



APLICAÇÃO DE ANÁLISE DE COMPONENTES PRINCIPAIS NO AGRUPAMENTO DE GENÓTIPOS DE CANA-DE-AÇÚCAR PROVENIENTES DE SEMENTES

Luiz Gustavo da Mata Borsuk¹, Bruna Sisti Michelan de Polli², Hugo Zeni Neto³

¹ Doutorando do Programa de Pós Graduação em Agronomia, Universidade Estadual de Maringá – UEM, Maringá - PR. Bolsista CAPES. lgborsuk@hotmail.com

² Doutoranda do Programa de Pós Graduação em Genética e Melhoramento, Universidade Estadual de Maringá – UEM, Maringá - PR. Bolsista CNPq. brunadepolli@gmail.com

³ Orientador, Doutor, Professor do Departamento de Agronomia, Universidade Estadual de Maringá - UEM. hzneto@uem.br

RESUMO

Três características quantitativas foram submetidas ao método PCA com o objetivo de avaliar a variabilidade genética existente entre quatorze genótipos de cana-de-açúcar. Baseado em tais informações é possível contribuir para que a seleção de genitores que visam o desenvolvimento de cultivares superiores seja feita de forma planejada, além de propiciar um aumento da base genética da cultura. Dois componentes principais foram capazes de explicar cerca de 75,72% da dissimilaridade genética que compõe as variedades utilizadas. A formação de três grupos contendo sete, cinco e dois genótipos indicam os indivíduos mais parecidos geneticamente. Isso significa que, a chance de ter um aumento na variabilidade genética gerada e consequentemente o sucesso dos programas de melhoramento genético de cana-de-açúcar.

PALAVRAS-CHAVE: Divergência Genética; Melhoramento Genético; Saccharum; Variabilidade.

1 INTRODUÇÃO

Os programas de melhoramento genético de cana-de-açúcar (*Saccharum ssp.*) têm como objetivo principal fornecer a melhor combinação de caracteres qualitativos e quantitativos afim de promover o desenvolvimento de cultivares geneticamente superiores (EMBRAPA, 2015). No Brasil, o Centro de Tecnologia Copersucar (CTC), o Instituto Agrônomo de Campinas (IAC) e a Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro (RIDESA) são responsáveis pelo desenvolvimento de novas tecnologias e seleção destas variedades que dependem, principalmente, da escolha de parentais adequados para aumentar a base genética da cultura (CARVALHO; FURTADO, 2013).

O estabelecimento de técnicas para a seleção de genitores é baseado no cruzamento entre parentais produtivos, complementares e geneticamente divergentes para se obter maior heterose e variabilidade genética nas gerações segregantes, o que resulta no aumento da probabilidade de encontrar indivíduos superiores dentro de uma população (BADALOO et al., 1998).

O processo de seleção de genitores é contínuo e iterativo, portanto, baseado nas características fenotípicas apresentadas pelas gerações subsequentes é possível planejar futuros cruzamentos que visem a incorporação de características de interesse à agroindústria canavieira, como aumento da produtividade e a qualidade do caldo. Esta etapa é altamente afetada pela qualidade da avaliação das características fenotípicas que são facilmente mascaradas por efeitos ambientais (DUTRA FILHO et al., 2011).

A constante necessidade de pesquisas que valorizem o conhecimento das constituições e divergências genéticas entre as variedades de cana-de-açúcar motivou a realização deste estudo. O objetivo principal foi analisar dados biométricos para verificar a variabilidade genética existente entre genótipos de cana-de-açúcar. Através dessa análise, buscamos contribuir com a seleção de genitores amplamente divergentes, que possam fornecer combinações genéticas favoráveis e, assim, contribuir para o aumento da base genética da cultura. Essa abordagem é fundamental para o avanço e aprimoramento da



cana-de-açúcar, bem como para o desenvolvimento de variedades mais produtivas, resistentes a doenças e adaptadas a diferentes condições ambientais.

2 MATERIAIS E MÉTODOS

O Os dados biométricos foram coletados a partir de quatorze genótipos de cana-de-açúcar (F01M14, F02M31, F02M35, F09M15, F09M20, F09M45, F13M11, F22M18, F27M09, F28M29, F29M28, F33M34, F35M02 e F42M41) propagados por sementes e oriundos de cruzamentos biparentais. As plantas estão localizadas no Centro Técnico de Irrigação (CTI), órgão vinculado ao Departamento de Agronomia (DAG) da Universidade Estadual de Maringá (UEM), situado na cidade de Maringá – PR (coordenadas 23°25'57" S, 51°57'08" W e 542 m).

O preparo de mudas foi realizado em bandejas de polietilenocontendo cerca de 30g de cariopses. Após as mudas alcançarem duas a três folhas verdadeiras foram transplantadas para o campo. As variáveis analisadas foram Diâmetro Médio dos Colmos (DC) por meio de paquímetro digital, o Grau BRIX (BRIX^o) por meio de um refratômetro e a produção de toneladas de cana por hectare (TCH) referente ao 12^o mês após o plantio.

Os dados coletados foram primeiramente submetidos ao teste de Shapiro-Wilk para normalidade dos dados. Confirmada a distribuição normal, foi realizada a análise de componentes principais, com os dados padronizados, com base nas distâncias euclidianas, para formação dos grupos foi utilizado o agrupamento K-means, pelo algoritmo de Hartigan-Wong (HARTIGAN; WONG, 1979), onde variação total dentro do cluster é definida como a soma das distâncias euclidianas quadráticas entre os itens e o centroide correspondente. Para inferir o número de grupos foram implementados um número de 2 a 14 grupos possíveis, com uma análise de reamostragem de bootstrap de 1000 iterações. Todas as análises foram realizadas pelo software R 4.3.1 (R CORE TEAM, 2023), com auxílio do pacote factoextra 1.0.7 (KASSAMBARA & MUNDT, 2017).

3 RESULTADOS E DISCUSSÕES

Após a confirmação da distribuição normal dos dados, foi possível verificar que dois componentes principais foram capazes de explicar cerca de 75,72% da variabilidade total presente na amostra (Tabela 1). Pode-se considerar que o método de análise multivariada (PCA) se mostrou eficiente para escolher a forma mais representativa do conjunto de dados uma vez que explica grande parte da variabilidade dos dados, sendo satisfatório para o estudo de diversidade genética (>50%) (OLIVEIRA, 2019).

Tabela 1: Autovalores correspondentes às percentagens de variação explicadas pelas componentes principais associados aos três caracteres avaliados: diâmetro de colmo, graus BRIX e tonelada de cana por hectare.

| Variáveis | Componentes Principais | | |
|--------------------------------|------------------------|-------------|------------|
| | CP1 | CP2 | CP3 |
| Diâmetro | 0.4882428 | -0.75396790 | 0.4394899 |
| ^o Brix | -0.6839682 | -0.01780582 | 0.7292945 |
| TCH | -0.5420392 | -0.65666990 | -0.5243836 |
| Variância | 1.3087516 | 0.9627920 | 0.7284563 |
| Variância Acumulada (%) | 43.6251 | 75.7181 | 100 |



Considerando a finalidade do método PCA em reduzir e eliminar sobreposições, observou-se que número de componentes principais se igualou ao número de variáveis analisadas no estudo. Este resultado já era esperado tendo em vista o número de variáveis analisadas, porém, mesmo nestas condições, a metodologia utilizada foi útil para resumir e explorar a estrutura dos dados bidimensionais, além de reduzir o ruído deles.

Diante do objetivo proposto pelo estudo, a dispersão gráfica dos resultados de PCA foi utilizada para avaliar a similaridade entre os genótipos de cana-de-açúcar avaliados. Os três grupos formados são compostos pelos genótipos mais similares entre si, ou seja, o cruzamento entre indivíduos de um mesmo grupo provavelmente não resultaria em genótipos superiores devido à sua semelhança genética considerando as variáveis Diâmetro, TCH e °BRIX (Figura 1).

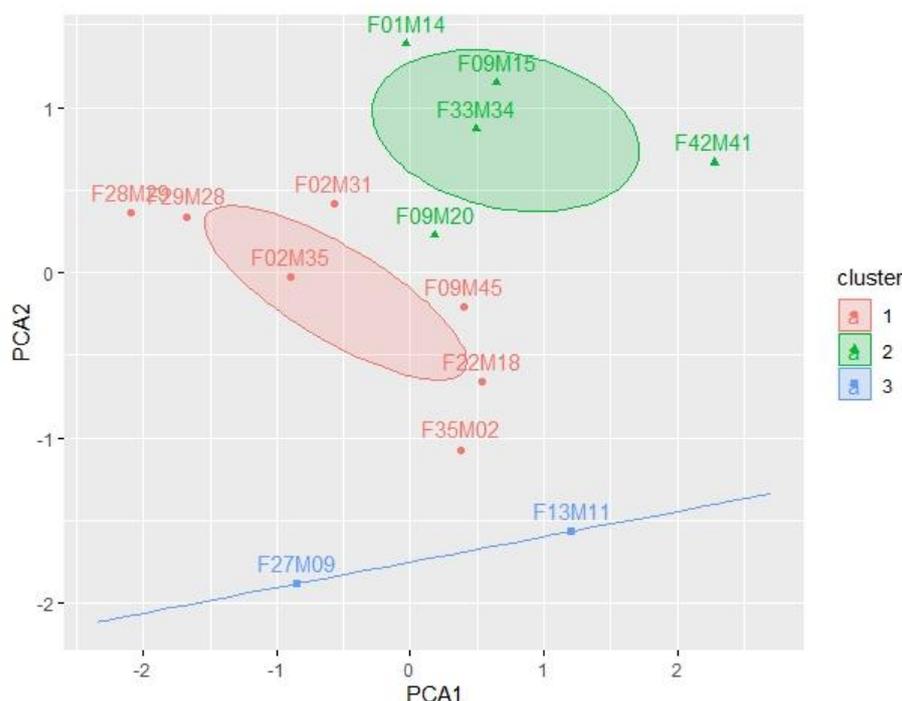


Figura 1 - Dispersão gráfica dos 14 genótipos de cana-de-açúcar em relação aos dois primeiros componentes principais, CP1 e CP2.

O cruzamento entre indivíduos de cana-de-açúcar é uma etapa crucial no processo de melhoramento genético, mas sua realização deve estar alinhada com as metas e objetivos dos melhoristas. É de extrema importância que o planejamento da escolha dos parentais inclua a consideração para evitar o pareamento entre indivíduos aparentados, a fim de prevenir a ocorrência de depressão endogâmica e o estreitamento da base genética (MATSUOKA et al., 2005). Para alcançar esse objetivo, é fundamental possuir conhecimento detalhado sobre a divergência genética entre os genótipos. Através dessa compreensão, os melhoristas podem fazer previsões mais precisas sobre o comportamento da população futura (SRIVASTAVA et al., 1999). Então se recomenda o cruzamento de famílias entre os grupos para se obter uma maior variabilidade no processo de melhoramento.

4 CONSIDERAÇÕES FINAIS

A análise da divergência genética entre os genótipos de cana-de-açúcar por meio do método de Componentes Principais (PCA) oferece uma base sólida para a tomada de



decisões informadas e estratégias de melhoramento genético mais eficientes. Ao evitar cruzamentos entre indivíduos estreitamente relacionados os melhoristas podem maximizar o processo de obtenção de uma maior variabilidade genética e conseqüentemente o um maior sucesso dos programas de melhoramento de cana-de-açúcar.

REFERÊNCIAS

BADALOO, G. H.; DOMAINGUE, R.; RAMDOYAL, K. A critical review of parental choice and cross prediction techniques in the Msiri sugar cane breeding programme. **Food and Agricultural Research Council**, Reduit, Mauritius. 1998.

CARVALHO, S. A. D. de.; FURTADO, A T. O melhoramento genético de cana-de-açúcar no Brasil e o desafio das mudanças climáticas globais. **Revista Gestão & Conexões**. Vitória, v. 2, n. 1, 2013.

DUTRA FILHO, J. D. A.; MELO, L. J. O. T. D.; RESENDE, L. V.; ANUNCIÇÃO FILHO, C. J. D.; BASTOS, G. Q. Aplicação de técnicas multivariadas no estudo da divergência genética em cana-de-açúcar. **Revista Ciência Agronômica**, v. 6, n. 3, p. 432-439, jul-set. 2011.

HARTIGAN, J. A.; WONG, M. A. Algorithm AS 136: A k-means clustering algorithm. **Journal of the royal statistical society. series c (applied statistics)**, v. 28, n. 1, p. 100-108, 1979.

KASSAMBARA, A.; MUNDT, F. Factoextra: extract and visualize the results of multivariate data analyses. **R Package Version**, v. 76, n. 2, 2017.

MATSUOKA, S.; GARCIA, A. A. F.; ARIZONO, H. Melhoramento da cana-de-açúcar. In: BORÉM, A. (Ed.). **Melhoramento de espécies cultivadas**. 2.ed. Viçosa, MG: UFV, 2005.

MORAIS, L. K.; CURSI, D. E.; SANTOS, J. M.; SAMPAIO, M.; CÂMARA, T. M. M.; ALBUQUERQUE E SILVA, P.; BARBOSA, G. V.; HOFFMANN, H. P. CHAPOLA, R. G.; FERNANDES JÚNIOR, A. R.; GAZAFFI, R. Melhoramento Genético da Cana-de-Açúcar. Aracajú: **Embrapa Tabuleiros Costeiros**, 2015. 38p.

OLIVEIRA, B. **Análise de Componentes Principais**. Disponível em: <https://site.statplace.com.br/blog/analise-de-componentes-principais/>. Acesso em: 04 ago. 2023.

R Core Team (2023). R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, Vienna, Austria. Disponível em: <https://www.R-project.org/>.

SRIVASTAVA, H. M.; SRIVASTAVA, S.; KUMAR, R.; MISRA, G. P. Genetic Divergence among Interspecific Hybrids of Sugarcane. **Sugar Tech**, v. 1, 1999.