

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

**ANÁLISE DA ESTRUTURA GENÉTICA DO BANCO DE
GERMOPLASMA DO PROGRAMA CANA - IAC POR MEIO
DE MARCADORES MOLECULARES**

João Ricardo Vieira Manechini

Biólogo

2017

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

**ANÁLISE DA ESTRUTURA GENÉTICA DO BANCO DE
GERMOPLASMA DO PROGRAMA CANA - IAC POR MEIO
DE MARCADORES MOLECULARES**

**João Ricardo Vieira Manechini
Orientadora: Profa. Dra. Luciana Rossini Pinto
Coorientador: Dr. Marcos Guimarães de Andrade Landell**

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Campus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Genética e Melhoramento de Plantas).

2017

M274a Manechini, João Ricardo Vieira
Análise da estrutura genética do banco de germoplasma do
Programa Cana - IAC por meio de marcadores moleculares /
Manechini, João Ricardo Vieira. -- Jaboticabal, 2017
x, 98 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2017
Orientadora: Luciana Rossini Pinto
Banca examinadora: Dilermando Perecin, Ana Lilia Alzate Marin
Bibliografia

1. *Saccharum* spp. 2. Banco de germoplasma. 3. Diversidade
genética. 4. Microssatélites. 5. Coleção nuclear. I. Título. II.
Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.52:633.61

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

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: ANÁLISE DA ESTRUTURA GENÉTICA DO BANCO DE GERMOPLASMA DO PROGRAMA CANA - IAC POR MEIO DE MARCADORES MOLECULARES

AUTOR: JOÃO RICARDO VIEIRA MANECHINI

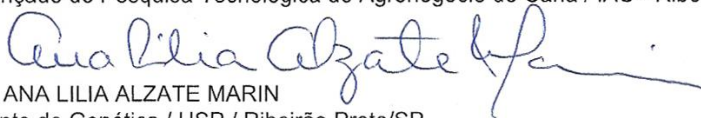
ORIENTADORA: LUCIANA ROSSINI PINTO

COORIENTADOR: MARCOS GUIMARÃES DE ANDRADE LANDELL

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



Pesquisadora Dra. LUCIANA ROSSINI PINTO
Centro Avançado de Pesquisa Tecnológica do Agronegócio de Cana / IAC - Ribeirão Preto/SP



Profa. Dra. ANA LILIA ALZATE MARIN
Departamento de Genética / USP / Ribeirão Preto/SP



Prof. Dr. DILERMANDO PERECIN
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Jaboticabal, 20 de junho de 2017

DADOS CURRICULARES DO AUTOR

João Ricardo Vieira Manechini, nascido em 25 de fevereiro de 1987 em Ribeirão Preto, SP, Brasil. Em 2011, ingressou no curso de Ciências Biológicas do Centro Universitário “Barão de Mauá”, em Ribeirão Preto. Em maio deste mesmo ano, iniciou estágio no Laboratório de Biologia Molecular do Centro de Cana do IAC, participando de projetos realizados pela Profa. Dra. Luciana Rossini Pinto (orientadora) e pelo Dr. Mauro Alexandre Xavier (coorientador), entre 2011 e 2012, sendo bolsista da Fundação de Apoio à Pesquisa Agrícola (FUNDAG), envolvido especialmente com a geração, organização e análise de dados biológicos de marcadores moleculares. Auxiliou vários projetos de alunos de mestrado e doutorado até 2014, quando recebeu o título de Bacharel em Ciências Biológicas. Neste mesmo ano, ingressou no curso de Mestrado do programa de Genética e Melhoramento de Plantas, do Departamento de Agronomia da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, também sob orientação da Profa. Dra. Luciana Rossini Pinto e coorientação do Dr. Marcos Guimarães de Andrade Landell, como bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

“Um livro é feito de uma árvore. É um conjunto de partes planas e flexíveis (ainda chamadas de "folhas") impressas com rabiscos escuros pigmentados. Um olhar para ele e você ouve a voz de outra pessoa, talvez alguém morto há milhares de anos. Através dos milênios, o autor está falando, clara e silenciosamente, dentro de sua cabeça, diretamente a você. Escrever é talvez a maior das invenções humanas, unindo pessoas, cidadãos de épocas distantes, que nunca se conheceram. Livros quebram os grilhões do tempo - prova de que os seres humanos podem fazer magia.”

Carl Sagan

Dedico aos meus pais Celio e Ana Maria e ao meu irmão João Paulo por serem meus exemplos máximos de amor, honestidade, respeito e bondade.

Ofereço à minha namorada Fernanda por todo amor, compreensão e disposição em me ajudar nas decisões e nos momentos difíceis.

AGRADECIMENTOS

Agradeço a **Deus** por todas as coisas.

Aos meus pais **Celio** e **Ana Maria** por tudo aquilo que fizeram (e ainda fazem) por mim, especialmente pela compreensão e entusiasmo a cada nova etapa da minha vida, me incentivando a perseguir meus sonhos ainda que estes pareçam intangíveis.

Ao meu irmão **João Paulo** pelo companheirismo, paciência e compreensão.

A toda a minha família e amigos próximos, em especial **Marco, Jane, Beto, Vera, Paulo, Roberto, Neide, Fernando, Marcelo, Mauri** e **Maria José** por serem pessoas tão amáveis e solícitas, estando sempre comigo, e a **Kita** e **Mauro** pelo bom humor e pela grande amizade sempre.

À minha namorada **Fernanda** por ser uma das melhores pessoas que conheço, e a toda a sua família, **Tida**, avó **Dirce, Guilherme, Brenda** e **Fernando** pela amizade e carinho.

Aos meus queridos professores do Centro Universitário “Barão de Mauá” por toda a base e conhecimento compartilhados conosco durante o curso, em especial para os professores **Marcelo Zini, Ana Rosa, Marcelo Mestriner, Patrícia, Monica, Lucila, Maria Helena, Liliane, Luiz Henrique, Odete, Lucimara, Luis Augusto “Pê”, Abbud, Daniel, Márcia** e **Gláucya**, e também aos meus queridos colegas da XXXVII turma de Ciências Biológicas, em especial **Cauê, Mario, Iuri, Lorena, Beatriz, Jéssica**, e tantos outros, por tornarem a jornada alegre e prazerosa.

Ao **Departamento de Genética e Melhoramento de Plantas** da **UNESP de Jaboticabal** pela oportunidade de cursar o mestrado.

Ao **Instituto Agrônomo de Campinas** pelas oportunidades de realizar meus estágios e minha pesquisa na pós-graduação.

À **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior** pela concessão da bolsa de estudo.

À **Fundação de Amparo à Pesquisa do Estado de São Paulo** (Projeto FAPESP 2013/22500-5) modalidade Auxílio Regular, pelo recurso financeiro para a realização da presente pesquisa.

À Profa. Dra. **Luciana Rossini Pinto** pela oportunidade de ser seu orientado, por toda a confiança, empenho, ensinamentos e conhecimentos compartilhados.

Ao Dr. **Mauro Alexandre Xavier** pelos conselhos, pela grande ajuda e por toda a confiança em mim depositada.

Ao Dr. **Marcos Guimarães de Andrade Landell** pela oportunidade de ser seu coorientado, por me proporcionar a oportunidade de integrar o quadro de estudantes do Centro de Cana do IAC, e por todo aprendizado e auxílio prestado.

À Dra. **Luciana Aparecida Carlini-Garcia** por toda a ajuda nas análises estatísticas do projeto.

A todos os amigos e colegas do Laboratório de Biologia Molecular do Centro de Cana do IAC pela amizade, auxílio e suporte, em especial **Thais, Maicon, Cibeli, Juliana, Maria Letícia, Maria Natália, Izadora, Marcel, Bruna, Fernanda, Alexandre, Larissa, Dr. Michael, Dra. Paula, Dra. Silvana, Roberto, Natália e Luis.**

Aos muitos profissionais do Centro de Cana do IAC que de alguma forma me auxiliaram neste trabalho, em especial ao **Daniel Nunes e Márcio Bidoia** pelo auxílio com a definição dos clones e cultivares deste estudo e ao **Jeremias Mendonça** pelo conhecimento compartilhado.

Aos meus irmãos musicais **Rômulo, André, Rodrigo, Raphael Guzzardi** (*in memoriam*), **Vitor, Júlio, Carol, Lucas, Gabriel, Eduardo e Leo** pelo auxílio na manutenção da minha saúde mental e emocional através da música.

E a todos aqueles que participaram de alguma forma na realização deste estudo.

Muito obrigado!

SUMÁRIO

	Página
RESUMO	iii
ABSTRACT	iv
CAPÍTULO 1 – Considerações gerais	1
Introdução	1
Revisão de Literatura	3
Canavicultura	3
Classificação botânica da Cana-de-açúcar	4
Histórico do início do melhoramento genético da cana-de-açúcar	5
Recursos Genéticos e Bancos de Germoplasma de Cana-de-açúcar	8
Marcadores moleculares e avaliação da variabilidade genética em cana-de-açúcar	11
Avaliação da variabilidade genética em cana-de-açúcar por SSR	14
Referências Bibliográficas	16
CAPÍTULO 2 – Análise da estrutura genética do banco de germoplasma do Programa Cana - IAC por meio de marcadores moleculares	21
Introdução	21
Materiais e Métodos	23
Material Vegetal	23
Extração de DNA e amplificação por PCR	23
Compilação da Matriz de Genotipagem	25
Análise da eficiência dos SSR	25
Análise de Estrutura Genética	27
Análise filogenética e de dissimilaridade	27
Obtenção da coleção nuclear	27
Resultados	28
Eficiência do conjunto de iniciadores SSR	28
Alelos exclusivos e distribuição da frequência alélica	29
Estrutura genética	30
Dissimilaridade e análise filogenética	34
Coleção nuclear	35
Discussão	43
Eficiência dos marcadores microssatélites	43
Distribuição da frequência alélica e alelos exclusivos entre grupos	43
Estruturação, dissimilaridades e análise filogenética	44
Construção da Coleção Nuclear	45
Referências Bibliográficas	46
CAPÍTULO 3 – Investigando a Estrutura Genética das Cultivares Comerciais Brasileiras de Cana-de-Açúcar Através de Marcadores Microssatélites	52
Resumo	52

Introdução	53
Materiais e Métodos	57
Material vegetal	57
Extração de DNA e reações de PCR	57
Análises de dados	58
Análises de eficiência dos iniciadores SSR	59
Análise de diversidade e estrutura genética	60
Análise filogenética e de dissimilaridade	60
Resultados	61
Análise da eficiência dos iniciadores SSR	61
Distribuição da frequência alélica e alelos exclusivos	62
Análise de estrutura e diferenciação genética	63
Tabela 3: Análise de variância molecular (AMOVA) para grupos predefinidos	64
Análise filogenética e de dissimilaridade	65
Tabela 4: Dissimilaridades entre grupos predefinidos de acordo com a função “Extreme Values” do software DARwin.	65
Discussão	66
Eficiência do conjunto de iniciadores SSR para gerenciamento de germoplasma de cana-de-açúcar e identificação de cultivares	66
Estrutura genética das cultivares comerciais brasileiras	67
Análise de agrupamento e de dissimilaridade genética	68
Referências Bibliográficas	71
Figuras e Material suplementar	78
APÊNDICE	88

ANÁLISE DA ESTRUTURA GENÉTICA DO BANCO DE GERMOPLASMA DO PROGRAMA CANA - IAC POR MEIO DE MARCADORES MOLECULARES

RESUMO – A cana-de-açúcar figura mundialmente como uma importante cultura econômica, com grande participação no suprimento da demanda global por energia e açúcar. Bancos ativos de germoplasma são repositórios *in vivo* de variabilidade genética para os programas de melhoramento, estando diretamente relacionados ao sucesso do programa. Para uma gestão adequada e eficiente do banco é necessário conhecer a magnitude e distribuição da variabilidade genética dentro e entre os grupos de acessos das respectivas espécies que o compõe bem como identificar os acessos duplicados ou erroneamente identificados. Marcadores moleculares microssatélites são adequados para tal finalidade. O presente trabalho teve como objetivo principal avaliar a estrutura genética de um grupo de 593 genótipos, constituído por germoplasma básico e melhorado, disponíveis no banco de germoplasma do Programa Cana – IAC. Além disso, a magnitude da diferenciação genética entre as principais cultivares brasileiras e o germoplasma básico de cana-de-açúcar também foi investigada. Para tanto, 12 pares de primers SSR de alto poder discriminatório foram utilizados. A análise da estrutura genética separou os genótipos em dois grupos principais, um constituído por germoplasma básico e o outro por genótipos melhorados. A partição da variabilidade genética mostrou que 14,44% ($P < 0,001$) da variabilidade total detectada pelos marcadores encontra-se entre estes dois grupos. Resultados semelhantes foram observados entre as principais cultivares brasileiras e o germoplasma básico, sendo que 17,91% ($P < 0,001$) da variabilidade presente encontra-se entre esses dois grupos. A análise filogenética agrupou os genótipos conforme suas genealogias, mantendo os materiais melhorados próximos aos acessos de *Saccharum officinarum*, confirmando relacionamento genético mais estreito. Os acessos de *S. spontaneum* apresentaram maior dissimilaridade com os materiais melhorados do que a espécie anterior. O estudo também permitiu estabelecer uma coleção nuclear com 156 genótipos (aproximadamente 30% do total) que apresentam diversidade genética equivalente a presente no grupo original. As informações obtidas poderão ser aplicadas tanto na gestão do banco, direcionando a importação de novos acessos, como também em futuros programas de hibridação para cana convencional ou cana bioenergia.

Palavras-chave: *Saccharum* spp., banco de germoplasma, diversidade genética, microssatélites, SSR, coleção nuclear.

GENETIC STRUCTURE ANALYSIS OF IAC SUGARCANE GERMPLASM BANK BY MOLECULAR MARKERS

ABSTRACT – Sugarcane is a crop with worldwide economic importance due to its large participation in supplying the global demand for energy and sugar. Active germplasm banks are *in vivo* repositories of genetic variability for breeding programs, being directly related to their success. For its adequate management, a constant routine of surveillance and control over the accessions is required. Microsatellite (SSRs) molecular markers are suitable for this purpose. The present work aims to evaluate the genetic structure of a group of 593 individuals, composed by basic germplasm accessions and breeding genotypes, available within the germplasm bank of the Campinas Agronomic Institute (IAC) Sugarcane Genetic Breeding Program, based on these markers, with a focus on the comparison between the genetic diversity of Brazilian clones and cultivars (IAC, CTC, SP and RB) with accessions of diverse species, genera and exotic hybrids. In addition, the magnitude of the genetic differentiation between the main Brazilian cultivars and the basic sugarcane germplasm was also investigated. For this purpose, 12 pairs of SSR primers with high discriminatory power were used. The genetic variability present in the group is structured in a way where breeding materials differ from wild accessions, constituting two groups of individuals. Genetic variability among them is 14.44%, while 85.56% is present within them ($P < 0.001$). Similar results were observed between the main Brazilian cultivars and the basic germplasm, with 17.91% ($P < 0.001$) of the present variability being found between these two groups. Phylogenetic analysis also grouped the individuals according to their genealogies, keeping the bred materials closer to the accessions of *Saccharum officinarum*, confirming a close genetic relationship. The accessions of *S. spontaneum* presented greater dissimilarity from bred materials than the previous species. The analysis of this molecular profile bank allowed the establishment of a core collection within the 593 individuals, comprising 153 genotypes (approximately 30% of the total) with equivalent representativeness of the genetic patrimony of the original group. The information obtained can be applied both in the management of the bank, directing the importation of new accessions, and in future hybridization programs for conventional or bioenergy cane.

Keywords: *Saccharum* spp., germplasm banks, genetic diversity, microsatellites, SSR, core collection.

ANÁLISE DA ESTRUTURA GENÉTICA DO BANCO DE GERMOPLASMA DO PROGRAMA CANA - IAC POR MEIO DE MARCADORES MOLECULARES

CAPÍTULO 1 – Considerações gerais

Introdução

A cana-de-açúcar é atualmente uma importante fonte de energia renovável, responsável por suprir uma crescente demanda por combustível, obtido através do processamento da sacarose pela indústria sucroalcooleira (etanol), e também por energia elétrica, obtida pela cogeração com a queima de biomassa. Assim, a cultura figura como uma sólida alternativa ao uso de combustíveis fósseis. Também possui grande relevância na produção de alimentos, como açúcar, xaropes, e de fármacos, como o ácido glicólico.

O desenvolvimento de cultivares pelos programas de melhoramento genético tem impactado positivamente o setor sucroenergético pela liberação de cultivares mais produtivas, resistentes às principais doenças e pragas, tolerantes aos estresses abióticos e adaptadas às diferentes condições edafoclimáticas do país. Os bancos de germoplasma dos programas de melhoramento genético são fundamentais para o desenvolvimento de novos materiais. Cultivares de cana-de-açúcar com perfil bioenergético tem sido obtidas, principalmente, pelo cruzamento de acessos de *Saccharum spontaneum* (os quais se caracterizam por apresentarem elevados teores de fibra e biomassa) com cultivares comerciais. A quantificação da magnitude da variabilidade genética e sua distribuição entre e dentro dos grupos de acessos constituintes dos bancos de germoplasma são fundamentais para sua gestão adequada. Dentre os países que detém programas de melhoramento genético de cana-de-açúcar, o Brasil é o que lança o maior número de novas cultivares anualmente.

Muito se questiona a respeito do estreitamento da base genética da cana-de-açúcar, decorrente da seleção pelo melhoramento genético. O grau deste estreitamento denota um problema importante para o futuro da cultura, já há muito estudado. Entretanto, a real magnitude deste estreitamento para os materiais

melhorados ainda não é totalmente conhecida. Essa informação pode auxiliar o direcionamento de futuros programas de melhoramento, no sentido de incorporar recursos genéticos ainda não explorados e que poderiam se mostrar valiosos para a continuidade da cultura e do setor sucroalcooleiro. Gêneros e espécies selvagens ainda não são completamente explorados nesse sentido, e possuem grande potencial para a ampliação da base genética. Determinar novas fontes de variabilidade e de recursos genéticos de interesse agrônomo, e estabelecer coleções nucleares dentro dos bancos de germoplasma permite aperfeiçoar e maximizar a exploração de seus recursos genéticos pelos programas de melhoramento.

O Instituto Agrônomo, através do Programa de Melhoramento de Cana, tem concentrado esforços na organização de uma Coleção Pública de Germoplasma representativa do Complexo *Saccharum*. Para atingir tal finalidade, diversas ações já foram iniciadas, como a importação de acessos selvagens de *Saccharum* e outros gêneros de importância (como *Erianthus* e *Miscanthus*) das coleções de Miami (Canal Point, FL, EUA) e Meringa (Austrália). Esses acessos têm grande potencial se utilizados em programas de introgressão genética para o desenvolvimento de cultivares de cana para bioenergia.

Neste sentido, o trabalho aqui apresentado teve como objetivos: i. avaliar a estrutura genética de 593 genótipos (entre acessos das principais espécies de *Saccharum* e gêneros correlatos, cultivares e clones elite do Programa Cana-IAC) disponíveis no Banco de Germoplasma do Programa CANA do Instituto Agrônomo de Campinas (IAC), por meio de marcadores moleculares do tipo microssatélites; e ii. avaliar a variabilidade genética das cultivares comerciais mais plantadas no Brasil e contrastá-la com um grupo de acessos do germoplasma básico de cana-de-açúcar de ampla variabilidade genética.

Revisão de Literatura

Canavicultura

A canavicultura, ou cultura da cana-de-açúcar (*Saccharum* spp.) tem importância econômica mundial principalmente para a produção de açúcar, combustível e eletricidade pela queima de biomassa. O Brasil é o segundo maior produtor de etanol e açúcar, com a segunda maior área plantada no mundo (estimada em 8,9 milhões de hectares) e um rendimento anual superior a 658 milhões de toneladas na safra 2015/16 (CONAB, 2015), o que gerou uma receita estimada em R\$113,27 bilhões para este período, que corresponde a aproximadamente 7% do PIB nacional. Atualmente, a cogeração de eletricidade nas usinas atende a 4% da demanda brasileira de energia elétrica (UNICA, 2016). Tamanho sucesso é atribuído não só a práticas de manejo agrônomo adequadas, mas também ao alto desempenho das cultivares comerciais brasileiras de cana-de-açúcar.

A principal matéria prima do setor sucroalcooleiro é o caldo rico em sacarose que é extraído mecanicamente dos colmos das cultivares comerciais. Os dois produtos principais da indústria sucroalcooleira são o açúcar e o etanol, obtidos pelo processamento desta sacarose. O subproduto do processo industrial – o bagaço – atualmente é queimado nos sistemas de cogeração nas usinas de etanol, gerando eletricidade. Esta eletricidade é capaz de alimentar a própria usina, e o excedente pode ser vendido para as companhias de eletricidade. Dessa forma, além da produção de alimento, a cana-de-açúcar se apresenta como uma sólida fonte renovável de energia limpa e alternativa imediata ao uso de derivados de petróleo e também à geração de eletricidade por usinas hidrelétricas e termelétricas (ANEEL, 2011). A crescente participação da indústria sucroalcooleira na matriz energética brasileira na última década atesta a sua capacidade de atender esta demanda (UNICA, 2010).

Novas práticas e tecnologias atualmente em pesquisa podem ocasionar uma futura mudança de paradigma no setor. Canas híbridas provenientes de cruzamentos entre cultivares atuais com acessos de *S. spontaneum*, intituladas

“canas bioenergia”, ideais para o uso na cogeração por apresentarem alta taxa fotossintética e grande produção de biomassa, podem no futuro substituir grandes áreas dos canaviais atuais para a produção exclusiva de eletricidade pela cogeração nas usinas. A cana bioenergia também apresenta outras características de interesse para a cultura, sendo mais resistentes a estresses bióticos e abióticos. Além disso, há o desenvolvimento da tecnologia do “etanol celulósico” ou “etanol de segunda geração”, que utilizaria um processo enzimático capaz de converter a celulose da parede celular em etanol (LIU et al., 2016; NOVACANA, 2017; SILVA, 2017).

Classificação botânica da Cana-de-açúcar

A cana-de-açúcar é uma gramínea pertencente ao gênero *Saccharum*, classificada na família *Poaceae*, tribo *Andropogoneae* e subtribo *Saccharinae*. O gênero possui 38 espécies validadas segundo The Plant List (2017), com seis espécies principais: *Saccharum officinarum*, *S. spontaneum*, *S. robustum*, *S. barberi*, *S. sinense* e *S. edule*, esta última não sendo considerada para o melhoramento por ser estéril e produzir inflorescências abortivas (D'HONT et al., 1998; GRIVET et al., 2004).

Um reservatório genético expandido pode ser caracterizado ao se considerar gêneros próximos, provavelmente com as mesmas origens genealógicas que originaram a cana-de-açúcar, mas que sofreram especiação ao longo do tempo, como *Erianthus* Michx. sect. *Ripidium* Henrard, *Miscanthus* Anderss. sect. *Diandra* Keng, *Narenga* Bor. e *Sclerostachya* (Hack.) A. Camus, que juntos constituem o Complexo *Saccharum* (AITKEN et al., 2007; DANIELS; ROACH, 1998; MUKHERJEE, 1957).

O gênero *Saccharum* é muito numeroso e diversificado, com representantes ocorrendo em diversas partes do mundo. Acredita-se que suas principais espécies tenham surgido e evoluído no Centro de Origem Indo-Malaio no continente Asiático, segundo classificação proposta por Vavilov (1992), sendo Índia, Indonésia e Nova Guiné os locais com maior riqueza e diversidade genética nativa (JANNOO et al., 1999). O gênero possui representantes por todo o continente Asiático, dispersos em uma faixa que se estende desde as ilhas de Papua-Nova Guiné e Indonésia, passando pela Índia e alcançando até o nordeste do continente Africano (LIU et al.,

2016). *Saccharum spontaneum* é considerada a espécie com maior distribuição geográfica no gênero sendo que “Glagah” era o termo usado pelos nativos que utilizavam a planta para fins diversos (MUTSAERS, 2007). *Saccharum barberi* é a forma da espécie presente na Índia, e *S. sinense* a presente na China.

As cultivares de cana-de-açúcar atuais são híbridos interespecíficos entre *Saccharum officinarum*, também chamada de “cana nobre” (decaplóide - $2n = 80$, $x = 10$ - com colmos coloridos, espessos e ricos em sacarose, altamente suscetível a estresses bióticos e abióticos) e espécies e gêneros selvagens, com ênfase na espécie *S. spontaneum*, octaplóide com cinco citótipos principais ($2n = 64, 80, 96, 112, 128$, $x = 8$) sem conteúdo de açúcar significativo, mas muito valorizado no melhoramento pelo seu alto vigor e rusticidade, sendo resistente a doenças, pragas e estresses abióticos diversos. (D’HONT et al., 1998; PAN; BURNER; LEGENDRE, 2000). É importante ressaltar que a espécie *Saccharum officinarum* não é conhecida em sua forma selvagem, sendo aceita a teoria de que a espécie fora domesticada por tribos nativas pré-históricas que habitavam o arquipélago Malaio, seu principal centro de diversidade (MUKHERJEE, 1957). Além disso, estudos sugerem que *S. officinarum* possa ter origem interespecífica, a partir de hibridações naturais entre *Saccharum robustum* e *S. spontaneum* (BROWN et al., 2007).

Histórico do início do melhoramento genético da cana-de-açúcar

O plantio de cana-de-açúcar no Brasil teve início logo após o seu descobrimento. Com a chegada do capitão Martin Afonso de Souza à recém-descoberta colônia, em 1531, vieram também as primeiras sementes e mudas de cana-de-açúcar e houve então a construção do primeiro engenho de açúcar do Brasil, chamado “Engenho do Senhor Governador”, na capitania de São Vicente, no atual litoral paulista (MACHADO, 2017). A cana-de-açúcar plantada durante todo o período imperial (entre os séculos XVI e XVIII) era conhecida como “Crioula” sendo considerada uma variedade pura de *Saccharum officinarum*, mas que segundo Bremer (1932) era um híbrido, provavelmente originado de um cruzamento interespecífico natural entre *S. officinarum* e *S. barberi*, que foi trazida ao Brasil da Ilha da Madeira pelo capitão. Esta mesma variedade foi muito utilizada em outras áreas do continente americano como México e algumas ilhas da América Central,

pela sua desejada capacidade de produzir açúcar, produto agrícola este cada vez mais requisitado pelo mercado europeu (JAMES, 2004). Apesar de ter apresentado queda de interesse nos séculos seguintes, o açúcar proveniente da cultura rendeu duas vezes mais riqueza para Portugal do que todo o ouro explorado durante o período (MACHADO, 2017).

Não há registros históricos sobre a utilização de outras espécies do gênero *Saccharum* no Brasil colonial. Ainda segundo James (2004), as cultivares plantadas eram conhecidas por outros nomes de acordo com os locais onde eram empregadas. Segundo a literatura, “Otaheite” era o nome da cana-de-açúcar nobre mais utilizada ao redor do mundo, chamada de “Crioula” no Brasil. “Lousier” foi o nome dado a ela nas Ilhas Maurício, nos anos 1730. “Bourbon” foi como ficou conhecida após ter sido levada das Ilhas Maurício para as Índias Ocidentais Francesas em 1793. “Lahaiana” é como era conhecida na Indonésia e no Havaí, e “Vellai” na Índia.

Ao final do século XIX e durante o início do século XX, devido à baixa resistência a doenças e pragas da variedade Otaheite e suas variantes, ocorreram surtos de doenças importantes para a cultura nos principais polos produtores ao redor do mundo, como as causadas pelos fungos *Phaeocystostroma sacchari* (“Podridão da Casca”) e *Glomerella tucumanensis* (“Podridão Vermelha”), que ocasionaram enormes perdas de produção, ameaçando o colapso total do setor. Com isso, novas cultivares melhoradas de cana nobre começam a ser plantadas. No Brasil, a variedade “Cheribon” (ou “Cana Transparente”, obtida em Java, Indonésia) foi plantada a partir da década de 1860, apresentando desempenho muito superior à sua antecessora, sendo mais resistente a doenças (JAMES, 2004). No entanto, os clones de cana nobre cultivados sempre apresentaram susceptibilidade a uma considerável gama de doenças, submetendo a cultura a adaptações (como plantio de sementes em viveiros isolados e posterior transporte das mudas para os canaviais, por exemplo) que aumentavam consideravelmente os custos ao ponto de inviabilizá-la (JESWIET, 1928).

O grande avanço para a canavicultura veio com a descoberta da fertilidade das suas sementes, que tornou então possível obter novas cultivares por meio do cruzamento entre dois parentais diferentes. A apreciação deste fato se deu

primeiramente por J. W. Parris, no ano de 1858 em Barbados, em seguida por Soltwedel na Indonésia, em 1885 e por Harrison e Bovell, também em Barbados, no ano de 1888, cada qual ocorrendo de forma independente. A princípio, os cruzamentos eram realizados com parentais do que então se acreditava serem acessos puros de *S. officinarum*. Entretanto, o conceito de hibridação interespecífica de cana-de-açúcar é atribuído a Soltwedel após tentativas infrutíferas de se cruzar a variedade de *Saccharum officinarum* “Loethers” com o acesso javanês de *Erianthus arundinaceus* “Glonggong” em 1885, e novamente “Loethers” com acessos de *S. spontaneum*, em 1887, na busca por clones resistentes ao mal de “sereh” (BERDING; ROACH, 1998; JESWIET, 1928; LANDELL; SILVA, 2004).

O êxito no desenvolvimento das primeiras cultivares híbridas ocorreu no período entre 1890 e 1911. Soltwedel, Wakker e Wilbrink ampliaram a gama de materiais utilizados em cruzamentos interespecíficos na busca por fontes de resistência às doenças mais importantes da cultura, como a gomose, o mosaico e o mal de “sereh”. Jeswiet e van Harreveld iniciaram experimentos no que seria conhecido como o processo de “nobilização sistemática”, no qual os híbridos interespecíficos F1 são retrocruzados com acessos de *S. officinarum* para recuperar o teor de sacarose de progênies híbridas entre *S. officinarum* e canas de outras espécies, então consideradas “canas selvagens” (DANIELS; ROACH, 1998).

A cultivar javanesa “POJ2878”, chamada de “cana maravilha de Java” foi obtida em 1921 com a hibridação entre o acesso de *S. officinarum* “EK28” e a híbrida “POJ2364” resultante do cruzamento entre “KASSOER” (Cheribon X *S. spontaneum*) e “POJ100” (Bandjarmasin Hitam X Loethers) (JESWIET, 1928). O sucesso decorrente do plantio desta cultivar em áreas acometidas por pragas e doenças comuns para as canas da época reestabeleceu a cultura a partir de 1930, até então à beira da falência (BERDING; ROACH, 1998; JAMES, 2004). Esta cultivar foi também a mais plantada, até 1950, na maior parte do continente Americano e na Indonésia, tendo ocupado mais de 90% da área plantada na Indonésia (MUTSAERS, 2007).

Apesar da notável diversidade genética do gênero *Saccharum*, poucos acessos das principais espécies (*S. officinarum*, *S. spontaneum*, *S. barberi*, *S. sinense* e *S. robustum*) foram exaustivamente utilizados nas primeiras hibridações,

com ênfase para as duas primeiras (DANIELS; ROACH, 1998; JANNOO et al., 1999). Assim, as cultivares nobilizadas tem *S. officinarum* e *S. spontaneum* como principais espécies genitoras (BERDING; ROACH, 1998). Arceneaux (1967) relata que apenas 19 acessos de *S. officinarum* (contando duplicatas como “Lahaiana” e “Vellai”, por exemplo), dois acessos de *S. spontaneum* (“KASSOER” e “COIMBATORE”), quatro de *S. sinense* (“CHUNNEE”, “SARETHA”, “UBA” e “TEKCHA”), um de *S. robustum* (“MOLOKAI1231”) e um de *Sorghum vulgare* (“DURA”) foram utilizadas como parentais na obtenção das primeiras cultivares híbridas na Indonésia, apesar de, por exemplo, já existirem mais de 350 acessos classificados de *S. officinarum* disponíveis (BERDING; ROACH, 1998).

A estreita base genética das cultivares de cana-de-açúcar atuais resultante deste fato é uma questão importante para o melhoramento no âmbito da perda de vigor por estagnação genética causada pela endogamia, e da subutilização de recursos genéticos que poderiam estar associados à resistência ou produtividade (JANNOO et al., 1999). Segundo Arceneaux (1967), a utilização de tamanho recurso genético de forma heurística representa uma tarefa desencorajadora considerando a quantidade de etapas envolvidas em um projeto de melhoramento e o tempo necessário para o seu desenvolvimento. Entretanto, para o mesmo autor, o uso restrito de acessos de *S. officinarum* pode ser justificado tanto pela já descrita capacidade destes acessos de produzir e acumular sacarose, quanto por já serem considerados bons parentais e também pelo simples fato de estarem disponíveis no local naquele momento.

Recursos Genéticos e Bancos de Germoplasma de Cana-de-açúcar

Os programas de melhoramento demandam a disponibilidade imediata de recursos genéticos para serem explorados. Sendo assim, é necessário que possuam acesso a coleções ou bancos de germoplasma. O termo “germoplasma” denota toda e qualquer unidade de conservação de material genético capaz de originar um indivíduo completo e viável. Para vegetais, são consideradas unidades de conservação (1) as suas sementes, (2) propágulos vegetativos capazes de originar um novo indivíduo geneticamente idêntico e (3) o próprio vegetal mantido *in vivo* em ambiente controlado, limpo e livre de pragas e doenças. Portanto, por definição,

banco ou coleção de germoplasma significa banco ou coleção de recursos genéticos prontamente disponíveis para exploração (BERDING; HOGARTH; COX, 2004).

Existem várias coleções de germoplasma de *Saccharum* sp. pelo mundo, mantidas por seus respectivos programas de melhoramento. Entre eles, a Coleção Mundial de Cana-de-Açúcar e Gramíneas Relacionadas - WCSRG ("World Collection of Sugarcane and Related Grasses") do Departamento de Agricultura dos Estados Unidos, em Miami, Flórida, e a Coleção Mundial de Cana-de-açúcar do Centro de Pesquisa de Kannur, em Coimbatore, na Índia, são consideradas as mais importantes coleções de germoplasma de *Saccharum* e gêneros relacionados atualmente em funcionamento. Existem ainda outros bancos e coleções de germoplasma, com significativa participação na pesquisa de cana-de-açúcar. No Brasil, os principais programas de melhoramento também mantêm bancos de germoplasma, como é o caso da RIDESA, CTC, GRANBIO, VIGNIS e IAC.

Além de conservar a riqueza genética para eventual uso no melhoramento, os bancos de germoplasma também possuem a função de promover intercâmbios de acessos, dado que muitos programas de melhoramento dependem da importação de material por estarem em regiões nas quais não ocorrem exemplares nativos do gênero *Saccharum*. Em 1974, foi fundado o Conselho Internacional de Recursos Genéticos de Plantas - IBPGR ("International Board for Plant Genetic Resources"), conhecido a partir de 1994 como Instituto Internacional de Recursos Genéticos de Plantas – IPGRI ("International Plant Genetic Resources Institute"), uma instituição científica autônoma que atua sob a égide do Grupo Consultivo sobre Pesquisa Agrícola Internacional, cuja função é promover uma rede internacional de centros de recursos genéticos, apoiando assim a coleta, conservação, documentação, avaliação e uso de germoplasma vegetal (BERDING; ROACH, 1998; CROFT et al., 1996).

As primeiras coleções de cana-de-açúcar eram constituídas majoritariamente por clones de cana nobre e alguns poucos genótipos contrastantes (ARCENEUX, 1967). A partir da década de 1920, com o sucesso das recém-desenvolvidas canas nobilizadas e o alto valor de melhoramento atribuído aos acessos de espécies selvagens, foram coletados gêneros e espécies diversas do Complexo *Saccharum* nas expedições realizadas nos centros de diversidade, principalmente Índia e Nova

Guiné (BERDING; HOGARTH; COX, 2004). Um exemplo importante disso foi a “Spontaneum Expedition Scheme” (SES), uma expedição de coleta, estudo e exploração completa das características de *S. spontaneum* e gêneros relacionados para a utilização em melhoramento (PANJE; BABU, 1960). Além de incorporar a riqueza genética destes novos acessos e disponibilizá-la no *pool* gênico usado no melhoramento, outro objetivo era o de salvaguardar estes acessos da extinção, uma vez que com o crescimento populacional e a degradação de ambientes selvagens e dos chamados “jardins nativos de biodiversidade”, tornava-se cada vez mais difícil obter tais materiais (BERDING; ROACH, 1998).

Com a movimentação da cana-de-açúcar ao redor do mundo pelas grandes navegações europeias na metade do segundo milênio, ou pelo intercâmbio de materiais entre programas de melhoramento sem os devidos cuidados fitossanitários, houve também a dispersão de pragas e doenças associadas à canavicultura. Várias delas, assim como a própria planta, encontraram ambientes favoráveis ao seu estabelecimento ao redor do mundo, causando desequilíbrio ambiental e sérias implicações para a continuidade da cultura em escala global (CROFT et al., 1996). Diretrizes para o transporte e importação e procedimentos de quarentena para materiais importados se tornaram obrigatórios para programas de melhoramento, visando identificar pragas e doenças e evitar sua disseminação no ambiente no qual serão inseridas (CROFT; THOMPSON; MAGAREY, 2011; FRISON; PUTTER, 1993).

Outro fator importante que reforça a necessidade de se manter bancos de germoplasma é a ocorrência de erosão gênica, um efeito colateral do melhoramento. A seleção exaustiva de certas características desejáveis em um organismo inevitavelmente tende a excluir outras do seu patrimônio genético. Certas regiões gênicas ou mesmo alelos completos e subconjuntos de genes podem ser perdidos e desaparecerem completamente com o tempo. De acordo com a União Internacional para Conservação da Natureza (IUCN), a conservação da diversidade genética é crucial para a manutenção da biodiversidade (FRANKHAM; BALLOU; BRISCOE, 2002; JORDAN et al., 1998). A perda de variabilidade está associada ao estreitamento da base genética, o que implica na perda de biodiversidade e aumenta o risco de ocorrência de doenças ou pragas inesperadas (GASHAW; MEKBIB;

AYANA, 2016). Portanto, bancos e coleções de germoplasma constituem um meio importante de conservar este patrimônio genético.

Marcadores moleculares e avaliação da variabilidade genética em cana-de-açúcar

A organização e identificação dos materiais são determinantes quanto à funcionalidade dos bancos de germoplasma e ao sucesso de seus programas de melhoramento. Geralmente, estes bancos possuem milhares de acessos. Apesar de desejável, a grande quantidade de materiais dificulta o armazenamento e manutenção. A variabilidade genética presente é reconhecidamente subexplorada (NAYAK et al., 2014). A falta de conhecimento a respeito da variabilidade genética total disponível no banco, como também eventuais misturas e duplicidades entre materiais, podem limitar e até mesmo prejudicar o planejamento de cruzamentos futuros (CORDEIRO; PAN; HENRY, 2003). A biotecnologia e a genética molecular oferecem ferramentas capazes de auxiliar o melhoramento de muitas maneiras, inclusive no gerenciamento de rotina dos bancos de germoplasma. Marcadores moleculares são as ferramentas preferidas para essas tarefas (BURNER; PAN; WEBSTER, 1997; PAN et al., 1997; PARIDA et al., 2010).

Considera-se um marcador biológico todo polimorfismo de um caráter capaz de distinguir dois indivíduos diferentes (RAMALHO et al., 2012), sejam morfológicos (fenotípicos) ou moleculares. A identificação oficial de clones e acessos de cana-de-açúcar é realizada de acordo os Descritores Mínimos para Cana-de-Açúcar (MAPA, 1998), baseada em 46 marcadores morfológicos que ocorrem na planta. Embora tenham um papel importante na classificação e organização de acessos, uma chance de erros pode ocorrer uma vez que os fenótipos são afetados pelas condições ambientais (SILVA et al., 2012). Muitos caracteres da cana-de-açúcar podem variar segundo condições edafoclimáticas e fitossanitárias. A classificação por descritores morfológicos nestes casos é comprometida, podendo inclusive ser equivocada e ocasionar misturas de materiais (FERREIRA; GRATTAPAGLIA, 1998; THOKOANE, 1999).

Os marcadores moleculares, por outro lado, detectam polimorfismos na molécula de DNA e são ferramentas complementares para a identificação oficial realizada por descritores morfológicos (MAPA, 2011; PAN, 2006). As diversas

formas dos caracteres de um indivíduo são condicionadas de acordo com a sua constituição genética em interação com o ambiente e fatores externos, ou seja, o fenótipo resulta da interação entre o genótipo com o ambiente. Dessa forma, indivíduos com genótipos iguais podem não expressar um mesmo fenótipo, porém indivíduos iguais sempre terão genótipos iguais. Mutações, adições ou deleções gênicas podem ocasionar mudanças no material genético. Porém, essas mudanças não necessariamente serão herdáveis ou fixadas no indivíduo adulto. Sendo assim, existe alto grau de confiança na identificação de dois indivíduos diferentes a partir de seus genótipos (MARCONI et al., 2011).

Com o descobrimento da molécula de DNA por Watson e Crick (1953) e das enzimas de restrição no final da década de 70, consolida-se o campo da Biotecnologia, que tornou possível acessar o material genético dos indivíduos e suas informações com o uso das suas técnicas. Neste trabalho, discutiremos o uso em especial dos marcadores moleculares. Para a cana-de-açúcar, estes apresentam grande potencial tanto para o estudo e investigação de ancestralidade (CAI et al., 2005; CHEN et al., 2009; DOS SANTOS et al., 2012), diversidade gênica (COSTA et al., 2011; DOS SANTOS et al., 2012; MARCONI et al., 2011; NAYAK et al., 2014; SHARMA et al., 2008), construção de mapas genéticos e grupos de ligação (AITKEN; JACKSON; MCINTYRE, 2005; OLIVEIRA et al., 2007; PALHARES et al., 2012), *fingerprinting* (HAMEED et al., 2012; MACCHERONI et al., 2009; SILVA et al., 2012; SILVA et al., 2008), entre outras funcionalidades, como no auxílio à manutenção de rotina e gerenciamento de bancos e coleções de germoplasma (JARNE; LAGODA, 1996; PALHARES et al., 2012; PERERA et al., 2012).

Existem várias técnicas de marcadores moleculares, cada qual com sua metodologia e finalidades específicas. Algumas das mais utilizadas são a de RFLP (*Restriction Fragment Length Polymorphism*, ou Polimorfismo de Cumprimento de Fragmento de Restrição), a de RAPD (*Random Amplified Polymorphic DNA*, ou Polimorfismo de Amplificação Arbitrária de DNA), que apesar de eficientes, eram técnicas de difícil realização e baixa reprodutibilidade, sendo gradualmente substituídas por técnicas mais modernas, como a de AFLP (*Amplified Fragment Length Polymorphism*, ou Polimorfismo de Tamanho de Fragmento de Restrição), os SSR (*Simple Sequence Repeats*, ou Repetições de Sequências Simples), também

conhecidos como marcadores microssatélites (*Microsatellite Markers*), e os SNP (*Single Nucleotide Polymorphism*, ou Polimorfismo de Nucleotídeo Único).

Os marcadores microssatélites são muito utilizados por serem marcadores codominantes (capazes de detectar todas as formas alélicas que compõe um dado loco gênico, diferenciando homo e heterozigose para organismos diploides), baseados em PCR (Polimerase Chain Reaction), altamente polimórficos (SILVA, 2001), reprodutíveis, com custo relativamente baixo e rapidez na execução do procedimento, com visualização de bandas em géis de agarose ou poliacrilamida (SILVA et al., 2008). As regiões de microssatélites são trechos de DNA compostos por motivos de até seis pares de base que se repetem (a quantidade de repetições destes motivos constitui o polimorfismo explorado pela técnica), não necessariamente inseridos em regiões codificadoras de produtos gênicos, flanqueados por sequências muito conservadas entre espécies e até mesmo entre gêneros relacionados (DOS SANTOS et al., 2012), o que denota a alta transferibilidade da técnica entre organismos diferentes e com certo grau de parentesco (KALIA et al., 2010).

É importante justificar que apesar do caráter codominante da técnica de SSR, para indivíduos poliplóides como os do gênero *Saccharum* sp., não é possível detectar o número de cópias ou doses de cada alelo, apenas os diferentes tipos de alelos apresentados pelo indivíduo no loco em questão (CORDEIRO; PAN; HENRY, 2003; FERREIRA; GRATTAPAGLIA, 1998). Assim, para *Saccharum* sp., SSR é considerada uma técnica de marcador molecular de caráter dominante, e a leitura e pontuação de um experimento de eletroforese de SSR é feita considerando presença (1) ou ausência (0) das marcas por indivíduo no gel, que representam seu haplótipo (HOANG et al., 2015).

Para a utilização da técnica, é necessário o conhecimento prévio da região-alvo a ser estudada. Por exemplo, o par de iniciadores de SSR deve ser complementar às regiões que flanqueiam o microssatélite desejado, de modo a ocorrer hibridação/anelamento entre o iniciador e a fita de DNA desnaturada, pelo processo de PCR (FERREIRA; GRATTAPAGLIA, 1998). Para se conhecer a região-alvo, é necessário o estudo genômico do organismo. O longo e complexo genoma da cana-de-açúcar não foi totalmente sequenciado até o momento. Porém,

marcadores microssatélites também podem ser desenhados por meio dos dados obtidos por prospecção de sequências de ESTs (*Expressed Sequence Tag*), ou Etiquetas de Sequências Expressas (PINTO et al., 2004). Para cana-de-açúcar, existe uma grande quantidade de marcadores SSR derivados de EST. No Brasil, foram desenvolvidos marcadores microssatélites específicos para cana-de-açúcar através do programa SUCEST (VETTORE et al., 2003), que foi um consórcio brasileiro fundado em 1999 com a finalidade de sequenciar ESTs de *Saccharum* sp. e identificar genes expressos em bibliotecas de cDNA criadas a partir de vários tecidos da planta.

Avaliação da variabilidade genética em cana-de-açúcar por SSR

Os marcadores microssatélites são amplamente aplicados no melhoramento genético de *Saccharum* pela sua rapidez, baixo custo e quantidade de informações biológicas geradas (PAN; SCHEFFLER; RICHARD, 2007). Estudos com estes marcadores foram realizados para investigar a variabilidade genética presente tanto nas principais espécies genitoras quanto nos materiais já melhorados, a fim de se determinar quão estreita é a sua base genética e determinar o tamanho dos repositórios de variabilidade genética ainda inexplorados dentre estas espécies e gêneros relacionados no Complexo *Saccharum*.

Chen et al. (2009) utilizaram 20 marcadores microssatélites para avaliar a variabilidade genética presente em 35 cultivares chinesas de cana-de-açúcar e comparar essa variabilidade com a presente nas espécies, determinando que as espécies que mais contribuíram com as cultivares foram, respectivamente, *S. officinarum*, *S. spontaneum* e *S. robustum*, em acordo com o caminho evolutivo especulado na literatura para o gênero *Saccharum*.

Dos Santos et al. (2012) utilizaram apenas três marcadores microssatélites muito polimórficos para a avaliação de diversidade genética de 47 cultivares da RIDESA (Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro), obtendo 124 fragmentos polimórficos capazes de gerar perfis únicos para todos os genótipos, encontrando assim uma estruturação genética complexa dentro da amostra, dado que não houve agrupamentos consistentes entre materiais aparentados, justificada pelas altas poliploidia e heterozigosidade dos materiais.

Nayak et al. (2014) utilizaram 36 marcadores microssatélites para a caracterização dos 1002 acessos presentes na coleção mundial de cana-de-açúcar de Miami, FL, nos Estados Unidos (WCSRG) e a construção de uma coleção nuclear com 300 genótipos que representem a maior diversidade genética possível com o menor número de redundâncias, formando um grupo reduzido com alta representatividade da variabilidade genética da coleção mundial, facilitando o gerenciamento e as questões de logística e armazenamento.

Marcadores microssatélites também foram intensivamente utilizados em estudos de taxonomia do gênero *Saccharum*. Brown et al. (2007) utilizaram marcadores microssatélites no estudo da variabilidade genética presente dentro e entre as principais seis espécies do gênero *Saccharum*, confirmando estudos anteriores que preconizavam o agrupamento taxonômico de todas estas espécies apenas como *S. officinarum* ou *S. spontaneum*, exceto *S. robustum*, por ser uma provável ancestral da primeira espécie. O estudo também aponta para a tese de que *S. barberi* e *S. sinense* seriam híbridos intergenéricos, oriundos do cruzamento entre *S. officinarum* e indivíduos de outros gêneros relacionados, pela grande distância genética apresentada com as outras quatro espécies do gênero *Saccharum*.

Tais informações têm importantes papéis no melhoramento por promoverem uma maior compreensão sobre os recursos genéticos a serem utilizados. A avaliação da variabilidade genética de um banco de germoplasma tem como objetivo determinar os acessos mais contrastantes e orientar o melhorista na escolha dos materiais mais adequados para serem cruzados, direcionando também a importação de materiais exóticos conforme as necessidades do programa de melhoramento ao qual o banco pertence (CROFT et al., 1996).

Referências Bibliográficas

- AITKEN, K. et al. Characterization of intergeneric hybrids of *Erianthus rockii* and *Saccharum* using molecular markers. **Genetic Resources and Crop Evolution**, v. 54, n. 7, p. 1395–1405, 2007.
- AITKEN, K. S.; JACKSON, P. A.; MCINTYRE, C. L. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. **TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik**, v. 110, n. 5, p. 789–801, mar. 2005.
- ANEEL. Fontes renováveis: Biomassa. In: **Atlas de energia elétrica do Brasil: Biomassa**. 1. ed. Brasil: ANEEL, 2011. p. 63–74.
- ARCENEUX, G. Cultivated Sugarcane of the World and their Botanical Derivation. **Proceedings ISSCT**, v. 12, p. 844–854, 1967.
- BERDING, N.; HOGARTH, M.; COX, M. Plant Improvement of Sugarcane. In: JAMES, G. (Ed.). **World Agriculture Series: Sugarcane**. 2. ed. Oxford, UK: Blackwell Publishing, 2004. p. 20–53.
- BERDING, N.; ROACH, B. T. Germplasm collection, maintenance and use. In: HEINZ, D. J. (Ed.). **Sugarcane Improvement through Breeding**. 1. ed. Amsterdam: Elsevier, 1998. v. 11p. 143–210.
- BREMER, G. On the somatic chromosome numbers of sugarcane forms of endogenous cane. **Proceedings ISSCT**, v. 4, n. 20, p. 1–3, 1932.
- BROWN, J. S. et al. Analysis of clonal germplasm from five *Saccharum* species: *S. barberi*, *S. robustum*, *S. officinarum*, *S. sinense* and *S. spontaneum*. A study of inter and intra species relationships using microsatellite markers. **Genetic Resources and Crop Evolution**, v. 54, n. 3, p. 627–648, 2007.
- BURNER, D. M.; PAN, Y.; WEBSTER, R. D. Genetic diversity of North American and Old World *Saccharum* assessed by RAPD analysis. **Genetic Resources and Crop Evolution**, v. 44, p. 235–240, 1997.
- CAI, Q. et al. Verification of the introgression of *Erianthus arundinaceus* germplasm into sugarcane.pdf. **Plant Breeding**, v. 124, p. 322–28, 2005.
- CHEN, P. et al. SSR marker-based analysis of genetic relatedness among sugarcane cultivars (*Saccharum* spp. hybrids) from breeding programs in China and other countries. **Sugar Tech**, v. 11, n. 4, p. 347–354, 2009.
- CONAB. Acompanhamento da Safra Brasileira - Cana-de-açúcar– Safra 2015/16, terceiro levantamento. p. 65, 2015.
- CORDEIRO, G. M.; PAN, Y. B.; HENRY, R. J. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. **Plant Science**, v. 165, n. 1, p. 181–189, 2003.

COSTA, M. L. M. DA et al. Assessment of Genetic Diversity in Contrasting Sugarcane Varieties Using Inter-Simple Sequence Repeat (ISSR) Markers. **American Journal of Plant Sciences**, v. 2, p. 425–432, 2011.

CROFT, B. J. et al. **Sugarcane Germplasm Conservation and Exchange**. 1. ed. Canberra, Australia: Australian Centre for International Agricultural Research, 1996.

CROFT, B.; THOMPSON, N.; MAGAREY, R. Introduction to Sugarcane Quarantine: Instructions for staff of research organizations involved in sugarcane research. **BSES Limited Publication**, p. 35, out. 2011.

D'HONT, A. et al. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. **Genome**, v. 41, n. 2, p. 221–225, 1998.

DANIELS, J.; ROACH, B. T. Taxonomy and Evolution. In: HEINZ, D. J. (Ed.). **Sugarcane Improvement through Breeding**. 1. ed. Amsterdam: Elsevier, 1998. v. 11p. 7–84.

DOS SANTOS, J. M. et al. Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. **Industrial Crops and Products**, v. 40, n. 1, p. 145–150, 2012.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3. ed. Brasília: EMBRAPA, 1998.

FRISON, E. A.; PUTTER, C. A. J. (Eds.) FAO/IBPGR Technical Guidelines for the Safe Movement of Sugarcane Germplasm, Food and Agriculture Organization of the United Nations. Roma, 1993.

FRANKHAM, R.; BALLOU, J. D.; BRISCOE, D. A. **Introduction to Conservation Genetics**. 1. ed. Cambridge, UK: Cambridge University Press, 2002.

GASHAW, E. T.; MEKBIB, F.; AYANA, A. Genetic Diversity among Sugarcane Genotypes Based on Qualitative Traits. **Advances in Agriculture**, v. 2016, Article ID 8909506, p. 8, 2016.

GRIVET, L. et al. A review of recent molecular genetics evidence for sugarcane evolution and domestication. **Ethnobotany Research & Applications**, v. 2, p. 9–17, 2004.

HAMEED, U. et al. Use of simple sequence repeat markers for DNA fingerprinting and diversity analysis of sugarcane (*Saccharum* spp) cultivars resistant and susceptible to red rot. **Genetics and molecular research : GMR**, v. 11, n. 2, p. 1195–204, jan. 2012.

HOANG, N. V. et al. Potential for Genetic Improvement of Sugarcane as a Source of Biomass for Biofuels. **Frontiers in Bioengineering and Biotechnology**, v. 3, November, p. 1–15, 2015.

JAMES, G. **World Agriculture Series: Sugarcane**. 2. ed. Oxford, UK: Blackwell Publishing, 2004.

JANNOO, N. et al. Molecular investigation of the genetic base of sugarcane cultivars. **Theoretical and Applied Genetics**, v. 99, n. 1–2, p. 171–184, 1999.

JARNE, P.; LAGODA, P. J. Microsatellites, from molecules to populations and back. **Trends in ecology & evolution**, v. 11, n. 10, p. 424–9, out. 1996.

JESWIET, J. Immunity and cross-fertilisation in the genus *Saccharum*. **Recueil des Travaux Botaniques Neerlandais**, v. 25A, p. 185–202, 1928.

JORDAN, D. R. et al. Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. **Euphytica**, v. 102, n. 1, p. 1–7, 1998.

KALIA, R. K. et al. Microsatellite markers: an overview of the recent progress in plants. **Euphytica**, v. 177, n. 3, p. 309–334, 9 nov. 2010.

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

LIU, X. et al. Phylogenetic Analysis of Different Ploidy *Saccharum spontaneum* Based on rDNA-ITS Sequences. **Plos One**, v. 11, n. 3. <https://doi.org/10.1371/journal.pone.0151524>, 2016.

MACCHERONI, W. et al. Development of a dependable microsatellite-based fingerprinting system for sugarcane. **Sugar Cane Intl**, v. 27, p. 8–13, 2009.

MACHADO, F. B. P. **Brasil, a doce terra - História do Setor**. Disponível em: <https://www.agencia.cnptia.embrapa.br/Repositorio/historia_da_cana_000fhc62u4b02wyiv80efhb2attuk4ec.pdf>. Acesso em: 31 jan. 2017.

MAPA. **Descritores Mínimos de Cana de Açúcar**. 1. ed. Brasília, DF, Brasil: Diário Oficial da União, 1998.

MAPA. **Proteção de Cultivares no Brasil**. 1. ed. Brasília, DF, Brasil: Biblioteca Nacional de Agricultura, 2011.

MARCONI, T. G. et al. Functional markers for gene mapping and genetic diversity studies in sugarcane. **BMC Research Notes**, v. 4, n. 1, p. 264, jan. 2011.

MUKHERJEE, S. K. Origin and Distribution of *Saccharum*. **Botanical Gazette**, v. 119, n. 1, p. 55–61, 1957.

MUTSAERS, H. **Peasants, Farmers and Scientists**. 1. ed. Dordrecht: Springer Netherlands, 2007.

NAYAK, S. N. et al. Promoting utilization of *Saccharum* spp. genetic resources through genetic diversity analysis and core collection construction. **PloS one**, v. 9, n. 10. <https://doi.org/10.1371/journal.pone.0110856>, jan. 2014.

NOVACANA. **Os processos para obtenção do etanol celulósico**. Disponível em: <<https://www.novacana.com/estudos/os-processos-para-obtencao-do-etanol-celulosico-241013/>>. Acesso em: 31 jan. 2017.

- OLIVEIRA, K. M. et al. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding**, v. 20, n. 3, p. 189–208, 6 abr. 2007.
- PALHARES, A. C. et al. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. **BMC genetics**, v. 13, p. 51, jan. 2012.
- PAN, Y. Highly Polymorphic Microsatellite DNA Markers for Sugarcane Germplasm Evaluation and Variety Identity Testing. **Sugar Tech**, v. 8, n. 4, p. 246–256, 2006.
- PAN, Y. B. et al. Analysis of primer-derived, nonspecific amplification products in RAPD-PCR. **BioTechniques**, v. 22, n. 6, p. 1071–4, 1076–7, jun. 1997.
- PAN, Y. B.; BURNER, D. M.; LEGENDRE, B. L. An assessment of the phylogenetic relationship among sugarcane and related taxa based on the nucleotide sequence of 5S rRNA intergenic spacers. **Genetica**, v. 108, n. 3, p. 285–95, 2000.
- PAN, Y.; SCHEFFLER, B.; RICHARD, J. High-throughput genotyping of commercial sugarcane clones with microsatellite (SSR) DNA markers. **Sugar Technology**, v. 9, p. 176–181, 2007.
- PANJE, R. R.; BABU, C. N. Studies in *Saccharum spontaneum* Distribution and Geographical Association Chromosome Numbers. **Cytologia**, v. 25, p. 152–172, 1960.
- PARIDA, S. K. et al. Functionally relevant microsatellites in sugarcane unigenes. **BMC plant biology**, v. 10, n. 1, p. 251, jan. 2010.
- PERERA, M. F. et al. Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits. **Euphytica**, v. 185, n. 3, p. 491–510, 2012.
- PINTO, L. R. et al. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome / National Research Council Canada - Genome / Conseil national de recherches Canada**, v. 47, n. 5, p. 795–804, 2004.
- RAMALHO, M. A. P. et al. **Genética na Agropecuária**. 5. ed. Lavras, MG: Editora UFLA, 2012.
- SHARMA, R. K. et al. Evaluation of rice and sugarcane SSR markers for phylogenetic and genetic diversity analyses in bamboo 1. v. 103, n. 732, p. 91–103, 2008.
- SILVA, D. C. D. et al. DNA fingerprinting based on simple sequence repeat (SSR) markers in sugarcane clones from the breeding program RIDESA. **African Journal of Biotechnology**, v. 11, n. 21, p. 4722–4728, 13 mar. 2012.
- SILVA, J. A. The Importance of the Wild Cane *Saccharum spontaneum* for Bioenergy. **Sugar Tech**, p. 1–39, 2017.

SILVA, J. A. et al. Sugarcane Variety Identification through DNA Fingerprinting with Microsatellites Markers. **Subtropical Plant Science**, v. 3388, p. 1–7, 2008.

SILVA, J. A. G. Preliminary analysis of microsatellite markers derived from sugarcane expressed sequence tags (ESTs). **Genetics and Molecular Biology**, v. 24, n. 1, p. 155–159, 2001.

The Plant List. Disponível em:

<<http://www.theplantlist.org/tpl1.1/search?q=Saccharum>>. Acesso em: 22 fev. 2017.

THOKOANE, L. N. Progress towards a fingertyping database for sugarcane varieties. **Proc S Afr Sug Technol Ass**, v. 73, p. 164-6, 1999.

UNICA. **Etanol e Bioeletricidade: A cana de açúcar no futuro da matriz energética**. 1. ed. São Paulo, Brasil: LUC Comunicação Integrada, 2010.

UNICA. **Acompanhamento quinzenal da safra 2015/16** Departamento de Economia e Estatística. São Paulo, 2016.

VAVILOV, N. I. The Phytogeographical Basis for Plant Breeding. In LÖVE, D (tradutor) **Origin and Geography of Cultivated Plants**. 1 ed. Cambridge University Press, Cambridge, UK, 1992, p 316.

VETTORE, A. L. et al. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome research**, v. 13, n. 12, p. 2725–35, dez. 2003.

WATSON, J.; CRICK, F. A Structure for Deoxyribose Nucleic Acid. **Nature**, v. 171, p. 737–738, 1953.

CAPÍTULO 2 – Análise da estrutura genética do banco de germoplasma do Programa Cana - IAC por meio de marcadores moleculares

Introdução

O estudo da riqueza e distribuição dos recursos genéticos presentes em um banco de germoplasma possibilita maximizar a sua exploração (CROFT et al., 1996). Os programas de melhoramento de *Saccharum* sp. contam com uma enorme quantidade de genótipos diferentes, englobando inclusive outras espécies que se inter cruzam e originam genótipos viáveis (MUKHERJEE, 1957), por vezes com características de alto interesse agrônomo. É comum a falta de conhecimento sobre a diversidade genética total detida dentro de uma coleção ou banco de germoplasma, o que resulta na sua subexploração ou subutilização e na limitação do potencial de programas de melhoramento (NAIR, 2012).

Análises de diversidade genética baseadas em marcadores moleculares são amplamente utilizadas para avaliar essa variabilidade (MARCONI et al., 2011; OLIVEIRA et al., 2007). Algumas classes de marcadores moleculares são mais adequadas para estudos de diversidade genética em grandes populações por serem altamente polimórficos, rápidos e produzirem grande quantidade de dados, como é o caso da RAPD (HARVEY; BOTHA, 1996; PAN et al., 2004; TABASUM et al., 2010) AFLP (AITKEN et al., 2007; HOARAU et al., 2001; ZABEAU; VOS, 2000) e dos SSR (BROWN et al., 2007; EDMÉ; GLYNN; COMSTOCK, 2006; OLIVEIRA et al., 2009; PAN, 2010). Para *Saccharum*, organismo com alta ploidia e heterozigosidade, os marcadores SSR ou microssatélites são os preferidos por serem altamente polimórficos, de caráter codominante, de rápida realização uma vez consolidado um protocolo de execução, e por acessarem regiões presentes em praticamente todo o seu genoma (PINTO et al., 2006). As informações geradas por meio da avaliação de variabilidade genética com marcadores moleculares são valiosas para o melhorista no intuito de direcionar a escolha de materiais a serem usados em futuros cruzamentos (NAYAK et al., 2014).

O Programa de Melhoramento Genético de Cana do Instituto Agrônomo de Campinas (IAC), localizado na cidade de Ribeirão Preto, SP, no Centro Avançado de

Pesquisa Tecnológica do Agronegócio da Cana (CAPTA Cana), realiza trabalhos com cana-de-açúcar desde os anos 1950. Este programa de melhoramento conta com um banco de germoplasma localizado em Uruçuca, Bahia, com mais de 1000 acessos, englobando as espécies originais de *Saccharum* (*S. officinarum*, *S. spontaneum*, *S. barberi* e *S. robustum*) e gêneros correlatos, como *Erianthus* e *Miscanthus*, além de cultivares comerciais e clones elites. Desde 2012, acessos de *S. spontaneum* estão sendo importados da Coleção Mundial de Cana-de-Açúcar e Gramíneas Relacionadas (WCSRG, ou “World Collection of Sugarcane and Related Grasses” do USDA, localizado em Miami, FL, nos Estados Unidos, para o desenvolvimento do programa Bioenergia, responsável por desenvolver genótipos provenientes de hibridações entre cultivares convencionais com acessos de *S. spontaneum*.

Neste contexto, o objetivo deste trabalho foi o de avaliar a estrutura genética de um grupo de genótipos do banco de germoplasma do programa Cana – IAC utilizando marcadores microssatélites. Este trabalho, inédito no Brasil para este tamanho de amostra, permitirá identificar quais genótipos exibem os perfis mais contrastantes com relação às principais cultivares brasileiras de cana-de-açúcar, para assim explorar mais a fundo toda a riqueza genética disponível no banco.

Materiais e Métodos

Material Vegetal

Um total de 593 genótipos, divididos em cinco grupos separados como: 1) híbridos de *Saccharum officinarum*, abrangendo 309 genótipos, com 228 clones e 81 cultivares brasileiras sendo 43 do Centro de Tecnologia Canavieira (sigla SP ou CTC), 9 do Instituto Agrônomo de Campinas (sigla IAC) e 29 da Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (sigla RB), 2) acessos de *Saccharum officinarum* (52 genótipos), 3) acessos de *S. spontaneum* (149 genótipos), 4) acessos com genealogias ou espécies desconhecidas (28 genótipos) e 5) acessos híbridos ou de espécies diversas (55 genótipos representando *Erianthus* sp., *Miscanthus* sp., *S. barberi*, *S. robustum* e *S. sinense*). Todos estes foram ou coletados da Coleção de Germoplasma de cana-de-açúcar do IAC, mantida no Centro de Cana, Ribeirão Preto, SP, ou importados do Banco de Germoplasma do CTC, em Uruçuca, BA, como também da WCSRG de Miami, FL, Estados Unidos (Tabelas de 1A a 5A do Apêndice, respectivamente). Acessos com nomes duplicados e sem informações sobre suas espécies ou genealogias foram mantidos na tentativa de desvendar suas origens de acordo com seus agrupamentos. As origens dos materiais foram obtidas na página da Rede de Informações de Recursos de Germoplasma (GRIN, ou “Germplasm Resources Information Network” - <https://www.ars-grin.gov>) do USDA, descritas nas Tabelas de 2A a 5A e na Figura 1A do Apêndice.

É válido citar que devido à notável contribuição da cultivar javanesa POJ2878 no melhoramento genético de cana-de-açúcar ao longo do século XX, ela foi incluída no estudo dentro grupo de híbridos de *Saccharum officinarum* (nº 228). É importante citar também que o indivíduo de nº 97 “àUS571415” foi importado com esta grafia e consta como híbrido de *Saccharum officinarum*, motivo pelo qual foi mantido neste grupo.

Extração de DNA e amplificação por PCR

O DNA genômico total foi extraído de folhas jovens ou de meristema apical de acordo com o método CTAB (ALJANABI; FORGET; DOOKUN, 1999). As reações de

PCR foram conduzidas num volume final de 16 µl contendo 30 ng de DNA molde, 0,2 µM de cada iniciador “forward” e “reverse” (marcado com corantes infravermelhos IR700 ou IR800), 100 µM de cada dNTP, 2 mM de MgCl₂, 10 mM de Tris-HCl, 50 mM de KCl e 0,5 unidade de Taq DNA polimerase (Invitrogen, São Paulo, Brasil). As reações foram amplificadas em um termociclador MyCycler® Thermal Cycler (BIO-RAD). As sequências de pares de iniciadores SSR com respectivas temperaturas de anelamento e intervalo esperado de alelos estão descritas na Tabela 1. Para os pares de iniciadores “CV” (*CanaVialis*: CV29, CV37, CV38, CV60, CV79, CV94, CV106), as condições de termociclagem seguiram o protocolo descrito por Maccheroni et al. (2009) e Oliveira et al. (2007). Para os pares de iniciadores “SC” (*SugarCane*: SCB213, SCB381, SCB312, SCB436 e SCC423), as condições de termociclagem foram as seguintes: desnaturação a 94,0°C durante 3 minutos, seguida de 30 ciclos de anelamento (94,0°C durante 1 minuto, temperatura de anelamento específica para cada iniciador), segundo protocolos utilizados por Marconi et al. (2011) e Oliveira et al. (2009), submetidos a pequenos ajustes. Os produtos de amplificação foram separados por eletroforese em gel de poliacrilamida a 6% num analisador de DNA NEN4300 (LI-COR Biosciences®).

Tabela 1: Relação de pares de iniciadores SSR com respectivas temperaturas de anelamento, motivos de repetição, sequências “forward” e “reverse” (5' a 3'), e gama de tamanhos de produtos amplificados (em pares de bases). ^{a, b, c} De acordo com Maccheroni et al. (2009), Canavialis (não publicado) e Marconi et al. (2011), respectivamente.

SSR	Temp. De Anelamento	Motivos repetitivos	Sequência “Forward” Sequência “Reverse”	Gama de Tamanho (PB)
CV29 ^a	60,0°C	(ATCT) ₁₄	TCGCGTCCACCAATGTAACC GCGTGCATCGCTTGTGTCTT	82-170
CV37 ^a	60,0°C	(TTTC) ₁₅	GGATGGACGACGTGTCCTGG ATAAAGTGGCCGCTTGGATTGA	106-177
CV38 ^a	60,0°C	(CTTTT) ₁₈	GAAGCAGGGGCTCAAGTTG GTCAAACAGGCGATCTGGCTC	96-210
CV60 ^b	60,0°C	(CTCTCC) ₅	AATCTGCACCCTGCCCTCTC CAGCTGGAGCATGGATGGAG	158-194
CV79 ^b	60,0°C	(CTATAT) ₁₁ (TATAGA) ₆	GGCACTGCTGGTGGTTGATTG TCCACATCAAGAGGCAGCTA	131-231
CV94 ^b	60,0°C	(AAAAAG) ₅ (CGT) ₅	GGCAGGCCAAGATGAATGAAG AGCACAGCGGAGGGTACGG	182-234
CV106 ^b	60,0°C	(GGC) ₈	AAACAGAGCATACTCGAGGCC ACGTTGCTGACGAGGTTTTCC	78-162
SCB213 ^c	62,0°C	(TCC) ₆	AGCCGTCAGGGGTCAGG ATTCGATGGAGCCTGAGTGAG	247-337
SCB312 ^c	64,7°C	(TCC) ₇	AGTCCGTCGCCGTAATCATCTTG GCACCTCCTCCTTTCTTCTTATT	190-278
SCB381 ^c	60,0°C	(TAC) ₈	TGGAGCTCCGTCTTCTTGTT GCTAGCCCGTACATGCGGTA	190-257
SCB436 ^c	60,0°C	(GAG) ₅	AGTACGCCTGAGTCCTGACG AGGTGCAAGGGCTGATAGAA	158-274
SCC423	----- DISPONÍVEL SOB CONSULTA -----			

Compilação da Matriz de Genotipagem

As imagens dos géis foram manualmente pontuadas e revisadas com o software Saga GT™ (LI-COR Biosciences®). Este software identifica o tamanho do alelo (marca), pontuando “1” como presença ou “0” como ausência do alelo. Utilizou-se o padrão de peso molecular 50-350pb (LI-COR Biosciences®) para calibrar os tamanhos dos alelos amplificados em cada corrida. Os dados binários (presenças e ausências) foram exportados e organizados em uma planilha, montando assim a matriz de genotipagem que foi utilizada em todas as análises.

Análise da eficiência dos SSR

O Conteúdo de Informação de Polimorfismo (PIC) foi calculado de acordo com Cordeiro, Pan e Henry (2003) segundo a expressão

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^n \sum_{j=i+1}^n 2p_i^2 p_j^2$$

na qual p é a frequência do $j^{\text{ésimo}}$ padrão para o alelo i e n é o número de alelos, utilizando o programa “PIC Calculator” (KEMP, 2002) também utilizado para calcular a Diversidade Gênica (D) pela expressão

$$D = 1 - \sum_{i=1}^n P_{ij}^2$$

na qual P_{ij} é a frequência do $j^{\text{ésimo}}$ padrão para o iniciador SSR i , somado entre n padrões (CORDEIRO et al., 2001; NEI, 1973).

A Razão Multiplex (n) foi estimada como o número médio de alelos amplificados por genótipo. A Razão Multiplex Efetiva (E) foi estimada pela expressão

$$E = n\beta$$

onde n é a razão multiplex e β a fração de locos polimórficos, estimada por

$$\beta = \frac{n_p}{(n_p + n_{np})}$$

sendo n_p a soma de marcas polimórficas e n_{np} a soma das não polimórficas. O Índice do Marcador (IM) foi obtido pela expressão $IM = PIC \times E$ (VARSHNEY et al., 2007).

O Poder de Resolução (PR) para cada par de iniciadores foi calculado a partir da matriz de genotipagem utilizando a expressão

$$PR = \sum Ib$$

onde Ib é a informação alélica dada por

$$Ib = 1 - (2 \times |0.5 - M|)$$

na qual M é a proporção dos genótipos totais que contêm o alelo (PREVOST; WILKINSON, 1999).

Análise de Estrutura Genética

O número de subpopulações (K) e a proporção de identidade entre os genótipos foram inferidos usando o software “Structure” v.2.3.4 (EVANNO; REGNAUT; GOUDET, 2005; PRITCHARD; STEPHENS; DONNELLY, 2000). O valor de K foi ajustado de 1 a 10 (“K-value”), com 10 iterações a um período de amostragem de 100.000 repetições, sob 200.000 repetições de “Markov chain Monte Carlo” (MCMC). Os melhores valores de K e ΔK foram obtidos a partir da prospecção dos resultados pelo software “Structure Harvester” (EARL; VON-HOLDT, 2010). As análises da variância molecular (AMOVA) entre os cinco grupos predefinidos e os grupos inferidos pela análise do “Structure” foram realizadas pelo software “Arlequin” v.3.5 (EXCOFFIER; LISCHER, 2010), também utilizado para determinar o índice de diferenciação (F_{ST}) entre os cinco grupos.

Análise filogenética e de dissimilaridade

A dissimilaridade entre genótipos foi estimada com base no coeficiente de dissimilaridade de Jaccard ($Diss = 1 - Jaccard$). A árvore filogenética foi construída de acordo com o método “UnWeighted Neighbor Joining” com 1000 amostragens usando o software “DARwin” v.6.0.13 (PERRIER; JACQUEMOUD-COLLET, 2006).

Obtenção da coleção nuclear

A função “Max Length Subtree” do software DARwin foi utilizada para definir a coleção nuclear (SCHOEN; BROWN, 1993) com base na riqueza alélica encontrada a partir dos marcadores microssatélites, e a partir disso, gerar as respectivas árvores filogenéticas para cada um dos grupos predefinidos (estratos), selecionando de 25 a 30% dos materiais presentes em cada um. Foram considerados fatores intrínsecos do manejo e da manutenção dos bancos de germoplasma, que implicam em grande disponibilidade de recursos humanos e financeiros. Nesse contexto, é válida a definição de um grupo menor de genótipos que apresente riqueza genética igual ou equivalente à do banco de germoplasma da qual foi selecionado, visando reduzir custos para programas de melhoramento de pequeno e médio porte.

Resultados

Eficiência do conjunto de iniciadores SSR

Os 12 iniciadores SSR geraram 292 alelos para os 593 genótipos. As estimativas de polimorfismo e demais informações referentes a cada iniciador SSR são mostrados na Tabela 2.

Tabela 2: Número total de marcas amplificadas (n) e estimativas de Razão Multiplex Efetiva (E), Conteúdo de Informação de Polimorfismo (PIC), Diversidade Gênica (D), Índice de Marcador (IM), Poder de Resolução (PR) e Perfis Identificados (P. Id.) com total e porcentagem, para cada iniciador SSR.

Iniciadores SSR	n	E	PIC	D	IM	PR	P. Id. (%)
CV29	23	6,686	0,919	0,924	6,180	10,600	459 (77,40%)
CV37	20	4,592	0,879	0,889	4,037	6,192	272 (45,87%)
CV38	28	6,894	0,937	0,940	6,462	11,902	522 (88,03%)
CV60	10	3,123	0,787	0,812	2,457	3,669	92 (15,51%)
CV79	27	5,573	0,905	0,911	5,042	7,855	392 (66,10%)
CV94	16	4,064	0,861	0,873	3,500	5,693	156 (26,31%)
CV106	15	4,196	0,868	0,880	3,643	6,445	210 (35,41%)
SCB213	32	5,567	0,930	0,934	5,175	10,374	469 (79,09%)
SCB312	30	6,179	0,930	0,934	5,748	10,351	525 (88,53%)
SCB381	24	6,214	0,916	0,922	5,694	9,821	477 (80,44%)
SCB436	24	4,293	0,875	0,885	3,755	6,108	240 (40,47%)
SCC423	43	7,277	0,932	0,936	6,780	11,420	528 (89,04%)
Total	292	-	-	-	-	-	-
Média	24,33	5,388	0,895	0,903	4,873	8,369	362 (61,02%)

O painel de iniciadores SSR mostrou bom desempenho na identificação dos 593 genótipos. O número de marcas amplificadas por iniciador variou de 43 (SCC423) a 10 (CV60), com média de 24,33 alelos por iniciador. Em geral, os iniciadores CV29, CV38, CV79, SCB213, SCB312 e SCC423 mostraram os maiores valores para as estimativas empregadas no estudo. A avaliação dos haplótipos revelou que com quatro iniciadores SSR (CV38, SCB312, SCB381 e SCC423) foi possível identificar mais de 80% dos genótipos, enquanto que com CV29, CV37, CV79, SCB213 e SCB436 foi possível diferenciar entre 40 e 80% deles. Os iniciadores SSR com menor capacidade de identificação foram CV60, CV94 e CV106, gerando perfis únicos para apenas 15,51%, 26,31% e 35,41% dos genótipos, respectivamente (Tabela 2).

As estimativas para o Conteúdo de Informação de Polimorfismo (PIC) variaram de 0,787 (CV60) a 0,937 (CV38) com um valor médio de 0,895. As estimativas para Diversidade Gênica (D) variaram de 0,804 (CV94) a 0,938 (CV38) com média de 0,903. Os valores da Razão Multiplex Efetiva variaram de 3,123 (CV60) a 7,227 (SCC423), com média de 5,388. Os valores do Índice de Marcador variaram de 2,457 (CV60) a 6,780 (SCC423) com média de 4,873.

O Poder de Resolução (PR) foi calculado para avaliar a capacidade de cada iniciador SSR de gerar perfis polimórficos para todos os genótipos. Os valores variaram de 3,669 (CV60) a 11,902 (CV38).

A melhor correspondência para a variação entre o número de perfis identificados e as medidas de polimorfismo para cada iniciador SSR (PIC, MI, E, D e PR) foi dada pelo Poder de Resolução.

Alelos exclusivos e distribuição da frequência alélica

O total de alelos amplificados em cada grupo foi 191 (65,41%) nos Híbridos de *Saccharum officinarum* (IAC, CTC, SP, RB, incluindo a cultivar histórica POJ2878), 166 (56,85%) para as cultivares brasileiras, 220 (75,34%) para os acessos de *Saccharum officinarum*, 273 (93,49%) para os acessos de *S. spontaneum*, 222 (76,03%) para os acessos com genealogia desconhecida, e 253 (86,64 %) para os híbridos exóticos e espécies diversas. Reportamos a ocorrência de alelos exclusivos entre os grupos (ou seja, alelos que ocorrem apenas em alguns genótipos dentro de cada grupo), sendo um (0,17%) nos Híbridos de *Saccharum officinarum* (CV38.0217), 19 (3,20%) nos acessos de *S. spontaneum* (CV29.0082, CV37.0173, CV37.0106, CV94.0217, CV94.0178, SCB213.0255, SCB213.0251, SCB213.0247, SCB213.0243, SCB213.0240, SCB312.0285, SCB312.0200, SCB312.0190, SCB381.0238, SCB381.0190, SCB436.0274, SCB436.0251, SCC423.0179 e SCC423.0170), e 10 (1,69%) nos híbridos exóticos e espécies diversas (CV37.0177, SCB436.0163, SCB436.0158, SCC423.0288, SCC423.0285, SCC423.0175, SCC423.0158, SCC423.0122, SCC423.0116, SCC423.0112).

Com relação aos genótipos que contém estes alelos exclusivos, reportamos apenas um Híbrido de *Saccharum officinarum* (IAC912205), 44 acessos de *S. spontaneum* (GLAGAHKLOET, SSPONT4593, IN8403336, US5601601, SES184A,

NG77176, IN8488(2), SH301, PONAPPEWILD, SES365, YACHENG12, SES323, SES113A, SES196, PBG841216, PURPLE32, SES2081, FORMOSA, FORMOSA429, GEHRABON, PHILLIPINEWILD, SES519, IS76177, PANAMA, SES07322, KAWANDA, CANE604417, 78217273, B68264, IN76086, IN8404237, HOLES1, IK76006, TAIWAN1140, KLOET, 51NG26, BROWN33, IN81014, IN8109M2, PCAV8469, PCAV8413, SES35379, M.MOENTAI e SES2342) e quatro híbridos exóticos ou espécies diversas (*Miscanthus*, CHUNEE, US571415(2) e 57NG2) contendo estes alelos exclusivos de cada grupo, podendo ocorrer nestes genótipos apenas um alelo ou mais de um. As distribuições alélicas para cada iniciador SSR se encontram nas Figuras 2A até 4A do Apêndice.

Estrutura genética

A análise de estrutura genética realizada pelo software “Structure” a partir da informação genotípica encontrada com os 12 iniciadores SSR nos 593 genótipos do estudo gerou 10 subconjuntos de agrupamentos para cada valor de K (totalizando 100 arquivos individuais), que foram então analisados pelo software “Structure Harvester”. Os resultados foram compactados em um arquivo “.zip” e submetidos diretamente no site do programa. Esta análise retornou K=2 (número de subpopulações, K=10 não mostrado) como o de maior significância para a inferência de estrutura na amostra ($\Delta K = 1454.96$), que significa que todos os genótipos podem ser divididos em duas subpopulações (Figura 1a). Este resultado foi confirmado por uma segunda análise, utilizando o “pipeline Best K” do software CLUMPAK (KOPELMAN et al., 2012).

Decidimos gerar as imagens de proporção de identidade tanto para o número hipotético de grupos predefinido na organização dos materiais (cinco subgrupos na população, ou K=5) quanto para o valor de K com maior significância segundo o “Structure Harvester” (Figura 1b). Na parte superior (K=5) é possível verificar que a estruturação genética detectada delimita alguns grupos em acordo com suas genealogias, especialmente os Clones, as Cultivares, e as espécies *S. officinarum* e *S. spontaneum*.

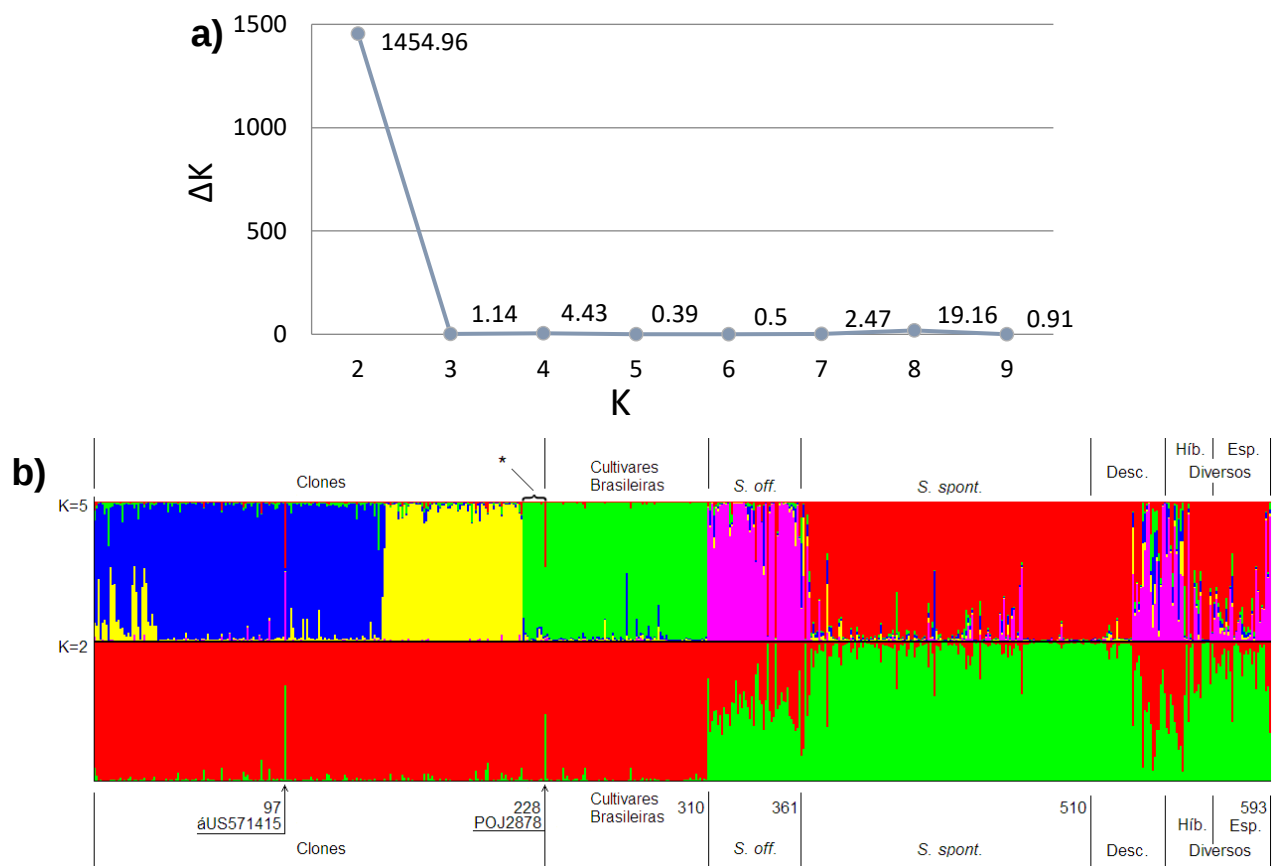


Figura 1: a) Análise do melhor valor de K para os 593 indivíduos, mostrando que o único valor significativo para a inferência dos agrupamentos está em $K=2$ ($\Delta K = 1454,96$); **b)** proporções de identidade entre os grupos para cada indivíduo. Cada linha vertical representa um indivíduo e as cores indicam a proporção de identidade de cada indivíduo para cada uma das subpopulações inferidas. Na parte superior da figura, o padrão de agrupamento considerando 5 grupos ($K=5$). Os indivíduos na pequena faixa identificada com o asterisco representam onze clones que possuem maior similaridade genética com as cultivares do que com os outros clones. Na parte inferior da figura, o resultado da análise para $K=2$, valor da estrutura genética inferida com maior significância.

Na parte inferior da Figura 1b ($K=2$), é possível distinguir claramente que os Clones e Cultivares agrupam-se de forma coesa e quase uniforme com relação às proporções de identidade de cada indivíduo. Os acessos de *S. officinarum*, *S. spontaneum*, os acessos com genealogias desconhecidas, os híbridos exóticos e espécies diversas (*Erianthus* sp., *Miscanthus* sp., *S. barberi*, *S. robustum* e *S. sinense*) formam o segundo grupo, mais heterogêneo e com proporções de identidade variadas, o que pode ser interpretado como a separação dos indivíduos entre genótipos melhorados e acessos básicos. Para o primeiro grupo, duas discrepâncias foram observadas. O genótipo de número 97 “àUS571415”, tido como

híbrido de *Saccharum officinarum* (o qual acreditamos ser um material trocado ou com genealogia diferente da disponível) difere significativamente dos demais genótipos do mesmo grupo. Da mesma forma, o de número 228, a cultivar histórica “POJ2878”, inserido no final do grupo de híbridos de *S. officinarum*, também apresenta uma proporção de identidade diferente dos demais indivíduos no grupo, provavelmente por se tratar de um material desenvolvido no início do melhoramento interespecífico da cana-de-açúcar (MUTSAERS, 2007), desenvolvido em estágios iniciais de retrocruzamento (Figura 2).

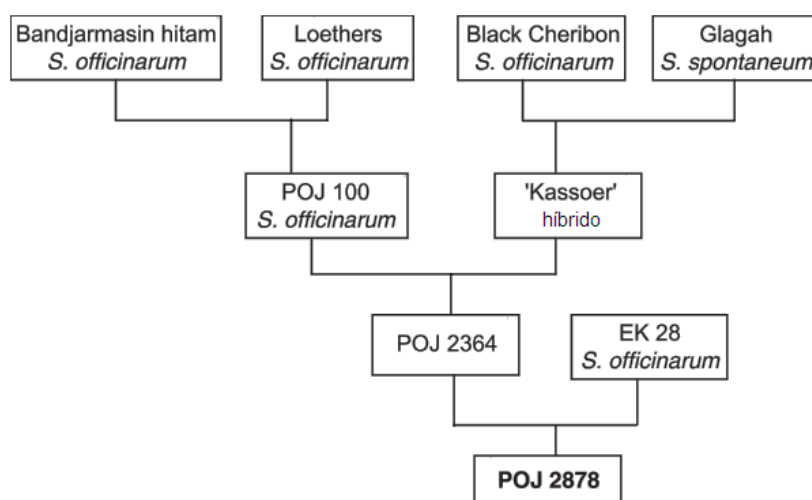


Figura 2: Genealogia da cultivar javanesa "POJ2878", desenvolvida entre 1920 e 1930, na Estação Experimental “Oost Java” localizada na Indonésia. Adaptado de Mutsaers (2007).

A Análise de Variância Molecular (AMOVA) foi realizada para os cinco grupos predefinidos e para os dois grupos inferidos na análise do “Structure”. No primeiro caso, a estimativa de diferenciação genética ($\widehat{\Phi_{ST}}$) obtida entre os cinco grupos foi 0,165 ($P < 0,001$), o que implica que 16,51% da diversidade total está distribuída entre os grupos, enquanto que o restante (83,49%) se distribui dentro dos grupos. Para o segundo caso, a estimativa obtida foi de 0,144 ($P < 0,001$), indicando que 14,44% e 85,66% de toda a variabilidade genética está distribuída entre e dentro dos dois grupos, respectivamente (Tabela 3).

Tabela 3: Análise de Variância Molecular (AMOVA) para os cinco grupos predefinidos (Clones e Cultivares; *S. officinarum*; *S. spontaneum*; Desconhecidos; Híbridos e Espécies Diversas) e para o inferido pela análise de estrutura ($P < 0,001$).

População	Fonte de Variação	Graus de Liberdade	Soma dos Quadrados	Estimativas de variação	Porcentagem de variação
Cinco grupos predefinidos	Entre as populações	4	2287,005	5,64008 Ve	16,51%
	Dentro das populações	588	16769,698	28,51989 Vd	83,49%
	Total	592	19056,703	34,15998	
	Índice de Fixação ($\widehat{\Phi ST}_1$):	0.1651			
Dois grupos inferidos pelo Structure	Entre as populações	1	1513,641	5,01120 Ve	14,44%
	Dentro das populações	591	17543,062	29,68369 Vd	85,56%
	Total	592	19056,703	34,69489	
	Índice de Fixação ($\widehat{\Phi ST}_2$):	0.1444			

O Índice de Diferenciação (F_{ST}) foi calculado entre os cinco grupos predefinidos (Tabela 4). Os valores variam de 0 a 1, sendo que 0 representa completa panmixia entre duas populações (total fluxo gênico), valores acima de 0 e abaixo de 1 indicam estruturação da diversidade genética entre as duas populações (fluxo gênico parcial), e 1 indica total separação entre as populações, portanto fluxo gênico inexistente (WEIR; COCKERHAM, 1984). Os resultados entre grupos distintos variaram entre 0.0494 (*S. spontaneum* X *S. officinarum*) e 0.2031 (Híbridos e espécies diversas X *S. spontaneum*), indicando a menor e maior distância genética entre grupos, respectivamente. Clones e Cultivares apresentaram as menores distâncias entre *S. officinarum* (0.0534) e *S. spontaneum* (0.0787).

Tabela 4: Índice de Diferenciação (F_{ST}) calculado pelo software "Arlequin" entre cada um dos cinco grupos predefinidos.

	Clones e Cultivares	<i>S. officinarum</i>	<i>S. spontaneum</i>	Desconhecidas	Híbridos e espécies diversas
Clones e Cultivares	0				
<i>Saccharum officinarum</i>	0.0534	0			
<i>Saccharum spontaneum</i>	0.0787	0.0494	0		
Desconhecidas	0.0942	0.1002	0.1784	0	
Híbridos e espécies diversas	0.1561	0.1383	0.2031	0.1854	0

Dissimilaridade e análise filogenética

A matriz de dissimilaridades foi gerada a partir da matriz de genotipagem pelo software “DARwin” (Figura 3). A partir dela, foi possível traçar uma árvore filogenética na qual constam todos os 593 indivíduos, agrupados de acordo com os seus dados de dissimilaridades, também pelo software “DARwin”. Pela construção da árvore filogenética, é possível identificar o agrupamento dos indivíduos conforme suas genealogias, formando dois grandes grupos de indivíduos, em consonância com os resultados da análise de estrutura. Entretanto, é possível identificar na Figura 3 agrupamentos menores dentro do Grupo 1, especialmente considerando as Cultivares Brasileiras, os Clones, acessos de *Saccharum officinarum* e genótipos com genealogias desconhecidas.

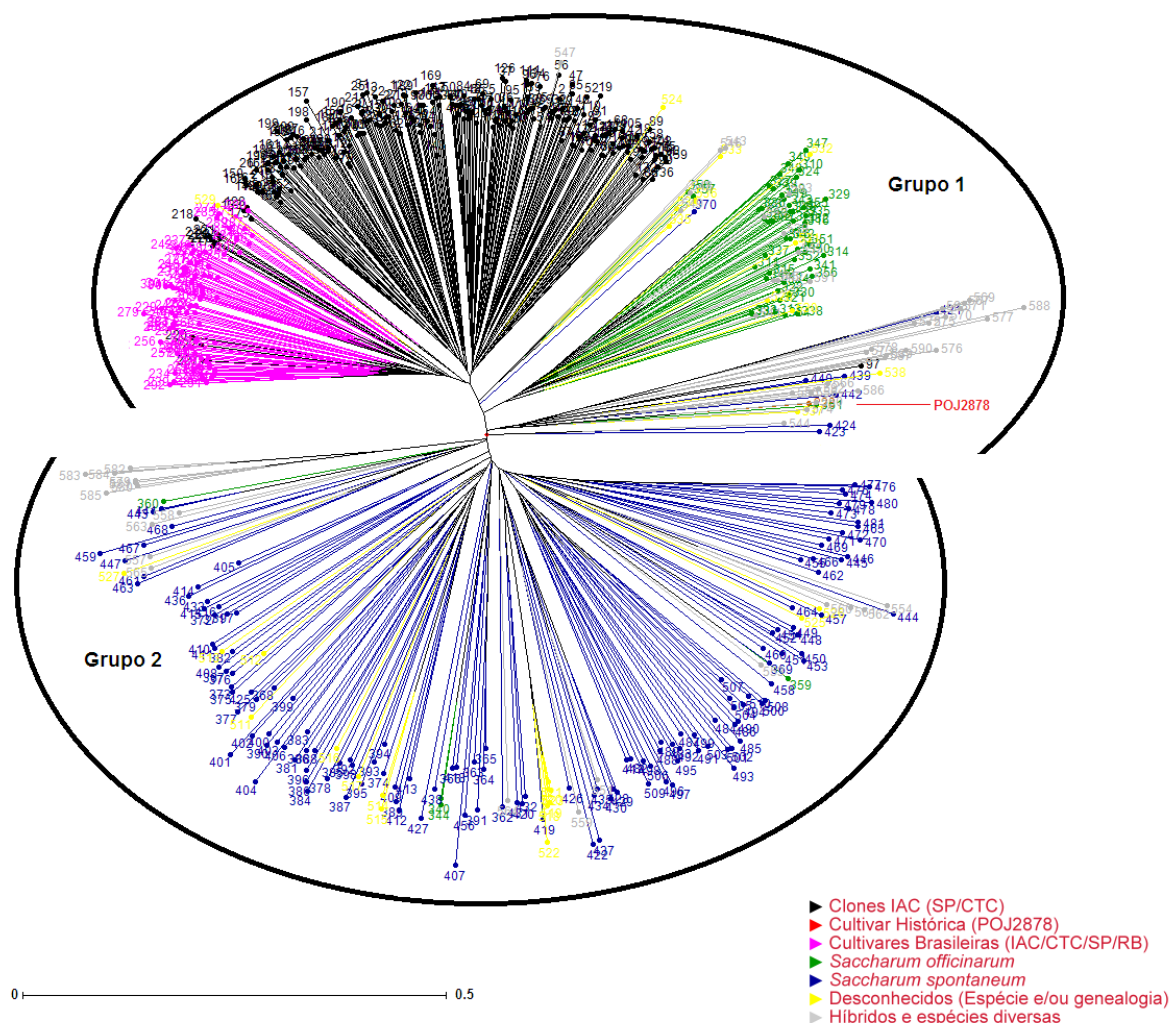


Figura 3: Árvore filogenética obtida pela análise “Neighbor Joining” feita no software “DARwin”.

Genótipos com genealogias desconhecidas (em amarelo) e híbridos exóticos e espécies diversas (em cinza) se espalharam ao longo da árvore, o que já era esperado. Porém, é importante trazer a atenção para algumas discrepâncias relatadas. Treze (13) Clones brasileiros se agruparam junto às cultivares (IACSP042286, IACSP013185, IACSP032030, IACSP982072, IACSP976610, IACSP993013, IACSP973331, IACSP983099, IACSP993039, IACSP993009, IACSP976615, IACSP993090 e IACSP956114), juntamente com um genótipo com genealogia desconhecida (HJ5741). Estes treze clones aparecem na região indicada com um asterisco na parte superior da Figura 1b, juntamente com o clone IACSP985010, que na árvore filogenética encontra-se no subgrupo dos clones. O híbrido exótico “ENDOR” (Uba X 28MQ674) e o acesso desconhecido “P571504” agruparam-se junto aos Clones. Os acessos em duplicidade IN8488(1) (*S. officinarum*) e IN8488(2) (*S. spontaneum*) agruparam-se de forma inversa às suas genealogias cadastradas, indicando que são trocas de material. O clone “àUS571415” agrupou-se fora do grupo dos Clones e próximo aos quatro acessos em duplicidade de *S. robustum* “US571415(1-4)”, confirmando uma suspeita de que é um material com erros tanto na grafia quanto na classificação genealógica. A cultivar histórica “POJ2878” agrupou-se fora dos grupos de cultivares, em um dos subgrupos formados no Grupo 1, juntamente com outros acessos de espécies diversas e híbridos exóticos, e como já esperado, muito próximo de uma de suas parentais, o acesso “KASSOER” que é um provável híbrido interespecífico natural. Os acessos de *S. officinarum* “MANTEIGA”, “IS76155” e “IN845” agruparam-se junto aos acessos de *S. spontaneum*, levantando dúvidas quanto às suas genealogias ou etiquetas.

Coleção nuclear

A função “Max Length Subtree” implementada pelo software DARwin selecionou entre 25 a 30% dos integrantes de cada grupo predefinido, totalizando 76 clones e cultivares (IACSP991305, IACSP992207, IACSP982025, IACSP012409, IACSP022065, POJ2878, IACCTC053516, IACSP023093, RB925345, IACSP015512, IACSP951017, SP832847, IACSP032071, IACSP022123, IACCTC053514, IACSP991018, IACSP987043, IACCTC052513, IACSP022018,

IACSP022006, IACCTC063709, IACSP966046, IACSP042286, RB935744, IACSP993427, IACSP023045, IACSP032092, IACSP991301, IACSP032035, IACCTC052534, IAC912205, CTC07, RB925268, IACSP012411, RB845210, IAC862480, IACCTC062757, IACSP013046, IACSP944004, CTC01, IACSP987061, IACCTC063741, IACSP982072, IACSP023168, IACSP993060, SP801836, IACSP977018, IACSP018002, IACSP955000, IACSP018152, IACSP013185, IACSP032086, IACSP993010, IACSP993088, IACSP976610, RB855511, IACSP012428, SP841201, IACSP023134, IACSP991306, IACCTC052505, IACSP993013, IACSP973353, IACSP032082, IACCTC053585, IACSP018034, SP87365, IACSP013127, IACSP993425, IACSP023219, IACSP022008, IACSP042515, IACSP994011, IACSP993334, SP87396 e IACSP983044 – o indivíduo àUS571415 foi selecionado pelo programa, mas foi retirado após a análise de agrupamento confirmar se tratar de um material que não pertencente a este grupo), 15 acessos de *S. officinarum* (SAC.OFF.8280, IN845, IN84105, IJ76470, ZWARTMANILLA, AJAX, NG57213, IJ76317, IJ76361, ZOPILOTA, BRAVADEPERICO, RP8, IS76155, IN8488(1) e NG2121), 40 acessos de *S. spontaneum* (SES090, SES033, PCAV8420, IS76173, PSOR8413, IN81101, IS76125, SES208, IN84011, IK6006, HOLES1, DACCA, SES184B, YACHENG12, US78519, SES196, US78523, US5601601, PCANOR8902, KLOET, BROWN33, SES27710, DJANTOER, US851008(2), SES113A, COIMBATORE, IN76086, SSPONT15, IND823184, IS761715, IS76177, TONGZA, SES35377, IN8403336, SES288(2), IN81014, US85108, IJ76386, NG21035 e IM76234), 10 acessos com genealogias desconhecidas (PIN8401, IND81170B, NG57222, P571504, IJ3852218, SLC9234, XMISCANTHUS, MUNDAU, CINCA77381 e NG517221) e 15 híbridos exóticos ou acessos de espécies diversas (IM76232(1), FIJI10, CHUNNEE, 57NG2, IJ76358, MATNASHAHJ, US571415(2), IN8458, US74069, CAYANA, ENDOR, KASSOER, NG5712, UK3739 e IJ76389), respectivamente apresentados nas Figuras de 4 a 8. A Figura 9 representa a árvore filogenética com todos os 593 indivíduos, identificados entre selecionados e não selecionados.

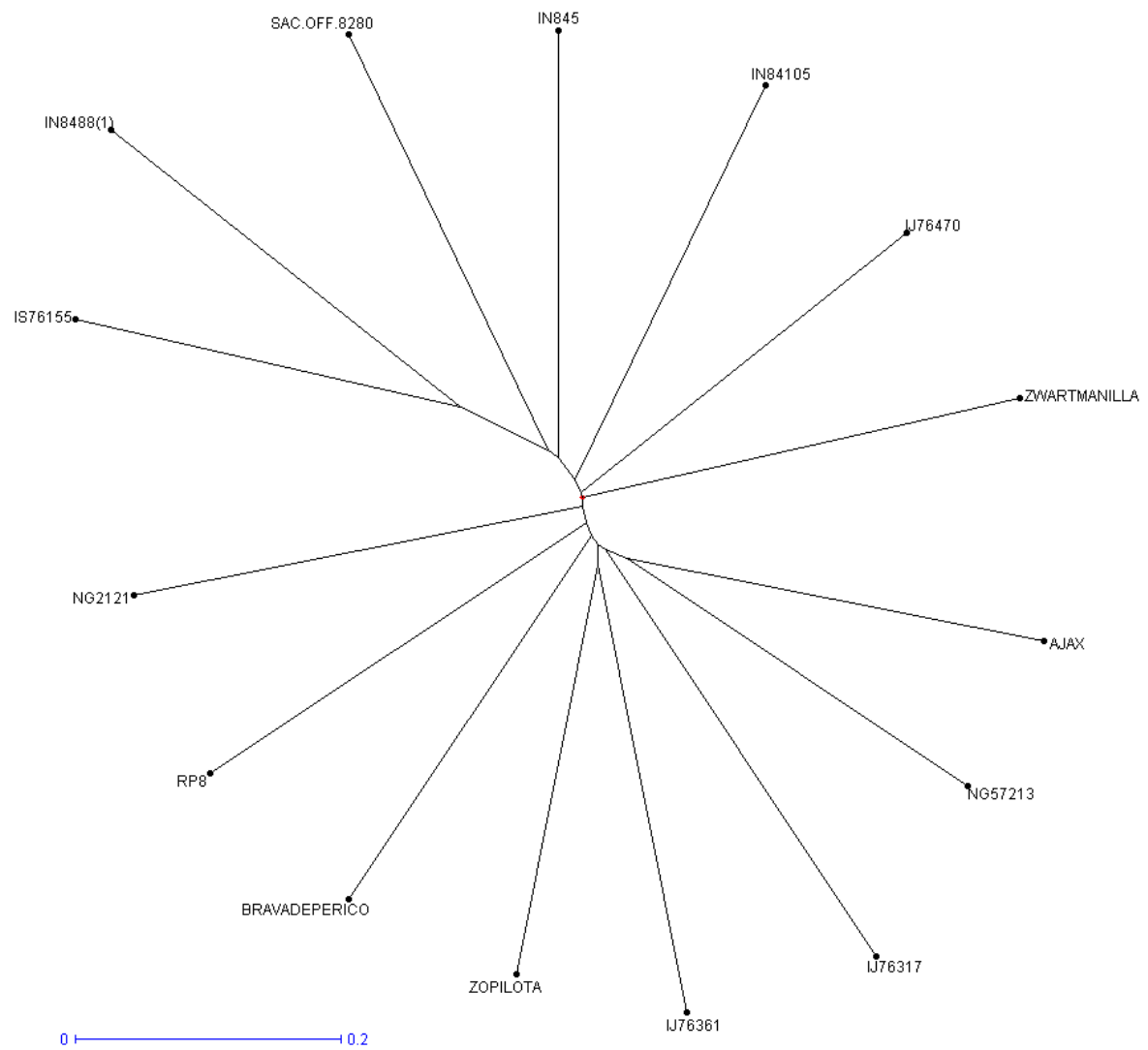


Figura 5: Árvore filogenética contendo os 15 acessos de *S. officinarum* selecionados para a coleção nuclear.

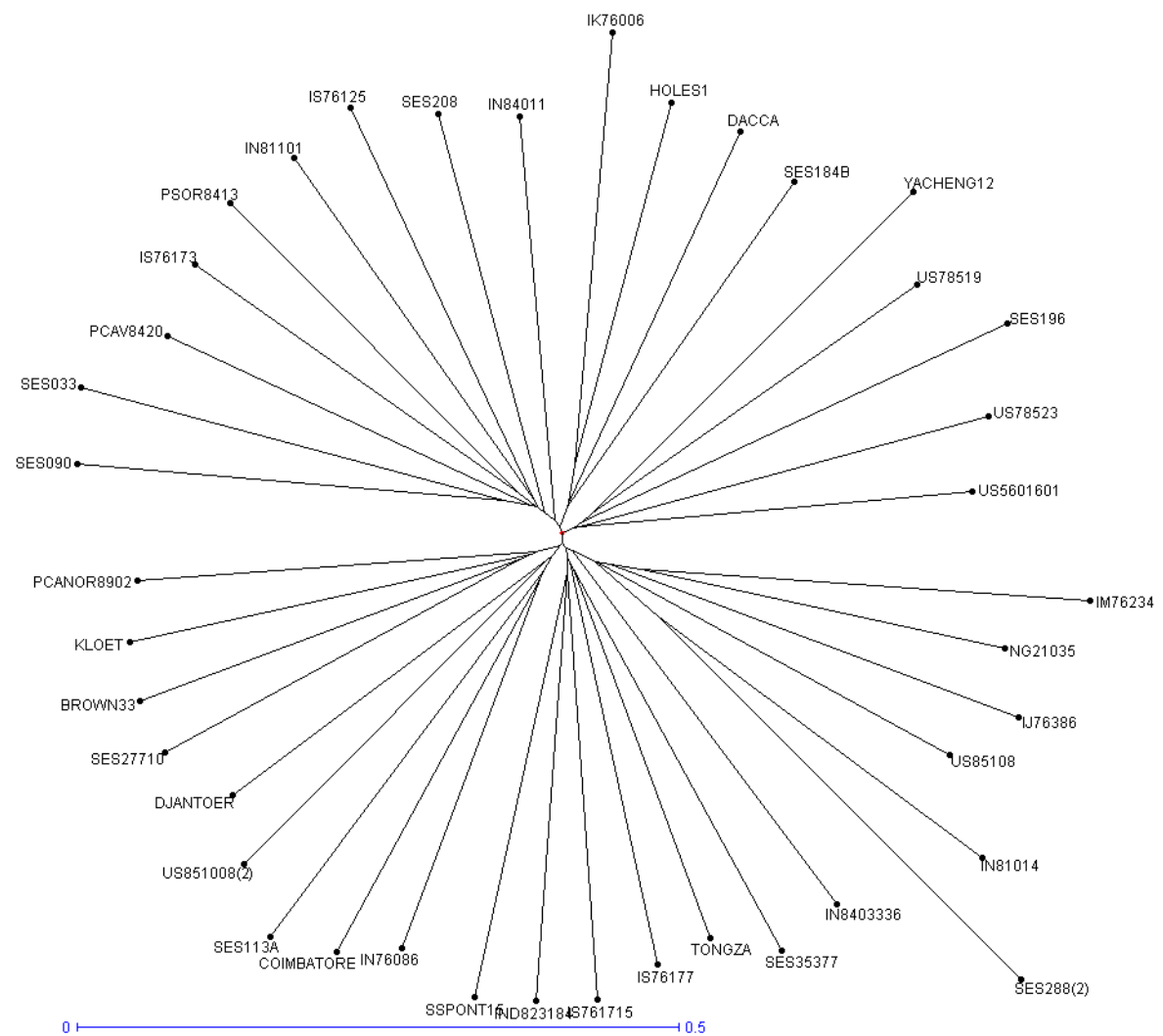


Figura 6: Árvore filogenética contendo os 40 acessos de *S. spontaneum* selecionados para a coleção nuclear.

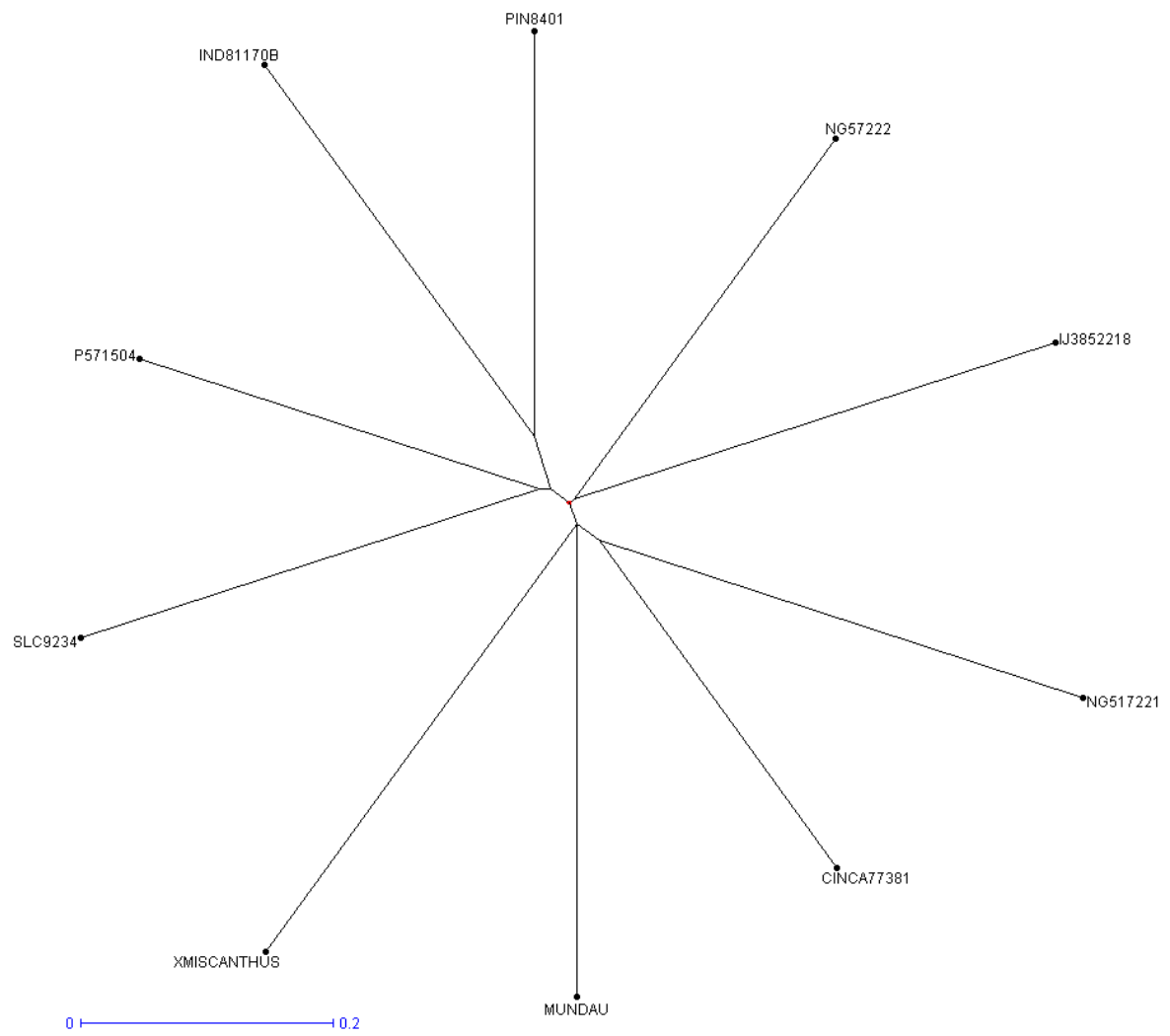


Figura 7: Árvore filogenética contendo os 10 acessos com genealogias desconhecidas, selecionados para a coleção nuclear.

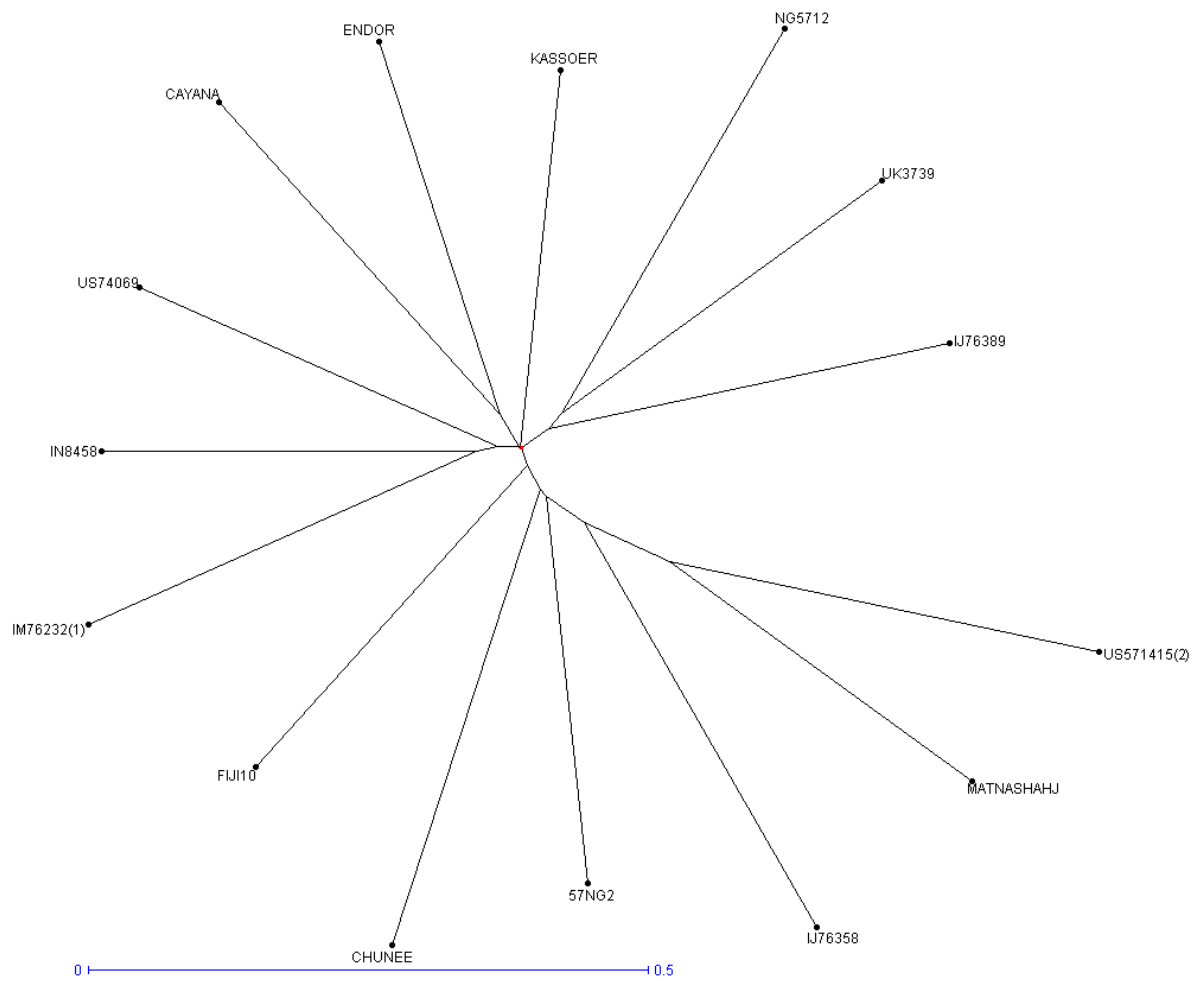


Figura 8: Árvore filogenética contendo todos os 15 acessos de espécies diversas e híbridos exóticos selecionados para a coleção nuclear.

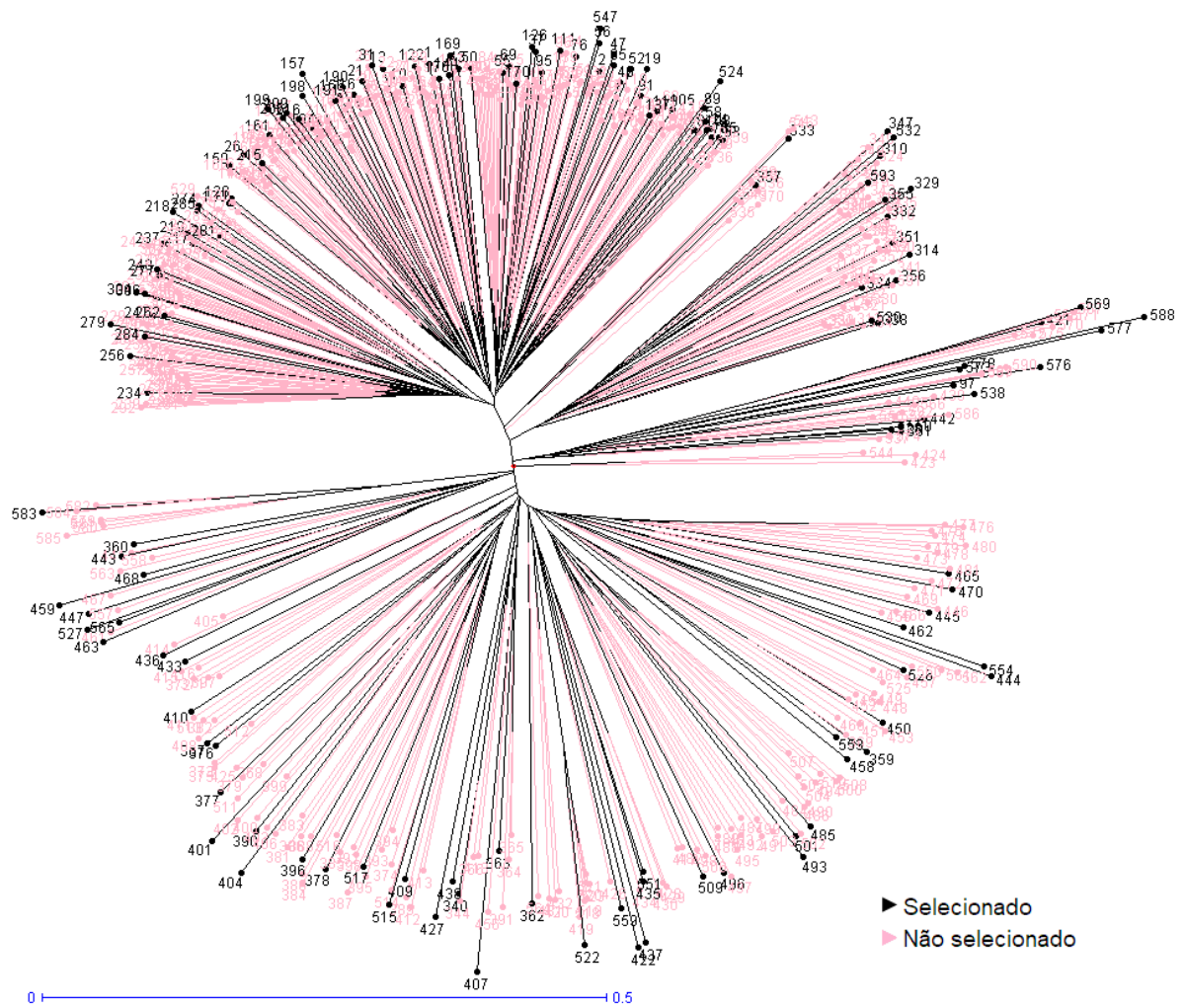


Figura 9: Árvore filogenética contendo os números de identificação dos indivíduos selecionados e não selecionados para a coleção nuclear.

Discussão

Eficiência dos marcadores microssatélites

O conjunto de doze marcadores microssatélites utilizados neste estudo diferenciou os 593 genótipos, inclusive os clones e cultivares brasileiros, que possuem estreita base genética. Os altos valores das estimativas da Razão Multiplex Efetiva (E), Conteúdo de Informação de Polimorfismo (PIC), Diversidade Gênica (D), Índice de Marcador (IM), Poder de Resolução (PR) e o número médio de Perfis Identificados (P. Id.) demonstram a utilidade destes iniciadores para estudos de estruturação genética em populações de *Saccharum* e gêneros relacionados. Estes valores superam os relatados em outros trabalhos que utilizaram SSR em *Saccharum*, uma vez que aqui foram utilizados apenas doze pares de iniciadores capazes de identificar toda a população, em contraste com o grande número de iniciadores SSR necessários para populações deste porte. Entretanto, para futuros trabalhos, é razoável a substituição dos três pares de iniciadores SSR que apresentaram menor desempenho e capacidade de gerar perfis únicos na população total (CV60, CV94 e CV106).

Distribuição da frequência alélica e alelos exclusivos entre grupos

Como esperado, o grupo com a maior ocorrência de alelos foi o de *S. spontaneum*, com 273 (93,49%), considerado pela literatura como o mais diverso grupo ancestral da cana-de-açúcar moderna (LIU et al., 2016). Em comparação com a ocorrência de alelos do grupo dos Clones (65,41%) e Cultivares (56,85%), no contexto da ampliação da base genética destes dois grupos, é possível afirmar que em tese, existe ainda grande potencial de contribuição de recursos oriundos desta espécie e das outras representadas no grupo de híbridos exóticos e espécies diversas (86,64%). Entretanto, é necessário considerar que o melhoramento vem ao longo dos anos direcionando a seleção genética para selecionar genes responsáveis pelo açúcar, oriundos de *S. officinarum*, e responsáveis por conferir resistência a estresses variados, oriundos de outras espécies e gêneros, e que dessa forma, toda essa variabilidade genética diversa entre esses grupos refere-se a traços que não foram selecionados. Há que se considerar também a herdabilidade cromossômica

em hibridações entre *S. officinarum* e outras espécies. No caso de *S. spontaneum*, ela ocorre de forma desbalanceada, sendo que *S. officinarum* transmite duas vezes mais material que esta outra espécie, na proporção $2n + n$ (RAMDOYAL; BADALOO, 2002). Dessa forma, no avanço das gerações de nobilitação, a proporção cromossômica original de *S. spontaneum* nos híbridos é reduzida pela metade, resultando em aproximadamente 80% da constituição genética vinda de *S. officinarum*, 10 a 15% de *S. spontaneum* e 10 a 5% como sendo cromossomos recombinantes entre as duas espécies (D'HONT et al., 1995; NAYAK et al., 2014). Este resultado possui também aplicação prática para auxiliar a escolha de genitores para cana energia, a qual é obtida com retrocruzamentos entre materiais melhorados (híbridos de *S. officinarum*) com acessos de *S. spontaneum*.

Estruturação, dissimilaridades e análise filogenética

A inferência de subgrupos para a população total indicou estrutura genética presente na forma de dois subgrupos (Figura 1), separando na amostra total os Clones e Cultivares dos demais genótipos. A alta significância desta inferência pelo software "Structure", em detrimento do baixo valor apresentado para todos os outros valores de K , revela a baixa dissimilaridade existente entre estes materiais melhorados e o seu distanciamento das espécies ancestrais que os originaram. Na figura 1b, para $K=2$, é possível perceber o caráter intermediário exibido pelo grupo de acessos de *S. officinarum* entre o grupo de Clones e Cultivares e o grupo de acessos de *S. spontaneum*, em concordância com os valores de F_{ST} obtidos entre os grupos predefinidos. Da mesma forma, é possível visualizar na árvore filogenética a forma como estes materiais se agrupam, na qual os Clones e Cultivares constituem grupos coesos e bem delimitados, indicando pouca distância entre seus indivíduos, e a forma como os acessos de *S. spontaneum* são permeados por acessos de todos os outros grupos, desconsiderando as discrepâncias descritas anteriormente nos Resultados. A considerável distância na árvore filogenética entre os acessos de *S. officinarum* e *S. spontaneum* está em acordo com o relatado por Brown et al. (2007) a respeito da origem da primeira espécie, uma especiação originada a partir de hibridações interespecíficas entre *S. spontaneum* e *S. robustum*.

Construção da Coleção Nuclear

Bancos de germoplasma reúnem centenas, talvez milhares de genótipos diferentes, que devem estar sempre prontos para a utilização em programas de melhoramento genético. Ademais, o investimento financeiro e humano necessário para a sua manutenção e gerenciamento é alto. A possibilidade de se possuir uma coleção reduzida de genótipos que apresente diversidade genética igual ou equivalente a presente no grupo original, na forma de uma Coleção Nuclear, se apresenta como uma alternativa com relação aos custos envolvidos. Além disso, com a ausência de redundâncias e duplicidades, a escolha dos materiais visando a máxima exploração dos recursos genéticos pode ser favorecida (PERSEGUINI et al., 2015).

Há de se ressaltar que os dados biológicos utilizados neste estudo compreendem apenas a genealogia entre os materiais, uma vez que os marcadores microssatélites aqui utilizados não apresentam correlações confirmadas com quaisquer fenótipos. Recomenda-se a utilização de outros critérios para a catalogação dos indivíduos, com relação às suas características fenotípicas, no âmbito de se consolidar uma lista mais abrangente e detalhada a respeito das características de interesse dos materiais para futuros cruzamentos e hibridações. Entretanto, a caracterização dos fenótipos será uma tarefa muito mais factível e viável para os 156 genótipos selecionados do que para todos aqueles que foram aqui estudados (NAYAK et al., 2014).

A estratégia adotada neste trabalho para a construção da Coleção Nuclear, isto é, estratificação da coleção em Cultivares e Clones, acessos de *S. officinarum*, acessos de *S. spontaneum*, acessos com genealogias desconhecidas, e híbridos exóticos e acessos com genealogias diversas, almejou manter as características de cada um dos estratos, de forma a selecionar os indivíduos mais diversos presentes dentro de cada grupo. O reduzido número de materiais selecionado (aproximadamente 30%) forneceu uma lista que pode auxiliar a escolha dos materiais para futuros programas de melhoramento para cana-de-açúcar convencional e para cana energia.

Referências Bibliográficas

- AITKEN, K. et al. Characterization of intergeneric hybrids of *Erianthus rockii* and *Saccharum* using molecular markers. **Genetic Resources and Crop Evolution**, v. 54, n. 7, p. 1395–1405, 2007.
- AITKEN, K. S.; JACKSON, P. A.; MCINTYRE, C. L. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. **TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik**, v. 110, n. 5, p. 789–801, mar. 2005.
- ALJANABI, S.; FORGET, L.; DOOKUN, A. An improved and rapid protocol for the isolation of polysaccharide-and polyphenol-free sugarcane DNA. **Plant Molecular Biology Reporter**, v. 17, p. 1–8, 1999.
- ARCENEUX, G. Cultivated Sugarcane of the World and their Botanical Derivation. **Proceedings ISSCT**, v. 12, p. 844–854, 1967.
- BERDING, N.; HOGARTH, M.; COX, M. Plant Improvement of Sugarcane. In: JAMES, G. (Ed.). **World Agriculture Series: Sugarcane**. 2. ed. Oxford, UK: Blackwell Publishing, 2004. p. 20–53.
- BERDING, N.; ROACH, B. T. Germplasm collection, maintenance and use. In: HEINZ, D. J. (Ed.). **Sugarcane Improvement through Breeding**. 1. ed. Amsterdam: Elsevier, 1998. v. 11, p. 143–210.
- BREMER, G. On the somatic chromosome numbers of sugarcane forms of endogenous cane. **Proceedings ISSCT**, v. 4, n. 20, p. 1–3, 1932.
- BROWN, J. S. et al. Analysis of clonal germplasm from five *Saccharum* species: *S. barberi*, *S. robustum*, *S. officinarum*, *S. sinense* and *S. spontaneum*. A study of inter- and intra-species relationships using microsatellite markers. **Genetic Resources and Crop Evolution**, v. 54, n. 3, p. 627–648, 2007.
- BURNER, D. M.; PAN, Y.; WEBSTER, R. D. Genetic diversity of North American and Old World *Saccharum* assessed by RAPD analysis. **Genetic Resources and Crop Evolution**, v. 44, p. 235–240, 1997.
- CAI, Q. et al. Verification of the introgression of *Erianthus arundinaceus* germplasm into sugarcane.pdf. **Plant Breeding**, v. 124, p. 322–28, 2005.
- CHEN, P. et al. SSR marker-based analysis of genetic relatedness among sugarcane cultivars (*Saccharum* spp. hybrids) from breeding programs in China and other countries. **Sugar Tech**, v. 11, n. 4, p. 347–354, 2009.
- CORDEIRO, G. M. et al. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to *Erianthus* and sorghum. **Plant science: an international journal of experimental plant biology**, v. 160, n. 6, p. 1115–1123, maio 2001.

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

COSTA, M. L. M. DA et al. Assessment of Genetic Diversity in Contrasting Sugarcane Varieties Using Inter-Simple Sequence Repeat (ISSR) Markers. **American Journal of Plant Sciences**, v. 2, p. 425–432, 2011.

CROFT, B. J. et al. **Sugarcane Germplasm Conservation and Exchange**. 1. ed. Canberra, Australia: Australian Centre for International Agricultural Research, 1996.

CROFT, B.; THOMPSON, N.; MAGAREY, R. Introduction to Sugarcane Quarantine: Instructions for staff of research organizations involved in sugarcane research. **BSES Limited Publication**, p. 35, out. 2011.

D'HONT, A. et al. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. **Genome**, v. 41, n. 2, p. 221–225, 1998.

D'HONT, A. et al. Identification and characterization of sugarcane intergeneric hybrids, *Saccharum officinarum* x *Erianthus arundinaceus*, with molecular markers and DNA in situ hybridization. **TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik**, v. 91, n. 2, p. 320–6, jul. 1995.

DA SILVA, J. A. The Importance of the Wild Cane *Saccharum spontaneum* for Bioenergy. **Sugar Tech**, p. 1–39, 2017.

DANIELS, J.; ROACH, B. T. Taxonomy and Evolution. In: HEINZ, D. J. (Ed.). **Sugarcane Improvement through Breeding**. 1. ed. Amsterdam: Elsevier, 1998. v. 11, p. 7–84.

DOS SANTOS, J. M. et al. Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. **Industrial Crops and Products**, v. 40, n. 1, p. 145–150, 2012.

EARL, D.; VON-HOLDT, B. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, v. 4, n. 2, p. 359–361, 2010.

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

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Molecular Ecology**, v. 14, p. 2611–20, 2005.

EXCOFFIER, L.; LISCHER, H. Arlequin suite ver. 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, p. 564–567, 2010.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3. ed. Brasília: EMBRAPA, 1998.

FRANKHAM, R.; BALLOU, J. D.; BRISCOE, D. A. **Introduction to Conservation Genetics**. 1. ed. Cambridge, UK: Cambridge University Press, 2002.

GASHAW, E. T.; MEKBIB, F.; AYANA, A. Genetic Diversity among Sugarcane Genotypes Based on Qualitative Traits. **Advances in Agriculture**, v. 2016, p. 8. Article ID 8909506. 2016.

GRIVET, L. et al. A review of recent molecular genetics evidence for sugarcane evolution and domestication. **Ethnobotany Research & Applications**, v. 2, p. 9–17, 2004.

HAMEED, U. et al. Use of simple sequence repeat markers for DNA fingerprinting and diversity analysis of sugarcane (*Saccharum* spp) cultivars resistant and susceptible to red rot. **Genetics and molecular research: GMR**, v. 11, n. 2, p. 1195–204, jan. 2012.

HARVEY, M.; BOTHA, F. C. Use of PCR-based methodologies for the determination of DNA diversity between *Saccharum* varieties. **Euphytica**, v. 89, n. 2, p. 257–265, 1996.

HOANG, N. V. et al. Potential for Genetic Improvement of Sugarcane as a Source of Biomass for Biofuels. **Frontiers in Bioengineering and Biotechnology**, v. 3, November, p. 1–15, 2015.

HOARAU, J. Y. et al. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. **Theoretical and Applied Genetics**, v. 103, n. 1, p. 84–97, 2001.

JAMES, G. **World Agriculture Series: Sugarcane**. 2. ed. Oxford, UK: Blackwell Publishing, 2004.

JANNOO, N. et al. Molecular investigation of the genetic base of sugarcane cultivars. **Theoretical and Applied Genetics**, v. 99, n. 1–2, p. 171–184, 1999.

JARNE, P.; LAGODA, P. J. Microsatellites, from molecules to populations and back. **Trends in ecology & evolution**, v. 11, n. 10, p. 424–9, out. 1996.

JESWIET, J. Immunity and cross-fertilisation in the genus *Saccharum*. **Recueil des Travaux Botaniques Neerlandais**, v. 25A, p. 185–202, 1928.

JORDAN, D. R. et al. Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. **Euphytica**, v. 102, n. 1, p. 1–7, 1998.

KALIA, R. K. et al. Microsatellite markers: an overview of the recent progress in plants. **Euphytica**, v. 177, n. 3, p. 309–334, 9 nov. 2010.

KEMP, S. **Pic Calculator**, 2002.

KOPELMAN, N. et al. **CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K**, 2012.

LIU, X. et al. Phylogenetic Analysis of Different Ploidy *Saccharum spontaneum* Based on rDNA-ITS Sequences. **Plos One**, v. 11, n. 3. <https://doi.org/10.1371/journal.pone.0151524>. 2016.

MACCHERONI, W. et al. Development of a dependable microsatellite-based fingerprinting system for sugarcane. **Sugar Cane Intl**, v. 27, p. 8–13, 2009.

MACHADO, F. B. P. **Brasil, a doce terra - História do Setor**. Disponível em: <https://www.agencia.cnptia.embrapa.br/Repositorio/historia_da_cana_000fhc62u4b02wyiv80efhb2attuk4ec.pdf>. Acesso em: 31 jan. 2017.

MAPA. **Descritores Mínimos de Cana de Açúcar**. 1. ed. Brasília, DF, Brasil: Diário Oficial da União, 1998.

MAPA. **Proteção de Cultivares no Brasil**. 1. ed. Brasília, DF, Brasil: Biblioteca Nacional de Agricultura, 2011.

MARCONI, T. G. et al. Functional markers for gene mapping and genetic diversity studies in sugarcane. **BMC Research Notes**, v. 4, n. 1, p. 264, jan. 2011.

MUKHERJEE, S. K. Origin and Distribution of *Saccharum*. **Botanical Gazette**, v. 119, n. 1, p. 55–61, 1957.

MUTSAERS, H. **Peasants, Farmers and Scientists**. 1. ed. Dordrecht: Springer Netherlands, 2007.

NAIR, N. V. Sugarcane Genetic Resources - Status, Potential and Role in Sugarcane Improvement. v. 2, p. 1–8, 2012.

NAYAK, S. N. et al. Promoting utilization of *Saccharum* spp. genetic resources through genetic diversity analysis and core collection construction. **PloS one**, v. 9, n. 10. <https://doi.org/10.1371/journal.pone.0110856>. 2014.

NEI, M. Analysis of Gene Diversity in Subdivided Populations. **Proc Nat Acad Sci**, v. 70, n. 12, p. 3321-3323, dez. 1973.

OLIVEIRA, K. M. et al. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding**, v. 20, n. 3, p. 189–208, 6 abr. 2007.

OLIVEIRA, K. M. et al. Characterization of new polymorphic functional markers for sugarcane. v. 209, p. 191–209, 2009.

PALHARES, A. C. et al. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. **BMC genetics**, v. 13, p. 51, jan. 2012.

PAN, Y. et al. An assessment of the genetic diversity within a collection of *Saccharum spontaneum* L. with RAPD-PCR. **Genetic Resources and Crop Evolution**, v. 51, p. 895–903, 2004.

PAN, Y. Highly Polymorphic Microsatellite DNA Markers for Sugarcane Germplasm Evaluation and Variety Identity Testing. **Sugar Tech**, v. 8, n. 4, p. 246–256, 2006.

PAN, Y.-B. Databasing Molecular Identities of Sugarcane (*Saccharum* spp.) Clones Constructed with Microsatellite (SSR) DNA Markers. **American Journal of Plant Sciences**, v. 1, n. 2, p. 87–94, 2010.

PAN, Y. B. et al. Analysis of primer-derived, nonspecific amplification products in RAPD-PCR. **BioTechniques**, v. 22, n. 6, p. 1071–4, 1076–7, jun. 1997.

PAN, Y. B.; BURNER, D. M.; LEGENDRE, B. L. An assessment of the phylogenetic relationship among sugarcane and related taxa based on the nucleotide sequence of 5S rRNA intergenic spacers. **Genetica**, v. 108, n. 3, p. 285–95, 2000.

PAN, Y.; SCHEFFLER, B.; RICHARD, J. High-throughput genotyping of commercial sugarcane clones with microsatellite (SSR) DNA markers. **Sugar Technology**, v. 9, p. 176–181, 2007.

PANJE, R. R.; BABU, C. N. Studies in *Saccharum spontaneum* Distribution and Geographical Association Chromosome Numbers. **Cytologia**, v. 25, p. 152–172, 1960.

PARIDA, S. K. et al. Functionally relevant microsatellites in sugarcane unigenes. **BMC plant biology**, v. 10, n. 1, p. 251, jan. 2010.

PERERA, M. F. et al. Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits. **Euphytica**, v. 185, n. 3, p. 491–510, 2012.

PERRIER, X.; JACQUEMOUD-COLLET, J. **DARwin software**, 2006.

PERSEGUINI, J. M. K. C.; SILVA, G. M. B.; ROSA, J. R. B. F.; GAZAFFI, R.; MARÇAL, J. F.; CARBONELL, S. A. M.; CHIORATO, A. F.; ZUCCHI, M. I.; GARCIA, A. A. F.; BENCHIMOL-REIS, L. L. Developing a common bean core collection suitable for association mapping studies. **Gen Mol Biol**, v. 38, n. 1, p. 67-78, 2015.

PINTO, L. R. et al. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome / National Research Council Canada = Genome / Conseil national de recherches Canada**, v. 47, n. 5, p. 795–804, 2004.

PINTO, L. R. et al. Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. **PLANT BREEDING**, v. 125, p. 378–384, 2006.

PREVOST, A.; WILKINSON, M. J. A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. **Theoretical and Applied Genetics**, v. 98, n. 1, p. 107–112, 1999.

PRITCHARD, J.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, p. 945–959, 2000.

RAMALHO, M. A. P. et al. **Genética na Agropecuária**. 5. ed. Lavras, MG: Editora UFLA, 2012.

RAMDOYAL, K.; BADALOO, G. Prebreeding in Sugarcane with an Emphasis on the Programme of the Mauritius Sugar Industry Research Institute. In: ENGELS, J. M. et al. (Eds.). **Managing Plant Genetic Diversity**. 1. ed. Oxford, UK: CABI Publishing, 2002. p. 307–321.

SCHOEN, D. J.; BROWN, A. H. D. Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. **Proc. Natl. Acad. Sci. USA**. v. 90, n. 22, p. 10623–7, 1993.

SHARMA, R. K. et al. Evaluation of rice and sugarcane SSR markers for phylogenetic and genetic diversity analyses in bamboo. **Genome**. v. 51, n. 2, p. 91–103, 2008. doi:10.1139/g07-101

SILVA, D. C. D. et al. DNA fingerprinting based on simple sequence repeat (SSR) markers in sugarcane clones from the breeding program RIDESA. **African Journal of Biotechnology**, v. 11, n. 21, p. 4722–4728, 13 mar. 2012.

SILVA, J. A. et al. Sugarcane Variety Identification through DNA Fingerprinting with Microsatellites Markers. **Subtropical Plant Science**, v. 60, p. 1–7, 2008.

SILVA, J. A. G. Preliminary analysis of microsatellite markers derived from sugarcane expressed sequence tags (ESTs). **Genetics and Molecular Biology**, v. 24, n. 1, p. 155–159, 2001.

TABASUM, S. et al. DNA profiling of sugarcane genotypes using randomly amplified polymorphic DNA. **Genetics and molecular research: GMR**, v. 9, n. 1, p. 471–83, jan. 2010.

The Plant List. Disponível em:

<<http://www.theplantlist.org/tpl1.1/search?q=Saccharum>>. Acesso em: 22 fev. 2017.

THOKOANE, L. N. Progress towards a fingertyping database for sugarcane varieties. **Proc S Afr Sug Technol Ass**, v. 73, p. 164–6, 1999.

VARSHNEY, R. K. et al. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. **Plant Science**, v. 173, n. 6, p. 638–649, 2007. <https://doi.org/10.1016/j.plantsci.2007.08.010>

VETTORE, A. L. et al. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome research**, v. 13, n. 12, p. 2725–35, dez. 2003.

WEIR, B.; COCKERHAM, C. Estimating F-Statistics for the Analysis of Population Structure. **Evolution**, v. 38, n. 6, p. 1358–70, 1984.

ZABEAU, M.; VOS, P. **Selective Restriction Fragment Amplification: Fingerprinting**, 2000.

CAPÍTULO 3 – Investigando a Estrutura Genética das Cultivares Comerciais Brasileiras de Cana-de-Açúcar Através de Marcadores Microsatélites

Resumo

A indústria brasileira de cana-de-açúcar tem uma participação enorme e crescente na oferta mundial de açúcar e etanol. A investigação da estrutura genética de cultivares comerciais atualmente em uso, e sua comparação com o “pool” genético de suas principais espécies ancestrais (ênfase para *S. officinarum* e *S. spontaneum*) pode contribuir para melhor gerenciamento de bancos de germoplasma e planejamento de cruzamentos futuros para os programas de melhoramento da cana, colaborando também com o estabelecimento de estratégias de conservação. No presente estudo, a estrutura genética de um painel de 137 cultivares comerciais brasileiras atualmente em uso por produtores de cana-de-açúcar (que por sua vez representa o “pool” genético real dos programas de melhoramento de cana-de-açúcar brasileiros) foi avaliada utilizando 12 marcadores microsatélites e contrastada com um grupo de germoplasma básico composto principalmente por acessos de *S. officinarum* e *S. spontaneum*. Como resultado da amplificação, foram observados 285 marcadores polimórficos (alelos). A avaliação da distribuição alélica revelou que 98,6% da diversidade das cultivares é similar à dos acessos do germoplasma básico. Além disso, cerca de 42% dos alelos de germoplasma básico estavam ausentes nas cultivares, sugerindo um importante conjunto de recursos genéticos não explorados. Cerca de 28% desses alelos ausentes são originais de *S. spontaneum*, 3% de *S. officinarum* e 14% de espécies diversas como *S. robustum*, *S. barberi* e *Miscanthus* sp. Encontramos menor variação significativa ($P < 0,001$) ao comparar cultivares e germoplasma básico ($\Phi_{ST} = 0,179$) do que ao comparar cultivares unicamente com acessos de *S. officinarum* ($\Phi_{ST} = 0,211$) e *S. spontaneum* ($\Phi_{ST} = 0,216$), mostrando que as cultivares comerciais são geneticamente mais distantes de *S. spontaneum*.

Introdução

A cultura de cana-de-açúcar tem importância econômica mundial principalmente para a produção de açúcar e etanol, sendo o Brasil o segundo maior produtor, com safra de mais de 658 milhões de toneladas em 2015/16 [1]. Além disso, a cogeração de energia elétrica atende atualmente 4% da demanda brasileira de energia elétrica [2]. Essas posições de destaque são em parte atribuídas a práticas de manejo agrônomo adequadas, e também ao alto desempenho das cultivares comerciais brasileiras de cana-de-açúcar, desenvolvidas pelos três principais programas de melhoramento de cana-de-açúcar do país (RIDESA: Rede Interuniversitária para o Desenvolvimento do setor Sucroenergético; CTC: Centro de Tecnologia Canavieira; IAC: Instituto Agrônomo de Campinas).

A cana-de-açúcar pertence ao gênero *Saccharum*, que está relacionado a outros gêneros como *Erianthus* sect. *Ripidium*, *Miscanthus* sect. *Diandra* Keng Narenga Bor, e *Sclerostachya* Hack A. Camus, os quais constituem o Complexo *Saccharum* [3].

O maior avanço para a cultura da cana-de-açúcar veio com a descoberta de que o gênero *Saccharum* produz descendência viável e ocasionalmente com fenótipos superiores a partir de cruzamentos interespecíficos [4]. Híbridações interespecíficas artificiais começaram no final do século XIX no continente Asiático, conduzidas por melhoristas holandeses [5]. Basicamente, duas espécies foram as mais utilizadas nesta abordagem: *S. officinarum*, decaplóide ($2n=80$, $x=10$) com talos espessos e suculentos, ricos em açúcar e conhecida como cana nobre, mas altamente suscetível a vários estresses bióticos e abióticos, e *S. spontaneum*, uma octaplóide ($2n=40-128$, $x=8$) sem conteúdo de açúcar significativo, mas altamente resistente, vigoroso [6,7] e eficiente na fixação de carbono devido à sua elevada taxa fotossintética [8]. Os primeiros híbridos interespecíficos foram repetidamente cruzados com acessos de *S. officinarum* para recuperar o teor de sacarose, num processo conhecido como “nobilização” [5]. Portanto, aproximadamente 80% do genoma das cultivares atuais de cana-de-açúcar vieram desta espécie ancestral, 10-15% de *S. spontaneum* ou outras espécies selvagens, e os restantes 5-10% representam cromossomos recombinantes [9]. Outras espécies de *Saccharum*, como *S. robustum*, *S. barberi* e *S. sinense*, embora menos utilizadas como espécies

ancestrais nas primeiras hibridizações interespecíficas artificiais, também contribuíram para o genoma das cultivares modernas de cana-de-açúcar [10].

Vale ressaltar que as cultivares de cana-de-açúcar das primeiras hibridações interespecífica impactaram positivamente a indústria de cana-de-açúcar do século XX, especialmente a cultivar “POJ2878”, conhecida como a "cana maravilha de Java", presente no pedigree da maioria das cultivares de cana-de-açúcar atuais em todo o mundo [11]. Apesar da natureza poliplóide da cana-de-açúcar e da grande variabilidade genética disponível no germoplasma de *Saccharum*, poucos acessos contribuíram geneticamente para os primeiros híbridos interespecíficos [12,13] levando à estreita base genética das cultivares atuais.

De acordo com a União Mundial para a Conservação (IUCN), a manutenção da biodiversidade depende diretamente da conservação da diversidade genética [14,15]. Por exemplo, a seleção exaustiva de certos traços de interesse para o melhoramento, durante várias gerações de hibridação, inevitavelmente tende a excluir outros traços indesejados. Certas regiões gênicas ou mesmo alelos completos e subconjuntos de genes podem ser perdidos e desapareceram completamente com o tempo. A perda de variabilidade e o estreitamento de base genética implicam em perda de biodiversidade, e intensificam o risco de susceptibilidade a uma doença ou praga inesperada [16].

A fim de preservar a variabilidade genética e a biodiversidade, bancos de germoplasma ativos são mantidos por programas de melhoramento como estratégia de conservação dos recursos genéticos disponíveis a serem utilizados em futuros cruzamentos, especialmente em programas de introgressão genética [17]. Existem vários bancos de germoplasma de cana-de-açúcar no mundo, sendo a World Collection of Sugarcane and Related Grasses (WCSRG) em Miami, FL, USA e o Centro de Pesquisa de Kannur, em Coimbatore, Índia, dois dos mais relevantes. No Brasil, um dos maiores bancos de germoplasma do mundo é mantido pelo CTC (Centro de Tecnologia Canavieira) em Urucuca, BA, Brasil [18]. Estes bancos englobam uma enorme quantidade de acessos de espécies de *Saccharum* e gêneros relacionados, incluindo os coletados nas primeiras expedições da cana-de-açúcar, que atuam em colaboração internacional [19].

No entanto, a variabilidade genética presente na maioria dos bancos de germoplasma é subexplorada. Os bancos de germoplasma de cana-de-açúcar têm milhares de acessos, o que traz dificuldades de logística e no planejamento de cruzamentos futuros. Organização e classificação são fatores-chave relacionados com a funcionalidade dos bancos de germoplasma. Embora os descritores morfológicos tenham um papel importante na classificação e organização dos acessos, erros podem ocorrer quando os fenótipos são afetados pelas condições ambientais. A obtenção do perfil molecular do acesso (*fingerprinting*), que é uma forma mais confiável de identificar um dado indivíduo [20,21], é um complemento à sua descrição morfológica. Para isto, vários marcadores moleculares têm sido utilizados para investigar e medir a diversidade genética no nível molecular de acessos de cana-de-açúcar e também para rastrear a relação genética entre as espécies de *Saccharum* [22,23].

Microssatélites ou Repetições de Sequência Única (SSR) estão entre os marcadores mais adequados e preferidos para estudos de estrutura genética e de conservação [24]. SSRs são trechos de DNA compostos por repetições de motivos de 1 a 6 pares de bases, que são flanqueados por sequências conservadas que permitem a amplificação através de PCR por pares de iniciadores específicos [25]. Além disso, a SSR é uma técnica altamente polimórfica, reproduzível, de caráter codominante e transferível entre espécies estreitamente relacionadas e gêneros [26]. Devido à alta ploidia da cana-de-açúcar, os marcadores microssatélites são considerados marcadores dominantes uma vez que a presença de uma dada banda amplificada não representa a sua dosagem entre os cromossomos, tornando assim impossível identificar se um locus é homo ou heterozigoto [27]. Dessa forma, os SSR podem gerar “*fingerprints*” genômicos, como impressões digitais únicas para identificação varietal [28]. Para cana-de-açúcar, uma grande quantidade de pares de primers SSR está disponível na literatura, desenvolvidos a partir de bibliotecas genômicas enriquecidas, como as do Consórcio Internacional de Microssatélites de Cana-de-açúcar (ISMC) [29] ou por etiquetas de sequências expressas (SUCEST – Sugarcane Expressed Sequence Tag Database), que são intensivamente aplicados para mapeamento genético, estudos de diversidade e *fingerprinting* [30–32].

A estreita base genética da cana-de-açúcar e suas implicações para garantir produção de açúcar e ganhos de produtividade são preocupações entre os melhoristas de cana-de-açúcar em todo o mundo [33,34]. No entanto, ainda existe uma considerável variabilidade genética entre as cultivares de cana-de-açúcar devido à sua heterozigosidade e natureza poliplóide [35]. Portanto, a investigação da estrutura genética de cultivares comerciais atualmente em uso e sua comparação com o “pool” genético de suas principais espécies ancestrais (*S. officinarum* e *S. spontaneum*) permite que melhoristas e curadores gerenciem melhor seus bancos de germoplasma e conduzam as hibridizações a serem realizadas por programas de melhoramento, permitindo a obtenção de novas cultivares. A caracterização de acessos de germoplasma básico de cana-de-açúcar tem merecido atenção por promover sua utilização em programas de melhoramento da cultura [10] principalmente por aqueles focados no desenvolvimento de cultivares nas quais *S. spontaneum* tem sido utilizado como parental recorrente em cruzamentos com cultivares comerciais para atender a demanda por produção de biomassa [36].

Devido à importância do setor sucroalcooleiro para a economia brasileira e à participação ativa dos programas de melhoramento de cana-de-açúcar no desenvolvimento e liberação de novas cultivares, quase todos os anos, há um grande número destas em uso. No entanto, a avaliação de sua estrutura genética e, portanto, do impacto da estreita base genética sobre a magnitude do “pool” genético da cana-de-açúcar atual brasileira ainda não foi completamente investigado.

No presente estudo, a estrutura genética de um conjunto de cultivares comerciais brasileiras atualmente em uso por produtores de cana-de-açúcar (que por sua vez representa o pool genético real dos programas de melhoramento de cana-de-açúcar brasileiros) foi avaliada por marcadores microssatélites e contrastada com um grupo de germoplasma básico composto principalmente por acessos de *S. officinarum* e *S. spontaneum*.

Materiais e Métodos

Material vegetal

Um total de 137 genótipos, abrangendo 81 cultivares brasileiras de cana-de-açúcar comercial e 56 acessos de germoplasma básico (19 *Saccharum officinarum*, 26 *S. spontaneum* × 5 híbridos exóticos e 6 acessos representando outras espécies como *S. barberi*, *S. robustum* e *Miscanthus*) foram coletados da Coleção de Germoplasma de Cana-de-Açúcar do IAC mantida no Centro de Cana, em Ribeirão Preto, SP. A maioria dos acessos de *S. spontaneum* e outras espécies foi importada da Coleção Mundial de Cana-de-Açúcar e Gramíneas Relacionadas do USDA (WCSR, Miami, FL, EUA). Os 56 acessos que compõem o grupo de germoplasma básico, foram previamente selecionados dentre 600 outros, como sendo os mais divergentes, não redundantes e historicamente relevantes do banco de dados do IAC, baseado em SSR. Alguns dos indivíduos historicamente relevantes foram escolhidos de acordo com sua contribuição no processo de nobilitação. Clones como “KASSOER”, “COIMBATORE” e “CHUNNEE” foram intensamente utilizados nas primeiras hibridizações interespecíficas e estão presentes no pedigree de várias cultivares comerciais em todo o mundo, bem como o clone POJ2878, que foram incluídos no trabalho devido à sua notável contribuição no desenvolvimento da maioria das cultivares atuais em todo o mundo [11, 37]. A lista detalhada de todos os materiais é fornecida na Tabela S1.

Extração de DNA e reações de PCR

DNA genômico total foi extraído de folhas jovens ou de tecido de rolo foliar de acordo com o método CTAB [38]. As reações de PCR foram conduzidas num volume final de 16µl, contendo 30ng de DNA molde, 0,2µM de cada iniciador “forward” e “reverse” (marcados com corantes infravermelhos IR700 ou IR800), 100µM de cada dNTP, 2,0mM de MgCl₂, 10mM de Tris-HCl, 50mM de KCl e 0,5 unidade de Taq DNA polimerase (Invitrogen, São Paulo, Brasil). As reações foram amplificadas em um termociclador MyCycler® Thermal Cycler (BIO-RAD). Para os pares de iniciadores CV (*CanaVialis*): CV29, CV37, CV38, CV60, CV79, CV94 e CV106 [28,30], as condições de termociclagem foram realizadas com adaptações adequadas [39]. Para os pares de primers SC (*SugarCane*): SCB213, SCB312,

SCB381, SCB423 e SCC436 [22,40] as condições de termociclagem foram as seguintes: desnaturação a 94,0°C durante 3 minutos, seguida de 30 ciclos de anelamento (94,0°C durante 1 minuto, temperatura de anelamento específica para cada iniciador durante 1 minuto, extensão a 72,0°C durante 1 minuto e um ciclo de alongamento final a 72,0°C durante 5 minutos). Informações sobre os pares de iniciadores microssatélites estão descritos na Tabela 1. Os produtos de amplificação foram separados por eletroforese em gel de poliacrilamida a 6% em um analisador de DNA NEN4300 (LI-COR Biosciences®). Os atributos dos marcadores SSR como ubiquidade, codominância e alta heterozigosidade [41], aliada à sua habilidade de identificar com sucesso genótipos de cana-de-açúcar [42] e também o baixo custo da técnica (em comparação com SNPs, por exemplo) foram considerados na sua escolha.

Tabela 1: Relação de pares de iniciadores SSR com respectivas temperaturas de anelamento, motivos de repetição, sequências “forward” e “reverse” (5' a 3'), e gama de tamanhos de produtos amplificados (em pares de bases). ^{a, b, c} De acordo com Maccheroni et al. (2009) [28], Canavialis (não publicado) e Marconi et al. (2011)[22], respectivamente.

SSR	Temp. de anelamento	Motivo de repetição	Sequência Forward Sequência Reverse	Gama de tamanho
CV29 ^a	60,0°C	(ATCT) ₁₄	TCGCGTCCACCAATGTAACC GCGTGCATCGCTTGTGTCTT	82-170
CV37 ^a	60,0°C	(TTTC) ₁₅	GGATGGACGACGTGTCTTGG ATAAAGTGGCCGCTTGGATTGA	106-177
CV38 ^a	60,0°C	(CTTTT) ₁₈	GAAGCAGGGGCTCAAGTTG GTCAAACAGGCGATCTGGCTC	96-210
CV60 ^b	60,0°C	(CTCTCC) ₅	AATCTGCACCCTGCCCTCTC CAGCTGGAGCATGGATGGAG	158-194
CV79 ^b	60,0°C	(CTATAT) ₁₁ (TATAGA) ₆	GGCACTGCTGGTGGTTGATTG TCCCACATCAAGAGGCAGCTA	131-231
CV94 ^b	60,0°C	(AAAAAG) ₅ (CGT) ₅	GGCAGGCCAAGATGAATGAAG AGCACAGCGGAGGGTACGG	182-234
CV106 ^b	60,0°C	(GGC) ₈	AAACAGAGCATACTCGAGGCC ACGTTGCTGACGAGGTTTCC	78-162
SCB213 ^c	62,0°C	(TCC) ₆	AGCCGTCAGGGGTCAGG ATTTCGATGGAGCCTGAGTGAG	247-337
SCB312 ^c	64,7°C	(TCC) ₇	AGTCCGTCGCCGTAATCATCTTG GCACCTCCTCCTTTCCTTCTTATT	190-278
SCB381 ^c	60,0°C	(TAC) ₈	TGGAGCTCCGTCTTCTTGTT GCTAGCCCGTACATTGGGTA	190-257
SCB436 ^c	60,0°C	(GAG) ₅	AGTACGCCCTGAGTCCCTGACG AGGTGCAAGGCTGATAGAA	158-274
SCC423	----- DISPONÍVEL SOB CONSULTA -----			

Análises de dados

As imagens de gel foram manualmente pontuadas e revisadas com o software Saga GT TM (LI-COR Biosciences®). Este software identifica o tamanho do

alelo (fragmento) para construir a matriz de genotipagem. Para marcadores de DNA codominantes como os SSR na cana-de-açúcar, dada a alta complexidade de seu genoma, a genotipagem é realizada considerando as marcas amplificadas como dominantes, de acordo com a ploidia elevada e a incapacidade de detectar homo ou heterozigose para cada loci. Assim, as bandas são marcadas como presença de alelo (1) ou ausência (0). Utilizou-se o padrão de peso molecular 50-350pb (LI-COR Biosciences®) para calibrar os tamanhos dos alelos amplificados em cada corrida.

Análises de eficiência dos iniciadores SSR

Conteúdo de Informação de Polimorfismo [43] foi calculado segundo a expressão $PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^n \sum_{j=i+1}^n 2p_i^2 p_j^2$ (na qual p é a frequência do $j^{ésimo}$ padrão para o marcador i e n é o número de alelos) usando o programa *PIC Calculator* [44] também usado para calcular a Diversidade Genética (D) segundo a expressão $D = 1 - \sum_{i=1}^n P_{ij}^2$, na qual P_{ij} é a frequência do $j^{ésimo}$ padrão para o marcador i , somado dentre n padrões[45].

A Razão Multiplex (n) foi estimada como o número medio de alelos amplificados por genótipo [46]. A Razão Multiplex Efetiva (E) foi estimada pela expressão: $E = n\beta$ na qual n é a Razão Multiplex e β a fração de loci polimórficos, estimada por $\beta = \frac{n_p}{(n_p + n_{np})}$, onde n_p é a soma dos loci polimórficos e n_{np} a soma dos não polimórficos. O Índice de Marcador (IM) foi obtido pela expressão $IM = PIC \times E$.

O Poder de Resolução (PR) para cada par de iniciadores SSR foi calculado a partir da matriz de genotipagem usando a expressão $PR = \sum Ib$, na qual Ib representa a informação alélica, sendo $Ib = 1 - (2 \times |0.5 - M|)$, onde M é a proporção do total de genótipos que contenham dado alelo [47].

Para verificar se o número de marcadores utilizados para estimar as dissimilaridades genéticas entre acessos foi adequado em termos de precisão, a técnica de reamostragem (“bootstrap”) [48] foi aplicada de forma similar a Tivang et al. (1994) [49], Hållden et al. (1994) [50] e outros. A matriz binária dos genótipos foi reamostrada considerando diferentes números de marcadores (3, 6, 9, ..., 285 marcadores). Para cada um dos 95 tamanhos de amostra considerados, foram realizados 1000 *bootstraps* e, para cada amostra, obteve-se uma matriz de

dissimilaridade, permitindo calcular o coeficiente médio de variação (CV%) das estimativas de similaridade (Jaccard) de cada tamanho de amostra. Uma função exponencial foi ajustada para estimar o número de marcadores necessários para assegurar que os CVs associados às estimativas de dissimilaridade fossem menores ou iguais a 10%, valor considerado aceitável nesta pesquisa. Como em Garcia et al. (2004) [51], a mediana (e não a média) das estimativas do coeficiente de variação foram utilizadas para avaliar a exatidão dos valores de dissimilaridade. Todos estes procedimentos foram realizados no software RStudio [52].

Análise de diversidade e estrutura genética

Com o objetivo de detectar a presença de algum padrão de agrupamento entre os genótipos avaliados, a Análise de Coordenadas Principais (PCoA) [53–55] foi realizada com base nas dissimilaridades (medidas do complemento da semelhança de Jaccard) entre cada par de acessos. A análise da variância molecular (AMOVA) entre os grupos predefinidos foi realizada pelo software Arlequin v.3.5 [56]. O número de subpopulações (k) e a proporção de identidade entre os indivíduos foram obtidos usando o software Structure v.2.3.4 [57,58]. O valor de K foi definido entre 1 a 10 (K-valor), com 10 iterações a 100,000 “burning period” e 200,000 repetições de “Markov chain Monte Carlo” (MCMC) [10]. Valores de K e ΔK foram obtidos pelo software Structure Harvester [59]. Melhores valores de K foram confirmados através da verificação dos resultados pelo pipeline “Best K” do software CLUMPAK [60].

Análise filogenética e de dissimilaridade

Como mencionado anteriormente, a dissimilaridade em pares foi calculada com base no complemento do coeficiente de similaridade de Jaccard ($Diss = 1 - Jaccard$) [40,61]. A árvore filogenética foi construída de acordo com o método Neighbor-Joining com 1000 reamostragens usando o software DARwin v.6.0.13 [62,63], usado também na identificação dos valores extremos de dissimilaridade.

Resultados

Análise da eficiência dos iniciadores SSR

Os 12 pares de iniciadores SSR geraram 285 alelos (marcadores) entre os 137 genótipos. O número de alelos variou de 10 (CV60) a 43 (SCC423) com uma média de 23,75 alelos por par de iniciadores (Tabela 2).

Em geral, os pares de iniciadores SSR CV29, CV38, CV79, SCB213, SCB312, SCB381 e SCC423 apresentaram os maiores valores para todos os índices avaliados (Tabela 2). Os pares de iniciadores CV37, CV60, CV94, CV106, SCB381 e SCB436 apresentaram polimorfismo de moderado a elevado, produzindo 10 a 24 alelos. A avaliação dos haplótipos revelou que oito pares de iniciadores (CV29, CV37, CV38, CV79, SCB213, SCB312, SCB381 e SCC423) foram capazes de identificar mais de 90% dos indivíduos, enquanto CV106 e SCB436 identificaram 64,2% e 59,9% e CV94 e CV60 identificaram 38,0% e 29,9%, respectivamente.

Tabela 2: Número total de alelos (n) e estimativas de PIC, Diversidade Genética (D), Razão Multiplex Efetiva (E), Índice de Marcadores (MI), Poder de Resolução (PR) e número de perfis identificados entre 137 indivíduos ("Todos" representa os resultados para todos os 137 acessos, "C" representa os resultados considerando apenas o grupo das Cultivares).

SSR	N	PIC	D	E	MI	R_p	Perfis Identificados (%)
		Todos / C	Todos / C	Todos / C	Todos / C	Todos / C	
CV29	23	0,906 / 0,913	0,913 / 0,919	7,387 / 8,272	6,800 / 7,552	10,478 / 8,938	131 (95,6%)
CV37	20	0,879 / 0,878	0,889 / 0,888	5,234 / 6,012	4,692 / 5,278	7,119 / 6,074	125 (91,2%)
CV38	27	0,934 / 0,903	0,938 / 0,910	6,693 / 7,037	6,218 / 6,357	10,94 / 7,654	134 (97,8%)
CV60	10	0,788 / 0,720	0,813 / 0,758	2,956 / 2,938	2,262 / 2,115	3,149 / 2,568	41 (29,9%)
CV79	27	0,901 / 0,874	0,908 / 0,885	5,949 / 5,951	5,358 / 5,203	7,448 / 5,877	125 (91,2%)
CV94	14	0,776 / 0,734	0,804 / 0,772	3,781 / 3,605	3,185 / 2,646	4,896 / 1,778	52 (38,0%)
CV106	14	0,809 / 0,848	0,831 / 0,863	4,175 / 4,519	3,616 / 3,831	6,746 / 5,605	88 (64,2%)
SCB213	30	0,919 / 0,883	0,924 / 0,892	5,569 / 4,728	5,148 / 4,175	9,746 / 7,284	131 (95,6%)
SCB312	29	0,906 / 0,922	0,913 / 0,927	7,044 / 7,531	6,565 / 6,946	11,746 / 10,642	133 (97,1%)
SCB381	24	0,894 / 0,916	0,902 / 0,921	7,255 / 8,457	6,668 / 7,743	9,761 / 8,543	128 (93,4%)
SCB436	24	0,818 / 0,850	0,838 / 0,832	4,628 / 5,235	3,973 / 4,448	4,955 / 2,889	82 (59,9%)
SCC423	43	0,925 / 0,906	0,929 / 0,912	6,285 / 5,765	5,854 / 5,222	11,000 / 8,198	133 (97,1%)
Total	285	-	-	-	-	-	-
Média	23,75	0,871 / 0,862	0,883 / 0,873	5,580 / 5,837	5,028 / 5,126	6,914 / 6,337	109 (79,25%)

As estimativas de PIC variaram de 0.776 (CV94) a 0.934 (CV38) com valor médio de 0.871 para todos os 137 acessos e de 0.720 (CV60) a 0.922 (SCB312) com média de 0.862 considerando somente as cultivares. As estimativas de Diversidade Genética variaram de 0,804 (CV94) a 0,938 (CV38) com média de 0,883 para todos os acessos e de 0,758 (CV60) a 0,927 (SCB312) com média de 0,873 para as cultivares. Os valores estimados de Razão Multiplex efetiva variaram de 2,956 (CV60) a 7,387 (CV29), com média de 5,580 para todos os acessos, e de 2,938 (CV60) a 8,457 (SCB381), com média de 5,837 para as cultivares. Os valores estimados do Índice de Marcadores variaram de 2,262 (CV60) a 6,800 (CV29) com média de 5,028 para todos os acessos e de 2,155 (CV60) a 7,743 (SCB381) com média de 5,166 para as cultivares.

Para avaliar a capacidade dos pares de iniciadores de detectar polimorfismo, calculou-se PR para todos os genótipos. Suas estimativas variaram de 3,149 (CV60) a 11,746 (SCB312). Além disso, RP foi calculado também considerando apenas as cultivares para identificar a capacidade de detectar polimorfismo em um grupo mais homogêneo, variando de 1,778 (CV94) a 10,642 (SCB312).

A melhor correspondência entre perfis identificados e as medidas de polimorfismo para cada iniciador SSR (PIC, MI, E, D e PR) foi dada pelo Poder de Resolução.

O número de marcadores utilizados nesta pesquisa foi suficiente para estimar as diferenças genéticas entre pares de acessos com nível aceitável de precisão. Considerando todos os 285 marcadores, o CV associado às estimativas de similaridade (Jaccard) foi de 6,86%, portanto, abaixo do limiar previamente estabelecido (10%). De fato, cerca de 140 marcadores seriam suficientes para obter uma estimativa de CV média em torno de 10% (Figura 1).

Distribuição da frequência alélica e alelos exclusivos

O número de alelos observados para as cultivares comerciais, *S. spontaneum* e *S. officinarum* foi de 166, 246 e 195, respectivamente (dados não apresentados). Trinta e três alelos (11,58%) foram exclusivos de *S. spontaneum* e quatro (1,43%) de *S. officinarum* quando considerados apenas o germoplasma básico. Quatro alelos (SCB381.0244, SCC423.0300, CV38.0210 e CV106.0135) que representam 1,43%

dos alelos totais, estavam presentes apenas em cultivares comerciais. Cento e dezenove alelos (41,75%) espalhados por acessos de germoplasma básico não estavam presentes em cultivares. Destes, quatro alelos estavam presentes apenas em *S. officinarum* (1,43%), 33 apenas em *S. spontaneum* (11,58%) e 18 apenas em outras espécies / híbridos acessos (6,34%, Figuras 2 a 4).

Com base nos alelos exclusivos putativos de *S. officinarum* e *S. spontaneum* previamente identificados, 15 cultivares (SP842025, SP901638, SP841201, SP845560, SP811763, SP841431, SP891115, CTC02, CTC05, CTC06, CTC08, CTC10, RB835054, RB865230, RB813804) mostraram 2 alelos exclusivos de *S. officinarum* (SCB436.0186 e CV37,0169) e 20 outras cultivares (SP841201, SP8642, SP87344, SP832847, SP891115, SP911049, CTC03, CTC10, CTC11, IAC873396, IAC911099, IACSP942101, IACSP953028, IACSP955000, RB925211, RB925345, RB931530, RB928064, RB835054, RB855536) mostraram 6 alelos exclusivos (SCB381.0236, SCB436.0199, SCB312.0208, CV79.0228, CV106.0132, CV106.0078) derivados de *S. spontaneum* (dados não mostrado). Não foram encontrados alelos exclusivos das espécies diversas nas cultivares. Os alelos SCB381.0244, SCC423.0300, CV38.0210 e CV106.0135 estavam ausentes no germoplasma básico, portanto, excluindo 12 cultivares (SP775181, SP803280, SP801836, RB925268, SP87396, SP842025, SP855077, IACSP942101, CTC13, CTC14, CTC17, SP813250).

Análise de estrutura e diferenciação genética

Os dois primeiros componentes principais (CPs) explicaram, em conjunto, 19,18% da variabilidade total expressa entre genótipos, com base na medida de dissimilaridade (Figura 5). De acordo com o primeiro CP, as cultivares são agrupadas em um agrupamento isolado dos outros. Observando o segundo CP, os acessos de *S. officinarum* formam um grupo distinto, enquanto que *S. spontaneum* e os outros genótipos (outras espécies e híbridos exóticos) tendem a se agrupar. Por outro lado, parece haver uma ligeira diferenciação entre estes dois grupos, uma vez que os acessos de *S. spontaneum* se concentram no primeiro quadrante do gráfico. O grau de estimativa de diferenciação genética (Φ_{ST}) obtido entre as cultivares comerciais e o germoplasma básico foi de 0,179 ($P < 0,001$), o que significa que

17,91% da variação total avaliada pelos SSRs é distribuída entre estes dois grupos enquanto 82,09 % Está dentro deles. A estimativa da diferenciação genética entre cultivares comerciais e acessos de *S. officinarum* foi de 0,211 ($P < 0,001$), o que foi ligeiramente inferior ao observado entre as cultivares comerciais e *S. spontaneum* ($\Phi ST = 0,216$; $P < 0,001$). Para comparação entre as duas espécies, a estimativa de diferenciação foi de 0,1270, indicando menor diferenciação entre estes dois grupos do que para as outras comparações, sendo que 12,70% da variabilidade genética está entre as duas espécies, e 87,30% encontra-se dentro delas.

Tabela 3: Análise de variância molecular (AMOVA) para grupos predefinidos.

População	Fonte de Variação	Graus de Liberdade	Soma dos quadrados	Componentes de Variância	% de variação
Germoplasma Básico e Cultivares	Entre	1	437.195	6.17473 Ve	17.91
	Dentro	135	3821.972	28.31091 Vd	82.09
	Total	136	4259.168	34.48564	
	(ΦST_1)	0.1790			
Cultivares e <i>S. officinarum</i>	Entre	1	278.349	7.25953 Ve	21.09
	Dentro	101	2742.699	27.15544 Vd	78.91
	Total	102	2021.049	34.41497	
	(ΦST_2)	0.2109			
Cultivares e <i>S. spontaneum</i>	Entre	1	412.637	8.51485 Ve	21.64
	Dentro	110	3392.113	30.83739 Vd	78.36
	Total	111	3804.750	39.35224	
	(ΦST_3)	0.2163			
<i>S. officinarum</i> e <i>S. spontaneum</i>	Entre	1	187.674	5.75541 Ve	12.70
	Dentro	51	2017.232	39.55456 Vd	87.30
	Total	52	2204.906	45.30897	
	(ΦST_4)	0.1270			

De acordo com a análise de estrutura, 10 subconjuntos de agrupamento resultaram para cada valor de entrada de K, que foram submetidos a análise no software Structure Harvester e ao “pipeline” Best K da CLUMPAK. O melhor valor de K foi 2 ($\Delta K = 318.196$, Figura 6a), sugerindo que os 137 genótipos podem ser agrupados em dois grupos (Figura 6b) correspondentes às cultivares e germoplasma básico. O segundo valor mais relevante para K é 3 ($\Delta K = 117.164$), sugerindo que o grupo formado pelo germoplasma básico pode ser subdividido de acordo com suas principais espécies (*S. officinarum* e *S. spontaneum*) capturando a maior parte da estrutura genética em três agrupamentos (Figura 6c).

Análise filogenética e de dissimilaridade

Os valores médios de dissimilaridade dentro das cultivares, *S. officinarum*, *S. spontaneum* e espécies diversas foram 0,176, 0,206, 0,248 e 0,285, respectivamente. Os valores médios de dissimilaridade entre as cultivares e *S. officinarum* foram de 0,247, entre as cultivares e *S. spontaneum* foi de 0,277, entre as cultivares e os acessos de espécies diversas foi de 0,272. Em relação às espécies, a média de dissimilaridade entre os acessos de *S. officinarum* e *S. spontaneum* foi de 0,279. A dissimilaridade média entre os acessos de espécies diversas, *S. officinarum* e *S. spontaneum* foi de 0,282 e 0,278, respectivamente (Tabela 4).

Tabela 4: Dissimilaridades entre grupos predefinidos de acordo com a função “Extreme Values” do software DARwin.

	Cultivares	<i>S. officinarum</i>	<i>S. spontaneum</i>	Diversos
Cultivares	0.176	-	-	-
<i>S. officinarum</i>	0.247	0.206	-	-
<i>S. spontaneum</i>	0.277	0.279	0.248	-
Diversos	0.272	0.282	0.278	0.285

Comparado ao grupo composto pelas cultivares comerciais, MIDAZ (*S. officinarum* x *S. barberi*) e CHUNNEE foram, respectivamente, os acessos geneticamente mais semelhantes (0,203) e divergentes (0,326) do grupo de germoplasma básico (Tabela S2).

Observou-se um baixo valor de dissimilaridade entre IAC873396 e SP87344 (0,077) que são irmão germanos oriundos do cruzamento entre CO740 e SP701143 (dados não mostrados). A avaliação dos seus haplótipos foi realizada para detectar erros de rotulagem ou mistura de material, confirmando diferentes padrões moleculares (dados não mostrados).

A árvore filogenética revelou a prevalência de dois grupos principais (cultivares e germoplasma básico), o que está de acordo com agrupamento apontado pelo software Structure ($K = 2$). Além disso, os acessos foram agrupados principalmente de acordo com suas espécies (*S. officinarum* e *S. spontaneum*) com os acessos das outras espécies e híbridos exóticos espalhados entre esses subgrupos, sugerindo a existência de três grupos (Figura 7).

Discussão

Eficiência do conjunto de iniciadores SSR para gerenciamento de germoplasma de cana-de-açúcar e identificação de cultivares

Sem dúvida, a disponibilidade de um conjunto eficiente e adequado de iniciadores SSR é de grande impacto para uma gestão rotineira de germoplasma de cana-de-açúcar e identificação de cultivares de maneira econômica. Vários índices estão disponíveis para avaliar a eficiência do iniciador e revelar o polimorfismo detectado [21,22,28,47]. De modo geral, os índices aqui utilizados para medir a eficácia dos iniciadores apresentaram valores elevados quer quando se consideram todos os genótipos quer apenas o painel de cultivares. Este resultado indica que uma grande quantidade de informação genética foi capturada por estes 12 iniciadores SSR em uma amostra altamente diversa de espécies do gênero *Saccharum*.

Microsatélites permitem a amplificação dos diferentes tipos de alelos em um loco. No entanto, o nível de ploidia da cana-de-açúcar, e consequentemente a heterozigosidade, implicam em mais de dois tipos de alelos por loco. Portanto, o padrão de bandas de cana-de-açúcar revela os diferentes tipos de alelos de um loco definido por um par de iniciadores SSR, embora a dosagem dos alelos ou o número de cópias não possa ser determinado [20,40,43]. Como consequência, pode ser observado um número considerável de alelos numa linha de gel para um único indivíduo com um único iniciador SSR. De fato, o número médio de alelos obtidos com os 12 iniciadores SSR estabelecidos foi promissor em comparação com os relatados por Nayak et al. (2014) [10] em que 209 alelos foram produzidos por 36 iniciadores SSR em 1002 acessos da Coleção Mundial de Cana-de-açúcar e Gramíneas Relacionadas (WCSRG), e por Pan (2010) [64], que obteve 144 alelos derivados de 21 iniciadores SSR em um conjunto de indivíduos diversos de *Saccharum*, utilizados na construção de um banco de dados moleculares.

O elevado número de alelos obtidos pelos 12 iniciadores de SSR também impactou positivamente os valores de PIC, as estimativas da Razão Multiplex Efetiva (E) e do Índice de Marcador (IM). Como cada par de primers SSR corresponde normalmente a um loco [46], a Razão Multiplex Efetiva é igual à Razão Multiplex (n), pois a fração de locus polimórfico (β) é 1. A maioria dos valores de PIC foi acima de

0,7 para o total de acessos ou apenas cultivares. O Poder de Resolução (PR) que revela a capacidade de um dado par de iniciadores para distinguir genótipos [47], foi elevado para a metade dos iniciadores SSR, e a maioria deles foi capaz de distinguir mais de 90% do total de acessos avaliados.

Estrutura genética das cultivares comerciais brasileiras

Um número muito limitado de acessos foi utilizado nos primeiros cruzamentos interespecíficos para desenvolver as primeiras cultivares comerciais [13,65]. Apenas 19 *Saccharum officinarum*, 2 *S. spontaneum* e 1 *S. sinense* foram intensivamente utilizados para constituir a base genética da cana-de-açúcar moderna durante as primeiras hibridizações interespecíficas [12]. Uma vez que o melhoramento começou com uma série de cruzamentos entre poucos clones nativos com alto rendimento de açúcar, provenientes de comunidades de ilhas asiáticas, é difícil determinar quais clones foram utilizados. Esta primeira proporção genética parece ser mantida com pouca adição de variação genética entre as gerações de programas de melhoramento, estreitando a base genética das cultivares comerciais atuais, com aumento da contribuição de *S. officinarum* pela nobilitação [66].

Do total de alelos exclusivos detectados no germoplasma básico, *S. spontaneum* apresenta a maior quantidade deles (58,21%), seguido pelos híbridos exóticos e espécies diversas (26,87%) e por *S. officinarum* (9,67%). De fato, *S. spontaneum* é considerada a espécie mais diversa do gênero *Saccharum* devido à sua grande distribuição ecogeográfica [67,68].

Como esperado, um número razoável de alelos exclusivos derivados do germoplasma básico (19,30%) não foi herdado pelas cultivares comerciais brasileiras. Relatamos a presença de alelos exclusivos derivados das cultivares comerciais, sugerindo que o germoplasma básico amostrado em nosso trabalho não abrange o “pool” genético total utilizado nos programas de melhoramento prévios, ou que podem ter ocorrido mutações nesses loci de microssatélites. Estes loci SSR nas cultivares comerciais poderiam ter, ao longo do tempo, sofrido substituição nucleotídica, adição ou deleção, pelo fato de as regiões de microssatélites terem altas taxas de mutação como resultado do deslizamento de polimerase de DNA ou de “crossing-over” desbalanceado [69, 70]. No entanto, o germoplasma básico

engloba a maior parte dos alelos que ocorrem em cultivares (98,60%), exibindo alta representatividade de sua riqueza alélica.

Na tentativa de rastrear a origem específica dos alelos presentes nas cultivares comerciais brasileiras, foram considerados alelos exclusivos putativos de *S. officinarum* e *S. spontaneum*. Dos poucos alelos exclusivos de origem putativa das espécies dentro das cultivares, 15,38% derivaram de *S. spontaneum* enquanto 33,30% de *S. officinarum* (dados não apresentados). Esse resultado pode se referir ao processo de nobilitação, já que quase 90% da constituição genômica das cultivares comerciais é oriunda de *S. officinarum*, indicando que a maioria dos alelos presentes foi compartilhada com esta espécie. Por outro lado, a proporção de alelos exclusivos de *S. spontaneum* encontrada nas cultivares está de acordo com a contribuição genômica dessa espécie (cerca de 10-15%) esperada [9]. Esse resultado também sugere que esta espécie teve uma notória contribuição para a variabilidade genética das cultivares utilizadas no Brasil. Além disso, os valores de diferenciação genética entre elas e *S. officinarum* foram ligeiramente menores (0,211) do que entre as cultivares e *S. spontaneum* (0,216). Os altos valores de diferenciação genética entre as cultivares e essas duas principais espécies, por sua vez, é um reflexo do efeito fundador, pois poucos clones dessas duas espécies progenitoras contribuíram para a base genética atual das cultivares. Além disso, como esperado, este resultado sinaliza que uma variabilidade genética considerável pode ser incorporada no “pool” genético, explorando estas duas espécies principais, embora se saiba que, apesar do número limitado de genótipos empregados nos primeiros programas de melhoramento, a base genética não é tão estreita quanto se pensava [61], especialmente se considerarmos a ploidia elevada e a diversidade genética das espécies ancestrais principais.

Análise de agrupamento e de dissimilaridade genética

Os quatro grupos previamente estabelecidos (Cultivares, *S. officinarum*, *S. spontaneum* e Híbridos e Espécies Diversas) são confirmados pelos resultados da Análise de Coordenadas Principais. As cultivares foram atribuídas a um cluster separado, o que era esperado, uma vez que se originaram de poucos genitores e foram selecionados para o mesmo perfil, visando atender a alta demanda para a

produção de açúcar e etanol. A espécie *S. officinarum* tem um perfil particular (alto teor de açúcar, baixo teor de fibras, susceptibilidade a estresses bióticos e não-bióticos, etc.), que difere das outras espécies relacionadas. De acordo com a diferenciação fenotípica, nessa pesquisa observou-se divergência a nível molecular entre os genótipos de *S. officinarum* e os demais espécimes, justificando a formação de um aglomerado isolado composto pelos genótipos desta espécie. *S. spontaneum* e as outras espécies e híbridos exóticos são mais semelhantes entre eles, e algumas sobreposições foram mostradas entre estes grupos. No entanto, há uma ligeira tendência para a separação entre eles, como pode ser observado na Figura 5. Todos esses resultados estão de acordo com os obtidos pelos outros métodos aplicados neste estudo.

Os acessos de *S. officinarum* apresentaram baixa dissimilaridade, revelando estreita relação genética. Maiores valores de dissimilaridade foram encontrados, principalmente, entre cultivares e acessos de *S. spontaneum* (0,277), o que concorda com a contribuição desta espécie nas primeiras hibridizações, e entre cultivares e os híbridos exóticos e espécies diversas, revelando o distanciamento genético entre esses grupos. Este resultado tem aplicação prática para os programas de melhoramento da cana-de-açúcar focados na cana energética, uma vez que a escolha de parentais mais divergentes (Tabela S2) amplifica a variabilidade genética na progênie.

Os valores médios de dissimilaridade obtidos nas cultivares foram 1,2 e 1,4 vezes inferiores aos de *S. officinarum* e *S. spontaneum*, respectivamente. O valor médio de dissimilaridade dentro das cultivares implica que, em média, elas compartilharam 82% das regiões genômicas avaliadas com SSR. Lima et al. 2002 [61] investigaram a similaridade genética avaliada por marcadores derivados de AFLP em um grupo de 79 genótipos compostos principalmente por importantes cultivares históricos e cultivares atuais utilizadas como pais em um programa de melhoramento brasileiro e também quatro genótipos de espécies de *Saccharum* (*S. sinense*, *S. Barberi* e dois de *S. officinarum*). O valor médio de similaridade genética foi de 0,47 e é inferior ao valor encontrado na presente pesquisa, na que apenas as cultivares comerciais foram avaliadas.

As análises filogenéticas e de agrupamento revelaram a relação genética entre os genótipos avaliados. Eles tendem a se agrupar de acordo com sua espécie e genealogia, formando grupos consistentes de acordo com os dados genealógicos disponíveis. No entanto, algumas discrepâncias foram observadas. Dois indivíduos de *S. spontaneum*, GLAGAH e IN8488, foram agrupados dentro de *S. officinarum*. De fato, apesar da sua classificação como *S. Spontaneum*, GLAGAH é um provável híbrido interespecífico natural. Embora IN8488 seja considerado um clone de *S. spontaneum*, a proporção de identidade mostrou alta correlação com *S. Officinarum*, sugerindo que seja também um híbrido.

POJ2878 e KASSOER são proximamente agrupados na árvore filogenética. Além disso, a análise da estrutura revelou quase a mesma proporção de identidade dos três grupos em ambos os indivíduos, revelando uma baixa dissimilaridade genética já esperada (0,0947), dado que a cultivar POJ2878 é um descendente F2 direto de KASSOER [11].

Referências Bibliográficas

1. CONAB. Acompanhamento da Safra Brasileira - Cana-de-açúcar– Safra 2015/16, terceiro levantamento. 2015; 65.
2. UNICA. Histórico de produção e moagem [Internet]. 2016 [cited 1 Jan 2016]. Available: <http://www.unicadata.com.br/historico-de-producao-e-moagem.php>
3. Daniels J, Roach BT. Taxonomy and Evolution. In: Heinz DJ, editor. Sugarcane Improvement through Breeding. 1st ed. Amsterdam: Elsevier; 1998. pp. 7–84.
4. Jeswiet J. Immunity and cross-fertilisation in the genus *Saccharum*. Warner MF, Sherman MA, editors. Recl des Trav Bot Neerl. Wageningen: USDA; 1928;25A: 185–202.
5. Berding N, Roach BT. Germplasm collection, maintenance and use. In: Heinz DJ, editor. Sugarcane Improvement through Breeding. 1st ed. Amsterdam: Elsevier; 1998. pp. 143–210.
6. Pan YB, Burner DM, Legendre BL. An assessment of the phylogenetic relationship among sugarcane and related taxa based on the nucleotide sequence of 5S rRNA intergenic spacers. *Genetica*. 2000;108: 285–95.
7. D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome*. 1998;41: 221–225. doi:10.1139/gen-41-2-221
8. Pinto LR, Oliveira KM, Marconi TG, Garcia AAF, Ulian EC, Souza AP. Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. *PLANT Breed*. 2006;125: 378–384.
9. Jannoo N, Grivet L, Seguin M, Paulet F, Domaingue R, Rao PS, et al. Molecular investigation of the genetic base of sugarcane cultivares. *Theor Appl Genet*. 1999;99: 171–184. doi:10.1007/s001220051222
10. Nayak SN, Song J, Villa A, Pathak B, Ayala-Silva T, Yang X, et al. Promoting utilization of *Saccharum* spp. genetic resources through genetic diversity analysis and core collection construction. *PLoS One*. 2014;9: e110856. doi:10.1371/journal.pone.0110856
11. Mutsaers H. Peasants, Farmers and Scientists. 1st ed. Springer. Dordrecht: Springer Netherlands; 2007. doi:10.1007/978-1-4020-6166-0

12. Ramdoyal K, Badaloo G. Prebreeding in Sugarcane with an Emphasis on the Programme of the Mauritius Sugar Industry Research Institute. In: Engels JM, Ramanatha R, Brown AH, Jackson MT, editors. Managing Plant Genetic Diversity. 1st ed. Oxford, UK: CABI Publishing; 2002. pp. 307–321.
13. Arceneaux G. Cultivated Sugarcane of the World and their Botanical Derivation. Proc ISSCT. 1967;12: 844–854.
14. Faulks LK, Kerezszy A, Unmack PJ, Johnson JB, Hughes JM. Going, going, gone? Loss of genetic diversity in two critically endangered Australian freshwater fishes, *Scaturiginichthys vermeilipinnis* and *Chlamydogobius squamigenus*, from Great Artesian Basin springs at Edgbaston, Queensland, Australia. Aquat Conserv Mar Freshw Ecosyst. 2016; doi:10.1002/aqc.2684
15. Frankham R, Ballou JD, Briscoe DA. Introduction to Conservation Genetics. 1st ed. Forest Ecology and Management. Cambridge, UK: Cambridge University Press; 2002. doi:10.1016/j.foreco.2003.12.001
16. Jordan DR, Tao YZ, Godwin ID, Henzell RG, Cooper M, McIntyre CL. Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. Euphytica. 1998;102: 1–7. doi:10.1023/A:1018311908636
17. Gashaw ET, Mekbib F, Ayana A. Genetic Diversity among Sugarcane Genotypes Based on Qualitative Traits. Adv Agric. 2016;2016: 8. doi:http://dx.doi.org/10.1155/2016/8909506
18. CTC. Variedades CTC [Internet]. 2017. Available: <http://www.ctcanavieira.com.br/melhoramentogenetico.html>
19. Berding N, Hogarth M, Cox M. Plant Improvement of Sugarcane. In: James G, editor. World Agriculture Series: Sugarcane. 2nd ed. Oxford, UK: Blackwell Publishing; 2004. pp. 20–53.
20. Silva DCD, Filho L, Santos J dos, Barbosa G de S, Almeida C, Sérgio L, et al. DNA fingerprinting based on simple sequence repeat (SSR) markers in sugarcane clones from the breeding program RIDESA. African J Biotechnol. 2012;11: 4722–4728. doi:10.5897/AJB11.2614
21. Pan Y-B. Molecular genotyping of sugarcane clones with microsatellite DNA markers.pdf. 2003.

22. Marconi TG, Costa EA, Miranda HR, Mancini MC, Cardoso-Silva CB, Oliveira KM, et al. Functional markers for gene mapping and genetic diversity studies in sugarcane. *BMC Res Notes*. BioMed Central Ltd; 2011;4: 264.
doi:10.1186/1756-0500-4-264
23. Aitken KS, Jackson PA, McIntyre CL. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theor Appl Genet*. 2005;110: 789–801.
doi:10.1007/s00122-004-1813-7
24. Jarne P, Lagoda PJ. Microsatellites, from molecules to populations and back. *Trends Ecol Evol*. 1996;11: 424–9.
25. Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, et al. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. 1996; 225–238.
26. Vieira MLC, Santini L, Diniz AL, Munhoz C de F. Microsatellite markers: What they mean and why they are so useful. *Genet Mol Biol*. 2016;39: 312–328.
doi:10.1590/1678-4685-GMB-2016-0027
27. dos Santos JM, Duarte Filho LSC, Soriano ML, da Silva PP, Nascimento VX, de Souza Barbosa GV, et al. Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. *Ind Crops Prod*. Elsevier B.V.; 2012;40: 145–150. doi:10.1016/j.indcrop.2012.03.005
28. Maccheroni W, Jordão H, Gaspari R, Moura G, Matsuoka S. Development of a dependable microsatellite-based fingerprinting system for sugarcane. *Sugar Cane Intl*. 2009;27: 8–13.
29. Pan Y. Highly Polymorphic Microsatellite DNA Markers for Sugarcane Germplasm Evaluation and Variety Identity Testing. *Sugar Tech*. 2006;8: 246–256.
30. Oliveira KM, Pinto LR, Marconi TG, Margarido GR a., Pastina MM, Teixeira LHM, et al. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. *Mol Breed*. 2007;20: 189–208. doi:10.1007/s11032-007-9082-1

31. Pinto LR, Oliveira KM, Ulian EC, Garcia AAF, de Souza AP. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome*. 2004;47: 795–804. doi:10.1139/g04-055
32. Pan Y, Scheffler B, Richard J. High-throughput genotyping of commercial sugarcane clones with microsatellite (SSR) DNA markers. *Sugar Technol*. 2007;9: 176–181.
33. Edmé SJ, Glynn NG, Comstock JC. Genetic segregation of microsatellite markers in *Saccharum officinarum* and *S. spontaneum*. *Heredity (Edinb)*. 2006;97: 366–375. doi:10.1038/sj.hdy.6800888
34. Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann HP, Garcia AAF, et al. Sugarcane improvement: How far can we go? *Curr Opin Biotechnol*. Elsevier Ltd; 2012;23: 265–270. doi:10.1016/j.copbio.2011.09.002
35. Aitken K, McNeill M. Diversity Analysis. In: Henry R, Kole C, editors. *Genetics, Genomics and Breeding of Sugarcane*. 1st ed. Enfield, New Hampshire: Science Publishers; 2010. pp. 19–42.
36. Scordia D, Consentino S, Jeffries T. Second generation bioethanol production from *Saccharum spontaneum* L. ssp. *aegyptiacum* (Willd.) Hack. *Bioresour Technol*. 2010;101: 5358–5365.
doi:http://dx.doi.org/10.1016/j.biortech.2010.02.036
37. Ming R, Moore PH, Wu K-K, D'Hont A, Glaszmann JC, Tew TL, et al. Sugarcane Improvement through Breeding and Biotechnology. In: Janick J, editor. *Plant Breeding Reviews*. 1st ed. Hoboken, NJ, USA: Wiler; 2006. pp. 15–118.
38. Aljanabi S, Forget L, Dookun a. An improved and rapid protocol for the isolation of polysaccharide-and polyphenol-free sugarcane DNA. *Plant Mol Biol Report*. 1999;17: 1–8. doi:10.1023/a:1007692929505
39. Palhares AC, Rodrigues-Morais TB, Van Sluys M-A, Domingues DS, Maccheroni W, Jordão H, et al. A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. *BMC Genet*. 2012;13: 51. doi:10.1186/1471-2156-13-51

40. Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas SM, et al. Characterization of new polymorphic functional markers for sugarcane. 2009;209: 191–209. doi:10.1139/G08-105
41. Silva JAG. Preliminary analysis of microsatellite markers derived from sugarcane expressed sequence tags (ESTs). Genet Mol Biol. 2001;24: 155–159.
42. Parida SK, Pandit A, Gaikwad K, Sharma TR, Srivastava PS, Singh NK, et al. Functionally relevant microsatellites in sugarcane unigenes. BMC Plant Biol. BioMed Central Ltd; 2010;10: 251. doi:10.1186/1471-2229-10-251
43. Cordeiro GM, Pan YB, Henry RJ. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. Plant Sci. 2003;165: 181–189. doi:10.1016/S0168-9452(03)00157-2
44. Kemp S. Pic Calculator. 2002.
45. Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to *Erianthus* and *Sorghum*. Plant Sci. 2001;160: 1115–1123.
46. Varshney RK, Chabane K, Hendre PS, Aggarwal RK, Graner A. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. Plant Science. 2007;173: 638-649. doi:10.1016/j.plantsci.2007.08.010
47. Prevost A, Wilkinson MJ. A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivares. Theor Appl Genet. 1999;98: 107–112. doi:10.1007/s001220051046
48. Efron B, Tibshirani RJ. An introduction to the bootstrap. CRC Press. 1994;
49. Tivang JG, Nienhuis J, Smith OS. Estimation of sampling variance of molecular marker data using the bootstrap procedure. Theor Appl Genet. 1994;89: 259–64.
50. Hållden C, Nilsson NO, Rading IM, Säll T. Evaluation of RFLP and RAPD markers in a comparison of *Brassica napus* breeding lines. Theor Appl Genet. 1994;88: 123–8.

51. Garcia AA, Benchimol LL, Barbosa AM, Geraldi IO, Souza-Jr CL, Souza APD. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genet Mol Biol.* 2004;27: 579–88.
52. Team Rs. RStudio: Integrated Development for R. Boston, MA; 2016.
53. Sneath P, Sokal R. Numerical taxonomy: the principles and practice of numerical classification. WH Freeman; 1973.
54. Dunn G, Everitt B. An introduction to numerical taxonomy. Cambridge, UK: Cambridge Univesity Press; 1982.
55. Manly B. Multivariate statistical methods: a primer. 2nd ed. Chapman & Hall; 1994.
56. Excoffier L, Lischer H. Arlequin suite ver. 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10: 564–567.
57. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155: 945–959.
58. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 2005;14: 2611–20.
59. Earl D, Von-Holdt B. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 2010;4: 359–361.
60. Kopelman N, Mayzel J, Jakobsson M, Rosenberg N, Mayrose I. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. 2012.
61. Lima MLA, Garcia AAF, Oliveira KM, Matsuoka S, Arizono H, De Souza CL, et al. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum spp.*). *Theor Appl Genet.* 2002;104: 30–38. doi:10.1007/s001220200003
62. Perrier X, Jacquemoud-Collet J. DARwin software. 2006.
63. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987;4: 406–25.

64. Pan Y-B. Databasing Molecular Identities of Sugarcane (*Saccharum* spp.) Clones Constructed with Microsatellite (SSR) DNA Markers. *Am J Plant Sci.* 2010;1: 87–94. doi:10.4236/ajps.2010.12011
65. Roach BT, Heinz PDJ. CYTOLOGICAL STUDIES IN *Saccharum*. CHROMOSOME TRANSMISSION IN INTERSPECIFIC AND INTERGENERIC CROSSES. 1969;
66. Pan Y, Burner D, Legendre B, Grisham M, White W. An assessment of the genetic diversity within a collection of *Saccharum spontaneum* L. with RAPD-PCR. *Genet Resour Crop Evol.* 2004;51: 895–903.
67. Liu X, Li X, Liu H, Xu C, Lin X, Li C, et al. Phylogenetic Analysis of Different Ploidy *Saccharum spontaneum* Based on rDNA-ITS Sequences. *PLoS One.* 2016;11: e0151524. doi:10.1371/journal.pone.0151524
68. Alwala S, Kimbeng C. Molecular Genetic Linkage Mapping in *Saccharum*: Strategies, Resources and Achievements. In: Henry R, Kole C, editors. *Genetics, Genomics and Breeding of Sugarcane*. Enfield, New Hampshire, New Hampshire: Science Publishers; 2010. pp. 69–96.
69. Costa MLM Da, Amorim LLB, Onofre AVC, Melo LJOT de, Oliveira MBM de, Carvalho R de, et al. Assessment of Genetic Diversity in Contrasting Sugarcane Varieties Using Inter-Simple Sequence Repeat (ISSR) Markers. *Am J Plant Sci.* 2011;2: 425–432. doi:10.4236/ajps.2011.23048
70. Schlotterer C, Tautz D. Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.* 1992;20: 211–215.

Figuras e Material suplementar

Tabela S1 - Lista dos genótipos utilizados neste estudo com as respectivas espécies, genitores femininos, masculinos e origem.

#	Genótipo	Espécie	Genitor Fem.	Genitor Masc.	Origem
1	SP87396	<i>S. officinarum</i> híbrido	Q76	?	Brasil - CTC
2	SP842025	<i>S. officinarum</i> híbrido	CO1007	SP716180	Brasil - CTC
3	SP903414	<i>S. officinarum</i> híbrido	SP801079	SP823544	Brasil - CTC
4	SP801816	<i>S. officinarum</i> híbrido	SP711088	H575028	Brasil - CTC
5	SP901638	<i>S. officinarum</i> híbrido	SP784601	?	Brasil - CTC
6	SP841201	<i>S. officinarum</i> híbrido	CO62175	?	Brasil - CTC
7	SP801842	<i>S. officinarum</i> híbrido	SP711088	H575028	Brasil - CTC
8	SP813250	<i>S. officinarum</i> híbrido	CP701547	SP711279	Brasil - CTC
9	SP87365	<i>S. officinarum</i> híbrido	SP773067	?	Brasil - CTC
10	SP855077	<i>S. officinarum</i> híbrido	H393633	SP721861	Brasil - CTC
11	SP8642	<i>S. officinarum</i> híbrido	SP701143	?	Brasil - CTC
12	SP775181	<i>S. officinarum</i> híbrido	HJ5741	?	Brasil - CTC
13	SP87344	<i>S. officinarum</i> híbrido	CO740	SP701143	Brasil - CTC
14	SP845560	<i>S. officinarum</i> híbrido	IAC52326	?	Brasil - CTC
15	SP803280	<i>S. officinarum</i> híbrido	SP711088	H575028	Brasil - CTC
16	SP811763	<i>S. officinarum</i> híbrido	CO775	NA5679	Brasil - CTC
17	SP841431	<i>S. officinarum</i> híbrido	H518029	?	Brasil - CTC
18	SP803480	<i>S. officinarum</i> híbrido	Q73	H611820	Brasil - CTC
19	SP832847	<i>S. officinarum</i> híbrido	HJ5741	SP701143	Brasil - CTC
20	SP801836	<i>S. officinarum</i> híbrido	SP711088	H575028	Brasil - CTC
21	SP891115	<i>S. officinarum</i> híbrido	CP731547	?	Brasil - CTC
22	SP86155	<i>S. officinarum</i> híbrido	SP783081	?	Brasil - CTC
23	SP853877	<i>S. officinarum</i> híbrido	H65606	?	Brasil - CTC
24	SP911049	<i>S. officinarum</i> híbrido	SP803328	SP813250	Brasil - CTC
25	SP835073	<i>S. officinarum</i> híbrido	SP711406	SP711088	Brasil - CTC
26	SP80185	<i>S. officinarum</i> híbrido	BO17	?	Brasil - CTC
27	SP801520	<i>S. officinarum</i> híbrido	H483166	SP711088	Brasil - CTC
28	CTC01	<i>S. officinarum</i> híbrido	LCP8110	?	Brasil - CTC
29	CTC02	<i>S. officinarum</i> híbrido	SP821176	?	Brasil - CTC
30	CTC03	<i>S. officinarum</i> híbrido	SP7652	?	Brasil - CTC
31	CTC04	<i>S. officinarum</i> híbrido	SP835073	?	Brasil - CTC
32	CTC05	<i>S. officinarum</i> híbrido	SP821176	?	Brasil - CTC
33	CTC06	<i>S. officinarum</i> híbrido	SP811763	SP791416	Brasil - CTC
34	CTC07	<i>S. officinarum</i> híbrido	RTT8618	?	Brasil - CTC
35	CTC08	<i>S. officinarum</i> híbrido	SP813493	?	Brasil - CTC
36	CTC09	<i>S. officinarum</i> híbrido	SP831483	?	Brasil - CTC
37	CTC10	<i>S. officinarum</i> híbrido	SP785495	?	Brasil - CTC
38	CTC11	<i>S. officinarum</i> híbrido	SP843478	?	Brasil - CTC
39	CTC12	<i>S. officinarum</i> híbrido	SP801842	SP811747	Brasil - CTC
40	CTC13	<i>S. officinarum</i> híbrido	SP842029	?	Brasil - CTC
41	CTC14	<i>S. officinarum</i> híbrido	SP801842	SP775181	Brasil - CTC
42	CTC15	<i>S. officinarum</i> híbrido	SP842025	?	Brasil - CTC
43	CTC17	<i>S. officinarum</i> híbrido	?	?	Brasil - CTC
44	IAC873396	<i>S. officinarum</i> híbrido	CO740	SP701143	Brasil - IAC
45	IAC911099	<i>S. officinarum</i> híbrido	RB785148	?	Brasil - IAC
46	IACSP944004	<i>S. officinarum</i> híbrido	SP826108	SP775181	Brasil - IAC
47	IACSP942101	<i>S. officinarum</i> híbrido	SP775181	RB785148	Brasil - IAC
48	IACSP942094	<i>S. officinarum</i> híbrido	SP847017	?	Brasil - IAC
49	IAC862480	<i>S. officinarum</i> híbrido	US71399	?	Brasil - IAC
50	IACSP953028	<i>S. officinarum</i> híbrido	SP842189	RB855156	Brasil - IAC
51	IACSP955000	<i>S. officinarum</i> híbrido	SP842066	SP80185	Brasil - IAC
52	IACSP933046	<i>S. officinarum</i> híbrido	SP791011	?	Brasil - IAC
53	RB935744	<i>S. officinarum</i> híbrido	RB835089	RB765418	Brasil - RIDESA
54	RB825336	<i>S. officinarum</i> híbrido	H533989	?	Brasil - RIDESA
55	RB925211	<i>S. officinarum</i> híbrido	RB855206	?	Brasil - RIDESA
56	RB925345	<i>S. officinarum</i> híbrido	H591966	?	Brasil - RIDESA
57	RB925268	<i>S. officinarum</i> híbrido	RB855511	?	Brasil - RIDESA
58	RB931530	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
59	RB928064	<i>S. officinarum</i> híbrido	SP701143	?	Brasil - RIDESA
60	RB93509	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
61	RB835054	<i>S. officinarum</i> híbrido	RB72454	NA5679	Brasil - RIDESA
62	RB855035	<i>S. officinarum</i> híbrido	L6014	SP701284	Brasil - RIDESA
63	RB855536	<i>S. officinarum</i> híbrido	SP701143	RB72454	Brasil - RIDESA
64	RB845197	<i>S. officinarum</i> híbrido	RB72454	SP701143	Brasil - RIDESA
65	RB842021	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
66	RB92579	<i>S. officinarum</i> híbrido	RB755126	RB72199	Brasil - RIDESA
67	RB867515	<i>S. officinarum</i> híbrido	RB72454	?	Brasil - RIDESA
68	RB865230	<i>S. officinarum</i> híbrido	RB72454	SP701143	Brasil - RIDESA
69	RB855453	<i>S. officinarum</i> híbrido	TUC717	?	Brasil - RIDESA
70	RB845257	<i>S. officinarum</i> híbrido	RB72454	SP701143	Brasil - RIDESA
71	RB855563	<i>S. officinarum</i> híbrido	TUC717	SP701143	Brasil - RIDESA

Tabela S1: Continuação

#	Genótipo	Espécie	Genitor Fem.	Genitor Masc.	Origem
72	RB855113	<i>S. officinarum</i> híbrido	SP701143	RB72454	Brasil - RIDESA
73	RB855463	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
74	RB813804	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
75	RB855036	<i>S. officinarum</i> híbrido	RB72454	SP701143	Brasil - RIDESA
76	RB855511	<i>S. officinarum</i> híbrido	?	?	Brasil - RIDESA
77	RB855546	<i>S. officinarum</i> híbrido	SP701143	RB72454	Brasil - RIDESA
78	RB845210	<i>S. officinarum</i> híbrido	RB72454	SP701143	Brasil - RIDESA
79	RB835486	<i>S. officinarum</i> híbrido	L6014	?	Brasil - RIDESA
80	RB966928	<i>S. officinarum</i> híbrido	RB855156	RB815690	Brasil - RIDESA
81	RB855156	<i>S. officinarum</i> híbrido	RB72454	TUC717	Brasil - RIDESA
82	AJAX	<i>S. officinarum</i>	-	-	Fiji
83	BADILA	<i>S. officinarum</i>	-	-	Austrália
84	BLACKBORNEO	<i>S. officinarum</i>	-	-	Indonésia
85	BOURBONSUR	<i>S. officinarum</i>	-	-	-
86	BRAVADEPERICO	<i>S. officinarum</i>	-	-	Cuba
87	CRIOLA	<i>S. officinarum</i>	-	-	Brasil
88	CAIANALISTRADA	<i>S. officinarum</i>	-	-	Brasil
89	CAIANAROXIA	<i>S. officinarum</i>	-	-	Brasil
90	CAIANEVERDADEIRA	<i>S. officinarum</i>	-	-	Brasil
91	CANAALHO	<i>S. officinarum</i>	-	-	Brasil
92	CERAMRED	<i>S. officinarum</i>	-	-	Índia
93	GREENGERMAN	<i>S. officinarum</i>	-	-	Índia
94	IJ76418RED	<i>S. officinarum</i>	-	-	Indonésia
95	KHAJURIA	<i>S. officinarum</i>	-	-	Índia
96	MUNTOKJAVA	<i>S. officinarum</i>	-	-	Indonésia
97	RP8	<i>S. officinarum</i>	-	-	-
98	ENDOR	<i>S. officinarum</i>	Uba	28MQ674	Austrália
99	CHUNNEE	<i>S. barberi</i>	-	-	Índia
100	KASSOER	<i>S. off. x S. spo.</i>	BLACKCHERI	<i>S. spontaneum</i>	Indonésia
101	MIDAZ	<i>S. off. x S. barberi</i>	TROJAN	VESTA	-
102	NG2121	<i>S. officinarum</i>	-	-	Papua-Nova Guiné
103	PITU	<i>S. officinarum</i>	-	-	Brasil
104	GLAGAH	<i>S. spontaneum</i>	-	-	Indonésia
105	HOLES1	<i>S. spontaneum</i>	-	-	-
106	IN8488	<i>S. spontaneum</i>	-	-	Indonésia
107	IN81101	<i>S. spontaneum</i>	-	-	Indonésia
108	IS76196	<i>S. spontaneum</i>	-	-	EUA - Havaí
109	M.MOENTAI	<i>S. spontaneum</i>	-	-	Indonésia
110	NG26011	<i>S. spontaneum</i>	-	-	Papua-Nova Guiné
111	PCAV8413	<i>S. spontaneum</i>	-	-	Filipinas
112	SES260	<i>S. spontaneum</i>	-	-	Índia
113	SES365	<i>S. spontaneum</i>	-	-	Índia
114	SES196	<i>S. spontaneum</i>	-	-	Índia
115	SH301	<i>S. spontaneum</i>	-	-	Índia
116	KRAKATAU	<i>S. spontaneum</i>	-	-	Indonésia
117	US851008	<i>S. spontaneum x ?</i>	-	-	Taiwan
118	NG5712	<i>S. robustum</i>	-	-	Papua-Nova Guiné
119	IND81170	<i>S. spontaneum</i>	-	-	Índia
120	UK3739	<i>Saccharum spp.</i>	-	-	-
121	IK76006	<i>S. spontaneum</i>	-	-	EUA - Havaí
122	IN8109M2	<i>S. spontaneum</i>	-	-	Indonésia
123	IN81014	<i>S. spontaneum</i>	-	-	Indonésia
124	PCAV84-6-9	<i>S. spontaneum</i>	-	-	Filipinas
125	SES35379	<i>S. spontaneum</i>	-	-	Índia
126	SES073-38	<i>S. spontaneum</i>	-	-	Índia
127	SES234-2	<i>S. spontaneum</i>	-	-	Índia
128	SES208-1	<i>S. spontaneum</i>	-	-	Índia
129	PURPLE-32	<i>S. spontaneum</i>	-	-	-
130	COIMBATORE	<i>S. spontaneum</i>	-	-	Índia
131	GLAGAH KLOET	<i>S. spontaneum</i>	-	-	Indonésia
132	51NG26	<i>S. spontaneum</i>	-	-	Papua-Nova Guiné
133	GANDACHENI	<i>S. barberi X ?</i>	-	-	Índia
134	57NG2	<i>S. robustum</i>	-	-	Papua-Nova Guiné
135	FIJI10	<i>Miscanthus</i>	-	-	Fiji
136	FIJI15	<i>Miscanthus</i>	-	-	Fiji
137	POJ2878	<i>S. officinarum</i> híbrido	-	-	Indonésia

CTC: Centro de Tecnologia Canavieira; RIDESA: Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético; IAC: Instituto Agrônomo de Campinas.

Tabela S2: Dissimilaridades médias entre cada acesso do germoplasma básico e o grupo de cultivares.

Acesso	Dissimilaridade média	Acesso	Dissimilaridade média
MIDAZ	0.203	GANDACHENI	0.264
NG2121	0.211	IND81170	0.266
ENDOR	0.222	AJAX	0.270
BOURBONSUR	0.224	NG26011	0.272
BADILA	0.226	IJ76418RED	0.273
IN8488	0.238	MUNTOKJAVA	0.274
CANAALHO	0.242	NG5712	0.276
IN81101	0.242	SES260	0.278
BLACKBORNEO	0.243	UK3739	0.278
GREENGERMAN	0.243	57NG2	0.281
KRAKATAU	0.245	KASSOER	0.281
IN81014	0.246	PCAV8413	0.284
CAIANALISTRADA	0.247	IK76006	0.286
CERAMRED	0.247	GLAGAH	0.286
PITU	0.249	FIJI15	0.287
CREOULA	0.250	SES196	0.288
SES365	0.251	51NG26	0.290
CAIANAROXIA	0.251	SES2081	0.290
KHAJURIA	0.251	PCAV8469	0.291
US85108	0.252	FIJI10	0.291
BRAVADEPERICO	0.252	SES35379	0.293
CAIANEVERDADEIRA	0.257	M.MOENTAI	0.294
POJ2878	0.259	SES2342	0.297
HOLES1	0.259	GLAGAHKLOET	0.298
COIMBATORE	0.260	IN8109M2	0.302
IS76196	0.260	PURPLE32	0.302
RP8	0.260	SES07338	0.313
SH301	0.263	CHUNNEE	0.326

Figura 1: Curva do Coeficiente de Variação (CV%) da similaridade (Jaccard) de acordo com os diferentes números de marcadores considerados.

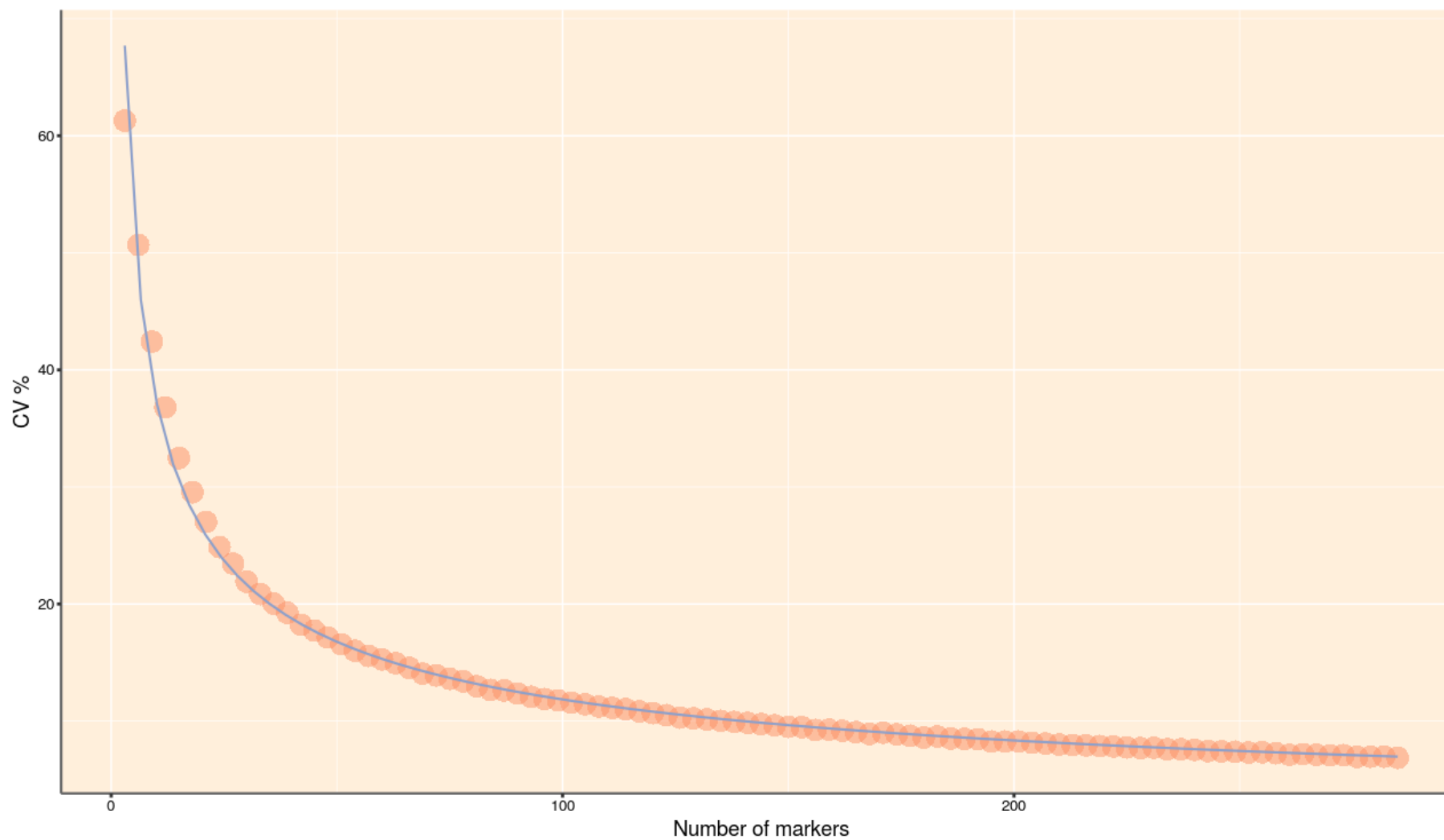


Figura 2: Distribuição gráfica dos alelos (%) dentro dos pares de iniciadores SCB213, SCB312, SCB381 e SCB436.

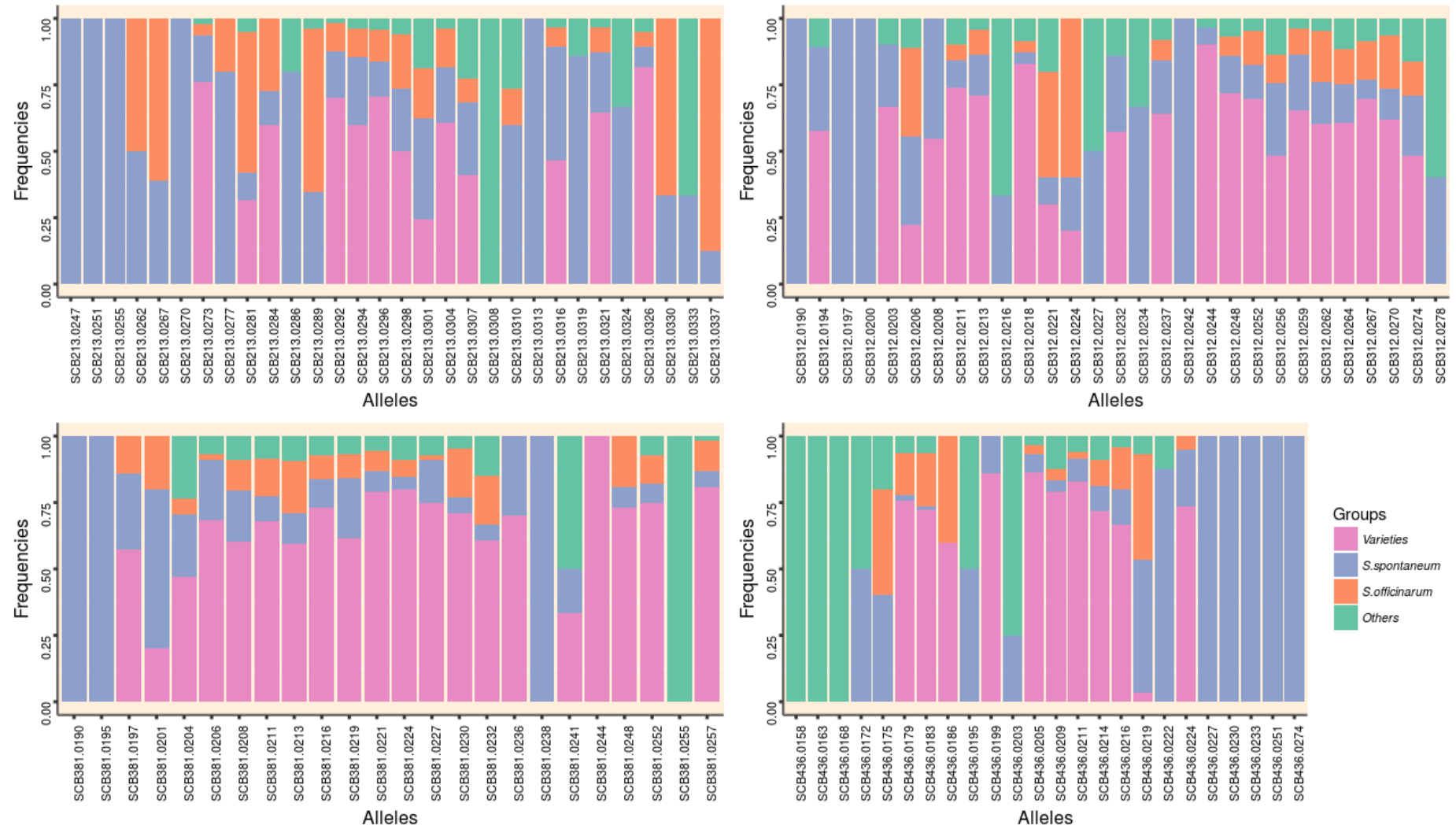


Figura 3: Distribuição gráfica dos alelos (%) dentro dos pares de iniciadores SCC423, CV29, CV37 and CV38.

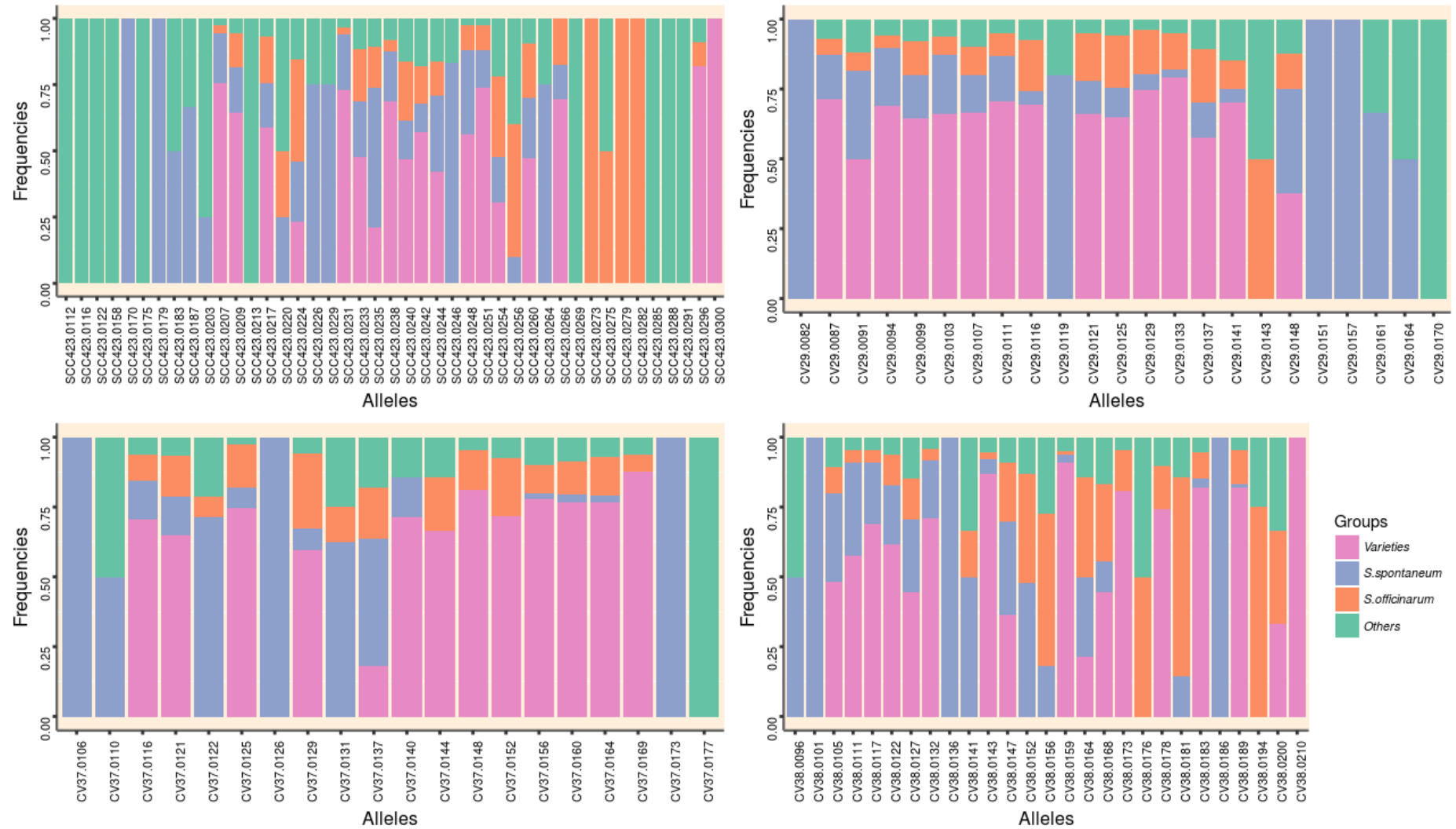


Figura 4: Distribuição gráfica dos alelos (%) dentro dos pares de iniciadores CV60, CV79, CV94 and CV106.

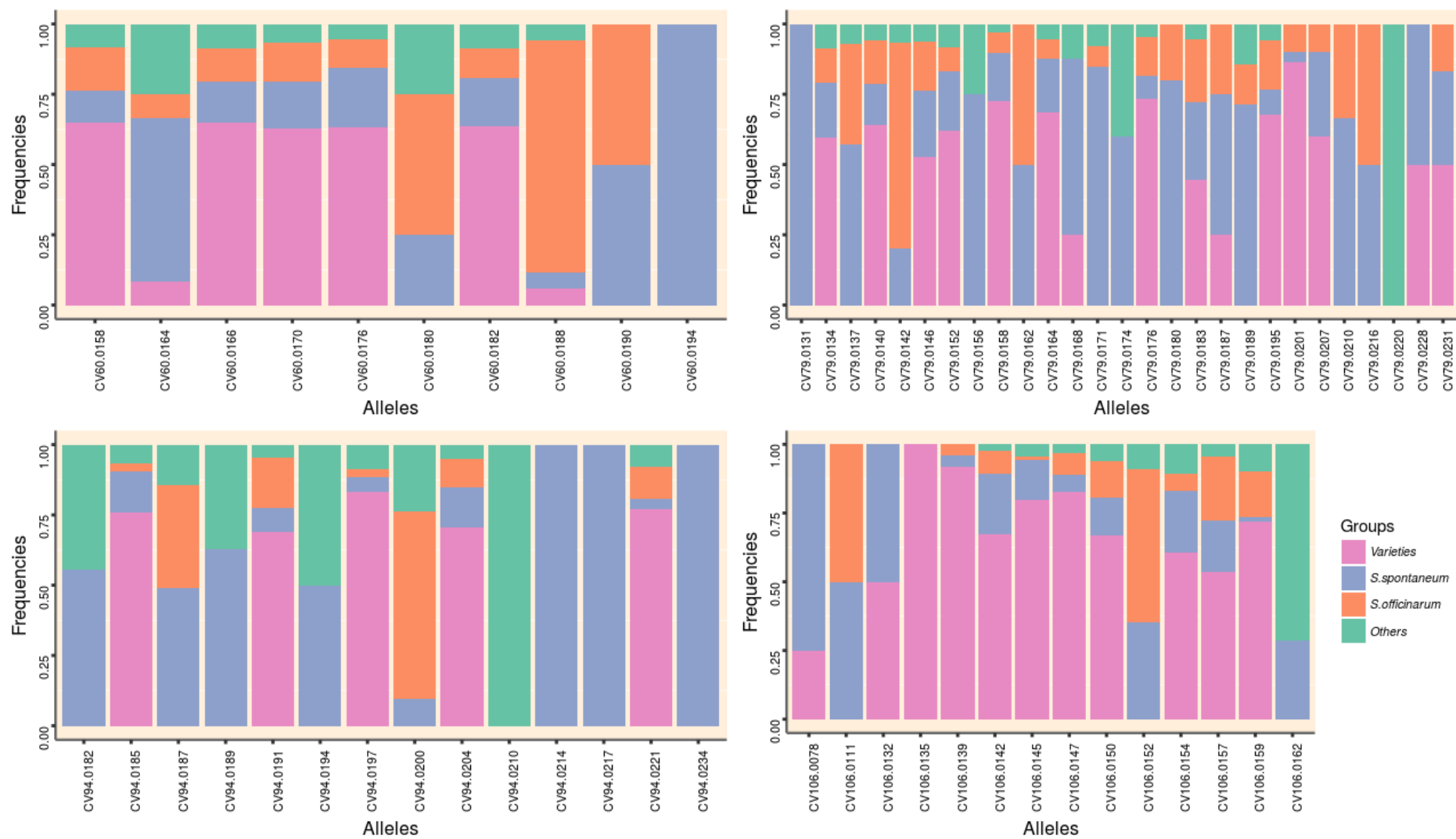


Figura 4: Análise de Coordenadas Principais dos acessos de cana-de-açúcar com base na matriz de dissimilaridade (1-Jaccard) obtida a partir de dados de SSR e calculada entre pares de genótipos. Na legenda, "Varieties" significa variedades (ou cultivares) e "Others" significa "outras espécies e híbridos exóticos".

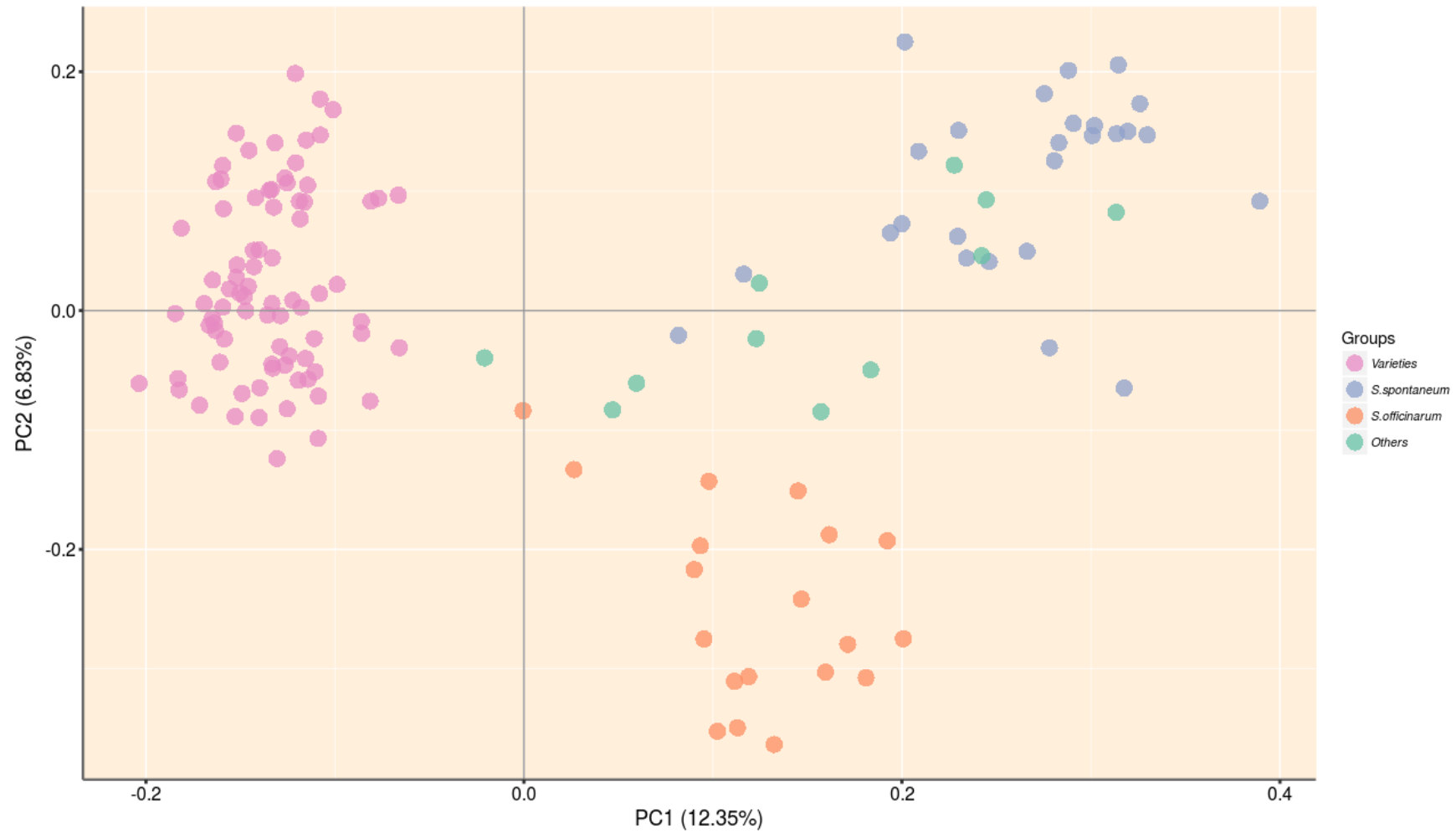


Figura 6: (A) Análise do melhor valor de K realizada pelo software CLUMPAK [60] mostrando valores de K de 2 a 9 (10 não mostrado); (B) proporção de identidade para cada genótipo feito pelo software Structure V.2.3.4: dois grupos distintos e quase homogêneos estão presentes dentro da população distinguindo indivíduos de 1 a 81 (cultivares) e 82 a 137 (germoplasma básico) para K = 2 ; (C) três grupos distintos, contrastando os primeiros 81 genótipos como um e o restante diferenciando em dois outros, compreendendo as principais espécies *S. officinarum* e *S. spontaneum*.

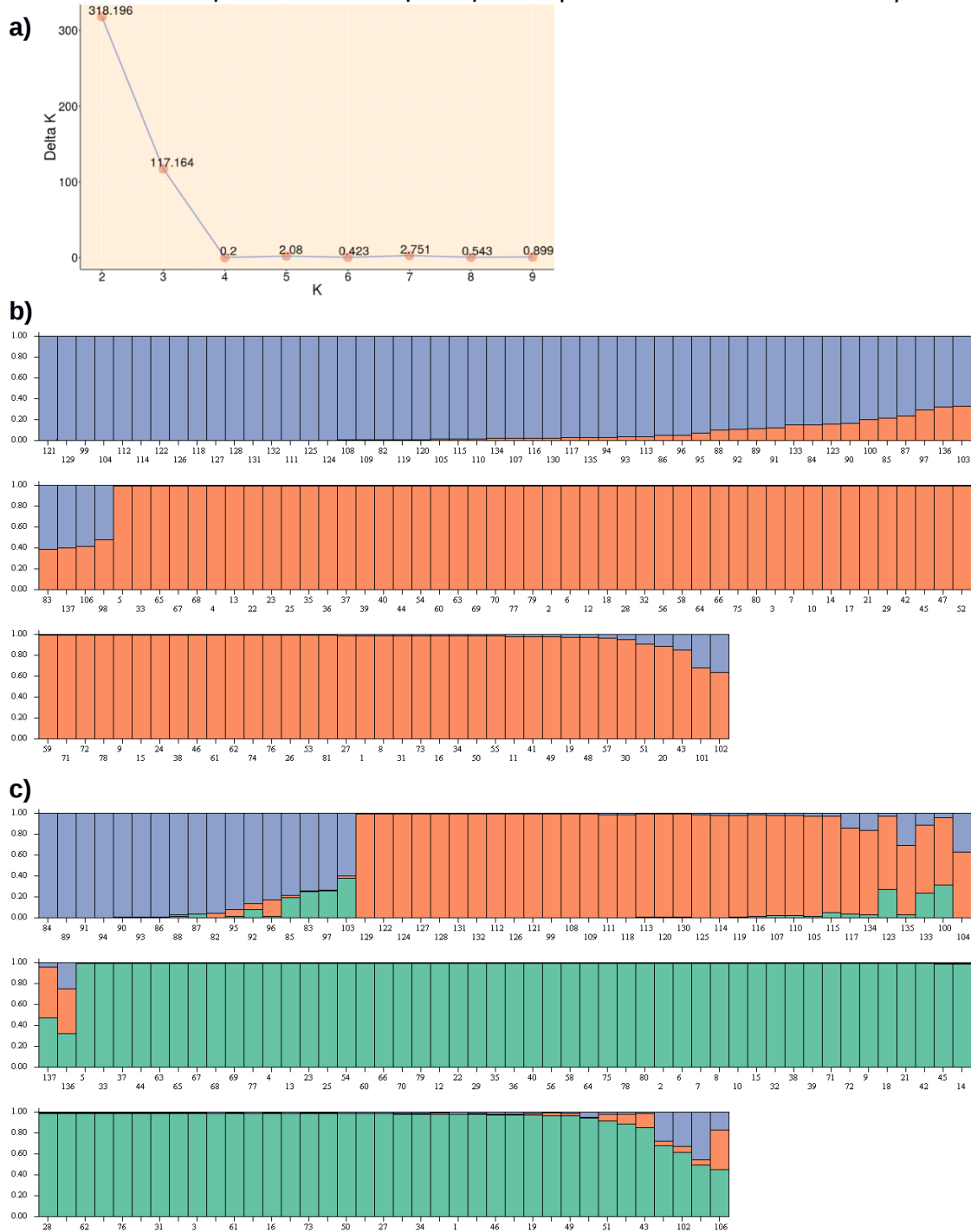
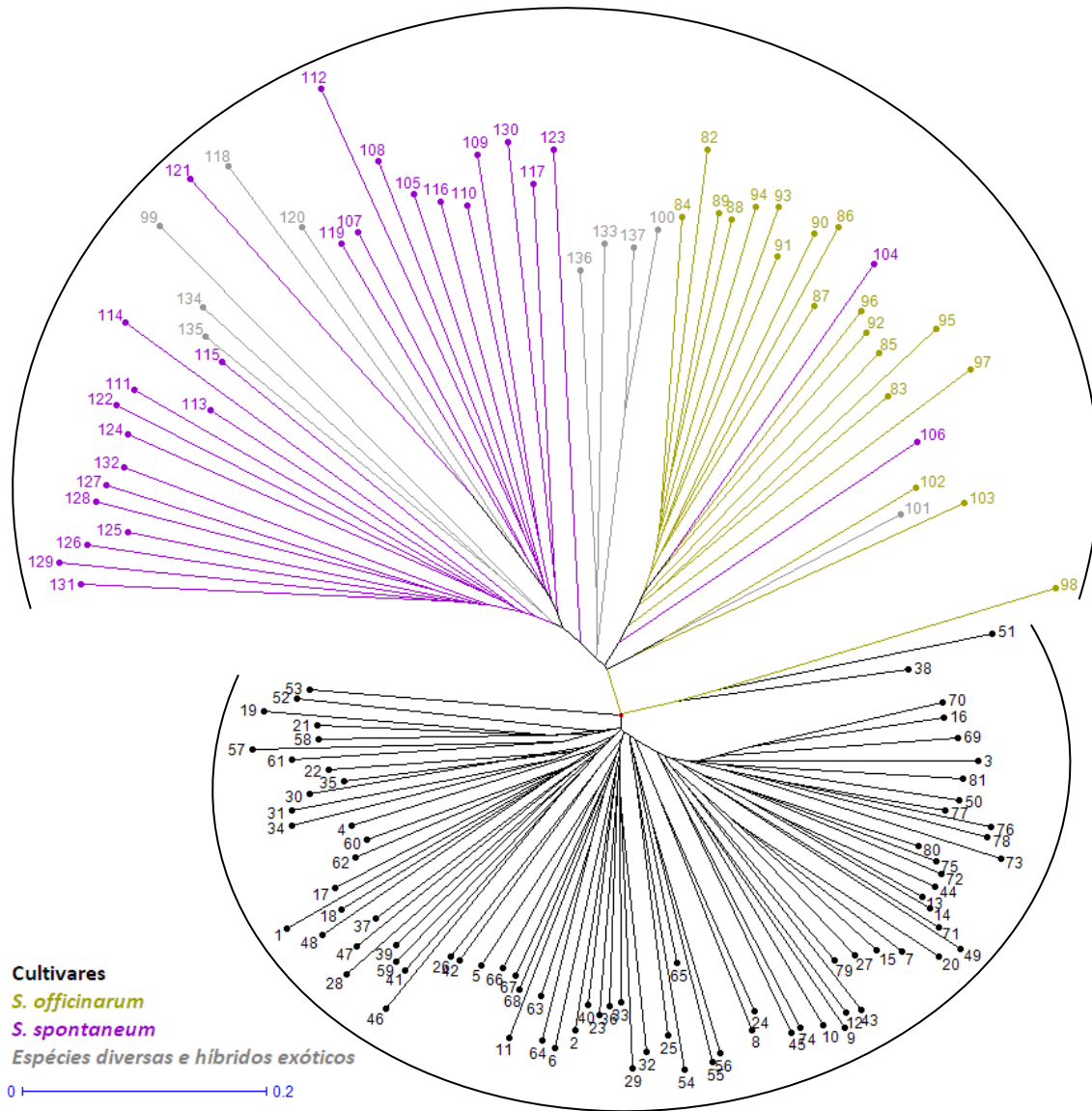


Figura 7: Análise de agrupamento realizada pelo software DARwin v.6 [45] revela dois grandes agrupamentos distintos: no topo da imagem, os acessos são agrupados de acordo com as suas principais espécies; na parte inferior da imagem, as cultivares agrupam-se em um grupo bem definido. Os indivíduos de espécies diversas e os híbridos exóticos estão espalhados dentro deste segundo grupo de acordo com suas genealogias (Método: Unweighted-Neighbor-Joining).



APÊNDICE

Tabela 1A: Listagem dos 309 híbridos de *Saccharum officinarum*. Sublinhados estão as 81 cultivares brasileiras. Os materiais com as siglas IAC, SP, RB e CTC são brasileiros. A cultivar POJ2878 (ID 228) é originária da Indonésia. O material “àUS571415” (ID 97) tem sua origem indeterminada.

ID	Indivíduo	ID	Indivíduo	ID	Indivíduo	ID	Indivíduo	ID	Indivíduo	ID	Indivíduo	ID	Indivíduo
1	IACSP991306	46	IACCTC063709	91	IACSP015512	136	IACSP952289	181	IACSP022033	226	IACSP993090	271	CTC17
2	IACCTC052513	47	IACCTC053585	92	IACSP012434	137	IACSP993010	182	IACSP023062	227	IACSP956114	272	IAC873396
3	IACSP991015	48	IACCTC053578	93	IACSP012433	138	IACSP022115	183	IACSP991308	228	POJ2878	273	IAC911099
4	IACCTC053550	49	IACCTC053612	94	IACSP016407	139	IACSP032058	184	IACSP022022	229	SP87396	274	IACSP944004
5	IACSP994010	50	IACCTC062757	95	IAC912205	140	IACSP993099	185	IACSP042065	230	SP842025	275	IACSP942101
6	IACCTC052508	51	IACCTC063746	96	IACSP983045	141	IACSP993091	186	IACSP994007	231	SP903414	276	IACSP942094
7	IACSP991418	52	IACSP977018	97	àUS571415*	142	IACSP012411	187	IACSP992014	232	SP801816	277	IAC862480
8	IACCTC053514	53	IACCTC063728	98	IACSP963130	143	IACSP993020	188	IACSP023157	233	SP901638	278	IACSP953028
9	IACCTC053511	54	IACCTC063738	99	IACSP012422	144	IACSP993427	189	IACSP017089	234	SP841201	279	IACSP955000
10	IACCTC052534	55	IACSP022008	100	IACSP983035	145	IACSP993100	190	IACSP991301	235	SP801842	280	IACSP933046
11	IACSP992239	56	IACCTC063741	101	IACSP012410	146	IACSP985024	191	IACSP994011	236	SP813250	281	RB935744
12	IACSP022067	57	IACSP017047	102	IACSP017016	147	IACSP993073	192	IACSP023098	237	SP87365	282	RB825336
13	IACCTC052505	58	IACSP018152	103	IACSP017012	148	IACSP011038	193	IACSP023029	238	SP855077	283	RB925211
14	IACCTC053546	59	IACCTC063743	104	IACSP017021	149	IACSP993085	194	IACSP987061	239	SP8642	284	RB925345
15	IACCTC053547	60	IACCTC063749	105	IACSP973353	150	IACSP032086	195	IACSP042002	240	SP775181	285	RB925268
16	IACSP042515	61	IACSP974048	106	IACSP972109	151	IACSP022128	196	IACSP994018	241	SP87344	286	RB931530
17	IACSP018027	62	IACSP974039	107	IACSP983053	152	IACSP993345	197	IACSP022004	242	SP845560	287	RB928064
18	IACSP018000	63	IACSP018046	108	IACSP012409	153	IACSP993321	198	IACSP022123	243	SP803280	288	RB93509
19	IACSP018034	64	IACSP996308	109	IACSP007320	154	IACSP993385	199	IACSP023134	244	SP811763	289	RB835054
20	IACSP018158	65	IACSP982025	110	IACSP017019	155	IACSP993018	200	IACSP022141	245	SP841431	290	RB855035
21	IACSP991018	66	IACSP977545	111	IACSP983044	156	IACSP032071	201	IACSP032105	246	SP803480	291	RB855536
22	IACCTC053543	67	IACSP022032	112	IACSP012408	157	IACSP032035	202	IACSP042015	247	SP832847	292	RB845197
23	IACSP018071	68	IACSP977532	113	IACSP013156	158	IACSP022139	203	IACSP013017	248	SP801836	293	RB842021
24	IACSP991305	69	IACSP993334	114	IACSP962041	159	IACSP012419	204	IACSP042401	249	SP891115	294	RB92579
25	IACSP022018	70	IACSP974047	115	IAC914168	160	IACSP993444	205	IACSP032107	250	SP86155	295	RB867515
26	IACSP023168	71	IACSP993390	116	IACSP983038	161	IACSP993088	206	IACSP023219	251	SP853877	296	RB865230
27	IACSP977608	72	IACSP993098	117	IACSP973307	162	IACSP993378	207	IACSP042289	252	SP911049	297	RB855453
28	IACCTC063747	73	IACSP993331	118	IACSP976634	163	IACSP042054	208	IACSP042081	253	SP835073	298	RB845257
29	IACSP022095	74	IACSP983104	119	IACSP951017	164	IACSP023183	209	IACSP022006	254	SP80185	299	RB855563
30	IACCTC052511	75	IACSP032056	120	IACSP952237	165	IACSP993145	210	IACSP994017	255	SP801520	300	RB855113
31	IACSP992207	76	IACSP993060	121	IACSP014414	166	IACSP042141	211	IACSP022014	256	CTC01	301	RB855463
32	IACCTC053557	77	IACSP993397	122	IACSP013046	167	IACSP012401	212	IACSP997251	257	CTC02	302	RB813804
33	IACCTC053552	78	IACSP032082	123	IACSP008045	168	IACSP012403	213	IACSP991419	258	CTC03	303	RB855036
34	IACCTC053522	79	IACSP993359	124	IACSP976645	169	IACSP023045	214	IACSP023175	259	CTC04	304	RB855511
35	IACSP992110	80	IACSP015504	125	IACSP963069	170	IACSP032092	215	IACSP987043	260	CTC05	305	RB855546
36	IACCTC063726	81	IACSP015501	126	IACSP966046	171	IACSP022030	216	IACSP022065	261	CTC06	306	RB845210
37	IACCTC053516	82	IACSP015514	127	IACSP983031	172	IACSP023204	217	IACSP982072	262	CTC07	307	RB835486
38	IACSP991020	83	IACSP013127	128	IACSP042286	173	IACSP997229	218	IACSP976610	263	CTC08	308	RB969228
39	IACCTC053549	84	IACSP017036	129	IAC913093	174	IACSP022020	219	IACSP993013	264	CTC09	309	RB855156
40	IACCTC063732	85	IACSP012428	130	IACSP941061	175	IACSP032030	220	IACSP973331	265	CTC10		
41	IACCTC063755	86	IACSP013106	131	IACSP962067	176	IACSP023093	221	IACSP983099	266	CTC11		
42	IACCTC053567	87	IACSP013139	132	IACSP042297	177	IACSP022118	222	IACSP993039	267	CTC12		
43	IACSP018002	88	IACSP012421	133	IACSP013185	178	IACSP022019	223	IACSP993009	268	CTC13		
44	IACCTC062758	89	IACSP993425	134	IACSP983055	179	IACSP987019	224	IACSP976615	269	CTC14		
45	IACSP022074	90	IACSP014406	135	IACSP001052	180	IACSP997236	225	IACSP985010	270	CTC15		

*indivíduo nº 97 “àUS571415” consta na listagem de materiais importados como um híbrido de *Saccharum officinarum*, possivelmente trocado ou mal etiquetado.

Tabela 2A: Listagem dos 52 acessos de *Saccharum officinarum*.

ID	Indivíduo	Origem	ID	Indivíduo	Origem
310	AJAX	Fiji (8)	336	IN84126	Indonésia (11)
311	BADILA	Austrália (1)	337	IN8446	Indonésia (11)
312	BLACKBORNEO	Indonésia (11)	338	IN84105	Indonésia (11)
313	BOURBONSUR	-	339	IS76116	EUA – Havaí (7)
314	BRAVADEPERICO	Cuba (5)	340	IS76155	EUA – Havaí (7)
315	CREOULA	Brasil (2)	341	KHAJURIA	Índia (10)
316	CAIANALISTRADA	Brasil (2)	342	LAM76427	-
317	CAIANARISCADA	Brasil (2)	343	LOUSIANA STRIPED	-
318	CAIANAROXIA	Brasil (2)	344	MANTEIGA	Taiwan (19)
319	CAIANEVERDADEIRA	Brasil (2)	345	MUNTOKJAVA	Indonésia (11)
320	CANAALHO	Brasil (2)	346	MZ151	-
321	CANABLANCA	Cuba (5)	347	NG57213	Papua-N. Guiné (15)
322	CERAMRED	Índia (10)	348	NG7718	Papua-N. Guiné (15)
323	FIJI19	Fiji (8)	349	NG7792	Papua-N. Guiné (15)
324	FIJI62	Fiji (8)	350	NG271124	Papua-N. Guiné (15)
325	FLORDECUBA	Cuba (5)	351	RP8	-
326	GREENGERMAN	Índia (10)	352	SABURA	-
327	IJ76418RED	Indonésia (11)	353	S.OFF.8272	Malásia (13)
328	IJ76313	Indonésia (11)	354	W.TRANSPARENT	Índia (10)
329	IJ76317	Indonésia (11)	355	ZOPILOTA	Costa Rica (4)
330	IJ76-315	Indonésia (11)	356	ZWARTMANILLA	Indonésia (11)
331	IJ76-319	Indonésia (11)	357	NG2121	Papua-N. Guiné (15)
332	IJ76361	Indonésia (11)	358	PITU	Brasil (2)
333	IJ76442	Indonésia (11)	359	IN8488(1)*	Indonésia (11)
334	IJ76470	Indonésia (11)	360	IN845	Indonésia (11)
335	IJ76566	Indonésia (11)	361	SAC.OFF.8280	Malásia (13)

*indivíduo nº 359 "IN8488(1)" é um material em duplicidade e consta como um acesso de *S. officinarum*, possivelmente trocado.

Tabela 3A: Listagem dos 149 acessos de *Saccharum spontaneum*.

ID	Indivíduo	Origem	ID	Indivíduo	Origem	ID	Indivíduo	Origem
362	DACCA	Índia (10)	412	SES295	Índia (10)	462	IS76177	EUA – Havaí (7)
363	DJANTOER	Indonésia (11)	413	SES297A	Índia (10)	463	IJ76386	Indonésia (11)
364	FORMOSA	Taiwan (19)	414	SES308A	Índia (10)	464	IJ76480	Indonésia (11)
365	GEHRABON	Índia (10)	415	SES323	Índia (10)	465	IS761715	EUA – Havaí (7)
366	GLAGAH	Indonésia (11)	416	SES365	Índia (10)	466	IS76201	EUA – Havaí (7)
367	HOLES1	-	417	SES367	Índia (10)	467	MOL4503	EUA – Havaí (7)
368	IK76049	EUA – Havaí (7)	418	SES519	Índia (10)	468	NG21035	Papua-N. Guiné (15)
369	IN8482	Indonésia (11)	419	SES602	Índia (10)	469	US7851939	Paquistão (16)
370	IN8488(2)*	Indonésia (11)	420	SES205A	Índia (10)	470	SSPONT15	-
371	IN84009	Indonésia (11)	421	SES288(2)*	Índia (10)	471	SSPONT37	-
372	IN84011	Indonésia (11)	422	SES196	Índia (10)	472	SSPONT4593	-
373	IN84021	Indonésia (11)	423	SH301	Índia (10)	473	PAGN84226	-
374	IN84022	Indonésia (11)	424	SM7916	-	474	PBG84523	Filipinas (9)
375	IN84089	Indonésia (11)	425	KRAKATAU	Indonésia (11)	475	PBG84525	Filipinas (9)
376	IN81101	Indonésia (11)	426	TAIWAN2N=96	Taiwan (19)	476	PPGN84813	Filipinas (9)
377	IS76125	EUA – Havaí (7)	427	COIMBATORE	Índia (10)	477	PGN84814	Filipinas (9)
378	IS76173	EUA – Havaí (7)	428	TAIWANSPONT64	Taiwan (19)	478	PBG841216	Filipinas (9)
379	IS76192	EUA – Havaí (7)	429	TAIWANSPONT72	Taiwan (19)	479	PCAV8469	Filipinas (9)
380	IS76196	EUA – Havaí (7)	430	THA83/71	Tailândia (18)	480	PCAV84630	Filipinas (9)
381	M.MOENTAI	Indonésia (11)	431	UM69015	Malásia (13)	481	PCANOR8436	Filipinas (9)
382	NG26011	Papua-N. Guiné (15)	432	US4633	EUA – Flórida (6)	482	PCAV891811	Filipinas (9)
383	NG77176	Papua-N. Guiné (15)	433	US5601601	Tailândia (18)	483	PCAV841812	Filipinas (9)
384	PAL8401	Filipinas (9)	434	US78518	Paquistão (16)	484	PCANOR84127	Filipinas (9)
385	PCASUR8401	Filipinas (9)	435	US78519	Paquistão (16)	485	PCANOR8902	Filipinas (9)
386	PCASUR8403	Filipinas (9)	436	US78523	Paquistão (16)	486	PANAMA	-
387	PCASUR8408	Filipinas (9)	437	YACHENG12	China (3)	487	SES35379	Índia (10)
388	PCASUR8413	Filipinas (9)	438	IN76086	Indonésia (11)	488	SES26920	Índia (10)
389	PCAV8413	Filipinas (9)	439	US721319(1)*	Taiwan (19)	489	SES26431	Índia (10)
390	PCAV8420	Filipinas (9)	440	US721319(2)*	Taiwan (19)	490	SES07338	Índia (10)
391	PHILLIPINEWILD	Filipinas (9)	441	US851008(1)*	Taiwan (19)	491	SES2342	Índia (10)
392	PONAPPEWILD	Japão (12)	442	US851008(2)*	Taiwan (19)	492	SES845824	Índia (10)
393	PPGN8408	Filipinas (9)	443	US85108**	-	493	SES35377	Índia (10)
394	PQ8404	Filipinas (9)	444	IK76006	EUA – Havaí (7)	494	SES07322	Índia (10)
395	PQ8409	Filipinas (9)	445	TONGZA	China (3)	495	SES2081	Índia (10)
396	PSOR8413	-	446	IN8109M2	Indonésia (11)	496	SES27710	Índia (10)
397	S66084	Taiwan (19)	447	IN81014	Indonésia (11)	497	SES2773	Índia (10)
398	S66097	Taiwan (19)	448	IN8401634	Indonésia (11)	498	TAIWAN1140	Taiwan (19)

Tabela 3A: Continuação

ID	Indivíduo	Origem	ID	Indivíduo	Origem	ID	Indivíduo	Origem
399	S66105	Taiwan (19)	449	IN84227	Indonésia (11)	499	THA82060	Tailândia (18)
400	SSPONT28	Taiwan (19)	450	IN8403336	Indonésia (11)	500	KAWANDA	Uganda (20)
401	SES033	Índia (10)	451	IN8404221	Indonésia (11)	501	KLOET	Indonésia (11)
402	SES084/58	Índia (10)	452	IN8404237	Indonésia (11)	502	PURPLE32	-
403	SES088C	Índia (10)	453	IN84088	Indonésia (11)	503	GLAGAHKLOET	Indonésia (11)
404	SES090	Índia (10)	454	IN840918	Indonésia (11)	504	CANE604417	-
405	SES092	Índia (10)	455	IND813819	Índia (10)	505	CANE604418	-
406	SES103	Índia (10)	456	IND818235	Índia (10)	506	51NG26	Papua-N. Guiné (15)
407	SES113A	Índia (10)	457	IND82257	Índia (10)	507	78217273	-
408	SES184B	Índia (10)	458	IND823184	Índia (10)	508	FORMOSA429	Taiwan (19)
409	SES184A	Índia (10)	459	IM76234	Austrália (1)	509	BROWN33	-
410	SES208	Índia (10)	460	IK7666	EUA – Havaí (7)	510	B68264	-
411	SES234	Índia (10)	461	IJ76111	Indonésia (11)			

*indivíduos com duplicidades;

**indivíduo com possível erro no nome e outra duplicidade do indivíduo "US851008" com grafia errada;

Tabela 4A: Listagem do grupo de 28 acessos com genealogias ou espécies desconhecidas.

ID	Indivíduo	Origem	ID	Indivíduo	Origem
511	IK7606	EUA – Havaí (7)	525	IA3135	-
512	IN8485	Índia (10)	526	NG57222	Papua-N. Guiné (15)
513	IND81170*	Indonésia (11)	527	IJ3852218	Indonésia (11)
514	IND81170A*	Indonésia (11)	528	SM8116	-
515	IND81170B*	Indonésia (11)	529	HJ5741	-
516	LONGCHUAN	China (3)	530	CINCA77381	-
517	PIN8401	Filipinas (9)	531	NG21021	Papua-N. Guiné (15)
518	SLC9232	Sri Lanka (17)	532	NG517221	Papua-N. Guiné (15)
519	SLC9262	Sri Lanka (17)	533	MUNDAU	-
520	SLC9277	Sri Lanka (17)	534	MIDAZ	-
521	SLC9285	Sri Lanka (17)	535	SAC.OFF.828	-
522	SLC9234	Sri Lanka (17)	536	SAC.OFF.8296	-
523	SLC9298	Sri Lanka (17)	537	YASAWA	Fiji (8)
524	P571504	Cuba (5)	538	XMISCANTHUS	-

*indivíduos com duplicidades;

Tabela 5A: Listagem do grupo de 55 acessos catalogados como diversos. Os acessos de espécies diversas aparecem sublinhados.

ID	Genealogia	Indivíduo	Origem	ID	Genealogia	Indivíduo	Origem
539	<i>S. officinarum</i> X ?	IJ76325	Indonésia (11)	567	<i>E. arudinaceus</i>	<u>IJ76364</u>	Indonésia (11)
540	<i>S. officinarum</i> X ?	IJ76560	Indonésia (11)	568	<i>E. arudinaceus</i>	<u>IJ76359(2)*</u>	Indonésia (11)
541	<i>S. officinarum</i> híbrido X ?	NG576	Papua-N. Guiné (15)	569	<i>E. arudinaceus</i>	<u>IJ76358</u>	Indonésia (11)
542	<i>S. officinarum</i> híbrido X ?	NG5750	Papua-N. Guiné (15)	570	<i>Erianthus</i>	<u>75II09(1)*</u>	-
543	<i>S. officinarum</i> híbrido X ?	IN84106	Indonésia (11)	571	<i>Erianthus</i>	<u>75II09(2)*</u>	-
544	<i>Saccharum</i> X <i>Miscanthus</i>	SM8136	-	572	<i>Erianthus</i>	<u>75II09(3)*</u>	-
545	<i>B. cheribon</i> x <i>Batjan</i>	D152	-	573	<i>Miscanthus</i>	<u>FIJ110</u>	Fiji (8)
546	405N1133 X 33MQ371	MALI	Mali (14)	574	<i>Miscanthus</i>	<u>FIJ115</u>	Fiji (8)
547	Uba X 28MQ674	ENDOR	Austrália (1)	575	<i>S. arudinaceum</i>	<u>IJ76410</u>	Indonésia (11)
548	<i>Saccharum</i> híbrido	NCO310	Índia (10)	576	<i>S. barberi</i>	<u>CHUNNEE</u>	Índia (10)
549	<i>Saccharum</i> híbrido	RAGNAR	EUA – Havaí (7)	577	<i>S. barberi</i>	<u>MATNASHAHJ</u>	Índia (10)
550	<i>Saccharum</i> híbrido	KASSOER	Indonésia (11)	578	<i>S. robustum</i>	<u>57NG2</u>	Papua-N. Guiné (15)
551	Híbrido	US74069	-	579	<i>S. robustum</i>	<u>IM76232(2)*</u>	Austrália (1)
552	Híbrido	ZEUS	-	580	<i>S. robustum</i>	<u>IM76229(1)*</u>	Austrália (1)
553	<i>S. spontaneum</i> X ?	IN8458	Indonésia (11)	581	<i>S. robustum</i>	<u>IM76229(2)*</u>	Austrália (1)
554	<i>S. robustum</i> x ?	NG5712	Papua-N. Guiné (15)	582	<i>S. robustum</i>	<u>IJ76293(1)*</u>	Indonésia (11)
555	<i>S. barberi</i> X ?	CHIN	Índia (10)	583	<i>S. robustum</i>	<u>IM76232(1)*</u>	Austrália (1)
556	<i>S. barberi</i> X ?	GANDACHENI	Índia (10)	584	<i>S. robustum</i>	<u>IN76229</u>	Indonésia (11)
557	<i>Saccharum</i> spp.	UK2215	-	585	<i>S. robustum</i>	<u>IJ76293(2)*</u>	Indonésia (11)
558	<i>Saccharum</i> spp.	UK3734	-	586	<i>S. robustum</i>	<u>IJ76293(3)*</u>	Indonésia (11)
559	<i>Saccharum</i> spp.	UK3739	-	587	<i>S. robustum</i>	<u>US571415(1)*</u>	-
560	<i>Saccharum</i> spp.	UK3748	-	588	<i>S. robustum</i>	<u>US571415(2)*</u>	-
561	<i>Saccharum</i> spp.	UK3759	-	589	<i>S. robustum</i>	<u>US571415(3)*</u>	-
562	<i>Saccharum</i> spp.	UK3765	-	590	<i>S. robustum</i>	<u>US571415(4)*</u>	-
563	<i>Saccharum</i> spp.	UK6913	-	591	<i>S. sinense</i>	<u>UBADEMERARA</u>	-
564	<i>E. arudinaceus</i>	<u>IJ76359(1)*</u>	Indonésia (11)	592	<i>S. sinense</i>	<u>UBADENATAL</u>	Cuba (5)
565	<i>E. arudinaceus</i>	<u>IJ76389</u>	Indonésia (11)	593	<i>S. sinense</i>	<u>CAYANA</u>	-
566	<i>E. arudinaceus</i>	<u>IJ76381</u>	Indonésia (11)				

*indivíduos com duplicidades.



Figura 1A: Mapa indicando todos os locais de coleta dos 593 genótipos confirmados em consulta ao GRIN (<https://www.ars-grin.gov>). Ao todo, as origens dos genótipos se distribuem entre 20 localidades de 19 países: Austrália (1), Brasil (2), China (3), Costa Rica (4), Cuba (5), EUA - Flórida(6) e Havaí (7), Fiji (8), Filipinas (9), Índia (10), Indonésia (11), Japão (12), Malásia (13), Mali (14), Papua-Nova Guiné (15), Paquistão (16), Sri Lanka (17), Tailândia (18), Taiwan (19) e Uganda (20).

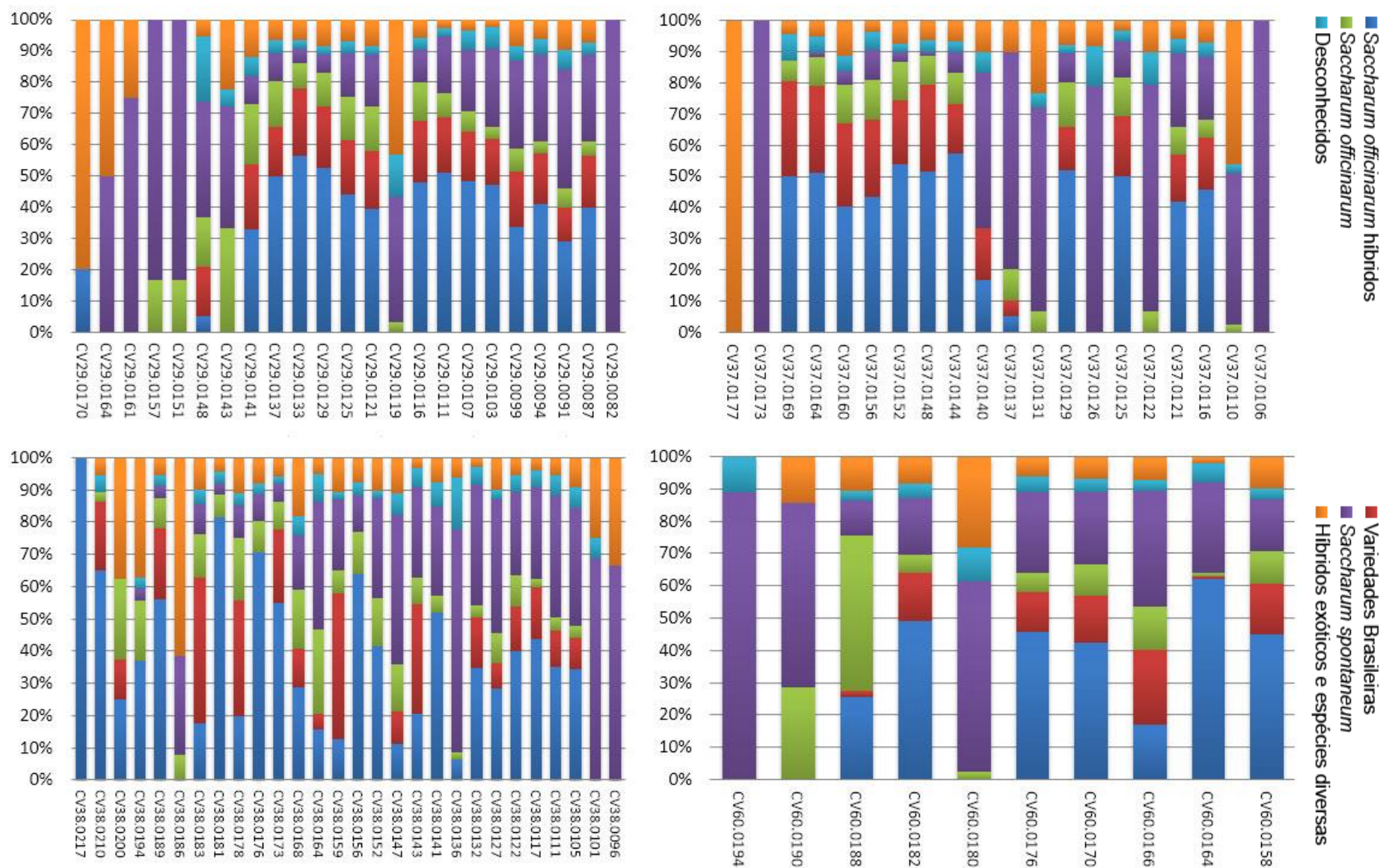


Figura 2A: Frequências alélicas para os primers CV29, CV37, CV38 e CV60.

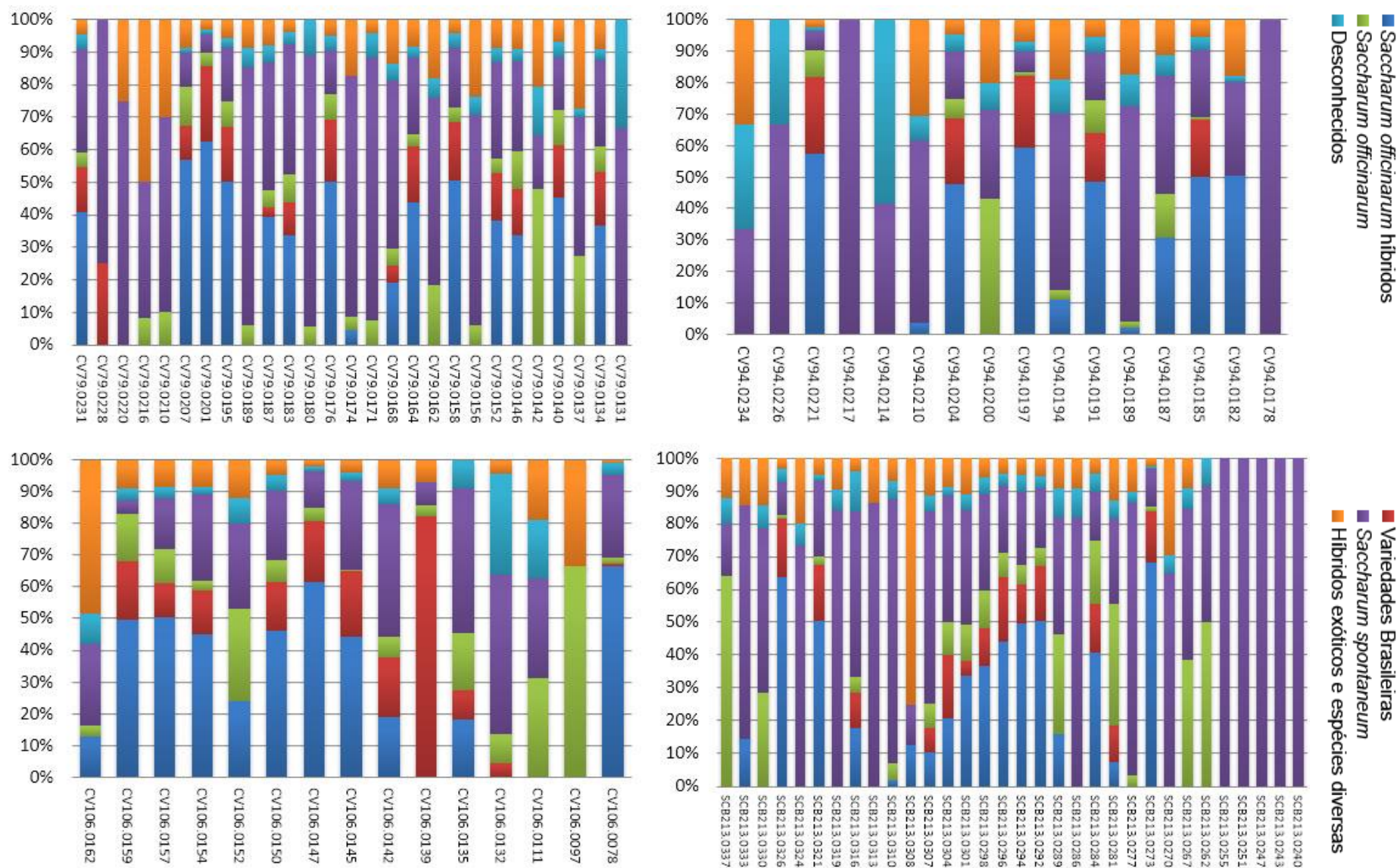


Figura 3A: Frequências alélicas para os primers CV79, CV94, CV106 e SCB213

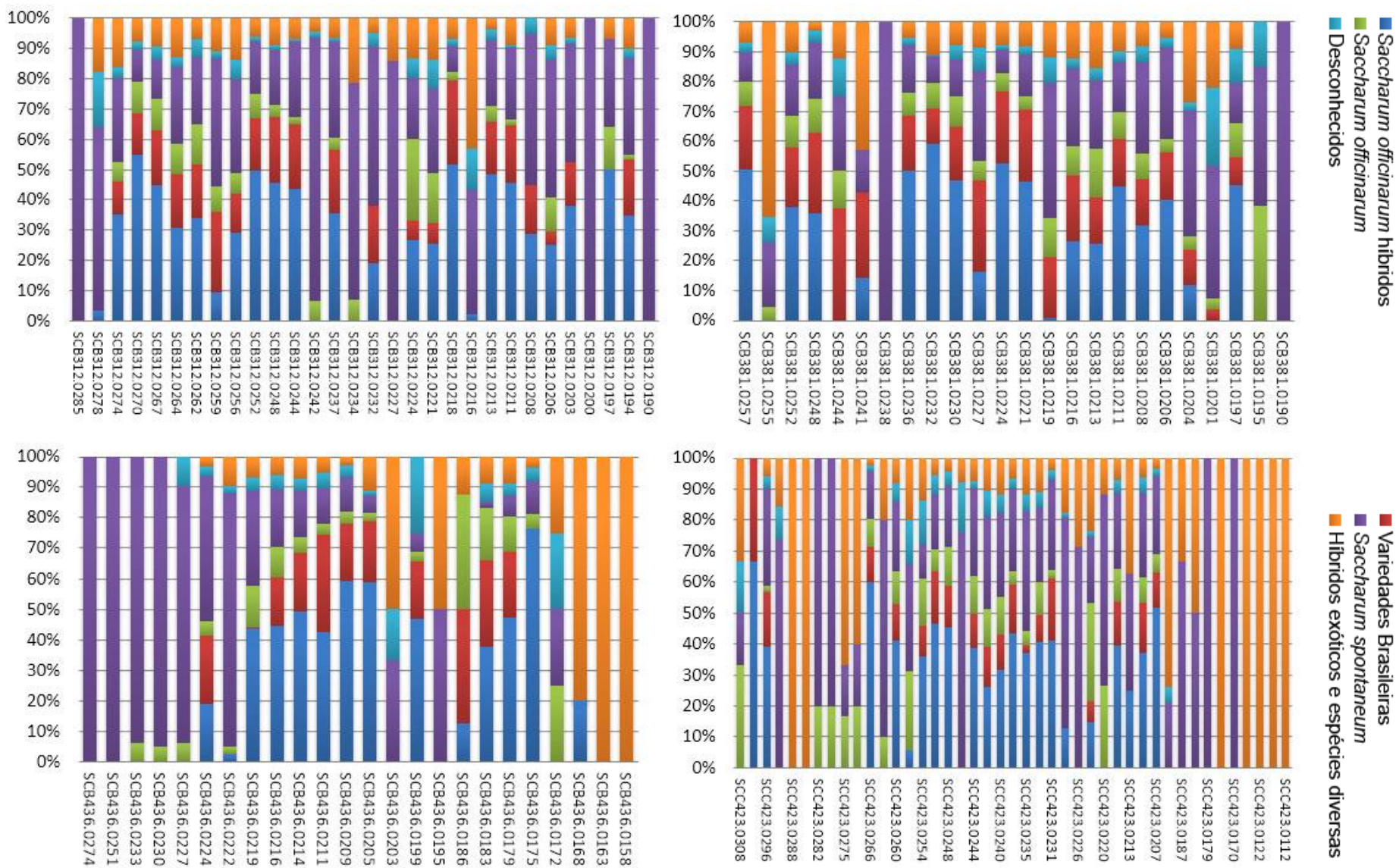


Figura 4A: Frequências alélicas para os primers SCB312, SCB381, SCB436 e SCC423.