

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA



HAMILTON JORDÃO JUNIOR

**“DESENVOLVIMENTO DE UM SISTEMA BASEADO EM
MARCADORES MOLECULARES DE DNA DO TIPO
MICROSSATÉLITES PARA IDENTIFICAÇÃO DE
VARIEDADES DE CANA-DE-AÇÚCAR”**

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
HAMILTON JORDÃO JUNIOR
e aprovada pela Comissão Julgadora.

Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas (UNICAMP) como parte dos requisitos exigidos à obtenção do título de Mestre em Biologia Funcional e Molecular na área de Bioquímica.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira

Campinas, SP

2009

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

J767d

Jordão Junior, Hamilton

Desenvolvimento de um sistema baseado em marcadores moleculares de DNA do tipo de microssatélites para identificação de acessos de cana-de-açúcar (*Saccharum* spp.) / Hamilton Jordão Junior. – Campinas, SP: [s.n.], 2009.

Orientador: Gonçalo Amarante Guimarães Pereira.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Cana-de-açúcar. 2. Marcadores genéticos. 3. DNA. 4. Microssatélites (Genética). 5. Melhoramento genético. I. Pereira, Gonçalo Amarante Guimarães Pereira. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Título em inglês: Development of a system based on DNA microsatellite molecular markers for identifying sugarcane accessions (*Saccharum* spp.).

Palavras-chave em inglês: Sugarcane; Genetic markers; DNA; Microsatellites (Genetics); Genetic improvement.

Área de concentração: Bioquímica.

Titulação: Mestre em Biologia Funcional e Molecular.

Banca examinadora: Gonçalo Amarante Guimarães Pereira, Walter Maccheroni Junior, Antônio Vargas de Oliveira Figueira.

Data da defesa: 04/02/2009.

Programa de Pós-Graduação: Biologia Funcional e Molecular.

Campinas, 04 de fevereiro de 2009

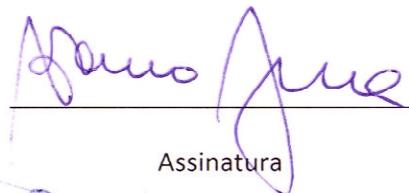
BANCA EXAMINADORA

Prof. Dr. Gonçalo Amarante Guimarães Pereira (Orientador)



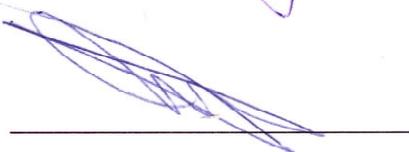
Assinatura

Prof. Dr. Antônio Vargas de Oliveira Figueira



Assinatura

Prof. Dr. Walter Maccheroni Júnior



Assinatura

Prof. Dr. Eduardo Galembeck



Assinatura

Prof. Dr. José Antônio Bressiani



Assinatura

Dedico este trabalho aos meus pais Hamilton e Elizete, exemplos de vida e incentivadores desse trabalho desde o início da minha formação.

À minha esposa Mariana, pelo constante apoio e compreensão.

Ao meu filho Mateus pela inspiração e alegria dada em todos os dias da minha vida desde a sua chegada.

AGRADECIMENTOS

A Deus.

A minha amada esposa Mariana e ao meu filho Mateus.

Aos meus pais e a minha irmã.

A CanaVialis pelo reconhecimento e pela oportunidade, oferecendo os recursos e o incentivo para a realização deste trabalho.

Ao professor Sizuo Matsuoka, Lee e Ricardo Madureira, pelo incentivo ao programa de pós-graduação da empresa e ao projeto de marcadores moleculares.

Ao professor Gonçalo, pela orientação, confiança, incentivo, cobranças desde ao período de graduação, iniciação científica até o mestrado.

Aos amigos do laboratório de biotecnologia da CanaVialis, Guilherme Moura, Rafaela, Fernanda, Cristiano, Gennari, Rebeca, André, Flávia, Annelise, Silvia, Alessandra, Mariana, Michele, Eduardo, Israel, Rafael, Veridiana, Cristiane pelo companheirismo e contribuição ao trabalho: todos os resultados e méritos deste trabalho são de vocês.

Ao amigo e orientador, Walter Maccheroni, por confiar a mim a responsabilidade deste trabalho, contribuindo como referência para minha formação profissional e acadêmica.

Aos amigos da Alellyx pela contribuição na execução deste projeto: Christian Greggio, Vagner Okura, Vivian e Almir.

Aos amigos do LGE.

Ao Welbe Oliveira Bragança, exemplo de vida e um dos responsáveis pelas atuais conquistas profissionais.

Ao meu amigo e primeiro orientador, Francesco Langone.

As professoras Maricilda, Edi e Marcela do CBMEG/Genética Humana, pela oportunidade dada há 10 anos para o início da minha carreira.

SUMÁRIO

| | |
|---|-----------|
| LISTA DE FIGURAS | 3 |
| LISTA DE TABELAS | 6 |
| RESUMO | 7 |
| ABSTRACT | 8 |
| 1. INTRODUÇÃO | 9 |
| 1.1. A CANA-DE-AÇÚCAR | 9 |
| 1.1.1. Aspectos econômicos | 9 |
| 1.1.2. Taxonomia | 11 |
| 1.1.3. Botânica | 13 |
| 1.1.4. Origem genética e geográfica do gênero <i>Saccharum</i> | 19 |
| 1.1.4.1. <i>Saccharum spontaneum</i> | 22 |
| 1.1.4.2. <i>Saccharum robustum</i> | 25 |
| 1.1.4.3. <i>Saccharum officinarum</i> | 26 |
| 1.1.4.4. <i>Saccharum sinense</i> e <i>Saccharum barberi</i> | 28 |
| 1.1.4.5. <i>Saccharum edule</i> | 30 |
| 1.1.4.6. <i>Miscanthus</i> | 30 |
| 1.1.4.7. <i>Erianthus</i> | 31 |
| 1.2. MELHORAMENTO GENÉTICO DE VARIEDADES | 33 |
| 1.3. MARCADORES MOLECULARES APLICADOS NO DESENVOLVIMENTO DE NOVAS VARIEDADES | 48 |
| 2. OBJETIVOS | 55 |
| 3. MATERIAL E MÉTODOS | 56 |
| 3.1. OBTENÇÃO DE MATERIAL VEGETAL | 56 |
| 3.2. EXTRAÇÃO DE DNA | 59 |
| 3.3. AMPLIFICAÇÃO DE DNA PELA PCR | 59 |
| 3.4. DETECÇÃO DE FRAGMENTOS | 60 |

| | | |
|--------|---|----|
| 3.5. | BANCO DE DADOS DE SEQÜÊNCIAS DE ESTs DE CANA-DE-AÇÚCAR ----- | 62 |
| 3.6. | SELEÇÃO <i>in silico</i> DE MARCADORES MICROSSATÉLITES ----- | 62 |
| 3.7. | DESENHO DE <i>PRIMERS</i> ----- | 63 |
| 3.8. | PROCESSO DE VALIDAÇÃO DOS MARCADORES MICROSSATÉLITES ----- | 63 |
| 4. | RESULTADOS E DISCUSSÃO ----- | 67 |
| 4.1. | DESENVOLVIMENTO DE UM PROCESSO DE SELEÇÃO DE MARCADORES MICROSSATÉLITES A PARTIR DE UM BANCO DE DADOS DE ESTs ----- | 67 |
| 4.1.1. | Seleção <i>in silico</i> de marcadores microssatélites candidatos----- | 67 |
| 4.1.2. | Avaliação dos marcadores selecionados <i>in silico</i> e seleção dos melhores candidatos para compor o sistema de identificação de variedades ----- | 72 |
| 4.2. | CARACTERIZAÇÃO DOS DEZ MARCADORES SELECIONADOS E VALIDAÇÃO DO SISTEMA DE IDENTIFICAÇÃO DE VARIEDADES DE CANA-DE-AÇÚCAR ----- | 80 |
| 4.2.1. | Principais características do sistema ----- | 80 |
| 4.2.2. | Avaliação do poder discriminatório dos marcadores selecionados----- | 86 |
| 4.2.3. | Reprodutibilidade do sistema de marcadores moleculares ----- | 90 |
| 4.2.4. | Aplicação dos marcadores moleculares no programa de melhoramento genético----- | 90 |
| 5. | CONCLUSÃO ----- | 93 |
| 6. | PRODUÇÃO CIENTÍFICA ----- | 93 |
| 7. | REFERÊNCIAS BIBLIOGRÁFICAS ----- | 97 |

LISTA DE FIGURAS

- Figura 1.** Touceira característica de uma variedade atual de cana-de-açúcar (a). Colmo característico de uma variedade de cana-de-açúcar, com detalhe para o internódio e nódio (b). Note a alternância da posição da gema entre internódios. Fonte (a e b): Subramanian, 1987.----- 13
- Figura 2.** Ilustração representando os tipos de nós, internódios e gemas observados em variedades de cana-de-açúcar, classificados de acordo com a forma. -----14
- Figura 3** Folha característica de variedade de cana-de-açúcar, com detalhe para bainha, colar (composto de *dewlap*, lígula e aurícula) e lâmina foliar. -----15
- Figura 4.** Ilustração representando os tipos de lígulas, *dewlap* e aurículas classificados de acordo com a forma. -----16
- Figura 5** A inflorescência da cana-de-açúcar. Variedades de cana-de-açúcar florescendo (a) com detalhe para a panícula em forma de flecha (b). -----17
- Figura 6.** Ilustração representando a formação de uma touceira de cana-de-açúcar com perfilhos formados a partir de brotos de gemas presentes no rizoma. -----18
- Figura 7.** Mapa indicando a região da Indonésia como provável centro de origem da cana-de-açúcar (círculo rosa). As principais regiões de cultivo da cana-de-açúcar estão coloridas de cinza. -----19
- Figura 8.** Esquema ilustrando a provável história evolutiva da cana-de-açúcar. -----21
- Figura 9.** Fotos mostrando plantas selvagens de *S. spontaneum* florescendo em três ambientes contrastantes na Índia. (a) Plantas de *S. spontaneum* crescendo em áreas rochosas; (b) Plantas de *S. spontaneum* em ambiente alagado; (c) Plantas de *S. spontaneum* em colinas. -----23
- Figura 10.** Dois tipos de plantas selvagens de *S. spontaneum*. (a) Acesso “SES113A” de *S. spontaneum* com touceira de hábito ereto; (b) Acesso “Coimbatore” de *S. spontaneum* com touceira na forma de tufo (tipo “capim”). -----24
- Figura 11.** Acesso “NG28-289” de *S. robustum* em banco de germoplasma nos EUA. Fonte: CanaVialis, 2005. -----26
- Figura 12.** (a) Acessos de *S. officinarum* em um banco de germoplasma na Índia. (b) Colmos de diferentes acessos de *S. officinarum* representando a beleza das plantas por sua diversidade de cores e formas dos colmos. -----27
- Figura 13.** Acesso de *S. sinense* (a) e *S. barberi* (b) em banco de germoplasma na Índia.-----29
- Figura 14.** Planta de *Miscanthus giganteus* em uma coleção na Alemanha.-----31

- Figura 15.** Plantas de *Erianthus* sp. em coleção de germoplasma nos EUA. -----32
- Figura 16.** Esquema mostrando os cruzamentos interespecíficos entre *S. officinarum* e *S. spontaneum* (a), formando progênie nobilizada em dois retrocruzamentos com *S. officinarum* (b, c e d) até a formação dos híbridos atuais de cana-de-açúcar. Fonte: Baseada em Ming *et al.*, 2006. -36
- Figura 17.** Esquema mostrando a organização cromossômica de um genoma de uma variedade atual de cana-de-açúcar. Em verde estão representados os cromossomos provenientes de *S. officinarum*. Em vermelho estão representados os cromossomos provenientes de *S. spontaneum*. -----37
- Figura 18.** Vista aérea do banco de germoplasma da CanaVialis em Maceió, AL, no período de florescimento (abril a junho). -----39
- Figura 19.** (a) Campânulas utilizadas em cruzamentos biparentais contendo inflorescências de um único progenitor masculino e de um progenitor feminino. Nestas campânulas as flores dos progenitores masculinos são posicionadas acima das flores dos progenitores femininos, para que o pólen caia sobre as flores e fecunde os óvulos; (b) Campânula utilizada para policruzamento contendo várias panículas de diversos progenitores; (c) grupo de sementes obtidas de uma panícula proveniente de um progenitor feminino. Fonte (a, b e c): CanaVialis 2007. -----40
- Figura 20.** Esquema para representar um microsatélite utilizando um organismo diplóide como exemplo. Dois indivíduos são apresentados. O motivo do microsatélite mostrado é a seqüência ATCC (tetranucleotídeo) e a variação no número de repetições deste motivo corresponde ao microsatélite. Os nucleotídeos em vermelho na seqüência correspondem aos *primers* desenhados para amplificar o microsatélite pela PCR. À direita, uma representação de como seria a visualização dos fragmentos amplificados em um sistema cromatográfico em gel. O conjunto de duas bandas em azul (alelos) para cada indivíduo corresponde a um haplótipo (perfil molecular) para um indivíduo. -----51
- Figura 21.** *Locus* de microsatélite para dois indivíduos diplóides [humanos] (a) e três indivíduos poliplóides [cana-de-açúcar] (b). O primeiro indivíduo (1) é homocigoto, com haplótipo formado por alelos de mesmo tamanho (visualização de uma única banda). O segundo indivíduo (2) é heterocigoto, com haplótipo formado por alelos de tamanho diferente (visualização de duas bandas). Os indivíduos (3) e (4) apresentam oito alelos e o indivíduo (5) apresenta seis alelos. ---52
- Figura 22.** Géis PAGE de três marcadores eliminados e um selecionado de acordo com os critérios de seleção propostos. Em (a) marcador de baixa amplificação, dificultando a identificação dos alelos (eliminado); (b) marcador monomórfico (eliminado); (c) marcador com grande número de

artefatos e amplificação desproporcional dos alelos (eliminado); (d) marcador sem artefatos e com número de alelos suficiente para discriminar os quatro acessos avaliados (selecionado).-----75

Figura 23. Exemplo de gel PAGE para um (CV38) dos dez marcadores selecionados pelos critérios de seleção propostos. A identificação dos alelos deste *loci* para cada amostra é clara. Em geral, a intensidade de amplificação de cada alelo é constante para todos os alelos e para todas as amostras. O perfil obtido para este *loci* é exclusivo para cada uma das 27 amostras, mostrando o excelente poder discriminatório do marcador.-----76

Figura 24. Perfis eletroforéticos dos 10 *loci* avaliados em seqüenciador de DNA (ABI 3730 XL) gerados pelo software STRAND. O número indicado acima de cada pico refere-se ao tamanho em pares de base (pb) de cada alelo. O nome entre parênteses corresponde ao nome do acesso.----77

Figura 25. Capacidade discriminatória do sistema de identificação de variedades de cana-de-açúcar em 54 acessos: separação dos acessos com pelo menos duas diferenças com o uso do conjunto de três *loci* (CV29, CV37 e CV38). Acessos nomeados como “Cana Pitú”, “Caiana”, “Cearense”, “Olhuda” e “Preta Kavangire” referem-se à *S. officinarum*; “IJ76-360” à *Erianthus arundinaceum*; “Khagzi” à *S. barberi*; e “Kavangire” à *S. sinense*.-----88

Figura 26. Capacidade do sistema de marcadores em discriminar 48 acessos originados de autofecundação (denominados de “selfing”) da variedade RB855002. Dendrograma gerado com o conjunto dos *loci* CV29, CV37, CV38, CV60, CV79, CV94 e CV106. -----89

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1. Lista das variedades brasileiras de cana-de-açúcar protegidas. ----- | 43 |
| Tabela 2. Comparação entre principais tipos de marcadores moleculares. ----- | 49 |
| Tabela 3. Relação dos 1.205 acessos de germoplasma utilizados no estudo distribuídos de acordo com o programa de melhoramento de origem e o respectivo país no caso de variedades híbridas e pelo nome no caso de espécies relacionadas. ----- | 57 |
| Tabela 4. Lista dos 27 acessos de cana-de-açúcar avaliados na segunda fase do processo de validação com indicação do <i>pedigree</i> correspondente. ----- | 64 |
| Tabela 5. Estatísticas da busca <i>in silico</i> de marcadores microssatélites no banco de dados de ESTs de cana-de-açúcar utilizando o software MISA. ----- | 69 |
| Tabela 6. Resultado do processo de seleção <i>in silico</i> de marcadores microssatélites. ----- | 71 |
| Tabela 7. <i>Loci</i> eliminados e selecionados para cada tipo de marcador microssatélite dos 150 avaliados. ----- | 74 |
| Tabela 8. Resultado da validação dos vinte marcadores avaliados. ----- | 78 |
| Tabela 9. Os 10 <i>loci</i> selecionados para compor o sistema de identificação de variedades de cana-de-açúcar. ----- | 81 |
| Tabela 10. Caracterização dos alelos que compõem o sistema de marcadores selecionado de cana-de-açúcar após análise de 1.205 acessos. ----- | 84 |
| Tabela 11. Provável distribuição dos dez marcadores microssatélites selecionados a partir da comparação com o genoma do sorgo. ----- | 92 |

RESUMO

Melhoristas de cana-de-açúcar tentam há muito tempo desenvolver um sistema confiável de identificação genética baseado em marcadores moleculares para ajudar programas de melhoramento genético. Além disso, a União Internacional para a Proteção de Novas Variedades de Plantas (UPOV) também procura por um sistema que atenda os padrões de um teste DHE (Distingüibilidade, Homogeneidade e Estabilidade) podendo ser usado para apoiar ou substituir o processo de caracterização morfológica convencional em processos de registro de variedades. Um processo de descoberta e validação de marcadores microssatélites usando um banco de dados de ESTs como única fonte de seqüências de DNA foi estabelecido no trabalho e um sistema de identificação genética baseada em marcadores microssatélites para cana-de-açúcar foi desenvolvido para apoiar o melhoramento de variedades de cana-de-açúcar e processos de registro de variedades. Um banco de dados de seqüências expressas (ESTs) com 352.122 seqüências foi avaliado para identificação de motivos de microssatélite e 150 *loci* com alto polimorfismo observado *in silico* foram selecionados e validados em dois sistemas de resolução: Polyacrylamide Gel Electrophoresis (PAGE) e Seqüenciador de DNA. Um conjunto de 10 *loci* foi selecionado de acordo com os valores de Conteúdo de Informação Polimórfica (PIC) e qualidade visual do perfil cromatográfico. A capacidade do sistema de discriminar indivíduos foi avaliada em 1.205 acessos de cana-de-açúcar e espécies relacionadas. Uma combinação de três *loci* foi suficiente para distinguir todos os acessos com pelo menos duas diferenças (alelos discriminatórios). A reprodutibilidade do sistema foi testada em um grande número de amostras obtidas de diversos tecidos, origens geográficas distintas e plântulas de dois métodos de propagação por cultura de tecido (meristema e calo), mostrando-se confiável. O sistema de identificação genética preenche todos os requisitos para distingüibilidade, homogeneidade e estabilidade (teste DHE). O sistema pode também ter muitas aplicações úteis em programas de melhoramento de cana-de-açúcar, tais como controle da identidade de acessos que compõem um banco de germoplasma, determinação de paternidade e rastreabilidade de clones em fase de seleção.

Palavras chave: Identificação genética; registro de cultivares; variedades de cana-de-açúcar; melhoramento; germoplasma.

ABSTRACT

Sugarcane breeders have been for long trying to develop a reliable molecular marker-based fingerprinting system that could aid their breeding programs. In addition, the International Union for the Protection of New Varieties of Plants (UPOV) has also been looking for such a system that, once it meets the standard of a DUS (Distinctness, Uniformity and Stability) test could be used to support or replace conventional morphological characterization used routinely to guarantee breeders property rights. A process of discovery and validation of microsatellites markers using EST database as the sole source for DNA sequences was established in this work and a microsatellite-based fingerprinting system for sugarcane was developed to support breeding of sugarcane varieties and property rights issues. An Expressed Sequence Tag (EST) database with 352,122 sequences was screened for microsatellite motifs and 150 *loci* with the highest polymorphism observed *in silico* were selected and validated in two fragment resolution systems, Polyacrylamide Gel Electrophoresis (PAGE) and automated DNA sequencer. A set of 10 *loci* was selected according to the Polymorphism Information Content (PIC) values and visual quality of chromatographic profiles. The capacity of the system to discriminate individuals was evaluated in 1,205 accessions of sugarcane and related species. A combination of three *loci* was sufficient to distinguish all accessions with the standard limit of at least two differences (discriminatory alleles). The reproducibility of the system was tested in large numbers of samples obtained from different tissues, distinct geographical origins, and plantlets from two tissue culture propagation methods (meristem and callus) and proved to be reliable. This fingerprinting system fulfills plant protection requirements for distinctness, uniformity, and stability (DUS test). The system may also have many useful applications in a sugarcane breeding program such as identification of mislabeled accessions in a germplasm bank, paternity determination and tracking of breeding populations.

Key Words: fingerprinting; property's rights, sugarcane varieties; breeding; germplasm.

1. INTRODUÇÃO

1.1. A CANA-DE-AÇÚCAR

1.1.1. Aspectos econômicos

A cana-de-açúcar é a principal matéria-prima para a fabricação de açúcar e contribui com 75% da produção mundial. A outra matéria prima utilizada para obtenção de açúcar é a beterraba (Ming *et al.*, 2006). Além da utilização do açúcar na alimentação humana, pode-se utilizá-lo na fermentação para obtenção de álcool combustível.

O Brasil é o maior produtor mundial de açúcar (18%), com domínio de 70% do atual comércio internacional (Licht, 2008). Da safra de 1990/1991 até a safra de 2007/2008 a produção brasileira de açúcar avançou de 7.365.344 toneladas a 30.760.165 toneladas (UNICA, 2008), representando um crescimento de aproximadamente 400% no período e rendendo ao país hegemonia no comércio internacional, com grandes perspectivas de ampliar essa participação caso haja retirada de subsídios praticados pela União Europeia para a produção de açúcar (UNICA, 2005).

Da fermentação do açúcar pode-se obter o álcool combustível, sendo este a principal alternativa energética no mundo a partir dos novos ideais de matriz energética que se estabelecem, principalmente por conta do agravamento da degradação ambiental pelo homem, dos perigos iminentes do aquecimento global e dos preços instáveis e exorbitantes do petróleo (Ortolan, 2006).

EUA e Brasil são os maiores produtores de etanol do mundo, responsáveis por 42% e 37% da produção mundial de etanol nos dias de hoje, respectivamente (UNICA, 2008). Embora os EUA tenham atualmente a maior produção de etanol no mundo, ela é baseada no uso do milho como matéria-prima. Em comparação com a cana-de-açúcar o milho possui um menor rendimento na produção de etanol (aproximadamente 4.000 litros por hectare),

além de afetar a disponibilidade de alimentos e de ração animal (English *et al.*, 2005). Com rendimento médio de 7.500 litros de etanol por hectare, a cana-de-açúcar é a melhor opção atualmente para produção deste biocombustível (Moreira, 2000).

De acordo com a União da Indústria de Cana-de-açúcar (UNICA), a exportação brasileira de etanol durante a safra 2006/2007 aumentou 20% e espera-se que atinja 5,5 bilhões de litros até 2012. O Brasil exportou aproximadamente 3,6 bilhões de litros de etanol na safra 2006/2007, representando um aumento na exportação de aproximadamente 300% desde 2003. Esta crescente demanda é resultado tanto da promoção de biocombustíveis de fontes renováveis, quanto das políticas de redução da dependência do petróleo proveniente de áreas de instabilidade política e da volatilidade de seus preços, exigindo produção de álcool combustível para os próximos 10 anos estimado em volumes superiores a 200 bilhões de litros anuais (Annual Energy Outlook 2008).

Nos últimos anos, grupos de pesquisa (públicos e privados) do mundo estão investindo no estabelecimento de tecnologias para a conversão do bagaço da cana-de-açúcar em açúcares fermentáveis, para produção de mais etanol (Demirbas, 2001). Embora ainda não possa ser aplicada em escala industrial, este tipo de tecnologia tem grande importância nesse contexto de crescente demanda (Badger, 2002). Por todos esses motivos o álcool produzido da cana-de-açúcar é o principal propulsor do setor sucroalcooleiro e atualmente movimenta grandes montantes de investimentos (Boyle *et al.*, 2008).

Nas usinas brasileiras, outro subproduto importante da cana-de-açúcar é a co-geração de energia elétrica a partir da queima do bagaço. Boa parte destas usinas são auto-suficientes do ponto de vista energético, criando grande expectativa na utilização da cana-de-açúcar como fonte de biomassa para a geração de energia elétrica, abastecendo cidades, indústrias etc (Scaramucci *et al.*, 2006).

1.1.2. Taxonomia

A cana-de-açúcar pertence ao gênero *Saccharum* (referência em latim ao açúcar) e foi descrita taxonomicamente pela primeira vez por Linnaeus em 1753 no livro “Species Plantarum” (Hitchcock, 1923). Em seu livro, Linnaeus constituiu o gênero *Saccharum* com duas espécies: *S. officinarum* e *S. spicalum* (Dillon *et al.*, 2007).

Em 1889, Eduard Hackel incluiu à classificação de Linnaeus quatro seções como subdivisões do gênero *Saccharum*: *Eusaccharum* (composta por seis espécies: *S. officinarum*, *S. robustum*, *S. arundinaceum*, *S. ciliare*, *S. narenga* e *S. porphyrocoma*), *Sclerostachya*, *Eriochrpsis* e *Leptosacharum*. No início do século XX coleções de espécies de *Saccharum* foram criadas e alguns pesquisadores propuseram-se a caracterizar estas espécies sugerindo outras formas de organização do gênero. Entre 1915 e 1919, Barber descreveu morfologicamente uma coleção de canas indianas, organizando-as em cinco grupos: Mungo, Saretha, Sunnabile, Nargori e Pansashi. Em 1925, Jeswiet realizou um estudo descritivo de órgãos vegetativos de diferentes espécies de cana-de-açúcar, propondo uma chave taxonômica que permitisse melhor reconhecimento das espécies que compunham o gênero (Rao, 1987).

Em revisão aos estudos, Dutt e Rao em 1950, sugeriram a composição do gênero *Saccharum* com cinco espécies: *S. officinarum*, *S. spontaneum*, *S. barberi* (que incluía os grupos Mungo, Saretha, Sunnabile e Nargori descritos por Barber), *S. sinense* (que incluía o grupo Pansashi) e *S. robustum*. Outras espécies que estavam descritas como plantas do gênero *Saccharum* foram re-classificadas nas cinco espécies citadas ou em outros novos gêneros [*Erianthus*, *Narenga* e *Sclerostachya*] (Grassl, 1964). Esta nova caracterização foi bem aceita, principalmente após os trabalhos de Mukherjee em 1957, mostrando a relação filogenética dos gêneros *Saccharum*, *Erianthus* seção *Ripidium*, *Sclerostachya* e *Narenga*, criando a expressão “Complexo *Saccharum*” para o grupo dos gêneros citados e descrevendo a capacidade destas espécies em realizar hibridações

entre si. Posteriormente, o gênero *Miscanthus* seção *Diandra* foi incluído a este grupo (Daniels *et al.*, 1975). Em 1981, o gênero *Saccharum* organizado na família Gramineae pela classificação de Engler em 1887 foi incluído à família Poaceae a partir de novas revisões taxonômicas por Cronquist (Scarpari e Beauclair, 2008).

Após todas as revisões taxonômicas citadas que ocorreram principalmente durante o século XX, atualmente a organização taxonômica mais utilizada (ou aceita) descreve a cana-de-açúcar como membro da família Poaceae, sub-família Panicoideae, tribo Andropogoneae; Sub-tribo Saccharinae, grupo Saccharastrae e gênero *Saccharum*. Ao gênero *Saccharum* inclui-se seis espécies: *S. officinarum*, *S. spontaneum*, *S. robustum*, *S. sinense*, *S. barberi* e *S. edule*. O “Complexo *Saccharum*” é constituído pelas espécies que compõem o gênero *Saccharum* juntamente com gêneros relacionados filogeneticamente [*Erianthus* seção *Ripidium*; *Sclerostachya*; *Narenga*; e *Miscanthus* seção *Diandra*] (Daniels e Roach, 1987). As atuais plantas cultivadas de cana-de-açúcar (variedades) são híbridos originados de cruzamentos interespecíficos feitos no início do século XX envolvendo principalmente plantas de *S. officinarum* e *S. spontaneum*. Estes híbridos são citados como *Saccharum* spp.

1.1.3. Botânica

A anatomia botânica de cana-de-açúcar se desenvolveu durante a fase de revisão taxonômica (essencialmente durante o século XX) e o principal trabalho neste assunto é do botânico Dillewjin, com o livro “Botany of Sugarcane” publicado em 1952 (Scarpari e Beauclair, 2008).

Espécies de cana-de-açúcar são perenes na sua forma natural e semiperenes na forma cultivada, pois exigem novos plantios após algumas colheitas (o número de colheitas depende do local e do manejo utilizado para o cultivo (Sizuo Matsuoka, comunicação pessoal). Geralmente as espécies de *Saccharum* possuem porte grande e sempre se desenvolvem na forma de touceiras. A parte aérea (projetada acima do solo) da planta é formada por colmos (com exceção de algumas plantas de *S. spontaneum*), folhas e inflorescências. A parte subterrânea é formada por raízes e rizomas.

Os colmos são cilíndricos e formados por sucessivos internódios (ou “entrenós”) ligados entre si pelos nós (ou “nós”). Cada nó possui uma gema axilar que se dispõe de maneira alternada entre cada internódio. Estas gemas axilares são utilizadas no cultivo extensivo para propagação das plantas e por isso, a cana-de-açúcar é frequentemente descrita como uma cultura de propagação vegetativa (figura 1).

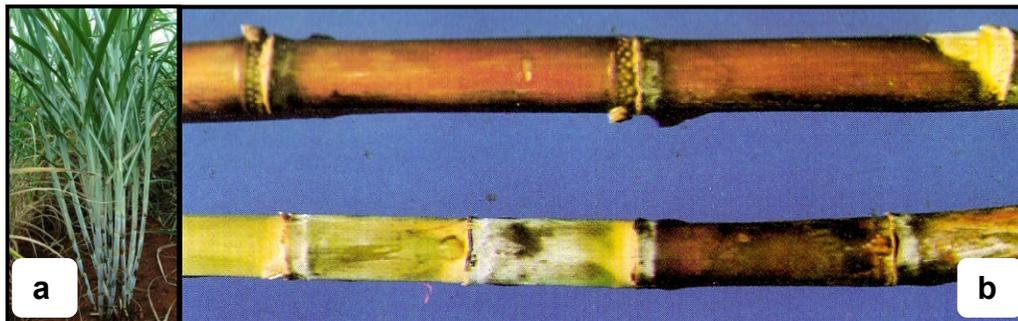


Figura 1. Touceira característica de uma variedade atual de cana-de-açúcar (a). Colmo característico de uma variedade de cana-de-açúcar, com detalhe para o internódio e nó (b). Note a alternância da posição da gema entre internódios. Fonte: (a) Sizuo Matsuoka, 2007 e (b) Subramanian, 1987.

Apenas na natureza ou para fins de melhoramento genético a cana-de-açúcar é propagada por semente. Internódios, nós e gemas que compõem um colmo possuem diversidade de cor, tamanho e forma, variando entre plantas (figura 2). Em plantas cultivadas os colmos são órgãos de reserva de sacarose.

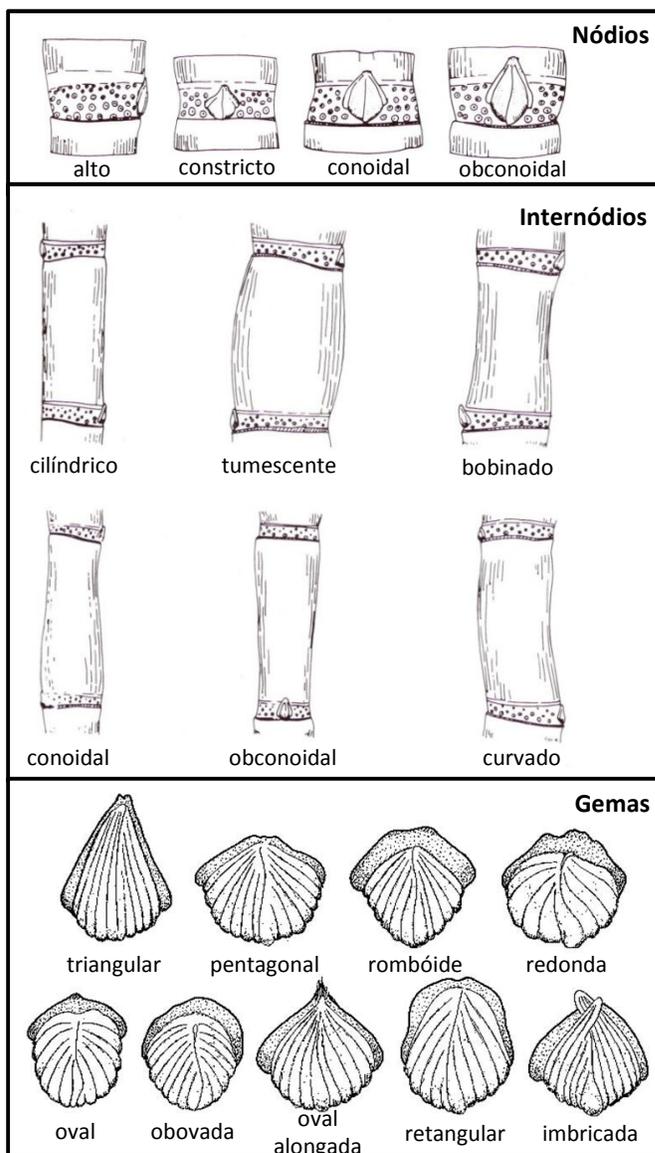


Figura 2. Ilustração representando os tipos de nós, internódios e gemas observados em variedades de cana-de-açúcar, classificados de acordo com a forma.

Fonte: Artschwager e Brandes, 1958.

As folhas estão fixadas à base dos nós e divididas em duas partes: bainha e lâmina foliar. A bainha é a uma região da folha de formato tubular que envolve o colmo e liga a lâmina foliar ao nó (no colmo). A lâmina foliar é a região sustentada pela nervura central que se projeta a partir da bainha na forma alongada e plana. A lâmina foliar pode se manter ereta até o topo, dobrada ou curva próximo ao topo ou curvada em sua altura média. A bainha pode apresentar coloração diferente da lâmina foliar, além da presença de pêlos (denominados de joçal) e cera. O ponto de ligação entre a bainha e a lâmina foliar é chamado de colar. O colar é dividido em três partes: *dewlap* (coletor de orvalho ou cotovelo), lígula e aurícula. A figura 3 apresenta em detalhe a bainha, colar e lamina foliar.

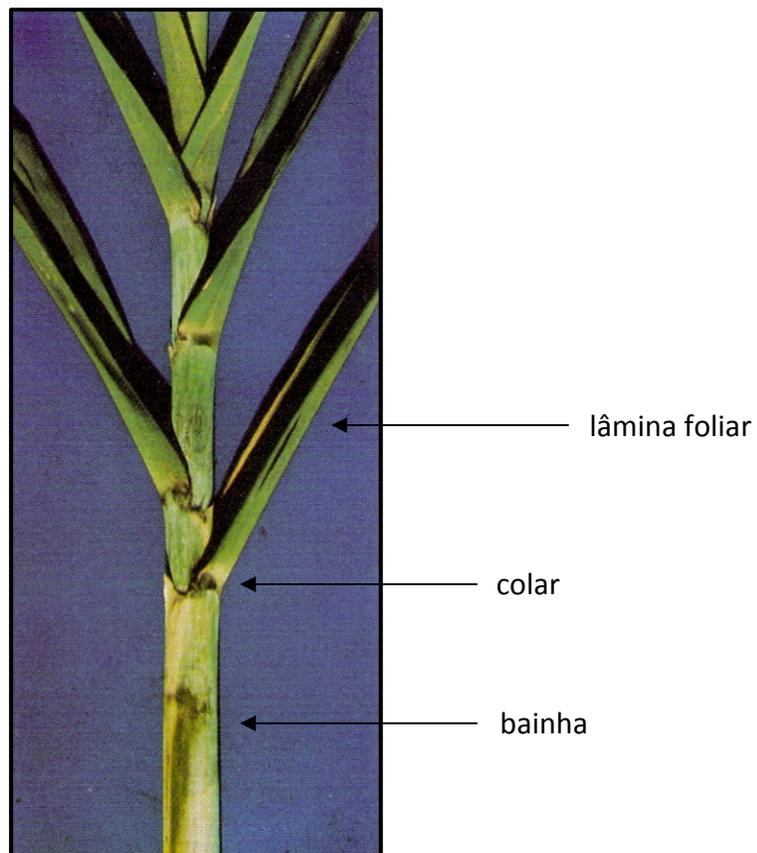


Figura 3 Folha característica de variedade de cana-de-açúcar, com detalhe para bainha, colar (composto de *dewlap*, lígula e aurícula) e lâmina foliar.

Fonte: Subramanian, 1987.

Entre espécies e híbridos há diversidade de cores, tamanhos e formas para todas as estruturas (figura 4).

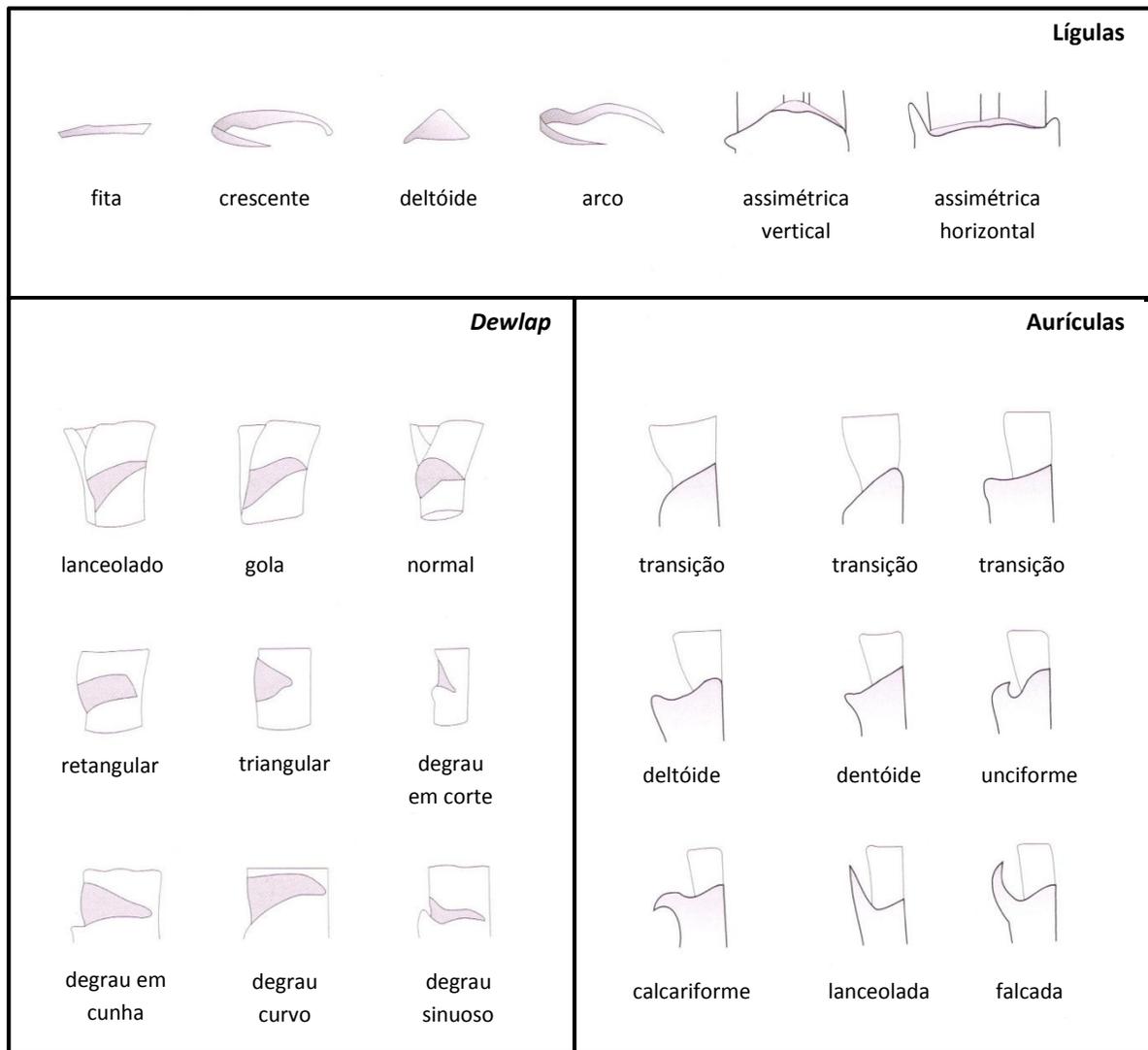


Figura 4. Ilustração representando os tipos de línguas, *dewlap* e aurículas classificados de acordo com a forma.

Fonte: Artschwager e Brandes, 1958.

A inflorescência da cana-de-açúcar é uma panícula aberta em forma de flecha (figura 5). As flores que compõem a inflorescência são muito pequenas. Uma panícula de cana-de-açúcar possui centenas a milhares de flores e cada colmo pode gerar uma panícula. A flor é hermafrodita e logo, pode ocorrer autofecundação, embora seja rara e uma mesma variedade ou planta pode ser utilizada em cruzamentos como progenitor masculino e/ou feminino (Van Dillewijn, 1952).



Figura 5 A inflorescência da cana-de-açúcar. Variedades de cana-de-açúcar florescendo (a) com detalhe para a panícula em forma de flecha (b).

Fonte: CanaVialis, 2009.

As raízes da cana-de-açúcar são do tipo fasciculado e se constitui de dois tipos: raízes adventícias (para absorção de água no estágio inicial de brotação das gemas) e permanentes [para absorção de água das camadas mais profundas, sobretudo no período da seca e sustentação da touceira já formada] (Mongelard, 1968; Thompson, 1976).

Rizomas são caules subterrâneos, espessados, ricos em reservas e de crescimento horizontal. Assim como nos colmos, os rizomas são constituídos de internódios, nós e gemas capazes de brotar e formar novos colmos. Os colmos derivados das gemas do rizoma são chamados de perfilhos. Cada perfilho

comporta-se de modo independente, porém todos possuem ligação entre si, constituindo a touceira. Os primeiros brotos formarão colmos (perfilho primário), com gemas que formarão novos colmos (perfilho secundário). As gemas deste novo colmo formarão novos colmos (perfilho terciário), assim sucessivamente e constituindo a touceira (figura 6). Em espécies como *S. spontaneum*, o crescimento lateral destes rizomas é bastante vigoroso, fazendo com que a planta produza muitos perfilhos e seja bastante invasiva. Nas demais espécies de *Saccharum* e nos atuais híbridos este crescimento horizontal é menos vigoroso e as touceiras formadas são mais compactas (sendo pouco ou não invasivas).

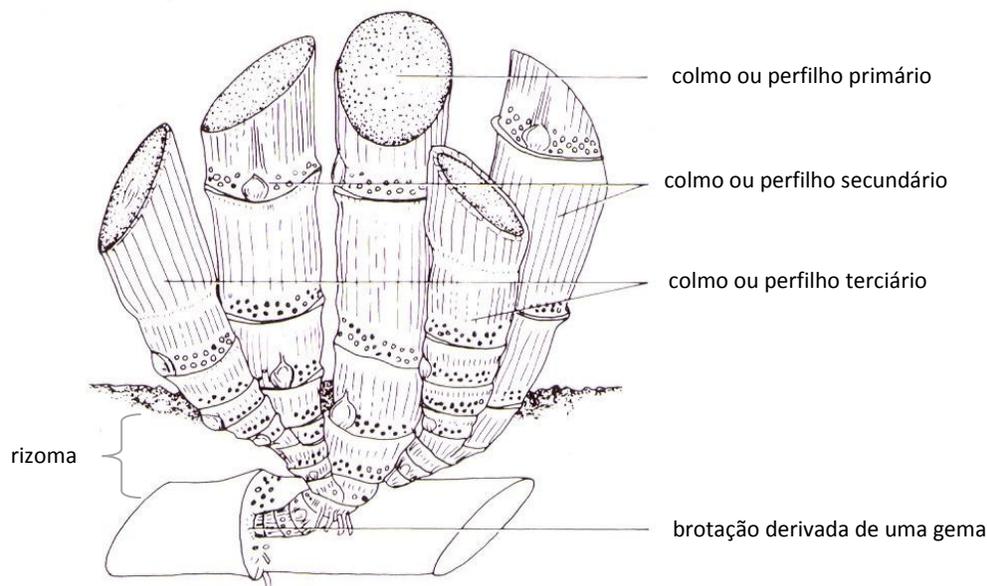


Figura 6. Ilustração representando a formação de uma touceira de cana-de-açúcar com perfilhos formados a partir de brotos de gemas presentes no rizoma.

Fonte: Artschwager e Brandes, 1958.

As estruturas apresentadas neste tópico, por apresentarem variações morfológicas entre espécies e híbridos do gênero *Saccharum*, são as estruturas comumente utilizadas em estudos de taxonomia ou em descrições para registro e caracterização de novas variedades.

1.1.4. Origem genética e geográfica do gênero *Saccharum*

Há incertezas sobre a origem da cana-de-açúcar, tanto na determinação exata do centro de origem, quanto na identificação de possíveis ancestrais.

Em relação à determinação do centro de origem da cana, há diferentes citações sobre a exata localização geográfica, com indicações para Índia, Indonésia, Papua Nova Guiné, China e ilhas da Polinésia, Fiji e Salomão. Embora existam divergências em relação aos países, não há contradições nas citações da Ásia como centro de origem [figura 7] (Mukherjee, 1957).

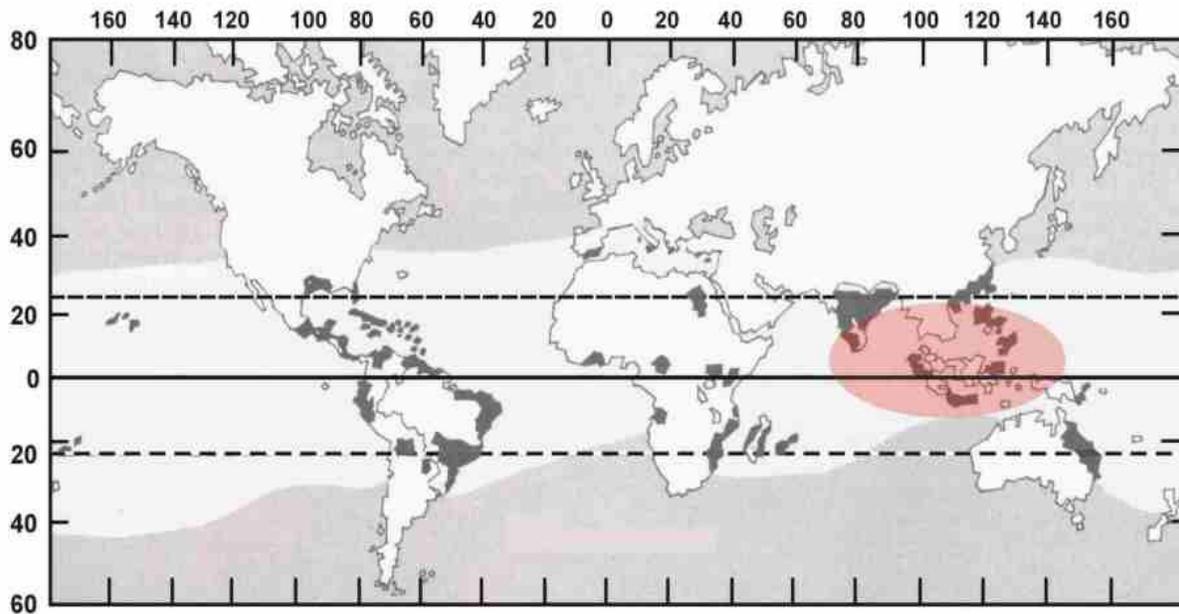


Figura 7. Mapa indicando a região da Indonésia como provável centro de origem da cana-de-açúcar (círculo rosa). As principais regiões de cultivo da cana-de-açúcar estão coloridas de cinza.

Fonte: Ometto, 1981.

Em relação à origem ancestral do gênero *Saccharum* a hipótese de um ancestral comum para os gêneros do “Complexo *Saccharum*” é tida como a mais provável (Barber, 1920). Estudos em citogenética para avaliação de progênies de cruzamentos realizados de plantas de *S. officinarum* e *Sclerostachya* (plantas que possuem número cromossômico básico de $x=10$) mostram que cinco dos dez cromossomos básicos existentes nas duas espécies são comuns sugerindo a existência de um único ancestral de número cromossômico básico de $x=5$ (Parthasarathy, 1948). Trabalhos semelhantes realizados na progênie gerada de cruzamentos intergenéricos envolvendo *Saccharum*, *Sclerostachya* e *Narenga* também sugerem a presença de um ancestral comum de número cromossômico básico de $x=5$ (Raghavan e Govindasamy, 1956).

Estes estudos coligados aos estudos de caracterização morfológica e ao fato de existir capacidade das espécies que compõem o “Complexo *Saccharum*” de se intercruzarem sugerem a existência de um ancestral comum para todas estas espécies. Das seis espécies que compõem o gênero *Saccharum*, apenas duas são consideradas espécies selvagens: *S. spontaneum* e *S. robustum* (Daniels e Roach, 1987).

Acredita-se que plantas de *S. officinarum* e *S. edule* surgiram a partir da seleção de populações selvagens de *S. robustum* e que *S. barberi* e *S. sinense* são espécies derivadas da hibridação natural de *S. officinarum* e *S. spontaneum*. Plantas atuais cultivadas surgiram principalmente da hibridação das espécies *S. officinarum* e *S. spontaneum* realizada pelos primeiros programas de melhoramento genético de cana-de-açúcar (Landell e Bressiani, 2008).

A figura 8 representa esquematicamente a provável origem e evolução das espécies de *Saccharum* e dos híbridos atuais cultivados.

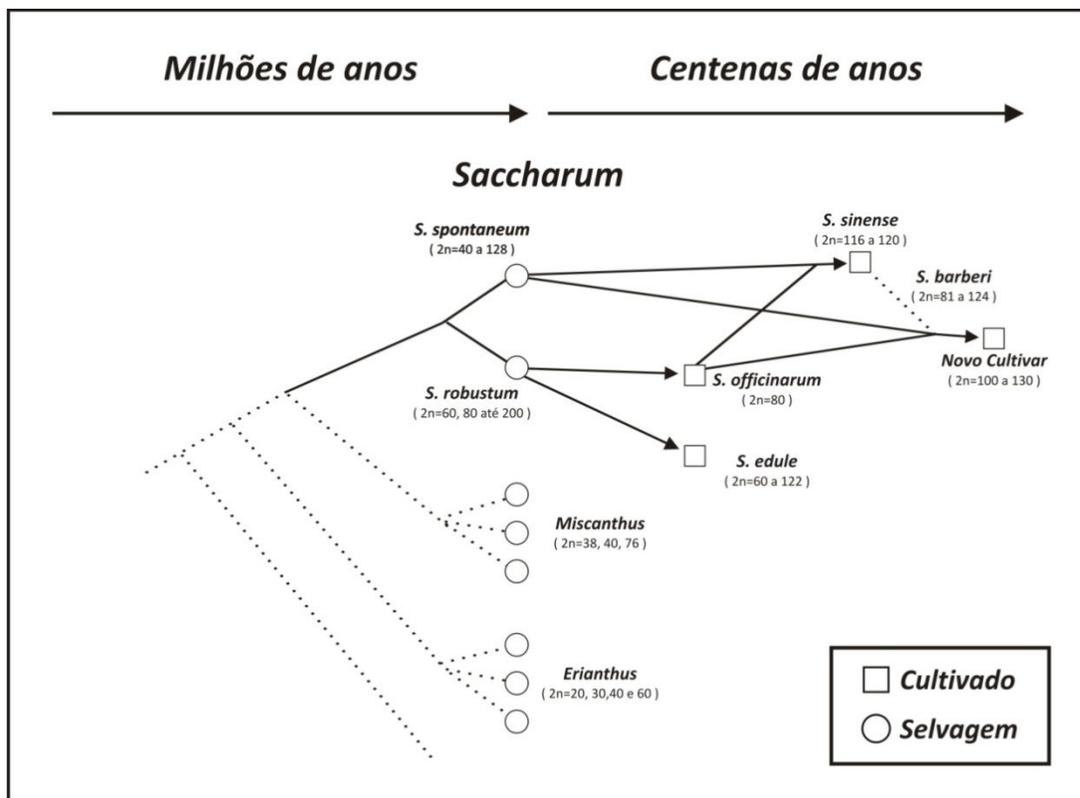


Figura 8. Esquema ilustrando a provável história evolutiva da cana-de-açúcar.
 Fonte: Baseado em Grivet *et al.*, 2006.

A seguir se discorrerá uma breve descrição de cada uma das espécies, de *Saccharum* apresentando detalhes de sua origem e diversidade (genética e geográfica), além de evidenciar alguns aspectos morfológicos que caracterizam cada espécie. Os gêneros *Miscanthus* e *Erianthus* serão descritos também, por serem de grande importância, não só na origem ancestral do gênero *Saccharum*, como na utilização dos atuais programas de melhoramento genético para o desenvolvimento de híbridos cultivados.

1.1.4.1. *Saccharum spontaneum*

Espécie altamente polimórfica. A variabilidade entre as plantas desta espécie é tão grande que somente na Índia, um dos principais centros de diversidade, já são conhecidos mais de 300 tipos. Plantas de *S. spontaneum* estão distribuídas desde o Japão e Indonésia/Papua-Nova Guiné até o Mediterrâneo e África, 8° S a 40° N a partir da linha do Equador. Formas desta espécie já foram coletadas em mais de 30 países [Argélia, Birmânia, Bangladesh, China, Congo, Egito, Ilha de Fiji, Índia, Indonésia, Iraque, Israel, Japão, Quênia, Malásia Cingapura, Ilhas Maurícius, Micronésia, Nova Guiné, Nepal, Nigéria, Paquistão, Filipinas, Itália (Sicília), Sri Lanka, Sudão, Tanzânia, Tailândia, Turquia, Turquemenistão, Uganda e Rússia] (Rao, 1987). A espécie *S. spontaneum* está organizada em duas subespécies: *indicum* e *aegyptiacum* (Pange, 1933).

A espécie possui número cromossômico total ($2n$) variando de 40 a 128 com $x=8$ (ou seja, todos os indivíduos que compõem a espécie são poliplóides, podendo ser classificados de pentaplóides (cinco cópias de cada um dos oito cromossomos) até hexadecaplóides (dezesseis cópias de cada um dos oito cromossomos). Plantas de *S. spontaneum* com $2n=40$ são encontradas na região do norte da Índia e há hipóteses de sua origem ter ocorrido a partir da hibridação natural de plantas de *Sclerostachya* e *Erianthus* seção *Ripidium* (Rao e Babu, 1955).

Plantas de *S. spontaneum* apresentam alta adaptabilidade, podendo ser encontrada em diversos ambientes (desertos, baixadas encharcadas, rochas, regiões litorâneas etc), altitudes (ao nível do mar às montanhas do Himalaia) e condições climáticas [de regiões de clima tropical a locais de inverno nevado] (Pange e Srinivasan, 1957). A figura 9 apresenta três fotos de plantas *S. spontaneum* em três ambientes contrastantes na Índia.



Figura 9. Fotos mostrando plantas selvagens de *S. spontaneum* florescendo em três ambientes contrastantes na Índia. (a) Plantas de *S. spontaneum* em áreas rochosas; (b) Plantas de *S. spontaneum* em ambiente alagado; (c) Plantas de *S. spontaneum* em colinas.

Fonte (a, b e c): Kandasami e Sreenivasan, 1983.

Morfológicamente as plantas podem se desenvolver em pequenas touceiras em forma de tufo (tipo “capim”) com ou sem colmos até em touceiras de hábito ereto com colmos de até 2,5 m de altura e diâmetro entre 5 a 13 mm (figura 10).

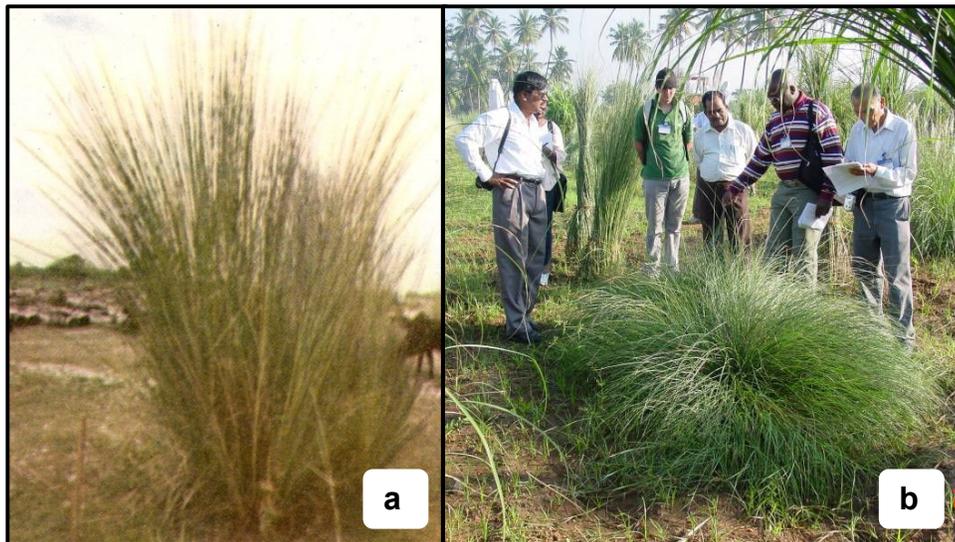


Figura 10. Dois tipos de plantas selvagens de *S. spontaneum*. (a) Acesso “SES113A” de *S. spontaneum* com touceira de hábito ereto; (b) Acesso “Coimbatore” de *S. spontaneum* com touceira na forma de tufo (tipo “capim”).

Fonte (a): Kandasami e Sreenivasan, 1983.

Fonte (b): CanaVialis, 2005.

Os colmos são ricos em fibras e são pobres no conteúdo de sacarose. Os internódios dos colmos são normalmente longos e os nós são sempre mais finos que os internódios. As cores dos colmos variam em tonalidade de amarelo a verde, de acordo com a planta e/ou estágio de desenvolvimento. As folhas variam de quase ausência de limbo, com apenas a nervura central, até limbos de 4 cm de largura. Em geral, são plantas de grande perfilhamento e possuem um rizoma vigoroso que se alastra no solo de modo bastante invasivo. A figura 9 ilustra dois tipos de *S. spontaneum*.

Por toda esta imensa variabilidade a espécie *S. spontaneum* é a que mais contribuiu com características de interesse (vigor, dureza, perfilhamento, capacidade de rebrota de soqueira, resistência a estresse hídrico e a salinidade, doenças e pragas) aos programas de melhoramento genético (Naidu e Sreenivasan, 1987), além de ser uma das principais fontes de fibra em programas de melhoramento genético que visam o desenvolvimento de variedades com grande potencial para a produção de biomassa (Roach, 1977).

1.1.4.2. *Saccharum robustum*

O provável centro de origem e diversidade é a região de Nova Guiné. A hipótese mais provável para a origem desta espécie é a de que surgiu a partir de hibridações naturais de *S. spontaneum* ($2n=80$) com outros gêneros [possivelmente *Erianthus* e *Miscanthus*] (Daniels *et al.*, 1975).

Em 1965, Price reconheceu cinco grupos de híbridos para *S. robustum* nos quais três deles apresentando número cromossômico $2n=60$ e os outros dois $2n=80$. Posteriormente foi relatado um sexto grupo, com $2n=114-205$.

Morfologicamente as plantas de *S. robustum* são extremamente vigorosas formando touceiras densas com colmos que podem chegar a 10 m de altura e diâmetro de 20 a 45 mm. As folhas são freqüentemente largas. Assim como em *S. spontaneum*, as plantas de *S. robustum* possuem colmos ricos em fibras e pouca sacarose (figura 11). São freqüentemente descritas como susceptíveis ao vírus do mosaico da cana-de-açúcar. Embora haja interesse potencial na utilização desta espécie em programas de melhoramento genético por conta do seu grande vigor, há poucas citações da utilização da espécie, exceto nos híbridos produzidos pelo programa de melhoramento no Havaí (Naidu e Sreenivasan, 1987; Roach e Daniels, 1987).



Figura 11. Acesso “NG28-289” de *S. robustum* em banco de germoplasma nos EUA. Fonte: CanaVialis, 2005.

1.1.4.3. *Saccharum officinarum*

Diferentemente das espécies *S. spontaneum* e *S. robustum* a espécie *S. officinarum* não é conhecida no estado selvagem. Embora o centro de origem não seja exatamente conhecido, há citações sugerindo sua origem nas regiões da Indonésia, Birmânia e China. O principal centro de diversidade está localizado em Papua-Nova Guiné. Há citações de outros centros de diversidade, como Índia, Norte da África, Malásia, ilha de Fiji e Polinésia (Rao, 1987).

A espécie possui número cromossômico $2n=80$ com $x=10$ (ou seja, todos os indivíduos que compõem a espécie são poliplóides, sendo classificados como octaplóides, o que significa ter oito cópias de cada um dos dez cromossomos). É a única espécie dentro do gênero *Saccharum* em que o número cromossômico não é variável entre os indivíduos.

É conhecida como “cana nobre”, por chamar a atenção pelos seus colmos grossos (diâmetro variando de 14 a 46 mm), suculentos, doces (com porcentagem de sacarose entre 7-22%) e diversidade de cores (figura 12). Geralmente apresenta folhas largas com limbos de 5 a 7 cm de largura e touceiras com poucos perfilhos.



Figura 12. (a) Acessos de *S. officinarum* em um banco de germoplasma na Índia. (b) Colmos de diferentes acessos de *S. officinarum* representando a beleza das plantas por sua diversidade de cores e formas dos colmos.

Fonte (a): CanaVialis, 2005.

Fonte (b): Sreenivasan e Nair, 1991.

Foi a espécie predominantemente cultivada no mundo até o início do século XX. No Brasil a espécie foi introduzida durante o período da colonização por Martim Afonso de Souza em 1532 através de mudas trazidas da Ilha da Madeira dando início ao ciclo do açúcar no Brasil (Figueiredo, 2008).

Enquanto a espécie foi cultivada, variedades de *S. officinarum* foram selecionadas de acordo com sua produtividade. A primeira variedade que se tornou conhecida e plantada em escala mundial foi a variedade “Creoula”. Em seguida, surgiu outra variedade denominada de “Otaheite” e que foi disseminada pelo mundo, recebendo diferentes nomes de acordo com a região em que foi cultivada, tais como, Bourbon, Lahaina e Vellai. Desde as primeiras décadas do século XX deixou de ser cultivada extensivamente no mundo por susceptibilidade a diversas doenças, tornando-se a espécie base nos programas de melhoramento genético juntamente com *S. spontaneum* na formação das atuais variedades de cana-de-açúcar. *S. officinarum* contribuiu principalmente como fonte genética para produção de sacarose, além de outras características importantes, tais como colmos suculentos, boa pureza do caldo e teor de fibra adequado para moagem etc (Naidu e Sreenivasan, 1987).

1.1.4.4. *Saccharum sinense* e *Saccharum barberi*

Por se tratarem de espécies de morfologia parecida e provável origem genética comum, as duas espécies serão descritas em um único tópico.

Embora não haja definição segura sobre a origem geográfica destas espécies, tem-se como hipótese mais provável a China e o norte da Índia como centros de origem para *S. sinense* e *S. barberi*, respectivamente (Daniels e Roach, 1987).

Trabalhos em citogenética mostram que tanto *S. sinense*, conhecida com “cana da Índia”, quanto *S. barberi* surgiram da hibridação natural de *S. officinarum* com *S. spontaneum* (D’Hont *et al.*, 2002). *S. sinense* possui número cromossômico $2n=80-124$ e *S. barberi* $2n=111-120$.

Assim como *S. officinarum*, as espécies *S. sinense* e *S. barberi* foram espécies cultivadas. *S. barberi* foi cultivada de modo mais restritivo no norte da Índia, possivelmente por ser adaptada e originária daquela região. *S. sinense* foi cultivada em algumas regiões do mundo, como Brasil (por volta de 1860), Ilhas Maurícius (por volta de 1870), EUA e Porto Rico (por volta de 1915). Há citações indicando tolerância à doença do vírus do mosaico da cana-de-açúcar em *S. sinense*, justificando a disseminação da espécie da China para as outras regiões citadas.

Apresentam colmos finos a médios e carecem de interesse no melhoramento atual, principalmente devido à dificuldade de florescimento e à esterilidade. A figura 13 apresenta fotos apresentando acessos das duas espécies citadas.



Figura 13. Acesso de *S. sinense* (a) e *S. barberi* (b) em banco de germoplasma na Índia.

Fonte (a e b): CanaVialis, 2005.

1.1.4.5. *Saccharum edule*

Possui centro de origem na região de Nova Guiné e ilha de Fiji e é freqüentemente citado como produto de hibridação de *S. officinarum* ou *S. robustum* com outro gênero (possivelmente *Miscanthus*). Possui número cromossômico $2n=60, 70$ ou 80 (Rao, 1987).

É uma espécie incomum que tem como principal característica a formação de inflorescências abortivas. Dessa forma, não há participação de *S. edule* em programas de melhoramento. Nas regiões citadas como centro de origem, a espécie é cultivada por nativos que utilizam sua inflorescência abortiva (de forma compacta) na alimentação (James, 2004).

1.1.4.6. *Miscanthus*

Possui centro de origem e diversidade no sudoeste asiático (China, Japão e ilhas da Polinésia), com algumas espécies ocorrendo na África. Por ter sido introduzido em toda a Europa e por existir freqüente hibridação interespecífica, há diversidade também no continente europeu. Atualmente, o gênero *Miscanthus* é composto por catorze espécies, que crescem em clima tropical, sub-tropical e temperado e em diferentes altitudes (podendo chegar a altitudes maiores de 3000 metros). Na Europa, espécies de *Miscanthus* foram introduzidas como plantas ornamentais. Por serem plantas de alta produtividade em fibra, com touceiras com grande quantidade de colmos de 10 mm de diâmetro e de 2-4 m de altura (figura 14), plantas de *Miscanthus giganteus* estão sendo utilizadas com freqüência na Europa como fonte de fibra para produção de energia elétrica. Algumas espécies são extremamente invasivas, sendo consideradas ervas-daninhas. Há espécies diplóides e poliplóides com número cromossômico básico de 19 e número cromossômico $2n=38$ (diplóide), 57 (triplóide) ou 76 (tetraplóide). Espécies de *Miscanthus* tem sido utilizadas nos programas de

melhoramento de cana-de-açúcar, visando principalmente a obtenção de plantas adaptadas a climas temperados (Scally *et al.*, 2000).



Figura 14. Planta de *Miscanthus giganteus* em uma coleção na Alemanha.

Fonte: Mendel Biotechnology, 2008.

1.1.4.7. *Erianthus*

Possui centro de origem na região entre China, Birmânia e Indonésia, mas espécies deste gênero se espalham também pela Índia e Nova Guiné. A taxonomia do gênero é extremamente complexo e divergente. Como já citado, até a revisão taxonômica de Dutt e Rao em 1950 as espécies do gênero *Erianthus* eram freqüentemente classificadas como *Saccharum*. Uma das características morfológicas que facilmente diferencia espécies do gênero *Erianthus* do gênero *Saccharum* é a ausência de colar (*dewlap*, lígula e aurícula) nas plantas do gênero *Erianthus*. Do ponto de vista de melhoramento genético, há diversos trabalhos

para introgressão de *Erianthus* (principalmente a espécie *E. arundinaceus*) em *Saccharum* visando aumentar a biomassa da plantas cultivadas de cana-de-açúcar. Espécies do gênero *Erianthus* destacam-se pelos seus colmos ricos em fibra de até 15 mm de diâmetro e de 5-7 m de altura (figura 15). Além disso, as espécies apresentam outras características importantes, tais como resistência ao frio, seca e algumas pragas. O gênero possui espécies diplóides e em sua maioria poliplóides com número cromossômico básico de $x=5, 8$ ou 10 e número cromossômico total de $2n = 20$ a 90 . Dentre as espécies que estão dentro da subdivisão *Ripidium* do gênero *Erianthus* (pertencente ao “Complexo *Saccharum*” estão *E. elephantinus* ($2n = 20$; $x=5$ ou 10), *E. arundinaceus* ($2n = 40$ ou 60 ; $x = 5$ ou 10), *E. giganteus* ($2n = 30$, $x=5$ ou 10), *E. procerus* ($2n = 40$; $x = 5, 8$ ou 10) e *E. ravennae* [$2n = 20$; $x=5$ ou 10] (Besse *et al.*, 1997; Besse e McIntyre, 1999).



Figura 15. Plantas de *Erianthus* sp. em coleção de germoplasma nos EUA.

Fonte (a): CanaVialis, 2005.

Fonte (b): CanaVialis, 2007.

1.2. MELHORAMENTO GENÉTICO DE VARIEDADES

Como já descrito, plantas de *S. officinarum* foram cultivadas de maneira extensiva em todo mundo até o início do século XX. Com a disseminação destas plantas pelo mundo, a busca por tipos mais produtivos e adaptados aos diferentes ambientes se tornou necessário. Diversos trabalhos realizados em dois grandes centros de pesquisa em cana-de-açúcar localizados em Java (atual Jakarta, Indonésia) e em Barbados (ilha do Caribe, América Central) na metade do século XIX divulgaram de maneira pioneira uma série de informações sobre o florescimento, viabilidade de pólen e formação de sementes de cana-de-açúcar (Rao, 1987).

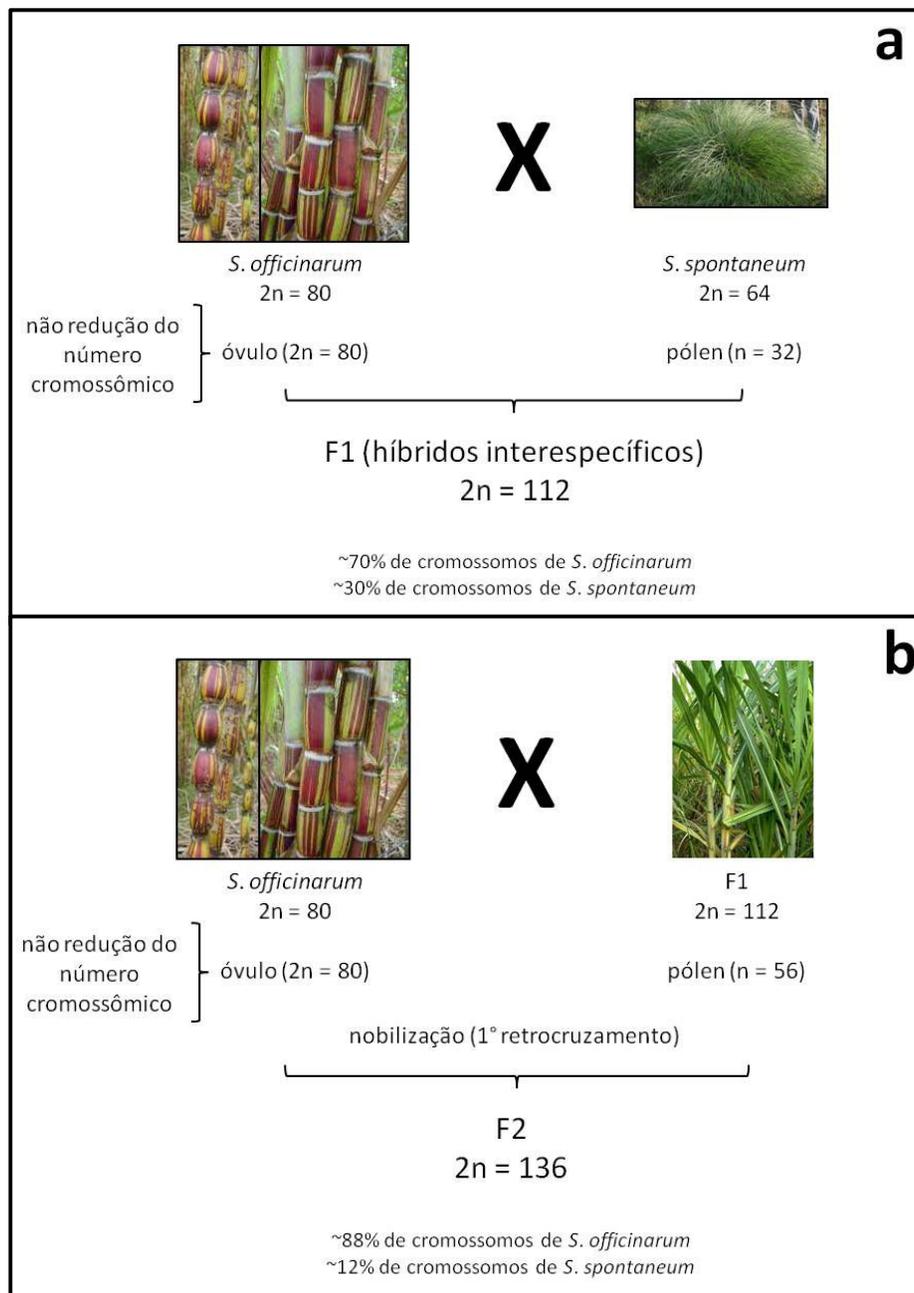
A determinação das condições ideais de florescimento, produção de pólen e realização de cruzamentos para obtenção de sementes viáveis foram os passos decisivos para a constituição dos primeiros programas de melhoramento genético de cana-de-açúcar. Desta forma, nos centros de pesquisa em Java e em Barbados foram criadas coleções de *S. officinarum* (bancos de germoplasma) e se estabeleceram condições para realização de hibridações entre plantas desta espécie. Por volta de 1888, ocorreram as primeiras hibridações entre diferentes acessos (plantas que compõem o banco de germoplasma) de *S. officinarum*. As sementes obtidas destas hibridações eram plantadas e a progênie era avaliada para seleção das plantas mais produtivas e/ou adaptadas. Diversas variedades de *S. officinarum* foram desenvolvidas a partir de hibridações e seleções realizadas em Java e Barbados. Logo, pesquisadores de outras localidades, como Havaí, Nova Guiné, Austrália e Índia, começaram a selecionar variedade pelo mesmo processo. Mesmo com a seleção de plantas mais produtivas e adaptadas, a espécie *S. officinarum* apresentava susceptibilidade a várias doenças e estresses abióticos (seca, por exemplo), tornando-se difícil o cultivo em determinadas localidades.

No início do século XX, pela necessidade de se obter variedades mais resistentes a diversas doenças (principalmente a uma doença conhecida como sereh, na região da Índia e de Java) e adaptadas a diversas condições ambientais, melhoristas dos programas de melhoramento de cana-de-açúcar de Java e da Índia (Coimbatore), desenvolveram híbridos interespecíficos de *S. officinarum* e *S. spontaneum*, na intenção de solucionar os problemas de susceptibilidade a doenças e adaptabilidade. Sendo assim, *S. officinarum* manteria a herança de riqueza em sacarose e *S. spontaneum* acrescentaria a herança de rusticidade nos híbridos derivados destes cruzamentos (Ming *et al.*, 2006).

Uma característica importante deste cruzamento é a transmissão $2n + n$ para formação da progênie interespecífica (F1). Esta transmissão só foi possível pela formação de gametas não reduzidos (gametas com número cromossômico ao da célula somática) nas plantas de *S. officinarum*. Estes gametas são resultado de um processo meiótico anormal [falha na fase I ou II da meiose] (Schifino-Wittmann e Dall'Agnol, 2001) em que a redução do número cromossômico não ocorre. Sendo assim, óvulos $2n$ (anormais) de *S. officinarum* tiveram capacidade de fusão com pólenes n (haplóides) de *S. spontaneum*, produzindo a progênie F1 (Allard, 1960).

À progênie destes cruzamentos interespecíficos realizaram-se sucessivos retrocruzamentos com *S. officinarum*, na intenção de conservar nos híbridos a riqueza em açúcar, adicionando-se apenas algumas características de rusticidade, originadas de *S. spontaneum*. O processo de retrocruzamento com plantas de *S. officinarum* ficou conhecido como **nobilização**, em referência ao nome dado as plantas de *S. officinarum*, cana nobre (Arceneaux, 1967). Desta forma, a maior parte do genoma de um híbrido de cana-de-açúcar vem de *S. officinarum* [80%] (Grivet *et al.*, 2006). Depois dos retrocruzamentos com *S. officinarum*, os híbridos gerados voltaram a ter capacidade de gerar cruzamentos do tipo $n + n$. Desde então, os híbridos cultivados até os dias atuais são híbridos interespecíficos e os programas de melhoramento genético espalhados pelo mundo desenvolvem novas variedades a partir da hibridação dos híbridos

melhorados. Como já descrito, tanto *S. officinarum* ($2n = 80$) quanto *S. spontaneum* ($2n = 40$ a 128) são espécies poliplóides. O cruzamento destas espécies com o processo de nobilitação gerou híbridos poliplóides de número cromossômico variável. Por conta disso, é freqüente a ocorrência de aneuploidia (processo em que ocorre o ganho ou a perda de cromossomos). A figura 16 apresenta um esquema para ilustrar o processo de nobilitação até a formação dos híbridos atuais.



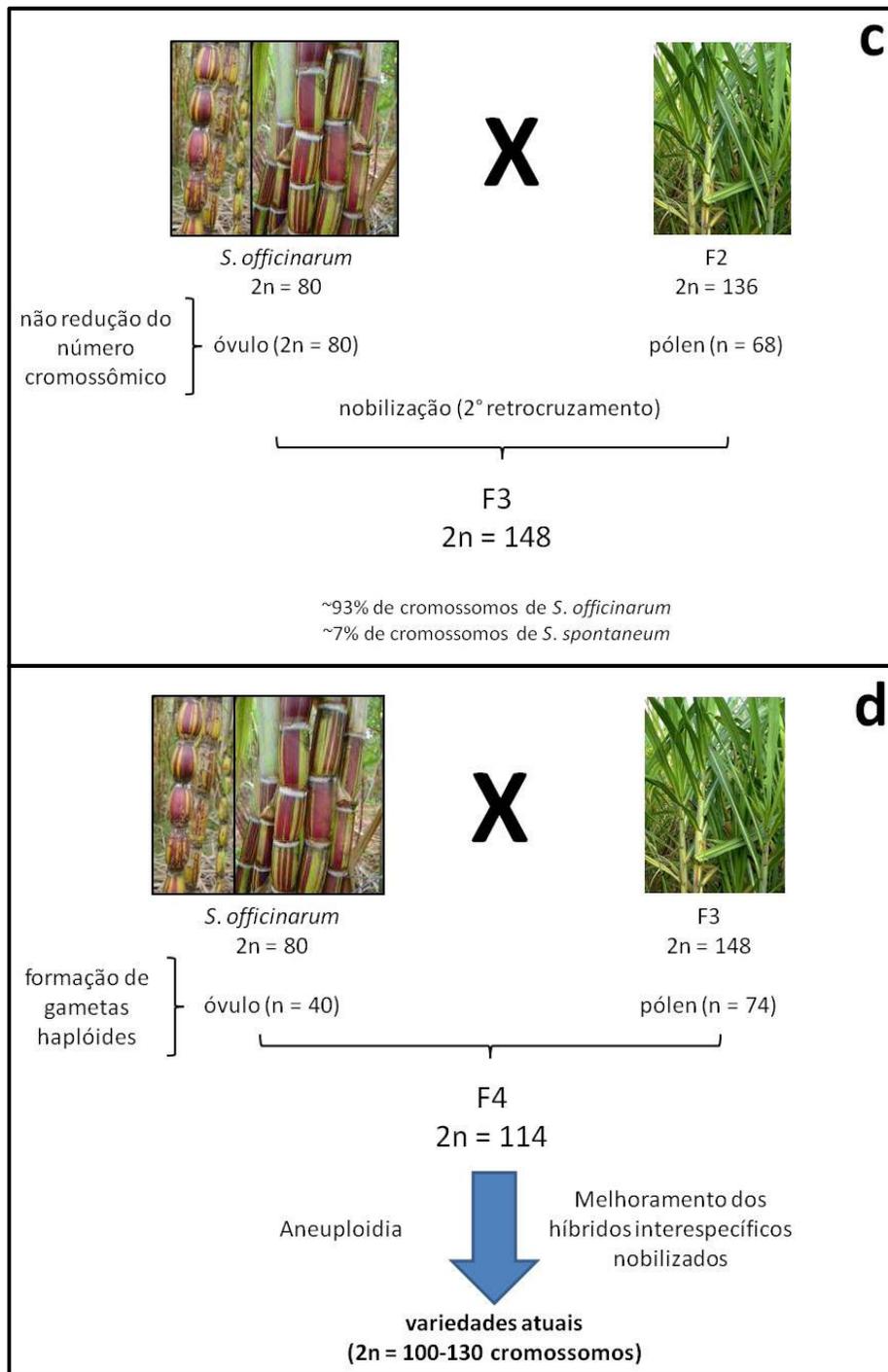


Figura 16. Esquema mostrando os cruzamentos interespecíficos entre *S. officinarum* e *S. spontaneum* (a), formando progênie nobilitada em dois retrocruzamentos com *S. officinarum* (b, c e d) até a formação dos híbridos atuais de cana-de-açúcar. Fonte: Baseada em Ming *et al.*, 2006.

Sendo assim, híbridos atuais possuem número cromossômico bastante variável, entre 100 a 130 cromossomos, com 80% destes cromossomos derivados de *S. officinarum*, 10% de *S. spontaneum* e 10% de cromossomos híbridos entre as espécies [figura 17] (D'Hont , 2005).

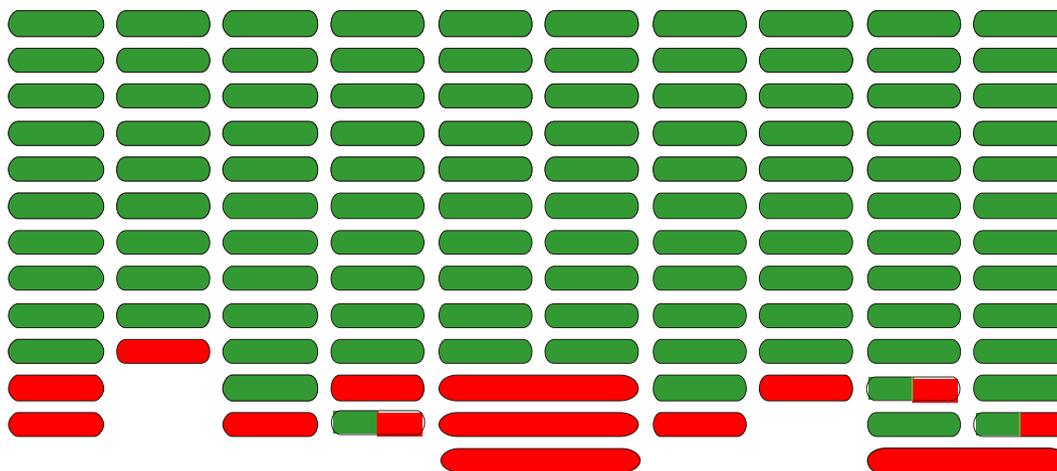


Figura 17. Esquema mostrando a organização cromossômica de um genoma de uma variedade atual de cana-de-açúcar. Em verde estão representados os cromossomos ou partes provenientes de *S. officinarum*. Em vermelho estão representados os cromossomos ou partes provenientes de *S. spontaneum*.

Fonte: Baseado em Grivet *et al.*, 1996.

As variedades produzidas pelo processo de nobilização se espalharam pelo mundo com muita rapidez e sucesso. A utilização destes híbridos na forma cultivada colocou fim no cultivo extensivo da cana-nobre. Dos novos híbridos desenvolvidos, destacou-se a variedade POJ2878 (o prefixo POJ refere-se às plantas produzidas pelo programa de melhoramento de Java), apelida na época como “cana maravilha de Java” (James, 2004). Em 1929, esta variedade ocupou aproximadamente 90% da área plantada de cana-de-açúcar no mundo. Nesta época, os dois principais programas de melhoramento genético no mundo eram os programas de Java (com o desenvolvimento de variedades nomeadas com o

prefixo POJ) e de Coimbatore na Índia (com o desenvolvimento de variedades nomeadas com o prefixo Co). Os híbridos destes dois programas foram a base de praticamente todos os outros programas de melhoramento e conseqüentemente das atuais variedades de cana-de-açúcar do mundo. No Brasil, as variedades POJ e Co foram introduzidas por volta de 1928, na intenção de adquirir e cultivar variedades resistentes à doença conhecida como mosaico (causada por um vírus).

Na década de 1920 a produção sucrocanavieira brasileira sofreu uma forte queda por conta de uma epidemia de mosaico. Com a introdução das variedades POJ e Co no país e impulsionados pelo sucesso relatado pelos outros programas de melhoramento no mundo que realizaram o desenvolvimento de novas variedades a partir dos híbridos nobilitados, o Instituto Agrônomo de Campinas (IAC) tornou-se pioneiro no melhoramento genético no país com o desenvolvimento das primeiras variedades brasileiras, nomeadas de IAC36-25 e IAC47-31 (Figueiredo, 2008).

Posteriormente, um novo programa de melhoramento foi criado na cidade de Campos, no estado do Rio de Janeiro, produzindo por volta de 1955 as variedades mais cultivadas da época. Estas variedades foram nomeadas com o prefixo CB (Campos, Brasil). Durante a década de 1950 estes dois programas de melhoramento foram importantes no desenvolvimento de variedades resistentes a doença conhecida como carvão, causada por um fungo. Uma grande epidemia desta doença na década de 1940 abalou pela segunda vez no século XX o setor sucrocanavieiro do país.

Por volta de 1970 outros dois programas de melhoramento foram criados no país: Planalsucar [que originou as variedades com prefixo RB (República do Brasil)] e do Centro de Tecnologia Coopersucar [que originou as variedades com prefixo SP (São Paulo)]. O programa Planalsucar foi transferido em 1992 para um conjunto de Universidades Federais, sendo renomeado de Ridesa. O Centro de Tecnologia Coopersucar foi reestruturado em 2004 e renomeado de Centro de Tecnologia Canavieira (originando as variedades com o prefixo CTC). Os programas RB e SP são marcos na história da canavicultura

brasileira, pelo desenvolvimento de uma série de variedades que ocuparam aproximadamente 90% da área cultivada até o início do século XXI, alavancando a produção de açúcar e álcool combustível no país e posicionando o Brasil como destaque no mundo. Em 2003 a partir de investimentos de capital privado surge a empresa CanaVialis S.A como o mais novo programa de melhoramento do país, com o objetivo principal de desenvolver e fornecer novas variedades mais produtivas e adaptadas às novas fronteiras agrícolas, no país e no mundo.

O processo de desenvolvimento de variedades de cana-de-açúcar em programas de melhoramento genético é extremamente demorado. Tradicionalmente, os programas brasileiros levam de 8 a 12 anos para desenvolver uma variedade. O processo de desenvolvimento de uma variedade comercial de cana-de-açúcar pode ser dividido em quatro etapas:

- a) Seleção dos progenitores, a partir de acessos de um banco de germoplasma, utilizados nas hibridações (figura 18).

Sendo assim, é fundamental que o banco de germoplasma utilizado contenha diversidade genética de acessos (com diferentes híbridos e espécies de cana-de-açúcar) para que se tenha grande variabilidade de características agrônômicas como substrato para se formar variedades híbridas superiores.



Figura 18. Vista aérea do banco de germoplasma da CanaVialis em Maceió, AL, no período de florescimento (abril a junho).

Fonte: CanaVialis, 2007.

b) Hibridação dos progenitores e produção de sementes viáveis.

Como já citado, a inflorescência da cana-de-açúcar é hermafrodita e sendo assim, o melhorista pode utilizar a estratégia de hibridação biparental, com a utilização de apenas dois progenitores no cruzamento, um como masculino e outro como feminino; ou hibridação multiparental (policruzamento) que utiliza diversos progenitores masculinos com apenas um feminino [figura 19] (Landell e Bressiani, 2008).



Figura 19. (a) Campânulas utilizadas em cruzamentos biparentais contendo inflorescências de um progenitor masculino e de um progenitor feminino. Nestas campânulas as flores dos progenitores masculinos são posicionadas acima das flores dos progenitores femininos, para que o pólen caia sobre as flores e fecunde os óvulos; (b) Campânula utilizada para policruzamento contendo várias panículas de diversos progenitores; (c) grupo de sementes obtidas de uma panícula proveniente de um progenitor feminino. Fonte (a, b e c): CanaVialis 2007.

A estratégia de policruzamento é bastante utilizada pelos programas de melhoramento de cana-de-açúcar. Variedades cultivadas de grande importância atualmente (RB865715, RB72454, RB835486 etc) são derivadas desta estratégia. O uso do policruzamento como método de hibridação é bastante utilizado na cultura da cana-de-açúcar porque para a obtenção de uma variedade, que apresente todas as características morfológicas e agronômicas desejadas, é necessário que o melhorista escolha a combinação ideal de progenitores e a determinação desta combinação é difícil, com poucas boas combinações (com capacidade de gerar uma variedade com todas as características desejadas) conhecidas (José Antônio Bressiani – comunicação pessoal).

Por conta disso, os melhoristas freqüentemente recorrem ao uso de policruzamentos, aumentando a probabilidade de combinação do progenitor feminino com os diversos progenitores masculinos e conseqüentemente aumentando a probabilidade de se obter uma boa variedade.

O principal problema desta estratégia é que após a seleção da variedade, não é possível repetir o cruzamento na intenção de selecionar mais variedades a partir da combinação de progenitores masculino e feminino estabelecida.

c) Seleção na progênie dos melhores indivíduos de acordo com as características desejadas.

Este processo tem duração de vários anos, pois para que uma planta de cana-de-açúcar se torne cultivada é necessário que ela possua um conjunto de características agronômicas (aproximadamente 40), tais como resistência a várias doenças, alto teor de sacarose, bom perfilhamento, boa plantabilidade etc. As avaliações de todas estas características são processos extremamente laboriosos e que devem ser executados com grande precisão e repetidos por vários anos. Sendo assim, o processo de avaliação e seleção na progênie dos indivíduos superiores é realizado ano a ano (cada ano para um conjunto de características) até a obtenção da(s) melhor(es) variedade(s);

d) Registro da variedade selecionada.

Este processo é importante, pois é o que assegura aos pesquisadores ou instituições de pesquisa em melhoramento o direito de explorar comercialmente as novas variedades desenvolvidas (Bernet *et al.*, 2003), garantindo a principal fonte de renda à estas instituições pelo recebimento de royalties pagos pelo produtor de acordo com a área ocupada pela nova variedade em questão.

No Brasil, a proteção varietal é garantida pela Lei de Proteção de Cultivares (Lei Nº 9.456), promulgada em 25 de abril de 1997. Por esta lei, o Serviço Nacional de Proteção de Cultivares (SNPC), órgão criado junto ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), é o responsável por garantir (concedendo e fiscalizando) a proteção das novas variedades de plantas desenvolvidas e cultivadas no país (para todas as culturas). Desde sua criação, foram protegidas pelo SNPC variedades 6810 variedades de 47 espécies diferentes. De cana-de-açúcar (*Saccharum spp.*), foram registradas 77 variedades até o momento [Tabela 1] (MAPA, 2008).

Tabela 1. Lista das variedades brasileiras de cana-de-açúcar protegidas.

| Nome da Variedade Registrada | Instituição de Pesquisa Responsável | Data Concessão | Validade da Concessão |
|-------------------------------------|--|-----------------------|------------------------------|
| CTC1 | Centro de Tecnologia Canavieira | 22/07/2005 | 22/07/2020 |
| CTC10 | Centro de Tecnologia Canavieira | 18/09/2006 | 18/09/2021 |
| CTC11 | Centro de Tecnologia Canavieira | 18/09/2006 | 18/09/2021 |
| CTC12 | Centro de Tecnologia Canavieira | 18/09/2006 | 18/09/2021 |
| CTC2 | Centro de Tecnologia Canavieira | 22/07/2005 | 22/07/2020 |
| CTC3 | Centro de Tecnologia Canavieira | 22/07/2005 | 22/07/2020 |
| CTC4 | Centro de Tecnologia Canavieira | 22/07/2005 | 22/07/2020 |
| CTC5 | Centro de Tecnologia Canavieira | 22/07/2005 | 22/07/2020 |
| CTC6 | Centro de Tecnologia Canavieira | 28/08/2006 | 28/08/2021 |
| CTC7 | Centro de Tecnologia Canavieira | 28/08/2006 | 28/08/2021 |
| CTC8 | Centro de Tecnologia Canavieira | 28/08/2006 | 28/08/2021 |
| CTC9 | Centro de Tecnologia Canavieira | 28/08/2006 | 28/08/2021 |
| CV32P2008 | CanaVialis S.A. | 27/10/2008 | 27/10/2023 |
| CV37P2008 | CanaVialis S.A. | 27/10/2008 | 27/10/2023 |
| CV41P2008 | CanaVialis S.A. | 17/07/2008 | 17/07/2023 |
| CV42P2008 | CanaVialis S.A. | 27/10/2008 | 27/10/2023 |
| CV49P2008 | CanaVialis S.A. | 27/10/2008 | 27/10/2023 |
| IAC86 2210 | Instituto Agronômico de Campinas | 17/11/1998 | 17/11/2013 |
| IAC87 3396 | Instituto Agronômico de Campinas | 17/11/1998 | 17/11/2013 |
| IACSP 933046 | Instituto Agronômico de Campinas | 04/06/2007 | 04/06/2022 |
| IACSP 942094 | Instituto Agronômico de Campinas | 04/06/2007 | 04/06/2022 |
| IACSP 942101 | Instituto Agronômico de Campinas | 04/06/2007 | 04/06/2022 |
| IACSP 944004 | Instituto Agronômico de Campinas | 04/06/2007 | 04/06/2022 |
| PAV 94-09 | AGROPAV AGROPECUÁRIA LTDA | 16/04/2001 | 16/04/2016 |
| PO88-62 | Usina da Barra S/A Açúcar e Álcool | 20/07/2004 | 20/07/2019 |
| RB835054 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |
| RB842021 | Universidade Federal de Alagoas * | 16/10/2001 | 16/10/2016 |
| RB845197 | Universidade Federal de São Carlos * | 03/09/2002 | 03/09/2017 |
| RB845210 | Universidade Federal de São Carlos * | 03/09/2002 | 03/09/2017 |
| RB845257 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |
| RB8495 | Universidade Federal de Alagoas * | 15/10/2001 | 16/10/2016 |
| RB855035 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |
| RB855036 | Universidade Federal de São Carlos * | 03/09/2002 | 03/09/2017 |
| RB855113 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |
| RB855463 | Universidade Federal de Alagoas * | 16/10/2001 | 16/10/2016 |
| RB855511 | Universidade Federal de Alagoas * | 16/10/2001 | 16/10/2016 |
| RB855536 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |
| RB855546 | Universidade Federal de São Carlos * | 16/04/1999 | 16/04/2014 |

| | | | |
|-----------|--------------------------------------|------------|------------|
| RB865230 | Universidade Federal de São Carlos * | 03/09/2002 | 03/09/2017 |
| RB867515 | Universidade Federal de Viçosa * | 16/04/2001 | 16/04/2016 |
| RB925211 | Universidade Federal de São Carlos * | 10/12/2007 | 10/12/2022 |
| RB925268 | Universidade Federal de São Carlos * | 10/12/2007 | 10/12/2022 |
| RB925345 | Universidade Federal de São Carlos * | 10/12/2007 | 10/12/2022 |
| RB92579 | Universidade Federal de Alagoas * | 23/07/2003 | 23/07/2018 |
| RB928064 | Universidade Federal de Viçosa * | 30/08/2002 | 30/08/2017 |
| RB931530 | Universidade Federal de Alagoas * | 03/03/2004 | 03/03/2019 |
| RB93509 | Universidade Federal de Alagoas * | 23/07/2003 | 23/07/2018 |
| RB935744 | Universidade Federal de São Carlos * | 10/12/2007 | 10/12/2022 |
| SP80-1816 | Centro de Tecnologia Coopersucar | 08/09/1998 | 08/09/2013 |
| SP80-185 | Centro de Tecnologia Coopersucar | 25/08/1998 | 25/08/2013 |
| SP80-3280 | Centro de Tecnologia Coopersucar | 25/08/1998 | 25/08/2013 |
| SP80-3480 | Centro de Tecnologia Coopersucar | 25/08/1998 | 25/08/2013 |
| SP83-2847 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP83-5073 | Centro de Tecnologia Coopersucar | 08/09/1998 | 08/09/2013 |
| SP84-1201 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP84-1431 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP84-2025 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP84-5560 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP85-3877 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP85-5077 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP86-155 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP86-42 | Centro de Tecnologia Coopersucar | 14/04/2000 | 14/04/2015 |
| SP87-344 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP87-365 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP87-396 | Centro de Tecnologia Coopersucar | 26/08/1999 | 26/08/2014 |
| SP89-1115 | Centro de Tecnologia Coopersucar | 27/03/2002 | 27/03/2017 |
| SP90-1107 | Centro de Tecnologia Coopersucar | 23/07/2003 | 23/07/2018 |
| SP90-1161 | Centro de Tecnologia Coopersucar | 23/07/2003 | 23/07/2018 |
| SP90-1638 | Centro de Tecnologia Coopersucar | 27/03/2002 | 27/03/2017 |
| SP90-3414 | Centro de Tecnologia Coopersucar | 27/03/2002 | 27/03/2017 |
| SP90-3723 | Centro de Tecnologia Coopersucar | 16/07/2001 | 16/07/2016 |
| SP91-1049 | Centro de Tecnologia Coopersucar | 16/07/2001 | 16/07/2016 |
| SP91-1285 | Centro de Tecnologia Coopersucar | 11/12/2003 | 11/12/2018 |
| SP91-1397 | Centro de Tecnologia Coopersucar | 11/12/2003 | 11/12/2018 |
| SP91-1597 | Centro de Tecnologia Coopersucar | 11/12/2003 | 11/12/2018 |
| SP91-3011 | Centro de Tecnologia Coopersucar | 11/12/2003 | 11/12/2018 |
| SP91-3440 | Centro de Tecnologia Coopersucar | 11/12/2003 | 11/12/2018 |

* Universidade Federal pertencente ao programa Ridesa.

No caso de uma variedade ser cultivada fora do país, o direito do melhorista é garantido, pois o governo brasileiro é signatário do tratado de 1991 da União Internacional para Proteção de Novas Variedades de Plantas (UPOV). A UPOV é um órgão internacional com sede em Genebra na Suíça criado por uma conferência diplomática no ano de 1961 com o objetivo de desenvolver e regulamentar sistemas efetivos de proteção de variedades de plantas, incentivando o desenvolvimento de novas variedades. Assim, a UPOV propõe mecanismos aos países signatários, regras e direitos para a proteção de cultivares de plantas. Sendo assim, todos os países que fazem parte da UPOV obrigam-se a proteger cultivares brasileiras e, em contrapartida, o Brasil também se obriga a proteger cultivares procedentes desses países, facilitando o intercâmbio de novos materiais gerados pela pesquisa brasileira e internacional. Até setembro de 2008, a UPOV possuía 65 países membros, dos quais o Brasil está incluído desde maio de 1999 (UPOV, 2008).

Há uma série de exigências, tanto da UPOV quanto do SNPC para que se conceda o direito de exploração comercial de uma nova variedade. O ponto fundamental para esta concessão está em conseguir distinguir a variedade em questão de todas as outras variedades já existentes. No Brasil (e na maioria dos outros países) o estudo que garante que a nova variedade é realmente diferente das demais existentes é realizado pelo próprio melhorista (ou instituição de pesquisa), o que significa que a proteção é fundamentada na declaração juramentada. Para a realização desse estudo, tanto SNPC quanto UPOV propõe testes específicos para cada cultura, baseados nos princípios de distingüibilidade, uniformidade e estabilidade. Por estes princípios, estes testes são conhecidos como teste DHE, em referência a Distingüibilidade, Homogeneidade e Estabilidade [ou *DUS test* (distinctiveness, uniformity e stability) em inglês] (CPVO, 2001).

Em relação ao princípio de distingüibilidade, testes são estabelecidos para que se permita a avaliação da nova variedade, diferenciando-a de todas as outras variedades já existentes. O princípio da uniformidade exige que os testes aplicados para distingüibilidade sejam suficientemente uniformes a ponto de

fornecerem sempre os mesmos resultados, independente do material (ou parte), técnica, ambiente e avaliador utilizados. A estabilidade exige que os testes aplicados para distingüibilidade sejam suficientemente estáveis independente do método de propagação da planta, idade ou condição fisiológica. Dessa forma, se uma variedade for identificada com diferente de qualquer outra por conta de uma característica que não é observada por qualquer avaliador, e/ou se manifesta em um ambiente ou condição fisiológica específica, e/ou depende de uma técnica (ou método) que não pode ser utilizado por qualquer outro grupo de interesse, não se pode afirmar que a variedade em questão é realmente distinta e portanto, não se pode conceder o direito de exploração comercial ao melhorista. Todos estes princípios são básicos para qualquer cultura e os testes são definidos de maneira específica para cada cultura com a aprovação do órgão local e/ou da UPOV.

De maneira geral, os testes DHE aplicados em processos de proteção varietal são baseados na avaliação de caracteres morfológicos e fisiológicos chamados de descritores e são realizados com a observação da variedade que se deseja registrar juntamente com a observação de variedades registradas anteriormente (Gallacher, 1997). Pela comparação dos descritores tenta-se encontrar características morfológicas e/ou fisiológicas específicas para cada variedade, sendo capazes de diferenciar a nova variedade das demais.

Como descritores morfológicos e fisiológicos são em sua maioria influenciados por fatores ambientais e nem sempre passíveis de avaliação biométrica (altura da planta, por exemplo), a avaliação desses descritores é freqüentemente subjetiva, além de laboriosa, demorada e cara (Giancola *et al.*, 2002). Geralmente o impacto da interação do ambiente nas características morfofisiológicas das plantas pode ser diminuído pela avaliação dos descritores em diversos ambientes e com várias repetições, exigindo grande disponibilidade de área e pessoal. Os problemas citados na avaliação de descritores morfológicos e fisiológicos constituem o principal problema no estabelecimento de métodos que atendam os princípios de estabilidade e uniformidade.

Para a cultura da cana-de-açúcar, o teste DHE para registro de uma nova variedade é baseada em critérios morfológicos. No Brasil, o SNPC publicou um documento (conhecido como “Descritores Botânicos de Cana-de-açúcar”) no Diário Oficial da União do dia 05/03/98 (nº 43 – Seção I, páginas 95 a 98) estabelecendo quais características devem ser avaliadas para registro varietal. Este documento serviu como base de outro documento (conhecido como “TG186-1 – Sugarcane Guideline”), dessa vez para ser usado como referência na UPOV. As características avaliadas nos dois documentos baseiam-se nos trabalhos do botânico Dillewijn, com o livro “Botany of Sugarcane” (Landell e Bressiani, 2008).

Apesar do método de caracterização varietal (baseado em descritores botânicos) ser empregado nos processos que envolvem direitos de propriedade intelectual (como no caso de registro de uma nova variedade), além de ser utilizado nos trabalhos de caracterização taxonômica e de gerenciamento de bancos de germoplasma, o uso exclusivo de tais descritores não atende todas as exigências dos princípios de distingüibilidade, uniformidade e estabilidade citados, sendo este um problema para praticamente todas as culturas (UPOV, 2005).

Descritores morfológicos não possuem, na maioria das vezes, variação (polimorfismo) suficiente para distinguir diferentes plantas (uma variedade ou acesso de um banco de germoplasma). Além disso, como já citado, descritores botânicos são sensíveis a mudanças ambientais e/ou fisiológicas, sendo instáveis e não uniformes.

A UPOV em diversos documentos relata esta preocupação na intenção de procurar alternativas a instabilidade e na laboriosidade dos descritores morfofisiológicos. Desta forma, diversos documentos tem sido publicados pela UPOV, sugerindo como complementação aos descritores morfofisiológicos o uso de descritores (ou marcadores) moleculares (UPOV, 2006). Estes marcadores são comumente utilizados em aplicações forenses em humanos, como por exemplo, testes de paternidade (Schumm *et al.*, 1998), mostrando-se adequados para as aplicações requeridas pela UPOV e SNPC e para as demandas citadas dentro dos programas de melhoramento.

1.3. MARCADORES MOLECULARES APLICADOS NO DESENVOLVIMENTO DE NOVAS VARIEDADES

Qualquer molécula (metabólitos, proteínas, seqüências de DNA e RNA etc) que apresente variações em quantidade ou em estrutura entre indivíduos é considerada um marcador molecular. Pelas variações que estas moléculas apresentam entre indivíduos, o fenótipo molecular oriundo destas moléculas será específico para um único indivíduo ou grupo, sendo portando, marcadores deste indivíduo ou grupo.

Dentre as moléculas que podem ser utilizadas como marcadores moleculares, o DNA é a que apresenta maior fonte de variação entre indivíduos (Caixeta *et al.*, 2006). Além disso, existem atualmente diversas técnicas para detecção destas variações, o que facilita a disseminação do uso destes marcadores.

Existem diversos tipos de marcadores moleculares de DNA. Cada marcador possui vantagens e desvantagens em seu uso, dependendo da aplicação que se deseja ou do investimento financeiro que possa dispor. Os principais tipos de marcadores de DNA existentes atualmente são os marcadores RAPD (Random Amplified Polymorphic DNA), RFLP (Restriction Fragment Length Polymorphism), AFLP (Amplified Fragment Length Polymorphism), microssatélites ou SSR (Simple Sequence Repeats) e SNP (Single-Nucleotide Polymorphism). O tipo ou a base de variação (polimorfismo) pode ser diferente entre cada marcador citado. Conseqüentemente os sistemas de detecção também serão diferentes. Para aplicações de identificação genética, os marcadores que mais se destacam são os microssatélites, amplamente utilizados em diversos organismos. Estes marcadores se destacam principalmente por apresentarem alta reprodutibilidade, além de apresentarem baixo custo de aplicação e de poderem ser utilizados em sistemas automatizados (ou semi-automatizados) de alta capacidade de amostras (tabela 2).

Tabela 2. Comparação entre principais tipos de marcadores moleculares.

| CARACTERÍSTICAS | PRINCIPAIS TIPOS DE MARCADORES MOLECULARES DE DNA | | | | |
|--|---|--------------------|--------------------|-----------------------------------|---------------------------|
| | RAPD | RFLP | AFLP | Microsatélite | SNP |
| Base do polimorfismo | sítio de anelamento de <i>primer</i> | sítio de restrição | sítio de restrição | número de repetições de um motivo | mutação de um nucleotídeo |
| Base do sistema de detecção | PCR | Hibridação | PCR | PCR | PCR ou Hibridação |
| Disponibilidade no genoma | Grande | Grande | Grande | Média | Grande |
| Presença do polimorfismo | Médio | Alto | Médio | Alto | Alto |
| Tipo do marcador | Dominante | Codominante | Dominante | Codominante | Codominante |
| Custos de desenvolvimento | Baixo | Alto | Médio | Médio/Alto | Alto |
| Custo de insumos para análise em 1.000 amostras (R\$) | 2.500,00 | 5.000,00 | 10.000,00 | 3.000,00 | 2.500,00 - 15.000,00 * |
| Nº de pessoas para aplicação em rotina de 1.000/mês | 2 | 4-5 | 4 | 2 | 1 |
| Reprodutibilidade | Baixa | Alta | Média/Alta | Alta | Alta |
| Automação | Baixa | Baixa | Média | Média/Alta | Alta |

* Dependendo da técnica que for utilizada na detecção. As técnicas mais modernas apresentam baixo custo nos insumos, mas exigem grandes investimentos em aquisição de equipamentos.

Microssatélites foram inicialmente descritos em humanos (Litt, 1989) e subsequentemente em outras espécies de mamíferos, como ratos (Love *et al.*, 1990), porcos (Johansson *et al.*, 1992) e bois (Kemp *et al.*, 1993). Em plantas os primeiros microssatélites foram identificados em arroz (Wu e Tanksley, 1993), cevada (Saghai *et al.*, 1994), trigo (Roder *et al.*, 1995) e eucalipto (Byrne *et al.*, 1996).

Microssatélites são regiões genômicas (*locus*) formados por unidades ou blocos compostos de 1 a 6 nucleotídeos (chamados de motivos) repetidos em seqüência e variando em tamanho entre indivíduos. Quando o motivo (ou unidade de repetição) é composto de um nucleotídeo, o marcador é chamado de mononucleotídeo. A repetição de motivos formados por dois nucleotídeos forma um microssatélite do tipo dinucleotídeos; a repetição de motivos formados por três, quatro, cinco e seis nucleotídeos formam microssatélites do tipo tri, tetra, penta e hexanucleotídeos, respectivamente. Microssatélites podem ser detectados pelo uso da técnica PCR (Polymerase Chain Reaction) e visualizados por diferentes métodos cromatográficos (Gupta *et al.*, 1996) O conjunto de fragmentos (ou alelos) detectado para um *locus* é chamado de haplótipo. Um ou mais haplótipos produzem um perfil molecular, podendo ser interpretados como códigos de barra biológicos e podendo gerar perfis exclusivos para cada indivíduo (figura 20).

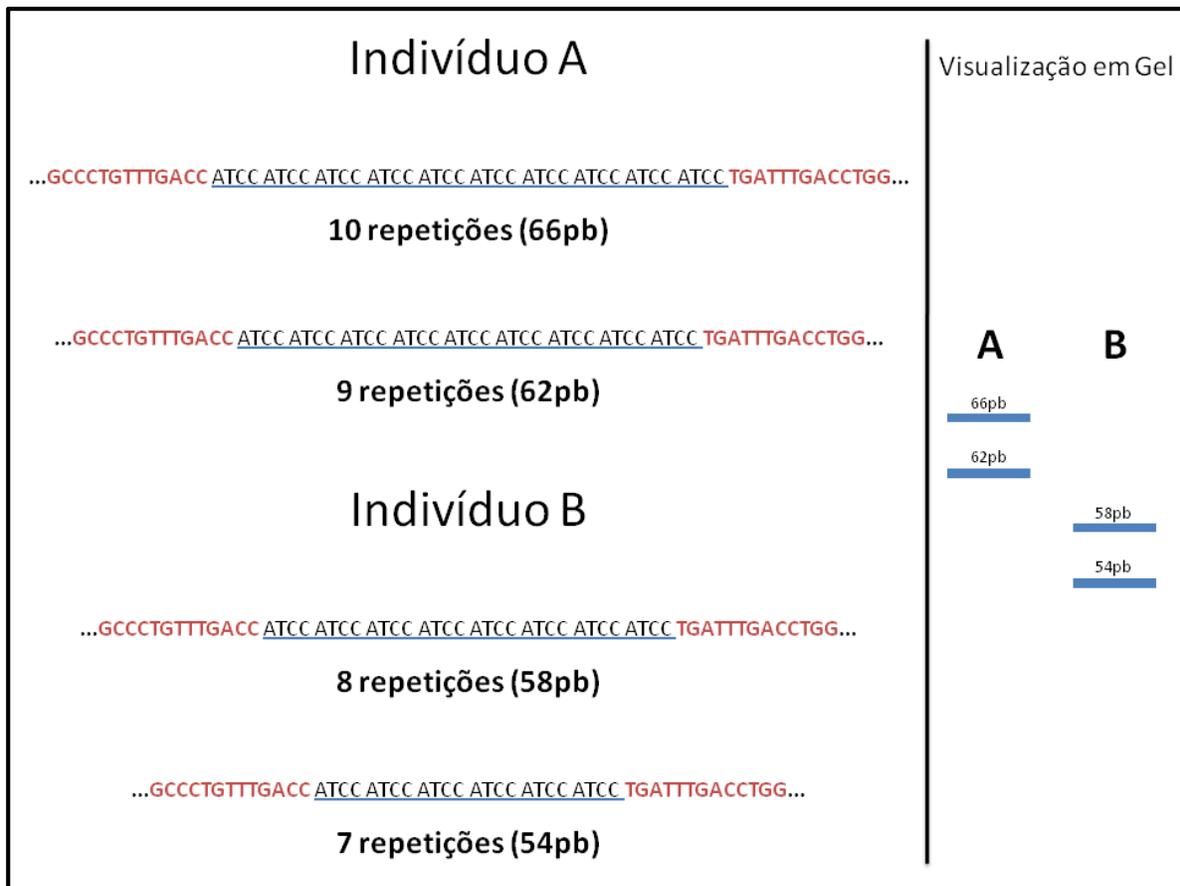


Figura 20. Esquema para representar um microssatélite utilizando um organismo diplóide como exemplo. Dois indivíduos são apresentados. O motivo do microssatélite mostrado é a seqüência ATCC (tetranucleotídeo) e a variação no número de repetições deste motivo corresponde ao microssatélite. Os nucleotídeos em vermelho na seqüência correspondem aos *primers* desenhados para amplificar o microssatélite pela PCR. À direita, uma representação de como seria a visualização dos fragmentos amplificados em um sistema cromatográfico em gel. O conjunto de duas bandas em azul (alelos) para cada indivíduo corresponde a um haplótipo (perfil molecular) para um indivíduo.

O número de alelos para cada indivíduo varia, de acordo com a ploidia do organismo e dose do alelo. O número máximo de alelos que pode ser visualizado para um *locus* de microssatélite é igual à ploidia do organismo (ou ao

número de cópias de cada cromossomo básico que um indivíduo apresenta). Em indivíduos diplóides, pode-se visualizar um ou dois alelos, dependendo se o indivíduo é homocigoto ou heterocigoto, respectivamente. Quando um indivíduo é homocigoto, as duas cópias do *locus* que contém o microssatélite possuem o mesmo número de repetições e, portanto o alelo (fragmento de DNA) amplificado apresentará o mesmo tamanho.

Em um organismo tetraplóide, o número máximo de alelos que pode ser visualizado por *locus* é quatro. Em cana-de-açúcar, como o número cromossômico é variável entre indivíduos, por conta da hibridação de duas espécies de número cromossômico diferente e da aneuploidia, é difícil estabelecer um número máximo de alelos que podem ser visualizados por *locus* (figura 21).

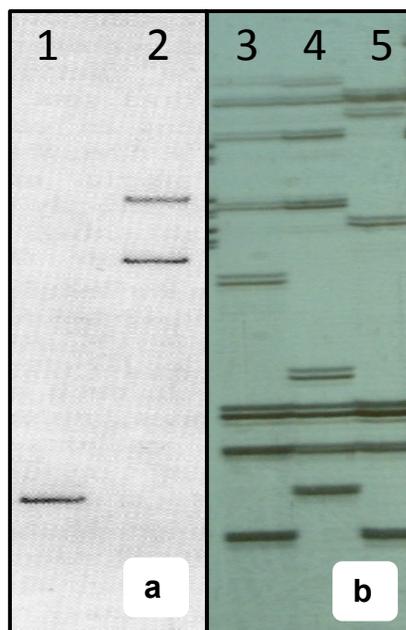


Figura 21. *Locus* de microssatélite para dois indivíduos diplóides [humanos] (a) e três indivíduos poliplóides [cana-de-açúcar] (b). O primeiro indivíduo (1) é homocigoto, com haplótipo formado por alelos de mesmo tamanho (visualização de uma única banda). O segundo indivíduo (2) é heterocigoto, com haplótipo formado por alelos de tamanho diferente (visualização de duas bandas). Os indivíduos (3) e (4) apresentam oito alelos e o indivíduo (5) apresenta seis alelos.

Em programas de melhoramento de plantas, marcadores moleculares podem ser utilizados de duas formas:

- a) Mapeamento genético: associação de características morfofisiológicas e agronômicas com marcadores moleculares. Neste caso, os marcadores moleculares servirão como ferramentas de apoio ao melhorista para seleção de progenitores e/ou de indivíduos superiores na progênie (Kole, 2007);
- b) Identificação genética (*fingerprinting*): a identificação varietal é um dos pontos críticos no processo de desenvolvimento de variedades em programas de melhoramento genético de plantas (Cooke, 1995). Como ferramenta auxiliar aos processos de identificação varietal tradicionais (baseado em descritores morfofisiológicos), marcadores moleculares são ferramentas importantes podem ser utilizados para ajudar os melhoristas em todas as etapas do programa de melhoramento, além de auxiliar no registro das variedades como citado anteriormente.

O controle da identidade dos acessos que compõem os bancos de germoplasma é tão importante quanto o número e a diversidade dos acessos da coleção. O gerenciamento correto da identidade dos acessos minimiza possíveis erros no momento da hibridação e/ou da escolha do progenitor (Berding e Roach, 1987); reduz custos de manutenção de germoplasma pela eliminação de acessos duplicados (Naidu e Sreenivasan, 1987); e ajuda na viabilização da permuta de germoplasma entre os programas de melhoramento no mundo (Gallacher, 1997).

A manutenção da identidade correta de uma planta é importante também nas fases de avaliação, propagação e seleção de indivíduos superiores (Roach, 1984), além auxiliar o melhorista na identificação de híbridos, caso a progênie seja proveniente da introgressão de espécies ou gêneros diferentes (Roach, 1977) e no registro de variedades (como já citado). Em policruzamentos, a determinação da paternidade é possível com a utilização de um sistema baseado em marcadores moleculares de DNA (Wang *et al.*, 2002).

Para a cana-de-açúcar, dentro de programas de melhoramento genético, vários pesquisadores em todo mundo tentaram estabelecer processos de identificação varietal, na intenção de atender as demandas citadas. Todos os processos até então estabelecidos basearam-se na avaliação de descritores botânicos ou morfofisiológicos e nenhum método foi totalmente adotado pelos melhoristas de cana-de-açúcar, seja por conta da instabilidade e/ou subjetividade de alguns descritores, seja por conta da laboriosidade dos processos propostos.

Microssatélites e outros marcadores de DNA têm sido usados com grande sucesso para identificação, classificação e gerenciamento de bancos de germoplasma, além da distinção de variedades, identificação de híbridos verdadeiros em programas de introgressão de características desejáveis e determinação de paternidade (Kirst *et al.*, 2005). Estes marcadores estão sob análise da UPOV (Union for the Protection of New Varieties of Plant) como ferramenta de apoio a discriminação varietal em processo de registro de variedades (Tommasini *et al.*, 2003).

Em cana-de-açúcar, diversos trabalhos já descreveram o desenvolvimento de métodos moleculares, baseados em isoenzimas (Almeida e Crócomo, 1994), marcadores de DNA do tipo RAPD (Huckett e Botha, 2004), RFLP e AFLP (Rodríguez *et al.*, 2003) e microssatélites (Cordeiro, 2001) com potencial de uso em sistemas de identificação varietal. Porém, nenhum dos trabalhos citados é utilizado rotineiramente, já que não foram validados para esse fim. O desenvolvimento de um sistema confiável de identificação varietal, com alto poder discriminatório, estável e uniforme será de grande importância para os programas de melhoramento de cana-de-açúcar no mundo e no futuro, poderá ser um método auxiliar aos descritores morfofisiológicos para formar um novo sistema para registro de cultivares, podendo ser adotado no país, através do SNPC ou mundialmente, através dos membros signatários da UPOV.

2. OBJETIVOS

- a) Desenvolver um processo técnico criterioso para a seleção de marcadores moleculares microssatélites com alta reprodutibilidade, baixo número de artefatos e alto poder discriminatório a partir de um banco de dados de ESTs (Expressed Sequence Tags);

- b) Selecionar e validar marcadores moleculares microssatélites para compor um sistema de identificação de variedades de cana-de-açúcar que possa ser utilizado em processos de registro varietal (de acordo com os critérios de distingüibilidade, estabilidade e uniformidade requeridos pela UPOV) e no apoio das atividades de programas de melhoramento genético de cana-de-açúcar.

3. MATERIAL E MÉTODOS

3.1. OBTENÇÃO DE MATERIAL VEGETAL

Para a realização do trabalho utilizaram-se as seguintes plantas:

- a) 1.205 híbridos interespecíficos e espécies relacionadas ao “Complexo *Saccharum*” obtidos em bancos de germoplasma nacionais e internacionais (tabela 3);
- b) 48 plantas de uma progênie produzida a partir da autofecundação (auto-polinização) de uma variedade comercial de cana-de-açúcar;
- c) 100 plântulas de variedades comerciais brasileiras obtidas a partir da multiplicação *in vitro* em laboratório de cultura de tecidos;
- d) 30 plantas transgênicas de variedades comerciais brasileiras e internacionais regeneradas a partir de calo;
- e) 30 plantas propagadas de maneira convencional (vegetativamente) que nunca passaram por nenhum método de multiplicação *in vitro* provenientes de diferentes fazendas da região central, noroeste, sudeste do Estado de São Paulo.

Tabela 3. Relação dos 1.205 acessos de germoplasma utilizados no estudo distribuídos de acordo com o programa de melhoramento de origem e o respectivo país no caso de variedades híbridas e pelo nome no caso de espécies relacionadas.

| Sigla do programa de melhoramento/nome da espécie | Número de acessos | País |
|--|--------------------------|-----------------|
| B | 1 | Barbados |
| CB | 3 | Brasil |
| CC | 2 | Colômbia |
| Co | 5 | Índia |
| CP | 10 | EUA |
| CV | 320 | Brasil |
| IAC/SP | 19 | Brasil |
| LCP | 1 | EUA |
| MZC | 1 | Cuba |
| NA | 1 | Argentina |
| PAV | 1 | Brasil |
| PO | 4 | Brasil |
| PR | 1 | Porto Rico |
| Q | 4 | Austrália |
| RB | 514 | Brasil |
| SP | 298 | Brasil |
| TUC | 1 | Argentina |
| VAT | 3 | Brasil |
| <i>Saccharum officinarum</i> | 11 | - |
| <i>S. spontaneum</i> | 1 | - |
| <i>S. barberi</i> | 1 | - |
| <i>S. sinense</i> | 1 | - |
| <i>S. edule</i> | 1 | - |
| <i>Erianthus bengalensis</i> | 1 | - |
| Total | 1.205 | 9 países |

Realizou-se a coleta de amostras de folha para todas as plantas utilizadas no estudo. Para as plantas provenientes de diferentes locais do Estado de São Paulo realizou-se a coleta de amostras de gemas e raízes, além das amostras de folha. Para padronização e uniformidade dos resultados estabeleceu-se um procedimento de coleta de material vegetal. Para o procedimento de coleta de folhas, deu-se preferência para coleta da folha -1 (figura 22) com a remoção da nervura central. Para a coleta de gemas deu-se preferência para gemas pertencentes aos internódios superiores e para a coleta de raízes, deu-se preferência para as raízes secundárias. Todo material vegetal coletado foi armazenado em tubo ou saco plástico identificado com código de barras. A identificação com código de barras se fez necessária para evitar erros de identificação durante o processamento devido ao grande número de amostras. Até o momento do processamento do material vegetal, todas as amostras foram mantidas em câmara fria a 4°C em caso de processamento em até três dias após a coleta ou ultrafreezer a -80°C em caso de processamento após três dias de coleta.

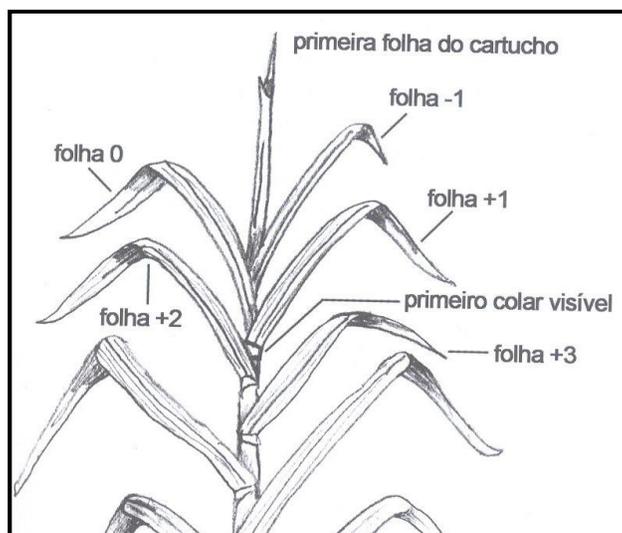


Figura 22. Ilustração das folhas mais jovens de uma planta de cana-de-açúcar com detalhe para a classificação das folhas pelo sistema Kuijper (Casagrande, 1991).

Fonte: Segatto *et al.*, 2006

3.2. EXTRAÇÃO DE DNA

DNA total foi extraído a partir de 100mg de tecido vegetal (folhas, gemas ou raízes). O material vegetal de cada amostra foi macerado em nitrogênio líquido utilizando o equipamento Freezer-Mill 6750 (SPEX Certiprep®). Em microtubos de 2mL adicionou-se aproximadamente 100 mg de tecido macerado e 700 µL de tampão de extração (100 mM Tris-HCl pH8, 20 mM EDTA, 2% CTAB, 1% PVP 40, 1,4 M NaCl, 0,3% 2-mercaptoetanol, 50 µg RNase A). A suspensão foi misturada vigorosamente em vortex e incubada a 65°C por 30 minutos. À mistura adicionou-se 800 µL de clorofórmio/álcool isoamílico (24:1). Após centrifugação (16.000 g por 10 minutos a temperatura ambiente) recuperou-se aproximadamente 450 µL de sobrenadante que foi transferido a um novo microtubo contendo 60 µL de solução de extração (10% CTAB, 1,4 M NaCl). À mistura adicionou-se 500 µL de clorofórmio/álcool isoamílico (24:1) e aproximadamente 400 µL sobrenadante foi recuperado após centrifugação nas condições descritas acima. Ao sobrenadante em um novo microtubo adicionou-se 400 µL de álcool isoamílico. Após agitação por inversão recuperou-se o *pellet* de DNA por centrifugação (16.000 g por 5 minutos a temperatura ambiente). O *pellet* foi lavado com 1 mL de etanol 70%, seco a temperatura ambiente e solubilizado em 50 µL de tampão TE (10 mM Tris-HCl pH8,0, 1 mM EDTA). A concentração de DNA de cada amostra foi quantificada em espectrofotômetro NanoDrop™ 8000 (www.nanodrop.com) e uma solução padrão foi ajustada para 10 ng/µL.

3.3. AMPLIFICAÇÃO DE DNA PELA PCR

Utilizando a técnica da Reação da Polimerase em Cadeira (PCR) realizou-se a amplificação de marcadores SSR em reações contendo 2 µL de DNA (20 ng de DNA), 4 µL de mix de reação [50 unidades/mL *Taq* DNA Polimerase (GoTaq® Flexi DNA Polymerase, catálogo M8295, Promega), 400 µM dATP, 400 µM dCTP, 400 µM dGTP, 400 µM dTTP (dNTP Set, catálogo R0182, Fermentas

Life Sciences), 100 mM Tris-HCl, pH9,0, 100 mM NaCl, 5 mM MgCl₂, 0,2% Igepal [IGEPAAL® CA-630, catálogo 9036-19-5, Sigma-Aldrich)], 0,5 µL de cada *primer* (10 µM) e 3 µL de água. No caso de análise dos marcadores SSR em seqüenciador de DNA, utilizou-se *primers forward* marcado com fluoróforos (NED, 6-FAM, PET ou VIC – Applied Biosystems). Amplificação foi realizada em termociclador GeneAmp 9700 (Applied Biosystems) com a seguinte programação: desnaturação inicial de 5 min a 94°C; 35 ciclos de 94°C por 30 segundos, 64°C por 30 segundos e 72°C por 30 segundos; e extensão final de 72°C por 60 minutos

3.4. DETECÇÃO DE FRAGMENTOS

Fragmentos de cada um dos marcadores foram analisados em PAGE (Polyacrylamide Gel Electrophoresis) e em dois sistemas de seqüenciador de DNA automatizados (ABI3730 XL, Applied Biosystems e MegaBACE1000, GE Healthcare Life Sciences).

Para detecção de fragmentos utilizando PAGE, géis de poliacrilamida desnaturantes foram feitos com 6% de poliacrilamida (PlusOne™ Acrylamide PAGE 40%, catálogo 17-1303-01, GE Healthcare Life Sciences) e 7 M uréia em tampão TBE 0.5 X (Sambrook, *et al.*, 1989). Amostras foram preparadas com 6 µL de reação de PCR, 3 µL de formamida e 1 µL de *loading buffer* (1 mM EDTA, 0,01% azul de bromofenol e 0,01% de xileno cianol), desnaturadas a 95°C por 5 minutos e mantidas em gelo até o momento de aplicá-las no gel. Utilizando o sistema de eletroforese Sequi-Gen GT Nucleic Acid Electrophoresis (Biorad) realizou-se pré-corrída por 60 minutos a 120 W. Amostras foram corridas por 3 horas a 90 W e os fragmentos foram visualizados após coloração com prata: o gel foi imerso em solução de fixação (10% etanol, 5% ácido acético) por 10 minutos, lavado com água deionizada, imerso em solução de pré-tratamento (1,5% ácido nítrico) por 3 minutos, lavado com água deionizada, imerso em solução de coloração (0,3% nitrato de prata) por 30 minutos, lavado duas vezes com água deionizada, imerso em solução de revelação (283 mM Na₂CO₃) até que os

fragmentos pudessem ser visualizados (aproximadamente 15-30 minutos). A identificação dos alelos de cada amostra para cada marcador foi feita pela inspeção visual do gel e o tamanho dos fragmentos estimado pela comparação com um padrão de peso molecular de 25 pb (25bp DNA Step Ladder, catálogo G4511, Promega).

Para detecção de fragmentos utilizando sistemas de seqüenciador de DNA utilizaram-se amostras amplificadas com *primers* marcados com fluorescência. A eletroforese das amostras foi feita utilizando dois sistemas diferentes de seqüenciador de DNA automatizados: ABI3730 XL e MegaBACE1000.

Para o seqüenciador de DNA ABI3730 XL a corrida das amostras foi feita em um arranjo de capilares de 50 cm preenchidos com polímero POP7, com tempo de injeção de amostras de 10 segundos a 8,5 kV e corrida de amostras de 2 horas a 8,5 kV e 60°C. Os dados de fragmentos gerados pelo seqüenciador foram visualizados com o software STRAND (Hughes, 1998). A identificação dos alelos de cada amostra foi feita pela inspeção visual e o tamanho dos fragmentos foi estimado pela comparação com o padrão de peso molecular GeneScan™ – 500 LIZ® Size Standard (catálogo 4322682, Applied Biosystems).

Para o seqüenciador de DNA MegaBACE1000 a corrida das amostras foi feita em um arranjo de capilares preenchidos com polímero MegaBACE Long Read Matrix, com tempo de injeção de amostras de 45 segundos a 3 kV e corrida de amostras de 40 minutos a 10 kV e 45°C. Os dados de fragmentos gerados pelo seqüenciador foram visualizados com o software Genetic Profiler v2.2 (GE Healthcare Life Sciences). A identificação dos alelos de cada amostra foi feita pela inspeção visual e o tamanho dos fragmentos foi estimado pela comparação com o padrão de peso molecular MegaBACE ET550-R Size Standards (catálogo 25-6550-01, GE Healthcare Life Sciences) ou pela comparação com o padrão de peso molecular desenvolvido pela empresa CanaVialis S.A. marcado com as fluorescências 6-FAM ou NED e com 11 fragmentos de tamanhos conhecidos.

3.5. BANCO DE DADOS DE SEQÜÊNCIAS DE ESTS DE CANA-DE-AÇÚCAR

Um banco de 352.122 seqüências (*reads*) de ESTs de cana-de-açúcar foi formado a partir das seqüências públicas disponíveis do projeto SUCEST (Vettore *et al.*, 2001) e de seqüências fornecidas pela empresa Alellyx Applied Genomics S.A. (www.allellyx.com.br).

3.6. SELEÇÃO *in silico* DE MARCADORES MICROSSATÉLITES

Utilizando o software MISA (Thiel *et al.*, 2003) buscou-se microssatélites no banco de dados de ESTs de cana-de-açúcar. O programa foi configurado para identificar motivos (unidades de repetição) com um (mono), dois (di), três (tri), quatro (tetra), cinco (penta) e seis (hexa) nucleotídeos em tamanho. O limite de repetição de cada motivo foi de dez repetições para mononucleotídeos, seis para dinucleotídeos e cinco para tetra, penta e hexanucleotídeos. Microssatélites com dois ou mais motivos repetidos separados por até 100 bases foram definidos como motivos compostos.

Os microssatélites identificados foram montados na forma de *clusters* (Telles e Silva, 2001) e cada *cluster* classificado pelo tipo de motivo (mono, di, tri, tetra, penta e hexanuclotídeo). Uma seleção de marcadores *in silico* foi realizada separadamente para cada tipo de motivo de acordo com três critérios:

a) *Clusters* formados por seqüências diferenciadas pelo número de repetições de cada motivo (polimorfismo *in silico*). Deu-se preferência para a seleção de marcadores em *clusters* que apresentaram maior número de “alelos *in silico*”, ou seja, número de repetições variável entre seqüências que compõem um *cluster*;

b) *Clusters* formados por seqüências com alto número de unidades de repetição (polimorfismo potencial);

c) Motivos do tipo tri, tetra, penta, hexa e hepta nucleotídeos, já que experimentalmente esses tipos de motivos apresentam menor quantidade de artefatos (Ewen *et al.*, 2000).

3.7. DESENHO DE *PRIMERS*

Um par de *primers* foi desenhado para cada *cluster* selecionado utilizando o software *Primer3* (Rozen e Skaletsky, 1998). O programa foi configurado para identificar nas seqüências de *clusters* selecionados pares de *primers* de tamanho entre 19-22 bases com temperatura de anelamento próxima de 64°C, posicionados aos flancos de cada motivo com pelo menos 5 bases de distância do motivo amplificando fragmentos de tamanho entre 100 e 250 pb.

3.8. PROCESSO DE VALIDAÇÃO DOS MARCADORES MICROSSATÉLITES

O processo de validação dos marcadores microssatélites selecionados *in silico* foi realizado em três etapas (as duas primeiras etapas foram eliminatórias e a última etapa classificatória).

Na primeira etapa realizou-se a amplificação pela PCR de quatro acessos de cana-de-açúcar com todos os pares de *primers* selecionados *in silico*. Os produtos amplificados para cada jogo de *primer* foram avaliados em PAGE (Polyacrylamide Gel Electrophoresis). Os acessos de cana-de-açúcar utilizados foram Caiana (*S. officinarum*), Q136 (variedade australiana) e as variedades brasileiras RB835054 e RB855036. Os padrões eletroforéticos obtidos foram analisados visualmente e os pares de *primers* foram avaliados e selecionados de acordo com dois critérios:

- a) Resolução: selecionou-se marcadores que apresentaram fragmentos de tamanho esperado, boa intensidade e baixo número de número

de artefatos do tipo *stutters* (Miller e Yuan, 1997), permitindo fácil identificação dos alelos presentes em cada acesso;

- b) Polimorfismo: selecionaram-se marcadores que fossem capazes de discriminar os quatro acessos testados.

Na segunda etapa realizou-se a amplificação pela PCR de vinte e sete acessos (tabela 4) de cana-de-açúcar com os pares de *primers* dos marcadores selecionados na etapa anterior, com avaliação dos produtos amplificados para cada jogo de *primer* em PAGE e em Seqüenciador de DNA (ABI3730 XL, Applied Biosystems).

Tabela 4. Lista dos 27 acessos de cana-de-açúcar avaliados na segunda fase do processo de validação com indicação do *pedigree* correspondente.

| Acessos* | Progenitor Feminino | Progenitor Masculino |
|---|----------------------------|-----------------------------|
| Caiana (<i>S. officinarum</i>) | Desconhecido | Desconhecido |
| Co740 | P3247 | P4775 |
| IAC86-2210 | CP52-48 | Co798 |
| Olhuda (<i>S. officinarum</i>) | Desconhecido | Desconhecido |
| Preta Kavangire (<i>S. officinarum</i>) | Desconhecido | Desconhecido n |
| Q136 | NCo310 | 54N7096 |
| RB72454 | CP53-76 | Desconhecido |
| RB835054 | RB72454 | NA56-79 |
| RB835486 | L60-14 | Desconhecido |
| RB855036 | RB72454 | SP70-1143 |
| RB855453 | TUC71-7 | Desconhecido |
| RB855536 | RB72454 | SP70-1143 |
| RB867515 | RB72454 | Desconhecido |
| RB925345 | H59-1966 | Desconhecido |
| RB928064 | SP70-1143 | Desconhecido |
| S1 * | RB835054 | RB855036 |
| S2 * | RB835054 | RB855036 |
| S3 * | RB835054 | RB855036 |
| S4 * | RB835054 | RB855036 |
| S5 * | RB835054 | RB855036 |

| | | |
|-----------|-----------|--------------|
| S6 * | RB835054 | RB855036 |
| S7 * | RB835054 | RB855036 |
| S8 * | RB835054 | RB855036 |
| SP80-1842 | SP71-1088 | H57-5028 |
| SP81-3250 | CP70-1547 | SP71-1279 |
| SP83-2847 | HJ5741 | SP70-1143 |
| SP89-1115 | CP73-1547 | Desconhecido |

* Acessos nomeados de S1 a S8 correspondem a plantas de uma progênie F1 do cruzamento entre as variedades meio-irmãs RB835054 e RB855036.

Os padrões eletroforéticos obtidos foram analisados visualmente e os pares de *primers* foram avaliados e selecionados de acordo com os mesmos critérios da primeira etapa, mas com avaliações mais rigorosas:

- a) Resolução: selecionou-se marcadores que apresentaram fragmentos de boa intensidade, fácil reconhecimento de artefatos do tipo stutters em seqüenciador de DNA e capacidade de automatização a partir de softwares específicos genotipagem;
- b) Polimorfismo: identificou-se os alelos de cada acesso (visualizados como bandas em PAGE ou picos em seqüenciador de DNA) anotando 1 (quando presentes) ou 0 (quando ausentes). Esse processo de identificação e anotação da presença ou ausência dos alelos foi realizado manualmente e por duas pessoas separadas com o objetivo de diminuir erros na identificação de alelos. Com essas informações construiu-se uma matriz binária para cada marcador e realizou-se o cálculo do Conteúdo de Informação Polimórfica ou “Polymorphism Information Content” (PIC). O PIC é um indicador da qualidade do marcador em estudos genéticos. Marcadores com valores superiores a 0,5 são considerados muito informativos, com valores entre 0,25 e 0,50 mediamente informativos, e com valores inferiores a 0,25, pouco informativos (Botstein *et al.*, 1980).

Utilizando-se essa classificação calculou-se o PIC de cada marcador de acordo com a fórmula $PIC_i = 1 - \sum_j P_{ij}^2$ (Weber, 1990) onde P_{ij} é a frequência do alelo j no genótipo i .

Na terceira etapa realizou-se a amplificação pela PCR de 1.205 acessos de cana-de-açúcar (indicados na tabela 1) com os pares de *primers* dos marcadores selecionados na etapa anterior e avaliou-se os produtos amplificados em Seqüenciador de DNA (ABI3730 XL, Applied Biosystems). Os padrões eletroforéticos obtidos foram analisados visualmente e os pares de *primers* foram classificados, de acordo com os critérios estabelecidos nas etapas anteriores:

- a) Resolução: como esse critério de seleção foi realizado nas duas primeiras etapas, os marcadores avaliados nessa etapa apresentaram excelente resolução. Mesmo assim, pela análise de uma grande quantidade de amostras, procurou-se caracterizar os alelos para cada marcador na intenção de identificar alelos de baixa amplificação em relação a outros alelos de um mesmo *locus*. Esse tipo de artefato, conhecido como alelo fraco (ou *weak allele*) pode comprometer a identificação correta dos alelos por ser de difícil detecção, já que exige análise em repetição (Hoffman e Amos, 2004).
- b) Polimorfismo: assim como na etapa anterior, realizou-se o cálculo do PIC dos marcadores selecionados com os dados de 1.205 acessos. Além do PIC, realizou-se o cálculo de probabilidade de identidade genética (Pid). A Pid é definida como a probabilidade de dois indivíduos selecionados aleatoriamente apresentarem o mesmo genótipo (Chakravarty e Li, 1983) e foi calculada usando a fórmula $Pid = \sum \chi^2$, onde χ é a frequência genotípica (Paetkau *et al.*, 1995). Quanto menor essa probabilidade, maior o poder discriminatório de um *locus* ou de um conjunto de *loci*.

4. RESULTADOS E DISCUSSÃO

4.1. DESENVOLVIMENTO DE UM PROCESSO DE SELEÇÃO DE MARCADORES MICROSSATÉLITES A PARTIR DE UM BANCO DE DADOS DE ESTs

O processo de seleção de marcadores proposto está dividido em três etapas: seleção *in silico* de marcadores a partir do banco de ESTs; avaliação e seleção dos marcadores selecionados na etapa anterior utilizando uma coleção de acessos de cana-de-açúcar e espécies relacionadas; e caracterização dos melhores marcadores. Em todas as etapas, o foco está voltado para a identificação de marcadores de boa resolução (garantindo a estabilidade e a reprodutibilidade do sistema) e polimórficos (garantindo o alto poder discriminatório).

4.1.1. Seleção *in silico* de marcadores microssatélites candidatos

Um banco de dados de 352.122 seqüências formado de ESTs de cana-de-açúcar foi utilizado para busca de microssatélites. A utilização exclusiva de seqüências expressas se deu pela disponibilidade imediata destes dados a partir dos bancos públicos (NCBI) e privado (Alellyx Applied Genomics). Dados de ESTs de cana-de-açúcar têm sido descritos como uma excelente fonte de microssatélite (Pinto *et al.*, 2004; Pinto *et al.*, 2006; Cordeiro *et al.*, 2001). Este banco combinado de ESTs foi então filtrado por meio do software MISA que identificou 33.324 seqüências (9,46% do total de seqüências analisadas) contendo 36.950 motivos de microssatélites. Este número é compatível com a freqüência verificada de clusters de ESTs que contêm microssatélites em outras plantas, como trigo (7,5%) e sorgo [8,2%] (Varshneya *et al.*, 2003).

As 352.122 seqüências utilizadas foram agrupadas em 79.526 *clusters* dos quais 14.291 continham microssatélites (18%). Desses *clusters*, 10.586 (75%)

apresentaram apenas um motivo e 3.705 (25%) apresentaram dois ou mais. A unidade de repetição mais freqüente encontrada foi trinucleotídeo (48%), seguido por mononucleotídeo (32%) e dinucleotídeo (16%). Tetra, penta e hexanucleotídeo foram identificados em 1.414 seqüências, o que corresponde a menos de 5% dos microssatélites presentes no banco de dados. A abundância de motivos do tipo trinucleotídeo em ESTs pode ser atribuída a fato desta mutação ser mais tolerável do que os demais tipos de SSR por ser uma mutação que não desloca (altera) o quadro de leitura (*frameshift*) em regiões codantes, não alterando os aminoácidos produzidos na proteína [apenas o tamanho da cadeia de aminoácidos] (Varshneya *et al.*, 2003).

Com a grande disponibilidade de motivos para serem selecionados, o uso de microssatélites localizados em regiões expressas torna-se interessante, pois estando regiões expressas tendem a ser mais conservadas do que regiões não expressas, aumentando a possibilidade de transferibilidade e utilização destes marcadores em outras espécies. Considerando que os marcadores selecionados poderão ser aplicados em acessos contidos em bancos de germoplasma de cana-de-açúcar (formado por híbridos interespecíficos, intergenéricos, espécies de *Saccharum* e de outros gêneros) a obtenção de marcadores em regiões conservadas aumenta a probabilidade de uso destes marcadores em todos os acessos que compõem os bancos de germoplasma de cana-de-açúcar.

A tabela 5 resume as estatísticas da filtragem do banco de dados após análise com o software MISA.

Tabela 5. Estatísticas da busca *in silico* de marcadores microsatélites no banco de dados de ESTs de cana-de-açúcar utilizando o software MISA.

| Estatísticas do Banco de Dados | Número total de seqüências | Frequência (%) |
|--|--|------------------------|
| Banco de dados | 352.122 | 100% |
| Seqüências com microsatélites | 33.324 | 9,46% |
| Motivos de microsatélites identificados | 36.950 | - |
| Seqüências contendo mais de um motivo | 2.936 | 0,83% |
| Motivos compostos identificados | 2.394 | - |
| Dados por <i>clusters</i> | Número total de <i>clusters</i> | Frequência (%)* |
| <i>Clusters</i> identificados | 79.526 | 100% |
| <i>Clusters</i> contendo microsatélites | 14.291 | 17,97% (100%) |
| <i>Clusters</i> contendo um microsatélite | 10.586 | 13,31% (75%) |
| <i>Clusters</i> contendo dois microsatélites | 2.185 | 2,74% (15%) |
| <i>Clusters</i> contendo três microsatélites | 737 | 0,93% (5%) |
| <i>Clusters</i> contendo mais de três microsatélites | 783 | 0,98% (5%) |
| Tipo de motivo | Número total de Seqüências | Frequência (%)* |
| Todos os tipos | 36.950 | 100% |
| 1 (mono) | 11.978 | 32,42% |
| 2 (di) | 5.846 | 15,82% |
| 3 (tri) | 17.712 | 47,93% |
| 4 (tetra) | 644 | 1,74% |
| 5 (penta) | 527 | 1,42% |
| 6 (hexa) | 243 | 0,66% |

* Os valores entre parênteses referem-se à distribuição dos microsatélites contidos em *clusters*.

Após a identificação dos *clusters* contendo microssatélites criou-se um novo banco de dados com estes clusters com o objetivo de selecionar os melhores microssatélites candidatos. Ao novo banco de dados adicionaram-se as seguintes informações:

1. nome dos *clusters* para cada seqüência;
2. tamanho da(s) unidade(s) de repetição contida(s) em cada seqüência;
3. número de repetições de cada motivo;
4. seqüência FASTA.

Inicialmente o banco de *clusters* foi ordenado pelo tamanho da unidade de repetição (mono, di, tri, tetra, penta e hexanucleotídeo). Seqüências contendo microssatélites do tipo mono e dinucleotídeo foram descartadas (48%) do banco de dados, já que a resolução em PAGE e seqüenciador de DNA desses tipos de microssatélites é ruim pela dificuldade em separar fragmentos de uma ou duas bases, além do grande número de artefatos que estes marcadores apresentam (Ewen *et al.*, 2000).

O restante dos microssatélites contidos no banco (tri, tetra, penta e hexa) foi agrupado pelo nome do *cluster*. Dois critérios foram utilizados para seleção dos melhores marcadores:

a) *Clusters* compostos por *reads* polimórficos: as seqüências que compõem cada *cluster* foram inspecionadas para identificação do número de repetições de cada motivo. A comparação entre o número de repetições presentes entre as seqüências de um *cluster* revelou o polimorfismo (*in silico*) de cada microssatélite. Cada microssatélite foi ordenado pelo número de polimorfismos *in silico* encontrado (número de repetições diferentes entre as seqüências de um *cluster*). Selecionou-se para este critério os 83 *clusters* com maior número de polimorfismos *in silico*.

b) *Clusters* com grande número de repetição de motivos: cada microssatélite foi ordenado pelo número de repetições

contido nas seqüências de um *cluster*. Selecionaram-se para esse critério os 67 *clusters* com maior número de repetições. *Clusters* contendo motivos com grande número de repetições têm maior probabilidade de apresentarem *reads* com número variável de repetições (polimorfismo potencial) [La Rota *et al.*, 2005].

A tabela 6 resume o resultado da seleção *in silico* de microssatélites.

Tabela 6. Resultado do processo de seleção *in silico* de marcadores microssatélites.

| Tipo de motivo | <i>Clusters</i> selecionados | Critério de seleção | |
|----------------|---------------------------------|-------------------------------|------------------------|
| | | Polimorfismo <i>in silico</i> | Polimorfismo potencial |
| Compostos | 12 | 5 (42%) | 7 (58%) |
| Tri | 75 | 53 (73%) | 22 (27%) |
| Tetra | 37 | 14 (48%) | 23 (52%) |
| Penta | 13 | 4 (31%) | 9 (69%) |
| Hexa | 13 | 7 (54%) | 6 (46%) |
| Total | 150 | 83 (56%) | 67 (44%) |

Dos *clusters* que apresentaram polimorfismo *in silico*, motivos do tipo compostos foram os que apresentaram maior média de *reads* polimórficos entre os tipos de selecionados (5). Isto já era esperado, considerando que microssatélites com motivos compostos possuem pelo menos dois tipos de motivo (ou seja, duas regiões com possibilidade de variação no número de repetições do motivo). Mesmo assim, o motivo do tipo composto foi o menos selecionado, pois dependendo da combinação dos motivos que compõem o microssatélite composto, podem-se obter alelos com variação de uma ou duas bases, sendo de

difícil análise e posterior automação [mesmo caso de mono e dinucleotídeos] (Zhang e Yeung, 1996). Em relação aos demais tipos de motivo, em relação ao critério de polimorfismo *in silico* e de polimorfismo potencial, a seleção de microssatélites seguiu a proporção de distribuição destes motivos no banco de dados.

4.1.2. Avaliação dos marcadores selecionados *in silico* e seleção dos melhores candidatos para compor o sistema de identificação de variedades

A avaliação dos marcadores selecionados *in silico* se deu em duas etapas, seguindo novamente procedimentos para obtenção de marcadores de alta resolução e polimorfismo.

Para a primeira etapa, pares de *primers* foram desenhados para os 150 *clusters* selecionados *in silico* e usados para a amplificação pela PCR de quatro acessos de cana-de-açúcar (Caiana, Q136, RB835054 e RB855036) e análise em PAGE. Optou-se pela utilização de quatro acessos para que nesta primeira triagem fossem identificados rapidamente marcadores monomórficos (sem polimorfismo) ou com baixa qualidade de amplificação na PCR (não amplificação; amplificação de fragmentos de tamanho não esperado; ou amplificação de fragmentos com alta presença de artefatos). A diversidade no conjunto de quatro acessos [com um acesso de *S. officinarum* (Caiana), um de um programa de melhoramento internacional (Q136) e dois de um programa de melhoramento nacional (RB835054 e RB855036)] e o *pedigree* dos acessos RB835054 e RB855036 (variedades “filhas” do mesmo progenitor masculino ou meio-irmãs) foram os motivos para escolha dos quatro acessos avaliados. Esperava-se com esta pequena população conhecer o maior número de alelos possível e identificar marcadores de alto polimorfismo (caso o marcador diferenciasse todos os acessos, inclusive dois acessos de alta similaridade genética). Nesta fase, optou-se pela avaliação dos 150 *loci* em PAGE pelo baixo custo que o sistema exige quando comparado com o custo de desenvolvimento de um seqüenciador de

DNA. Além do mais, uma vez selecionado em PAGE e comprovada sua eficiência, um marcador poderia ser utilizado por outros laboratórios que não dispusessem de um seqüenciador de DNA. Vale salientar que este trabalho se iniciou com a premissa de desenvolver um sistema de identificação de variedades de cana-de-açúcar que fosse público, permitindo assim seu uso pelo SNPC e UPOV no registro e regulamentação de variedades, bem como para intercâmbio de germoplasma entre programas internacionais e nacionais.

De acordo com os critérios de seleção propostos (resolução e polimorfismo) selecionou-se 20 pares de *primers (loci)* para avaliação na segunda etapa. Baixa resolução (baixa qualidade de amplificação pela PCR e/ou presença de artefatos) foi responsável pela eliminação de 63% dos *loci* candidatos (95) e baixo polimorfismo foi responsável por 23% (35). Um total de 20 *loci* (14%) com representantes dos seis tipos de motivos (compostos, tri, tetra, penta e hexa) foram selecionados para a segunda etapa de validação.

A estratégia utilizada no processo de seleção *in silico* de microssatélites onde se excluiu motivos com mono e dinucleotídeos dado a sua relação com artefatos e utilizou-se critérios de busca por polimorfismo *in silico* ou potencial, se mostrou eficiente, pois dos 150 *loci* avaliados, apenas 26% foram eliminados por conta da presença de artefatos (11%) ou baixo número de alelos (15%). O grande número de *loci* descartados por baixa qualidade da PCR (61%), dado à baixa intensidade ou ausência de amplificação, é visto como algo comum para marcadores microssatélites em seqüências de plantas derivadas de ESTs, por conta da presença de íntrons e/ou presença de regiões ricas em GC (Temnykh *et al.*, 2001).

A tabela 7 detalha para cada tipo de microssatélite o critério pelo qual os marcadores avaliados foram eliminados na primeira etapa.

Tabela 7. *Loci* eliminados e selecionados para cada tipo de marcador microssatélite dos 150 avaliados.

| Tipo do motivo (número de <i>loci</i>) | <i>Loci</i> eliminados | | | <i>Loci</i> selecionados |
|--|------------------------|------------------------|------------------|-----------------------------|
| | Baixa qualidade da PCR | Baixo número de alelos | Artefatos de PCR | |
| Compostos (12) | 5 (41%) | 2 (17%) | 3 (25%) | 2 (17%) |
| Tri (75) | 53 (70%) | 11 (15%) | 5 (7%) | 6 (8%) |
| Tetra (37) | 23 (62%) | 4 (10%) | 5 (14%) | 5 (14%) |
| Penta (13) | 3 (23%) | 4 (31%) | 3 (23%) | 3 (23%) |
| Hexa (13) | 7 (53%) | 2 (16%) | 0 (0%) | 4 (31%) |
| Total (150) | 91 (61%) | 23 (15%) | 16 (11%) | 20 (13%) |

A figura 22 mostra exemplos de géis PAGE obtidos nessa etapa a partir de três marcadores eliminados (pelos critérios de baixa qualidade da PCR, baixo número de alelos e polimorfismo e grande quantidade de artefatos) e de um marcador selecionado de acordo com os critérios indicados.

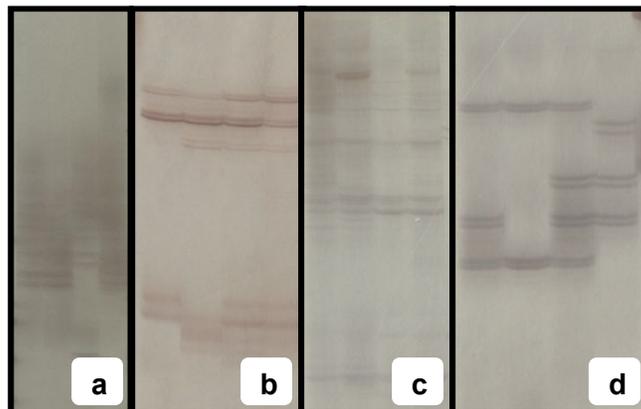


Figura 22. Géis PAGE de três marcadores eliminados e um selecionado de acordo com os critérios de seleção propostos. Em (a) marcador de baixa amplificação, dificultando a identificação dos alelos (eliminado); (b) marcador monomórfico (eliminado); (c) marcador com grande número de artefatos e amplificação desproporcional dos alelos (eliminado); (d) marcador sem artefatos e com número de alelos suficiente para discriminar os quatros acessos avaliados (selecionado).

A segunda etapa do processo de validação de marcadores foi realizada com o objetivo de selecionar os 10 melhores pares de *primers (loci)* para serem usados no sistema de identificação de variedades de cana-de-açúcar nos dois sistemas de detecção avaliados [PAGE e em seqüenciador de DNA (ABI 3730 XL e MegaBACE1000)].

Nesta etapa, o DNA de 27 acessos de cana-de-açúcar (tabela 4) foi amplificado com os 20 pares de *primers (loci)* selecionados. Obteve-se o perfil eletroforético das amostras em PAGE (figura 23) e seqüenciador de DNA (figura 24). Após avaliação visual os marcadores foram classificados (ordenados) de acordo com os critérios de qualidade de amplificação, baixo número de artefatos e poder discriminatório.

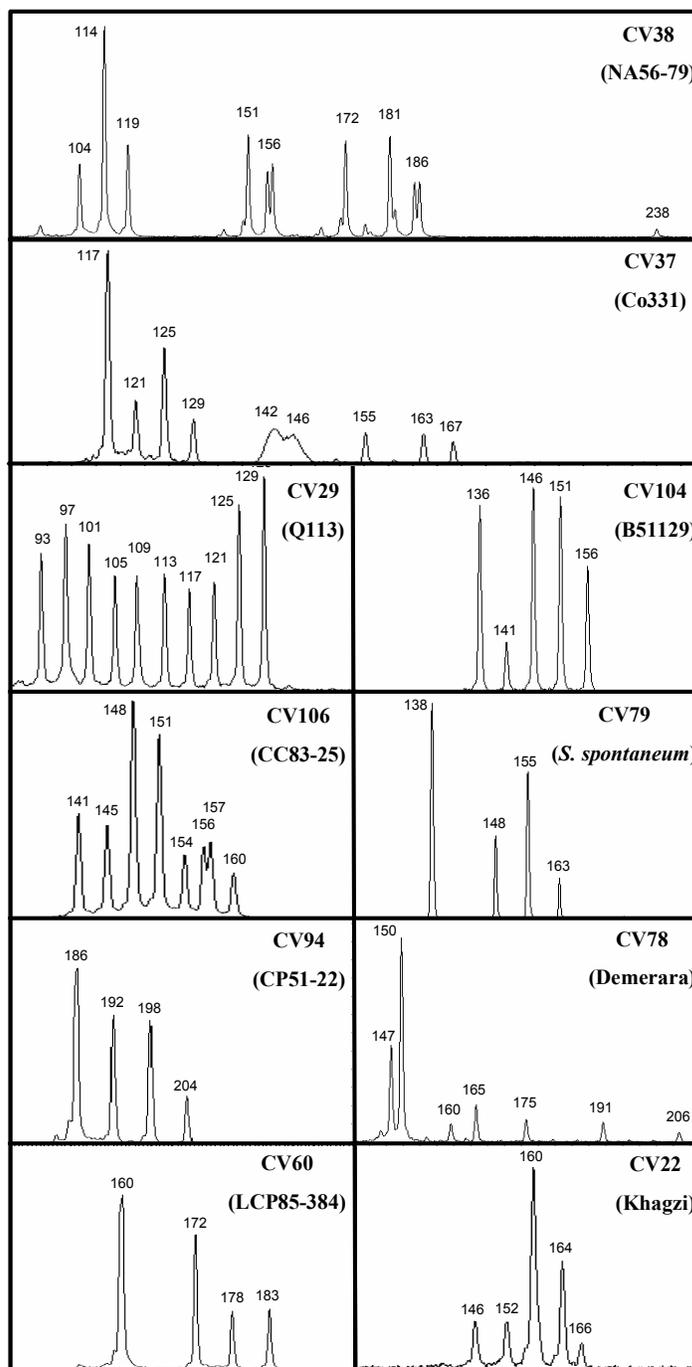


Figura 24. Perfis eletroforéticos dos 10 *loci* avaliados em seqüenciador de DNA (ABI 3730 XL) gerados pelo software STRAND. O número indicado acima de cada pico refere-se ao tamanho em pares de base (pb) de cada alelo. O nome entre parênteses corresponde ao nome do acesso.

Para cada *locus* calculou-se o valor de PIC (tabela 8). O valor do PIC para cada *locus* foi calculado a partir da frequência dos alelos identificados nos 27 acessos de cana-de-açúcar. Considerando que os valores de PIC obtidos para os *loci* avaliados foram altos (0,74-0,91), o critério de resolução dos perfis foi o principal na classificação dos melhores marcadores. Depois que os perfis foram avaliados e classificados, os 10 melhores *loci* foram selecionados para compor o sistema de identificação de variedades de cana-de-açúcar.

Tabela 8. Resultado da validação dos vinte marcadores avaliados.

| <i>Loci</i> | Tipo de Motivo | Critérios de Seleção | | | Status |
|-------------|----------------|----------------------|----------------------|-----------|-----------------|
| | | Amplificação | Presença de artefato | Valor PIC | |
| CV29 | Tetra | ótima | baixo | 0,91 | Selecionado |
| CV38 | Penta | ótima | baixo | 0,90 | Selecionado |
| CV37 | Tetra | ótima | baixo | 0,87 | Selecionado |
| CV106 | Tri | ótima | baixo | 0,83 | Selecionado |
| CV104 | Penta | ótima | baixo | 0,82 | Selecionado |
| CV60 | Hexa | ótima | baixo | 0,76 | Selecionado |
| CV79 | Hexa | ótima | baixo | 0,75 | Selecionado |
| CV94 | Composto | ótima | baixo | 0,74 | Selecionado |
| CV78 | Penta | ótima | médio | 0,89 | Selecionado |
| CV22 | Tetra | ótima | médio | 0,80 | Selecionado |
| CV133 | Tri | ótima | médio | 0,82 | Não selecionado |
| CV98 | Tri | ótima | médio | 0,80 | Não selecionado |
| CV128 | Tri | ótima | médio | 0,79 | Não selecionado |

| | | | | | |
|-------|----------|-----|-------|------|-----------------|
| CV80 | Tetra | boa | baixo | 0,81 | Não selecionado |
| CV86 | Tetra | boa | baixo | 0,80 | Não selecionado |
| CV58 | Hexa | boa | médio | 0,89 | Não selecionado |
| CV140 | Hexa | boa | médio | 0,84 | Não selecionado |
| CV135 | Tri | boa | médio | 0,83 | Não selecionado |
| CV51 | Tri | boa | médio | 0,75 | Não selecionado |
| CV23 | Composto | boa | alto | 0,91 | Não selecionado |

A seleção dos dez marcadores microssatélites para compor o sistema de identificação de variedades priorizou o critério de alta resolução, já que os vinte *loci* avaliados apresentaram alto poder discriminatório (PIC médio = 0,83) e sendo, portanto, altamente informativos. Os valores de PIC do sistema selecionado são compatíveis com outros sistemas (validados e consagrados) de identificação genética de outros organismos, como eucalipto [PIC médio para 6 marcadores = 0,83] (Kirst *et al.*, 2005) e humanos [PIC médio para 16 marcadores = 0,78] (Schumm *et al.*, 1998). Os altos valores de PIC obtidos demonstram mais uma vez a eficiente seleção dos marcadores na fase de seleção *in silico* e na triagem dos 150 *loci*.

Desta forma, na intenção de selecionar marcadores de alta reprodutibilidade (atendendo os princípios de uniformidade e estabilidade requeridos pela UPOV) selecionaram-se os dez marcadores que apresentaram ótima amplificação pela PCR com menor número de artefatos.

4.2. CARACTERIZAÇÃO DOS DEZ MARCADORES SELECIONADOS E VALIDAÇÃO DO SISTEMA DE IDENTIFICAÇÃO DE VARIEDADES DE CANA-DE-AÇÚCAR

4.2.1. Principais características do sistema

Uma coleção de 1.205 acessos de cana-de-açúcar e espécies relacionadas apresentada na tabela 1 (Material e Métodos) foi utilizada para caracterizar em mais detalhes os dez pares de *primers (loci)* selecionados (tabela 9) e para validar sua utilização prática na identificação de variedades. Neste caso, o sistema de detecção utilizado foi o seqüenciador de DNA ABI3730 XL e o seqüenciador de MegaBACE 1000, por se tratarem de sistemas de alta capacidade. O uso de dois modelos de seqüenciador de DNA teve o objetivo de avaliar se o uso de dois modelos de equipamentos fornece resultados consistentes e reprodutíveis. Um total de 150 alelos foi identificado para os dez *loci* analisados (média de 15 alelos por *locus*). A análise foi feita com amostras em duplicata para cada seqüenciador de DNA (para diminuir erros de genotipagem) e os tamanhos de cada fragmento (correspondente a cada alelo) foram estimados com o uso de um padrão de peso molecular.

A seleção dos marcadores em PAGE e em dois modelos de seqüenciador de DNA permitiu a seleção de dez *loci* de grande potencial, tanto do ponto de vista de resolução (boa qualidade de amplificação pela PCR e baixo número de artefatos) como de poder discriminatório (altos valores de PIC). A avaliação dos *loci* em diferentes técnicas (PAGE e Seqüenciador de DNA) permitiu a identificação de um conjunto de marcadores de alta reprodutibilidade independente do sistema de detecção, atendendo uma das exigências do princípio da uniformidade requerido pela UPOV. Foram identificados os mesmos alelos para todos os acessos avaliados nos dois modelos de seqüenciador de DNA. Isto é importante, já que é descrito para outros organismos (principalmente em humanos) há existência de *loci* de microssatélites com alelos de baixa

amplificação que podem ou não ser detectados, por conta da sensibilidade do modelo de seqüenciador de DNA utilizado (Koumi *et al.*, 2004).

Tabela 9. Os 10 *loci* selecionados para compor o sistema de identificação de variedades de cana-de-açúcar.

| <i>Loci</i> | Identificação do <i>Cluster</i> * | Seqüência dos <i>Primers</i> (5'-3') | Fluorôforo |
|-------------|-----------------------------------|--|------------|
| CV22 | AAO24905 ($7e^{-23}$) | Forward: CACTACTCGCCCCGATTTTC Reverse: CGAGTGCTTCTCCATCTGC | PET |
| CV29 | AAM34401 ($8e^{-52}$) | Forward: TCGCGTCCACCAATGTAACC Reverse: GCGTGCATCGCTTGTGTCTT | NED |
| CV37 | AY111193 ($2e^{-06}$) | Forward: GGATGGACGACGTGTCCTGG Reverse: ATAAAGTGGCCGCTTGGATTGA | PET |
| CV38 | NP_001065801 ($3e^{-23}$) | Forward: GAAGCAGGGGCCTCAAGTTG Reverse: GTCAAACAGGCGATCTGGCTC | 6-FAM |
| CV60 | NP_001055708 ($2e^{-37}$) | Forward: AATCTGCACCCTGCCCTCTC Reverse: CAGCTGGAGCATGGATGGAG | NED |
| CV78 | No hits found | Forward: ACGAGGCCACCATAGAACATG Reverse: GCAATTGGGAGGAGAGGAATG | 6-FAM |
| CV79 | NP_001052290 ($3e^{-160}$) | Forward: GGCACTGCTGGTGGTTGATTG Reverse: TCCCACATCAAGAGGCAGCTA | PET |

| | | | | |
|-------|--------------------------------|----------|-----------------------|-------|
| CV94 | No hits found | Forward: | GGCAGGCCAAGATGAATGAAG | NED |
| | | Reverse: | AGCACAGCGGAGGGTACGG | |
| CV104 | CAD32696 (8e ⁻⁴⁵) | Forward: | GATTTTCGACTGTGCGCTTGG | NED |
| | | Reverse: | AAGTTCTCTGCCGAGCAAAC | |
| CV106 | NP_568710 (2e ⁻³⁰) | Forward: | AAACAGAGCATACTCGAGGCC | 6-FAM |
| | | Reverse: | ACGTTGCTGACGAGGTTTTCC | |

* Seqüências de *primers* foram desenhadas baseado na seqüência consenso de cada *cluster*. A identidade de cada *cluster* (indicada entre parênteses) refere-se ao número de acesso da proteína no GenBank identificado pela ferramenta TBLASTX (www.ncbi.nlm.nih.gov/BLAST/) com a maior similaridade (expressa pelo *E-value*) para as seqüências dos *clusters* traduzidos.

A identificação dos alelos que compõem o sistema de dez marcadores, tanto em PAGE quanto em seqüenciador de DNA, foi clara para a maioria dos alelos dos *loci* avaliados. Em PAGE (figura 23) nota-se que a intensidade das bandas geradas pela amplificação e/ou método de coloração com prata dos alelos é constante (algo que foi observado para os dez *loci*). Em seqüenciador de DNA (figura 24), por ser um sistema de detecção mais sensível, verificou-se a presença de picos de diferentes tamanhos, mostrando que existe diferença de amplificação entre alelos de um *locus* para uma mesma amostra.

Por se tratar de um organismo poliplóide, alguns alelos podem estar representados por mais de uma cópia, sendo uma das hipóteses para esta diferença de amplificação entre os alelos de um *locus*. Porém, nota-se que as diferenças nos tamanhos dos picos (amplificação dos alelos) não é proporcional entre os alelos, além de nem sempre está diferença ser reprodutível, na avaliação em duplicata das amostras.

Como os *primers* utilizados para amplificação pela PCR foram desenhados com base na seqüência consenso, variações entre as cópias do cromossomo onde está localizado o marcador (*locus*) podem causar diferenças na especificidade do anelamento dos *primers*, gerando diferenças na amplificação

dos alelos (Primmer *et al*, 1995; Jarne e Lagoda, 1996). Outra hipótese observada para alguns marcadores é a amplificação preferencial por alelos de menor tamanho (Wattier *et al*, 1998). Devido à natureza competitiva da PCR, alelos de menor tamanho freqüentemente amplificam mais eficientemente do que alelos de maior tamanho.

As diferenças de amplificação (intensidade da banda ou tamanho de pico) entre os alelos de um *locus* foram identificadas, na intenção de conhecer as posições de alelos que possam ser fonte de erro na hora de identificação do perfil molecular de uma amostra. Quando um alelo sempre apresenta baixa amplificação, este é conhecido como alelo fraco (*weak allele*). Este tipo de artefato pode ser uma fonte de erros na identificação do perfil molecular de uma amostra, pois pode não ser detectado.

Outro tipo de artefato comum é a presença de falsos alelos, conhecidos como *stutters*. *Stutters* são bandas secundárias geradas por um erro da polimerase durante a PCR. Estas bandas sempre ocupam a posição anterior ou posterior de um alelo verdadeiro amplificado (Miller e Yuan, 1997). Algumas posições apresentam este tipo de artefato em maior freqüência e, portanto, podem ser mapeadas e excluídas da análise de identificação de perfil molecular, evitando erros.

A tabela 10 apresenta a caracterização de alelos para cada *locus*, indicando a freqüência de cada alelo e posições identificadas como problemáticas, por conta da presença de artefatos do tipo *stutter* e/ou alelos fracos.

Tabela 10. Caracterização dos alelos que compõem o sistema de marcadores selecionado de cana-de-açúcar após análise de 1.205 acessos.

| Motivo | Tri | | Tetra | | | Penta | | | Hexa | | Composto |
|----------------|-----------------------|-----------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------|-----------------------|------------------------|--|----------|
| | (GGC) ₆ | (GAGG) ₅ | (ATCT) ₁₄ | (TTTC) ₁₅ | (CTTTT) ₁₈ | (CTGTG) ₉ | (TCCTG) ₆ | (CTCTCC) ₅ | (CTATAT) ₁₁ | (AAAAAG) ₅ /(CGT) ₅ | |
| Nº de Alelos * | CV106 | CV22 | CV29 | CV37 | CV38 | CV78 | CV104 | CV60 | CV79 | CV94 | |
| 1 | 141 ^{3,80%} | 132 ^{0,04%} | 85 ^{9,14%} | 114 ^{0,16%} | <u>93</u> ^{0,21%} | 138 ^{0,55%} | 131 ^{10,63%} | 160 ^{26,80%} | 113 ^{0,30%} | 87 ^{0,06%} | |
| 2 | 145 ^{9,93%} | 137 ^{0,73%} | 89 ^{4,23%} | 117 ^{18,31%} | <u>98</u> ^{0,69%} | 141 ^{11,98%} | 136 ^{21,27%} | 166 ^{14,22%} | 133 ^{0,30%} | 115 ^{0,03%} | |
| 3 | 148 ^{13,24%} | 139 ^{7,74%} | 93 ^{10,02%} | 118 ^{0,11%} | <u>102</u> ^{0,90%} | 144 ^{1,72%} | 141 ^{8,91%} | 168 ^{0,60%} | 136 ^{10,39%} | 135 ^{0,03%} | |
| 4 | 151 ^{22,30%} | 142 ^{0,42%} | 97 ^{9,23%} | 121 ^{18,31%} | 104 ^{7,96%} | 147 ^{15,87%} | 146 ^{18,50%} | 172 ^{25,45%} | 138 ^{2,37%} | 144 ^{0,03%} | |
| 5 | 154 ^{11,21%} | 144 ^{1,81%} | 101 ^{5,23%} | 125 ^{14,09%} | 109 ^{6,04%} | 150 ^{17,83%} | 151 ^{19,96%} | 178 ^{15,49%} | 142 ^{26,71%} | 167 ^{0,03%} | |
| 6 | 156 ^{5,98%} | 146 ^{0,12%} | 105 ^{5,30%} | 129 ^{10,66%} | 114 ^{11,23%} | 152 ^{4,36%} | 156 ^{20,67%} | 183 ^{16,62%} | 144 ^{0,30%} | 170 ^{0,03%} | |
| 7 | 157 ^{16,95%} | 148 ^{13,28%} | 109 ^{6,25%} | 138 ^{0,05%} | 119 ^{10,71%} | 155 ^{1,44%} | 162 ^{0,04%} | 189 ^{0,82%} | 148 ^{19,58%} | 181 ^{0,03%} | |
| 8 | 160 ^{16,80%} | 150 ^{0,22%} | 113 ^{9,29%} | <u>142</u> ^{0,11%} | 124 ^{3,80%} | <u>158</u> ^{0,81%} | 167 ^{0,02%} | | 155 ^{14,54%} | 183 ^{0,14%} | |
| 9 | | 152 ^{15,48%} | 117 ^{4,66%} | <u>146</u> ^{2,57%} | 130 ^{1,05%} | 160 ^{12,85%} | | | 161 ^{19,29%} | 186 ^{28,37%} | |
| 10 | | 156 ^{5,06%} | <u>121</u> ^{5,98%} | 151 ^{11,55%} | 135 ^{0,06%} | 163 ^{1,88%} | | | 163 ^{0,30%} | 189 ^{0,06%} | |
| 11 | | 158 ^{1,63%} | 125 ^{12,73%} | 155 ^{8,71%} | 140 ^{4,68%} | 165 ^{1,44%} | | | 167 ^{4,15%} | 192 ^{22,35%} | |

| | | | | | | | | | | |
|------------------------------|-----------------------|-----------------------|----------------------|-----------------------------|-----------------------------|----------------------|-----------------------|------|------|------|
| 12 | 160 ^{20,42%} | 129 ^{12,54%} | 159 ^{5,48%} | 145 ^{0,32%} | 168 ^{5,75%} | 174 ^{0,30%} | 198 ^{25,59%} | | | |
| 13 | 164 ^{15,70%} | 133 ^{3,84%} | 163 ^{2,64%} | 151 ^{4,09%} | 170 ^{4,96%} | 179 ^{0,59%} | 204 ^{18,27%} | | | |
| 14 | 166 ^{17,25%} | 137 ^{0,99%} | 167 ^{5,38%} | 156 ^{12,24%} | 173 ^{9,29%} | 185 ^{0,89%} | 221 ^{4,99%} | | | |
| 15 | 168 ^{0,10%} | 142 ^{0,21%} | 171 ^{1,86%} | 161 ^{0,50%} | 175 ^{4,14%} | | | | | |
| 16 | | 147 ^{0,35%} | | 167 ^{0,15%} | 178 ^{5,12%} | | | | | |
| 17 | | 151 ^{0,01%} | | 171 ^{0,08%} | 180 ^{1,85%} | | | | | |
| 18 | | | | 172 ^{6,79%} | 183 ^{0,09%} | | | | | |
| 19 | | | | 176 ^{0,63%} | 186 ^{7,34%} | | | | | |
| 20 | | | | 181 ^{13,03%} | 189 ^{0,11%} | | | | | |
| 21 | | | | 186 ^{10,73%} | 191 ^{15,84%} | | | | | |
| 22 | | | | 192 ^{0,34%} | 196 ^{0,81%} | | | | | |
| 23 | | | | 199 ^{0,13%} | 201 ^{12,48%} | | | | | |
| 24 | | | | 204 ^{0,04%} | <u>206</u> ^{1,60%} | | | | | |
| 25 | | | | 209 ^{1,26%} | 212 ^{0,02%} | | | | | |
| 26 | | | | 214 ^{0,01%} | | | | | | |
| 27 | | | | <u>238</u> ^{2,41%} | | | | | | |
| Total^b | 8/0 | 15/0 | 17/1 | 15/2 | 27/4 | 25/2 | 8/0 | 7/0 | 14/0 | 14/0 |
| PIC value^c | 0,85 | 0,85 | 0,92 | 0,89 | 0,92 | 0,92 | 0,82 | 0,79 | 0,80 | 0,77 |

* Os alelos sublinhados foram identificados como alelos de baixa amplificação ou do tipo *stutter*.

Os *loci* CV38 e CV78 revelaram um grande número de alelos (27 e 25, respectivamente). Identificou-se de 14-17 alelos para os *loci* CV29, CV37, CV22, CV79 e CV94 e de 78 alelos para os *loci* CV104, CV106 e CV60. Dentro de cada *locus* observou-se grande variabilidade a partir das freqüências alélicas. O grande número de alelos de cada *locus* avaliado está fortemente correlacionado com os valores observados de PIC (marcadores com maior número de alelos apresentam maior valor de PIC).

Com a identificação de alelos de baixa amplificação (indicados na tabela 10) foi importante para mapear alelos (posições) que possam gerar erros de genotipagem. Nesse caso, para o processo de identificação de variedades, esses alelos foram descartados da análise. Excluindo-se esses alelos do sistema de identificação de variedades (9) o conjunto de 10 *loci* fornece 141 alelos confiáveis.

Um dos *loci* selecionados possui motivo imperfeito (CV94), representando um problema em sistemas automatizados, como já citado. Uma alternativa a este problema pode ser o desenvolvimento de um padrão de peso molecular com fragmentos de mesmo tamanho dos alelos (*allelic ladders*), facilitando a identificação de “alelos imperfeitos” (Zhang e Yeung, 1996).

4.2.2 Avaliação do poder discriminatório dos marcadores selecionados

Após a avaliação dos dez marcadores selecionados na população de 1.205 plantas determinou-se a capacidade discriminatória de cada *locus*, expressa pela probabilidade de identidade genética e pela porcentagem de acessos que apresentaram genótipo único para um *locus* ou para uma combinação de *loci*.

Nenhum dos 10 *loci* foi capaz de discriminar todos os acessos individualmente. O conjunto dos 10 *loci* combinados discrimina todos os acessos com pelo menos três diferenças (presença de pelo menos três alelos que diferenciam um par de genótipos ou acessos). Um conjunto de 3 *loci* dos 10 avaliados foi capaz de diferenciar todos os acessos com pelo menos duas diferenças.

Dois alelos de diferença entre dois acessos diferentes é adotado como número mínimo para outros sistemas de identificação genética já validados [humanos, por exemplo] (Schumm *et al.*, 1998). Quando a comparação entre dois acessos apresenta apenas um alelo de diferença é desejável que se utilize mais marcadores (outros *loci*) no estudo, já que a diferença de um alelo pode representar uma mutação específica naquele *locus*, para aquele acesso. Embora o mesmo raciocínio possa ser utilizado para a diferença de dois alelos, considera-se um evento extremamente raro e, portanto, suficiente para afirmar que dois acessos são distintos.

Embora o *locus* CV78 tenha sido capaz de discriminar uma alta porcentagem de acessos (60%) ele não foi escolhido para integrar a combinação de três *loci* por conta do grande número de artefatos do tipo “alelo fraco” (alelos sublinhados indicados na tabela 10).

Sendo assim, a combinação dos *loci* CV38 (44%), CV29 (33%) e CV37 (8%) mostrou-se a combinação mais informativa dos 10 *loci* avaliados para a discriminação dos 1.205 acessos avaliados.

A capacidade discriminatória da combinação dos 3 *loci* pode ser demonstrada na figura 25. O dendrograma contém 54 acessos presentes na coleção avaliada, dos quais 21 são variedades comerciais ou clones do programa de melhoramento RB originadas dos mesmos progenitores (RB72454 e SP70-1143) e logo, plantas irmãs. Outros 21 acessos são representantes de outros 9 programas de melhoramento. Os demais (10) são espécies relacionadas. O conjunto dos três *loci* diferenciou os 54 acessos com pelo menos cinco diferenças.

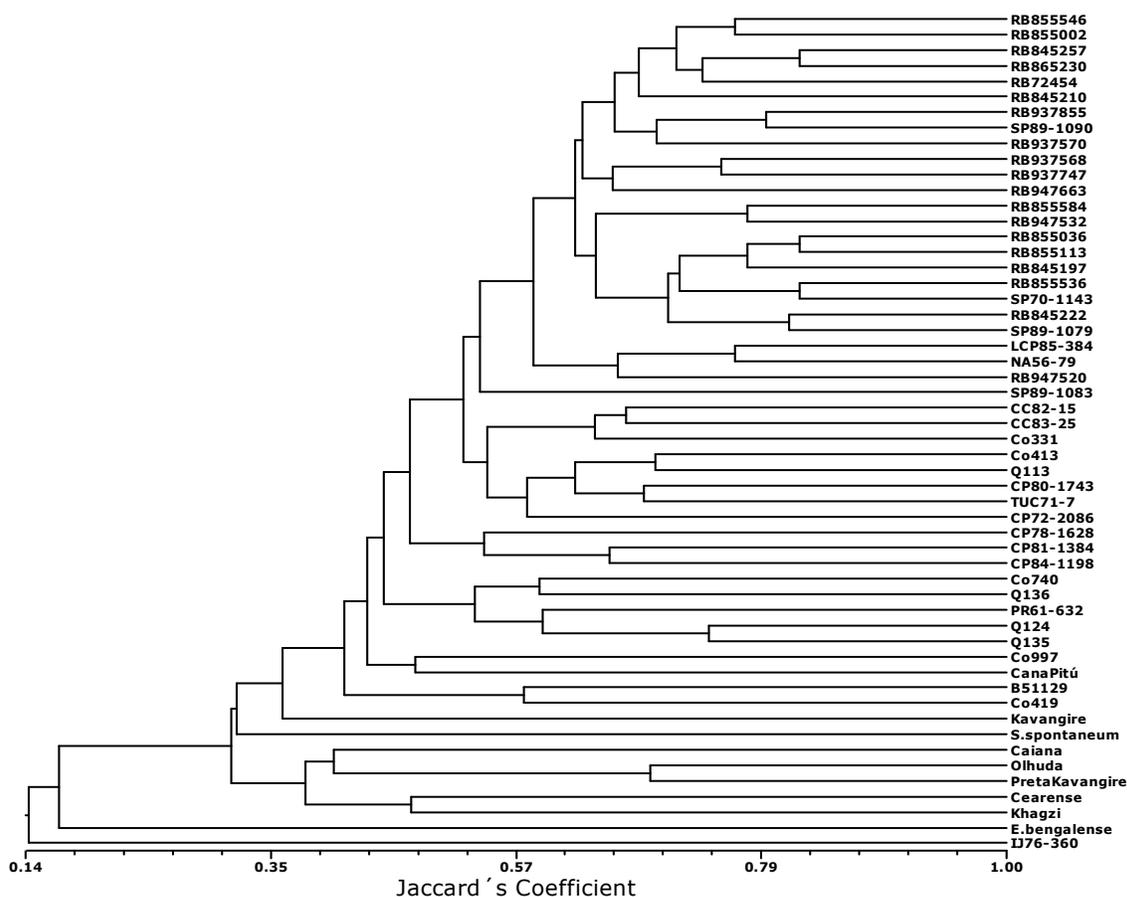


Figura 25. Capacidade discriminatória do sistema de identificação de variedades de cana-de-açúcar em 54 acessos: separação dos acessos com pelo menos duas diferenças com o uso do conjunto de três *loci* (CV29, CV37 e CV38). Acessos nomeados como “Cana Pitú”, “Caiana”, “Cearense”, “Olhuda” e “Preta Kavangire” referem-se à *S. officinarum*; “IJ76-360” à *Erianthus arundinaceum*; “Khagzi” à *S. barberi*; e “Kavangire” à *S. sinense*. O dendrograma foi feito com o programa NTSys v.2.2 (www.exetersoftware.com) utilizando o coeficiente de *Jaccard* e agrupamento UPGMA (Unweighted Pair Group Method with Arithmetic mean).

Outra importante avaliação do poder discriminatório foi a análise dos marcadores na população de 48 acessos provenientes da autofecundação da variedade RB855002. A utilização de uma progênie oriunda de autofecundação de uma variedade se deu, pois a discriminação desta progênie seria o caso mais

difícil de ser resolvido pelos marcadores moleculares, por conta da alta similaridade genética dos indivíduos envolvidos. Neste caso, o conjunto de três *loci* foi capaz de diferenciar apenas 79% (36) dos 48 acessos avaliados, sendo necessário o uso de sete *loci* (CV29, CV37, CV38, CV60, CV79, CV94 e CV106) para que todos os acessos fossem discriminados com pelo menos duas diferenças (figura 26).

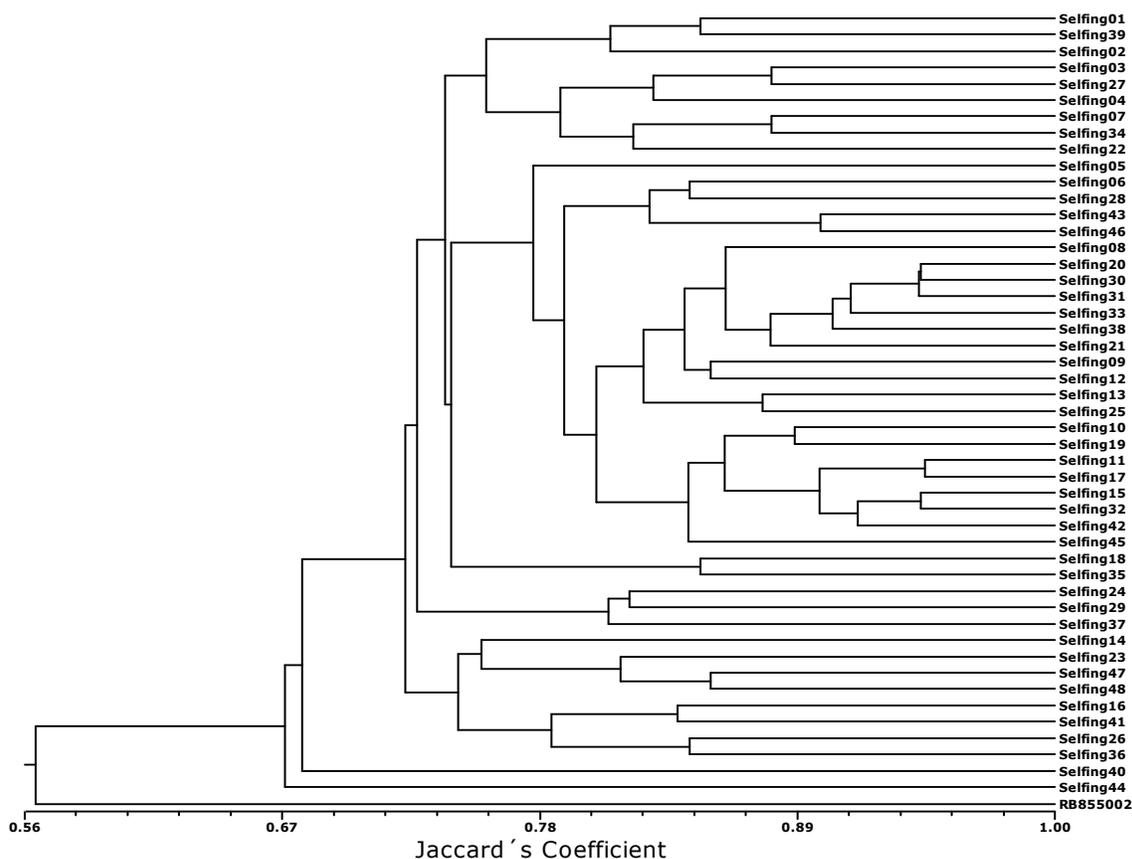


Figura 26. Capacidade do sistema de marcadores em discriminar 48 acessos originados de autofecundação (denominados de “selfing”) da variedade RB855002. Dendrograma gerado com o conjunto dos *loci* CV29, CV37, CV38, CV60, CV79, CV94 e CV106. O dendrograma foi feito com o programa NTSys v.2.2 (www.exetersoftware.com) utilizando o coeficiente de *Jaccard* e agrupamento UPGMA (Unweighted Pair Group Method with Arithmetic mean).

4.2.3 Reprodutibilidade do sistema de marcadores moleculares

Para avaliar a reprodutibilidade do sistema de identificação varietal realizou-se a amplificação do DNA extraído com os 10 *loci* avaliados dos seguintes materiais:

- a) 100 plântulas de variedades comerciais brasileiras obtidas a partir da multiplicação *in vitro* em biofábrica pela regeneração de tecido meristemático;
- b) 30 amostras de folhas transgênicas de variedades comerciais brasileiras e internacionais regeneradas a partir de calo;
- c) 30 amostras de folhas, gemas e raízes propagadas de maneira convencional (vegetativamente) que nunca passaram por qualquer método de multiplicação *in vitro* e provenientes de diferentes fazendas da região central, noroeste, sudeste do Estado de São Paulo.

A avaliação desses três grupos de plantas também foi em duplicata e por dois operadores diferentes. Em nenhum dos 10 *loci* avaliados observou-se diferença de perfil molecular, aumento de artefatos ou qualquer outra modificação que descaracterizasse a identidade genética do material, o que confirma a uniformidade e estabilidade do sistema.

4.2.4. Aplicação dos marcadores moleculares no programa de melhoramento genético

Dentro da história e do cultivo da cana-de-açúcar, uma das variedades mais importantes é a NA56-79. Esta variedade foi desenvolvida pelo melhorista argentino Roberto Ullivari que trabalhando em Porto Rico, obteve sementes da famosa variedade indiana Co419. Nesta época, o melhorista descreveu que as sementes obtidas eram provenientes da autofecundação da variedade Co419. As sementes foram germinadas e a variedade NA56-79 foi selecionada na estação de

melhoramento na Argentina [a sigla NA refere-se a “Norte Argentina”] (Pedrozo, 2006).

A variedade foi introduzida no Brasil em 1964, alcançando grande importância durante a expansão do programa Pró-Álcool e representando um grande avanço na produtividade média dos canaviais brasileiros. Esta variedade, posteriormente, deixou de ser cultivada, por conta da sua alta susceptibilidade à doença conhecida como carvão (Figueiredo, 2008).

Na história do melhoramento genético da cana-de-açúcar este foi, para nosso conhecimento, o único caso descrito de uma variedade cultivada oriunda de um cruzamento de autofecundação. Alguns autores já colocaram esta informação em dúvida, a partir de algumas análises de similaridade genética (Lima *et al.*, 2002), mas nunca se conseguiu provar se a variedade NA56-79 é realmente um produto de autofecundação ou não.

Sendo assim, utilizou-se o conjunto de dez marcadores microssatélites selecionados e identificou-se o perfil molecular das variedades NA56-79 e Co419. Pela comparação dos alelos identificados na variedade Co419 com os alelos identificados na variedade NA56-79, identificaram-se nove alelos presentes na NA56-79 e ausentes na Co419, sugerindo que estes alelos foram transmitidos por outro progenitor e que a NA56-79 não foi derivada da autofecundação da variedade indiana Co419. Este caso demonstra o potencial dos marcadores moleculares como ferramenta de apoio aos programas de melhoramento genético.

Além disso, pela capacidade de identificar com precisão variedades comerciais, este sistema de marcadores pode ser utilizado na organização de banco de germoplasmas, identificação de misturas varietais em biofábricas de produção de mudas e em viveiros de mudas e canaviais comerciais de produtores.

Outra possível aplicação dos marcadores moleculares é o seu uso para estimar a diversidade genética de acessos que compõem um banco de germoplasma, auxiliando o melhorista na identificação de progenitores. Para isto, é importante o uso de marcas moleculares que representem da melhor forma possível a diversidade genômica do acesso. Sendo assim, a boa distribuição dos

marcadores moleculares no genoma da cana-de-açúcar é importante para fornecer estimativas de diversidade genética confiáveis. Pela indisponibilidade de seqüências genômicas da cana-de-açúcar, utilizou-se o genoma do sorgo (*Sorghum bicolor*), devido à alta homologia existente entre os genomas de *Sorghum* e *Saccharum*, para avaliar a provável distribuição dos dez marcadores selecionados no genoma. Assim como em *Saccharum*, o genoma de *Sorghum* apresenta dez cromossomos básicos. Porém, *Sorghum* é diplóide (Ming *et al.*, 1998).

Pelo BLAST das seqüências dos *clusters* dos dez marcadores com as seqüências disponíveis do genoma do sorgo, identificou-se a presença de nove regiões dentro do genoma do sorgo de alta similaridade com as seqüências dos *clusters* (tabela 11).

Tabela 11. Provável distribuição dos dez marcadores microsatélites selecionados a partir da comparação com o genoma do sorgo.

| Marcador | Cromossomo de sorgo | Identidade do cluster com a seqüência do genoma do sorgo (%) |
|----------|---------------------|--|
| CV104 | 3 | 84,56 |
| CV106 | 3 | 98,49 |
| CV29 | 1 | 94,00 |
| CV37 | 1 | 95,18 |
| CV38 | 5 | 96,14 |
| CV60 | 9 | 93,15 |
| CV78 | 1 | 100,00 |
| CV79 | 6 | 92,91 |
| CV94 | 9 | 94,30 |
| CV22 | Não identificado | - |

5. CONCLUSÃO

Os objetivos iniciais deste trabalho visavam o desenvolvimento de um processo criterioso para a seleção de marcadores moleculares microssatélites que tivessem alta reprodutibilidade e alto poder discriminatório. Numa segunda etapa, este processo deveria ser utilizado para a seleção e validação dos marcadores com a qualidade desejada e que fossem capazes de formar um sistema de identificação de variedades de cana-de-açúcar.

O processo de seleção de marcadores envolvendo a avaliação de polimorfismo *in silico* e potencial se mostrou bastante eficiente. Isso é demonstrado pela eliminação, por baixo polimorfismo, de apenas 15% dos *loci* selecionados. Nesta etapa de seleção, a escolha de marcadores com motivos maiores (tri, tetra, penta e hexa) foi acertada e contribuiu para a qualidade visual dos marcadores selecionados. Na segunda etapa de seleção a utilização de PAGE como método de resolução de fragmentos foi laboriosa, mas levou à identificação de marcadores com alta qualidade visual. A utilização de dois métodos de detecção de fragmentos aumentou em muito a capacidade de decisão dos melhores marcadores para formar o sistema de identificação de variedades.

O sistema formado pelos 10 *loci* selecionados mostrou um alto poder discriminatório, diferenciando uma grande coleção de acessos de cana-de-açúcar e espécies relacionadas, além de ter diferenciado uma progênie de autofecundação. Testes em diversas origens de material vegetal mostram que o sistema tem as características de uniformidade e estabilidade. A exclusão de falsos alelos (*stutters*) e daqueles com amplificação irregular (alelos fracos) aumentaram ainda mais a precisão do sistema. Sendo assim, o sistema desenvolvido neste trabalho atende aos requisitos exigidos pela UPOV, podendo ser credenciado como uma ferramenta auxiliar em processos de registro de variedades. Sua capacidade de discriminar acessos também tem grande importância nas diversas etapas de um programa de melhoramento genético de

cana-de-açúcar, controlando a identidade de acessos em bancos de germoplasma, em clones em fase de seleção, investigação de paternidade, identificação de híbridos verdadeiros etc.

6. PRODUÇÃO CIENTÍFICA

- 1) Pôster apresentado no 52º Congresso Brasileiro de Genética (2006).
Trabalho intitulado: Desenvolvimento de um sistema de marcadores moleculares microssatélite para identificação de variedades de cana-de-açúcar.
Autores: Hamilton Jordão Junior; Guilherme Lacava de Moura; Rafaela De Gaspari; Sizuo Matsuoka; Walter Maccheroni.

- 2) Resumo apresentado no XXVI International Society of Sugar Cane Technologists (ISSCT) Congress (2007).
Trabalho intitulado: Development of a dependable microsatellite-based fingerprinting system for sugarcane
Autores: Walter Maccheroni; Hamilton Jordão Junior; Rafaela De Gaspari; Sizuo Matsuoka.

- 3) Patente
WO/2008/083456 (aplicação internacional)
Data de Publicação: 17.07.2008

- 4) Pôster apresentado no II Simpósio Brasileiro de Recursos Genéticos (2008).
Trabalho intitulado: Desenvolvimento de um sistema de identificação de variedades de cana-de-açúcar pelo uso de marcadores moleculares.
Autores: Hamilton Jordão Junior; Rafaela De Gaspari; Rebeca Siqueira; Walter Maccheroni.
Prêmio: Este trabalho foi premiado como 2º melhor trabalho da área de recursos genéticos vegetal.

5) Publicação de artigo em periódico

Título do artigo: Development of a dependable microsatellite-based fingerprinting system for sugarcane.

Periódico: Sugar Cane International.

Autores: Walter Maccheroni; Hamilton Jordão Junior; Guilherme Lacava de Moura; Rafaela De Gaspari; Sizuo Matsuoka.

Data da publicação: Janeiro/2009

7. REFERÊNCIAS BIBLIOGRÁFICAS

Allard, R.W. (1960). Backcross breeding. In: (Allard, R.W) (Ed.). *Principles of plant breeding*. New York: John Wiley and Sons.

Almeida, M. e Crócomo, O.J. (1994). Caracterização bioquímica de cultivares de cana-de-açúcar (*Saccharum* spp.): isoenzimas, proteína solúvel e valor brix. *Scientia Agricola*. 51: 422-429.

Arceneaux, G. (1967). Cultivated sugarcane of the world and their botanical derivation. *Proceedings of the International Society of Sugar Cane Technologists*. 12: 844-854.

Artschwager, E. e Brandes, E.W. (1958) Sugarcane (*Saccharum officinarum* L.): origin, classification, characteristics, and descriptions of representative clones. *Agriculture Handbook N°122*. Washington: Agricultural Research Service, Crops Research Division.

Annual Energy Outlook (2008). www.eia.doe.gov/oiaf/aeo/. Acessado em 10/07/2008.

Badger, P. C. (2002). Ethanol from cellulose: a general review. In: Janick, J.; Whipkey, A. (Ed.). *Trends in new crops and new uses*. Alexandria, VA: ASHS Press.

Barber, C.A. (1920). The origin of sugarcane. *International Sugar Journal*. 22.

Berding, N. e Roach, B.T. (1987). Germplasm collection, maintenance, and use. In: Heinz DJ (Ed.). *Sugarcane Improvement through Breeding*. Amsterdam: Elsevier Press.

Bernet, G.P; Bramardi, S.; Calvache, D.; Carbonell, E.A. e Asins, M.J. (2003). Applicability of molecular markers in the context of protection of new varieties of cucumber. *Plant Breeding*. 122: 146-152.

Besse, P.; McIntyre, L.C. e Berding, N. (1997). Characterisation of *Erianthus* sect. *Ripidium* and *Saccharum* germplasm (Andropogoneae - Saccharinae) using RFLP markers. *Euphytica*. 93: 283–292.

Besse, P. e McIntyre, L.C.; (1999). Chromosome in situ hybridisation of ribosomal DNA in *Erianthus* sect. *Ripidium* species with varying chromosome numbers confirms $x = 10$ in *Erianthus* sect. *Ripidium*. *Genome*. 42: 270–273.

Botstein, D.; White, R.L. e Skolmick, H. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human*. 32: 314-331.

Boyle, R.; Greenwood, C.; Hohler, A.; Liebreich, M.; Sonntag-O'Brien, V.; Tyne, A. e Usher, E. (2008). Global Trends in Sustainable Energy Investment 2008 Report. United Nations Foundation. http://sefi.unep.org/fileadmin/media/sefi/docs/publications/Global_Trends_2008.pdf Acesso em 28/07/2008.

Byrne, M.; Marques-Garcia, M.I.; Uren, T.; Smith, D.S. e Moran, G.F. (1996). Conservation and genetic diversity of microsatellite *loci* in the genus *Eucalyptus*. *Australian Journal of Botany*. 44: 331-341.

Caixeta, E.T.; Oliveira, A.C.B.; Brito, G.G. e Sakiyama, N.S. (2006). Tipos de Marcadores Moleculares. In: Borém, A.; Caixeta, E.T. (Ed.). *Marcadores Moleculares*. Viçosa, MG: UFV.

Casagrande, A.A. (1991). Tópicos de morfologia e fisiologia da cana-de-açúcar. Jaboticabal: FUNEP.

Chakravarty, A e Li, C.C. (1983). The effect of linkage on paternity calculation. Inclusion probabilities in parentage testing. In: (Chakravarty, A; Li, C.C.) (Ed.). *American Association of Blood Banks*. Arlington: AABB.

Cooke, J. R. (1995). Varietal identification of crop plants. Wallingford, Reino Unido: CAB International.

Cordeiro, G.M. (2001). Molecular marker systems for sugarcane germplasm analysis. In: Henry, R.J. (Ed.). *Plant Genotyping: the DNA Fingerprinting*. New York: CAB International.

Cordeiro, G.M.; Casu, R.; McIntyre, C.L.; Manners, J.M. e Henry, R.J. (2001). Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to *Erianthus* and *Sorghum*. *Plant Science*. 160: 1115-1123.

CPVO. Community Plant Variety Office (2001). Protocol for distinctness, uniformity and stability tests. CPVO-TP/Sugarbeet/1-final.

Daniels, J.; Smith, P.; Paton, N. e Williams, C.A. (1975). The origin of the genus *Saccharum*. *Sugarcane Breeding Newsletter*. 36: 24-39.

Daniels, J. e Roach, B.T. (1987). Taxonomy and evolution in sugarcane. In: Heinz DJ (Ed.). *Sugarcane Improvement through Breeding*. Amsterdam: Elsevier Press.

Demirbas, A. (2001). Biomass resource facilities and biomass conversion processing for fuels and chemicals. *Energy Conversion and Management*. 42: 1357-1378.

D'Hont, A. (2005). Unravelling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenetics Genome Research*. 109: 27-33.

D'Hont, A.; Paulet, F. e Glaszmann, J.C. (2002). Oligoclonal interspecific origin of North Indian and Chinese sugarcanes. *Chromosome research*. 10: 253-262.

Dillon, S.L.; Shapter, F.M.; Robert, H.J.; Cordeiro, G.; Izquierdo, L. e Lee, S.L. (2007). Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). *Annals of Botany*. 5: 975-989.

English, B.C.; Menard, J.R.; Ugarte, D.T. e Walsh, M. (2005). Economic impacts of ethanol production from maize stover in selected midwestern states. In: Outlaw, J. (Ed.). *Agriculture as a Producer and Consumer of Energy*. Knoxville: CABI Publishing.

Ewen, K.R.; Bahlo, M.; Treolar, S.A.; Levinson, D.F.; Mowry, B.; Barlow, J.W. e Foote, S.J. (2000). Identification and analysis of error types in high-throughput genotyping. *American Journal of Human Genetics*. 67: 727-736.

Licht, F.O. (2008). World Ethanol & Biofuels Report.

Figueiredo, P. (2008). Breve história da cana-de-açúcar e do papel do instituto agrônomo no seu estabelecimento no Brasil. In: Dinardo-Miranda, L.L.; Vasconcelos, A.C.M.; Landell, M.G.A. (Ed.). *Cana-de-açúcar*. Ribeirão Preto: IAC.

Gallacher, D.J. (1997). Optimised descriptors recommend for Australian sugarcane germplasm (*Saccharum* spp. hybrid). 48: 775-779.

Giancola, S.; Marcucci, S.P.; Poltri, S.; Lacaze, P. e Hopp, H.E. (2002). Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. *Euphytica*. 127: 95-113.

Grassl, C.O. (1964). Problems relative to the origin and evolution of wild and cultivated *Saccharum*. *Indian Journal of Sugarcane Research and Development*. 8: 106-116.

Grivet L.; Glaszmann J.C. e D'Hont A. (2006). Molecular evidence of sugarcane evolution and domestication. In: Motley, T.J.; Zerega, N.; Cross, H. Darwin's harvest: New approaches to the origins, evolution, and conservation of crops. New York: Columbia University Press.

Grivet, L.; D'Hont, A.; Roques, D.; Feldmann, P.; Lanaud, C. e Glaszmann, J.C. (1996). RFLP mapping in cultivated sugarcane (*Saccharum* spp.): Genome organization in a highly polyploid and aneuploid interspecific hybrid. *Genetics*. 142: 987-1000.

Gupta, P.K.; Balyan, H.S.; Sharma, P.C. e Ramesh, B. (1996). Microsatellites in plants: a new class of molecular markers. *Current Science*. 70: 45-54.

Hitchcock, A.S. (1923). Type species of the first 100 genera of Linnaeus "Species Plantarum". *Botanical Society of America*: 510-517.

Hoffman, J.I. e Amos, W. (2004). Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*. 14: 599-612.

Huckett, B.I. e Botha, F.C. (2004). Stability and potential use of RAPD markers in a sugarcane genealogy. *Euphytica*. 86: 117-125.

Hughes, S. (1998). STRand Nucleic Acid Analysis Software: University of California Davis. <http://www.vgl.ucdavis.edu/STRand>. Acessado em 12.05.2005.

James, G.L. (2004). An introduction to sugarcane. In: (James, G.L.) (Ed.). *Sugarcane*. Oxford: Blackwell Publishing.

Jarne, P. e Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*. 11: 424-429.

Johansson, M.; Ellegren, H. e Andersson, L. (1992). Cloning and characterization of highly polymorphic porcine microsatellites. *Journal of Heredity*. 83: 196-198.

Kandasami, P.A. e Sreenivasan, T.V. (1983). Catalogue *Saccharum spontaneum* L. Coimbatore: ICAR.

Kemp, S.J.; Brezinsky, L. e Teale, J.A. (1993). A panel of bovine, ovine and caprine polymorphic microsatellites. *Animal Genetics*. 24: 363-365.

Kirst, M.; Cordeiro, C.M.; Rezende, G.D. e Grattapaglia, D. (2005). Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. *Journal of Heredity*. 96: 161-166.

Kole, C. (2007). Genome mapping and molecular breeding in plants. Berlin, Alemanha: Springer.

Koumi, P.; Green, E.H.; Hartley, H.; Jordan, D.; Lahec, S.; Richard, L.J.; Kam, T.W. e Ward, D.M. (2004). Evaluation and validation of the ABI 3700, ABI 3100, and the MegaBACE 1000 capillary array electrophoresis instruments for use with short tandem repeat microsatellite typing in a forensic environment. *Electrophoresis*. 25: 2227-2241.

La Rota, M.; Kantety, R.V.; Yu, J.K. e Sorrells, M.E. (2005). Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley. *BMC Genomics*. 6: 23

Landell, M.G.A. e Bressiani, J.A. (2008). Melhoramento genético, caracterização e manejo varietal. In: Dinardo-Miranda, L.L.; Vasconcelos, A.C.M.; Landell, M.G.A. (Ed.). *Cana-de-açúcar*. Ribeirão Preto: IAC.

Lima, M.L.A.; Garcia, A.A.F.; Oliveira, K.M.; Matsuoka, S.; Arizono, H.; Souza Jr, C.L. e Souza, A.P. (2002). Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum* spp.). *Theoretical and Applied Genetics*. 104: 30-38.

Litt, M. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*. 44: 397-401.

Love, J.M., Knight, A.M.; McAleer, M.A. e Todd, J.A. (1990). Toward construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. *Nucleic Acids Research*. 18: 4123-4130.

MAPA. Ministério da Agricultura, Pecuária e Abastecimento (2008). www.agricultura.gov.br/portal/page?_pageid=33,976096&_dad=portal&_schema=PORTAL. Acessado em 18.09.2008

Miller, M.J. e Yuan, B.Z. (1997). Semiautomated Resolution of Overlapping Stutter Patterns in Genomic Microsatellite Analysis. *Analytical Biochemistry*. 251: 50-56.

Ming, R.; Moore, P.H.; Wu, K.; D'Hont, A.; Glazmann, J.C.; Tew, T.L.; Mirkov, T.E.; da Silva, J.; Jifon, J.; Rai, M.; Schnell, R.J.; Brumbley, S.M.; Lakshmanan, P.; Comstock, J.C. e Paterson, A.H. (2006). Sugarcane improvement through breeding and biotechnology. *Plant Breeding Reviews*. 27:15-118.

Ming, R.; Liu, S.-C.; Lin, Y.-R.; Silva, J.; Wilson, W.; Braga, D.; van Deynze, A.; Wenslaff, T.F.; Wu, K.K.; Moore, P.H.; Burnquist, W.; Sorrells, M.E.; Irvine, J.E. e Paterson, A.H. (1998). Detailed Alignment of *Saccharum* and *Sorghum* Chromosomes: Comparative Organization of Closely Related Diploid and Polyploid Genomes. *Genetics*. 150: 1663-1682.

Mongelard, J.E. (1968). The effect of different water regimes on the growth of two sugar cane varieties. *International Society Sugar Technologists*, Prod. 13th. Congress, 643-651.

Moreira, J. C. (2000). Sugarcane for energy: recent results and progress in Brazil. *Energy for Sustainable Development*. 4: 43-54.

Mukherjee, S.K. (1957). Origin and distribution of *Saccharum*. *Botanical Gazette*. 119: 55-61.

Naidu, K.M. e Sreenivasan, T.V. (1987). Conservation of sugarcane germplasm. In: *Coopersucar International Breeding Workshop*. Piracicaba: Coopersucar.

Ometto, J.C. (1981). *Bioclimatologia vegetal*. São Paulo: Agronômica Ceres.

Ortolan, M.C.A. (2006). Perspectivas para o setor sucroalcooleiro brasileiro. In: Segato, S.V. et al (Ed.). *Atualização em produção de cana-de-açúcar*. Piracicaba: Esalq/USP.

Paetkau, D.; Calvert, W.; Stirling, I. e Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*. 4: 347-358.

Pange, R.R. (1933). *Saccharum spontaneum* L. A comparative study of the forms grown at the imperial Sugarcane Breeding Station, Coimbatore. *Indian Journal of Agriculture Science*. 3: 1-10.

Pange, R.R. e Srinivasan, K. (1957). Studies in *Saccharum spontaneum*. The geographical distribution of spikelet length. *Indian J. Sugarcane Research & Development*. 1: 1-8.

Parthasarathy, N. (1948). Origin of noble sugarcane (*Saccharum officinarum* L.). *Nature*. 161: 608.

Pedrozo, C.A. (2006) Eficiência da seleção em fases iniciais no melhoramento da cana-de-açúcar. Dissertação de mestrado. Universidade Federal de Viçosa.

Pinto L.R.; Oliveira, K.M.; Marconi, T.; Garcia, A.A.F.; Ulian, E.C. e Souza, A.P. (2006). Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. *Plant Breeding*. 125: 378-384.

Pinto, L.R.; Oliveira, K.M.; Ulian, E.C.; Garcia, A.A. e Souza, A.P. (2004). Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome*. 47: 795-804.

Primmer, C.R.; Moller, A.P. e Ellegren, H. (1995). Resolving genetic relationships with microsatellite markers: a parentage testing system for the swallow *Hirundo rustica*. *Molecular Ecology*. 4: 493-498.

Raghavan, T.S. e Govindasamy, S. (1956). Phylogeny of *Saccharum* and related genera. *Proceedings of the International Society of Sugar Cane Technologists*. 9: 695-708.

Rao, J.T. (1987). Sugarcane origin, taxonomy, breeding and varieties. In: (Rao, J.T.) (Ed.). *Sugarcane varietal improvement*. Coimbatore, India: ICAR.

Rao, J.T. e Babu, C.N. (1955). Chromosome numbers in *Saccharum spontaneum* L. *Current Science*. 24: 53-54.

Roach, B.T. (1977). *Saccharum spontaneum*: Utilization of wild germplasm in breeding. *Proceedings of the International Society of Sugar Cane Technologists*. 16: 233-238.

Roach, B.T. (1984). Conservation and use of the genetic resources of sugar cane. *Sugar Cane*. 2: 7-11.

Roder, M.S.; Plaschke, J.; Konig, S.U.; Borner, A. e Sorrels, M.E. (1995). Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics*. 246: 327-333.

Rodríguez, M.; Hernández, I.; Coto, O.; Canales, E. e Cornide, M.T. (2003). RFLP and AFLP polymorphism analysis in sugarcane varieties. *Revista CENIC Ciencias Biológicas*. 9: 221-229.

Rozen, S. e Skaletsky, H.J. (1998). *Primer3* on the WWW for general users and for biologist programmers. In: (Ed.). *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Totowa: Human Press.

Saghai, M.A.M.; Biyashev, R.M.; Yan, G.P.; Zhang, Q. e Allard, R.W. (1994). Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. *Proceedings of the National Academy of Sciences*: 5: 5466-5470.

Scally, L.; Hodkinson, T. e Jones, M.B. (2000). Origin and Taxonomy of *Miscanthus*. In: Walsh, M. (Ed.). *Miscanthus: for energy and fibre*. Dublin, Irlanda: James&James.

Scaramucci, J.A.; Perin, C.; Pulino, P.; Orlando F.J.G.; Bordoni, M.P.C. e Cortez, L.A.B. (2006). Energy from sugarcane bagasse under electricity rationing in Brazil: a computable general equilibrium model. *Energy Policy*. 34: 986-992.

Scarpari, M.S. e Beauclair, E.G.F. (2008). Anatomia e Botânica. In: Dinardo-Miranda, L.L; Vasconcelos, A.C.M.; Landell, M.G.A. (Ed.). *Cana-de-açúcar*. Ribeirão Preto: IAC.

Schifino-Wittmann, M.T. e Dall'Agnol, M. (2001). Gametas não reduzidos no melhoramento de plantas. *Ciência Rural*. 31: 169-175.

Schumm, J.W., Micka, K.A.; James, W.; Katherine, A. e Rabbach, R.D. (1998). Multiplex amplification of short tandem repeat *loci*. Patente: 5.843.660, EUA.

Segatto, S.V.; Pinto, A.S.; Jendiroba, E. e Nóbrega, J.C.M. (2006). *Atualização em produção de cana-de-açúcar*. Piracicaba: Esalq/USP.

Sreenivasan, T.V. e Nair, N.V. (1991). Catalogue on Sugarcane Genetic Resources-III (*Saccharum officinarum* L.). Coimbatore, Índia: ICAR

Subramanian, K. (1987). Morphological descriptors and agriculture characteristics. In: Sankaranarayanan, P.; Natarajan, B.V. (Ed.). *Sugarcane varieties in India*. Coimbatore, Índia: ICAR.

Telles, G.P. e Silva, F.R. Trimming and clustering sugarcane ESTs. *Genetics and Molecular Biology*. 24: 17-23.

Temnykh, S.; DeClerck, G.; Lukashova, A.; Lipovich, L.; Cartinhour, S. e McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Research*. 11:1441-1452.

Thiel, T.; Michalek, W.; Varshney, R.K. e Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*. 106: 411-422.

Thompson, G.D. (1976). Water use by sugarcane. *South African Sugar Journal*. 60: 627-635.

Tommasini, L.; Batley, G.M.; Arnold, G.M.; Cooke, R.J.; Donini, P.; Lee, D.; Law, J.R.; Lowe, C.; Moule, C.; Trick, M. e Edwards, K.J. (2003). The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. *Theoretical and Applied Genetics*. 106: 1091-1101.

Van Dillewijn, C. (1952). *Botany of sugarcane*. Waltham: Mass.

Varshneya, R.K.; Granera, A. e Sorrells, M.E. (2004). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*. 23: 48-55.

Vettore, A.L.; Silva, F.R.; Kemper, E.L. e Arruda, P. (2001). The libraries that made SUCEST. *Genetics and Molecular Biology*., 24: 1-7.

UNICA. União da Indústria da Cana-de-açúcar (Dados e Cotações, 2008). www.unica.com.br. Acessado em 10/07/2008.

UNICA. União da Indústria da Cana-de-açúcar (Notícias: Retirada dos subsídios: quando e como?, 2005). www.unica.com.br. Acessado em 14/09/2008.

UPOV. União Internacional para Proteção de Novas Variedades de Plantas (2005). Essential derivation and diversity issues in winter wheat. BMT/9/10.

UPOV. União Internacional para Proteção de Novas Variedades de Plantas (2006). Progress report of the work of the technical committee, the technical working parties and the working group on biochemical and molecular techniques, and DNA-profiling in particular. C/40/10.

UPOV. União Internacional para Proteção de Novas Variedades de Plantas (2008). www.upov.int. Acessado em 15/09/2008.

Wang, Z.; Zemetra, R.S.; Hansen, J.; Hang, An; Mallory-Smith, C.,A. e Burton, C. (2002) Determination of the Paternity of Wheat (*Triticum aestivum* L) x Jointed Goatgrass (*Aegilops cylindrica* Host) BC1 Plants by Using Genomic In Situ Hybridization (GISH) Technique. *Crop Science*. 42: 939-943.

Wattier, R.; Engel, C.R.; Saumitou-Laprade, P. e Valero, M. (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite *loci*: experimental evidence at the dinucleotide *locus* Gv1CT in *Gracilaria gracilis* (*Rhodophyta*). *Molecular Ecology*. 7: 1569-1573.

Weber, J.L. (1990). Informativeness of human (dC-dA)_n(dG-dT)_n polymorphisms. *Genomics*. 7: 524-530.

Wu, K.S. e Tanksley, S.D. (1993). Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics*. 241: 225-235.

Zhang, N. e Yeung, E.S. (1996). Genetic typing by capillary electrophoresis with the allelic ladder as an absolute standard. *Analytical Chemistry*. 68:2927-2931.