

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CAMPUS DE JABOTICABAL

PARÂMETROS GENÉTICOS E MAPEAMENTO
ASSOCIATIVO DE CARACTERÍSTICAS TECNOLÓGICAS
DE CANA DE AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA

Alisson Esdras Coutinho
Engenheiro Agrônomo

2017

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CAMPUS DE JABOTICABAL

PARÂMETROS GENÉTICOS E MAPEAMENTO
ASSOCIATIVO DE CARACTERÍSTICAS TECNOLÓGICAS
DE CANA DE AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA

Alisson Esdras Coutinho

Orientador: Prof. Dr. Dilermando Perecin

Coorientadores: Dr^a Luciana Rossini Pinto

Dr. Roberto Carvalheiro

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

2017

C871p Coutinho, Alisson Esdras
Parâmetros genéticos e mapeamento associativo de características tecnológicas de cana-de-açúcar em três épocas de colheita / Alisson Esdras Coutinho. -- Jaboticabal, 2017
xii, 115 p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2017
Orientador: Dilermando Perecin
Coorientadores: Luciana Rossini Pinto, Roberto Carneiro
Banca examinadora: Rodrigo Gazaffi, Marcos Guimarães de Andrade Landell, Gustavo Vitti Mouro, Andreia da Silva Meyer
Bibliografia

1. Estrutura populacional. 2. Análise bayesiana. 3. Pol. 4. Maturação. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.52:633.61

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

CERTIFICADO DE APROVAÇÃO

TÍTULO: PARÂMETROS GENÉTICOS E MAPEAMENTO ASSOCIATIVO DE CARACTERÍSTICAS
TECNOLÓGICAS DE CANA-DE-AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA

AUTOR: ALISSON ESDRAS COUTINHO

ORIENTADOR: DILERMANDO PERECIN

COORIENTADORA: LUCIANA ROSSINI PINTO

COORIENTADOR: ROBERTO CARVALHEIRO

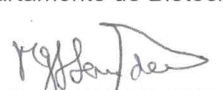
Aprovado como parte das exigências para obtenção do Título de Doutor em AGRONOMIA
(GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:



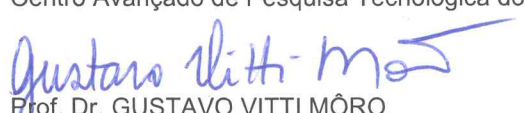
Prof. Dr. DILERMANDO PERECIN
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal



Prof. Dr. RODRIGO GAZAFFI
Departamento de Biotecnologia e Produção Vegetal e Animal / UFSCar - Araras, SP



Prof. Dr. MARCOS GUIMARÃES DE ANDRADE LANDELL
Centro Avançado de Pesquisa Tecnológica do Agronegócio de Cana / IAC - Ribeirão Preto, SP



Prof. Dr. GUSTAVO VITTI MÔRO
Departamento de Produção Vegetal / FCAV / UNESP - Jaboticabal



Profa. Dra. ANDREIA DA SILVA MEYER
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Jaboticabal, 23 de janeiro de 2017.

DADOS CURRICULARES DO AUTOR

ALISSON ESDRAS COUTINHO – nascido em 7 de abril de 1983 em Jaboatão dos Guararapes (PE), Brasil. Ingressou em 2005, no curso de Agronomia na Universidade Federal Rural de Pernambuco. Durante a graduação estagiou no Laboratório de Biotecnologia, onde desenvolveu o projeto de iniciação científica sob a orientação da Prof^a. Dr.^a Luciane Vilela Resende. Foi monitor da disciplina Estatística Aplicada à Agricultura sob a orientação do Prof. Dr. Mario de Andrade de Lira Junior e com o qual também executou o trabalho de conclusão de curso. Em fevereiro de 2011 recebeu o grau de Engenheiro Agrônomo. No mesmo ano ingressou no Programa de Pós-Graduação em Melhoramento Genético de Plantas da UFRPE (Universidade Federal Rural de Pernambuco) e concluiu o mestrado em fevereiro de 2013.

⁸Por lo demás, Hermanos, fíjense en todo lo que encuentren de verdadero, noble justo y limpio; en todo lo que es fraternal, en todos los valores morales que merecen alabanza. ⁹Pongan en práctica todo lo que han aprendido, recibido y oído de mí, todo lo que me han visto hacer, y Díos de la paz estará con ustedes.

(Filipenses 3, 8-9)

Ao meu pai Antonio Faustino Coutinho (*in memorian*) e à minha mãe Maria José de Araújo Coutinho pelo grande esforço, dedicação e amor proporcionados para que eu concretizasse todos os meus sonhos.

DEDICO

À minha querida irmã Amanda Nelma Coutinho pelo seu amor, força e coragem transmitidos.

OFEREÇO

AGRADECIMENTOS

A Deus, por todo aprendizado proporcionado mesmo nos momentos mais difíceis e a tudo que a vida nos proporciona de maneira fantástica.

À minha mãe, Maria José de Araújo Coutinho, minha tia Elizabete e à minha irmã, Amanda Coutinho pela grande contribuição para que eu pudesse vencer esta longa caminhada.

Ao meu sobrinho, Luiz Henrique, pela alegria transmitida. E aos todos os meus outros familiares paternos e maternos. Tenho muito orgulho de todos.

À Universidade Federal Rural de Pernambuco pela base profissional adquirida que até hoje rende frutos.

Ao meu orientador, Prof. Dilermando Perecin, pelos ensinamentos dentro e fora de sala de aula, pela compreensão e oportunidade de trabalhar com temas tão desafiadores.

Dr^a Luciana Rossini Pinto, pela coorientação dia a dia, atenção e preocupação para que tudo ocorresse da melhor maneira possível. Sua contribuição foi imprescindível para a realização deste trabalho.

Ao Dr. Roberto Carvalheiro, pela coorientação e contribuição que foram fundamentais nas análises estatísticas.

À Universidade Estadual Paulista pela oportunidade de formação e à CAPES pela concessão de bolsa.

Ao Dr. Mauro Alexandre Xavier pela grande contribuição nas várias etapas do experimento desde a formação da população de mapeamento até a colheita do experimento.

Ao companheiro de laboratório, Maicon Volpin, pela sua imensa dedicação e ajuda no dia a dia e por sua grande amizade.

Às técnicas e amigas de laboratório, Taís e Juliana pela ajuda, aprendizado e momentos de alegria vivenciados.

À amiga Dr^a Fernanda Raquel Camilo dos Santos pelo esclarecimento de dúvidas acerca do trabalho realizado.

À amiga Lívia Maria Carvalho pela revisão do texto.

A todos os funcionários do Centro de Cana pelo ótimo trabalho que desempenham.

A todos os outros pesquisadores do Centro de Cana que trabalham duramente proporcionando progresso científico contribuindo para um Brasil mais próspero.

Ao amigo Alysson Jalles pela grande contribuição nas análises.

Aos amigos de doutorado: Arturo Samuel, Gustavo Hugo, Lucas Santos, Hudson Rabelo, Carlos Caixeta, Guilherme Diniz, Silvan Gomes, Filipe Matias, Andreia Bastos, Miquéias Gomes, Marília Gabriela e Arturo Samuel pelos bons momentos compartilhados.

Às companheiras de Laboratório de Biotecnologia do Centro de Cana do Instituto Agronômico de Campinas (IAC): Maria Leticia, Maria Natália, Izadora Pastore, Marcel, Carol Neuber, Alexandre, Rafael Peixoto, Natália Takahashi, Jéssica Freiria, Michael Brito pelos conhecimentos compartilhados.

Muito obrigado!

SUMÁRIO

RESUMO.....	iii
ABSTRACT.....	iv
CAPÍTULO 1 –	1
1 INTRODUÇÃO GERAL.....	1
2 REVISÃO DE LITERATURA	3
2.1 Importância econômica da cana-de-açúcar	3
2.2 O gênero <i>Saccharum</i>	4
2.3 Melhoramento da cana-açúcar.....	6
2.4 Marcadores moleculares em cana-de-açúcar	9
2.5 Mapeamento associativo em cana-de-açúcar	12
REFERÊNCIAS.....	17
CAPÍTULO 2 – ESTIMATIVA DE PARÂMETROS GENÉTICOS PARA CARACTERÍSTICAS TECNOLÓGICAS DE CANA-DE-AÇÚCAR POR MEIO DO REML/BLUP E ANÁLISE MULTIVARIADA PARA DIFERENTES ÉPOCAS DE COLHEITA	28
RESUMO.....	28
ABSTRACT.....	29
1 INTRODUÇÃO	30
2 MATERIAL E MÉTODOS	32
2.1 Material vegetal e dados fenotípicos.....	32
2.2 Seleção de genótipos.....	33
2.3 Estimação dos parâmetros genéticos	37
3 RESULTADOS E DISCUSSÃO	38
4 CONCLUSÃO	56
REFERÊNCIAS.....	57
CAPÍTULO 3 – MAPEAMENTO ASSOCIATIVO DE GENES RELACIONADOS A POL E À FIBRA DE CANA-DE-AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA.....	61
RESUMO.....	61
1 Introdução.....	63
2. Material e métodos.....	66
2.1. Material vegetal	66
2.2 Avaliação fenotípica	69
2.3. Extração e quantificação do DNA genômico	69

2.4.Reações de PCR - Marcadores AFLP (<i>Amplified Fragment Length Polymorphism</i>)	69
3. Análise de dados	71
3.2.Análise da estrutura genética da população	71
3.2.1.Análise pelo software STRUCTURE	71
3.2.2.Análise via componentes principais	72
4 Análise de associação	73
4.1 Método Bayes C _π	73
5 RESULTADOS E DISCUSSÃO	74
5.1 Análise da estrutura da população.....	74
5.2 Análise de estrutura pelo programa STRUCTURE	80
5.2 Análise de associação	83
6 CONCLUSÕES	93
REFERÊNCIAS.....	94
APÊNDICES	100

PARÂMETROS GENÉTICOS E MAPEAMENTO ASSOCIATIVO DE CARACTERÍSTICAS TECNOLÓGICAS DE CANA-DE-AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA

RESUMO – O conhecimento dos teores de Pol e fibra tem por finalidade avaliar a qualidade da matéria-prima, seja para a produção de açúcar ou co-geração de energia na própria usina ou produção de etanol de segunda geração. As cultivares de cana-de-açúcar possuem diferentes períodos de maior acúmulo de sacarose, podendo ser no início, meio ou fim de safra. Como o período de safra se estende de abril até novembro, há influência das estações de ano (outono, inverno e primavera) provocando mudanças de temperatura e regime de chuvas. Portanto, a avaliação de experimentos em diferentes épocas de colheita é importante para identificar em qual período da safra os genótipos possuem maior acúmulo de sacarose. Diante deste contexto, o ranqueamento dos genótipos por época de colheita baseado nos valores genéticos preditos de Pol e fibra é de grande utilidade para fins de seleção. Além disto, a identificação de genes associados a características sob diferentes condições é de grande contribuição para a obtenção de novas cultivares pelos programas de melhoramento. Com isto, por meio do mapeamento associativo é possível detectar genes associados aos fenótipos utilizando métodos estatísticos. Sendo assim, o presente trabalho teve os seguintes objetivos: prever os valores genéticos de 100 genótipos e estimar os parâmetros genéticos de Pol e fibra em três épocas de colheita e identificar marcadores AFLP associados a Pol e fibra usando o método Bayes Cπ. Foram conduzidos três experimentos, cada um correspondente a uma época de colheita nos agrícolas de 2014/2015 e 2015/2016. O ordenamento dos genótipos por cada época de colheita foi realizado mediante os valores genéticos preditos. Para a variável fibra do grupo 1, o IN8458 foi o que obteve o maior teor de fibra para as três épocas de colheita. Enquanto que o US571415 apresentou o maior teor de Pol. No grupo 2, a IACSP972084 apresentou o maior teor de fibra e a IACSP043150, o maior valor de Pol para a análise conjunta. Para a identificação das associações entre marcador e fenótipo, foi avaliada a estrutura genética aplicando marcadores AFLP ao painel composto pelos 100 genótipos de cana-de-açúcar pertencentes ao banco de germoplasma do Centro de Cana do Instituto Agrônomo de Campinas. A análise de associação foi conduzida pelo método Bayes Cπ a partir das 640 marcas polimórficas obtidas pela genotipagem. Do total de marcas significativamente associadas com Pol, 39 marcas tiveram efeito ao menos em duas épocas de colheita. Já para fibra, observou-se 12 marcas com efeito ao menos em duas épocas. As marcas AFLP 418, AFLP 435 e AFLP 700 tiveram associações em mais de uma época para a variável Pol. Enquanto que para fibra, houve apenas a marca AFLP 520 com associação para três épocas. Quanto a variação fenotípica explicada pelos marcadores, foram 99,51%, 99,75%, 97,56% e 94,49 para as épocas 1, 2 e 3 (ano agrícola 2014/2015) e época 3 (2015/2016), respectivamente. As marcas significativamente associadas às épocas 1, 2, 3 e 3 (2016) explicaram 97,83%, 94,82%, 68,44% e 74,66%, respectivamente da variação fenotípica.

Palavras-chave: estrutura populacional, análise bayesiana, Pol, fibra, maturação

GENETIC PARAMETERS AND ASSOCIATION MAPPING IN TECHNOLOGICAL CHARACTERISTICS OF SUGARCANE IN THREE HARVEST SEASONS

ABSTRACT – The knowledge of Pol and fiber contents aims to assess the quality of the raw material either for production of sugar or co-generation of energy in the industry itself or production of second generation ethanol. The cultivars of sugarcane have different periods of higher accumulation of sucrose and may be at the beginning, middle or end of harvest. As the harvest period extends from April to November, there is influence of the seasons (autumn, winter and spring) caused by changes in temperature and rainfall. Therefore the evaluation of experiments in different harvesting times is important to identify in which period of the crop the genotypes have greater accumulation of sucrose. In this context the ranking of genotypes per harvest season based on predicted genetic values of Pol and fiber is a great utility for selection purposes. In addition, the identification of genes associated with characteristics under different conditions is a great contribution, in terms of to obtain new cultivars by breeding programs. Thus, through association mapping it is possible to detect genes associated with phenotypes using statistical methods. The present study aimed to predict the genetic values of 100 genotypes and to estimate the genetic parameters of Pol and fiber at three harvest seasons and to identify AFLP markers associated to Pol and fiber using the Bayes C π methodology. Three different experiments, each one correspond a different harvesting season, were carried out in two crops season, 2014/2015 and 2015/2016. The ordering of the genotypes for each harvesting season was performed using predicted genetic values. For the group 1 fiber variable IN8458 was the one that obtained the highest fiber content for the three harvest periods. While US571415 presented the highest Pol content. In group 2 IACSP972084 had the highest fiber content and IACSP043150 the highest Pol value for the joint analysis. For the identification of the associations between marker and phenotype the genetic structure was evaluated by applying AFLP markers to the panel composed of 100 sugarcane genotypes belonging to the germplasm bank of the Cane Center of the Agronomic Institute of Campinas. The association analysis was conducted by the Bayes C π method from the 640 polymorphic tags obtained by genotyping. Of the total of markers significantly associated with Pol 39 brands had effect in at least two harvesting seasons. For fiber we observed 12 marks with effect in at least two periods. The AFLP 418 AFLP 435 and AFLP 700 brands had associations in more than one season for the Pol variable. Whereas for fiber there was only the AFLP 520 brand with association for three seasons. The phenotypic variation explained by the markers was 99.51% 99.75% 97.56% and 94.49 for seasons 1 2 and 3 (agricultural year 2014/2015) and season 3 (2015/2016), respectively. Significantly associated with seasons 1 2 3 and 3 (2016) explained 97.83% 94.82% 68.44% and 74.66% respectively of the phenotypic variation.

Keywords: Population structure, Bayesian analysis, Pol, fiber, maturation

CAPITULO 1 –

1 INTRODUÇÃO GERAL

O Brasil, seguido por Índia e China são os maiores produtores de cana-de-açúcar com 661 milhões toneladas, 353 milhões toneladas e 130 milhões toneladas, respectivamente (OECD-FAO, 2016). Segundo dados da FAO (2016), a área mundial cultivada com cana-de-açúcar é de 22 milhões de hectares representada por cerca de 120 países. Dessa forma, a cultura da cana-de-açúcar possui grande importância no setor sucroalcooleiro, sendo responsável por mais da metade do açúcar comercializado no mundo. Além disto, o uso de biocombustíveis como alternativa energética contribui para o destaque desta cultura, colocando o Brasil como segundo maior produtor de etanol com 31% da produção mundial, estimada para 2024, logo atrás dos Estados Unidos com 42% (OECD-FAO, 2015).

Dados de produção apresentados pela Conab (2016) no último levantamento, preveem produção para a safra 2016/2017 de 694.544,8 mil toneladas de cana, um aumento de 4,4% em relação à safra anterior, em que foram colhidas 665.586,2 milhões de toneladas. Baseado nestas estimativas de aumento de produção de cana, é importante ressaltar a qualidade da matéria-prima para a indústria sucroalcooleira. Sendo assim, na história do melhoramento sempre buscou-se a obtenção de genótipos ricos em sacarose ao invés de fibra, como no século XIX, período em que variedades como a caiana e manteiga possuíam altos teores de sacarose, mas apresentavam pouca fibra e suscetibilidade ao mosaico. Assim, as características tecnológicas que definem a qualidade da cana são: brix, Pol, açúcares redutores, fibra e pureza (LAVANHOLI, 2010).

A Pol representa a porcentagem de sacarose aparente contida em uma solução açucarada de peso normal determinada pelo desvio provocado pela solução no plano de vibração da luz polarizada (LAVANHOLI, 2010). Quanto mais elevado o teor de Pol, maior é a quantidade de açúcar obtido pela indústria. Já com a relação à fibra sua importância está direcionada à geração de energia pela própria usina por meio do uso do bagaço da cana. Além disso, podem ser utilizados para a produção de etanol celulósico ou de segunda geração.

A quantificação de Pol e fibra são importantes para dois perfis distintos de cultivares: aquelas ricas em teor de sacarose, destinadas para a produção de açúcar ou etanol de primeira geração, e as que apresentam altos teores de fibra, selecionadas para a geração de energia ou produção de etanol celulósico. Na seleção para características agronômicas ou de importância econômica pode-se buscar o uso dos marcadores moleculares como ferramentas para tornar o processo de seleção mais eficiente Lande e Thompson (1990).

O emprego de técnicas utilizando marcadores moleculares aliadas a métodos estatísticos possibilita a identificação de genes para diversas características agronômicas. Esses caracteres em sua maioria são de herança quantitativa e controlados por dezenas, centenas ou milhares de loci de pequeno efeito, denominados de loci de característica quantitativa (Quantitative Trait Loci –QTL) (MACKAY et al., 2009).

O mapeamento genético é um exemplo de técnica que utiliza marcadores moleculares e métodos estatísticos, ele permite identificar a posição do QTL por meio de mapas de ligação, além de estimar seus efeitos e dissecar caracteres de herança complexa. Por outro lado, o mapeamento de QTL apresenta determinadas limitações como: baixa variação genética, baixa resolução do mapa e necessidade de tempo e recursos para a obtenção das populações de mapeamento.

Com o objetivo de superar essas limitações foi proposto o mapeamento associativo também conhecido por mapeamento por desequilíbrio de ligação, inicialmente aplicado no estudo de doenças humanas para o entendimento da base genética como os que se observam a herança suscetibilidade. A diferença para o mapeamento de QTL é que permite também o uso populações naturais ou painéis constituídos por bancos de germoplasma como populações de mapeamento, que proporciona vantagem como o menor tempo para identificar a associação marcador-característica e maior poder na detecção de associações.

Esses estudos de associação funcionam como estratégia para identificar novos marcadores ligados a importantes caracteres em espécies cultivadas para então serem utilizados em seleção assistida por marcadores (SAM), o que pode tornar o processo de seleção mais eficiente para um programa de melhoramento. Sendo assim, o objetivo do trabalho foi analisar a estrutura genética da população, realizar a fenotipagem

quanto a Pol e fibra dos genótipos de cana-de-açúcar e identificar por meio do método Bayes Ст, marcadores associados a Pol e fibra em três épocas de colheita.

2 REVISÃO DE LITERATURA

2.1 Importância econômica da cana-de-açúcar

A cadeia produtiva de cana-de-açúcar é responsável por 21,82% do PIB dos agronegócios segundo o levantamento realizado pelo Centro de Estudos Avançados em Economia Aplicada da Universidade de São Paulo (CEPEA, 2016). A Região Centro-Sul é a maior produtora de cana do País com 626.770 milhões toneladas de cana colhidas na safra 2015/2016. Estima-se que a produção nacional de açúcar ficará em 39.814,8 milhões de toneladas com acréscimo de 18,9% em relação à safra anterior, enquanto que para a produção de etanol deverá acontecer uma diminuição não ultrapassando os 27,9 bilhões de litros (CONAB, 2016).

O Brasil é o maior produtor e exportador de açúcar, com 20,3% da produção mundial. Deste total da produção nacional, 32 % é destinado ao consumo interno e 68% para a exportação (USDA, 2016). O açúcar também pode ser obtido a partir de outras culturas como a beterraba e o milho, mas o açúcar oriundo da cana tem menor custo e maior rendimento, sendo responsável por 80% da produção mundial. O etanol é outro produto importante na indústria sucroalcooleira brasileira ficando atrás somente dos Estados Unidos que é líder na produção de etanol obtido de milho, embora 60% do etanol produzido no mundo seja originado da cana-de-açúcar. Outro fato importante é que 49,2% da produção foi destinada aos Estados Unidos em 2015, o que contribuiu para o superávit da balança comercial brasileira.

O setor sucroalcooleiro brasileiro é favorecido também pelo aumento estimado de área colhida para safra 2016/2017 que é de 456,1 mil hectares (CONAB, 2016). Além disto, comparado a outros países produtores, as usinas no Brasil não são apenas produtoras de açúcar, o que permite destinar a cana para a produção de açúcar ou álcool conforme a demanda do mercado.

2.2 O gênero *Saccharum*

A cana-de-açúcar pertence à família Poaceae e tem como representantes as seguintes espécies que constituem o complexo *Saccharum*: *Saccharum officinarum* L., *Saccharum Robustum* Brandes e Jeswiet. Ex Grassl, *Saccharum barberi* Jeswiet, *Saccharum sinense* Roxb. e *Saccharum edule* Hassk. Além dos gêneros *Erianthus* seção *Ripidium*, *Sclerostachia*, *Narenga* e *Miscanthus* seção *Diandra*, os quais são relacionados filogeneticamente (DANIELS e ROACH, 1987). Uma característica em comum entre as espécies é que são altamente poliploides, possuindo de 6 a 10 cópias de cada um dos cromossomos, além de apresentarem aneuploidia (D'HONT et al., 1998; D'HONT, 2005; HA et al., 1999). Estas espécies são utilizadas em diversos estudos pelos programas de melhoramento para muitas finalidades, que compreendem a produção de açúcar, etanol e geração de energia.

A *S. officinarum* L. segundo Daniels e Roach (1987) apud Landell et al. (2010) é derivada de hibridações interespecíficas entre *S. spontaneum*, *Miscanthus*, *Erianthus arundinaceus* e *S. robustum*. É uma espécie octaploide ($2n=80$, $x=10$), ou seja, possui oito cópias de cada um dos dez cromossomos. Em relação às características morfológicas apresentam: colmos largos com diâmetro variando de 14 a 46 mm, alto teor de sacarose (7 a 22%), folhas largas (5 a 7 cm) e touceiras com poucos perfilhos (FERRARI, 2010). Quanto a sua origem, não se tem uma confirmação, considerada a Nova Guiné como principal centro de diversidade e distribuída geograficamente pelo Sudoeste da Ásia. Sua distribuição se deu da Nova Guiné para Malásia, Índia, Polinésia e do Havaí até provavelmente o Oeste da África (MING et al., 2006).

A *S. spontaneum* tem a maior distribuição geográfica comparada às outras espécies, estendendo-se do oeste do Afeganistão até a península do Malaia, Taiwan e no oeste das ilhas do Pacífico Sul e tem a Índia como centro de diversidade (SREENIVASAN; AHLOOWALIA; HEINZ, 1987). É considerada a mais primitiva dentre as espécies do complexo *Saccharum* e provavelmente pode ter surgido de regiões frias subtropicais da Índia (ROACH, 1989). Esta espécie é a que apresenta maior diversidade, além disso, constituem a maioria das variedades modernas por meio dos cruzamentos realizados com a *S. officinarum* (SINGH et al., 2013). Já quanto

ao número de cromossomos, os clones desta espécie possuem um número variado de cromossomos (40 a 128) (SREENIVASAN; AHLOOWALIA; HEINZ, 1987). Em relação às características agronômicas apresentam: alto vigor, boa capacidade de rebrota, resistência a pragas e doenças. É de pequeno porte, possui colmos curtos e finos, baixo teor de sacarose, sistema radicular bem desenvolvido e perfilhamento abundante. Quanto à adaptabilidade, está presente em diversas condições, como: ambientes de seca, incluindo desertos, lugares com excessiva umidade como regiões litorâneas e de inverno prolongado ou até de grande altitude (ROACH, 1989).

Até o ano de 1925 no Brasil, os campos de cultivo eram constituídos apenas por cultivares de *S. officinarum*. Estes genótipos apresentavam alto teor de sacarose, mas possuíam baixa resistência à doenças. Diante disto, como medida de controle, foi dado início pelos programas de melhoramento as hibridações com outras espécies do gênero *Saccharum* com o objetivo de introduzir características como rusticidade e resistência a fatores bióticos e abióticos. A *S.spontaneum* e *S. officinarum* são as principais espécies na formação das atuais cultivares visando atender a esses problemas, unindo características (rusticidade e vigor), além de atributos para melhorar a qualidade da matéria prima, maior teor de fibra, visando cultivares com perfil bioenergético e maior teor de sacarose, boa pureza de caldo, colmos suculentos no desenvolvimento de cultivares para a produção de açúcar (NAIDU; SREENIVASAN, 1987; PATERSON; MOORE; TEW 2013).

A *S. robustum* tem como maior centro de diversidade a Nova Guiné e sua origem pode estar relacionada a hibridações de *S.spontaneum* com *Erianthus* e *Miscanthus*. Ademais é considerado o parente selvagem mais próximo da *S. officinarum*. Apresentam touceiras densas, colmos com até 10 m de altura, de 20 a 45 mm e com folhas largas (DANIELS et al.,1975). Quanto ao número de cromossomos identificado por um estudo realizado por Price (1965), varia de $2n=60$ a 100, porém, os citótipos mais frequentes foram: $2n=60$ e $2n=80$. Foram ainda identificados clones apresentando entre 140 a 200 cromossomos.

A *S.sinense* ($2n=116-120$) e *S. barberi* ($2n=81-124$) são canas indígenas e foram cultivadas na Índia e China até o início do século XX. Quanto à morfologia, são de colmos finos com alto teor de fibra, baixo teor de sacarose, sistema radicular bem

desenvolvido e são melhores adaptadas a ambientes subtropicais (GRIVET; GLASZMANN; D'HONT, 2006)

. A *S. edule* é conhecida por apresentar abortamento de flores, o que dificultando seu uso no melhoramento e é usada como cultura de subsistência para a alimentação humana na Nova Guiné. É originada da hibridação de *S. officinarum* ou *S. robustum*, provavelmente com *Miscanthus*. Quanto ao seu genoma foram identificados genótipos com diferentes números de cromossomos como: $2n=60$, $2n=70$ e $2n=80$ (ROACH, 1972).

2.3 Melhoramento da cana-de-açúcar

As primeiras hibridações de cana-de-açúcar tiveram início nas ilhas de Java e Barbados nos anos de 1858 e 1859, respectivamente, favorecidas pelas condições de florescimento. As primeiras cultivares por estes programas foram identificadas como variedades POJ (Proefstation Oost Java). Outro programa importante foi o de Coimbatore (1912), Índia, no desenvolvimento de materiais com ampla adaptabilidade, a exemplo das cultivares Co e NCo. O principal objetivo desses programas era obter genótipos ricos em teor de sacarose. Sendo assim, a história do melhoramento é descrita em quatro estágios: melhoramento entre canas nobres, mobilização a partir de cruzamentos interespecíficos e retrocruzamentos com *S. officinarum*, melhoramento de canas nobres por meio de cruzamentos de cultivares e introgressão para ampliar a base genética (MING et al., 2006).

Exemplo de um desses períodos, o cruzamento entre canas nobres como: Otaheiti, Lahaina e Bourbon por meio de polinização aberta originaram algumas das primeiras cultivares nobres: como EK em Java, H109 no Havaí, B176 em Barbados, Q813 na Austrália originada. Esses genótipos contribuíram para a produção de açúcar no início do século XX. Porém, eram suscetíveis a pragas e doenças. Em virtude disto, houve a necessidade de se fazer hibridações com outra espécie de maior rusticidade, a *S. spontaneum* que possibilitou a obtenção de genótipos resistentes e mais adaptáveis (STEVENSON, 1965).

Dessa forma, o processo de nobilização é realizado após a hibridação entre *S. officinarum* e *S. spontaneum* e são realizados sucessivos retrocruzamentos com *S.*

officinarum com o objetivo de recuperar as características de interesse (alto teor de sacarose, colmos espessos e succulentos e baixo teor de fibra) e assim atender às exigências do produtor (ARCENEUX, 1967). Podem ser destacadas como exemplos deste processo a seleção da POJ2878 na ilha de Java em 1921 e obtenção das cultivares indianas, Co213, Co281 e Co290, resultantes da nobilitação entre *S. barberi* e *S. spontaneum* com *S. officinarum*. Estas cultivares foram bem aceitas, além da Índia, nos seguintes países: África do Sul, Austrália, Louisiana, Argentina, Brasil e para regiões com condições desfavoráveis com solos mais pobres (LANDELL e BRESSIANI, 2010).

Detalhe a ser destacado no processo de nobilitação é que na F_1 e o RC_1 os gametas do genitor feminino (*S. officinarum*) são transferidos de forma não reduzida ($2n=80$) juntamente com o conteúdo gamético reduzido de genitor masculino (*S. spontaneum*) (MING et al., 2006). Sendo assim, os estes híbridos que são representados pelas modernas cultivares têm seu genoma em maior parte constituído de *S. officinarum*, em torno de 80% e de 10 a 20% pertencente a *S. spontaneum* (D'HONT, 2005). Cultivares oriundas de várias gerações de cruzamento iniciadas a partir das primeiras hibridações são altamente poliploides e apresentam aneuploidia (100 a 130 cromossomos) que consiste no ganho ou perda de cromossomos e tamanho do em torno de 10.000 Mb (D'HONT; GLASZMANN, 2001), uma vez que existem genótipos de *S. spontaneum* com $2n= 40$ a 128 que contribuem para este fenômeno (MING et al., 2006).

A partir do período de nobilitação foram obtidas cultivares de grande importância para a produção de açúcar nas três décadas seguintes, como a POJ3016 e POJ3067, obtidas a partir de cruzamentos da POJ2878 com outras do grupo "POJ". Tais cultivares foram amplamente cultivadas ocupando mais de 85% da área de cultivo na ilha de Java em 1960. Além dessas cultivares, a H32-8560 (Co312 x POJ2978) ocupou cerca de 65% da área cultivada no Havaí em 1945. Mais adiante, foram obtidas importantes cultivares por meio de nobilitação por outros programas de melhoramento, a exemplo do Barbados e Coimbaitore e cultivadas em países África, Ásia e das Américas do Sul de do Norte (LANDELL; BRESSIANI, 2010).

No Brasil, o melhoramento de cana iniciou-se com a introdução de germoplasmas oriundos de outros programas de melhoramento por causa do

surgimento de doenças como o mosaico, que provocam grande queda de produção. Na época surgiram os primeiros programas de melhoramento: no Estado de Pernambuco, com a Estação de Escada, Estação de Experimental de Barreiros e Escolar Superior de Agricultura e no Rio de Janeiro, com a Estação de Campos. A partir daí foram obtidos as cultivares EB4 e EB10 na Estação de Barreiros e em Campos, as cultivares CB1300 e CB3600. No Estado de São Paulo, os primeiros estudos foram realizados em 1894 e 1895 pelo Instituto Agrônomo de Campinas conduzindo ensaios de competição com cultivares nobres. Só na década de 1930 foram obtidas as primeiras cultivares IAC36-25 e IAC47-31 (LANDELL; BRESSIANI, 2010).

É importante não deixar de destacar que a maioria das cultivares brasileiras são derivadas de menos de 20 genótipos, sendo a maioria dos programas de Java e Coimbatore, da Índia, o que proporcionou o estreitamento de base genética. Em virtude disto, foram estabelecidas medidas para a ampliação da base genética por programas de melhoramento de outros países como Barbados e Louisiana, por meio da introdução de novos materiais e cruzamentos entre canas selvagens e cultivares comerciais. No Brasil, a estratégia para aumentar a base genética teve início em 1989 com a Copersucar, selecionando parentais no segundo retrocruzamento com híbridos comerciais (LANDELL; BRESSIANI, 2010).

Outros programas surgiram após a década de 1950, como o Programa Nacional de Melhoramento de Cana (Planalsucar) que mais adiante em 1992 foi transferido para as Universidades Federais, tornando-se a Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa) e o Copersucar com sigla SP para as cultivares. Em 2003, foi criado o programa de melhoramento de cana da Cana Vialis, grupo privado com foco na viabilidade econômica sustentada pela política de royalties (GOES et al., 2009). E em 2004, o Copersucar passou a se chamar Centro de Tecnologia Canavieira, obtendo as futuras cultivares CTC com a participação de mais usinas e associação de produtores (LANDELL e BRESSIANI, 2010).

O processo para a obtenção de cultivares é um processo longo, levando de 9 a 13 anos (BARBOSA; SILVEIRA, 2010). A quantidade de etapas varia em função do programa de melhoramento, tomando como o exemplo o IAC, as fases são: hibridação para a obtenção das sementes; seleção na população segregante; fase de seleção 1

(F1); seleções clonais (F2 e F3); F4, denominada ensaios regionais e a F5, denominada como ensaio nacional para posteriormente liberação da cultivar (LANDELL; BRESSIANI, 2010).

Na primeira fase de seleção (F1) são avaliados milhares de genótipos e como são oriundos de sementes, cada um é representado por uma única planta. Portanto, nesta etapa os ensaios são conduzidos em um único ambiente (LANDELL e BRESSIANI, 2010). Nesta fase, a eficiência de seleção é menor, se comparada às demais etapas, em virtude dos caracteres avaliados serem de baixa herdabilidade no sentido amplo, além do tempo requerido e custos para a avaliação de um grande número de genótipos. Na F2 são adotadas maiores parcelas (duas linhas de quatro metros) e avaliados caracteres fenotípicos além do Brix, em ciclos de cana planta e cana soca. Na F3, os ensaios são conduzidos em parcelas de cinco metros, com duas repetições em distintas, através das quais são avaliadas características tecnológicas. E por fim, a F4 e F5, que consistem nos ensaios regionais e nacionais, respectivamente (XAVIER, 2011).

O objetivo das fases de seleção é o desenvolvimento de cultivares que sejam produtivas para em açúcar ou que tenham altos teores de fibra para maior geração de energia na indústria associada as outras importantes características como: resistência a pragas e doenças; selecionar genótipos tolerantes a fatores abióticos como seca, ausência de florescimento e isoporização, capacidade de brotação sob palha, hábito ereto de crescimento, brotação rápida e prolongada no caso das soqueiras edespalha fácil ou espontânea e elevada (LANDELL; BRESSIANI, 2010).

2.4 Marcadores moleculares em cana-de-açúcar

Os marcadores moleculares são sequências de DNA capazes de diferenciar dois ou mais indivíduos de uma determinada população revelando assim a diversidade genética existente. Segundo Ferreira e Grattapaglia (1998), marcadores moleculares são definidos como qualquer fenótipo molecular oriundo de um gene expresso ou de um segmento específico de DNA. Além desta definição, Milach (1998) define que são características de DNA que diferenciam dois ou mais indivíduos e são herdados geneticamente. Assim, o uso no melhoramento genético de plantas apresenta

vantagens, quais sejam: obtenção de grande número de polimorfismos, identificação direta do genótipo sem influência do ambiente, são de herança mendeliana, contribuindo para seleção de genótipos superiores, seleção precoce para indivíduos em que seus caracteres se manifestam em estádios fenológicos mais avançados, como o exemplo de plantas perenes (FALEIRO, 2007).

O início da aplicação dos marcadores no melhoramento de plantas se deu na década de 1980, com as isoenzimas, denominados marcadores bioquímicos (TANKSLEY; RICK, 1980; TANKSLEY; STEVEN 1983a). Um exemplo de sua aplicação foi na introgressão de características desejáveis em cultivares a partir de espécies selvagens por meio de retrocruzamentos (TANKSLEY; STEVEN, 1983b). Outra técnica de marcador utilizada para o melhoramento de plantas foi o RFLP (Restriction Fragment Length Polymorphism) (BECKMANN; SOLLER, 1986). O RFLP é baseado no uso de enzimas de restrição e foi o primeiro marcador a detectar as diferenças de indivíduos diretamente no DNA. Esta variabilidade detectada é devida à criação ou eliminação de sítios de restrição, mutação, deleções, inserções, inversões ou translocações ocorridas nas fitas de DNA, alternando a distância entre dois sítios de restrição adjacentes (CAIXETA et al., 2006). A sua aplicação em cana-de-açúcar é também vista em mapeamento de QTLs para o teor de sacarose, fibra e produção (PINTO et al., 2010). Além destas duas técnicas existem os marcadores que são baseados na técnica de PCR (Polymerase Chain Reaction), tais como: RAPD (Random Amplified), STS, (Sequence Tagged Sites), SCAR (Sequence Characterized Amplified Regions), marcadores microssatélites (SSR – Simple Sequence Repeats) e o AFLP (Amplified Fragment Length Polymorphisms).

Os marcadores SSR por meio da avaliação de germoplasma de cana de açúcar revelaram-se úteis para a identificação da constituição genética de modernas cultivares de origens interespecíficas (CORDEIRO et al., 2003).

Na aplicação dos marcadores no estudo de caracterização de cultivares de cana-de-açúcar realizado por Cordeiro et al. (2000) foi observada polimorfismo em 91% dos marcadores utilizados. Assim, em decorrência dos microssatélites revelarem um alto grau de polimorfismo e a cana possuir um genoma poliploide, houve então a necessidade de maiores informações de tais marcadores para diversas aplicações em genética e melhoramento (PARIDA et al., 2009). Os marcadores microssatélites

possuem as seguintes características: reprodutibilidade, técnica simples, pequena quantidade de DNA requerida e alto nível de polimorfismo que contribuem para o uso nos estudos de divergência genética (CAIXETA et al., 2006).

Os marcadores AFLP (*Amplified Fragment Length Polymorphism*) consistem na amplificação do DNA via PCR para detectar diferenças de fragmentos selecionados e digeridos com enzima de restrição (CAIXETA et al., 2006). Esta técnica foi proposta por Vos et al. (2005) e possui as seguintes etapas: digestão do DNA genômico, ligação dos adaptadores, pré-amplificação e amplificação seletiva. A terceira etapa foi proposta com a finalidade de se reduzir o número de fragmentos não específicos utilizando primers com um nucleotídeo seletivo adicional. Assim, os fragmentos produzidos servirão como DNA molde para a amplificação seletiva com três nucleotídeos seletivos.

A técnica AFLP possui as seguintes vantagens: alta reprodutibilidade, receptividade para qualquer complexidade de genoma, possibilidade de gerar um grande número de marcadores sem a necessidade de alguma informação prévia de sequência. Estas características são importantes para a análise da estrutura genética de populações em estudos de mapeamento associativo. Ademais, é considerado importante para pesquisas em que se utiliza uma grande quantidade de marcadores cobrindo todo o genoma sobre baixo custo por data point (MCNEIL, et al., 2011).

A utilidade destes tipos de marcadores vão desde a detecção da variabilidade genética em diversos organismos (plantas, vírus, bactérias, fungos), até os estudos de genes de resistência, seleção assistida, construção de mapas genéticos, fingerprinting. É vasto o exemplo de suas aplicações na literatura, como na piramidação de alelos de resistência a ferrugem, mancha angular e antracnose (ALZATE-MARIN, 2005), mapeamento de QTLs (WÜRSCHUM, 2012; GUO; YE, 2014; EL-SODA et al., 2014), seleção assistida (BERNARDO, 2008), seleção genômica (GOUY et al., 2013), mapeamento associativo (FARFAN et al., 2015; VISIONI et al., 2013) e estudos de diversidade e herança (LERCETEAU; SZMIDT, 1999).

2.5 Mapeamento associativo em cana-de-açúcar

A maioria das características de importância agrônômica são de herança quantitativa. Como exemplo, produção de grãos, tempo de florescimento em milho, conteúdo de sacarose em cana-de-açúcar, os quais são governados por muitos genes. As técnicas de mapeamento genético com o uso de marcadores moleculares e métodos estatísticos possibilitam a identificação de QTLs (Quantitative Trait Loci) no genoma relacionados a importantes características bem como a estimação de seus efeitos. Segundo Resende (2008), QTL é apenas uma associação estatística entre uma região do genoma e o caráter.

O uso do mapeamento é visto em diversas espécies cultivadas e em especial para cana-de-açúcar, seja para a produção ou variáveis tecnológicas que definem a qualidade da matéria-prima (MARGARIDO et al., 2015; COSTA et al., 2016; ANONI, 2011; SINGH et al., 2013; ALWALA et al., 2009; HOUARAU et al., 2001) e para a resistência a doenças (SANTOS et al., 2015; AITKEN et al., 2008). Para a condução de mapeamento de QTLs são utilizadas populações de cruzamentos controlados, tais como: F_1 , F_2 , populações de retrocruzamento, linhagens endogâmicas recombinantes, linhagens isogênicas (NILs) e duplo haploide. Cada uma dessas populações possui vantagens e desvantagens que devem ser levadas em consideração dependendo do objetivo considerado.

Embora o mapeamento de QTLs tenha sua contribuição no melhoramento genético de plantas, conforme mostrado nos trabalhos citados anteriormente, ele como qualquer outra técnica tem suas limitações. Segundo Malosetti et al. (2007), a variação genética é baixa em virtude da pouca segregação, uma vez que no caso de espécies diploides segregam no máximo dois alelos por locus. Como a maioria das populações de mapeamento são formadas a partir dois genitores, implica que nem sempre são representativas do germoplasma utilizado no melhoramento. A baixa precisão do QTL também é um ponto fraco por causa do número limitado de recombinações desde a geração parental, fazendo com que marcadores localizados a distâncias relativamente grandes dos QTLs mostrem associações significativas. Por fim, o custo e tempo necessários para a obtenção das populações de mapeamento são um obstáculo para o mapeamento de QTLs (LOPES, 2011).

Em virtude das limitações do mapeamento de QTLs houve o surgimento do mapeamento associativo no estudo de genes relacionados a doenças humanas. Griffiths et al. (2012) definem mapeamento associativo como um método para identificar associações estatísticas entre marcadores moleculares e a variação fenotípica para características complexas. Conhecido também como mapeamento por desequilíbrio de ligação (DL), a existência deste evento é de fundamental importância para a identificação de marcadores associados a genes relacionados a importantes características e consiste na associação não aleatória entre alelos em diferentes loci (FLINT-GARCIA et al., 2003). Os seguintes trabalhos revelam forte DL existente em cana-de-açúcar contribuindo para o sucesso dos estudos de associação: Rosa (2011), Raboin et al. (2008), Wei et al. (2010). Além dos QTLs, esta metodologia também permite identificar o polimorfismo gênico causador da diferença entre dois fenótipos alternativos, haplótipos e blocos haplótipos representando diferentes alelos de um gene (GUPTA et al., 2005).

A diferença entre o mapeamento de QTLs e o mapeamento associativo é que o primeiro se baseia em populações de cruzamentos controlados, citadas anteriormente, resultando na identificação de QTLs em blocos gênicos maiores por causa da pouca recombinação ocorrida. No mapeamento associativo, pelo contrário como são utilizadas populações naturais, coleções germoplasma ou linhagens em desenvolvimento há um maior histórico de recombinação, fazendo com a que identificação de associações entre o marcador e gene seja em blocos gênicos menores (ROSA, 2011).

Autores como Rafalski (2002), Malosetti et al. (2007) destacam o mapeamento associativo como uma ferramenta para identificar associações entre marca e gene pelas seguintes vantagens: a população não necessariamente tem que ser obtida de cruzamentos controlados, podendo ser coleções de bancos de germoplasma, populações naturais ou linhagens em desenvolvimento. Isto é vantajoso porque aproveita todo o histórico de recombinações ocorridas no decorrer de várias gerações de cruzamentos, proporcionando maior poder de detectar associações. Ainda se tem estimativas mais realistas dos efeitos dos QTLs, uma vez que são avaliados em um conjunto de genótipos relevantes para o melhoramento, além da redução de custos e tempo para a obtenção de populações experimentais.

No estudo de mapeamento associativo existem duas abordagens: a identificação de um gene candidato ou a busca por associações de marcas com genes específicos para determinadas características, a qual se denomina “Genome –Wide-Association-Study” (GWAS). A aplicação desta técnica de mapeamento já é utilizado há mais de uma década em grandes culturas como a de cevada (KRAAKMAN et al., 2006; ELLIS et al., 2002; TOOJINDA et al., 2000), trigo (SINGH; NELSON; SORRELLS, 2000) e milho (GUPTA; RUSTGI; KULWAL, 2005).

Estudos de mapeamento associativo em cana-de-açúcar foram conduzidos para diversas finalidades: na identificação de resistência a doenças como podridão da raiz, escaldadura das folhas, Fijivirus e o carvão conduzidos por Wei et al. (2006), na avaliação de características agrônômicas (BILAL et al., 2015), no estudo de características tecnológicas avaliadas em diferentes ambientes (WEI et al., 2010; LOPES, 2011). Racedo et al. (2016) conduziram o mapeamento associativo utilizando 1535 marcadores DArT e 103 TRAP com um painel constituído de 88 acessos de cana-açúcar oriundos da Estación Experimental Agroindustrial Obispo Colombres de Túcuman na Argentina. O estudo foi conduzido em cana planta, primeira soca e segunda soca. Foram identificados 43, 42 e 41 incluindo marcadores DArT e TRAP para a produtividade para os três ciclos de cultivo, respectivamente. Enquanto que para Pol: 38, 34 e 47 marcadores (DArT e TRAP) foram identificados como significativamente associados a Pol.

Um outro fator que contribui para a identificação da associação marcador-caráter é a existência de desequilíbrio de ligação (DL), que consiste na associação não aleatória entre alelos de diferentes loci com uma frequência maior do que a esperada, caso estes alelos se combinassem de forma independente (FLINT-GARCIA et al., 2003). Segundo Gupta et al. (2005), estatisticamente, o DL representa a covariância dos polimorfismos exibidos por dois marcadores. Um detalhe importante a ser destacado é que nem sempre que dois loci estão em DL estejam obrigatoriamente ligados. E da mesma maneira dois loci fisicamente ligados em determinado cromossomo ou haplótipo não significa que estão em DL (ROSA, 2011).

O DL para espécie poliploide, como a cana-de-açúcar, apresenta diferença comparado às diploides em relação à quantidade de cromossomos, que dependerá

da ploidia da espécie ou também da ploidia do grupo de homologia, uma vez que eventos aneuploides durante a segregação podem ser frequentes (ROSA, 2011).

Na análise de associação um fator que pode influenciar na confiabilidade dos resultados é a existência da estrutura da população, ou seja, a presença de subgrupos dentro da população em estudo. A execução da análise desconsiderando sua presença pode levar a associações espúrias ou falsos positivos. Para evitar este problema, é recomendável que seja realizada dentro de cada subgrupo com a finalidade de remover as correlações genéticas devido ao efeito dos subgrupos. Existem métodos estatísticos, como controle genômico (DEVLIN; ROEDER, 1999), inferência bayesiana, análise de componentes principais (YU et al., 2006), regressão logística (SETAKIS; STIRNADEL; BALDING, 2006) que são utilizados para controlar o efeito da esterificação nas populações.

Além dos métodos para remover o efeito de confundimento causado pela estrutura da população são utilizadas outras metodologias para detectar os efeitos dos marcadores no controle do caráter. Os métodos são aplicados tanto na seleção genômica ampla com em GWAS e dentre os mais utilizados estão: modelos mistos (YU et al., 2006) e métodos bayesianos (Bayes A, Bayes B, Bayes C, Bayes C π , Bayes D π). Segundo Resende (2008), os métodos bayesianos estão associados a sistemas de equações não lineares e as predições não lineares podem ser melhores quando os efeitos de QTL não são normalmente distribuídos devido a presença de genes maiores.

Na análise bayesiana, diferente da estatística frequentista, os termos sobre variável dependente são denominadas distribuições de probabilidades: distribuição a priori $f(y|\theta)$, θ consiste em uma quantidade desconhecida e necessária para descrever a variável resposta. Já a distribuição a posteriori é obtida a partir da informação sobre os dados inserida na análise por meio da função de verossimilhança e da distribuição a priori para os parâmetros de interesse.

O método Bayes C π foi descrito por Habier et al. (2011) e para o mapeamento de QTLs propicia a informação sobre a arquitetura genética do caráter quantitativo e identifica as posições de QTL por modelagem da frequência de SNPs. O que caracteriza o método é a distribuição π uniforme (0,1), o que equivale ao método RR-BLUP com seleção de covariáveis implementado via MCMC. A modelagem de π é

importante para a análise de associação, pelo fato de a maioria das marcas não estarem em DL com os genes.

REFERÊNCIAS

AITKEN, K. S.; HERMANN, S.; KARNO, K.; BONNETT, G. D.; MCINTYRE, L. C.; JACKSON, P. A. Genetic control of yield related stalk traits in sugarcane. **Theoretical and Applied Genetics**, v. 117, n. 7, p. 1191-1203, 2008.

ALWALA, S.; KIMBENG, C. A.; VEREMIS, J. C.; GRAVOIS, K. A. Identification of molecular markers associated with sugar-related traits in a *Saccharum* interspecific cross. **Euphytica**, v. 167, n. 1, p. 127-142, 2009.

ALZATE-MARIN, A. L.; CERVIGNI, G. D.; MOREIRA, M. A.; BARROS, E. G. Seleção assistida por marcadores moleculares visando ao desenvolvimento de plantas resistentes a doenças, com ênfase em feijoeiro e soja. **Fitopatologia Brasileira**, v. 30, n. 4, p. 333-342, 2005.

ANONI, C. D. O. **Mapeamento associativo para múltiplos ambientes e múltiplos locos, visando tolerância à seca em milho**. 2016. 64 f. Tese de doutorado - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2016.

ANONI, C. D. O. **Mapeamento de QTLs em progênies de irmãos-completos de cana-de-açúcar utilizando imputação múltipla**. 2011. 100 f. Dissertação de mestrado - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2011.

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

BARBOSA, M.H.P; SILVEIRA, L.C.I. Melhoramento genético e recomendação de cultivares. In: SANTOS, F.; BORÉM, A.; CALDAS, C. (Eds). **Cana-de-açúcar - Bioenergia, Açúcar e Alcool – Tecnologias e Perspectivas**. Viçosa: UFV, 2010. p. 313-331.

BECKMANN, J.; SOLLER, M. Restriction fragment length polymorphisms and genetic improvement of agricultural species. **Euphytica**, v. 35, n. 1, p. 111-124, 1986.

BERNARDO, R. Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years **Crop Science**, v. 48, n. 5, p. 1649-1664, 2008.

BILAL, M.; SAEED, M.; NASIR, I. A.; TABASSUM, B.; ZAMEER, M.; KHAN, A.; TARIQ, M.; JAVED, M. A.; HUSNAIN, T. Association mapping of cane weight and tillers per plant in sugarcane. **Biotechnology & Biotechnological Equipment**, v. 29, n. 4, p. 617-623, 2015.

CAIXETA, E. T.; OLIVEIRA, A. C. B.; BRITO, G. G.; SAKIYAMA, N. S. Tipos de marcadores moleculares. In: BORÉM, A.; CAIXETA, E. T. **Marcadores moleculares**. Viçosa: Universidade Federal de Viçosa, 2006. p. 09-78.

CEPEA, Centro de Estudos Avançados em Economia Aplicada da Esalq/USP. **PIB Agronegócio de Cadeias, 3º trimestre de 2016**. São Paulo: Cepea 2016.

CONAB, Companhia Nacional de Abastecimento. **Acompanhamento de safra brasileira: cana-de-açúcar, terceiro levantamento, dezembro/2016** – Companhia Nacional de Abastecimento. – Brasília: Conab 2016.

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

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

COSTA, E. A.; ANONI, C. O.; MANCINI, M. C.; SANTOS, F. R. C.; MARCONI, T. G.; GAZAFFI, R.; PASTINA, M. M.; PERECIN, D.; MOLLINARI, M.; XAVIER, M. A.; PINTO, L. R.; SOUZA, A. P.; GARCIA, A. A. F. QTL mapping including codominant SNP markers with ploidy level information in a sugarcane progeny. **Euphytica**, v. 211, n. 1, p. 1-16, 2016.

D'HONT, A. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. **Cytogenetic and genome research**, v. 109, n. 1-3, p. 27-33, 2005.

D'HONT, A.; GLASZMANN, J. C. Sugarcane genome analysis with molecular markers, a first decade of research. In: Proc Int Soc Sugarcane Technol, 2001, p. 556-559.

DANIELS, J.; ROACH, B.T. Taxonomy and evolution. In: Heinz, D.J. (Ed.). **Sugarcane improvement through breeding**. Elsevier, Amsterdam, 1987. p.7-84.

DANIELS, J.; SMITH, P.; PATON, N.; WILLIAMS, C.A. (1975). The origin of the genus *Saccharum*. *Sugarcane Breeding Newsletter* 36: 24-39.

DEKKERS, J. C. Commercial application of marker-and gene-assisted selection in livestock: strategies and lessons. **Journal of Animal Science**, v. 82, n.13 suppl, p. E313-E328, 2004.

DEUBER, R. Maturação da cana-de-açúcar na região sudeste do Brasil. In: SEMINÁRIO DE TECNOLOGIA DA COPERSUCAR, 1988. Piracicaba. **Anais**. Piracicaba: Copersucar, 1988. p. 33-40.

DEVLIN, B.; ROEDER, K. Genomic control for association studies. **Biometrics**, v. 55, n. 4, p. 997-1004, 1999.

D'HONT, A.; ISON, D.; ALIX, K.; ROUX, C.; GLASZMANN, J. C. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. **Genome**, v. 41, n. 2, p. 221-225, 1998.

ELLIS, R. P.; FORSTER, B. P.; GORDON, D. C.; HANDLEY, L. L.; KEITH, R. P.; LAWRENCE, P.; MEYER, R.; POWELL, W.; ROBINSON, D.; SCRIMGEOUR, C. M.; YOUNG, G.; THOMAS, W. T. B. Phenotype/genotype associations for yield and salt tolerance in a barley mapping population segregating for two dwarfing genes. **Journal of Experimental Botany**, v. 53, n. 371, p. 1163-1176, 2002.

EL-SODA, M.; MALOSETTI, M.; ZWAAN, B. J.; KOORNNEEF, M.; AARTS, M. G. M. Genotype × environment interaction QTL mapping in plants: lessons from Arabidopsis. **Trends in Plant Science**, v. 19, n. 6, p. 390-398, 2014.

FALEIRO, F. G. **Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos**. 1ª edição. Planaltina: Embrapa Cerrados, 2007. 102 p.

FAO. Agricultural Data. Faostat Database Results, 2016. Disponível em: <<http://faostat.fao.org/>>. Acesso em: 27 nov. 2016.

FARFAN, I. D. B.; DE LA FUENTE, G. N.; MURRAY, S. C.; ISAKEIT, T.; HUANG, P.-C.; Warburton, M.; WILLIAMS, P.; WINDHAM, G. L.; KOLOMIETS, M. Genome Wide Association Study for Drought, Aflatoxin Resistance, and Important Agronomic Traits of Maize Hybrids in the Sub-Tropics. **PLOS ONE**, v. 10, n. 2, p. e0117737, 2015.

FERNANDES, A. C. **Cálculos na Agroindústria da cana de açúcar**. Piracicaba, STAB: Açúcar, Álcool e Subprodutos, 2000, 193p.

FERRARI, F. Caracterização cromossômica em cana-de-açúcar (*Saccharum* spp., Poaceae). 2010.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética 2ª**. Brasília Embrapa Recursos Genéticos e Biotecnologia 1998. 219 p.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análises genéticas**. Brasília: Embrapa-Cenargen, 1998. 220 p.

FLINT-GARCIA, S. A.; THORNSBERRY, J. M.; IV, B. Structure of linkage disequilibrium in plants*. **Annual review of plant biology**, v. 54, n. 1, p. 357-374, 2003.

GOES, T.; ARAÚJO, M. D.; MARRA, R. Novas fronteiras tecnológicas da cana-de-açúcar no Brasil. **Revista de Política Agrícola**, v. 18, n. 1, p. 50-59, 2009.

GONÇALVES, B. S. **Identificação de marcadores moleculares relacionados à severidade de sintomas da ferrugem marrom em painel associativo de cana-de-açúcar (*Saccharum* spp.)**. 2012. 79 f. Tese de doutorado -Universidade Estadual Paulista, Jaboticabal, 2012.

GOUY, M.; ROUSSELLE, Y.; BASTIANELLI, D.; LECOMTE, P.; BONNAL, L.; ROQUES, D.; EFILE, J.-C.; ROCHER, S.; DAUGROIS, J.; TOUBI, L.; NABENEZA, S.; HERVOUET, C.; TELISMART, H.; DENIS, M.; THONG-CHANE, A.; GLASZMANN, J. C.; HOARAU, J.-Y.; NIBOUCHE, S.; COSTET, L. Experimental assessment of the accuracy of genomic selection in sugarcane. **Theoretical and Applied Genetics**, v. 126, n. 10, p. 2575-2586, 2013.

GRIFFITHS, A. J. F.; WESSLER, S. R.; CARROLL, S. B.; JOHN, D. **Introduction to genetic analysis**. Tenth. New York: W. H. Freeman and Company, 2012. 862 p.

GRIVET, L.; GLASZMANN, J.-C.; D'HONT, A. 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, 2006. cap. 3, p. 49-63.

GUO, L.-B.; YE, G.-Y. Use of Major Quantitative Trait Loci to Improve Grain Yield of Rice. **Rice Science**, v. 21, n. 2, p. 65-82, 2014/03/01, 2014.

GUPTA, P. K.; RUSTGI, S.; KULWAL, P. L. Linkage disequilibrium and association studies in higher plants: present status and future prospects. **Plant molecular biology**, v. 57, n. 4, p. 461-485, 2005.

HA, S.; MOORE, P. H.; HEINZ, D.; KATO, S.; OHMIDO, N.; FUKUI, K. Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence in situ hybridization and imaging methods. **Plant molecular biology**, v. 39, n. 6, p. 1165-1173, 1999.

HA, S.; MOORE, P. H.; HEINZ, D.; KATO, S.; OHMIDO, N.; FUKUI, K. Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence in situ hybridization and imaging methods. **Plant molecular biology**, v. 39, n. 6, p. 1165-1173, 1999.

HABIER, D.; FERNANDO, R. L.; KIZILKAYA, K.; GARRICK, D. J. Extension of the Bayesian alphabet for genomic selection. **BMC bioinformatics**, v. 12, n. 1, p. 1, 2011.

HE, S.; OHM, H.; MACKENZIE, S. Detection of DNA sequence polymorphisms among wheat varieties. **Theoretical and Applied Genetics**, v. 84, n. 5-6, p. 573-578, 1992.

HIRSCHHORN, J. N.; DALY, M. J. Genome-wide association studies for common diseases and complex traits. **Nature Reviews Genetics**, v. 6, n. 2, p. 95-108, 2005.

HOARAU, J. Y.; OFFMANN, B.; D'HONT, A.; RISTERUCCI, A. M.; ROQUES, D.; GLASZMANN, J. C.; GRIVET, L. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. **Theoretical and Applied Genetics**, v. 103, n. 1, p. 84-97, 2001.

KO, H.; COWAN, D.; HENRY, R.; GRAHAM, G.; BLAKENEY, A.; LEWIN, L. Random amplified polymorphic DNA analysis of Australian rice (*Oryza sativa* L.) varieties. **Euphytica**, v. 80, n. 3, p. 179-189, 1994.

KRAAKMAN, A. T. W.; MARTINEZ, F.; MUSSIRALIEV, B.; VAN EEUWIJK, F. A.; NIKS, R. E. Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. **Molecular Breeding**, v. 17, n. 1, p. 41-58, 2006.

KRAAKMAN, A. T. W.; MARTINEZ, F.; MUSSIRALIEV, B.; VAN EEUWIJK, F. A.; NIKS, R. E. Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. **Molecular Breeding**, v. 17, n. 1, p. 41-58, 2006.

LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, v. 124, n. 3, p. 743-756, 1990.

LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, v. 124, n. 3, p. 743-756, 1990.

LANDELL, M. G. A.; BRESSIANI, J. A. Melhoramento genético, caracterização e manejo varietal. In: DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M.; LANDELL, M. G. A. **Cana-de-açúcar**. Campinas: Instituto Agrônômico 2010. cap. 5, p. 101-155.

LAVANHOLI, M. D. G. D. P. Qualidade da cana-de-açúcar como matéria-prima para produção de açúcar e álcool. In: DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M.; LANDELL, M. G. D. A. **Cana-de-açúcar**. 1ª. Campinas: Instituto Agrônômico de Campinas, 2010. cap. 8, p. 882.

LERCETEAU, E.; SZMIDT, A. E. Properties of AFLP markers in inheritance and genetic diversity studies of *Pinus sylvestris* L. **Heredity**, v. 82, n. 3, p. 252-260, 1999.

LOPES, F. C. C. **Mapeamento genético de cana-de-açúcar (*Saccharum* spp.) por associação empregando marcadores SSR e AFLP**. 2011. 144 f. Tese de doutorado -Universidade de São Paulo - Escola Superior de Agricultura "Luiz de Queiroz" Piracicaba, 2011.

MACKAY, T. F. C.; STONE, E. A.; AYROLES, J. F. The genetics of quantitative traits: challenges and prospects. **Nat Rev Genet**, v. 10, n. 8, p. 565-577, 2009.

MALOSETTI, M.; VAN DER LINDEN, C. G.; VOSMAN, B.; VAN EEUWIJK, F. A. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. **Genetics**, v. 175, n. 2, p. 879-889, 2007.

MARGARIDO, G. R. A.; PASTINA, M. M.; SOUZA, A. P.; GARCIA, A. A. F. Multi-trait multi-environment quantitative trait loci mapping for a sugarcane commercial cross provides insights on the inheritance of important traits. **Molecular Breeding**, v. 35, n. 8, p. 175, 2015.

MARGARIDO, G. R. A.; PASTINA, M. M.; SOUZA, A. P.; GARCIA, A. A. F. Multi-trait multi-environment quantitative trait loci mapping for a sugarcane commercial cross provides insights on the inheritance of important traits. **Molecular Breeding**, v. 35, n. 8, p. 175, 2015.

MCNEIL, M. D.; HERMANN, S.; JACKSON, P. A.; AITKEN, K. S. Conversion of AFLP markers to high-throughput markers in a complex polyploid, sugarcane. **Molecular Breeding**, v. 27, n. 3, p. 395-407, 2011.

MILACK, S. C. K. **Marcadores moleculares em plantas**. Porto Alegre: UFRGS, 1998. 141 p.

MING, R.; MOORE, P. H.; WU, K.; D HONT, A.; GLASZMANN, J. C.; TEW, T. L.; MIRKOV, T. E.; DA SILVA, J.; JIFON, J.; RAI, M. Sugarcane improvement through breeding and biotechnology. **Plant breeding reviews**, v. 27, p. 15, 2006.

NAIDU, K. M.; SREENIVASAN, T. Conservation of sugarcane germplasm. In: Proceedings of the Copersucar International Sugarcane Breeding Workshop. Copersucar Technology Centre, Piracicaba-SP, Brazil, 1987, p. 33-53.

NORDBORG, M. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. **Genetics**, v. 154, n. 2, p. 923-929, 2000.

OECD/FAO (2015), "Biofuels", in *OECD-FAO Agricultural Outlook 2015*, OECD Publishing, Paris.

OECD/FAO (2016), "Sugar", in *OECD-FAO Agricultural Outlook 2016-2025*, OECD Publishing, Paris.

PARIDA, S. K.; KALIA, S. K.; KAUL, S.; DALAL, V.; HEMAPRABHA, G.; SELVI, A.; PANDIT, A.; SINGH, A.; GAIKWAD, K.; SHARMA, T. R.; SRIVASTAVA, P. S.; SINGH, N. K.; MOHAPATRA, T. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. **Theoretical and Applied Genetics**, v. 118, n. 2, p. 327-338, 2009.

PASTINA, M.; MALOSETTI, M.; GAZAFFI, R.; MOLLINARI, M.; MARGARIDO, G.; OLIVEIRA, K.; PINTO, L.; SOUZA, A.; VAN EEUWIJK, F.; GARCIA, A. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theoretical and Applied Genetics**, v. 124, n. 5, p. 835-849, 2012.

PATERSON, A. H.; MOORE, P. H.; TEW, T. L. The gene pool of *Saccharum* species and their improvement. In: **Genomics of the Saccharinae**. Springer, 2013. p. 43-71.

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

PRICE, S. Interspecific hybridization in sugarcane breeding. In: Proc Int Soc Sugar Cane Technol, 1965, p. 1021-1026.

RABOIN, L.-M.; PAUQUET, J.; BUTTERFIELD, M.; D'HONT, A.; GLASZMANN, J.-C. Analysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. **Theoretical and Applied Genetics**, v. 116, n. 5, p. 701-714, 2008.

RACEDO, J.; GUTIÉRREZ, L.; PERERA, M. F.; OSTENGO, S.; PARDO, E. M.; CUENYA, M. I.; WELIN, B.; CASTAGNARO, A. P. Genome-wide association mapping of quantitative traits in a breeding population of sugarcane. **BMC Plant Biology**, v. 16, n. 1, p. 142, 2016.

RAFALSKI, A. Applications of single nucleotide polymorphisms in crop genetics. **Current opinion in plant biology**, v. 5, n. 2, p. 94-100, 2002.

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

RESENDE, M. D. V. **Genética quantitativa e seleção no melhoramento de plantas perenes e animais**. Colombo: Embrapa Florestas, 2008. 330 p.

RIPOLI, T. C. C.; RIPOLI, M. L. C. **Biomassa de cana-de-açúcar: colheita, energia e ambiente**. Piracicaba: Barros & Marques Ed. Eletrônica, 2004. 302 p.

ROACH, B. Chromosome numbers in *Saccharum* edule. **Cytologia**, v. 37, n. 1, p. 155-161, 1972.

ROACH, B. ORIGIN AND IMPROVEMENT OF THE GENETIC BASE of SUGARCANE. In: Proceedings of Australian Society of Sugar cane Technology, 1989, Australia.

ROBBINS, M.; WITSENBOER, H.; MICHELMORE, R. W.; LALIBERTE, J.-F.; FORTIN, M. Genetic mapping of turnip mosaic virus resistance in *Lactuca sativa*. **Theoretical and Applied Genetics**, v. 89, n. 5, p. 583-589, 1994.

ROSA, J. R. B. F. **Análise do desequilíbrio de ligação e da estrutura populacional do germoplasma brasileiro de cana-de-açúcar**. 2011. 97 f. Tese de doutorado - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2011.

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

SEEMA, N.; KHAN, I. A.; RAZA, S.; YASMEEN, S.; BIBI, S.; NIZAMANI, G. S. Assessment of genetic variability in somaclonal population of sugarcane. **Pak. J. Bot**, v. 46, n. 6, p. 2107-2111, 2014.

SETAKIS, E.; STIRNADEL, H.; BALDING, D. J. Logistic regression protects against population structure in genetic association studies. **Genome research**, v. 16, n. 2, p. 290-296, 2006.

SINGH, R. P.; NELSON, J. C.; SORRELLS, M. E. Mapping Yr28 and other genes for resistance to stripe rust in wheat. **Crop Science**, v. 40, n. 4, p. 1148-1155, 2000.

SINGH, R.; SINGH, R.; SINGH, S.; MISHRA, N.; RASTOGI, J.; SHARMA, M.; KUMAR, A. Genetic diversity among *Saccharum spontaneum* clones and commercial hybrids through SSR markers. **Sugar Tech**, v. 15, n. 2, p. 109-115, 2013.

SREENIVASAN, T. V.; AHLOOWALIA, B. S.; HEINZ, D. J. Chapter 5 - Cytogenetics. In: DON, J. H. **Developments in Crop Science**. Elsevier, v. Volume 11, 1987. p. 211-253.

STEVENS, M.; HEINY, D.; RHOADS, D.; GRIFFITHS, P.; SCOTT, J. A linkage map of the tomato spotted wilt virus resistance gene Sw-5 using near isogenic lines and an

interspecific cross. **Tospoviruses and Thrips of Floral and Vegetable Crops 431**, p. 385-392, 1995.

STEVENSON, G. C. (1965). *Genetics and Breeding of Sugar Cane*. Longmans, Green and Co, London.

TANKSLEY, S. D. Introgression of Genes from Wild Species. In: STEVEN, D. T.; THOMAS, J. O. **Developments in Plant Genetics and Breeding**. Elsevier, .Volume 1, Part A, 1983a. p. 331-337.

TANKSLEY, S. D. Molecular markers in plant breeding. **Plant Molecular Biology Reporter**, v. 1, n. 1, p. 3-8, 1983b.

TANKSLEY, S.; RICK, C. Isozymic gene linkage map of the tomato: applications in genetics and breeding. **Theoretical and Applied Genetics**, v. 58, n. 2, p. 161-170, 1980.

TOOJINDA, T.; BROERS, L. H.; CHEN, X. M.; HAYES, P. M.; KLEINHOF, A.; KORTE, J.; KUDRNA, D.; LEUNG, H.; LINE, R. F.; POWELL, W.; RAMSAY, L.; VIVAR, H.; WAUGH, R. Mapping quantitative and qualitative disease resistance genes in a doubled haploid population of barley (*Hordeum vulgare*). **Theoretical and Applied Genetics**, v. 101, n. 4, p. 580-589, 2000.

UDOP, União dos Produtores de Bioenergia, 2016 Disponível em: <<http://www.udop.com.br>>. Acesso em: 10 out 2016.

USDA, United States Department of Agriculture, 2016. Disponível em: <https://www.usda.gov>. Acesso em: 27 nov. 2016.

VISIONI, A.; TONDELLI, A.; FRANCA, E.; PSWARAYI, A.; MALOSETTI, M.; RUSSELL, J.; THOMAS, W.; WAUGH, R.; PECCHIONI, N.; ROMAGOSA, I.; COMADRAN, J. Genome-wide association mapping of frost tolerance in barley (*Hordeum vulgare* L.). **BMC Genomics**, v. 14, n. 1, p. 424, 2013.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITERS, A.; POT, J.; PALEMAN, J.; KUIPER, M. AFLP: a new technique for DNA fingerprinting. **Nucleic acids research**, v. 23, n. 21, p. 4407-4414, 1995.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITERS, A.; POT, J.; PALEMAN, J.; KUIPER, M. AFLP: a new technique for DNA fingerprinting. **Nucleic acids research**, v. 23, n. 21, p. 4407-4414, 1995.

WANG, G. J.; CASTIGLIONE, S.; ZHANG, J.; FU, R. Z.; MA, J. S.; LI, W. B.; SUN, Y. R.; SALA, F. HYBRID RICE (ORYZA-SATIVA L) - IDENTIFICATION AND PARENTAGE DETERMINATION BY RAPD FINGERPRINTING. **Plant Cell Reports**, v. 14, n. 2-3, p. 112-115, 1994.

WEI, X.; JACKSON, P. A.; HERMANN, S.; KILIAN, A.; HELLER-USZYNSKA, K.; DEOMANO, E. Simultaneously accounting for population structure, genotype by environment interaction, and spatial variation in marker-trait associations in sugarcane. **Genome**, v. 53, n. 11, p. 973-981, 2010/11/01, 2010.

WEI, X.; JACKSON, P. A.; MCINTYRE, C. L.; AITKEN, K. S.; CROFT, B. Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. **Theoretical and Applied Genetics**, v. 114, n. 1, p. 155-164, 2006.

WÜRSCHUM, T. Mapping QTL for agronomic traits in breeding populations. **Theoretical and Applied Genetics**, v. 125, n. 2, p. 201-210, 2012.

XAVIER, M. A. Ponto de mérito de família, parâmetros genéticos e identificação de genitores masculinos em policruzamentos de cana-de-açúcar. 2011. 92 f. Tese de doutorado -Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2011.

YU, J.; PRESSOIR, G.; BRIGGS, W. H.; BI, I. V.; YAMASAKI, M.; DOEBLEY, J. F.; MCMULLEN, M. D.; GAUT, B. S.; NIELSEN, D. M.; HOLLAND, J. B. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. **Nature genetics**, v. 38, n. 2, p. 203-208, 2006.

CAPÍTULO 2 – ESTIMATIVA DE PARÂMETROS GENÉTICOS PARA CARACTERÍSTICAS TECNOLÓGICAS DE CANA-DE-AÇÚCAR POR MEIO DO REML/BLUP E ANÁLISE MULTIVARIADA PARA DIFERENTES ÉPOCAS DE COLHEITA

RESUMO - A estimativa de parâmetros genéticos são de grande contribuição para um programa de melhoramento, uma vez que auxiliam na escolha do método mais adequado, nos caracteres a serem selecionados no início e em fases avançadas de melhoramento, bem como na predição dos ganhos genéticos com a seleção. Diante disto, o objetivo do trabalho foi estimar os parâmetros genéticos e prever os valores genotípicos de 100 genótipos de cana-de-açúcar por meio do REML/BLUP (Máxima Verossimilhança Restrita/Melhor Preditor Linear não Viesado) e análise de componentes principais para selecionar genótipos superiores para dois perfis agrônômicos distintos: indivíduos com altos teores de fibra (grupo 1) e indivíduos com potencial para acumular sacarose (grupo 2). Foram conduzidos três experimentos correspondentes às três épocas de colheita: maio, julho e setembro em delineamento de blocos casualizados com quatro repetições. Em cada experimento foram realizadas análises para as seguintes variáveis tecnológicas: Brix, Pol, pureza, açúcar redutor (AR) e fibra. Foi realizada uma análise de variância multivariada (Manova) considerando as cinco variáveis tecnológicas e posteriormente uma análise de componentes principais com o propósito de discriminar os grupos de genótipos para produção de açúcar e fibra. Os genótipos diferiram entre si considerando todas as variáveis de uma vez. Na análise conjunta dos três experimentos para o grupo 1, observou-se que o genótipo IN8458 apresentou o maior valor genotípico predito de fibra e o IACSP043150 o maior valor genotípico de Pol. Quanto aos parâmetros genéticos, as herdabilidades de média de genótipos de Pol e fibra apresentaram alta magnitude, indicando possibilidade de seleção com alta acurácia. No grupo 2, houve mudanças de ordenamento entre genótipos de acordo com as épocas e colheita. Além disto, houve genótipos com curto a longo período útil de industrialização.

Palavras-chave: *Saccharum* spp, modelo misto, componentes principais, valores genotípicos, fibra

CHAPTER 2 - ESTIMATION OF GENETIC PARAMETERS FOR TECHNOLOGICAL CHARACTERISTICS OF THE SUGARCANE BY REML/BLUP AND MULTIVARIATE ANALYSIS FOR DIFFERENT HARVEST SEASONS

ABSTRACT - The estimation of genetic parameters are a great contribution to a breeding program since they help out in terms of the choosing the most appropriate breeding method, characters to be selected at the beginning and in the advanced stages of breeding program as well as in the prediction of the genetic gains with the selection. The aim of this work was to estimate the genetic parameters and to predict the genotypic values of 100 sugarcane genotypes using REML/BLUP (Maximum Restricted Likelihood/Best Unbiased Linear Predictor) and principal component analysis to select Superior genotypes for two distinct agronomic profiles individuals with high fiber contents (group 1) and individuals with potential to accumulate sucrose (group 2). Three experiments were conducted corresponding to the three harvest periods: May July and September in a randomized complete block design with four replicates. For each experiment the following technological variables were analyzed Brix, Pol, purity, reducing sugar (AR) and fiber. A multivariate analysis of variance (Manova) was performed considering the five technological variables and later an analysis of principal components to discriminate the groups of genotypes for sugar and fiber production. The genotypes differed by considering all variables at one time. In the joint analysis of the three experiments for group 1 it was observed that genotype IN8458 presented the highest genotypic value of fiber and IACSP043150 the highest genotypic value of Pol. Regarding the genetic parameters the heritability of genotypes of Pol and Fiber presented high magnitude indicating possibility of selection with high accuracy. In group 2 there were changes of ordering among genotypes between the seasons and harvest. In addition, there were genotypes with short to long useful periods of industrialization.

Keywords: *Saccharum* spp, mixed model, principal componente, genotipic values, fiber

1 INTRODUÇÃO

O cultivo da cana-de-açúcar possui destaque tanto para a produção de álcool quanto para a produção de açúcar. Segundo levantamento da Conab (2016) a produção estimada para o ano agrícola 2016/2017 será de aproximadamente 694.544,8 milhões toneladas de cana comparado aos 665.586,2 milhões de toneladas da safra 2015/2016. Já a produção de açúcar foi de 33.489,1 toneladas de açúcar, enquanto que a de etanol total, incluindo o etanol anidro e o hidratado foi 30.461.524,5 milhões de litros.

A colheita da cana no Brasil é realizada de abril a novembro na Região Centro-Sul (SILVA; JERONIMO; LÚCIO, 2008). Devido ao longo período de colheita, mudanças climáticas, bem como disponibilidade de água e alterações de temperatura influenciam diretamente no acúmulo de sacarose dos genótipos sendo que estas alterações estão relacionadas às diferentes estações do ano, como: outono, inverno e primavera. Dessa forma, é importante para a indústria sucroalcooleira colher a cana na época de maior acúmulo de sacarose proporcionando maior retorno econômico. Por essa, a avaliação de genótipos ao longo de toda a safra é fundamental para o estudo de interação de genótipo x época de colheita em diferentes safras e locais.

Os atributos tecnológicos que definem a qualidade da cana-de-açúcar utilizada como matéria-prima para a fabricação de açúcar e álcool são: Brix caldo, Pol cana, Pureza e ART (Açúcares redutores totais) (LAVANHOLI, 2010).

O Brix representa o teor de sólidos solúveis contidos em uma solução pura de sacarose na solução expressa em porcentagem (peso/peso). A Pol consiste na porcentagem aparente de sacarose contida em uma solução de açúcares (caldo da cana) e é calculado com base no teor de Pol extraído do caldo corrigido e pela porcentagem de fibra. E a pureza é a porcentagem de sacarose contida nos sólidos solúveis.

A fibra consiste na matéria insolúvel em água contida na cana. Segundo Fernandes (2003), sua importância está no aspecto agrícola, em que cultivares ricas em fibra possuem maior resistência ao tombamento e incidência de pragas que afetam os colmos. Quanto ao aspecto industrial, a sua importância está referida cogeração à de energia que é utilizada na própria usina. Genótipos de cana-de-açúcar que

possuem altos teores de fibra têm se tornado importantes na obtenção de cultivares para a produção de etanol celulósico ou etanol de segunda geração. Enquanto que a maior quantidade de Pol é um parâmetro tecnológico para selecionar materiais destinados à produção de açúcar.

No processo de seleção, a estimativa de componentes de variância e dos parâmetros genéticos são informações importantes porque permitem saber a natureza dos genes envolvidos no controle dos caracteres mensurados, identificar qual método mais adequado a ser adotado em um programa de melhoramento para obtenção dos ganhos genéticos com manutenção da base genética adequada na população (CRUZ; CARNEIRO; REGAZZI, 2014). Entre os parâmetros genéticos de maior importância destacam-se as variâncias genéticas aditiva e não aditiva, as correlações e as herdabilidades. Com relação às variâncias genéticas, as amostras referentes aos genótipos e aos ambientes de experimentação precisam ser bem representativas da população e da área geográfica de interesse (ROBINSON; COCKERHAM, 1965 apud MORAIS et al., 1996).

Diversos trabalhos mostram o emprego da metodologia REML (Máxima Verossimilhança Restrita) na estimação dos parâmetros genéticos juntamente com o BLUP (Melhor Preditor Linear não Viesado) na predição dos valores genéticos para selecionar indivíduos superiores em cana de açúcar (BARBOSA et al., 2005; OLIVEIRA et al., 2008; OLIVEIRA et al., 2011; XAVIER et al., 2014).

Em virtude dos experimentos de cana-de-açúcar serem conduzidos em vários locais, anos e épocas de colheita, a metodologia de modelos mistos é adequada permite predições mais acuradas do valor genotípico para os efeitos de interação, tornando a seleção mais eficiente.

Mediante o exposto, o objetivo do presente trabalho foi estimar os parâmetros genéticos e prever os valores genotípicos de 100 genótipos de cana-de-açúcar pela REML/BLUP e análise de componentes principais para selecionar genótipos superiores para dois perfis agrônômicos distintos: indivíduos com potencial para acumular sacarose e indivíduos com altos teores de fibra.

2 MATERIAL E MÉTODOS

2.1 Material vegetal e dados fenotípicos

Utilizou-se 100 genótipos de cana-de-açúcar, incluindo cultivares comerciais e clones elite contrastantes para Pol e fibra. (Tabela 1). Estes genótipos pertencem ao banco de germoplasma do Centro de Cana do Instituto Agrônomo de Campinas (IAC).

A população utilizada no experimento foi constituída por genótipos de diferentes programas de melhoramento, tais como: IAC, Centro de Tecnologia Canavieira (CTC), genótipos obtidos da hibridação entre genitores do IAC e CTC (IACCTC), Rede Interinstitucional de Desenvolvimento do Setor Sucroalcooleiro (Ridesa - RB) e antigo Copersucar (SP). Os seguintes genótipos são oriundos de programas da Nova Guiné: NG5712 (*Saccharum robustum*), NG57213 (*S. officinarum*), IJ76293 e IS76155. Além dos genótipos IN8482, IN8458 e IN8488 oriundos da Indonésia. Foi incluído também um genótipo selvagem, SES205 (*S. spontaneum*). E por fim, outros genótipos introduzidos: US571415 (*S. robustum*), Krakatau (*S. spontaneum*) e IJ76293 (*S. robustum*), além das seguintes cultivares IACSP953028, CTC 15, IACSO93-3046, IACSP96-3076, IACSP97-4039, RB85-5156, RB93-5744 e SP832847 e os demais 81 clones elite do Brasil.

Foram instalados três experimentos após o fim da safra, em dezembro de 2013 no Centro de Cana do Instituto Agrônomo de Campinas, em Ribeirão Preto, São Paulo. Cada experimento foi correspondente à uma época de colheita totalizando três épocas. O primeiro experimento corresponde à colheita realizada em maio, o segundo em julho e o terceiro em setembro. No ano de 2014 foram realizados três cortes, um em cada experimento, sendo o primeiro experimento cortado em maio, o segundo em julho e o terceiro em setembro, para que no ano de 2015, cada um dos três experimentos fossem colhidos com 12 meses a partir do corte de uniformização, ou seja, o primeiro colhido em maio, o segundo em julho e o terceiro em setembro, além desses foi colhido também a época 3 em setembro de 2016.

Os experimentos foram conduzidos em delineamento de blocos casualizados com 4 repetições e 100 tratamentos. Cada parcela foi constituída por uma linha de 1,5

metros com espaçamento de 1,5 m entre fileiras e 0,5 m entre mudas, contendo 3 mudas pré-brotadas por parcela. O clima é do tipo tropical com inverno seco, classificado de acordo como o Sistema Internacional de Classificação de Köppen, como Aw, a pluviometria média anual é de 1.422 mm e latossolo vermelho distroférico. A região onde foi conduzido o experimento refere-se ao ambiente de produção B2, com média de produtividade de 90t/ha em cinco cortes (Prado 2008).

As análises das variáveis tecnológicas: fibra, Pol cana, pureza, AR (Açúcar Redutor) e Brix (caldo da cana) foram realizadas a partir de uma amostra de dez colmos de cada parcela e adotada a metodologia segundo as Normas Técnicas de Determinação da Qualidade da Cana-de-açúcar e Álcool do Estado de São Paulo (CONSECANA, 2006).

2.2 Seleção de genótipos

Foi realizada uma análise de variância multivariada (Manova) em cada uma das três épocas de colheita para verificar se os 100 genótipos diferem entre si considerando as cinco variáveis tecnológicas no PROC GLM utilizando SAS 9.4 (SAS, 2013). O modelo adotado da Manova foi o seguinte: $Y_{ijr} = \mu_r + g_{ir} + b_{jr} + e_{ijr}$, em que Y_{ijr} é o valor observado da r -ésima variável do i -ésimo genótipo no j -ésimo bloco, μ_r é a média geral da r -ésima variável, g_{ir} é o efeito do i -ésimo genótipo na r -ésima variável, b_{jr} é o efeito do j -ésimo bloco na r -ésima variável e e_{ijr} é o efeito aleatório associado à observação a Y_{ijr} .

Para a verificação das hipóteses de igualdade dos genótipos foram utilizados os seguintes testes: teste de Wilks (Wilks, 1932), Traço de Pillai, Traço de Hottelling-Lawley e a Maior Raiz característica de Roy. Estes testes servem para justificar a análise por componentes principais (ACP), desde que pelo menos três sejam significativos.

Tabela 1: Lista dos 100 genótipos de cana-de-açúcar (*Saccharum* spp.) utilizados nos três experimentos, colhidos em 2015 e suas respectivas identificações.

Nº	Genótipo	Genitor feminino	Genitor masculino	Origem	Nº	Genótipo	Genitor feminino	Genitor masculino	Origem
1	US571415	Desconhecido	Desconhecido	EUA	51	IACCTC069708	IACSP962042	CT014455	Brasil
2	CANA ALHO	Desconhecido	Desconhecido		52	IACCTC069713	IACSP955000	CT961307	Brasil
3	IACSP012410	R570	IAC873396	Brasil	53	IACCTC069741	IACSP955000	RB72454	Brasil
4	IACSP046007	PO8862	CTC9	Brasil	54	IACSP012417	RB835486	SP775181	Brasil
5	IACSP046059	IACSP953018	SP963092	Brasil	55	IACSP012430	IAC912195	SELF	Brasil
6	IACSP972053	SP842066	IAC863154	Brasil	56	IACSP015501	SP913059	CTC957	Brasil
7	IACSP953028	IACSP953028	RB855156	Brasil	57	IACSP015519	RB835486	SP775181	Brasil
8	NG5712	<i>S.robustum</i>	Desconhecido	Indonésia	58	IACSP018046	R570	IAC873396	Brasil
9	IACSP042516	IAC912288	SP80185	Brasil	59	IACSP018082	IAC913093	SP87365	Brasil
10	KRAKATAU	<i>S.spontaneum</i>	Desconhecido	Indonésia	60	IACSP018158	RB835486	SP775181	Brasil
11	IN8482	<i>S.spontaneum</i>	Desconhecido	Indonésia	61	IACSP022067	SP973060	SP701143	Brasil
12	IACSP042508	SP913059	IACSP972077	Brasil	62	IACSP022125	SP901161	SP901644	Brasil
13	IACSP955037	SP842066	IAC873396	Brasil	63	IACSP023025	SP924230	SP88813	Brasil
14	IACCTC059630	IACSP972053	RB85453	Brasil	64	IACSP023168	IAC914216	SP803280	Brasil
15	IACSP046053	IACSP956114	IACSP933046	Brasil	65	IACSP042503	PO8862	CTC9	Brasil
16	IACSP993009	CP701133	SP832847	Brasil	66	IACSP042504	IACSP955037	Desconhecido	Brasil
17	IS76155	Desconhecido	Desconhecido	Indonésia	67	IACSP042509	SP913059	IACSP973391	Brasil
18	IACSP012421	RB835486	SP775181	Brasil	68	IACSP042510	IAC913093	CTC9638	Brasil
19	IACSP042521	IAC912288	SP963092	Brasil	69	IACSP043123	IACSP953028	SP832847	Brasil
20	IACSP046058	IACSP953018	SP963092	Brasil	70	IACSP043148	IACSP953018	IACSP933046	Brasil
21	IACSP991020	SP851682	SP80185	Brasil	71	IACSP043150	IACSP953018	IACSP933046	Brasil

22	IACSP045065	IACSP956114	IACSP953050	Brasil	72	IACSP043259	IACSP966026	RB855453	Brasil
23	IACCTC069767	IACSP963069	IACSP952078	Brasil	73	IACSP045081	IACSP966026	RB855453	Brasil
24	IACSP985046	SP813251	SP823697	Brasil	74	IACSP046032	IACSP953028	SP775181	Brasil
25	IACSP018034	IAC913093	SP87432	Brasil	75	IACSP046035	SP913059	IACSP966026	Brasil
26	IACSP993357	SP921584	Desconhecido	Brasil	76	IACSP046073	IACSP953028	SP924221	Brasil
27	IACCTC059578	WHITETRANS	SP803280	Brasil	77	IACSP046077	IACSP972109	IACSP956114	Brasil
28	IACCTC055580	CTC4	IACSP933046	Brasil	78	IACSP046152	IAC911099	SP931124	Brasil
29	IACSP976628	RB835486	SELF	Brasil	79	IACSP933046	SP791011	Desconhecido	Brasil
30	IACSP976680	SP842066	SP847017	Brasil	80	IACSP953018	SP842189	SP801842	Brasil
31	IACCTC059534	IACSP953028	CT961414	Brasil	81	IACSP956114	IAC873187	CTC9019	Brasil
32	IACSP972084	RB855035	SP80185	Brasil	82	IACSP962008	SP80144	Desconhecido	Brasil
33	CTC15	SP842025	Desconhecido	Brasil	83	IACSP963056	SP826108	Desconhecido	Brasil
34	IAC914168	SP813137	Desconhecido	Brasil	84	IACSP963069	SP803280	Desconhecido	Brasil
35	IACBIO232	RB855465	GLAGAH	Brasil	85	IACSP963076	SP847017	Desconhecido	Brasil
36	IACBIO241	RB855465	GLAGAH	Brasil	86	IACSP973384	RB855113	RB855453	Brasil
37	IACBIO257	NG26011	Desconhecido	Brasil	87	IACSP974039	RB835486	RB855453	Brasil
38	IACBIO266	NG26011	Desconhecido	Brasil	88	IACSP974048	SP842066	IAC863154	Brasil
39	IACBIO270	SES069	Desconhecido	Brasil	89	IACSP982053	SP847017	SP801842	Brasil
40	IACBIO271	SES069	Desconhecido	Brasil	90	IACSP983011	IAC913093	SP823530	Brasil
41	IACBIO273	SES069	Desconhecido	Brasil	91	IACSP993369	SP901616	Desconhecido	Brasil
42	IACBIO275	SES069	Desconhecido	Brasil	92	IACSP994011	SP842025	RB855113	Brasil
43	IACBIO277	SES069	Desconhecido	Brasil	93	IJ76293	<i>S.robustum</i>	Desconhecido	Java
44	IACBIO279	SES069	Desconhecido	Brasil	94	IN8458	<i>S.spontaneum</i>	Desconhecido	Indonésia
45	IACCTC053616	IAC915155	CT961414	Brasil	95	IN8488	<i>S.spontaneum</i>	Desconhecido	Indonésia
46	IACCTC056518	SP911049	CT931231	Brasil	96	NG57213	Desconhecido	Desconhecido	Nova Guiné
47	IACCTC059552	IACSP955011	CT961414	Brasil	97	RB855156	RB72454	TUC717	Brasil

48	IACCTC059607	IACSP972053	SP88721	Brasil	98	RB935744	RB835089	RB765418	Brasil
49	IACCTC059634	IACSP972053	RB855453	Brasil	99	SES205A	Desconhecido	Desconhecido	
50	IACCTC061050	CT943165	IAC911099	Brasil	100	SP832847	HJ5741	SP701143	Brasil

O modelo estatístico univariado para um delineamento em blocos para cada época de colheita foi: $Y = Xr + Zg + e$, em que Y é o vetor de dados, r é vetor do efeito fixo de blocos somados à média geral, g é vetor dos efeitos de genótipos (assumidos como aleatórios), e é o vetor de erros ou resíduos (aleatórios) e X e Z são as matrizes de incidência dos vetores: r e g , respectivamente.

Como os genótipos são pertencentes a diferentes origens, como diversos programas de melhoramento do Brasil e de outros países, foi realizada antes a discriminação dos genótipos via análise de componentes principais para a estimação dos parâmetros genéticos levando-se em conta as seguintes variáveis tecnológicas: Brix, Pol, pureza, fibra, e AR.

A análise de componentes principais (ACP) foi realizada gerando os autovalores e autovetores em que um gráfico biplot foi gerado para verificar a relação entre cada genótipo e as variáveis avaliadas. Os autovalores são as variâncias associadas aos componentes principais. Os autovetores fornecem os coeficientes das variáveis padronizadas (média= zero, $\sigma^2=1$) e avaliadas para calcular as combinações lineares dos componentes principais. O biplot é uma generalização do gráfico de dispersão para visualizar e estudar a interrelação entre observações (genótipos) com variáveis avaliadas em experimentos (HAIR, 2009). A análise de componentes principais como o gráfico biplot foram realizadas por meio software R (R Development Core Team, 2015).

2.3 Estimação dos parâmetros genéticos

A estimativa dos parâmetros genéticos como herdabilidade e correlação genotípica foi realizada por meio de modelos mistos, com a metodologia REML/BLUP. O modelo do presente experimento determina que o efeito de genótipos é aleatório, já os de blocos e épocas de colheita, são efeitos fixos. Adotou-se modelo que utiliza a seguinte equação matricial em um delineamento de blocos ao acaso: $Y = Xr + Zg + Wge + e$, em que Y é o vetor de dados, r é vetor de efeitos fixos (média de blocos através das épocas) somados à média geral, g é vetor dos efeitos genotípicos (assumidos como aleatórios), ge é o vetor dos efeitos da interação genótipo x épocas

(assumidos como aleatórios) e e é o vetor de erros ou resíduos (aleatórios) e X , Z e W são as matrizes de incidência dos vetores: r , g e ge , respectivamente.

Para o ranqueamento individual dos genótipos, utilizou-se os valores genotípicos preditos de Pol e fibra, considerado o valor mais conveniente do ponto de vista de ganho genético. A distribuição de médias e variâncias foi conforme Resende e Duarte (2007).

Os valores genotípicos preditos na análise conjunta com interação são compostos por: $\mu_{ei} + g + gem$, nos quais μ representa a média de todas as épocas para determinada variável tecnológica, g é efeito do genótipo e gem é o efeito médio da interação genótipo x época. Para cada época os valores genotípicos preditos são compostos por $\mu_{ei} + g$, neles μ_{ei} é a média para cada época, g_e (efeito do genótipo).

A análise de variância multivariada foi feita utilizando o PROC GLM, ao passo que a estimação dos efeitos fixos e predição dos efeitos aleatórios dos cálculos de modelos mistos foram executadas no PROC MIXED utilizando o SAS 9.4 (SAS, 2013).

3 RESULTADOS E DISCUSSÃO

Os resultados obtidos por meio de análise multivariada da variância (Manova), mostraram que os genótipos se diferem significativamente pelos testes de Lambda de Wilks, Traço de Pillai, Traço de Hotteling-Lawley e a Maior Raiz de Roy (Tabela 2).

As análises de componentes principais para as características tecnológicas em cada uma das três épocas de colheita (maio, julho e setembro) explicaram mais do que 90% da variância contida nas variáveis originais. No entanto, na época colhida em setembro de 2015 a discriminação dos grupos foi superior. O componente principal 1 (CP1) explicou 93,7% da variação total contida nas variáveis originais e discriminou melhor os genótipos com maiores teores de fibra (grupo 1) dos genótipos com maiores teores de sacarose (grupo 2). Para o mesmo componente, os autovetores variaram de -0,38 a 0,42, isto é, apresentaram poder discriminatório semelhante em módulo. Já o componente principal 2 (CP2) contribuiu com 4,2% da variância total contida nas variáveis originais e a fibra foi a que apresentou maior autovetor, -0,8566 (Tabela 3).

O grupo 1 foi formado por 18 genótipos, em que destacam-se aos genótipos 39, 40, 41, 42, 43 e 44 que são meios-irmãos, assim como o 35 e 36 por possuírem

similaridade genética já que têm o mesmo genitor feminino listados na Tabela 1. A maior dispersão observada neste grupo, conforme o biplot, foi provavelmente por descenderem das espécies *Saccharum spontaneum* (Figura 1). Além disto, os altos teores apresentados por estes genótipos foram acima de 17%, superiores às principais cultivares utilizadas no Brasil, segundo Silveira et al. (2015), com perfil voltado para a geração de energia dependendo dos teores dos componentes da fibra como lignina ou para a obtenção de etanol celulósico dependendo dos teores de celulose ou hemicelulose. Os outros 82 genótipos mais ricos para os atributos que conferem a maior produção de açúcar (Pol, pureza e Brix) foram alocados no grupo 2 e descendem de *Saccharum officinarum*.

Tabela 2. Modelo da análise multivariada de variância (Manova) com quatro os testes de hipóteses para os 100 genótipos colhidos na 3ª época de (Setembro de 2015).

Estatística	Valor	F	Num DF	Den DF	Pr > F
Teste de Wilks	0,0069	4,53*	495	1309,4	<,0001
Traço de Pillai	2,7135	3,18*	495	1325	<,0001
Traço de Hotelling -Lawley	14,6501	7,68*	495	1203,2	<,0001
Maior Raiz de Roy	10,9762	29,38*	99	265	<,0001

*Significativo ao nível de 5% pelo teste F.

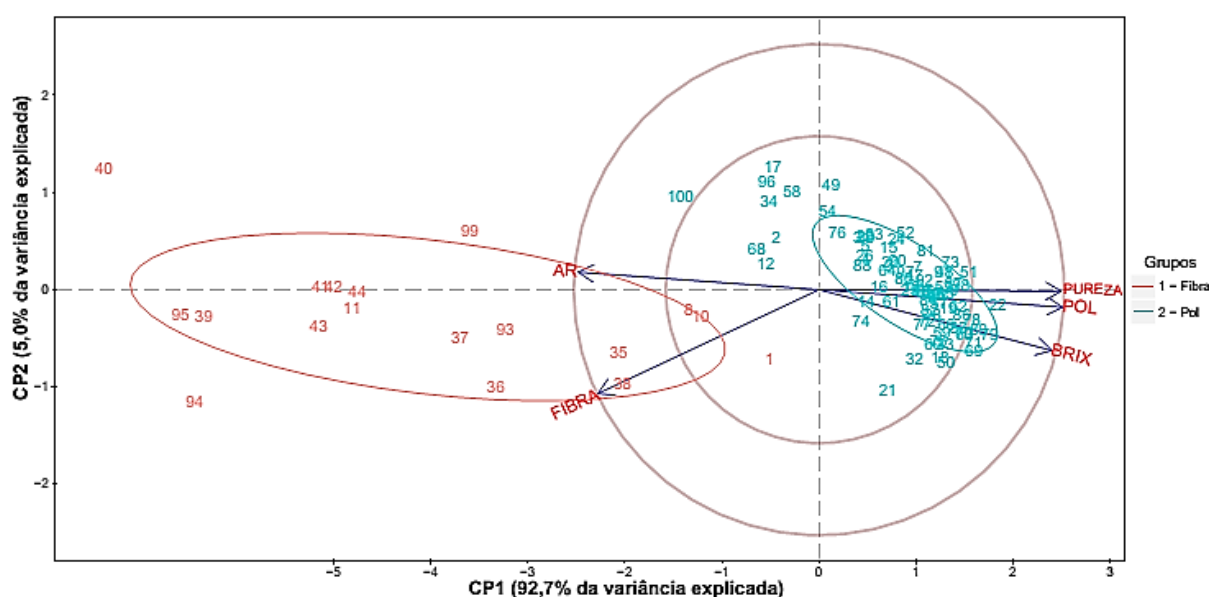


Figura 1. Análise de componentes principais das características tecnológicas para os 100 genótipos de cana-de-açúcar na 3ª época (colhida em setembro de 2015).

Tabela 3. Autovetores obtidos da matriz de correlação para os cinco atributos tecnológicos: Brix (caldo), fibra, pureza, Pol (pol cana) eAR (Açúcar Redutor) para os 100 genótipos de cana-de-açúcar na 3ª época (colhida em setembro de 2015).

Atributos	Autovetores para cada componente principal (CP)	
	CP1	CP2
Brix	0,4407	-0,4941
Fibra	-0,4187	-0,8456
Pureza	0,4585	-0,0119
Pol%cana	0,4610	-0,1425
AR	-0,4557	0,1429

O biplot mostra que os genótipos que se situam mais abaixo (negativos) no CP2 e à esquerda no CP1 (negativos), são mais semelhantes quanto ao teor de fibra e mais contrastantes em relação aos genótipos do grupo característico para o teor de Pol (Figura 1). Os genótipos mais à direita do CP1, localizados no interior da elipse menor, são mais responsivos à produção de açúcar (grupo 2). No entanto foi observado que o SP832847 (100) localizado fora do primeiro círculo à esquerda do CP1 e acima do CP2 apresentou o menor teor de Pol dentre os 82 genótipos na análise conjunta das três épocas de colheita. O genótipo IACSP991020 (21), situado fora da menor elipse, à direita do CP1 e abaixo do CP2, apresentou o segundo maior teor de fibra e o genótipo IACCTC059634 (49), acima do CP2 e mais próximo de zero tomando como referência o CP1, apresentou a menor porcentagem de fibra abaixo dos dez primeiros, dados não mostrados no artigo.

No CP2, a variável fibra foi a que melhor discriminou os genótipos do grupo 1. O genótipo IN8458 (94) que está fora da elipse do grupo 1, abaixo do CP2 e à esquerda do CP1, apresentou o maior teor de fibra na análise conjunta e está entre os dez primeiros para todas as épocas de colheita (Tabela 4). Enquanto que o US571415 (1), no interior da maior elipse e mais próximo de zero, apresentou o menor teor de fibra e a maior Pol na análise conjunta e valores entre os dez primeiros para todas as épocas de colheita contrastando com o genótipo IACBIO271 (40), fora da maior elipse com menor valor dentre os 18 genótipos, não listado na Tabela 4.

Santchurn et al. (2012) também mostraram o uso da ACP na identificação e seleção de genótipos de cana-de-açúcar derivados de *S.officinarum*, *S.spontaneum*, *S. robustum*, *S. sinense*, *Miscanthus* e *Erianthus* para diferentes finalidades, seja para

a produção de açúcar, obtenção de fibra ou para ambos os perfis. Foram observados 77% da variação total explicada pelos dois primeiros CPs, a partir da análise de 18 variáveis, incluindo variáveis tecnológicas, morfológicas e de produção.

Os parâmetros genéticos estimados para o grupo 1 indicaram um coeficiente de variação residual (CV_r) variando de 8,70% a 14,79% para fibra e de 25,61% a 47,99% para Pol listados nas tabelas 4 e 5, respectivamente. Comparado aos teores de 5,94% e 6,68% de fibra e Pol encontrados por Moraes et al. (2010), os valores foram considerados altos, uma vez que alguns genótipos apresentaram valores contrastantes para as duas variáveis, contribuindo assim para um alto valor do CV_r . No grupo 2, foram obtidos valores de CV_r , variando de 5,19% a 7,22% para fibra e de 7,94% a 10,66% para Pol, valores que são considerados altos aos apresentados por Moraes et al.(2010). Comparando aos resultados de Oliveira (2007) com relação à seleção de famílias para a maturação precoce, as variâncias genéticas observadas foram abaixo de 10% para Pol e fibra todas as épocas de colheita.

Tabela 4. Estimativas dos componentes de variância e parâmetros genéticos para fibra para o grupo 1 (18 genótipos) e grupo 2 (82 genótipos) nas três épocas de colheita.

Parâmetros genéticos	Época 1	Época 2	Época 3	Época 3 (2016)	Época 1	Época 2	Época 3	Época 3 (2016)	Análise conjunta	
	G1*				G2*				G1*	G2
σ_f^2	5,81	7,75	10,39	3,74	1,31	1,00	0,95	1,23	6,96	1,12
σ_g^2	1,93	3,08	2,30	0,78	0,69	0,35	0,59	0,64	1,98	0,51
σ_{ge}^2									0,00	0,04
C_g^2									0,00	0,04
σ_r^2	3,87	4,67	8,09	2,96	0,62	0,65	0,37	0,59	4,98	0,56
$\hat{h}_g^2$	33,27	39,75	22,15	20,90	52,52	34,97	61,59	51,72	0,28	0,46
h_{mc}^2	66,60	72,52	53,23	51,38	81,56	68,26	86,51	81,08	86,44	91,74
$A_{cgen}(\%)$	81,61	85,16	72,96	71,68	90,31	82,62	93,01	90,04	92,97	95,78
$CV_g\%$	8,16	9,91	7,89	4,47	7,25	5,30	6,57	6,43	7,67	6,15
$CV_r\%$	11,55	12,20	14,79	8,70	6,89	7,22	5,19	6,21	12,16	6,42
CV_g/CV_r	0,71	0,81	0,53	0,51	1,05	0,73	1,27	1,04	0,63	0,96
$\hat{r}_{gloc}$									1,00	0,92
PEV	0,65	0,85	1,08	0,38	0,13	0,11	0,08	0,12	0,27	0,04
SEP	0,80	0,92	1,04	0,62	0,36	0,33	0,28	0,35	0,52	0,21
MG	17,04	17,71	19,23	19,78	11,42	11,19	11,66	12,41	18,36	11,67

: G1= Grupo 1, G2= Grupo 2. σ_f^2 : variância fenotípica individual; σ_g^2 : variância genotípica; σ_{ge}^2 : variância da interação genótipo x época; C_g^2 : coeficiente de determinação dos efeitos da interação genótipo x época; σ_r^2 : variância residual; h_{mc}^2 : herdabilidade da média de genótipo; $\hat{h}_g^2$: herdabilidade individual no sentido amplo; A_{cgen} : acurácia da seleção de genótipos; CV_g : coeficiente de variação genético; CV_r : coeficiente de variação relativa; $\hat{r}_{gloc}$: correlação genotípica entre o desempenho dos genótipos nas várias épocas; PEV: variância do erro de predição; SEP: desvio padrão do valor predito dos genótipos; MG: média geral.

Tabela 5. Estimativas dos componentes de variância e parâmetros genéticos para Pol para o grupo 1 (18 genótipos) e grupo 2 (82 genótipos) nas três épocas de colheita.

Parâmetros genéticos	Época 1	Época 2	Época 3	Época 3-2016	Época 1	Época 2	Época 3	Época 3-2016	Análise conjunta	
	G1				G2				G1	G2
σ_f^2	6,92	13,91	17,24	8,93	2,50	3,13	4,50	3,07	11,71	3,30
σ_g^2	3,28	4,97	5,90	4,82	1,29	0,48	1,72	1,21	4,13	1,12
σ_{ge}^2									0,49	0,08
C_g^2									0,04	0,02
σ_r^2	3,63	8,94	11,33	4,12	1,21	2,65	2,78	1,86	7,08	2,10
$\hat{h}_g^2$	47,48	35,71	34,25	53,93	51,59	15,42	38,32	39,45	0,35	0,34
h_{mc}^2	78,33	68,96	67,57	82,40	81,00	42,18	71,31	72,27	87,96	88,03
$A_{cgen}(\%)$	88,51	83,04	82,20	90,78	90,00	64,95	84,44	85,01	93,79	93,82
$CV_g\%$	25,56	29,65	34,57	27,71	8,20	4,55	8,02	6,60	27,30	6,81
$CV_r\%$	26,89	39,79	47,90	25,61	7,94	10,66	10,17	8,18	35,73	9,34
CV_g/CV_r	0,95	0,75	0,72	1,08	1,03	0,43	0,79	0,81	0,76	0,73
$\hat{r}_{gloc}$									0,89	0,93
PEV	0,71	1,54	1,91	0,85	0,25	0,28	0,49	0,34	0,50	0,13
SEP	0,84	1,24	1,38	0,92	0,50	0,53	0,70	0,58	0,71	0,37
MG	7,09	7,52	7,03	7,92	13,86	15,27	16,37	16,66	7,45	15,52

: G1= Grupo 1, G2= Grupo 2. σ_f^2 : variância fenotípica individual; σ_g^2 : variância genotípica; σ_{ge}^2 : variância da interação genótipo x época; C_g^2 : coeficiente de determinação dos efeitos da interação genótipo x época; σ_r^2 : variância residual; h_{mc}^2 : herdabilidade da média de genótipo; $\hat{h}_g^2$: herdabilidade individual no sentido amplo; A_{cgen} : acurácia da seleção de genótipos; CV_g : coeficiente de variação genético; CV_r : coeficiente de variação relativa; $\hat{r}_{gloc}$: correlação genotípica entre o desempenho dos genótipos nas várias épocas; PEV: variância do erro de predição; SEP: desvio padrão do valor predito dos genótipos; MG: média geral.

Os coeficientes de variação genético (CV_g) para fibra do grupo 1 foram menores que 10% para todas as épocas de corte (Tabela 4). O atributo tecnológico Pol apresentou valores de CV_g superiores a 10%, o que indica que há presença de variabilidade genética entre genótipos com possibilidade de seleção (RESENDE, 2002). Valores inferiores a 10% foram observados por Moraes et al. (2010) no estudo de famílias oriundas de cruzamentos biparentais de cana-de-açúcar no Litoral Norte de Pernambuco variando de 2% a 5% para atributos tecnológicos, exemplificando que há maior porcentagem de variação ambiental do que variação genética, reduzindo o sucesso com a seleção.

O CV_g obtido no grupo 2 para fibra e Pol também foi menor que 10%, indicando haver pouca variabilidade genética e com isso reduzindo a possibilidade de sucesso na seleção, conforme listados nas tabelas 4 e 5. Oliveira et al. (2005) encontraram valores de CV_g acima de 10% em três ambientes na seleção de genótipos de açúcar para toneladas por hectare. Com isso, fica demonstrada a possibilidade efetiva de seleção entre genótipos. Valores de CV_g para época 3 de colheita no ano agrícola 2015/2016 para o grupo 1 foi menor que 10% em fibra, enquanto para o teor de sacarose (Pol) foi 27,71%. No grupo 2 observou-se CV_g menor que 10% para fibra e Pol. A relação CV_g/Cv_e foi menor que 1 para a análise conjunta dos grupos 1 e 2, indicando haver alta influência do ambiente sobre a variação, dessa forma não existindo condição favorável para a seleção.

As herdabilidades individuais no sentido amplo (h_g^2) foram inferiores às herdabilidades da média de clone h_{mc}^2 para o grupo 1 e grupo 2. Os valores estimados à herdabilidade individual variaram de 20,90% a 39,75% para fibra no grupo 1 e de 34,97% a 61,59% no grupo 2, sendo considerada herdabilidade média de acordo com Resende (2002) (Tabela 4). Na Pol, no grupo 1, os valores variaram de 34,25% a 53,93% e no grupo 2 de 15,42% a 51,59%. Os valores estimados de h_g^2 na análise conjunta das épocas de colheita nos grupos 1 e 2 foram inferiores comparados aos apresentados por Mancini et al. (2010) em cana soca, enquanto que foram superiores aos apresentados por Oliveira (2007) avaliando famílias de irmãos germanos de cana-de-açúcar. Já a época 3, colhida em 2015, a h_g^2 estimada para fibra foi superior aos resultados apresentados por Xavier et al. (2014). A variação entre esses coeficientes

de herdabilidade para os mesmos caracteres está relacionada à população em estudo e das condições do ambiente onde foram avaliadas.

Os valores estimados de h_{mc}^2 para fibra variaram de 51,38% a 72,52% no grupo 1 e de 68,26% a 86,51% para o grupo 2 (Tabela 4). Para Pol, a variação foi de 67,57% a 82,40% (grupo 1) e de 42,18% a 81% (Tabela 5). Este tipo de herdabilidade é utilizada quando se usam médias como unidades de avaliação e seleção. A alta h_{mc}^2 observada no presente trabalho indica pouca influência do ambiente na expressão de fibra e Pol (DABHOLKAR, 1999). Segundo Resende (2002), estes valores são considerados altos e proporcionam boa acurácia, pois indicam maior porcentagem de componentes genéticos em relação ao ambiental, o que resulta em alta acurácia, favorecendo o sucesso com seleção.

A acurácia de seleção dos genótipos é calculada por meio da raiz quadrada da herdabilidade de média de genótipos. Este parâmetro é fundamental no aspecto da avaliação genotípica porque nos fornece o quanto o valor predito estimado de um está mais próximo do real valor genético (HENDERSON, 1984). O valor desta correlação varia de 0 a 1 e quanto mais próximo de 1 ou 100%, maior a acurácia na seleção e mais precisa é a inferência do valor predito. Portanto, os valores de acurácia variaram de 0,8 a 0,9 indicando uma ótima qualidade experimental e segurança na seleção, nos grupos 1 e 2 (Tabelas 4 e 5).

Os valores para a correlação genotípica dos genótipos nas três épocas de colheita $\hat{r}_{gloc}$ para o grupo 1 foram 1 e 0,84 e para o grupo 2: 0,90 e 0,89. Estes valores correspondem a fibra e pol, respectivamente. Estas correlações foram de alta magnitude, quando comparadas ao valor de correlação genotípica de 0,62 e 0,49 encontrado por Oliveira et al. (2005) e Bastos et al. (2007) avaliando genótipos de cana em vários ambientes, os quais revelam moderado nível de interação genótipo x ambiente.

A predição dos valores genéticos via BLUP (Melhor Preditor Linear não Viesado) dentro de cada grupo permitiu o ordenamento dos genótipos por cada época considerando as variáveis Pol e fibra. Foram apresentados os 10 genótipos com os maiores g (BLUPS) para Pol e para fibra. Os g (BLUPS) são os valores genéticos preditos que somados a média ajustada resultam no valor genotípico. Valores com asterisco (*) diferem de zero ($p < 0,05$) pelo teste t.

Na avaliação da fibra no ano de 2015 para o grupo 1, formado por 18 genótipos (responsivos para o atributo fibra), o IN8458 (94) apresentou os maiores valores preditos para a 1ª e 3ª época, porém na 2ª época houve alteração no ordenamento, permanecendo o genótipo 43 (IACBIO277) como de maior valor genotípico. Com relação ao valor predito da análise conjunta das três épocas de colheita, além da colheita realizada em setembro de 2015, o genótipo 94(IN8458), também apresentou o maior valor de teor de fibra, enquanto que o genótipo 99 (SES205A) obteve o menor valor predito dentre dez listados na Tabela 6.

Segundo Rao e Weerathaworn (2009), o teor de fibra aumenta paralelamente ao aumento da idade da cana. Os mesmos autores identificaram em um trabalho de desenvolvimento de cultivares para aptidão tanto para fibra como para produção de sacarose que a cultivar tailandesa MPT99-648 no nono mês apresentou 16,5% de fibra, mas mostrou incremento de 2% do 11º para o 12º mês. Já resultados contrários, segundo Elibox (2011), é que teores de fibra têm se mantido estáveis ao longo de 4 a 12 anos por genótipo nos repetitivos ensaios durante várias socas, diferentes locais em cinco países, são eles: Barbados, Belize, República Dominicana, Guiana e Jamaica. Entre os genótipos usados pelo mesmo autor no estudo de herança, também foi verificado que os teores de fibra são estáveis para os genótipos nos experimentos em diversos locais e anos.

Tabela 6: Ranking dos dez melhores genótipos do grupo 1 de cana-de-açúcar e predição dos valores genotípicos para o atributo fibra.

Ranking	Análise conjunta		Época 1		Época 2		Época 3		Época 3*	
	Gen	g	Gen	g	Gen	g	Gen	g	Gen	g
1	94	3.0633*	94	2,2254*	43	2,534*	94	2,8245*	94	1,0927 ^{ns}
2	43	2.0991*	41	1,5161 ^{ns}	41	2,1569*	39	1,2955 ^{ns}	43	0,9180 ^{ns}
3	39	1.3328*	43	1,348 ^{ns}	94	2,0184 ^{ns}	95	1,2276 ^{ns}	40	0,8242 ^{ns}
4	95	0.9526 ^{ns}	39	1,1931 ^{ns}	8	1,909 ^{ns}	43	1,1065 ^{ns}	95	0,4492 ^{ns}
5	41	0.8345 ^{ns}	11	0,8959 ^{ns}	44	1,283 ^{ns}	36	0,4065 ^{ns}	39	0,3569 ^{ns}
6	36	0.2673 ^{ns}	95	0,4522 ^{ns}	39	0,5269 ^{ns}	11	0,3572 ^{ns}	36	0,2398 ^{ns}
7	44	0.2602 ^{ns}	37	0,07922 ^{ns}	99	0,01022 ^{ns}	44	-0,0446 ^{ns}	99	0,0070 ^{ns}
8	11	-0.01796 ^{ns}	42	-0,0723 ^{ns}	95	-0,0411 ^{ns}	38	-0,207 ^{ns}	8	0,0000 ^{ns}
9	40	-0.3303 ^{ns}	99	-0,0723 ^{ns}	36	-0,0478 ^{ns}	40	-0,2376 ^{ns}	44	-0,0864 ^{ns}
10	99	-0.3891 ^{ns}	36	-0,2421 ^{ns}	93	-0,4863 ^{ns}	37	-0,38 ^{ns}	11	-0,1296 ^{ns}
Media Geral		18,3565	17,04105		17,7109		19,2264		19,7782	

*: significativo ao nível de 5% pelo teste t; **ns**: não significativo; **g**= BLUP. **8**=NG5712; **11**=IN8482; **36**=IACBIO241; **37**=IACBIO257; **38**=IACBIO266; **39**=IACBIO270; **40**=IACBIO271; **41**=IACBIO273; **42**=IACBIO275; **43**=IACBIO277; **44**=IACBIO279; **93**=IJ76293; **94**=IN8458; **95**=IN8488; **99**=SES205A. **Época 3 *** (colhida em 2016); **Gen**= Genótipos

Os resultados do grupo 1 destacam as combinações híbridas favoráveis para os genótipos com maior teor de fibra entre *S. officinarum* x *S. spontaneum* e *S. robustum* x *S. officinarum*. Genótipos pertencentes a *S. spontaneum* possuem uma alta porcentagem de alelos relacionados a teores de fibra, destinados para a obtenção de cultivares de “cana energia”. Isto indica que estes cruzamentos podem gerar genótipos com valores genéticos positivos para fibra.

Viator e Richard (2012) avaliando efeitos de data de plantio no teor de fibra em Louisiana, Estados Unidos, comparando cultivares para a produção com a L79-1002 (cana energia) detectaram valores de 2400 Kg/ha a mais de fibra no mês em relação aos plantios de setembro e outubro.

Os genótipos 41 (IACBIO273), 44 (IACBIO279), 11 (IN8482), 99 (SES205A), 8 (NG5712), 38 (IACBIO266) e 42 (IACBIO275) não estão entre os dez melhores para todas as épocas, indicando que a época de colheita influencia no ranking dos genótipos (Tabela 6). O genótipo 42 esteve apenas entre os dez melhores na época 1, o 8 e 93 na época 2 e o 38 e 40, na época 3. Embora a análise de variância não revele significância para a interação genótipo x época para fibra do grupo 1, poucos genótipos sofrem interferência da época de colheita como: 42, 8, 93, 38 e 40 que estão apenas em entre os dez primeiros em uma das três épocas colhidas em 2015. Os teores de fibra são superiores aos encontrados por Xavier et al. (2014) na seleção de progênie, indicando que podem ser utilizados na obtenção de cultivares para a produção de bioenergia.

Para a mesma variável (fibra), no grupo 2 que contém 82 indivíduos mais responsivos em relação ao teor de sacarose (Pol), observou-se maior valor predito na análise conjunta e na primeira época de colheita para o genótipo 32 (IACSP972084), o qual também esteve dentre os dez primeiros nas épocas 2 e 3, além da época 3 colhida em 2016. Os seguintes genótipos estiveram presentes entre os dez apenas na época 1 (Tabela 7): 16 (IACSP993009), 9 (IACSP042516), 61 (IACSP022067) e 100 (SP832847). Os genótipos 57 (IACSP015519), 34 (IAC914168) e 80 (IACSP953018) apresentaram valores preditos dentre os dez apenas para a segunda época. Já os genótipos 68 (IACSP042510) e 18 (IACSP012421), na época 3. As mudanças no ordenamento dos genótipos destacam a existência da interação das épocas de colheita no valor da variável.

Tabela 7: Ranking dos dez melhores genótipos do grupo 2 de cana-de-açúcar e predição dos valores genotípicos para o atributo fibra.

Ranking	Análise conjunta		Época 1		Época 2		Época 3		Época 3*	
	Gen	g	Gen	g	Gen	g	Gen	g	Gen	g
1	32	1.951*	32	2,1084*	57	1,8376*	21	2,6055*	96	1,7736*
2	21	1.8008*	21	1,511*	50	1,353*	32	1,6993*	60	1,7023*
3	50	1.6287*	60	1,4641*	32	1,3444*	74	1,6734*	21	1,6922*
4	60	1.4458*	50	1,3784*	74	0,83*	12	1,2559*	50	1,6192*
5	74	1.3329*	16	1,2194*	14	0,7374*	50	1,2343*	32	1,607*
6	14	0.963*	9	0,9767*	60	0,7181*	60	1,1694*	74	1,5097*
7	33	0.9381*	33	0,8747*	34	0,6176 ^{ns}	33	0,9705*	14	0,9423*
8	12	0.8971*	12	0,8547*	21	0,6106 ^{ns}	14	0,836*	33	0,8307*
9	100	0.7193*	61	0,7891*	33	0,5234 ^{ns}	68	0,7845*	65	0,8064*
10	57	0.6695*	100	0,7075*	80	0,4946 ^{ns}	18	0,7836*	100	0,7719*
Média geral		11,6697	11,4225		11,1929		11,6557		12,41292	

*: significativo ao nível de 5% pelo teste t ;^{ns}: não significativo; g= BLUP. **9**= IACSP042516; **12**=IACSP042508; **14**=IACCTC059630; **16**=IACSP993009; **18**=IACSP012421; **21**=IACSP991020; **32**=ACSP972084; **33**=CTC15; **34**=IAC914168; **50**=IACCTC061050; **57**=IACSP015519; **60**=IACSP018158; **61**=IACSP022067; **68**=IACSP042510; **74**=IACSP046032; **80**=IACSP953018; **100**=SP832847. **Época 3 *** (colhida em 2016).

O crescimento vegetativo é favorecido por quantidade de água e temperatura adequadas. As necessidades estão em torno de 1200 a 1300 mm e temperatura variando de 22°C a 30°C. Com o fim do estágio de crescimento vegetativo inicia-se a maturação no mês de abril a maio para a região Sudeste. Esta fase é estimulada com redução de chuvas e temperatura até a metade do ano. No ano agrícola de 2014/2015 e 2015/2016 as baixas temperaturas contribuíram para o acúmulo de sacarose (Figura 2).

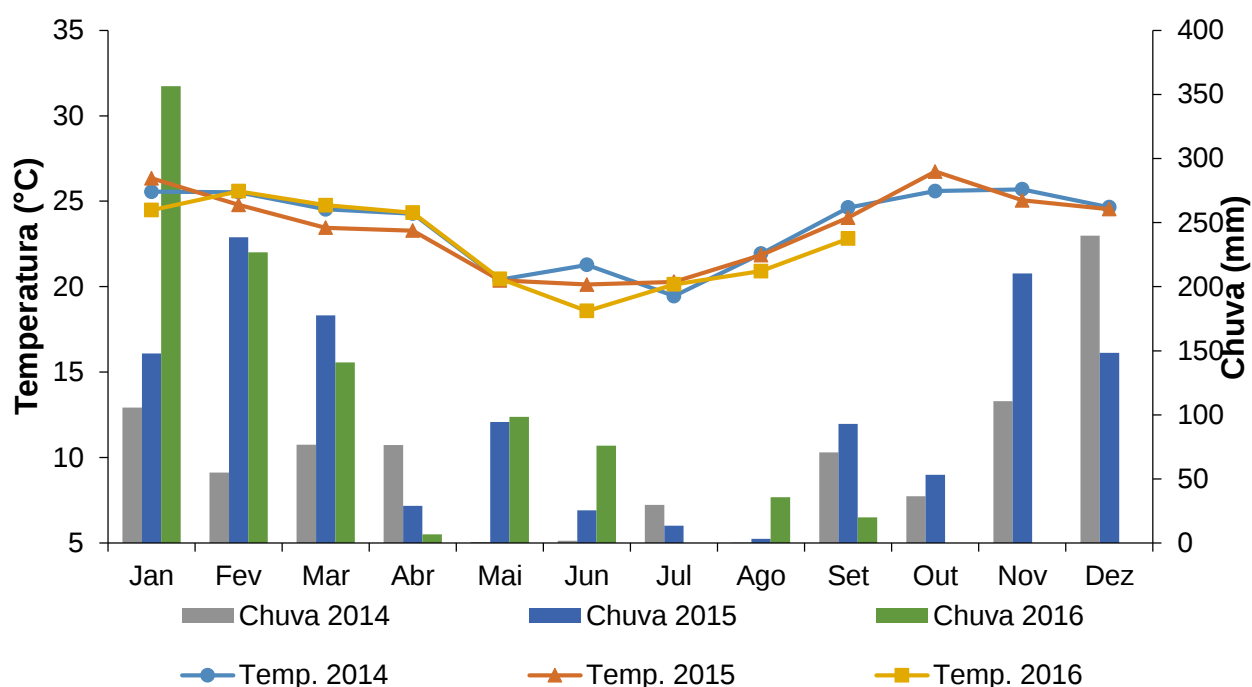


Figura 2 -Valores médios de temperatura máxima e mínima (°C) e pluviosidade mensais de janeiro de 2014 a outubro de 2016 fornecidos pelo Centro de Cana/IAC.

Para o atributo Pol no grupo 1, foi observado que os genótipos 1 (US571415), 35 (IACBIO232), 36 (IACBIO241), 37 (IACBIO257), 38 (IACBIO257) e 93 (IJ76293) estiveram entre os dez melhores incluindo a época 3 colhida no ano de 2016 (Tabela 8). Os genótipos 8 (NG5712), 44 (IACBIO279) e 99 (SES205A) estiveram entre os dez primeiros na primeira e terceira época de colheita, no início e fim da safra de 2015. O 41 (IACBIO273) esteve presente entre os dez na análise conjunta, primeira e segunda época de colheita. Enquanto que o 39 (IACBIO270), 11 (IN8482) e 95 (IN8488) na

segunda época e 10 (Krakatau) na análise conjunta e terceira época no ano de 2015. Os teores de sacarose (Pol cana) apresentados pelos genótipos do grupo 1 considerados abaixo do padrão para a produção de açúcar é justificado pelo fato de esses genótipos terem sido desenvolvidos para a produção de níveis mais elevados de fibra.

Tabela 8: Ranking dos dez melhores genótipos do grupo 1 de cana-de-açúcar e predição dos valores genotípicos para o atributo Pol (cana).

Ranking	Análise conjunta		Época 1		Época2		Época 3		Época 3*	
	Gen	g	Gen	g	Gen	g	Gen	g	Gen	g
1	1	3,7918*	37	3,2056*	1	3,646*	10	3,1351*	35	3,428*
2	38	2,8957*	38	2,7289*	38	2,9405*	1	3,0613 ^{ns}	1	3,077*
3	35	2,5039*	8	2,1187*	37	2,8141*	35	2,3334 ^{ns}	38	1,913 ^{ns}
4	37	2,3796*	1	1,2807 ^{ns}	35	1,7137 ^{ns}	38	2,2528 ^{ns}	37	1,879 ^{ns}
5	36	1,2596 ^{ns}	36	1,2561 ^{ns}	36	0,8167 ^{ns}	8	1,2837 ^{ns}	36	1,497 ^{ns}
6	8	1,1414 ^{ns}	35	1,0346 ^{ns}	39	0,3787 ^{ns}	36	0,8462 ^{ns}	93	0,746 ^{ns}
7	93	0,3054 ^{ns}	93	0,1571 ^{ns}	11	0,3105 ^{ns}	93	0,8297 ^{ns}	44	0,271 ^{ns}
8	10	-0,4833 ^{ns}	44	-0,0101 ^{ns}	95	-0,2056 ^{ns}	99	0,412 ^{ns}	39	0,259 ^{ns}
9	44	-0,8066 ^{ns}	99	-0,3775 ^{ns}	41	-0,3298 ^{ns}	37	0,4076 ^{ns}	95	0,020 ^{ns}
10	41	-0,8484 ^{ns}	41	-0,5985 ^{ns}	93	-0,5688 ^{ns}	44	-0,6488 ^{ns}	8	0,000 ^{ns}
Media Geral		7,4473	7,0881		7,5164		7,0275		7,9216	

*: significativo ao nível de 5% pelo teste t; **ns**: não significativo; **g**= BLUP. **1**=US571415; **8**=NG5712; **10**=KRAKATAU; **11**=IN8482; **35**=IACBIO232; **36**=IACBIO241; **37**=IACBIO257; **38**=IACBIO266; **39**=IACBIO270; **41**=IACBIO273; **44**=IACBIO279; **93**=IJ76293; **95**=IN8488; **99**=SES205A. *Época 3= (colhida em 2016); **Gen**= Genótipos

Embora apresentem níveis de sacarose dentre os dez melhores no início, meio e fim de safra, esses genótipos não possuem condições tecnológicas adequadas para serem colhidos, porque exibem teor de sacarose abaixo do nível considerado para o período útil de industrialização (PUI) que é acima de 12,257% de Pol (cana) (LAVANHOLI, 2010).

No grupo 2 (responsivo para teor de sacarose), somente o genótipo 71(IACSP043150) apresentou valores genotípicos entre os dez genótipos para as três épocas de colheita e análise conjunta. Isto mostra que ele possui um longo período útil industrialização, podendo ser colhido em qualquer uma das três épocas de colheita. Ademais, seus valores genéticos estimados (BLUPS) foram positivos e significativos nas épocas 1 e 3, ou seja, diferiram de zero. Os genótipos 46 (IACCTC056518), 91 (IACSP993369) e 60 (IACSP018158) apresentaram valores preditos para Pol entre os dez na primeira e segunda época, isto é, no início e meio de safra (Tabela 9); O que indica que devem ser colhidos nesses períodos de safra por estarem em plena maturação.

Tabela 9: Ranking dos dez melhores genótipos do grupo 2 de cana-de-açúcar e predição dos valores genotípicos para o atributo Pol (cana).

	Análise conjunta		Época 1		Época 2		Época 3		Época 3*	
Ranking	Gen	g	Gen	g	Gen	g	Gen	g	Gen	g
1	71	1,4731*	80	1,9469*	71	0,7845 ^{ns}	22	1,6824*	51	1,5258*
2	51	1,46*	18	1,5236*	22	0,7227 ^{ns}	79	1,6138*	87	1,4227*
3	18	1,284*7	46	1,5223*	91	0,708 ^{ns}	89	1,401*	18	1,3708*
4	91	1,2626*	91	1,4761*	19	0,7018 ^{ns}	69	1,3913*	55	1,2757*
5	87	1,2257*	71	1,2963*	46	0,6825 ^{ns}	71	1,3864*	65	1,2129*
6	22	1,1183*	51	1,2763*	62	0,5839 ^{ns}	63	1,1627 ^{ns}	28	1,0822 ^{ns}
7	55	1,0104*	45	1,2275*	25	0,5573 ^{ns}	78	1,1538 ^{ns}	79	1,0063 ^{ns}
8	78	0,974*	60	1,2183*	60	0,5178 ^{ns}	70	1,1018 ^{ns}	86	0,9853 ^{ns}
9	28	0,9347*	82	1,165*	87	0,4974 ^{ns}	51	1,094 ^{ns}	63	0,9663 ^{ns}
10	82	0,926*	55	1,1646*	64	0,4945 ^{ns}	83	1,0173 ^{ns}	45	0,957 ^{ns}
Média geral		15,5227		13,8616		15,2681		16,3745		16,6599

*: significativo ao nível de 5% pelo teste t ;^{ns}: não significativo; **g**= BLUP; **Época 3 *** (colhida em 2016). **18**= IACSP012421; **19**=IACSP042521; **22**= IACSP045065; **25**= IACSP018034; **45**= IACCTC053616; **46**= IACCTC056518; **51**= IACCTC069708; **55**= IACSP012430; **60**= IACSP018158; **62**= IACSP022125; **63**=IACSP023025; **64**= IACSP023168; **69**= IACSP043123; **70**= IACSP043148; **71**= IACSP043150; **78**= IACSP046152; **79**= IACSP933046; **80**=IACSP953018; **82**= IACSP962008; **83**= IACSP963056; **87**= IACSP974039; **89**= IACSP982053; **91**=IACSP993369; **Gen**= Genótipos

O 22 (IACSP045065) foi o único que apresentou valor genotípico na segunda e terceira época para Pol, indicando que possui perfil de maturação para o meio e fim de safra, os quais são os períodos de maior volume de produção (LANDELL, 2007). Os demais genótipos: 80 (IACSP953018), 18 (IACSP012421), 45 (IACCTC053616), 82 (IACSP962008) e 55 (IACSP012430) apresentaram valores genotípicos entre os dez apenas na primeira época. O que indica que possuem maior acúmulo de sacarose na primeira época e devem ser colhidos no início de safra, fator contribui para um período cujo volume de produção é menor. Estes valores foram inferiores aos dez maiores (14%) encontrados por Oliveira (2007) na seleção de famílias de cana com maturação precoce para o estado do Paraná.

O 19 (IACSP042521), 62 (IACSP022125), 25 (IACSP018034), 87 (IACSP974039) e 64 (IACSP023168) possuem maior acúmulo de sacarose dentre os dez melhores, somente na segunda época, deixando claro que recomenda-se que devem ser colhidos no meio de safra. E por fim, os genótipos 79 (IACSP933046), 89 (IACSP982053), 69 (IACSP043123), 63 (IACSP023025), 78 (IACSP046152), 70 (IACSP043148) e 83 (IACSP963056) possuem perfil de maturação para fim de safra, uma vez que apresentaram maiores valores genotípicos apenas na terceira época, portanto devem ser colhidos no mês de setembro (Tabela 9).

A mudança no ordenamento pode ser visualizada, por exemplo, para o genótipo 51 (IACCTC069708), presente na análise conjunta, primeira e terceira época, mas abaixo dos dez primeiros na segunda época. O mesmo ocorreu com 46 (IACCTC056518), 60 (IACSP018158) e 91 (IACSP993369), presentes entre os dez na análise conjunta, primeira e segunda época, sofreram influência em relação a época de colheita apresentando valores genotípicos na terceira abaixo dos dez primeiros genótipos. No entanto, fica demonstrado que a época de colheita provoca mudança no ranking.

No ano agrícola de 2015/2016 foi avaliada a terceira época, colhida em setembro para os dois grupos discriminados. No grupo 1, para a variável fibra, oito dos dez genótipos melhores do ranking de 2014/2015 estiveram presentes apresentando os maiores valores genotípicos preditos no fim de safra de 2016. Os materiais foram: 94 (IN8458); 43 (IACBIO277), 40 (IACBIO271), 95 (IN8488), 39 (IACBIO270), 36 (IACBIO241), 44 (IACBIO279) e 11 (IN8452) (Tabela 6). Quanto a

Pol, foi observada também a presença de oito genótipos da época 3 colhida em 2015, 35 (IACBIO232), 1 (US571415), 38 (IACBIO266), 37 (IACBIO257), 36 (IACBIO241), 93 (IJ76293), 44 (IACBIO279) e 8 (NG5712). Estes materiais estão ordenados pelos seus respectivos valores genotípicos e se confirmam na segunda colheita de cana soca como materiais tardios apresentando maior teor de sacarose na terceira colheita (Tabela 8).

No grupo 2, 60 (IACSP018158), 21 (IACSP991020), 50 (IACCTC061050), 32 (IACSP972084), 74 (IACSP046032), 14 (IACCTC059630) e 33 (CTC15) apresentaram maior teor de fibra para a época 3 comparada ao ano de 2015 (Tabela 7). Enquanto que 65 (IACSP042503) e 96 (NG57213) em 2015 não estiverem entre os dez melhores para nenhuma época de colheita. O genótipo 100 (SP832847) em 2015 teve percentual de fibra dentre os dez somente na primeira época de colheita. Para Pol, 51 (IACCTC069708), 79 (IACSP933046) e 63 (IACSP023025) também estiveram dentre os dez na época 3 de 2016 (Tabela 9). Os demais genótipos: 45 (IACCTC053616) e 87 (IACSP974039) estiveram entre os dez nas épocas 1 e 2, respectivamente, 18 (IACSP012421) e 45 (IACCTC053616) na época 1. O 55 (IACSP012430), 65 (IACSP042503), 28 (IACCTC055580) e 86 (IACSP973384) não estiveram entre os dez com maior Pol em nenhuma época colhida no ano de 2015.

Ao longo da história do melhoramento de cana sempre foram selecionados genótipos ricos em teor de sacarose e contra fibra (SANTCHURN et al., 2012). Além disso, há uma correlação negativa entre teor de açúcar e fibra e isto fez com que a seleção para uma característica proporcionasse ganho negativo para outra (JACKSON, 2005). Com relação à cana de dupla aptidão, aquelas que se situam mais ao centro dos dois CPs possuem teor de açúcar e fibra mais próximos. As dez melhores do grupo 1 para a variável Pol apresentam níveis abaixo do exigido para a indústria. Porém, Rao e Weerathaworn (2009) identificaram sete genótipos considerados de dupla aptidão, com teores de fibra variando de 14,88% a 15,93% e para Pol de 11,59% a 14,07%.

A importância do uso de modelos mistos em cana-de-açúcar é visualizada na seleção de famílias, em que são preditos os valores genéticos dos cruzamentos, identificando assim genótipos potenciais para novos cruzamentos. Assim é possível assim descartar aqueles genótipos com baixos valores genéticos para Pol em uma

determinada época de colheita. Os genótipos que apresentaram BLUP negativo não foram significativos, ou seja, podem ser iguais a zero, indicando que os indivíduos podem não possuir valor genotípico abaixo da média geral e assim não serem eliminados. A alta correlação verificada por Barbosa et al. (2005) entre BLUP individual e BLUP individual simulado (BLUPIS) indica o número de indivíduos a ser selecionado por família, o número total de clones a ser avançado e o número de famílias a contribuir com indivíduos selecionados. Portanto, a aplicação de valores genéticos preditos pode contribuir para eficiência da seleção clonal.

Outro fator importante a ser considerado é o conhecimento dos ambientes de produção que consiste nas condições físicas, hídricas, químicas e mineralógicas dos solos, associadas ao preparo e manejo adequado do solo, precipitação pluviométrica, temperatura, radiação solar e evaporação (PRADO, 2008). O conhecimento dos ambientes de produção serve para alocar as cultivares corretamente em função do seu potencial genético, colhendo-se a cana-de-açúcar na época mais indicada (CAVALCANTE; PRADO, 2010).

4 CONCLUSÃO

A herdabilidade ampla da média de genótipos apresentou alta magnitude (50% a 80%), o que indica possibilidade de seleção com alta acurácia.

Os genótipos do grupo 1 possuem perfil para a obtenção de cultivares de “cana energia” mediante os altos teores de fibra.

Houve genótipos com curto e longo período útil de industrialização, podendo sere colhidos em épocas específicas ou em qualquer período da safra. Isto confirma a prática usual de especificidades para a colheita.

REFERÊNCIAS

BARBOSA, M. H. P.; DE RESENDE, M. D. V.; BRESSIANI, J. A.; DA SILVEIRA, L. C. I.; PETERNELLI, L. A. Selection of sugarcane families and parents by Reml/Blup. **CROP BREEDING AND APPLIED TECHNOLOGY**, v. 5, n. 4, p. 443, 2005.

BASTOS, I. T.; BARBOSA, M. H. P.; RESENDE, M. D. V.; PETERNELLI, L. A.; SILVEIRA, L. C. I.; DONDA, L. R.; FORTUNATO, A. A.; COSTA, P. M. A.; FIGUEIREDO, I. C. R. Avaliação da interação genótipo x ambiente em cana-de-açúcar via modelos mistos. **Pesquisa Agropecuária Tropical**, v. 37, n. 4, p. 195-203, 2007.

CAVALCANTE, E. P.; PRADO, H. AMBIENTES DE PRODUÇÃO DE CANA-DE-AÇÚCAR DE LATOSSOLOS NA REGIÃO DE ARAXÁ MG. **Nucleus**, v. 7, n. 2, 2010.

CONAB, Companhia Nacional de Abastecimento. **Acompanhamento de safra brasileira: cana-de-açúcar, terceiro levantamento, dezembro/2016** – Companhia Nacional de Abastecimento. – Brasília: Conab 2016.

Consecana. 2006. Conselho dos produtores de cana-de-açúcar, açúcar e álcool do estado de São Paulo. Manual de Instruções, Piracicaba.

CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **MODELOS BIOMÉTRICOS Aplicados ao MELHORAMENTO GENÉTICO**. 3ª. Viçosa: UFV, 2014. 668 p.

ELIBOX, W. Inheritance for fibre in sugar cane (*Saccharum* spp. hybrids). **International sugar journal**, v. 113, n. 1353, p. 666-671, 2011.

FERNANDES, A. C. Cálculos na agroindústria canavieira. **Piracicaba: STAB**, v. 240, 2003.

HAIR, J. F. Multivariate data analysis. 2009.

HENDERSON, C. 1984-Guelph. 1984.

JACKSON, P. A. Breeding for improved sugar content in sugarcane. **Field Crops Research**, v. 92, n. 2, p. 277-290, 2005.

JACKSON, P. A. Breeding for improved sugar content in sugarcane. **Field Crops Research**, v. 92, n. 2, p. 277-290, 2005.

LANDELL, M. D. A. **Variedades de cana-de-açúcar para o Centro-Sul do Brasil**. Instituto Agrônomo, 2005. p.

LANDELL, M. G. A.; SILVA, D. N.; ANJOS, A. I.; MENDONÇA, J. R.; BIDOIA, M. A. P.; CAMPANA, M. P.; XAVIER, M. A.; FIGUEIREDO, P. **Variedades de cana-de-açúcar para o Centro-Sul do Brasil**. Instituto Agrônomo de Campinas. Campinas, p.37. 2007. (201).

LAVANHOLI, M. D. G. D. P. Qualidade da cana-de-açúcar como matéria-prima para produção de açúcar e álcool. In: DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M.; LANDELL, M. G. D. A. **Cana-de-açúcar**. 1ª. Campinas: Instituto Agrônomo de Campinas, 2010. cap. 8, p. 882.

MANCINI, M. C. **Mapeamento genético e identificação de marcadores aflu, microssatélites genômicos e funcionais associados a parâmetros agroindustriais em cana-de-açúcar**. 2010. 113 f. Dissertação de mestrado - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2010.

MCVEAN, G. A Genealogical Interpretation of Principal Components Analysis. **PLoS Genetics**, v. 5, 2009.

MENDONÇA, J. R.; NOCITI, P. R. H.; DEOTTI, R. C. Estudo de diferentes épocas de corte em cana-de-açúcar. **Boletim Técnico COPERSUCAR**, 1985.

MORAES, M. F. D.; BASTOS, G. Q.; ANUNCIAÇÃO FILHO, C. J.; MELO, L. J. O. T.; REIS, O. V. D. Avaliação agroindustrial e parâmetros genéticos de progênies de cana-de-açúcar em fase inicial na Zona Canavieira do Litoral Norte de Pernambuco. **Ciênc. agrotec.,(Impr.)**, v. 34, n. 5, p. 1086-1092, 2010.

MORAIS, O. P.; SILVA, J. C.; CRUZ, C. D.; REGAZZI, A. J.; NEVES, P. D. C. F. Estimativa dos parâmetros genéticos da população de arroz irrigado CNA-IRAT **Pesquisa Agropecuária Brasileira (Brasil)**. **Abr.**, v. 32, n. 4, p. 421-433, 1997.

OLIVEIRA, R. A. **Seleção de famílias de maturação precoce em cana-de-açúcar via REML/BLUP**. 2007. 142 f. Tese de doutorado -Universidade Federal do Paraná, Curitiba, 2007.

OLIVEIRA, R. A.; DAROS, E.; BESPALHOK FILHO, J. C.; ZAMBON, J. L. C.; IDO, O. T.; WEBER, H.; RESENDE, M. D. V.; NETO, H. Z. Seleção de famílias de cana-de-açúcar via modelos mistos. **Scientia Agraria**, v. 9, n. 3, p. 269-274, 2008.

OLIVEIRA, R. A.; DAROS, E.; RESENDE, M. D. V.; BESPALHOK-FILHO, J. C.; ZAMBON, J. L. C.; SOUZA, T. R.; LUCIUS, A. S. F. Procedimento Blupis e seleção massal em cana-de-açúcar. **Bragantia**, v. 70, n. 4, p. 796-800, 2011.

OLIVEIRA, R. A.; RESENDE, M. D. V.; DAROS, E.; FILHO, J. C. B.; ZAMBON, J.; IDO, O.; WEBER, H.; KOEHLER, H. S. Genotypic evaluation and selection of sugarcane clones in three environments in the state of Paraná. **CROP BREEDING AND APPLIED TECHNOLOGY**, v. 5, n. 4, p. 426, 2005.

PRADO, H. **Solos e ambientes de produção: cana-de-açúcar**. Campinas: Instituto Agrônômico, 2008. p.179-204.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL <https://www.R-project.org/>. Vienna, Austria.
RAO, M. S.; WEERATHAWORN, P. Diversification of breeding program to develop multipurpose sugarcane cultivars. **Sugar Tech**, v. 11, n. 1, p. 77-79, 2009.

RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Embrapa Informação Tecnológica, Colombo: Embrapa Florestas, 2002. p.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v. 37, n. 3, p. 182-194, 2007.

ROBINSON, H.; COCKERHAM, C. Estimación y significado de los parámetros genéticos. **Fitotecnia latinoamericana**, v. 2, n. 4/2, p. 23-38, 1965 apud MORAIS, O. P.; SILVA, J. C.; CRUZ, C. D.; REGAZZI, A. J.; NEVES, P. D. C. F. Estimacao dos parâmetros genéticos da população de arroz irrigado CNA-IRAT 4/0/3¹. **Pesquisa Agropecuária Brasileira (Brasil)**. **Abr.**, v. 32, n. 4, p. 421-433, 1997.

SANTCHURN, D.; RAMDOYAL, K.; BADALOO, M. G. H.; LABUSCHAGNE, M. From sugar industry to cane industry: investigations on multivariate data analysis techniques in the identification of different high biomass sugarcane varieties. **Euphytica**, v. 185, n. 3, p. 543-558, 2012.

SAS, Institute. (2013). SAS 9.4 Output Delivery System: User's Guide. SAS institute, Cary, NC.

SILVA, M. A.; JERONIMO, E. M.; LÚCIO, A. D. C. Perfilhamento e produtividade de cana-de-açúcar com diferentes alturas de corte e épocas de colheita. **Pesq. agropec. bras., Brasília**, v. 43, n. 8, p. 979-986, 2008.

SILVEIRA, L. C. I. D.; BRASILEIRO, B. P.; KIST, V.; DAROS, E.; PETERNELLI, L. A. Genetic diversity and coefficient of kinship among potential genitors for obtaining cultivars of energy cane1. **Revista Ciência Agronômica**, v. 46, p. 358-368, 2015.

VIATOR, R. P.; RICHARD, E. P. Sugar and energy cane date of planting effects on cane, sucrose, and fiber yields. **Biomass and Bioenergy**, v. 40, p. 82-85, 2012.

Wilks, S.S. (1932). Certain generalizations in the analysis of variance. *Biometrika*: 471-494.

XAVIER, M. A.; PERECIN, D.; TOLEDO ALVIM, K. R.; ANDRADE LANDELL, M. G.; ARANTES, F. C. Seleção de famílias e progênies de irmãos completos de cana-de-açúcar para atributos tecnológicos e de produção pelo método de REML/BLUP. **Bragantia**, v. 73, n. 3, p. 253-262, 2014.

CAPÍTULO 3 – MAPEAMENTO ASSOCIATIVO DE GENES RELACIONADOS A POL E À FIBRA DE CANA-DE-AÇÚCAR EM TRÊS ÉPOCAS DE COLHEITA

RESUMO - A cana-de-açúcar é uma cultura semiperene e o uso de marcadores moleculares no melhoramento dessas espécies de ciclo longo pode reduzir o tempo e custo para a obtenção de novas cultivares. A aplicação do mapeamento associativo permite identificar marcadores associados a caracteres de importância econômica em vários ambientes, anos e colheitas. Assim, o estudo teve como objetivo detectar marcadores associados às variáveis tecnológicas (Pol e fibra) utilizando método Bayes C π em três épocas de colheita: início, meio e fim de safra. Para o estudo foram utilizados 100 genótipos de cana-de-açúcar, incluindo cultivares comerciais e acessos IAC. Foram conduzidos três ensaios e avaliados em cana soca. Após o plantio foi realizada a extração de DNA de cada um dos materiais e posteriormente genotipagem com marcadores AFLP (polimorfismo de comprimento de fragmento amplificado). Posteriormente foi realizada a análise de estrutura da população via abordagem bayesiana pelo software STRUCTURE e também por análise de componentes principais. Foram identificados três subgrupos, enquanto que na análise de componentes principais e de agrupamento (k-means) foi revelada a existência de dois subgrupos. As análises de associação foram realizadas utilizando 640 marcas polimórficas após o controle de qualidade. Para cada época de colheita houve quantidades diferentes de marcas significativamente associadas com Pol e fibra. Em cada uma das épocas avaliadas, as marcas explicaram mais de 90% da variação fenotípica dos dados. Além disto, quando testou-se marcas significativas de uma época para prever o efeito em outra, observou-se R^2 acima de 50%, exceto para fibra na predição da época 3 colhida em 2016 quando utilizou-se marcas da época 3 de 2015. Dentre as marcas associadas a Pol, as marcas AFLP 418, 435 e 700 estiveram associadas às três épocas. Para fibra, apenas a marca AFLP 520 esteve associada às três épocas de colheita. Sendo assim, as marcas podem ser validadas na seleção para a obtenção de novos materiais.

Palavras-chave: estrutura genética, Bayes C π , interação QTL x época, marcador AFLP

CHAPTER 3 - ASSOCIATION MAPPING OF GENES RELATED TO POL AND FIBER OF SUGAR CANE IN THREE HARVEST SEASONS

ABSTRACT –The sugarcane is a semiperennial crop and the use of molecular markers in the breeding of these long-cycle species can reduce the time and cost to obtain new cultivars. The application of association mapping allows the identification of markers associated with characters of economic importance in various environments years and harvests. The aim of this study was to detect markers associated to the technological variables (Pol and fiber) using Bayes C π method in three harvesting periods beginning, middle and end of harvest. For the study, 100 sugarcane genotypes were used including commercial cultivars and IAC accessions. Three trials were conducted and evaluated in cane ratoon. After the planting DNA extraction of each materials was carried out and later genotyping with AFLP markers (amplified fragment length polymorphism). Later the analysis of the population structure was performed using the Bayesian approach by STRUCTURE software and by principal components analysis. Three subgroups were identified, whereas in the analysis of principal components and grouping (k-means) was revealed the existence of two subgroups. Association analyzes were performed using 640 polymorphic tags after quality control. For each harvesting season there were different amounts of markers significantly associated with Pol and fiber. In each of the evaluated harvest season, the markers explained more than 90% of the phenotypic variation of the data. Furthermore, when significant marks of one season were tested to predict the effect in another, R^2 was observed above 50% except for fiber in the prediction of season 3 harvested in 2016, when using marks from season 3 of 2015. Among the markers associated with Pol AFLP markers 418 435 and 700 were associated with three seasons. For fiber only the AFLP 520 marker was associated with the three harvest seasons. Therefore, these markers can be validated in the selection to obtain new materials.

Keywords: Genetic structure, Bayes C π , QTL x harvest season, AFLP marker

1 Introdução

O cultivo da cana-de-açúcar possui destaque nacional e internacional pela produção de açúcar e etanol. Os primeiros programas de melhoramento genético da cana objetivavam cultivares com alto teor de sacarose. No entanto, os programas atuais estão também focando no desenvolvimento de materiais ricos em fibra, os quais são direcionados para a geração de energia na própria usina e/ou para a produção do etanol de segunda geração.

A cana-de-açúcar é uma espécie propagada vegetativamente e permanece no campo produzindo durante 5 a 10 anos, por sucessivos ciclos de colheita. O primeiro ano de corte é chamado de planta, e os demais de cana soca. A safra anual é dividida em três épocas: safra de outono, safra de inverno e safra de primavera. Na região Centro-Sul do Brasil a safra se inicia no mês de abril e se estende até novembro. Por ser uma espécie de ciclo longo, a cana é exposta a diferentes condições climáticas durante os estágios fenológicos (LANDELL, 2007). Sendo assim, as cultivares possuem diferentes períodos de ponto de maturação ao longo da safra, que consiste no estágio em que a espécie apresenta maiores teores de sacarose (Pol). Segundo Lavanholi (2010), a cana está em condições adequadas de ser industrializada quando apresenta teor de sacarose acima de 12,257%. Cultivares que atingem este valor no início de safra (mês de abril) são classificadas como precoce, em julho, medianas e a partir de setembro, tardias. Este período no qual as cultivares apresentam condições tecnológicas mínimas exigidas é denominado de período útil de industrialização.

A Pol que significa porcentagem de sacarose aparente, é uma das variáveis tecnológicas e assim como fibra é mensurada para a avaliação da qualidade da matéria prima. Essas duas características são as principais para a obtenção de materiais destinados a produção de açúcar e álcool ou mais recentemente na seleção de cana “energia” para a co-geração de energia na própria indústria ou fabricação de etanol celulósico.

Estudos genômicos são ferramentas importantes que auxiliam um programa de melhoramento na seleção mais eficiente de materiais para cana convencional ou para cana com perfil bioenergético. Sendo assim, o uso de marcadores moleculares juntamente com métodos de maior poder estatístico pode contribuir para se chegar ao

melhor conhecimento acerca dos genes ou loci relacionados a caracteres tecnológicos.

Nesse sentido, técnicas de associação como mapeamento associativo são úteis no estudo em espécies de importância agronômica como milho (YU; BUCKLER, 2006), soja (SHI et al., 2010; SIGRIST, 2012) e cana-de-açúcar (GONÇALVES, 2012). Estudos com mapeamento associativo são úteis por identificarem marcadores associados a características de importância agronômica englobando informações de genótipos, ambientes a interação genótipos por ambiente ou condições de cultivo (DEBIBAKAS et al., 2014; SANTOS et al., 2016; XUE et al., 2013).

Na cultura da cana vários trabalhos demonstram a importância do mapeamento associativo relacionados a diferentes características como resistência a doenças (GOUY et al., 2015; WEI et al., 2006) e relacionados aos teores de sacarose (AITKEN et al., 2008; JACKSON, 2005 ; RACEDO et al., 2016) e de fibra (LOPES, 2011).

O mapeamento associativo, conhecido também como mapeamento por desequilíbrio de ligação, surgiu como alternativa ao mapeamento de QTLs por possuir maior flexibilidade na escolha da população de mapeamento que pode ser formada por populações biparentais, coleções de bancos de germoplasma, linhagens em desenvolvimento e populações naturais; e por considerar todo o histórico de recombinações oriundo de sucessivas gerações de cruzamento proporcionando aumento do poder de associações (GUPTA; RUSTGI; KULWAL, 2005).

Um fator importante para adquirir sucesso na condução do mapeamento associativo é a existência do desequilíbrio de ligação (DL). A quantidade de marcadores utilizados varia em função da presença do alto nível de desequilíbrio de ligação, e quanto maior este nível, maior a chance de mais marcadores estarem associados aos loci de interesse. Fatores como mistura de populações, mutação, seleção efeito fundador e deriva genética podem influenciar o desequilíbrio de ligação (FLINT-GARCIA; THORNSBERRY; IV, 2003).

Nos estudos de associações há a ocorrência de falsas associações provenientes da estruturação da população, isto é, da existência de grupos dentro da população em estudo. Em virtude disto, a identificação da estrutura pode ser feita por meio de métodos estatísticos, tais como: abordagem bayesiana (PRITCHARD; STEPHENS; DONNELLY, 2000), análises de componentes principais (PRICE et al.,

2006). Porém, os métodos que excluem o efeito de confundimento causado nessas populações estratificadas podem vir a remover também a associação verdadeira com o fenótipo (MA et al., 2012).

Para detectar as associações entre marcadores e fenótipos é necessário o uso de métodos estatísticos como modelos mistos que incorporam fatores fixos e aleatórios em seu modelo. Sua finalidade está em descrever as relações entre o fenótipo e algumas covariáveis nos dados que foram medidas ou observadas junto com a variável resposta (fenótipo) (LOPES, 2011). Por meio desses modelos é possível a inclusão da informação de parentesco e de estrutura da população (ZHOU; STEPHENS, 2012). Sua aplicação têm apresentado sucesso na análise de dados fenotípicos provenientes de experimentos conduzidos em vários locais, colheitas e anos agrícolas e que apresentam desbalanceamento (VAN EEWIJK et al., 2007).

A abordagem bayesiana diferente dos métodos frequentistas utiliza uma distribuição a priori baseada em experimentos ou conhecimentos anteriores para cada parâmetro de interesse. Por meio de uma distribuição a posteriori são obtidas as inferências dos parâmetros populacionais. Existem vários métodos como: Bayes A e Bayes B propostos por Meuwissen et al. (2001). Estes métodos proporcionam maiores acurácias da predição do efeito dos marcadores. Habier et al. (2007) exemplificam a superioridade da acurácia do Bayes B em relação ao RR-BLUP na predição de efeitos de marcadores. Outros métodos como o Bayes C, Bayes C π e Bayes D π foram propostos por Habier et al. (2011), destacando-se o Bayes C π por propiciar informações sobre a arquitetura genética do caráter quantitativo e identificar as posições do QTL pela modelagem da frequência do single nucleotide polymorphism (SNP) não nulos, além de considerarem simultaneamente todas as marcas.

O objetivo do trabalho foi avaliar a estrutura genética da população de cana-de-açúcar constituída de genótipos ricos em sacarose e em fibra, e além de identificar marcas associadas a Pol e fibra em três épocas de colheita utilizando o método Bayes C π .

2. Material e métodos

2.1. Material vegetal

O painel de associação utilizado no presente estudo foi constituído por 100 genótipos de cana-de-açúcar contrastantes para Pol e fibra incluindo clones e cultivares selecionados do banco de germoplasma do Centro de Cana do Instituto Agrônomo de Campinas. Desses 100, as cultivares são: IACSP95-3028, CTC 15, IACSP93-3046, IACSP96-3076, IACSP97-4039, RB85-5156, RB93-5744 e SP832847.

Os genótipos avaliados pertencem a diferentes programas de melhoramento do Brasil e de outros países. Do Brasil, 89 genótipos são pertencentes aos seguintes programas de melhoramento: IAC (Instituto Agrônomo de Campinas), CTC (Centro de Tecnologia Canavieira), SP (Copersucar – Coopertativa de Produtores de Cana-de-Açúcar, Açúcar e Álcool do Estado de São Paulo) e RB (Ridesa – Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro). Os outros genótipos com nomenclatura “NG” são oriundos da Nova Guiné, região de centro de origem da *S. officinarum*. Os genótipos IS, IN, IJ e Krakatau são oriundos de ilhas na Indonésia e o US51415 é oriundo dos Estados Unidos. Ao painel também foi adicionado o SES205A, genótipo selvagem pertencente a *S. spontaneum*.

Tabela 1. Painel composto por 100 genótipos de cana-de-açúcar selecionados do banco de germoplasma de Centro de Cana/IAC para mapeamento associativo de Pol e fibra

Nº	Genótipo	Genitor feminino	Genitor masculino	Origem	Nº	Genótipo	Genitor feminino	Genitor masculino	Origem
1	US571415	Desconhecido	Desconhecido	EUA	51	IACCTC069708	IACSP962042	CT014455	Brasil
2	CANA ALHO	Desconhecido	Desconhecido		52	IACCTC069713	IACSP955000	CT961307	Brasil
3	IACSP012410	R570	IAC873396	Brasil	53	IACCTC069741	IACSP955000	RB72454	Brasil
4	IACSP046007	PO8862	CTC9	Brasil	54	IACSP012417	RB835486	SP775181	Brasil
5	IACSP046059	IACSP953018	SP963092	Brasil	55	IACSP012430	IAC912195	SELF	Brasil
6	IACSP972053	SP842066	IAC863154	Brasil	56	IACSP015501	SP913059	CTC957	Brasil
7	IACSP953028	IACSP953028	RB855156	Brasil	57	IACSP015519	RB835486	SP775181	Brasil
8	NG5712	<i>S. robustum</i>	Desconhecido	Indonésia	58	IACSP018046	R570	IAC873396	Brasil
9	IACSP042516	IAC912288	SP80185	Brasil	59	IACSP018082	IAC913093	SP87365	Brasil
10	KRAKATAU	<i>S. spontaneum</i>	Desconhecido	Indonésia	60	IACSP018158	RB835486	SP775181	Brasil
11	IN8482	<i>S. spontaneum</i>	Desconhecido	Indonésia	61	IACSP022067	SP973060	SP701143	Brasil
12	IACSP042508	SP913059	IACSP972077	Brasil	62	IACSP022125	SP901161	SP901644	Brasil
13	IACSP955037	SP842066	IAC873396	Brasil	63	IACSP023025	SP924230	SP88813	Brasil
14	IACCTC059630	IACSP972053	RB85453	Brasil	64	IACSP023168	IAC914216	SP803280	Brasil
15	IACSP046053	IACSP956114	IACSP933046	Brasil	65	IACSP042503	PO8862	CTC9	Brasil
16	IACSP993009	CP701133	SP832847	Brasil	66	IACSP042504	IACSP955037	Desconhecido	Brasil
17	IS76155	Desconhecido	Desconhecido	Indonésia	67	IACSP042509	SP913059	IACSP973391	Brasil
18	IACSP012421	RB835486	SP775181	Brasil	68	IACSP042510	IAC913093	CTC9638	Brasil
19	IACSP042521	IAC912288	SP963092	Brasil	69	IACSP043123	IACSP953028	SP832847	Brasil
20	IACSP046058	IACSP953018	SP963092	Brasil	70	IACSP043148	IACSP953018	IACSP933046	Brasil
21	IACSP991020	SP851682	SP80185	Brasil	71	IACSP043150	IACSP953018	IACSP933046	Brasil
22	IACSP045065	IACSP956114	IACSP953050	Brasil	72	IACSP043259	IACSP966026	RB855453	Brasil
23	IACCTC069767	IACSP963069	IACSP952078	Brasil	73	IACSP045081	IACSP966026	RB855453	Brasil
24	IACSP985046	SP813251	SP823697	Brasil	74	IACSP046032	IACSP953028	SP775181	Brasil

25	IACSP018034	IAC913093	SP87432	Brasil	75	IACSP046035	SP913059	IACSP966026	Brasil
26	IACSP993357	SP921584	Desconhecido	Brasil	76	IACSP046073	IACSP953028	SP924221	Brasil
27	IACCTC059578	WHITETRANS	SP803280	Brasil	77	IACSP046077	IACSP972109	IACSP956114	Brasil
28	IACCTC055580	CTC4	IACSP933046	Brasil	78	IACSP046152	IAC911099	SP931124	Brasil
29	IACSP976628	RB835486	SELF	Brasil	79	IACSP933046	SP791011	Desconhecido	Brasil
30	IACSP976680	SP842066	SP847017	Brasil	80	IACSP953018	SP842189	SP801842	Brasil
31	IACCTC059534	IACSP953028	CT961414	Brasil	81	IACSP956114	IAC873187	CTC9019	Brasil
32	IACSP972084	RB855035	SP80185	Brasil	82	IACSP962008	SP80144	Desconhecido	Brasil
33	CTC15	SP842025	Desconhecido	Brasil	83	IACSP963056	SP826108	Desconhecido	Brasil
34	IAC914168	SP813137	Desconhecido	Brasil	84	IACSP963069	SP803280	Desconhecido	Brasil
35	IACBIO232	RB855465	GLAGAH	Brasil	85	IACSP963076	SP847017	Desconhecido	Brasil
36	IACBIO241	RB855465	GLAGAH	Brasil	86	IACSP973384	RB855113	RB855453	Brasil
37	IACBIO257	NG26011	Desconhecido	Brasil	87	IACSP974039	RB835486	RB855453	Brasil
38	IACBIO266	NG26011	Desconhecido	Brasil	88	IACSP974048	SP842066	IAC863154	Brasil
39	IACBIO270	SES069	Desconhecido	Brasil	89	IACSP982053	SP847017	SP801842	Brasil
40	IACBIO271	SES069	Desconhecido	Brasil	90	IACSP983011	IAC913093	SP823530	Brasil
41	IACBIO273	SES069	Desconhecido	Brasil	91	IACSP993369	SP901616	Desconhecido	Brasil
42	IACBIO275	SES069	Desconhecido	Brasil	92	IACSP994011	SP842025	RB855113	Brasil
43	IACBIO277	SES069	Desconhecido	Brasil	93	IJ76293	<i>S. robustum</i>	Desconhecido	Java
44	IACBIO279	SES069	Desconhecido	Brasil	94	IN8458	<i>S. spontaneum</i>	Desconhecido	Indonésia
45	IACCTC053616	IAC915155	CT961414	Brasil	95	IN8488	<i>S. spontaneum</i>	Desconhecido	Indonésia
46	IACCTC056518	SP911049	CT931231	Brasil	96	NG57213	Desconhecido	Desconhecido	Nova Guiné
47	IACCTC059552	IACSP955011	CT961414	Brasil	97	RB855156	RB72454	TUC717	Brasil
48	IACCTC059607	IACSP972053	SP88721	Brasil	98	RB935744	RB835089	RB765418	Brasil
49	IACCTC059634	IACSP972053	RB855453	Brasil	99	SES205A	Desconhecido	Desconhecido	
50	IACCTC061050	CT943165	IAC911099	Brasil	100	SP832847	HJ5741	SP701143	Brasil

2.2 Avaliação fenotípica

Foram conduzidos três experimentos no Centro de Cana do Instituto Agrônomo de Campinas, localizado no Município de Ribeirão Preto-SP, (latitude 21° 12' 42" S e longitude 47° 48' 24"). O clima é do tipo tropical com inverno seco, classificado de acordo como o Sistema Internacional de Classificação de Köppen, como Aw. A pluviometria média anual é de 1.422 milímetros.

Cada experimento correspondeu a uma época de colheita, sendo a primeira época em maio, a segunda em julho e a terceira em setembro de 2015 e 2016. Os três experimentos foram colhidos com 12 meses, após os cortes de uniformização realizados em cana planta, respectivamente, nos meses de maio, julho e setembro de 2014. Os experimentos foram conduzidos em delineamento de blocos casualizados com 4 repetições e 100 genótipos. Cada parcela foi constituída por uma linha de 1,5 m com espaçamento de 1,5 m entre fileiras e 0,5 m entre mudas, contendo 3 mudas pré-brotadas por parcela.

As variáveis tecnológicas, Pol (% cana) e fibra, foram analisadas seguindo as normas técnicas do Conselho de Produtores de Cana-de-Açúcar do Estado de São Paulo (CONSECANA, 2006).

2.3. Extração e quantificação do DNA genômico

Amostras foram coletadas de primórdios foliares de cada um dos 100 materiais escolhidos de acordo com a metodologia proposta por Aljanabi et al. (1999). O DNA extraído foi quantificado tomando como referência padrões de concentrações, conhecidos como DNA do fago lambda, em gel de agarose 0,8 % (p.v⁻¹), corado com brometo de etídio.

2.4. Reações de PCR - Marcadores AFLP (*Amplified Fragment Length Polymorphism*)

Para as reações de PCR com os marcadores AFLP, foram utilizadas 300ng do DNA extraído de cada genótipo para serem digeridos com duas enzimas de restrição (*EcoRI* e *MspI*) em um volume de 20 µL, contendo tampão 10X e água ultrapura

esterilizada. As amostras foram incubadas por três horas a 37°C e a 70°C por cinco minutos em termociclador. Após a digestão procedeu-se a ligação dos adaptadores contendo 20 µL da reação de digestão, 8 µL do tampão (Buffer – 5X), 1µL do adaptador EcoRI a 5µM, 1µL do adaptador MspI (Hpa II) a 50 µM, 1 µL de ATP (10mM) e 0,51 µL da enzima T4 DNA ligase (67U/µL) e completou-se o volume para 40µL com água ultrapura. A reação de ligação foi realizada em termociclador a 37°C por duas horas e a 16°C por 16 horas.

Anteriormente à da etapa de amplificação seletiva, foi realizada pré-amplificação com primers complementares aos adaptadores de *EcoRI* e *MspI*(*HpaII*) que não possuem nucleotídeos seletivos na extremidade 3'. Foram diluídos os produtos de ligação dos adaptadores seis vezes em água ultrapura para o preparo da reação de amplificação. Em seguida adicionou-se 2µL de cada amostra diluída 6X, 1 µL do primer EcoRI+0 (50 µM), 1µL do primer Msp+0 (50µM) 1 µL de dNTPs (2,5mM), 1,2mM MgCL₂ (25 mM), 0,2 µL de Taq DNA polimerase (5U/µL) e completou-se o volume da reação para 15 µL com água ultrapura. As reações foram submetidas a 29 ciclos de 94°C por 30 segundos, 56°C por 1 minuto e 72°C por 1 minuto. Todas as amostras foram dispostas em termociclador, sob os seguintes ciclos repetidos: 29 ciclos de 94°C por 30 segundos, 56°C por 1 minuto, 72°C por 1 minuto.

A última etapa consistiu na amplificação seletiva que utiliza dois primers: um referente à sequência do adaptador da sequência de corte raro que tem acrescentado um ou mais nucleotídeos aleatórios na extremidade 3'; e o outro refere-se à sequência do adaptador da enzima de corte frequente seguida pelos nucleotídeos arbitrários. Em seguida, para a amplificação seletiva foi feita uma diluição 10x das reações de pré-amplificação e adicionado o primer seletivo para três nucleotídeos.

Foram utilizadas 35 combinações de primers EcoRI/MspI (Apêndice 1A). Os primers EcoRI700 e EcoRI800 foram marcados com os fluoróforos IRDye700 ou IRDye800. Realizou-se duas reações, referentes aos primers IRDye700 e IRDye800 foram com o seguinte protocolo: 2,5µL da pré-amplificação diluída 10x, 1,0 µL do tampão 10X, 0,5µL do primer Eco700 (1 µM) ou 0,7 µL do primer Eco800 (1 µM) marcado, 0,25 µL do primer não marcado (10 µM), 1,07 µL dNTPs (2,5 µM), 0,8 µL de MgCL₂ (25 mM), 1,5 µL de Taq DNA polimerase (5U/ µL) e completou-se o volume para 10,5 µL com água ultrapura.

Para as reações das combinações com os primers EcoRI800, os ciclos das reações de amplificação foram os seguintes: 12 ciclos a 94°C por 30 segundos, 65°C por 30 segundos (decrecendo 0,7°C/ciclo) e 72°C por 1 minuto, seguidos por 23 ciclos de 94°C por 30 segundos, 56°C por 30 segundos e 72°C por 1 minuto.

As condições de amplificação utilizadas constaram de 13 ciclos a 94°C por 30 segundos, 65°C (-0,7 °C/ ciclo) por 30 segundos e 72°C por 1 minuto, seguido por 23 ciclos de 94°C por 30 segundos, 56°C por 30 segundos e 72°C por 1 minuto. Os produtos de PCR foram visualizados em gel de poliacrilamida com o uso do Licor 4300 DNA analyzer seguindo as recomendações do fabricante.

3. Análise de dados

3.2. Análise da estrutura genética da população

3.2.1. Análise pelo software STRUCTURE

A análise de estrutura de população para verificar a existência de grupos foi efetuada por meio do software STRUCTURE v.2.3.4 desenvolvido por Pritchard; Stephens e Donnelly (2000), analisam as amostras empregando um algoritmo bayesiano, que permite a identificação de subgrupos dentro da população. Nesta análise foram testadas hipóteses dos valores de variando de 1 a 10 no total de repetições. Para cada hipótese de K foi utilizado um período de burn-in de 100.000 iterações e Cadeia de Markov e Monte Carlo (MCMC, Monte Carlo Markov Chain) de 200.000 iterações para a convergência da MCMC. Foram determinados os seguintes parâmetros: nível de ploidia igual a 1, modelo de não mistura ("no admixture") que consiste em um indivíduo que possui ancestrais comuns em mais de uma população; e frequências correlacionadas entre as populações para o modelo de frequências alélicas, uma vez que os genótipos possuem ancestrais comuns entre si. A determinação do nível mais provável de subgrupos (K) dentro da população é baseada no Delta K (ΔK) que é uma média dos valores absolutos do logaritmo de K ($L(K)$) do número de corridas realizadas dividido pelo desvio padrão de $L(K)$, sendo o maior ΔK

relacionado ao valor mais provável de K (EVANNO; REGNAUT; GOUDET, 2005), o qual foi obtido por meio da análise conduzida no STRUCTURE HARVESTER v.0.6.94 dos resultados transferidos das corridas (EARL; VONHOLDT, 2012).

3.2.2. Análise via componentes principais

Análises conduzidas utilizando conhecimento a priori do pedigree

Dos 100 genótipos, 80 foram classificados como grupo Pol, 14 como fibra e 6 como dúvida (1, 8, 10, 17, 93 e 96). Em seguida, foram identificadas as marcas AFLP com maior diferença na frequência de ausência entre os grupos Pol e fibra para avaliar o poder das mesmas em discriminar estes grupos. Para a identificação das marcas foi considerada a diferença de frequência maior que 0,5, por exemplo, a marca AFLP 1 com frequência no grupo Pol, igual a 0,21 e no grupo fibra com 0,78, portanto a diferença maior que 0,5. Depois foi construída uma matriz de parentesco genômica com base nas marcas identificadas e a partir delas foi realizada uma análise de componentes principais (ACP), definindo a matriz de parentesco como variável original. Utilizando a matriz de parentesco genômico e assumindo $K=2$, foi realizada uma análise de agrupamento (K-means). Esses resultados foram combinados com os da ACP para a discriminação dos grupos formando um “scatter plot”, gráfico usado para observar a relação entre dois conjuntos de dados. Todas as análises foram realizadas usando o programa R (R Core Team, 2015).

Análises conduzidas com os dados fenotípicos

Os mesmos procedimentos foram aplicados à média dos dados fenotípicos das 4 repetições de Pol e fibra para cada uma das três épocas de colheita. Para análise de componentes principais foi utilizada a função `prcomp` do software R que usa o método de decomposição singular. Os valores dos componentes principais (CPs) de cada genótipo foram utilizados para realizar a análise de agrupamento considerando $K=2$, sendo estes grupos formados de acordo com os CPs dos fenótipos para Pol e

fibra. Depois de definido os grupos foi construída uma matriz de parentesco genômico com base nos marcadores que melhor os discriminaram, e então realizada uma análise de componentes principais utilizando a matriz de parentesco como variável original. Foram gerados scatter plots referentes à informação do fenótipo e à combinação dos resultados da ACP e da análise de agrupamento baseado nos marcadores.

4 Análise de associação

4.1 Método Bayes C π

Previamente à análise associação os dados (100 genótipos e 814 marcas) sete genótipos e 46 marcas AFLPs foram excluídos por apresentarem taxa de leitura (“call rate”) abaixo de 90%. Foram excluídos também 128 marcas AFLPs com frequência abaixo de 5% ou acima de 95%, uma vez que podem ser falsos positivos. Com isso, restaram-se 93 genótipos e 640 marcas.

O efeito de cada marca AFLP foi estimado utilizando o método Bayes C π proposto por Habier et al. (2011) por meio do software GS3 (LEGARRA et al., 2011) aplicando o seguinte modelo:

$$y = 1_n\mu + \sum_{i=1}^n g_i b_i \delta_i + e$$

em que y é o vetor de fenótipos (Pol ou fibra), 1_n é um vetor de 1_s ; μ é a média geral; g_i é o vetor contendo os genótipos das amostras para a i -ésima marca, codificado como 1 para presença e 0 para ausência da marca, b_i é o efeito da i -ésima marca, onde $b \sim N(0, \sigma_b^2)$, δ_i é uma variável indicadora (0 ou 1) amostrada de uma distribuição binomial com parâmetros n e π , nos quais n é o número de marcadores e π é fração de marcadores não incluídos no modelo, e e é o vetor dos resíduos, onde $e \sim N(0, \sigma_e^2)$.

Foi assumida para π uma distribuição β com parâmetros $\alpha=10^8$ e $\beta=10^9$ de valor de 0,90, ou seja, 10% das marcas sendo significativas (LEGARRA; RICARD; FILANGI, 2012). Uma distribuição a priori de qui-quadrado escalonada invertida foi assumida para variância genética aditiva pelos marcadores (σ_b^2) e variância residual

(σ_e^2), com $v_a = 4$ graus de liberdade e parâmetro de escala S_a^2 com probabilidade marginal a priori $(1-\pi)$. Este parâmetro foi calculado pela seguinte equação:

$$S_a^2 = \frac{\tilde{\sigma}_a^2(v_a-2)}{v_a}$$

na qual a $\tilde{\sigma}_a^2$ é a variância aditiva do efeito de locus amostrados aleatoriamente. A análise foi realizada por meio do método de amostragem de Gibbs, um tipo de método de Monte Carlo via Cadeias de Markov (MCMC), utilizando 1,1 milhão de iterações, burn-in de 100.000 iterações, intervalo de thin de 1000 iterações e a priori usada foi de distribuição do tipo admixture.

Para a determinação dos efeitos das AFLPs, foi usado o fator de Bayes (BF – Bayes Factor) por meio da seguinte fórmula:

$$BF = \frac{[p/(1-p)]}{[\pi/(1-\pi)]}$$

na qual p e π representam a probabilidade a posteriori e a priori, respectivamente, de um marcador ser incluído na análise.

5 RESULTADOS E DISCUSSÃO

5.1 Análise da estrutura da população

Análise de componentes principais

A importância para a realização deste tipo de análise é detectar marcas que discriminem melhor os genótipos, ou seja, marcas que provavelmente estejam mais associadas a Pol ou à fibra com base nas diferenças de frequências de ausência entre os grupos (Pol e fibra) classificados de acordo com a informação a priori do pedigree.

Para a discriminação dos grupos nas épocas 1, 2 e 3 foram identificadas 69 marcas AFLP com diferença maior que 0,5 para a ausência das marcas entre os genótipos classificados como Pol ou fibra, baseando-se na informação a priori do

pedigree. A partir daí, por meio da análise de k-means dos componentes principais, foram discriminados dois grupos para as épocas colhidas em 2015 (Figura1). Foram classificados 85 indivíduos no grupo 1 (Pol) e 15 no grupo 2 (fibra) pela análise de k-means. Observou-se que houve alta concordância entre a discriminação dos grupos baseados tanto na informação a priori do pedigree como na análise de k-means das 69 marcas.

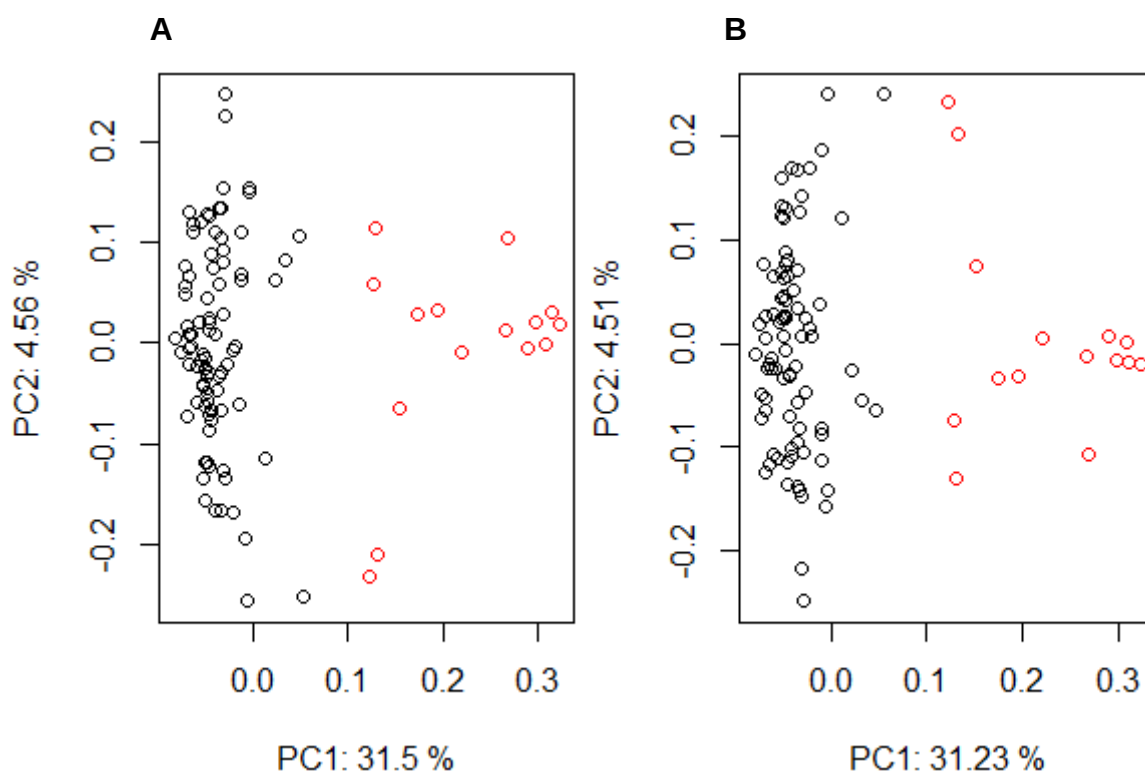


Figura 1. Análise de agrupamento (K-means) dos 100 genótipos de cana-de-açúcar baseada nos marcadores. A (Épocas 1, 2 e 3) e B (Época 3 colhida em 2016). Os genótipos de cor preta constituem o grupo Pol e os de cor vermelha, o grupo fibra.

É importante destacar também que quando utilizou-se a informação dos 814 marcadores não houve coincidência dos grupos definidos com base no pedigree e pela análise de componentes principais e de k-means. As mesmas análises anteriores foram aplicadas à época 3 colhida no ano de 2016 e identificaram 70 marcas com maior diferença para ausência entre os grupos Pol ou fibra (Figura 1B), não havendo mudança no número de grupos identificados. Pela análise de k-means assumindo $k=2$, observou-se que 83 genótipos de cana foram agrupados para Pol (grupo 1) e 15 para

fibra (grupo 2). Nesta época foram eliminados dois genótipos em função de parcelas perdidas, isto é, a análise com 98 ao invés de 100 genótipos.

Utilizando a média das 4 repetições dos fenótipos das variáveis Pol e fibra de cada uma das épocas avaliadas foram discriminados dois grupos via análise de componentes principais. O primeiro componente (CP1) apresentou para a variável Pol, um autovetor de 0,70 e para a variável fibra, o valor foi de -0,70 (Tabela 2). Na época 1, somente o CP1 explicou 89,62% da variação total contida nas variáveis originais, enquanto que o CP2, 10,38%. A análise de agrupamento (K-means) assumindo K=2 gerada de acordo com os CPs dos fenótipos para Pol e fibra classificou 17 genótipos para fibra (grupo 1) e 83 como Pol (grupo 2) (Figura 2A).

Tabela 2. Autovetores obtidos da matriz de correlação para as variáveis Pol e a fibra dos 100 genótipos em todas as épocas de colheita, incluindo a época 3 colhida em 2016.

Variáveis	CP1	CP2
POL	0,7071	0,7071
FIBRA	-0,7071	0,7071

Após a discriminação dos grupos baseados no fenótipo, 47 marcas foram identificadas com base na diferença na frequência de ausência das marcas entre os grupos Pol e fibra para a primeira época de colheita. A partir da matriz de parentesco genômico baseada nas 47 marcas, foram obtidos scatter plots de análise de agrupamento (k-means) baseada na ACP classificando 16 genótipos no grupo 1 (Fibra) e 84 no grupo 2 (Pol) (Figura 3A).

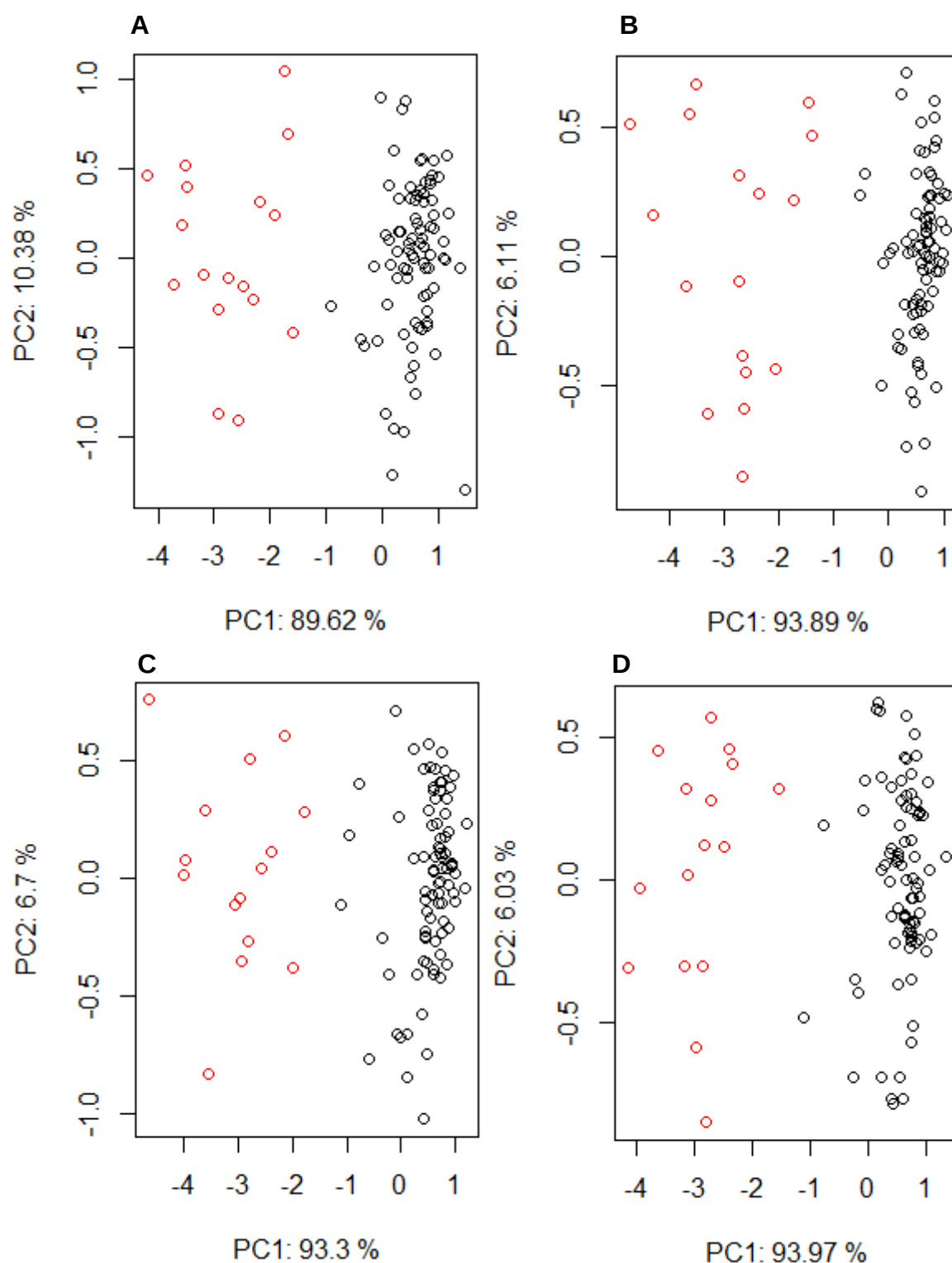


Figura 2. Análise de agrupamento (K-means) dos fenótipos Pol e fibra. A (Época 1, colheita de maio de 2015), B (Época 2, colheita de julho de 2015), C (Época 3, colheita de setembro de 2015), D (Época 3, colheita de setembro de 2016). Genótipos de cor vermelha constituem o grupo 1 (fibra) e de cor preta, o grupo 2 (Pol).

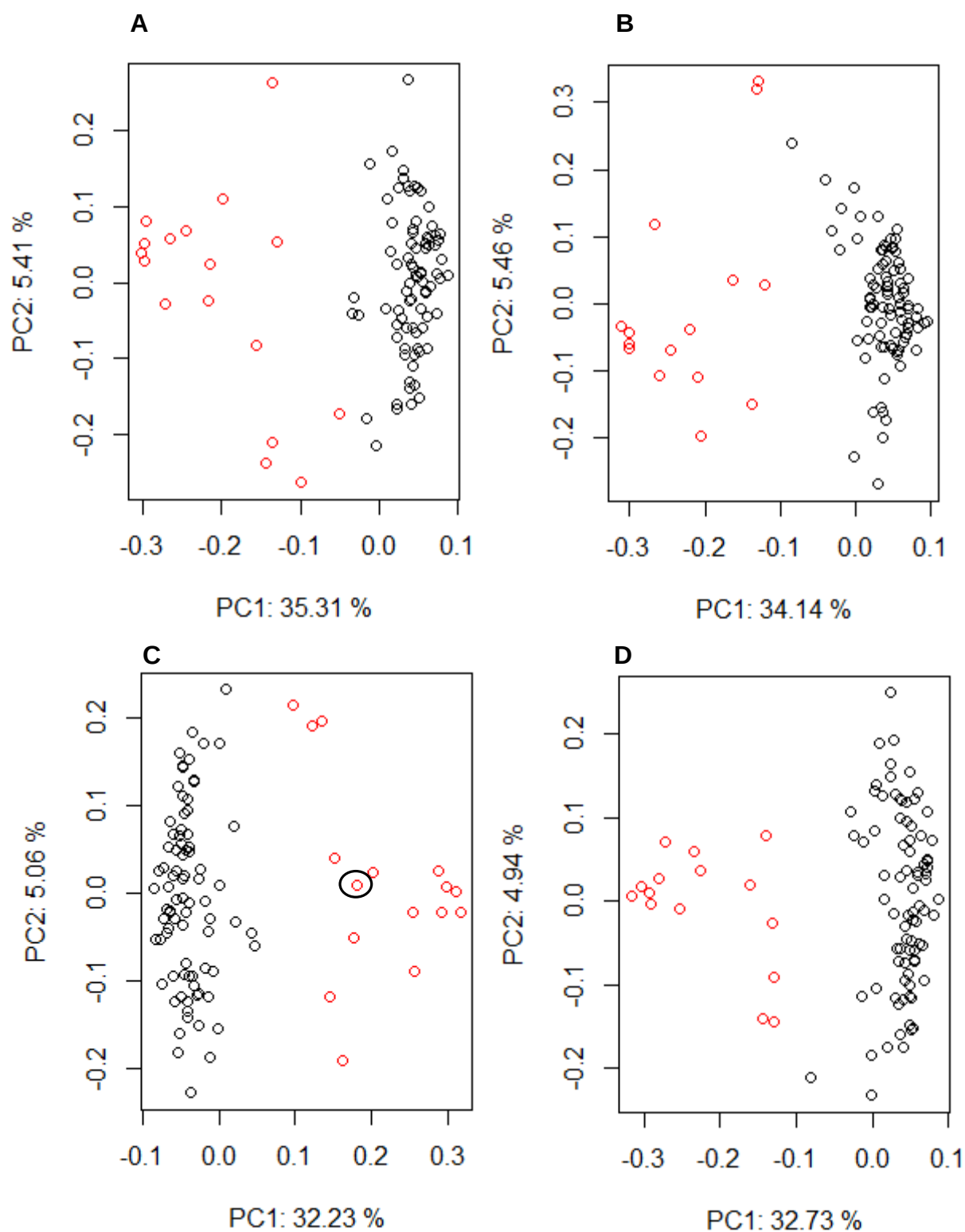


Figura 3. Análise de agrupamento (K-means) baseada nos marcadores com a informação a priori do fenótipo. A (Época 1, 47 marcadores), B (Época 2, 50 marcadores), C (Época 3, 62 marcadores) e D (Época 3 de 2016, 62 marcadores). Os genótipos de cor vermelha constituem o grupo 1 (fibra) e os de cor preta, o grupo 2 (Pol).

Para a segunda época, o CP1 explicou 93% da variação total dos dados e o segundo, 6,11%. A análise de agrupamento (K-means) assumindo K=2 gerada de acordo com os CPs dos fenótipos para Pol e fibra classificou 17 genótipos como fibra (grupo 1) e 83 genótipos como Pol (grupo 2) (Figura 2B). Identificou-se 50 marcas com maior diferença para a ausência entre os grupos Pol e fibra, as quais discriminaram 15 genótipos para o grupo 1 (Fibra) e 85 para o grupo 2 (Pol) no experimento colhido no meio de safra (Figura 3B).

Na terceira época (colhida em setembro de 2015), o CP1 explicou 93,3% da variação total dos dados e o segundo, 6,7%. A análise de agrupamento (K-means) assumindo K=2 gerada de acordo com os CPs dos fenótipos para Pol e fibra classificou 15 genótipos como fibra (grupo 1) e 85 como Pol (grupo 2) (Figura 2C). Identificou-se 62 marcas com maior diferença para a ausência entre os grupos Pol e fibra, as quais discriminaram dois grupos, grupo 1 (fibra) com 16 genótipos e o grupo 2 (Pol) com 84 indivíduos para o experimento colhido no fim de safra (Figura 3C). Para a mesma época colhida em 2016, o CP1 apresentou 93,9% da variação total contida nos dados e o CP2 6,03%. A análise de agrupamento (K-means) assumindo K=2 gerada de acordo com os CPs dos fenótipos classificou 16 genótipos como grupo 1 (fibra) e 82 como grupo 2 (Pol) (Figura 2D). Identificou-se 62 marcas, as quais discriminaram dois grupos, grupo 1 (fibra) com 15 indivíduos e grupo 2 (Pol) com 83 para o experimento colhido no fim de safra (Figura 3D).

A análise de componentes principais e de k-means dos componentes principais baseada nas matrizes de parentesco com base nos marcadores que discriminaram os grupos em cada uma das épocas mostraram que o CP1 explicou mais de 30% da variação existente. Estes valores foram superiores aos obtidos por Brown et al. (2007), Gonçalves (2012) e Lopes (2011) e representam um adequado ajuste do modelo em função da variabilidade detectada pela discriminação dos grupos existentes na população e ao grande número de marcadores utilizado. Com a mesma finalidade do que o algoritmo bayesiano para analisar a estrutura da população, métodos multivariados como a ACP são uma opção para grandes conjuntos de dados sem a informação prévia sobre o modelo genético da população (HABIER et al., 2011; JOMBART; DEVILLARD; BALLOUX, 2010).

Como a cana-de-açúcar apresenta períodos de acúmulo de sacarose distintos, ou seja, parte dos genótipos possuem maior acúmulo no início de safra (maio), no meio de safra (julho) ou no fim de safra (setembro), além daqueles que apresentam longo período de industrialização, isto é, alto teor de sacarose durante todo o período de safra. Isso pode ter ocasionado a diferença na distribuição dos indivíduos nos grupos discriminados em cada época de colheita podendo ser visualizadas nos scatter plots.

Em relação as classificações definidas de acordo com o pedigree, o fenótipo e as análises de k-means da matriz do parentesco genômico que melhor discriminaram os grupos com base no fenótipo e no pedigree, houve coincidência entre as classificações das diferentes análises, exceto para os genótipos oriundos do policruzamento em que não se tem informação do genitor masculino, além de apresentarem valores de fibra e Pol próximos e localizados mais ao centro dos eixos dos componentes principais. Alguns genótipos apresentaram divergência como o genótipo Krakatau destacado na figura 3C (Época 3, colhida em 2015) que é derivada de policruzamento, ou seja, não se conhece o genitor masculino e além disso, apresentou teores de Pol e fibra muito próximos.

5.2 Análise de estrutura pelo programa STRUCTURE

As análises de estrutura da população detectaram 3 grupos mais prováveis, dados pelo valor de ΔK , conforme a figura 1 (EVANNO; REGNAUT; GOUDET, 2005). A classificação dos indivíduos para os grupos foi baseada na proporção acima de 50% em um determinado cluster. A distribuição dos genótipos nos três grupos mostrou que 58 genótipos da população ficaram no agrupamento 2 com mais de 90% da proporção genômica, enquanto que no agrupamento 2 e 3 abrangeram 12 e 30 genótipos, respectivamente (Figura 4). Os genótipos: IACBIO273, IACBIO275, IACBIO241, IACBIO279, IACBIO257, IACBIO266, IACBIO270, IACBIO271 que foram classificados no agrupamento 2, são meios irmãos, além de terem percentual de fibra muito próximos (Tabela 1). Segundo Lopes (2011), a interferência da estrutura se dá quando o fenótipo é mais expresso em determinados indivíduos da população do que em outros, juntamente com a alta frequência do alelo, ocasionando assim uma

associação estatística positiva entre este alelo e o caráter sem que, necessariamente exista real causa genética. A diferença entre as frequências alélicas dos grupos existentes dentro de uma população é o que resulta na estruturação. Sendo assim, como cada grupo possui sua própria composição genética, esta pode não estar relacionada à característica avaliada.

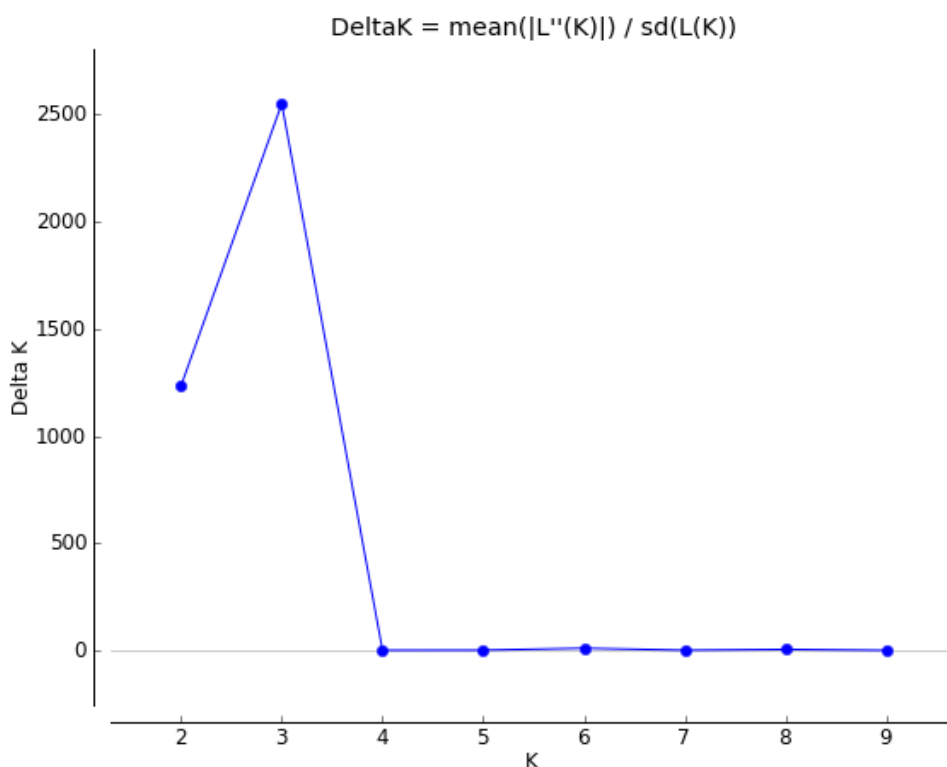


Figura 4. Valor de Delta K= 3 indicando o número de subgrupos mais prováveis existentes na população de mapeamento.

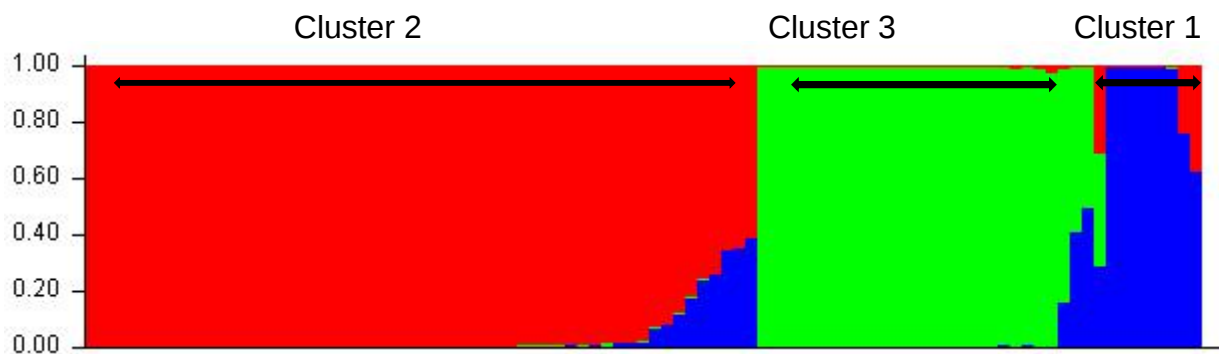


Figura 5. Constituição dos grupos existentes dentro da população de mapeamento dos genótipos. Os genótipos estão representados por linhas verticais divididas em segmentos coloridos proporcionais à fração de seus genomas (valor Q).

As análises para discriminação dos grupos pelas diferentes metodologias apresentadas tiveram o mesmo objetivo de evitar a ocorrência de falsos positivos, os quais podem ser um fator de confundimento no estudo de associação com a característica de interesse.

O resultados obtidos por meio ACP e a de K-means identificaram a existência de dois subgrupos (Pol e fibra), diferentemente da análise realizada no STRUCTURE, que revelou a existência de três subgrupos. Isto é justificado porque na análise de ACP e K-means foi realizada com os marcadores que melhor discriminaram os grupos para Pol e fibra em cada época dentro da população e que tiveram uma alta adesão com grupos definidos com a informação a priori do pedigree (Figura 1). Além disto, quando usou-se a informação das variáveis, Pol e fibra, para cada época de colheita, houve também a identificação de dois grupos e novamente alta adesão com os grupos definidos baseado nos marcadores. Já no STRUCTURE, utilizou-se as 640 marcas e 93 indivíduos após o tratamento prévio das marcas com frequência menor do que 5% e maior do que 95%, além das amostras com frequência menor do que 90% e dos indivíduos com taxa de leitura menor do que 90% (“call rate”) com o objetivo de eliminar potenciais erros nos resultados de associação.

5.2 Análise de associação

O uso de métodos estatísticos é muito discutido por sua finalidade de corrigir a estrutura da população para que não haja a ocorrência de falsos positivos ou associações espúrias (FALUSH; STEPHENS; GOUY et al., 2015; LI et al., 2010 ; PATTERSON; PRICE; REICH, 2006 ; PRICE et al., 2006; PRICE et al., 2010; PRITCHARD et al., 2000; ZHU et al., 2008). Por outro lado, a eficiência na remoção do efeito de confundimento em populações estratificadas pode juntamente excluir efeitos associativos verdadeiros à característica avaliada. Como exemplo, pode-se citar o mapeamento associativo para o tempo de florescimento em população estratificada de *Arabidopsis* sp., em que efeitos verdadeiros a genes candidatos foram eliminados para o florescimento tardio. Além disto, há o fator de que os genótipos serem contrastantes a um nível genômico amplo. Assim, qualquer alelo compartilhado por todos ou pela maioria pode vir a ser removido com o auxílio de métodos estatísticos para reduzir o efeito da estrutura da população (ZHAO et al., 2007).

Ma et al. (2012) compararam três métodos de correção: Associação esperada do eficiente modelo misto (EMMAX – Efficient Mixed-Model Association Expedited), quadrados mínimos generalizados e análises de componentes principais (ACP). Os mesmos autores relataram que os três métodos removeram efeitos verdadeiros associados com seleção em gado de leite, sendo que o processo de remoção é específico para cada método. Portanto, a análise de associação pode ser conduzida sem a correção para a estrutura da população (ZHAO et al., 2007; MA et al., 2012).

Outro fato que tem influência nos estudos de associação é o desequilíbrio de ligação que consiste na associação não aleatória entre alelos de dois ou mais loci que podem estar ou não fisicamente ligados, estando no mesmo ou em diferentes cromossomos. O elevado nível de desequilíbrio entre as marcas demonstram a existência de associação com o caráter avaliado. Fatores como: mistura de populações, deriva, mutação, seleção e recombinação com a prática melhoramento podem favorecer surgimento dessas associações preferenciais (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003). Em cana-de-açúcar, altos níveis de desequilíbrio de ligação são confirmados por Gonçalves (2012); Henry; Kole (2010) e

Jannoo et al.(1999), o que contribui na identificação de associações importantes relacionadas ao teor de açúcar e de fibra.

Nos casos de estudos de associação com espécies poliploides, como a cana, que possui mais de 4 cromossomos por grupo homólogo não é possível estimar o número de cópias ou a dosagem de cada alelo utilizando marcadores como AFLP, SSR (Simple Sequence Repeat) ou RFLP (Restriction Length Polymorphism), os quais foram desenvolvidos para aplicações em organismos diploides. Marcadores dominantes como AFLP ou SSR se tornam limitados ao contrário dos SNPs porque ao serem usados em cana-de-açúcar apenas informam a presença de ao menos uma cópia do alelo ou a ausência do mesmo, o qual pode-se atribuir que o determinado locus seja recessivo (GARCIA et al., 2013).

Neste estudo, foram consideradas como significativas as marcas que tiveram um fator de Bayes (BF) acima de 0,20 (Vidal et al., 2005). Todas as marcas significativas apresentaram valor da probabilidade de inclusão a posteriori acima de 0,60. Os Manhattan plots ilustram a PIP para todas as 640 marcas. Os pontos mais acima do eixo x representam as marcas com maior probabilidade de serem incluídas a posteriori na cadeia de Markov durante as iterações (Figuras 5 e 6). Este valor consiste na probabilidade que uma marca esteja verdadeiramente associada com o fenótipo (STEPHENS; BALDING, 2009). Yang et al. (2015) definem com que a frequência através da qual um determinado QTL é incluído no modelo serve para avaliar a importância do mesmo. Van Den Berg; Fritz e Boichard (2013), por meio de simulações mostraram que em cenários de baixa herdabilidade e grande número de QTLs são obtidos altos valores de Probabilidade de inclusão a Posteriori (PIP).

A identificação de marcadores moleculares associados às características agronômicas, incluindo experimentos em diferentes ambientes relacionadas ao período de safra em sucessivos anos, tem ganhado importância com diversas culturas na última década. Para Pol foram identificadas 52 marcas associadas significativamente (época 1), 70 marcas (época 2), 51 marcas (época 3 de 2015) e 69 marcas (época 3 colhida em 2016). Para fibra, 30 marcas (época 1), 26 marcas (época 2), 7 marcas (época 3 colhida em 2015) e 61 marcas (época 3 colhida em 2016).

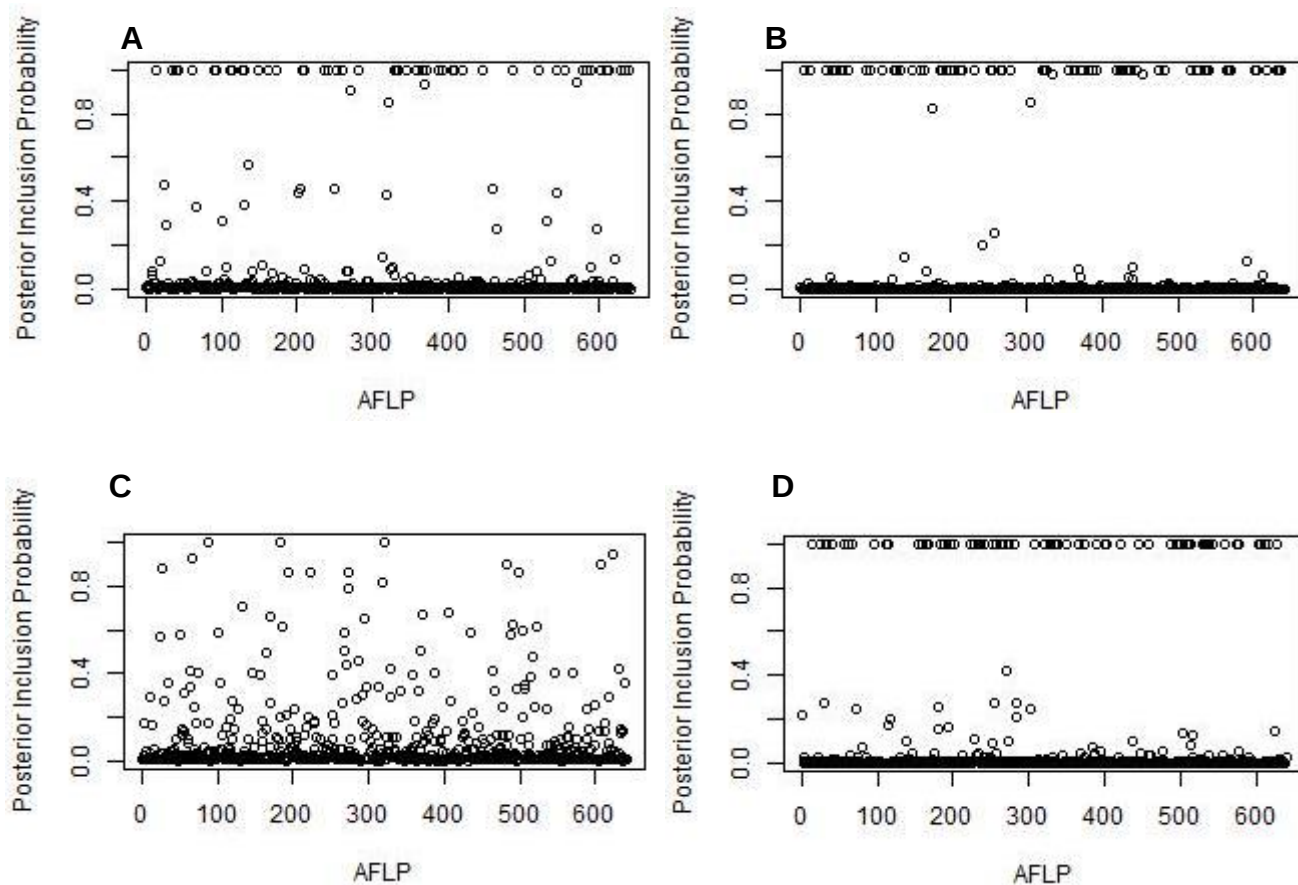
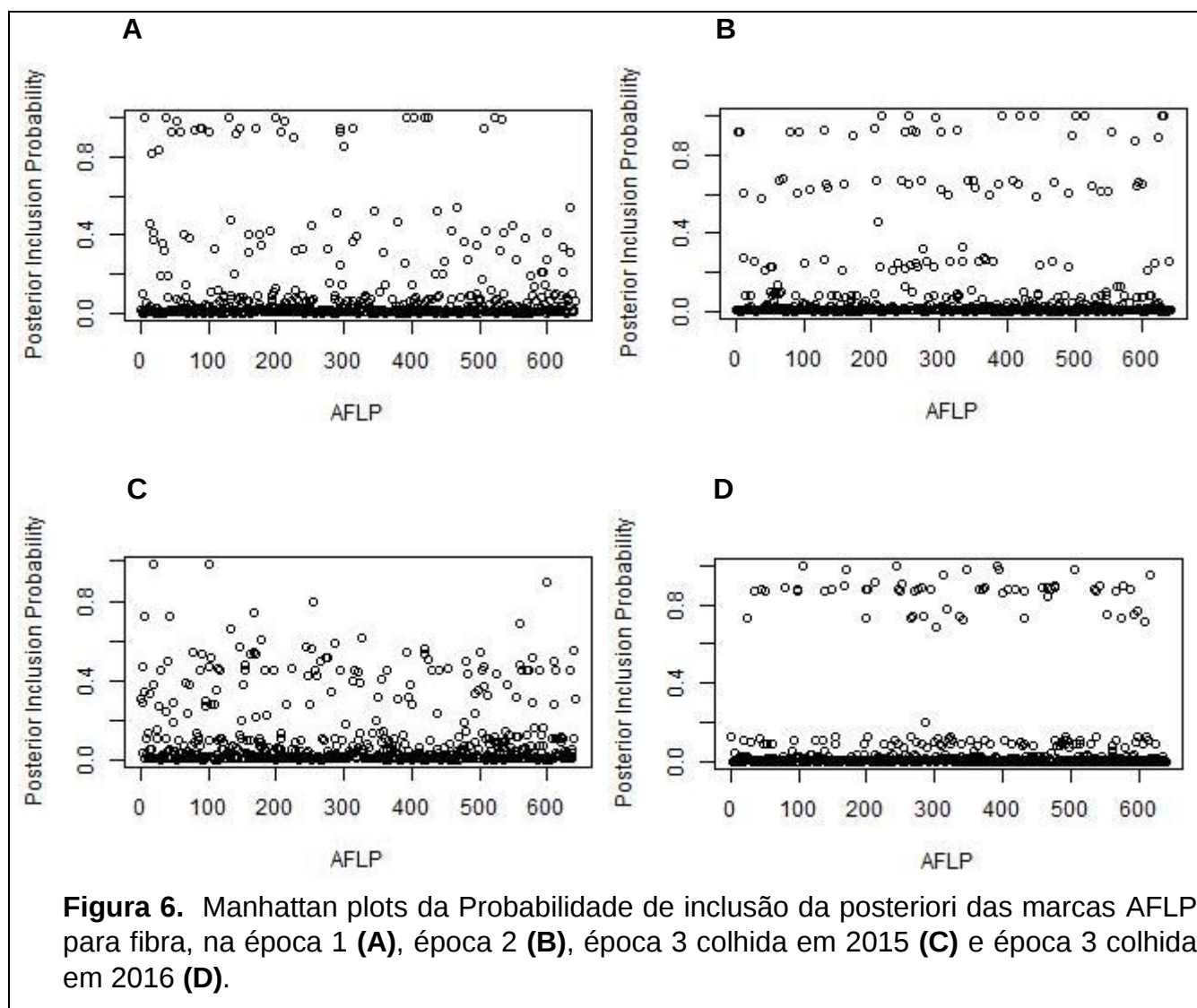


Figura 5. Manhattan plots da Probabilidade de inclusão da posteriori das marcas AFLP para Pol, na época 1 **(A)**, época 2 **(B)**, época 3 colhida em 2015 **(C)** e época 3 colhida em 2016 **(D)**.



Para verificar a validação dos efeitos associativos das marcas foram comparadas as épocas duas a duas com o propósito de identificar quais marcas apresentavam associação em mais de uma época.

Para a variável Pol, na comparação da época 1 com a época 2, 8 marcas foram coincidentes, com a época 3 de 2015, apenas 1 e com a época 3 de 2016, 12 marcas. Entre a época 2 e 3 de 2015, 8 marcas, e com a época 3 de 2016, 9 marcas. Na comparação entre as épocas 3 de 2015 e 2016, 5 marcas foram coincidentes (Tabela 3).

Para fibra, 2 marcas na comparação entre a primeira e segunda época, 1 marca entre a primeira e terceira de 2015, 9 entre a primeira e terceira de 2016. Para a

segunda e a terceira de 2015 e 2016, apenas 1 marca e entre a terceira de 2015 e 2016, nenhuma marca coincidente (Tabela 4).

A marca que teve efeito positivo associada a Pol para ambas as épocas 1 e 2 foi a AFLP 513 com 100% de probabilidade de inclusão a posteriori e está fortemente associada ao aumento de Pol, enquanto que as marcas, AFLP 159, AFLP 326, AFLP 435 e AFLP 773 tiveram efeito negativo e estão associadas ao decréscimo de açúcar. As marcas AFLP 435, AFLP 476, AFLP 549 apresentaram efeitos divergentes, isto é, positivo para uma época e negativo para outra. Isso significa que para uma determinada época a marca esteve associada a um locus que tenha efeito negativo e para outra pode estar associada a um locus com expressão de efeito positivo sobre o fenótipo. Para as épocas 1 e 3 uma única marca (AFLP 118) teve efeito positivo para ambas. Já com as épocas 1 e 3 de 2016, 6 marcas (AFLP 21, AFLP 291, AFLP 435, AFLP 579, AFLP 703, AFLP 742 e AFLP 783) mostraram efeitos positivos ligados ao aumento de sacarose. As marcas AFLP 84 e 429 apresentaram efeito negativo ligadas à redução de sacarose. As marcas AFLP 57 e 417 exibiram efeitos distintos entre as duas épocas, negativo na época 1 e positivo para a época 3 de 2016, ao contrário da marca AFLP 280 com efeito positivo na época 1 e negativo para a época 3 de 2016.

Observando-se os efeitos para as épocas 2 e 3 conjuntamente na variável Pol, as marcas AFLP 62, AFLP 327, AFLP 418 e AFLP 807 apresentaram efeitos positivos, e AFLP 554 e AFLP 811, efeito negativo. Com efeitos divergentes, surgiram as marcas AFLP 422 e 700 apresentando efeitos negativos na época 2 e positivos para a época 3 de 2015. Na comparação com a época 3 de 2016, houveram 3 marcas com efeito positivo (AFLP 75, AFLP 418 e AFLP 696) e com efeito negativo, quatro marcas (AFLP 228, 700, 776 e 789). Com efeitos divergentes, a marca AFLP 200 apresentou efeito positivo para a época 2 e negativo para a época 3 de 2016. Por fim, na verificação para os anos de 2015 e 2016 da época 3, as seguintes marcas tiveram efeito positivo: AFLP 418, AFLP 647 e AFLP 700. Não houve nenhuma marca com efeito negativo para os dois anos, mas com efeito negativo para 2015 e positivo para 2016. Das marcas citadas anteriormente, as marcas AFLP 418 e 700 apresentaram efeitos para as épocas 2 e 3 de 2015 e para a época 3 colhida em 2016, enquanto apenas as marcas: AFLP 435, AFLP 418 e AFLP 700 foram as únicas que apresentaram efeito para três épocas.

Tabela 3. Efeitos de marcas coincidentes associadas a Pol em pelo menos duas épocas de colheita de cana-de-açúcar.

Marcas	Época 1	Época 2	Época 3	Época 3 (2016)	Combinação	Tamanho em pares de base (pb)
159	-0,4937	-1,6225	*	*	MSPTTG/ECOACA-IRD700	200
326	-0,2469	-0,0868	*	*	MSPACC/ECOACC-IRD700	365
476	-0,7725	1,0593	*	*	MSPGAG/ECOACC-IRD700	183
513	0,8089	1,6343	*	*	MSPACA/ECOACG-IRD700	395
549	0,8058	-0,3778	*	*	MSPTCG/ECOACG-IRD700	268
773	-0,7155	-0,0998	*	*	MSPTAG/ECOAG-IRD800	151
735	-0,1722	-0,0534	*	*	MSPACA/ECOAG-IRD800	189
118	1,0443	*	0,3218	*	MSPACA/ECOACA-IRD700	322
21	0,4534	*	*	1,9495	MSPTTG/ECOAGA-IRD800	231
57	-0,3563	*	*	0,8550	MSPTAG/ECOAGA-IRD800	266
84	-0,3183	*	*	-1,3366	MSPGAA/ECOAGA-IRD800	292
280	0,4433	*	*	-0,8190	MSPGAG/ECOACA-IRD700	252
291	0,6113	*	*	0,5536	MSPGAG/ECOACA-IRD700	67
429	-1,0254	*	*	-0,4901	MSPACT/ECOACC-IRD700	175
435	1,0343	-0,5952	*	0,4339	MSPACT/ECOACC-IRD700	145
579	1,2693	*	*	0,6943	MSPTAG/ECOACG-IRD700	152
703	1,4039	*	*	1,6960	MSPACT/ECOAGG-IRD800	116
742	1,3441	*	*	1,6657	MSPACA/ECOAG-IRD800	149
783	0,2882	*	*	0,5400	MSPACT/ECOAGG-IRD800	83
417	-0,2094	*	*	0,6168	MSPACT/ECOACC-IRD700	302
62	*	0,2478	0,1974	*	MSPTAG/ECOAGA-IRD800	175
327	*	1,5032	0,2397	*	MSPACC/ECOACC-IRD700	358
422	*	-1,3798	0,5461	*	MSPACT/ECOACC-IRD700	235
554	*	-1,6570	-0,6083	*	MSPTCG/ECOACG-IRD700	240
807	*	0,6544	0,1728	*	MSPCGC/ECOAG-IRD800	89
811	*	-0,7149	-0,1276	*	MSPCGC/ECOAG-IRD800	73
75	*	0,5980	*	0,8856	MSPCGC/ECOAGA-IRD800	188
200	*	0,2919	*	-1,1454	MSPACC/ECOACA-IRD700	168
228	*	-0,3930	*	-0,3250	MSPTAG/ECOACA-IRD700	293
418	*	0,4430	0,2282	1,0140	MSPACT/ECOACC-IRD700	284
518	*	-0,3849	*	1,5834	MSPACA/ECOACG-IRD700	350
696	*	0,7554	*	0,4958	MSPACT/ECOAGG-IRD800	164
700	*	-0,2440	0,0605	-0,1288	MSPACT/ECOAGG-IRD800	151
776	*	-0,3695	*	-0,1871	MSPTAG/ECOAG-IRD800	135
789	*	-0,3356	*	-0,9765	MSPCGC/ECOAG-IRD800	208
223	*	*	-0,5683	-1,5242	MSPTAG/ECOACA-IRD700	323
647	*	*	0,2028	0,2548	MSPACC/ECOAGG-IRD800	167
142	*	*	-0,0675	0,2527	MSPTTG/ECOACA-IRD700	342

*Ausência de associação nas épocas.

Na variável fibra, foram identificadas para a época 1, 30 marcas significativas, para a época 2, 26 marcas significativas, 7 marcas na época 3 do ano de 2015 e 61 marcas no ano de 2016. Conforme tabela 4, entre as épocas 1 e 2, somente uma marca (AFLP 163) esteve presente em ambas épocas com efeito positivo muito fortemente associado com teor de fibra. Enquanto que a marca AFLP 520 exibiu efeito negativo e está relacionada ao decréscimo ao teor de fibra. Entre as épocas 1 e 3 houve apenas efeito negativo associada a AFLP 13. Para as épocas 1 e 3 colhida em 2016, tiveram efeito positivo, as marcas AFLP 548, AFLP 252, AFLP 207, AFLP 409, AFLP 652 e AFLP 381. Enquanto os efeitos negativos foram exibidos pela AFLP 239 e AFLP 520. Apenas a marca AFLP 63 apresentou efeito positivo na época 1 e negativo para a época 3 de 2016. Nas épocas 2 e 3 de 2015, houve apenas efeito negativo associado à marca AFLP 327. O mesmo aconteceu com as épocas 2 e 3, e que a marca AFLP 520 apresentou efeito negativo, além disso, a mesma marca apresentou também efeito para três épocas.

Foram identificadas poucas marcas coincidentes para todas as épocas de colheita. No entanto, as marcas com efeito específico para cada época podem oscilar de acordo com as condições ao longo dos anos (MARGARIDO et al., 2015). Os mesmos autores observaram que muitos efeitos de QTLs são apenas para um único locus e colheita e além disto, foram sempre consistentemente positivos ou negativos acerca dos locais e colheitas. Lopes (2011) utilizando modelos mistos testando quatro modelos de associação para seis características de cana-de-açúcar em três locais identificou três marcas associadas para todas as regiões avaliadas. Para Ribeirão Preto, utilizando marcadores AFLP e SSR identificou 8 marcas associadas a Pol%cana. Hu et al. (2014) também identificaram apenas duas marcas SSR associadas a produtividade em soja em dois locais.

Racedo et al. (2016) utilizando marcadores DArTs e TRAP para Pol detectaram 38, 34 e 41 marcadores com efeito associativo significativo para cana planta, primeira soca e segunda soca, respectivamente. Do total de marcas associadas, apenas uma marca esteve associada aos três ciclos. Diferentemente dos resultados apresentados por Racedo et al. (2016), não foi identificada nenhuma associada a ambas

características (fibra e Pol). Sendo assim, mais uma vez fica evidenciado a existência de poucas marcas associadas ao fenótipo em vários ambientes.

Como pode ser observado, para a variável Pol houve uma quantidade de marcas significativas variando de 50 a 70 no total das épocas avaliadas, porém poucas tiveram efeitos coincidentes em todas épocas. Isto pode ser a provável existência de falsos positivos. Gouy et al. (2015) avaliando 13 características, entre agronômicas, morfológicas, tecnológicas e resistência a doença em cinco locais também obtiveram apenas 11 marcadores significativamente associados com 3 das 13 características. Os mesmos observaram para fibra em que uma única marca esteve associada com mais de duas épocas. Esta baixa frequência de coincidência de associação marcador com a característica foi também observada em outros estudos de associação para o teor de sacarose em diferentes épocas, locais e anos como resultados obtidos por Racedo et al. (2016), nos quais um marcador, M64, teve efeito significativo em relação ao teor de sacarose para os três anos e 12 marcadores com associação para primeira e segunda soca, enquanto 20 foram associados para produção em 2 anos, entre cana planta, primeira ou segunda soca para a produção de açúcar.

Hoarau et al. (2002) identificaram em três anos apenas um QTL com efeito por dois anos para características relacionadas a produção de açúcar como peso de colmo, comprimento de colmo e número de colmo. Enquanto que Aitken et al. (2008) conduzindo mapeamento de QTLs em população biparentais identificaram 18 alelos com efeito consistente em mais de uma ano e ao menos em uma característica.

Na experimentação agrícola, o ambiente consiste em locais, anos agrícolas, épocas de plantio, épocas de colheita, sistemas de cultivo (sequeiro ou irrigado). A consistência da associação da marca com o fenótipo avaliado é um fator observado quando experimentos são conduzidos em vários ambientes. A presença de associações nos ambientes utilizados quando presentes indica a não existência da interação QTL x ambiente, enquanto a associação da marca presente apenas em um ou poucos ambientes é uma evidência de interação (LYNCH e WALSH, 1998).

Quanto à variação fenotípica explicada pelas marcas significativamente associadas, na época 1, as 52 marcas associadas ao conteúdo de sacarose (Pol) explicaram conjuntamente 99,51% da variância fenotípica. Na época 2, as 70

associadas explicaram 99,75%, na época 3, 51 marcas (97,56%) e na época 3 colhida em 2016 (69 marcas), 73,92% da variância fenotípica.

Para fibra, as marcas significativamente associadas às épocas 1 (30 marcas), 2 (26 marcas), 3 (7 marcas) e 3 colhida em 2016 (61 marcas) explicaram 97,83%, 94,82%, 68,44% e 74,66%, respectivamente da variação fenotípica dos teores de fibra. Estes valores correspondem ao coeficiente de determinação (R^2) que consiste na porcentagem de variação fenotípica explicada pela variação dos efeitos de associação das marcas com o fenótipo.

Valores de R^2 apresentados por Glauce (2005) foram inferiores na explicação da variação fenotípica da produtividade de milho em vários locais com vários testadores. A proporção da variação fenotípica explicada individualmente por cada marca apresenta baixo valor, como por exemplo, nos resultados obtidos por Toledo (2006) utilizando análise bayesiana no mapeamento de QTL para produtividade de grãos em milho em que a proporção variou de 0,09% a 0,69%, enquanto que para os cinco QTLs conjuntamente, a explicação foi 1,93%.

Quando se utilizou as marcas significativas de uma época para predizer o efeito sobre o mesmo fenótipo em outra época, todos os R^2 ficaram acima de 0,50, ou seja, conseguiram explicar mais de 50% da variação fenotípica em outra época.

Para fibra, apenas quando se utilizou as marcas significativas da época 3 para predizer o fenótipo na época 3 colhida em 2016 observou-se um coeficiente de determinação abaixo de 50%, apenas 23,93% da variação fenotípica existente na época 3 de 2016 foi explicada pelas marcas significativas identificadas em 2015.

Foram analisadas também as correlações das marcas entre as épocas de colheita para Pol e fibra e selecionadas apenas de valores acima 0,7 e de -0,7. As correlações identificadas como positivas entre épocas distintas revelam que para trabalhos posteriores de seleção pode-se utilizar tanto uma marca como outra para uma dada característica sob avaliação. A magnitude das correlações negativas variaram de -0,86 a -0,71, enquanto que nas correlações positivas foram identificadas valores de 0,70 a 0,89 (Apêndice 2A).

Para fibra, as correlações variaram de -0,81 a 0,70, totalizando seis correlações negativas, enquanto que nove marcas apresentaram correlações positivas de 0,71 a 0,88, (Apêndice 3A).

Tabela 4. Efeitos de marcas coincidentes associadas à fibra em quatro épocas de colheita de cana-de-açúcar.

Marcas	Fibra				Combinação	Tamanho em pares de base (pb)
	Época 1	Época 2	Época 3	Época 3 (2016)		
AFLP163	0,6804	0,5678	*	*	MSPTTG/ECOACA-IRD700	179
AFLP520	-0,6957	-1,4095	*	-0,5746	MSPACA/ECOACG-IRD700	334
AFLP13	-1,1886	*	-0,4347	*	MSPTTG/ECOAGA-IRD800	297
AFLP548	0,7253	*	*	0,9959	MSPTCG/ECOACG-IRD700	270
AFLP239	-0,5728	*	*	-1,4514	MSPTAG/ECOACA-IRD700	207
AFLP252	0,4036	*	*	0,9785	MSPCGC/ECOACA-IRD700	324
AFLP207	0,4487	*	*	0,6610	MSPACC/ECOACA-IRD700	116
AFLP409	0,4344	*	*	0,3541	MSPGAG/ECOACC-IRD700	178
AFLP652	0,5758	*	*	0,6819	MSPACC/ECOAGG-IRD800	113
AFLP381	0,5143	*	*	0,6567	MSPCGC/ECOACC-IRD700	145
AFLP63	0,6381	*	*	-0,4330	MSPTAG/ECOAGA-IRD800	173
AFLP327	*	-1,8448	-0,5004	*	MSPACC/ECOACC-IRD700	358

*Ausência de efeito

Um fator que interfere na associação da marca com o fenótipo é o desequilíbrio de ligação (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003). Quanto maior a extensão do DL, mais associado está a marca com o locus relacionado ao fenótipo. Como o genoma da cana possui 18.000 cM de comprimento, dificulta os trabalhos de mapeamento por causa da necessidade de grande quantidade de marcadores necessários para cobrir todos os cromossomos (ZHANG et al., 2009). Os resultados mostrados por Gonçalves (2012), Jannoo et al. (1999) e Raboin et al. (2008) com cultivares comerciais brasileiras e de outros países confirmam o alto nível de desequilíbrio de ligação existente na espécie. Isso justifica a ideia que, mesmo utilizando a quantidade de 640 marcas polimórficas após a exclusão daquelas com menos de 5% e mais de 95% de frequência há a possibilidade de se identificar associações com o fenótipo.

6 CONCLUSÕES

As marcas AFLP 418, 435 e 700 tiveram associações em três épocas de colheita para Pol. Para a variável fibra, a marca AFLP 520 foi a única com efeito consistente para três épocas (época 1, 2 e 3 de 2016). Estas marcas são importantes para serem validadas em outros materiais no processo de seleção de um programa de melhoramento. A maioria das marcas significativas apresentou efeito específico para cada época, mas as mesmas marcas conjuntamente de uma época servem para prever o efeito em relação ao fenótipo para outra época de colheita.

REFERÊNCIAS

AITKEN, K. S.; HERMANN, S.; KARNO, K.; BONNETT, G. D.; MCINTYRE, L. C.; JACKSON, P. A. Genetic control of yield related traits in sugarcane. **Theoretic and Applied Genetics**, v. 117, n. 7, p. 1191-1203, 2008.

BROWN, J. S.; SCHNELL, R. J.; POWER, E. J.; DOUGLAS, S. L.; KUHN, D. N. Analysis of clonal germplasm from five *Saccharum* species: *S. barberi*, *S. robustum*, *S. officinarum*, *S. sinense* and *S. spontaneum*. A study of inter- and intra species relationships using microsatellite markers. **Genetic Resources and Crop Evolution**, v. 54, n. 3, p. 627-648, 2007.

Consecana. 2006. Conselho dos produtores de cana-de-açúcar, açúcar e álcool do estado de São Paulo. Manual de Instruções, Piracicaba.

DEBIBAKAS, S.; ROCHER, S.; GARSMEUR, O.; TOUBI, L.; ROQUES, D.; D'HONT, A.; HOARAU, J.-Y.; DAUGROIS, J. H. Prospecting sugarcane resistance to Sugarcane yellow leaf virus by genome-wide association. **Theoretic and Applied Genetics**, v. 127, n. 8, p. 1719-1732, 2014.

EARL, D. A.; VONHOLDT, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, v. 4, n. 2, p. 359-361, 2012.

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

FLINT-GARCIA, S. A.; THORNSBERRY, J. M.; IV, B. Structure of linkage disequilibrium in plants*. **Annu review of plant biology**, v. 54, n. 1, p. 357-374, 2003.

FUSH, D.; STEPHENS, M.; PRITCHARD, J. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. **Genetics**, v. 164, 2003.

GARCIA, A. A.; MOLLINARI, M.; MARCONI, T. G.; SERANG, O. R.; SILVA, R. R.; VIEIRA, M. L.; VICENTINI, R.; COSTA, E. A.; MANCINI, M. C.; GARCIA, M. O. SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. **Scientific reports**, v. 3, 2013.

GONÇALVES, B. S. **Identificação de marcadores moleculares relacionados à severidade de sintomas da ferrugem em painel associativo de cana-de-açúcar (*Saccharum* spp.)**. 2012. 54 f. Dissertação de mestrado -Universidade Estadual Paulista Jabotical, 2012.

GOUY, M.; ROUSSELLE, Y.; THONG CHANE, A.; ANGLADE, A.; ROYAERT, S.; NIBOUCHE, S.; COSTET, L. Genome wide association mapping of agro-morphologic and disease resistance traits in sugarcane. **Euphytica**, v. 202, n. 2, p. 269-284, 2015.

GUPTA, P. K.; RUSTGI, S.; KULWAL, P. L. Linkage disequilibrium and association studies in higher plants: present status and future prospects. **Plant molecular biology**, v. 57, n. 4, p. 461-485, 2005.

HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. **Genetics**, v. 177, n. 4, p. 2389-2397, 2007.

HABIER, D.; FERNANDO, R. L.; KIZILKAYA, K.; GARRICK, D. J. Extension of the Bayesian phabet for genomic selection. **BMC bioinformatics**, v. 12, n. 1, p. 1, 2011.

HENRY, R. J.; KOLE, C. **Genetics, genomics and breeding of sugarcane**. CRC Press, 2010. p.

HOARAU, J.-Y.; GRIVET, L.; OFFMANN, B.; RABOIN, L.-M.; DIORFLAR, J.-P.; PAYET, J.; HELLMANN, M.; D'HONT, A.; GLASZMANN, J.-C. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.).II. Detection of QTLs for yield components. **Theoretic and Applied Genetics**, v. 105, n. 6, p. 1027-1037, 2002.

JACKSON, P. A. Breeding for improved sugar content in sugarcane. **Field Crops Research**, v. 92, n. 2, p. 277-290, 2005.

JANABI, S. M.; FORGET, L.; DOOKUN, A. An improved and rapid protocol for the isolation of polysaccharide-and polyphenol-free sugarcane DNA. **Plant Molecular Biology Reporter**, v. 17, n. 3, p. 281-281, 1999.

JANNOO, N.; GRIVET, L.; DOOKUN, A.; D'HONT, A.; GLASZMANN, J. Linkage disequilibrium among modern sugarcane cultivars. **Theoretic and applied genetics**, v. 99, n. 6, p. 1053-1060, 1999.

JOMBART, T.; DEVILLARD, S.; BLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **BMC Genetics**, v. 11, n. 1, p. 94, 2010.

LANDELL, M. G. A.; SILVA, D. N.; ANJOS, A. I.; MENDONÇA, J. R.; BIDOIA, M. A. P.; CAMPANA, M. P.; XAVIER, M. A.; FIGUEIREDO, P. **Variedades de cana-de-açúcar para o Centro-Sul do Brasil**. Instituto Agronômico de Campinas. Campinas, p.37. 2007. (201).

LAVANHOLI, M. D. G. D. P. Qualidade da cana-de-açúcar como matéria-prima para produção de açúcar e álcool. In: DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M.; LANDELL, M. G. D. A. **Cana-de-açúcar**. 1ª. Campinas: Instituto Agronômico de Campinas, 2010. cap. 8, p. 882.

LEGARRA, A.; RICARD, A.; FILANGI, O. GS3. **Genomic selection, Gibbs sampling, Gauss Seidel and Bayes Cpi**. Disponível <http://snp.toulouse.inra.fr/~alegarra> 2012. Acesso em: 20 de out. 2016.

LEGARRA, A.; ROBERT-GRANIÉ, C.; CROISEAU, P.; GUILLAUME, F.; FRITZ, S. Improved Lasso for genomic selection. **Genetics research**, v. 93, n. 01, p. 77-87, 2011.

LI, M.; REILLY, M. P.; RADER, D. J.; WANG, L.-S. Correcting population stratification in genetic association studies using a phylogenetic approach. **Bioinformatics**, v. 26, n. 6, p. 798-806, 2010.

LOPES, F. C. C. **Mapeamento genético de cana-de-açúcar (*Saccharum spp.*) por associação empregando marcadores SSR e AFLP**. 2011. 144 f. Tese de doutorado -Universidade de São Paulo - Escola Superior de Agricultura "Luiz de Queiroz" Piracicaba, 2011.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 980 p.

MA, L.; WIGGANS, G. R.; WANG, S.; SONSTEGARD, T. S.; YANG, J.; CROOKER, B. A.; COLE, J. B.; VAN TASSELL, C. P.; LAWLOR, T. J.; DA, Y. Effect of sample stratification on dairy GWAS results. **BMC Genomics**, v. 13, n. 1, p. 536, 2012.

MARGARIDO, G. R. A.; PASTINA, M. M.; SOUZA, A. P.; GARCIA, A. A. F. Multi-trait multi-environment quantitative trait loci mapping for a sugarcane commerci cross provides insights on the inheritance of important traits. **Molecular Breeding**, v. 35, n. 8, p. 175, ,2015.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. **Genetics**, v. 157, n. 4, p. 1819-1829, 2001.

MORAES, M. F. D.; BASTOS, G. Q.; ANUNCIAÇÃO FILHO, C. J.; MELO, L. J. O. T.; REIS, O. V. D. Aviação agroindustri e parâmetros genético de progênies de cana-de-açúcar em fase inici na Zona Canavieira do Litor Norte de Pernambuco. **Ciênc. agrotec.,(Impr.)**, v. 34, n. 5, p. 1086-1092, 2010.

PATTERSON, N.; PRICE, A. L.; REICH, D. Population structure and eigenanysis. **PLoS genet**, v. 2, n. 12, p. e190, 2006.

PRICE, A. L.; PATTERSON, N. J.; PLENGE, R. M.; WEINBLATT, M. E.; SHADICK, N. A.; REICH, D. Princip components anysis corrects for stratification in genome-wide association studies. **Nature Genetics**, v. 38, 2006.

PRICE, A. L.; ZAITLEN, N. A.; REICH, D.; PATTERSON, N. New approaches to population stratification in genome-wide association studies. **Nature Reviews Genetics**, v. 11, n. 7, p. 459-463, 2010.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of Population Structure Using Multilocus Genotype Data. **Genetics**, v. 155, n. 2, p. 945-959, 2000.

PRITCHARD, J. K.; STEPHENS, M.; ROSENBERG, N. A.; DONNELLY, P. Association mapping in structured populations. **The American Journ of Human Genetics**, v. 67, n. 1, p. 170-181, 2000.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL <https://www.R-project.org/>. Vienna, Austria.

RABOIN, L.-M.; PAUQUET, J.; BUTTERFIELD, M.; D'HONT, A.; GLASZMANN, J.-C. Anysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. **Theoretic and Applied Genetics**, v. 116, n. 5, p. 701-714, 2008.

RACEDO, J.; GUTIÉRREZ, L.; PERERA, M. F.; OSTENGO, S.; PARDO, E. M.; CUENYA, M. I.; WELIN, B.; CASTAGNARO, A. P. Genome-wide association mapping of quantitative traits in a breeding population of sugarcane. **BMC Plant Biology**, v. 16, n. 1, p. 142, 2016.

RAO, M. S.; WEERATHAWORN, P. Diversification of breeding program to develop multipurpose sugarcane cultivars. **Sugar Tech**, v. 11, n. 1, p. 77-79, 2009.

SANTOS, F. R. C.; ZUCCHI, M. I.; PARK, J.-W.; BENATTI, T. R.; DA SILVA, J. A.; SOUZA, G. M.; LANDELL, M. G. A.; PINTO, L. R. New Sugarcane Microsatellites and Target Region Amplification Polymorphism Primers Designed from Candidate Genes Related to Disease Resistance. **Sugar Tech**, p. 1-6, 2016.

SHI, A.; CHEN, P.; ZHANG, B.; HOU, A. Genetic diversity and association analysis of protein and oil content in food-grade soybeans from Asia and the United States. **Plant Breeding**, v. 129, n. 3, p. 250-256, 2010.

SIGRIST, M. S. **Mapeamento associativo de locos relacionados à produtividade de grãos em soja**. 2012. f. -Escola Superior de Agricultura "Luiz de Queiroz, 2012.

VAN DEN BERG, I.; FRITZ, S.; BOICHARD, D. QTL fine mapping with Bayes C (π): a simulation study. **Genetics Selection Evolution**, v. 45, n. 1, p. 1, 2013.

WEI, X.; JACKSON, P. A.; MCINTYRE, C. L.; AITKEN, K. S.; CROFT, B. Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. **Theoretic and Applied Genetics**, v. 114, n. 1, p. 155-164, 2006.

XUE, Y.; Warburton, M. L.; SAWKINS, M.; ZHANG, X.; SETTER, T.; XU, Y.; GRUDLOYMA, P.; GETHI, J.; RIBAUT, J.-M.; LI, W.; ZHANG, X.; ZHENG, Y.; YAN, J. Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. **Theoretic and Applied Genetics**, v. 126, n. 10, p. 2587-2596, 2013.

YANG, D.; HAN, S.; JIANG, D.; YANG, R.; FANG, M. Inference of posterior inclusion probability of QTLs in Bayesian shrinkage analysis. **Genetics research**, v. 97, p. e6, 2015.

YU, J.; BUCKLER, E. S. Genetic association mapping and genome organization of maize. **Current Opinion in Biotechnology**, v. 17, n. 2, p. 155-160, 2006.

ZHANG, L. Y.; MARCHAND, S.; TINKER, N. A.; BELZILE, F. Population structure and linkage disequilibrium in barley assessed by DArT markers. **Theoretic and applied genetics**, v. 119, n. 1, p. 43-52, 2009.

ZHAO, K.; ARANZANA, M. J.; KIM, S.; LISTER, C.; SHINDO, C.; TANG, C.; TOOMAJIAN, C.; ZHENG, H.; DEAN, C.; MARJORAM, P. An Arabidopsis example of association mapping in structured samples. **PLoS Genet**, v. 3, n. 1, p. e4, 2007.

ZHOU, X.; STEPHENS, M. Genome-wide efficient mixed-model analysis for association studies. **Nature genetics**, v. 44, n. 7, p. 821-824, 2012.

ZHU, X.; LI, S.; COOPER, R. S.; ELSTON, R. C. A unified association analysis approach for family and unrelated samples correcting for stratification. **The American Journal of Human Genetics**, v. 82, n. 2, p. 352-365, 2008.

APÊNDICES

Apêndice 1A. Combinações dos primers AFLP utilizados na genotipagem da população. As sequências abaixo correspondem aos três nucleotídeos seletivos posicionados na extremidade 3' dos iniciadores *EcoRI* e *MspI*.

Nº da combinação	Sequência 3' do primer		Nº da combinação	Sequência 3' do primer	
	<i>EcoRI</i>	<i>MspI/Hpa</i>		<i>EcoRI</i>	<i>MspI/Hpa</i>
1	AGA-IRD800	ACA	26	ACC-IRD700	ACT
2	AGA-IRD800	TTG	33	AAC-IRD700	CGC
3	AGA-IRD800	ACC	34	AAC-IRD700	GAG
5	AGA-IRD800	TAG	36	AAC-IRD700	GAA
6	AGA-IRD800	CGC	37	ACG-IRD700	ACA
9	AGA-IRD800	GAA	40	ACG-IRD700	TCG
10	ACA-IRD700	ACA	41	ACG-IRD700	TAG
11	ACA-IRD700	TTG	42	ACG-IRD700	CGC
12	ACA-IRD700	ACC	48	AGG-IRD800	ACC
14	ACA-IRD700	TAG	49	AGG-IRD800	TCG
15	ACA-IRD700	CGC	51	AGG-IRD800	CGC
16	ACA-IRD700	GAG	53	AGG-IRD800	ACT
18	ACA-IRD700	GAA	54	AGG-IRD800	GAA
20	ACC-IRD700	TTG	55	AAG-IRD800	ACA
21	ACC-IRD700	ACC	56	AGG-IRD800	TTG
22	ACC-IRD700	TCG	59	AGG-IRD800	TAG
24	ACC-IRD700	CGC	60	AGG-IRD800	CGC
25	ACC-IRD700	GAG			

Apêndice 2A. Lista das correlações das marcas entre as épocas de colheita para Pol com valores acima de -0,70 e 0,7.

Marca 1	Marca 2	Correlação
291	441	-0,8620
291	424	-0,8400
190	291	-0,7655
88	441	-0,7442
291	513	-0,7402
89	291	-0,7395
88	190	-0,7205
88	424	-0,7174
88	89	-0,7098
424	794	-0,7095
190	246	0,7016
89	349	0,7044
195	441	0,7123
89	424	0,7170
21	325	0,7366
424	513	0,7556
89	441	0,7577
59	474	0,7608
88	291	0,7682
89	190	0,7759
190	424	0,7811
89	513	0,8052
441	513	0,8071
190	441	0,8312
190	513	0,8801
424	441	0,8916

Apêndice 3A. Lista das correlações das marcas entre as épocas de colheita para fibra com valores acima de -0,70 e 0,7

Marca 1	Marca 2	Correlação
36	69	0,7783
36	368	0,7831
36	616	-0,7060
36	764	0,8801
69	345	0,7169
69	368	0,8261
69	616	-0,7424
69	764	0,8883
73	616	0,7155
368	616	-0,7251
368	764	0,8859
381	703	-0,7058
616	719	0,7261
616	764	-0,8145
719	764	-0,7804

Apêndice 4A: Médias ajustadas para o atributo fibra pelo Least Squares Means (LSMEANS) para as três épocas de colheita em Ribeirão Preto-SP.

Época 1		Época 2		Época 3		Época 3*	
Genótipos	Médias	Genótipos	Médias	Genótipos	Médias	Genótipos	Médias
94	20,38	8	22,66	94	24,53	94	21,91
41	19,32	43	21,21	39	21,66	43	21,57
43	19,07	94	20,70	95	21,53	40	21,38
39	18,83	41	20,69	43	21,31	95	20,65
11	18,53	44	19,48	36	19,99	36	20,25
95	17,72	39	18,44	11	19,52	39	19,88
37	17,16	99	17,73	44	19,14	99	19,86
42	16,93	95	17,60	38	18,84	44	19,61
99	16,93	36	17,59	40	18,78	42	19,43
36	16,68	93	16,93	37	18,51	38	19,40
38	16,48	42	16,64	41	18,50	37	19,21
44	16,40	38	16,47	42	18,43	93	19,09
1	16,14	40	16,44	93	18,33	41	18,79
93	15,97	35	16,29	35	17,36	10	18,68
10	15,93	37	16,11	99	16,49	11	18,15
40	15,30	10	15,96	1	15,75	35	17,37
35	14,43	11	15,58	8	15,60	1	16,13
32	14,01	57	13,89	10	15,53	96	14,74
8	13,55	1	13,83	21	14,67	60	14,51
21	13,28	50	13,18	32	13,62	21	14,50
60	13,22	32	13,16	74	13,59	50	14,41
50	13,11	74	12,55	12	13,11	32	14,40
16	12,92	34	12,40	50	13,08	74	14,28
9	12,62	14	12,40	60	13,01	14	13,76
12	12,54	60	12,25	33	12,78	33	13,44
33	12,50	21	12,09	14	12,74	65	13,41
61	12,39	33	12,05	18	12,68	100	13,37
100	12,29	80	11,92	68	12,56	18	13,21
14	12,26	19	11,89	100	12,53	82	13,21
25	12,19	12	11,86	69	12,47	30	13,19
56	12,18	71	11,80	75	12,46	69	13,16
82	12,15	30	11,72	77	12,45	64	13,15
68	12,13	75	11,71	72	12,43	68	13,14
64	12,12	66	11,70	59	12,42	12	13,13
59	12,11	100	11,63	66	12,28	67	13,05
74	12,09	83	11,58	61	12,24	61	13,03
66	12,02	22	11,57	70	12,22	66	12,99
70	11,99	61	11,55	83	12,17	72	12,91

63	11,96	6	11,51	82	12,16	63	12,87
19	11,94	64	11,50	63	12,15	29	12,87
78	11,88	77	11,44	71	12,15	78	12,85
26	11,85	4	11,44	28	12,00	55	12,78
30	11,80	68	11,44	16	11,99	77	12,77
34	11,80	70	11,41	78	11,86	83	12,77
18	11,79	82	11,39	89	11,83	56	12,77
67	11,79	18	11,36	2	11,80	89	12,77
72	11,76	65	11,35	79	11,79	59	12,73
89	11,74	29	11,35	65	11,78	34	12,73
75	11,72	2	11,31	86	11,75	88	12,72
47	11,72	45	11,30	64	11,74	71	12,69
83	11,71	78	11,27	57	11,74	70	12,63
71	11,71	67	11,25	25	11,73	75	12,57
45	11,69	63	11,22	45	11,71	19	12,55
55	11,69	98	11,22	19	11,69	45	12,48
65	11,69	88	11,19	56	11,67	57	12,40
2	11,67	26	11,19	34	11,65	58	12,36
57	11,58	62	11,18	67	11,65	9	12,32
98	11,55	59	11,14	6	11,61	86	12,31
69	11,55	20	11,11	20	11,59	25	12,28
88	11,51	54	11,11	4	11,59	79	12,26
46	11,48	79	11,09	55	11,57	87	12,23
77	11,46	56	11,09	27	11,53	26	12,23
80	11,46	89	11,08	97	11,52	28	12,16
28	11,46	16	11,08	58	11,47	84	12,12
79	11,42	90	11,06	88	11,46	20	12,09
29	11,40	58	11,06	3	11,44	98	12,07
62	11,28	47	11,04	62	11,44	85	12,05
20	11,28	28	11,03	46	11,42	16	12,04
86	11,28	25	11,01	76	11,42	22	12,02
97	11,15	86	10,97	26	11,41	47	11,99
90	11,14	53	10,96	85	11,38	92	11,93
84	11,04	97	10,96	13	11,37	62	11,87
27	11,03	92	10,87	92	11,34	48	11,85
58	10,99	85	10,85	80	11,30	52	11,84
22	10,91	69	10,85	47	11,29	73	11,83
96	10,84	46	10,84	29	11,27	13	11,79
15	10,83	27	10,82	84	11,20	54	11,75
91	10,82	9	10,79	9	11,15	97	11,74
54	10,81	72	10,78	98	11,12	81	11,72
92	10,78	87	10,77	30	11,10	27	11,70

7	10,75	91	10,75	87	11,02	4	11,69
87	10,74	13	10,72	15	11,01	90	11,68
4	10,68	73	10,58	7	11,01	53	11,67
85	10,64	48	10,56	48	10,99	80	11,67
31	10,60	15	10,54	96	10,92	7	11,57
24	10,56	81	10,54	22	10,86	3	11,55
48	10,56	31	10,53	53	10,85	91	11,52
73	10,55	17	10,51	54	10,84	46	11,49
52	10,48	76	10,49	90	10,81	17	11,43
81	10,46	55	10,49	5	10,79	76	11,42
13	10,45	51	10,46	31	10,79	6	11,16
6	10,44	96	10,44	91	10,74	15	11,13
76	10,44	84	10,44	24	10,66	49	11,11
53	10,42	52	10,38	73	10,60	51	11,08
3	10,12	24	10,34	81	10,60	23	11,06
5	10,12	49	10,33	23	10,49	24	10,98
17	10,06	5	10,26	52	10,36	31	10,94
51	9,96	23	9,66	49	10,31	5	10,72
23	9,91	7	9,63	51	10,31	pd	pd
49	7,72	3	9,44	17	10,12	pd	pd

*Época 3= (colhida em 2016); pd: parcela perdida

Apêndice 5A: Médias ajustadas para o atributo Pol pelo Least Squares Means (LSMEANS) para as três épocas de colheita em Ribeirão Preto-SP.

Época 1		Época 2		Época 3		Época 3*	
Genótipos	Médias	Genótipos	Médias	Genótipos	Médias	Genótipos	Médias
80	16,27	19	17,30	22	18,73	51	18,77
18	15,74	71	17,13	79	18,64	87	18,63
46	15,74	22	16,98	89	18,34	18	18,61
91	15,68	91	16,95	69	18,33	55	18,43
71	15,46	46	16,89	71	18,32	65	18,34
51	15,44	4	16,86	63	18,01	28	18,16
45	15,38	62	16,65	78	17,99	79	18,05
60	15,37	51	16,62	70	17,92	86	18,02
82	15,30	25	16,59	51	17,91	63	18,00
55	15,30	60	16,50	18	17,85	45	17,98
25	15,23	87	16,45	83	17,80	91	17,98
28	15,18	64	16,44	86	17,80	9	17,94
78	15,13	83	16,40	62	17,76	69	17,90
21	15,12	80	16,36	50	17,72	71	17,88
87	15,11	50	16,30	33	17,56	82	17,86
86	15,07	84	16,25	53	17,55	83	17,68
52	15,00	65	16,24	9	17,54	78	17,64
73	14,96	21	16,18	75	17,53	89	17,53
72	14,96	82	16,17	98	17,52	64	17,51
97	14,90	52	16,13	19	17,52	84	17,45
65	14,87	67	16,09	87	17,48	5	17,41
19	14,85	77	16,09	59	17,46	57	17,41
20	14,80	98	15,99	66	17,46	66	17,36
70	14,66	76	15,99	85	17,34	62	17,31
63	14,58	89	15,95	28	17,24	4	17,28
14	14,54	32	15,93	57	17,23	52	17,25
58	14,50	9	15,88	91	17,22	60	17,19
32	14,48	45	15,82	84	17,21	22	17,19
64	14,43	55	15,80	73	17,21	50	17,17
29	14,38	18	15,79	60	17,21	7	17,14
69	14,37	28	15,77	67	17,11	70	17,10
67	14,36	78	15,76	48	17,08	21	17,09
47	14,34	33	15,74	55	17,08	23	17,09
50	14,32	88	15,74	45	17,06	33	17,05
62	14,32	90	15,71	27	17,04	72	17,03
22	14,27	73	15,70	72	16,99	92	16,96
79	14,24	63	15,67	77	16,97	3	16,88
31	14,19	53	15,50	82	16,97	47	16,88

84	14,13	69	15,49	46	16,92	80	16,83
90	14,09	20	15,45	32	16,89	25	16,83
83	14,00	30	15,44	6	16,85	81	16,81
33	13,95	81	15,43	92	16,73	75	16,74
7	13,86	56	15,36	65	16,71	14	16,71
89	13,83	29	15,30	13	16,64	46	16,70
77	13,82	85	15,28	81	16,58	56	16,70
57	13,77	47	15,24	7	16,55	59	16,69
16	13,76	59	15,23	56	16,52	88	16,65
54	13,73	70	15,19	47	16,51	73	16,63
75	13,72	48	15,13	25	16,47	53	16,58
48	13,68	75	15,01	21	16,42	13	16,57
59	13,66	61	14,94	80	16,36	97	16,57
56	13,55	86	14,85	90	16,20	77	16,55
81	13,53	27	14,84	97	16,16	90	16,50
23	13,50	12	14,77	61	16,14	98	16,47
76	13,50	72	14,77	52	16,02	85	16,46
98	13,43	66	14,73	24	15,84	20	16,45
2	13,37	23	14,69	3	15,81	61	16,44
88	13,31	58	14,66	16	15,81	48	16,43
66	13,29	31	14,65	64	15,76	27	16,36
24	13,28	16	14,63	15	15,70	67	16,35
85	13,19	79	14,47	20	15,68	58	16,24
92	13,15	97	14,47	14	15,61	76	16,15
61	13,13	13	14,46	23	15,52	29	16,08
15	13,09	92	14,26	4	15,43	12	16,03
12	13,03	54	14,26	5	15,38	32	15,94
49	12,98	6	14,16	88	15,36	24	15,87
30	12,93	34	14,16	74	15,31	30	15,86
74	12,88	96	14,16	26	15,28	19	15,84
4	12,69	49	14,16	31	15,23	26	15,78
27	12,58	15	14,00	30	15,07	31	15,62
53	12,58	14	13,85	29	14,91	16	15,61
3	12,28	7	13,65	76	14,41	74	15,56
13	12,28	24	13,62	54	14,21	49	14,79
26	12,27	57	13,59	49	13,92	6	14,71
9	12,19	17	13,57	1	13,72	15	14,29
34	11,48	26	13,56	2	13,44	17	14,20
5	11,41	68	13,53	17	13,06	54	13,98
37	11,18	74	13,49	12	13,03	1	13,78
6	11,16	2	13,40	68	12,89	68	13,75
96	10,95	1	13,20	58	12,83	100	13,71

68	10,95	3	13,18	96	12,69	34	12,75
100	10,89	5	12,86	34	12,65	35	12,08
38	10,57	100	12,34	10	12,53	96	11,16
17	10,44	38	11,78	8	11,56	38	10,24
8	10,16	37	11,60	100	10,93	37	10,20
1	9,14	35	10,13	35	10,48	36	9,74
36	8,69	36	8,69	38	10,36	93	9,07
35	8,41	39	8,07	36	8,28	39	9,04
93	7,29	11	7,86	93	8,26	44	8,25
44	7,08	95	7,06	99	8,06	95	7,95
99	6,61	41	7,04	37	7,63	99	7,31
41	6,32	42	6,58	44	6,07	41	7,13
43	6,13	93	6,48	42	5,98	94	7,04
42	5,85	94	6,40	11	5,85	42	6,31
95	5,70	40	6,14	41	5,44	10	6,15
39	5,52	10	5,39	43	5,32	11	5,92
40	5,28	44	4,73	94	3,74	43	5,03
94	4,75	99	4,47	39	3,69	40	3,82
10	4,65	8	4,06	95	3,48		
11	4,47	43	3,98	40	2,37		

*Época 3= (colhida em 2016); **pd**: parcela perdida