterizadas por apresentarem, entre outros, elevado teor de fibra e vigor (ORIOLO, 2016).

Dado esse contexto, o melhoramento genético pode contribuir, de modo eficiente, por meio da obtenção de cultivares de cana com alto desempenho e com características favoráveis, para atender às demandas do mercado (ZAN et al., 2020), contribuindo, para elevar a produção de açúcar e etanol, com menor custo de produção e redução de impactos ambientais.

### 3.2 ORIGEM DA CANA-DE-AÇÚCAR

A cana-de-açúcar (*Saccharum spp.*) pertence ao gênero *Saccharum*, tribo Andropogonae, família Poaceae, ordem Cyperales, class Liliopsida, divisão Magnoliophyta, reino Plantae (CRONQUIST, 1981). Trata-se de uma gramínea tropical, com metabolismo do tipo C4, bastante eficiente em converter CO<sub>2</sub>, e em acumular quantidade elevada de sacarose nos colmos, a qual, após processamento, é convertida principalmente em açúcares como o refinado e mascavo, por exemplo, (NEVES; CONEJERO, 2010a).

A cana é originária da Nova Guiné, tendo se difundido pelo mundo por meio da migração humana. Foi introduzida ao Brasil em 1520 após a chegada dos portugueses, tornando-se, desde então, uma das culturas mais importantes do país (NOVA CANA, 2021).

Cheavegatti-Gianotto, et al., (2011), mencionam o fato de que não se conhece ao certo o número de espécies pertencentes ao gênero *Saccharum*. É bastante comum, contudo, utilizar a classificação segundo a qual há quatro espécies cultivadas (*S. officinarum*, *S. barberi*, *S. sinense* e *S. edule*) e duas selvagens (*S. spontaneum* e *S. robustum*), que compõem o referido gênero, o qual juntamente com espécies dos gêneros correlatos *Narenga*, *Sclerostachya*, *Erianthus* e *Miscanthus*, resultam o Complexo *Saccharum* (MUKERJEE, 1957; DANIELS; DANIELS, 1975; ROACH; DANIELS, 1987). Segundo Aitken; McNeil (2010), mesmo que essa não seja a classificação mais correta, tem sido utilizada com sucesso para fins de manutenção e uso dos acessos que compõem a Coleção Mundial de Cana e Espécies Correlatas.

Segundo Bremer (1932), a variedade Crioula é resultante da hibridação interespecífica, ocorrida naturalmente entre *S. barberi* e *S. officinarum*. Foi introduzida no Brasil e, depois de dois séculos e meio sendo cultivada nas Américas, foi substituída por variedades com maior teor de açúcar, selecionadas a partir da espécie *S. officinarum* (LANDELL; SILVA, 2004). Porém, como esta espécie é vulnerável a várias doenças, no final do século XIX, cruzamentos interespecíficos artificiais entre *S. officinarum* e outras espécies do gênero passaram a ser realizados, visando obter genótipos que apresentassem elevado teor de sacarose, mas também resistência/tolerância a doenças e, com isso, iniciou-se o melhoramento genético da cultura (D'HONT, 1996; LANDELL; SILVA, 2004).

As variedades atuais de cana-de-açúcar são definidas como organismos interespecíficos, autopoliplóides complexos, incluindo aneuploidia (GRIVET et al., 1996; HOARAU et al., 2001; GARCIA et al., 2013), com a presença de diferentes níveis de ploidia entre grupos de ligação num mesmo indivíduo (GARCIA et al., 2013). O processo de

autoploidização consiste num importante mecanismo dentro do processo evolutivo, sendo muito frequente em plantas (ADAM; WENDEL, 2005; BURGESS, 2015).

Observe-se que, na literatura, em cana, muitas vezes o uso dos termos variedade, cultivar e clones se confundem. Segundo Fehr (1987), o termo variedade é utilizado para designar uma subdivisão taxonômica de espécie, enquanto que cultivar se refere a uma variedade cultivada de uma espécie. Ou seja, uma espécie pode apresentar diversas variedades e, dentre elas, há as variedades que são utilizadas para cultivo, que são as cultivares. Ainda segundo esse autor, na literatura, o termo clone se refere ao indivíduo resultante de reprodução assexuada e que se caracteriza por ser uma cópia geneticamente idêntica de seu genitor. Em cana, existem variedades que ocorrem naturalmente e variedades que foram resultantes de cruzamentos realizados visando obtenção de cultivares. A cana é uma cultura que se reproduz por meio de cruzamentos e também vegetativamente (produzindo clones). Atualmente, há um grande número de cultivares de cana, que são utilizadas pelos produtores e que são multiplicadas e comercializadas como clones, mantendo a uniformidade da cultura. Dessa maneira, resumidamente, nos programas de melhoramento, são realizados cruzamentos para obtenção de novos genótipos, que são multiplicados de modo clonal e passam por diversas fases de seleção, podendo resultar em novas cultivares de cana (também chamadas de variedades comerciais de cana), que são multiplicadas vegetativamente (reprodução clonal).

### 3.3 MELHORAMENTO GENÉTICO DE CANA-DE-AÇÚCAR

As cultivares de cana moderna resultam do trabalho dos melhoristas durante décadas. São caracterizadas por apresentarem alta produtividade, resistência/tolerância a estresses bióticos e abióticos, alto perfilhamento, ampla adaptação ambiental, entre outros.

A espécie *S. officinarum*, embora suscetível a várias doenças, pragas e estresses abióticos, continua sendo predominantemente empregada nas hibridações em programas de melhoramento genético de cana, o que se deve ao fato de ela apresentar alta concentração de açúcar e reduzido teor de fibras (BREMER, 1961; SCARPARI; BEAUCLAIR, 2008), características importantes para a indústria sucroalcooleira. A espécie *S. spontaneum*, em menores proporções, também é usada nessas hibridações, pois, segundo esses mesmos autores, conferem às variedades comerciais características como rusticidade, vigor, maiores capacidade de rebrota e perfilhamento, além de resistência a diferentes tipos de estresses. Como exemplo, pode se mencionar a pesquisa de D'Hont et al. (1996), que ao estudarem a

composição genética da variedade R570, observaram que cerca de 80% de seu genoma é proveniente de *S. officinarum*, 10% de *S. spontaneum* e 10% da recombinação entre o genoma dessas duas espécies. Além das duas espécies mencionadas, *S. sinense*, *S. robustum* e *S. barberi*, também têm contribuições, ainda que menos expressivas, na formação genômica das variedades comerciais brasileiras (MATSUOKA et al., 1999).

A base genética da cana cultivada é estreita, uma vez que poucos híbridos interespecíficos *S. officinarum* e *S. spontaneum*, deram origem a diversas variedades comerciais e espécies relacionadas (ARCENEUX, 1967). Esse problema vem sendo parcialmente contornado pelo elevado nível de ploidia das cultivares, sendo que, quando são realizados cruzamentos biparentais ou poliparentais, entre genitores, que são variedades ou clones comerciais, podem ocorrer muitas recombinações entre diferentes alelos em cada loco. Como mencionado por Aitken; McNeil (2010), a introgressão de alelos, advindos de outras espécies do Complexo *Saccharum*, colabora para aumentar a variabilidade genética presente na cultura, o que traz vantagens para os melhoristas e também reduz a vulnerabilidade da mesma.

Como mencionado anteriormente, o início do melhoramento da cultura data do final do século XIX. Há programas de melhoramento de cana em diversas localidades do mundo, como na Índia, nos EUA, na Austrália, na Argentina, no México, entre outros. Mais especificamente, no Brasil existem três programas de melhoramento genético da cultura: o da Rede Interuniversitária de Desenvolvimento do setor Sucroalcooleiro RIDESA (Variedades RB), o do centro de Tecnologia Canaveira – CTC e o Programa Cana do Instituto Agrônomo IAC, sendo dois deles públicos (RIDESA e IAC) e um privado (CTC).

Embora cada programa tenha suas peculiaridades, de modo geral, as etapas dos programas de melhoramento de cana, segundo Landell; Silva (2004) e Landell; Bressiani (2008) são: a) fase de hibridação, por meio de realização de cruzamentos bi ou poliparentais, entre variedades comerciais e espécies relacionadas. Nessa fase, ocorre a ampliação da variabilidade genética e também é possível a transferência de alelos favoráveis de espécies relacionadas para os descendentes resultantes dos cruzamentos; b) várias fases de avaliação e seleção de híbridos, em experimentos realizados em diversos ambientes de produção. Os clones promissores são selecionados e propagados vegetativamente durante o processo de avaliação e seleção, podendo resultar em novas cultivares.

Os programas de melhoramento de cana-de-açúcar visam obter variedades com alta produtividade, alto teor de açúcar e baixa porcentagem de fibras, entre outros, visando produção de açúcar e álcool. Por isso, a contribuição de *S. officinarum* para o genoma das

cultivares de cana é elevada (BREMER, 1961); (SCARPARI; BEAUCLAIR, 2008); (D'HONT et al., 1996; 1998). Contudo, nos últimos anos e em menor proporção, também tem havido a demanda por genótipos de cana com perfil energético, caracterizados por apresentarem elevada concentração de biomassa, para produzir etanol celulósico, com concomitante geração de energia elétrica (LANDELL; BRESSIANI, 2008).

Apesar de o Programa Cana do Centro de cana do Instituto Agrônômico (IAC/APTA/SAA-SP) atuar nessas duas frentes do melhoramento, o foco principal é obtenção de variedades de cana-de-açúcar e não de cana energia. O centro de Cana do IAC vem investindo, ao longo dos anos, em organizar uma Coleção Pública de Germoplasma, que seja representativa do Complexo *Saccharum*. No banco de germoplasma de cana, há acessos de variedades comerciais, de clones em fase de avaliação e de outras espécies tanto do gênero *Saccharum*, quanto de gêneros relacionados como *Miscanthus* e *Erianthus*. A partir de 2013, esse banco de germoplasma passou a abrigar uma cópia da Coleção Mundial de cana e Espécies Relacionadas, tornando-se a terceira instituição mantenedora de tal coleção (ORIOLO, 2016).

O melhoramento genético de uma cultura depende da variabilidade genética disponível aos melhoristas. Contudo, para que a mesma seja eficientemente utilizada, é preciso que esteja descrita. Por isso, acessos do Banco de Germoplasma de Cana do IAC, vêm passando por um processo de caracterização no decorrer dos últimos anos: a) avaliações de características agrônômicas e tecnológicas de genótipos a partir de experimentos com repetições em campo; b) “screening” para diversas doenças, pragas que acometem a cultura; c) avaliações citológicas; d) obtenção dos perfis moleculares (“fingerprints”) dos acessos, com emprego de marcadores SSRs.

Destaque-se que a genotipagem dos acessos por meio dos marcadores SSRs têm sido realizada com sucesso no Programa Cana do IAC, e esse protocolo tem atendido a diversas funções, tais como: acompanhar e identificar os genótipos em fases de avaliação no campo, como em Jannoo et al. (2001) e Pan et al. (2003; 2010); descrever a estrutura genética do banco de germoplasma e, com isso, permitir o estabelecimento de coleções nucleares, auxiliando, inclusive, no direcionamento de hibridações. Embora, atualmente, os SNPs (Single nucleotide polymorphism) sejam marcadores mais adequados para diferentes tipos de estudo em Genética, inclusive por permitirem a dosagem de alelos (GARCIA et al., 2013), para os fins descritos aqui, o emprego de marcadores SSRs ainda continua bastante satisfatório. Como verificado por Cordeiro et al. (2003) SSRs são uma ferramenta bastante eficiente na identificação da constituição genética das cultivares modernas de cana, quando

comparados aos marcadores RAPD (Random Amplified Polymorphic DNA); AFLP e RFLP (Restriction Fragment Length Polymorphism). Isso, no entanto, não invalida o uso de outros marcadores para o mesmo fim. Inclusive, marcadores AFLP continuam sendo utilizados também na caracterização molecular do banco de germoplasma de cana do IAC.

A partir de dados dos marcadores, podem ser calculadas, entre pares de indivíduos, suas medidas de similaridades, como por exemplo, Jaccard (JACCARD, 1901), Coincidência Simples (SOKAL; MICHENER, 1958), Dice (DICE, 1945), ou de distâncias, como a Euclideana e Euclideana quadrada, a de Nei - 72 (NEI, 1972), a de Rogers (ROGERS, 1972).

O conhecimento da similaridade e da estrutura genética entre acessos de um banco de germoplasma é muito importante inclusive na definição das estratégias do programa de melhoramento a ele vinculado. Nas últimas décadas, diversos tipos de marcadores moleculares foram desenvolvidos e, no caso de cana-de-açúcar, há uma grande quantidade de trabalhos que avaliaram a similaridade/diversidade entre indivíduos, inclusive realizados no Centro de Cana do IAC. Podem-se citar, por exemplo, pesquisas em cana com RFLPs (LU et al., 1994), RAPDs (HARVEY et al., 1994; KAWAR et al., 2009; KHAN et al.; 2010), AFLPs (BESSE et al. 1998; LIMA et al., 2002; AITKEN et al., 2006; SELVI et al., 2006; SELVI et al., 2008; CRESTE et al, 2010), SSRs (CORDEIRO et al. 2003; PINTO et al., 2004; PINTO et al., 2006; OLIVEIRA et al., 2009; DUARTE FILHO et al., 2010; SINGH et al., 2010; CRESTE et al., 2010; MARCONI et al., 2011, MANECHINI et al., 2018), TRAPs - "Target Region Amplification Polymorphism" (ALWALA et al., 2006; CRESTE et al., 2010; SINGH et al., 2017); DArTs - "Diversity Arrays Technology" (HELLER-USZYNSKA et al., 2011); SNPs - "Single Nucleotide Polymorphisms" (DEVARUMATH et al., 2013; BELLO et al., 2019), entre outros.

Além da similaridade/dissimilaridade ou distância entre indivíduos, quando possível, é interessante também obter as estimativas do coeficiente de parentesco entre eles, que podem ser calculadas tanto a partir do conhecimento da genealogia desses indivíduos, quanto a partir de informações obtidas com base em marcadores moleculares. A definição do coeficiente de parentesco foi proposta por Wright (1922). Segundo o autor considerando-se organismos diploides, o coeficiente de parentesco é a probabilidade de que dois alelos, amostrados de modo aleatório, a partir de dois indivíduos, sejam idênticos por descendência (ibd - "identical by descent").

No caso de organismos aloploidos, pode-se aplicar essa mesma ideia, uma vez que a herança é do tipo dissômica na meiose. Por outro lado, em organismos

autopoliploides, deve-se considerar que, na meiose, a herança é polissômica e, o cálculo do coeficiente de parentesco deve ser adaptado para essa circunstância (AMADEU et al., 2020). Nesse contexto, Kerr et al. (2012) propuseram um algoritmo para calcular o parentesco entre indivíduos autopoliploides, levando em consideração a herança polissômica e a possibilidade de ocorrência de redução dupla na meiose. Como previamente mencionado, o parentesco entre indivíduos também pode ser estimado a partir de dados de marcadores moleculares, sendo que, quando indivíduos apresentam um mesmo alelo para um loco marcador, diz-se que tais indivíduos apresentam alelo idêntico em estado (ibs - "identical by state"). Segundo Lynch; Ritland (1999), o cálculo do parentesco entre pares de indivíduos tem contribuído para melhorar a qualidade da predição genômica tanto em estudos de associação, como de seleção genômica.

O parentesco com base em dados moleculares pode ser calculado em diplóides, a partir de locos bialélicos, com os procedimentos propostos por VanRaden (2008) e por Yang et al. (2010), ou multialélicos de acordo com método de Lynch; Ritland, (1999). No caso de organismos autopoliploides, há os métodos propostos por Huang et al.(2014, 2015), para locos multialélicos e por Ashraf et al. (2016) e por Slater et al. (2016) para marcadores bialélicos.

Por meio de simulações Amadeu (2020), calculou o parentesco entre genótipos, a partir de dados moleculares, utilizando oito métodos de estimação do parentesco entre pares de indivíduos, desenvolvidos para autopoliploides, sendo cinco deles apresentados por Huang et al. (2014, 2015), um por Ashraf et al. (2016) e dois propostos por Slater et al. (2016). Depois, comparou esses resultados com os obtidos a partir de genealogia, de acordo com o método de Kerr et al. (2012), que foi implementado por Amadeu et al. (2016) no pacote computacional AGHmatrix V 2.0. Concluíram que, no momento, é recomendável utilizar mais de 100 marcadores bialélicos do tipo SNP, com dosagem de alelos, e o método de VanRaden (2008),estendido por Ashraf et al. (2016), para obter estimativas do coeficiente de parentesco em autotetraploides.

O programa Cana do IAC possui um conjunto de 83 acessos de cana, incluindo cultivares e clones, cujas genealogias são conhecidas, e que foram genotipados, por meio de marcadores moleculares AFLPs e SSRs, por Gonçalves (2012). Com o intuito de realizar mapeamento associativo entre esses marcadores e ferrugem marrom em cana, Gonçalves (2012) estudou a estrutura genética desse material, a partir dos dados de marcadores moleculares, calculou as dissimilaridades complementares ao coeficiente de Jaccard entre os pares de genótipos e os agrupou, com base nessas estimativas de dissimilaridade, com o

emprego do método de Neighbor-Joining e pela análise de coordenadas principais. Com parte dos dados dos marcadores e considerando marcas com fortemente ligadas entre si, Gonçalves (2012) fez também análise da estrutura genética por método bayesiano.

No presente trabalho, os dados genealógicos desses 83 genótipos, bem como suas avaliações moleculares obtida por Gonçalves (2012), foram utilizados para estimar o parentesco entre os genótipos, baseado em genealogias e em marcadores moleculares, utilizando, respectivamente, os métodos de Kerr (2012) e de VanRaden (2008) estendido por Ashraf et al. (2016). Foram estimados os coeficientes de similaridade de Jaccard entre os pares de genótipos também a partir dos dados de marcadores. Tanto o parentesco molecular, quanto os coeficientes de Jaccard foram obtidos considerando os dados de AFLPs e SSRs conjunta e separadamente. Os resultados obtidos com as diferentes abordagens foram comparados entre si.

#### 4. MATERIAL E MÉTODOS

O material utilizado nessa pesquisa é constituído por 83 genótipos do BAG do IAC, dentre os quais 32 são variedades e 51 são clones (Tabela 1). Com o intuito de realizar o mapeamento associativo para ferrugem marrom (*Puccinia melanocephala*) nesse painel de genótipos, Gonçalves (2012), avaliou a resposta desse material ao fungo, em dois anos (2011 e 2012), em experimento instalado no Centro de Cana do IAC, em Ribeirão Preto, SP.

Além da fenotipagem, Gonçalves (2012) realizou a genotipagem do painel por meio de marcadores moleculares dos tipos AFLP (amplified fragment length polymorphism) e microssatélites (SSR - single sequence repeat), estes obtidos a partir de bibliotecas genômicas (gSSR) (CORDEIRO et al., 1999; CORDEIRO et al., 2000), ou do banco de sequências expressas (EST-SSR) do SUCEST (PINTO et al., 2004; OLIVEIRA et al., 2009; JORDÃO JR., 2009). Todos os protocolos utilizados para genotipagem estão descritos em Gonçalves (2012) e foram baseados no trabalho de Al-Janabi et al. (1999), para extração de DNA, em Schuelke (2000), Creste et al. (2001), Garcia et al. (2006), para obtenção dos SSRs, e em Vos et al. (1995) para os AFLPs.

Dos 83 genótipos avaliados, 61 têm os genitores masculino e feminino conhecidos (cruzamentos biparentais) e 22 foram obtidos de policruzamentos, sendo que, nesse caso, somente o genitor feminino é identificado (GONÇALVES, 2012). Esse mesmo autor listou os genitores dos 83 materiais, a partir de pesquisa de literatura e do programa computacional CAIANA do Centro de Cana do IAC, que é um banco de dados dos acessos do banco de germoplasma do referido Centro. Na presente pesquisa, foram utilizados dados até os tataravôs dos clones e variedades sob avaliação, fornecidos pelo CAIANA, quem tem, continuamente, suas informações sendo complementadas e atualizadas. Essas genealogias foram ordenadas das variedades e clones mais antigos para os mais novos para a realização dos cálculos. Note-se que, no CAIANA, nem todas as genealogias estão completas até a geração de tataravós, devido à falta de informação disponível.

Tabela 1. Genótipos de cana-de-açúcar considerados neste trabalho.

VARIEDADES		CLONES		
IAC52-150	IAC862480	IACSP973331	IACSP98-5046	IACSP96-7569
IAC86-2210	IAC91-1099	IACSP97-3384	IACSP98-6202	IACSP97-2000
IAC87-3396	IAC91-2195	IACSP97-3391	IACSP98-7043	IACSP97-2023
IACSP94-2101	IAC91-2218	IACSP97-3406	IACSP96-1005	IACSP97-2020
IACSP94-4004	IACSP93-2060	IACSP97-4039	IACSP99-1305	IACSP97-2028
IACSP95-1218	IACSP93-3046	IACSP97-6682	IACSP96-1066	IACSP97-2053
IACSP95-3028	IACSP93-6006	IACSP97-7018	IACSP99-1308	IACSP97-2055
IACSP95-5000	IACSP94-2094	IACSP97-7065	IACSP99-3009	IACSP97-2084
IACSP96-2042	IACSP95-5094	IACSP97-7074	IACSP96-2019	IACSP97-2098
POJ2878	IACSP96-3060	IACSP97-7077	IACSP99-3085	IACSP97-2109
R570	RB72454	IACSP97-7543	IACSP99-4007	IACSP97-3313
RB855156	RB835486	IACSP98-2030	IACSP96-3048	IACSP95-2078
RB855536	RB867515	IACSP98-2053	IACSP99-4008	IACSP96-2008
SP79-1011	RB925211	IACSP98-3020	IACSP99-4010	IACSP 96-3056
SP80-3280	SP70-1143	IACSP98-3021	IACSP99-4011	IACSP96-3076
CTC9	SP86-42	IACSP98-3099	IACSP99-4013	IACSP96-7586
		IACSP98-5010	IACSP96-7506	IACSP02-2074

Fonte: GONÇALVES (2012).

Além das informações genealógicas, 922 marcas polimórficas (AFLPs e SSRs), obtidas por Gonçalves (2012), foram utilizadas para estimar os coeficientes de parentesco molecular, par a par, entre os 83 genótipos avaliados, utilizando o método de VanRaden (2008) para marcadores bialélicos, estendido por Ashraf et al. (2016). Também foram calculadas as similaridades de acordo com o coeficiente de Jaccard (JACCARD, 1908) entre os pares de genótipos. Todas essas análises foram realizadas considerando os marcadores AFLPs e SSRs concomitante e separadamente. Deve-se ressaltar que, embora os

marcadores SSRs sejam codominantes, em cana, por se tratar de organismo autoploiploide, são lidos como marcadores dominantes, considerando-se a presença ou ausência de cada banda e para cada alelo em cada locus polimórficos, como é feito com os marcadores AFLPs, que são originalmente do tipo dominante.

A matriz de parentesco entre os mesmos genótipos, com base nos dados de genealogia, foi obtida por meio do método proposto por Kerr et al. (2012), considerando ploidia igual a oito e ausência de redução dupla. Para tanto, foi utilizado o pacote *AGHmatrix* (AMADEU et al., 2016), desenvolvido para o ambiente R (R CORE TEAM, 2020).

Os dados moleculares foram dispostos em uma matriz em que os genótipos foram representados nas linhas e os marcadores nas colunas. Os dados da matriz foram compostos por zeros (ausência da banda de um dado marcador para um dado genótipo), uns (presença da banda de um dado marcador para um dado genótipo) e -9 (para dados faltantes). Foram produzidas três matrizes: a) All - uma matriz com todos os marcadores; b) AFLPs - uma matriz somente com os marcadores AFLPs; c) SSRs - uma matriz somente com marcadores SSRs. Com base em cada uma das três matrizes, foram obtidas as estimativas dos parentescos moleculares e dos coeficientes de Jaccard entre pares de genótipos, gerando, desse modo, seis conjuntos de resultados diferentes, baseados em marcadores moleculares.

Os coeficientes de parentesco molecular foram calculados de acordo com o método de VanRaden (2008) estendido por Ashraf et al. (2016), com uso do pacote *AGHMatrix*, como mencionado previamente, enquanto os coeficientes de dissimilaridade de Jaccard foram estimados com uso da função *dist* do pacote *stats* (R CORE TEAM, 2020), sendo posteriormente transformados em coeficientes de similaridade de Jaccard, calculados como complemento da dissimilaridade.

É importante mencionar que as estimativas dos coeficientes de Jaccard entre os genótipos estudados nessa pesquisa já foram previamente obtidas por Gonçalves (2012), considerando todas as marcas (AFLPs e SSRs) conjuntamente. No entanto, esse cálculo foi repetido no presente trabalho para possibilitar a comparação das matrizes de parença obtidas de diferentes maneiras, a saber, com base em informações genealógicas, no parentesco molecular e no coeficiente de similaridade de Jaccard, considerando AFLPs e/ou SSRs, como descrito anteriormente. Gonçalves (2012) visou estimar a estrutura populacional desses genótipos para utilizar as informações no contexto de mapeamento associativo de QTLs, enquanto no presente trabalho, o foco é na comparação de diferentes métodos de estimação das medidas de parença entre os genótipos.

As sete matrizes geradas (parentesco (A), parentesco molecular considerando os dois marcadores conjuntamente (G\_All), somente os AFLPs (G\_AFLPs) e somente os SSRs (G\_SSRs) e as matrizes com os coeficientes de Jaccard, considerando os dois marcadores concomitantemente (J\_All) e somente os AFLPs (J\_AFLPs) e somente os SSRs (J\_SSRs)) foram comparadas entre si por meio de correlações de Pearson e a significância dessas correlações foi testada com uso do teste de Mantel (MANTEL, 1967). Para tanto, foi utilizada a função *mantel* do pacote *vegan* (OKSANEN et al. 2022), com 9999 permutações.

Foram obtidos os gráficos das correlações entre as matrizes, com uso da função *chart.Correlation*, do pacote *PerformanceAnalytics* (PETERSON; CARL, 2020) do R (R CORE TEAM, 2020). Foram feitos gráficos do tipo “heatmap”, utilizando-se a função *heatmap* do pacote *stats* (R CORE TEAM, 2020).

Todas as análises foram realizadas no ambiente R (R CORE TEAM, 2020), utilizando a interface do RStudio (RSTUDIO TEAM, 2021).

## 5. RESULTADOS E DISCUSSÃO

Na presente pesquisa, o nível de ploidia considerado para realização dos cálculos dos parentescos, a partir de dados genealógicos, foi oito. Foram observadas estimativas de coeficientes de parentesco variando de zero, para vários pares de genótipos, a 1,5144, no caso da estimativa do parentesco do genótipo IACSP97-3331 com ele mesmo (Figura 1). O par de genótipos mais próximos entre si (1,0155), de acordo com as informações genealógicas, foi RB83-5486 e IACSP97-3331.

As estimativas de parentescos do genótipo IAC86-2480 com todos os demais genótipos foi zero. Isso ocorreu porque, em sua genealogia, está identificado somente um dos genitores (US71399), que não se encontra na genealogia de nenhum dos outros genótipos avaliados na presente pesquisa.

No caso das estimativas dos coeficientes de parentesco entre pares de genótipos, obtidas a partir de dados de marcadores moleculares, os valores variaram de -0,2436 (POJ2878 x SP70-1143) a 0,3869 (IACSP97-2053 x IACSP97-7065), considerando ambos os marcadores concomitantemente, de -0,2597 (POJ2878 x SP70-1143) a 0,4089 (RB85-5156 x IACSP95-3028), considerando somente os AFLPs e de -0,2633 (SP80-3280 x IAC87-3396) a 0,7340 (IACSP99-4007 x IACSP99-4008), levando em conta apenas os SSRs (Figuras 2 a 4). Não houve grande variação nas amplitudes das estimativas, considerando AFLPs e/ou SSRs, embora, aparentemente, os resultados referentes ao parentesco molecular quando obtidos a partir dos AFLPs sejam um pouco mais parecidos com os obtidos a partir de ambos os marcadores em conjunto, mesmo sendo o número de marcadores AFLP (427) menor que o de SSR (495). Os maiores valores estimados de coeficiente de parentesco molecular foram 1,0976 (AFLPs e SSRs conjuntamente) e 1,1174 (AFLPs), ambos para o genótipo IACSP96-7506 com ele mesmo. Esse genótipo possui em sua genealogia duas repetições do genótipo NA5679 (um bisavô e um trisavô) por parte de um dos genitores, o que pode justificar os elevados valores das estimativas de parentesco. No caso dos marcadores SSRs, a estimativa foi de 1,1650 para o genótipo IACSP97-3331 com ele mesmo, coincidindo com o observado no caso do parentesco baseado em genealogias, ressaltando-se que esse genótipo é oriundo do cruzamento de um genitor com ele mesmo, ou seja, o genitor masculino e o feminino são coincidentes.

Observe-se que, os genótipos POJ2878 e SP70-1143, apresentaram menores valores de estimativas do coeficiente de parentesco molecular (considerando ambos os marcadores, ou somente os AFLPs) são aparentados, sendo que POJ2364 é bisavô de SP70-

1143 e um dos genitores de POJ2878. Embora o valor das estimativas de parentesco sejam menores que zero, isso ocorre porque os cálculos são realizados em relação ao valor médio dos parentescos que é zero, nesse método. Isso significa, portanto, que esse dois genótipos são menos aparentados entre si que os demais pares de genótipos. SP80-3280 e IAC87-3396, os genótipos menos aparentados, com base em SSRs, têm em comum, em sua genealogia, o POJ2364, que é três vezes tataravô do primeiro e uma vez trisavô e uma vez tataravô do segundo (Figuras 2 a 4).

As estimativas de similaridade de Jaccard entre todos os pares de genótipos variaram de 0,3321 (RB85-5196 x IACSP94-2101) a 0,6118 (IACSP97-2020 x IACSP97-7074), considerando os dois marcadores conjuntamente, de 0,2780 (IAC91-2195 x IACSP96-3076) a 0,6201 (IACSP97-2020 x IACSP97-7074), somente com base em AFLPs e de 0,3199 (IACSP96-7506 X IACSP97-7018) a 0,6500 (SP79-1011 x IACSP93-3046) considerando somente os SSRs. Além disso, como esperado de acordo com o coeficiente de Jaccard, as estimativas das similaridades de cada genótipo com ele mesmo foi 1.

RB85-5196 e IACSP94-2101, que apresentaram a menor estimativa do índice de Jaccard (AFLPs e SSRs concomitantemente), têm em comum os genótipos POJ2878, como tataravô e trisavô, respectivamente, e POJ2364 como tataravô. Considerando somente os AFLPs, os genótipos menos similares não apresentam ancestrais comuns até a geração dos tataravôs, enquanto os genótipos menos similares, com base nos microssatélites, apresentam em comum, em mais de uma geração, os ancestrais POJ2878 e CO419, sendo este último obtido pelo cruzamento POJ2878 x CO290.

Por outro lado, os pares de genótipos mais similares de acordo com as informações genealógicas foram RB83-5486 e IACSP97-3331, o que se justifica pelo fato de IACSP97-3331 ser resultante do cruzamento de RB83-5486 com ela mesma. De acordo com o parentesco molecular (AFLPs e SSRs conjuntamente), o par de genótipos mais próximo é composto por IACSP97-2053 e IAC 97-7065, que são irmãos-germanos. Enquanto que o parentesco molecular baseado em AFLPs evidencia como mais próximos os genótipos RB85-5156 x IACSP95-3028, sendo o primeiro um dos genitores do segundo e, finalmente, considerando somente os SSRs, os genótipos mais aparentados são os irmãos-germanos IACSP99-4007 e IACSP99-4008, que obtiveram também a segunda maior estimativa de parentesco molecular quando todas as marcas foram consideradas (Figuras 1 a 4).

De acordo com o índice de Jaccard, os genótipos mais similares, com base em todas as marcas, ou somente nos AFLPs, são os irmãos-germanos IACSP97-2020 e IACSP97-7074, enquanto que, considerando somente os SSRs, SP79-1011 e IACSP93-3046 são os

genótipos mais similares, sendo que SP79-1011 é um dos genitores de IACSP93-3046 (dados não apresentados).

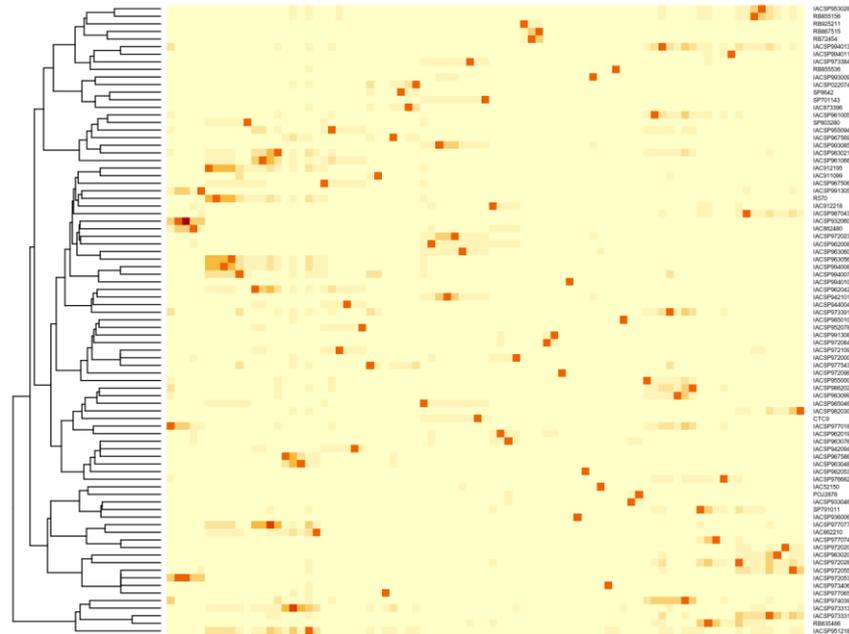


Figura 1. Heatmap e dendrograma (UPGMA) a partir das estimativas dos coeficientes de parentesco, calculadas a partir de informações genealógicas.

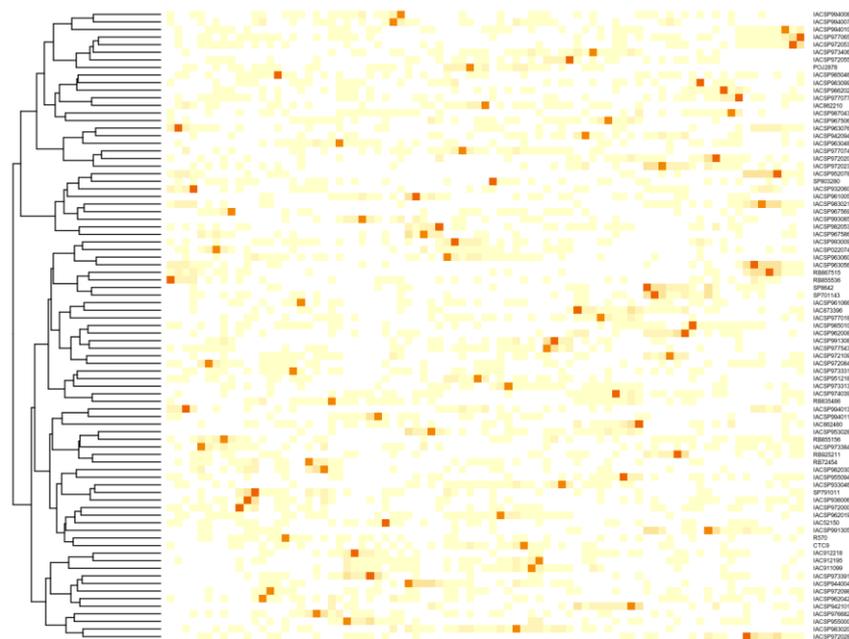


Figura 2. Heatmap e dendrograma (UPGMA) a partir das estimativas dos coeficientes de parentesco, calculados com dados de marcadores moleculares (AFLPs e SSRs).

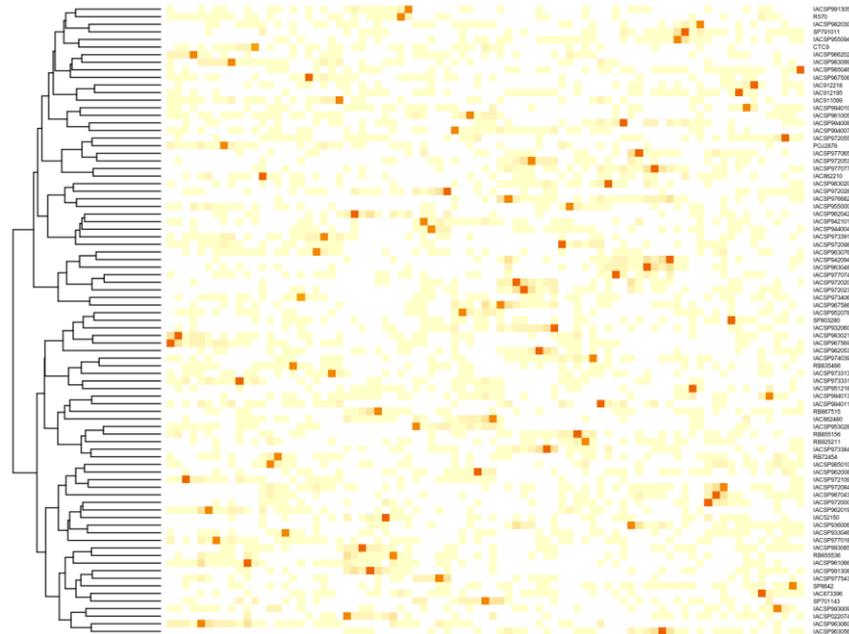


Figura 3. Heatmap e dendrograma (UPGMA) a partir das estimativas dos coeficientes de parentesco, calculados com dados de marcadores moleculares (AFLPs).

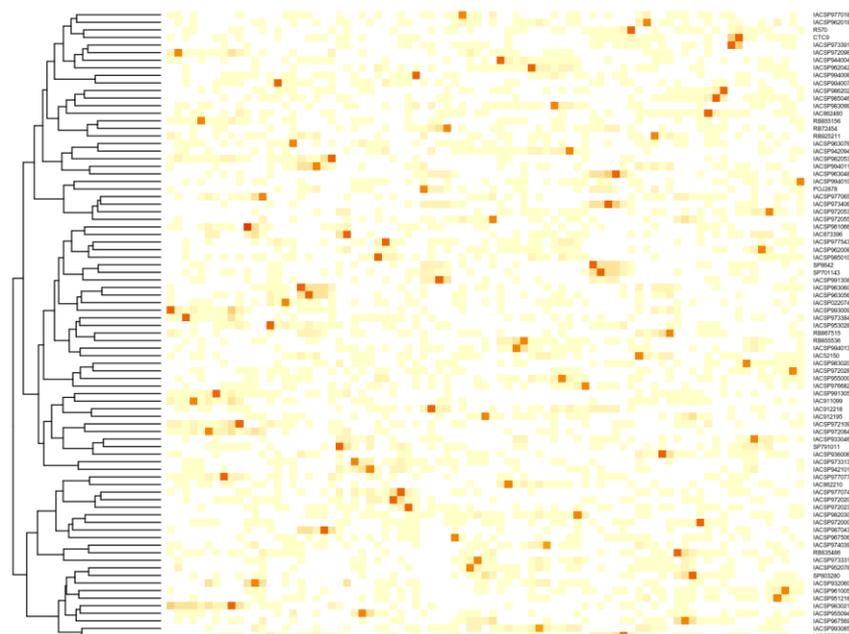


Figura 4. Heatmap e dendrograma (UPGMA) a partir das estimativas dos coeficientes de parentesco, calculados com dados de marcadores moleculares (SSRs).

São observados alguns agrupamentos a partir das estimativas de parentesco. No caso dos dados genealógicos, por exemplo, observa-se um grupo que contém os genótipos SP80-3280, IACSP95-5094, IACSP96-1005, IACSP96-7569, IACSP98-3021 e IACSP99-

3085, o que se justifica pelo fato de os cinco últimos acessos serem meios-irmãos, tendo como um dos genitores a variedade SP80-3280 (Figura 1). Verifica-se que um agrupamento parecido com esse se forma com base nas informações dos marcadores, que contém os acessos SP80-3280, IACSP95-2078, IACSP93-2060, IACSP96-1005, IACSP96-7569, IACSP98-3021 e IACSP99-3085 (Figura 2). É interessante notar que os genótipos SP80-3280, IACSP96-7569 e IACSP98-3021 são classificados no mesmo grupo, de acordo com as sete formas de avaliar as medidas de parença entre os genótipos, consideradas nessa pesquisa. Os genótipos IACSP96-1005 e IACSP93-2060 só não são agrupados com SP80-3280, IACSP96-7569 e IACSP98-3021 quando são avaliadas as matrizes G\_AFLPs e A, respectivamente.

Outro grupo coeso de genótipos é formado por RB83-5486, IACSP95-1218, IACSP97-3313, IACSP97-3331 e IACSP97-4039, sendo esses últimos quatro genótipos têm RB83-5486 como pelo menos um de seus genitores. Esses cinco genótipos estão sempre muito próximos de acordo com as matrizes A, G\_All, A\_AFLPs, J\_All e J\_AFLPs (dados não mostrados). De acordo com a matriz de parentescos baseada em marcadores SSRs, no entanto, os genótipos IACSP95-1218 e IAC97-3313 não estão tão próximos dos demais. O mesmo ocorre ao avaliar a matriz de similaridades de Jaccard, sendo que, nesse caso, o genótipo IACSP97-3331 também está mais distante de RB83-5486 e IACSP97-4039. Isso pode ser uma evidência de que o tipo de variabilidade detectada por diferentes marcadores pode não ser totalmente coincidente, o que é esperado, uma vez que são marcadores com características diferentes também.

As correlações estimadas entre os pares de matrizes de medidas de parença variaram 0,4268 entre as matrizes A e J\_SSRs a 0,9057 entre G\_All e G\_SSRs (Tabela 2, Figura 5), sendo todas elas altamente significativas ( $\alpha \leq 0,0001$ ), de acordo com o teste de Mantel.

Tabela 2. Comparação das matrizes de parecência, estimativas das correlações entre matrizes e de seus valores de  $p$ , estimados a partir do teste de Mantel.

Matrizes <sup>1</sup> comparadas	Estimativa do coeficiente de correlação de Pearson	Percentil <sup>2</sup>		Valor de $p$ <sup>3</sup>
		0,95	0,99	
A e G_All	0,5206	0,0266	0,0387	$10^{-04}$
A e G_AFLPs	0,4585	0,0270	0,0391	$10^{-04}$
A e G_SSRs	0,4702	0,0268	0,0394	$10^{-04}$
A e J_All	0,4922	0,0453	0,0644	$10^{-04}$
A e J_AFLPs	0,4354	0,0492	0,0699	$10^{-04}$
A e J_SSRs	0,4268	0,0472	0,0679	$10^{-04}$
G_All e G_AFLPs	0,8779	0,0297	0,0429	$10^{-04}$
G_All e G_SSRs	0,9057	0,0290	0,0419	$10^{-04}$
G_All e J_All	0,8617	0,0246	0,0364	$10^{-04}$
G_All e J_AFLPs	0,7300	0,0292	0,0435	$10^{-04}$
G_All e J_SSRs	0,7754	0,0249	0,0364	$10^{-04}$
G_AFLPs e G_SSRs	0,5921	0,0294	0,0426	$10^{-04}$
G_AFLPs e J_All	0,7555	0,0248	0,0365	$10^{-04}$
G_AFLPs e J_AFLPs	0,8276	0,0243	0,0348	$10^{-04}$
G_AFLPs e J_SSRs	0,5125	0,0248	0,0348	$10^{-04}$
G_SSRs e J_All	0,7814	0,0252	0,0358	$10^{-04}$
G_SSRs e J_AFLPs	0,5010	0,0244	0,0355	$10^{-04}$
G_SSRs e J_SSRs	0,8512	0,0248	0,0360	$10^{-04}$
J_All e J_AFLPs	0,8636	0,0567	0,0815	$10^{-04}$
J_All e J_SSRs	0,8905	0,0528	0,0757	$10^{-04}$
J_AFLPs e J_SSRs	0,5401	0,0585	0,0818	$10^{-04}$

<sup>1</sup> Matrizes comparadas: A (parentesco baseado em dados genealógicos), G\_All (parentesco molecular considerando AFLPs e SSRs), G\_AFLPs (parentesco molecular considerando AFLPs), G\_SSRs (parentesco molecular considerando SSRs), J\_All (índice de similaridade de Jaccard considerando AFLPs e SSRs), J\_AFLPs (índice de similaridade de Jaccard considerando AFLPs), J\_SSRs (índice de similaridade de Jaccard considerando SSRs).

<sup>2</sup> São os valores de  $p$  correspondentes aos níveis de significância de 0,05 e 0,01, considerando 9999 permutações.

<sup>3</sup> Valor de  $p$ , considerando 9999 permutações.

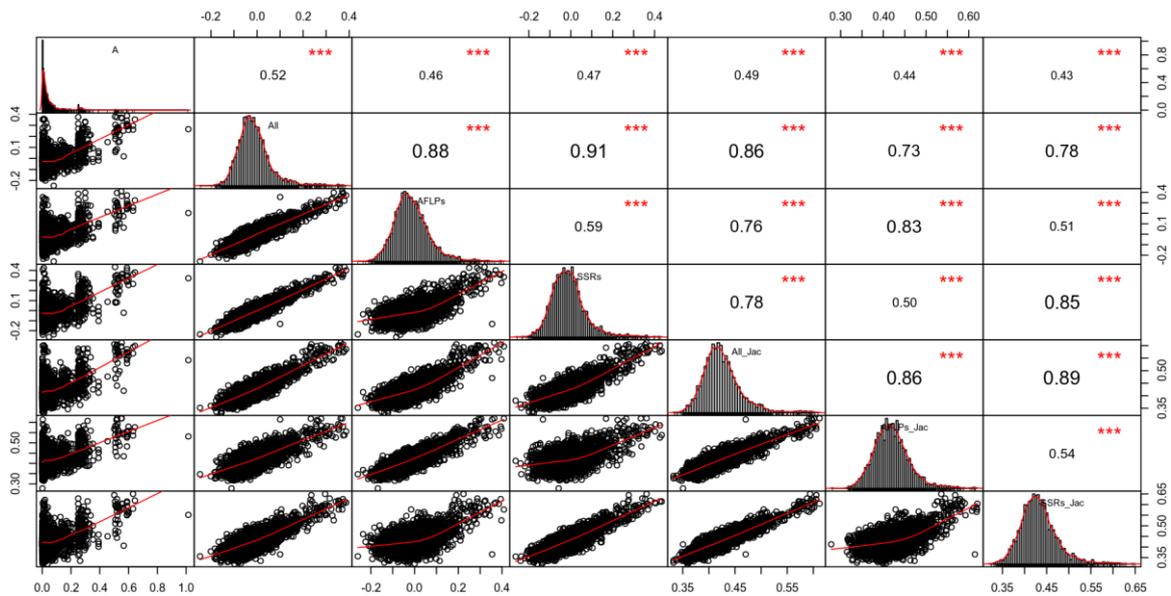


Figura 5. Correlações entre as estimativas de parentesco baseadas em dados genealógicos e moleculares, obtidas par a par, entre os 83 genótipos avaliados. \*\* Significativo com ( $\alpha \leq 0,01$ ); \*\*\* Significativo com ( $\alpha \leq 0,001$ ).

As menores correlações foram observadas entre os parentescos calculados a partir de dados genealógicos em relação a todas as outras medidas de parentesco. Isso pode ter ocorrido em função da diferença da natureza dos dados genealógicos e moleculares. As maiores correlações envolvendo parentesco genealógico foram obtidas em relação ao parentesco molecular e ao índice de Jaccard considerando todos os marcadores moleculares concomitantemente, enquanto as menores correlações ocorreram nas comparações com os índices de Jaccard calculados com apenas um dos marcadores (Figura 5). Isso indica que, para se chegar mais próximo das estimativas dos coeficientes de parentesco calculados com informações genealógicas, deve-se utilizar as estimativas de parentesco calculadas com marcadores AFLPs e SSRs simultaneamente.

As correlações mais elevadas ocorreram entre os parentescos moleculares obtidos com todas as marcas e os obtidos com marcadores somente com os SSR e, em seguida, com os parentescos obtidos somente com os AFLPs. Dessa forma, parece haver uma boa consistência no cálculo dos coeficientes de parentesco moleculares desses genótipos se forem utilizados ambos os marcadores ou apenas um deles. Quando consideradas as estimativas dos coeficientes de Jaccard, a maior correlação foi observada quando o cálculo foi realizado com ambos os marcadores e quando os coeficientes foram estimados somente com SSRs. Contudo, a correlação entre J\_All e J\_SSRs também foi alta. Novamente, houve uma

boa consistência entre as estimativas dos coeficientes de Jaccard usando ambos ou somente um dos marcadores.

As estimativas correlações envolvendo os dois tipos de marcadores separadamente (G\_AFLPs e G\_SSRs; G\_AFLPs e J\_SSRs; G\_SSRs e J\_AFLPs; J\_AFLPs e J\_SSRs) foram as mais baixas dentre as correlações entre coeficientes parecidos obtidos a partir de dados moleculares, evidenciando que as informações fornecidas por diferentes tipos de marcadores apresentam semelhanças e divergências entre si. Essas diferenças são esperadas, pois diferentes tipos de marcadores, cada qual com suas características particulares, medem diferentes tipos de variabilidade também.

Amadeu et al. (2020) verificaram que o uso de marcadores bialélicos é mais indicado para estimar o parentesco que os marcadores multialélicos e enfatizam que métodos desenvolvidos para poliploides devem ser utilizados. Observaram também que é mais fácil inferir sobre o parentesco entre genótipos de espécies poliploides quando o nível de endogamia é reduzido. Sugerem que, em autoploiploides, para se atingir elevada acurácia na estimativa dos parentescos, devem ser usados mais de 100 SNPs com dosagem conhecida e o método de VanRaden estendido. Na presente pesquisa, o número de marcadores foi 922, sendo todos interpretados como bialélicos para a realização dos cálculos de parentescos molecular de acordo com o método de estendido de VanRaden, estando, nesses quesitos, em acordo com as recomendações feitas por Amadeu et al. (2020) e, desse modo, acredita-se que as estimativas dos coeficientes de parentesco obtidas sejam acuradas.

Evidentemente, estudos mais abrangentes, envolvendo simulações, inclusive para se quantificar a perda de informações ao se considerar um marcador codominante como dominante nesses cálculos seria interessante, bem como testar outros níveis de ploidia, incluindo ou não a possibilidade de ocorrência de redução dupla na meiose.

Os resultados obtidos com o presente estudo são importantes para o programa de melhoramento do IAC. Informações sobre as relações de parentesco entre os genitores é fundamental para o planejamento de hibridações. Muitas vezes, ocorre que se têm disponíveis informações genealógicas, ou de genotipagem molecular. Com base na avaliação desse painel de acessos de cana do Programa Cana do IAC, que é representativo da parte do banco de germoplasma da instituição, há indicativos de que ambos os tipos de dados podem ser utilizados para obter a informação sobre o parentesco (ou similaridade) entre os acessos. No caso dos marcadores moleculares, parece ser indicado utilizar estimativas feitas a partir dos dois tipos de marcadores, simultaneamente, por apresentarem correlações mais elevadas quando comparados ao método do cálculo do parentesco a partir de dados genealógicos.

## 6. CONCLUSÕES

As estimativas dos coeficientes de parentesco obtidas com método estendido de VanRaden e considerando ambos os tipos de marcadores conjuntamente foram as que mais se aproximaram das estimativas de parentesco baseadas em dados genealógicos. Desse modo, recomenda-se que, considerando dados de marcadores moleculares, sejam utilizados os dois tipos de marcadores e também método de estimação adaptado para espécies poliploides.

Quando foi utilizado um único marcador, a concordância dos resultados obtidos com dados genealógicos e os obtidos com os dados AFLPs ou SSRS foram similares.

Quando utilizado somente um marcador, tanto no caso dos parentescos moleculares quanto nas similaridades de Jaccard, as correlações obtidas com marcadores de tipos diferentes foram sempre menores que as calculadas a partir do mesmo tipo de marcadores, indicando que as informações fornecidas por diferentes tipos de marcadores são parcialmente distintas.

As correlações entre os resultados obtidos com as diferentes abordagens de estimação de medidas de parença são relativamente satisfatórias. Essa informação é de grande valia para os programas de melhoramento. Entretanto, os resultados aqui obtidos são restritos ao painel avaliado, sendo necessários estudos adicionais envolvendo outros conjuntos de dados, bem como de simulações.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAM, K.L.; WENDEL, J.F. Polyploidy and genome evolution in plants. **Current Opinion in Plant Biology**, v.8, p.135-141, 2005. DOI 10.1016/j.pbi.2005.01.001.
- AITKEN, K.S.; LI, J.-C.; JACKSON, P.; PIPERIDIS, G.; MCINTYRE, C.L. AFLP analysis of genetic diversity within *Saccharum officinarum* and comparison with sugarcane cultivars. **Australian Journal of Agricultural Research**, v.57, p.1167-1184, 2006. <https://doi.org/10.1071/AR05391>.
- AITKEN, K.; McNEIL, M. Diversity Analysis. In: HENRY, R.; CHITTARANJAN, K. (Eds.) **Genetics, Genomics and Breeding of Sugarcane**. CRC Press, 2010. Chap. 3, p.19-42.
- AL-JANABI, S.M.; FORGET, L.; DOOKUN, A. An improved and rapid protocol for the isolation of polysaccharide- and polyphenol-free sugarcane DNA. **Plant Molecular Biology Reports**, Netherlands, v. 17, n. 3, p. 1-8, 1999.
- ALWALA, S.; SUMAN, A.; ARRO, J.A.; VEREMIS, J.C.; KIMBENG, C.A. Target Region Amplification Polymorphism (TRAP) for Assessing Genetic Diversity in Sugarcane Germplasm Collections. **Crop Science**, v.46, p.448-455, 2006. doi:10.2135/cropsci2005.0274.
- AMADEU, R.R.; CELLON, C.; OLMSTEAD, J.W.; GARCIA, A.A.F.; RESENDE JR., M.F.R. ; MUÑOZ, P.R. AGHmatrix: R to Construct Relationship Matrices for Autotetraploid and Diploid Species: A Blueberry Example. **The Plant Genome**, v.9, n.3, 10 p. 2016. doi:10.3835/plantgenome2016.01.0009.
- AMADEU, R.R.; LARA, L.A.C.; MUNOZ, P.; GARCIA, A.A.F. Estimation of Molecular Pairwise Relatedness in Autopolyploid Crops. **G3-Genes/Genomes/Genetics**. Vol. 10, 2020. <https://doi.org/10.1534/g3.120.401669>.
- ARCENEUAUX, G. Cultivated sugarcane of the worlds and their botanical derivation. Proc. **Int. Soc. Sugar Cane Technol.** v.12, p. 844-854. 1967.
- ASHRAF, B.H.; BYRNE, S.; FÉ, D.; CZABAN, A.; ASP, T.; PEDERSEN, M.G.; LENK, I.; ROULUND, N.; DIDION, T.; JENSEN, C.S. et al., Estimating genomic heritabilities at the level of family-pool samples of perennial ryegrass using genotyping-by-sequencing. **Theoretical Applied Genetics**, v.129, p.45-52, 2016.
- BELLO, E.B.; RASCO, J.L.S.; SENDON, P.M.D.; CUEVA, F.M.D.; LALUSIN, A.G.; LAURENA, A.C. Genetic Diversity Analysis of Selected Sugarcane (*Saccharum* spp. Hybrids) Varieties Using DaRT-Seq Technology. **Philippine Journal of Science**, v.148, S1, p.103-114, 2019. (Special Issue on Genomics).
- BESSE, P.; TAYLOR, G.; CARROLL, B.; BERDING, N.; BURNER, D.; MCINTYRE, C.L. Assessing genetic diversity in a sugarcane germplasm collection using an automated AFLP analysis. **Genetica**, v. 104, p.143-153, 1998.
- BREMER, G. On the somatic chromosome numbers of sugarcane forms of endogenous cane. **Proc ISSCT**, v.4, p.30, 1932.

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

BURGESS, D. Polyploid gains. **Nature Reviews Genetics**, v.16, n.15, p.196, 2015.  
doi:10.1038/nrg3925.

CHEAVEGATTI-GIANOTTO, A.; ABREU, H.M.C; ARRUDA, P., et al. Sugarcane (*Saccharum X officinarum*): A Reference study for the regulation of genetically modified cultivars in Brazil. **Tropical Plant Biology**, v.4, n.1, p.62–89, 2011.

CONAB. Observatório Agrícola: Acompanhamento da Safra Brasileira. Cana-de-açúcar. Safra 2022/23. Segundo Levantamento. Monitoramento Agrícola - Cana-de-açúcar. Companhia Nacional de Abastecimento, v.9, n.1, p 8-10, 2021.  
(<https://www.conab.gov.br/info-agro/safras/cana>). Acesso em 30/05/2022.

CORDEIRO, G.M.; MAGUIRE, T.L.; HENRY, R.J.; EDWARDS, K.J. Optimisation of a microsatellite enrichment technique in *Saccharum* spp. **Plant Molecular Biology Reporter**, Vol. 17, pag. 225-229, 1999.

CORDEIRO, G.M.; PAN, Y.-B.; HENRY, R.J. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. **Plant Science**, v.165, p.181-189, 2003.  
doi:10.1016/S0168-9452(03)00157-2.

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

CRESTE, S.; SANSOLI, D.M.; TARDIANI, A.C.S.; SILVA, D.N.; GONÇALVES, F.K.; FÁVERO, T.M.; MEDEIROS, C.N.F.; FESTUCCI, C.S.; CARLINI-GARCIA, L.A.; LANDELL, M.G.A.; PINTO, L.R. Comparison of AFLP, TRAP and SSRs in the estimation of genetic relationship in sugarcane. **Sugar Tech**, v.12, n.2, p.150-154, 2010.

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

CRONQUIST, A. **An Integrated System of Classification of Flowering Plants**. New York: Columbia University Press, 1981.

DANIELS, J.; DANIELS, C.A. Geographical, Historical and Cultural Aspects of the Origin on the Indian and Chinese Sugarcane *S. barberi* and *S. sinense*. **Sugarcane Breeding Newsletter**, v.36, p.4-23, 1975.

DEVARUMATH, R.M.; SACHIN B. KALWADE, S.B.; BUNDOCK, P.; ELIOTT, F.G.; HENRY, R. Independent target region amplification polymorphism and single nucleotide polymorphism marker utility in genetic evaluation of sugarcane genotypes. **Plant Breeding**, v.132, n.6, p.736-747, 2013. <https://doi.org/10.1111/pbr.12092>.

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

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

DICE, L. R. Measures of the amount of ecologic association between species. **Ecology**, v.26, n.3, 297-302, 1945.

DUARTE FILHO, L.S.C.; SILVA, P.P.; SANTOS, J.M.; BARBOSA, G.V.S.; RAMALHONETO, C.E.; SOARES, L.; ANDRADE, J.C.F.; ALMEIDA, C. Genetic similarity among genotypes of sugarcane estimated by SSR and coefficient of parentage. **Sugar Tech** 12, 145–149, 2010. doi:10.1007/s12355-010-0028-2. <https://link.springer.com/article/10.1007/s12355-010-0028-2>.

FEHR, W.R. **Principles of Cultivar Development. Theory and Technique**. Vol.1. 1987. 536p.

GARCIA, A.A.F.; KIDO, E.A.; MEZA, A.N.; SOUZA, H.M.B.; PINTO, L.R.; PASTINA, M.M.; LEITE, C.S.; DA SILVA, J.A.G.; ULIAN, E.C.; FIGUEIRA, A.; DE SOUZA, A.P. Development of a integrated genetic map o a sugarcane (*Saccharum* spp.) commercial cross, based on maximum-likelihood approach for estimation of linkage and linkage phases. **Theoretical and Applied Genetics**, Berlin, v. 112, n. 2, p. 288-314, 2006.

GARCIA, A.A.F.; et al., SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. **Scientific Reports**, v.3, 2013.

GONÇALVES, B.S. **Identificação de marcadores moleculares relacionados à severidade de sintomas da ferrugem marrom em painel associativo de cana-de-açúcar (*Saccharum* spp.)**. 68f. Dissertação (Mestrado em Agronomia) - Faculdade de Ciências Agrárias e Veterinárias – UNESP. Jaboticabal - SP, 2012.

GRIVET, L.; D'HONT, A.; ROQUES, D.; FELDMAN, P.; LANAUD, C.; GLASZMANN, J. C. RFLP Mapping in Cultivated Sugarcane (*Saccharum* spp.): Genome Organization in a Highly Polyploid and Aneuploid Interspecific Hybrid. **Genetics**, v.142, n.3, p. 987-1000, 1996.

HARVEY, M.; HUCKETT, B.I., BOTHA, F.C. Use of the polymerase chain (PCR) and Random Amplification of Polymorphic DNAs (RAPDs) for determination of genetic distances between 21 sugarcane varieties. **Proceedings of The South African Sugar Technologists' Association** - p.36-40, 1994. [https://sasta.co.za/mdocs-posts/1994\\_harvey\\_use-of-the-polymerase/](https://sasta.co.za/mdocs-posts/1994_harvey_use-of-the-polymerase/).

HELLER-USZYNSKA, K., USZYNSKI, G., HUTTNER, E.; EVERS, M.; CARLIG, J.; CAIG, V.; AITKEN, K.; JACKSON, P.; PIPERIDIS, G.; COX, M.; GILMOUR, R.; D'HONT, A.; BUTTERFIELD, M. GLASZMANN, J.-C.; KILIAN, A. Diversity Arrays Technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. **Molecular Breeding**, v.28, p.37-55, 2011. doi:10.1007/s11032-010-9460-y.

- HOARAU, J.Y.; OFFMANN, B.; D'HONT, A.; RISTERUCCI, A.M.; ROQUES, D.; GLASZMANN, J.C.; GRIVET, L. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. **Theoretical Applied Genetics**, v.103, p.84-97, 2001. <https://doi.org/10.1007/s001220000390>.
- HUANG, K.; GUO, S.; SHATTUCK, M.; CHEN, S.; QI, X.; ZHANG, P.; LI, B. A maximum-likelihood estimation of pairwise relatedness for autopolyploids. **Heredity**, v.114, p.133-142, 2015.
- HUANG, K., RITLAND, K; GUO, S.; SHATTUCK, M.; LI, B. A pairwise relatedness estimator for polyploids. **Molecular Ecology Resources**, v.14, p.734-744, 2014.
- JACCARD, P. Etude comparative de la distribution florale dans une portion des Alpes et du Jura. **Bulletin Society Vaudoise Science Nature**, v.37, p. 547-579, 1901.
- JACCARD, P. Nouvelles recherches sur la distribution florale. **Bulletin Society Vaudoise Science Nature**, v.44, p.223-270, 1908.
- JANNOO, N.; FORGET, L.; DOOKUN, A. Contribution of microsatellites to sugarcane breeding program in Mauritius. In: XXIV Congress, 2001, Brisbane. **International Society of Sugar Cane Technologists**. Brisbane, 2001.
- JORDÃO Jr., H. **Desenvolvimento de um sistema baseado em marcadores moleculares de DNA do tipo microssatélites para identificação de variedades de cana-de-açúcar**. 2009. 106p. Dissertação (Mestrado em Biologia Funcional e Molecular) - Universidade Estadual de Campinas, Campinas, 2009.
- KAWAR, P.G.; DEVARUMATH, R.M.; NERKAR, Y. Use of RAPD markers for assessment of genetic diversity in sugarcane cultivars. **Indian Journal of Biotechnology**, v.8, p.67-71, 2009. [http://nopr.niscair.res.in/bitstream/123456789/2955/1/IJBT%208\(1\)%2067-71](http://nopr.niscair.res.in/bitstream/123456789/2955/1/IJBT%208(1)%2067-71).
- KERR, R.J.; LI, L.; TIER, B.; DUTKOWSKI, G.W.; MCRAE, T.A. Use of the numerator relationship matrix in genetic analysis of autopolyploid species. **Theoretical Applied Genetics**, v.124, p. 1271-1282, 2012.
- KHAN, I.A.; BIBI, S.; YASMIN, S.; KHATRI, A.; SEEMA, N.; AFGHAN, S. Pak. Genetic variability in mutated population of sugarcane clone NIA-98 through molecular markers (RAPD and TRAP). **Journal Botanical**, v.42, n.1, 605-614, 2010.
- LANDELL, M.G.A.; BRESSIANI, J.A. Melhoramento genético e manejo varietal. In: DINARDO-MIRANDA, L.L.; VASCONCELOS, A.C.M.; LANDELL, M.G.A.(Org.). **Cana-de-Açúcar**. 1ed. Campinas: Instituto Agrônomo de Campinas, 2008, v. 1, p. 101-156.
- LANDELL, M. G. A.; SILVA, M. A. As estratégias de seleção de cana em desenvolvimento no Brasil. **Visão Agrícola**, v.1, p. 18-23, 2004.
- LIMA, M.L.A.; GARCIA, • A.A.F.; OLIVEIRA, K.M.; MATSUOKA, S.; ARIZONO, H.; SOUZA JR., C.L.; SOUZA, A.P. Analysis of genetic similarity detected by AFLP and

coefficient of parentage among genotypes of sugar cane (*Saccharum* spp.). **Theoretical Applied Genetics**, v.104, p.30-38, 2002.

LU, Y.H.; D'HONT, A.; WALKER, D.I.T.; RAO, P.S.; FELDMANN, P.; GLASZMANN, J.C. Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. **Euphytica**, v.78, n.7, 1994. <https://doi.org/10.1007/BF00021393>.

LYNCH, M.; RITLAND, K. Estimation of pairwise relatedness with molecular markers. **Genetics**, v.152, p.1753-1766, 1999.

MANECHINI, J.R.V.; COSTA, J.B.; PEREIRA, B.T., et al. Unraveling the genetic structure of Brazilian commercial sugarcane cultivars through microsatellite markers. **Plos One**. 2018. <https://doi.org/10.1371/journal.pone.0195623>.

MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, **27**, 209–220, 1967.

MAPA, 2020. Disponível em:

<http://www.agricultura.gov.br/assuntos/sustentabilidade/agroenergia/agroenergia-no-brasil> (acesso em 31/05/2022)

MARCONI, T.G.; COSTA, E.A.; MIRANDA, H.R.C.A.N., MANCINI, M.C.; CARDPSP-SILVA, C.B.; OLIVEIRA, K.M.; PINTO, L.R.; MOLLINARI, M.; GARCIA, A.A.F.; SOUZA, A.P. Functional markers for gene mapping and genetic diversity studies in sugarcane. **BMC Research Notes**, v.4, n.264; 10p. 2011. <http://www.biomedcentral.com/1756-0500/4/264>.

MATSUOKA, S.; GARCIA, A. A. F.; ARIZONO, H. Melhoramento da cana-de-açúcar. In: BORÉM, A.(Ed), **Melhoramento de espécies cultivadas**. 2 ed. Viçosa, UFV, 1999, p. 205-251.

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

NEI, M. Genetic distance between populations. **The American Naturalist**, v.106, n.949, p.283-292, 1972.

NEVES, M.F.; CONEJERO, M.A. O negócio do açúcar no Brasil e no mundo. In: **Estratégias para a cana no Brasil: um negócio classe mundial**. São Paulo: Atlas, 2010a, p. 89-112.

NOVA CANA.COM. Curitiba-Paraná, 2013-2021. Disponível em: <https://www.novacana.com/cana-de-acucar>. Acesso em: 05.08.2021.

OKSANEN, J. et al. vegan: Community Ecology Package. R package version 2.6-2, 2022. <https://CRAN.R-project.org/package=vegan>.

OLIVEIRA, K.M.; PINTO, L.R.; MARCONI, T.G.; MOLLINARI, M.; ULIAN, E.C.; CHABREGAS, S.M.; FALCO, M.C.; BURNQUIST, W.; GARCIA, A.A.F.; SOUZA, A.P.

Characterization of new polymorphic functional markers for sugarcane. **Genome**, v.52, p.191-209, 2009. doi: 10.1139/G08-105.

ORIOLO, F. K.M.Y. **Genotipagem molecular de acessos da Coleção Mundial de Cana-de-açúcar e Gramíneas Correlatas, com aplicações na obtenção de variedades com elevada concentração de biomassa e resistência a fatores bióticos e abióticos**. 2016. Dissertação (Mestrado em Sanidade, Segurança Alimentar e Ambiental no Agronegócio) – Programa de Pós-Graduação, Instituto Biológico, São Paulo, 2016. Disponível em: <http://www.biológico.sp.gov.br/pos/uploads/files/pdf/2016/fernanda.pdf>. Acesso em: 28/11/2020.

PAN, Y. B. Databasing Molecular Identities of Sugarcane (*Saccharum* spp.) Clones constructed with microsatellite (SSR) DNA Markers. **American Journal of Plant Sciences**, v.1, p.87-94, 2010.

PAN, Y.-B., MILLER, J.D., SCHNELL, R.J., RICHARD JR, E.P., WEI, Q. Application of microsatellite and RAPD fingerprints in the Florida sugarcane variety program. **Sugar Cane International**. p.19-28, 2003.

PETERSON, B. G.; CARL, P. PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis. R package version 2.0.4. 2020. <https://CRAN.R-project.org/package=PerformanceAnalytics>

PINTO, L.R.; OLIVEIRA, K.M.; MARCONI, T.; GARCIA, A.A.F.; ULIAN, E.C.; SOUZA, A.P. Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. **Plant Breeding**, v.125, n.4, p.378-384, 2006. <https://doi.org/10.1111/j.1439-0523.2006.01227.x>

PINTO, L.R.; OLIVEIRA, K.M.; ULIAN, E.C.; GARCIA, A.A.F.; SOUZA, A.P. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome**, v.47, n.5, pp. 795-804, 2004. <https://doi.org/10.1139/g04-055>.

R CORE TEAM. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 2020- <https://www.r-project.org/>.

RStudio TEAM (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. 2021. <http://www.rstudio.com/>.

ROACH, B.T.; DANIELS, J. (1987). A review of the origin and improvement of sugarcane. In: **Copersucar International Sugarcane Breeding Workshop**, 1: 1-31.

ROGER, J. S. Measure of genetic similarity and genetic distance. Studies in genetics VII. **University of Texas publication**, v.7213, p.145-153, 1972.

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

SCHUELKE, M. An economic method for fluorescent labeling of PCR fragments. **Nature Biotechnology**, New York, v. 18, p. 233-234, 2000.

SELVI, A.; MUKUNTHAN, N.; SHANTHI, R.M.; GOVINDARAJ, P.; SINGARAVELU, B.; PRABU, T.K. Assessment of genetic relationships and marker identification in sugarcane cultivars with different levels of top borer resistance. **Sugar Tech**, v.10, p.53-59, 2008.

SELVI, A.; NAIR, N.V.; NOYER, J.L.; SINGH, N.K.; BALASUNDARAM, N.; BANSAL, K.C.; KOUNDAL, K.R.; MOHAPATRA, T. AFLP Analysis of the Phenetic Organization and Genetic Diversity in the Sugarcane Complex, *Saccharum* and *Erianthus*. **Genetic Resources and Crop Evolution**, v.53, p.831-842, 2006. <https://doi.org/10.1007/s10722-004-6376-6>.

SINGH, R.B.; SINGH, B.; SINGH, R.K. Study of genetic diversity of sugarcane (*Saccharum*) species and commercial varieties through TRAP molecular markers. **Indian Journal Plant Physiology** v.22, p.332-338, 2017. doi:10.1007/s40502-017-0314-z.

SINGH, R.K.; MISHRA, S.K.; SINGH, S.P.; MISHRA, N.; SHARMA, M.L. Evaluation of microsatellite markers for genetic diversity analysis among sugarcane species and commercial hybrids. **Australian Journal of Crop Science**, v.4, n.2, p.115-124, 2010.

SLATER, A.T.; COGAN, N.O.; FORSTER, J.W.; HAYES, B.J.; DAETWYLER, H.D. Improving genetic gain with genomic selection in autotetraploid potato. **Plant Genome**, v.9, p.1-15, 2016.

SOKAL, R.R.; MICHENER, C.D. A statistical method for evaluating systematic relationships. **University Kansas Science Bulletin**, v.38, p.1409-1438, 1958.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **Journal Dairy Science**, v.91: p.4414-4423, 2008.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; LEE, T.V.; HORNES, M.; FRIJTERS, A.; POT, J.; PELEMAN, J.; KULPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, Oxford, v. 23, n. 1, p. 4407-4414, 1995.

WRIGHT, S. Coefficients of inbreeding and relationship. **American Naturalist**, v.56, p.330-338, 1922.

YANG, J.; BENYAMIN, B.; MCEVOY, B.P.; GORDON, S.; HENDERS, A.K.; NYHOLT, D.R.; MADDEN, P.A.; HEATH, A.C.; MARTIN, N.G.; MONTGOMERY, G.W.; GODDARD, M.E.; VISSCHER, P.M. Common SNPs explain a large proportion of the heritability for human height. **Nature Genetics**, v.42, p.565-569, 2010.

ZAN, F.; ZHANG, Y.; WU, Z.; ZHAO, J.; WU, C., et al. Genetic analysis of agronomic traits in elite sugarcane (*Saccharum* spp.) germplasm. **PLOS ONE**, CHINA. 2020. Disponível em: <https://doi.org/10.1371/journal.pone.0233752> pmid: 32526769. Acesso em: 22/08/2020.