

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**ANÁLISE MULTIVARIADA PARA CARACTERIZAÇÃO E
DIVERGÊNCIA DE GENÓTIPOS E CORRELAÇÃO ENTRE
CARACTERES EM MILHO**

**Sophia Mangussi Franchi Dutra
Engenheira Agrônoma**

2018

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**ANÁLISE MULTIVARIADA PARA CARACTERIZAÇÃO E
DIVERGÊNCIA DE GENÓTIPOS E CORRELAÇÃO ENTRE
CARACTERES EM MILHO**

Sophia Mangussi Franchi Dutra

Orientador: Prof. Dr. Gustavo Vitti Môro

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

2018

D978a Dutra, Sophia Mangussi Franchi
Análise multivariada para caracterização e divergência de
genótipos e correlação entre caracteres em milho / Sophia Mangussi
Franchi Dutra. -- Jaboticabal, 2018
ix, 48 p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de
Ciências Agrárias e Veterinárias, 2018

Orientador: Gustavo Vitti Mõro

Banca examinadora: Sandra Helena Uneda Trevisoli, Rinaldo
Cesar de Paula, Maria Elisa Ayres Guidetti Paterniani, João Antônio
da Costa Andrade

Bibliografia

1. Melhoramento de plantas. 2. Agrupamento. 3. Correlação. I.
Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.52:633.15



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

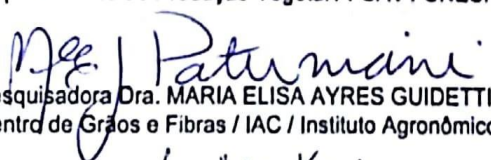
TÍTULO DA TESE: ANÁLISE MULTIVARIADA PARA CARACTERIZAÇÃO E DIVERGÊNCIA DE GENÓTIPOS E CORRELAÇÃO ENTRE CARACTERES EM MILHO

AUTORA: SOPHIA MANGUSSI FRANCHI DUTRA

ORIENTADOR: GUSTAVO VITTI MÔRO


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


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


Pesquisadora Dra. MARIA ELISA AYRES GUIDETTI ZAGATTO PATERNIANI
Centro de Grãos e Fibras / IAC / Instituto Agronômico de Campinas


Profa. Dra. SANDRA HELENA UNÊDA TREVISOLI
Departamento de Produção Vegetal / FCAV / UNESP - Jaboticabal


Prof. Dr. JOÃO ANTONIO DA COSTA ANDRADE
Departamento de Biologia e Zootécnica / Faculdade de Engenharia de Ilha Solteira


Prof. Dr. RINALDO CESAR DE PAULA
Departamento de Produção Vegetal / FCAV / UNESP - Jaboticabal

Jaboticabal, 22 de março de 2018

DADOS CURRICULARES DO AUTOR

SOPHIA MANGUSSI FRANCHI DUTRA - filha de Eloisa Mangussi Franchi Dutra e Paulo Roberto Franchi Dutra e irmã de Beatriz Mangussi Franchi Dutra, nascida aos 21 de janeiro de 1988, natural de Jundiaí, interior do estado de São Paulo, Brasil. Coursou o ensino fundamental e médio na Escola Particular “Universitário”. Em fevereiro de 2007 ingressou no curso de Agronomia na UFLA – Universidade Federal de Lavras. Foi integrante do DCE – Diretório Central dos Estudantes e CAA – Centro Acadêmico de Agronomia, participando da organização de diversos eventos. Estagiou nas áreas de Microbiologia Vegetal e Sementes. Foi integrante do NESEM – Núcleo de Estudos em Sementes, participando da organização de cursos e eventos. Obteve o título de Engenheira Agrônoma em janeiro de 2012. Ingressou no curso de Pós-graduação em Agronomia (Fitotecnia/Produção e Tecnologia de Sementes), nível de Mestrado, em agosto de 2012, pela UFLA – Universidade Federal de Lavras, sendo bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), tendo como orientadora a Professora Dra. Édila Vilela de Resende Von Pinho, obtendo o título em julho de 2014, trabalhando com expressão de genes relacionados à tolerância a altas temperaturas durante a germinação de sementes de milho. Ingressou no curso de Pós-graduação em Agronomia (Genética e Melhoramento de Plantas), nível de Doutorado, em agosto de 2014, pela Universidade Estadual Paulista UNESP – Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal (FCAV), sendo bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), tendo com orientador o Professor Dr. Gustavo Vitti Môro. É integrante do NEGEMM – Núcleo de Estudos em Genética e Melhoramento de Milho. No ano de 2017, realizou intercâmbio no departamento de Crop Science da University of Illinois, por um período de quatro meses.

A Deus, por iluminar meu caminho durante toda minha trajetória e por me permitir ter fé
em todos os momentos.

Aos meus pais, minha gratidão e amor eterno, pela confiança e dedicação.

DEDICO

AGRADECIMENTOS

À Universidade Estadual Paulista – Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal – FCAV/UNESP, pelos ensinamentos oferecidos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão de bolsa de estudos de doutorado.

Ao Professor Dr. Gustavo Vitti Môro, por toda atenção, ensinamento e competência.

Aos professores do Departamento de Sementes/UFLA, em especial à Professora Édila Vilela de Resende Von Pinho, por terem me apoiado a seguir o caminho em que estou hoje.

Aos membros da Banca Examinadora Profa. Dra. Sandra Helena Uneda Trevisoli, Dr. Rinaldo Cesar de Paula, Pesquisadora Dra. Maria Elisa Ayres Guidetti Paterniani e Prof. Dr. João Antônio da Costa Andrade, pelas valiosas contribuições para melhoria do artigo científico e da tese.

Aos docentes do curso de Pós-Graduação em Agronomia (Genética e Melhoramento de Plantas) pela importante contribuição em meu crescimento, tanto como ser humano quanto como pesquisadora.

Aos funcionários da Fazenda de Ensino, Pesquisa e Extensão - FEPE da FCAV, Campus de Jaboticabal, pelo apoio na condução dos trabalhos de campo e pelo companheirismo.

Aos funcionários do Departamento de Produção Vegetal pelo apoio, ensinamento e amizade.

Aos meus amigos do Núcleo de Estudos em Genética e Melhoramento de Milho - NEGEMM, pelas risadas nos momentos de tensão, pelos desabafos na “salinha” e pelo auxílio nas horas de dificuldade.

À gloriosa República Poucas & Boas, pelo companheirismo e irmandade desde 2007. As maiores e melhores amigas, Cristina, Rebeca, Elise e Rayana, pelo companheirismo, apoio e parceria de sempre. Agradeço a Lele e o Antônio pela amizade.

À Família Mangussi, em especial a Tia Ivone, Tia Maria e Tia Lourdes, pelo apoio e carinho.

SUMÁRIO

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS	1
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	2
2.1 Grupos heteróticos	2
2.2 Análise multivariada	3
2.2.1 Análise de agrupamento	4
2.2.1.1 Métodos de agrupamento	5
2.3 Análise de correlação	7
2.3.1 Análise de correlação linear simples (coeficiente de pearson)	8
2.3.2 Análise bayesiana	9
3 REFERENCIAL BIBLIOGRÁFICO.....	10
CAPÍTULO 2 - ABORDAGEM MULTIVARIADA NA ALOCAÇÃO DE GENÓTIPOS DE MILHO EM GRUPOS HETERÓTIPOS.....	14
INTRODUÇÃO	15
MATERIAL E MÉTODOS.....	177
RESULTADOS E DISCUSSÃO.....	199
CONCLUSÕES	344
REFERÊNCIAS.....	355
CAPÍTULO 3 - CORRELAÇÃO ENTRE CARACTERES DE GENÓTIPOS DE MILHO..	39
INTRODUÇÃO	40
MATERIAL E MÉTODOS.....	41
RESULTADOS E DISCUSSÃO.....	42
CONCLUSÕES	45
REFERÊNCIAS.....	45
ANEXO.....	477

ANÁLISE MULTIVARIADA PARA CARACTERIZAÇÃO E DIVERGÊNCIA DE GENÓTIPOS E CORRELAÇÃO ENTRE CARACTERES EM MILHO

RESUMO- Os objetivos principais com a pesquisa foram: alocar genótipos de milho em grupos divergentes pela abordagem multivariada e correlacionar caracteres agronômicos com caracteres relacionados à qualidade de sementes de milho. Para alocação em grupos heteróticos, foram utilizados dois métodos multivariados, método *K-means* e método Tocher. A partir da avaliação de caracteres agronômicos (altura de planta, posição relativa da espiga, acamamento, quebramento, prolificidade e produtividade de grãos), foram alocadas 229 linhagens parcialmente endogâmicas (S_3), avaliadas em três safras. Já para correlação entre caracteres foram utilizados os métodos Redes Bayesiana e correlação linear simples de Pearson. Os caracteres avaliados, de 31 genótipos de milho, foram caracteres agronômicos (altura de planta, posição relativa da espiga, acamamento, quebramento, prolificidade e produtividade de grãos), avaliados em três safras, e caracteres de sementes (primeira contagem de germinação, germinação final, primeira contagem de emergência, emergência final e índice de velocidade de germinação). As análises estatísticas foram realizadas utilizando-se o programa computacional SAS, Selegen e RStudio. Na análise de variância conjunta foi constatado efeito significativo para safra em todas as características analisadas, caracterizando a influência das safras e indicando diferença entre os ambientes de avaliação. Foi possível alocar os genótipos de milho em grupos heteróticos a partir dos métodos de agrupamento. A correlação linear de Pearson variou de -0,511 (entre quebramento e produtividade) a 0,987 (entre primeira contagem de emergência e emergência final). A partir do *Directed Acyclic Graph* é possível observar a correlação entre caracteres agronômicos e caracteres relacionados à qualidade de sementes de milho. Assim, os genótipos de milho foram alocados quanto à divergência a partir das análises multivariadas de agrupamento K-médias e Tocher. Foi possível identificar sete grupos divergentes pelo método K-médias e 21 pelo método Tocher. E também, observa-se pouca correlação existente entre os caracteres agronômicos e os caracteres relacionados à qualidade de sementes de milho.

Palavras-chave - agrupamento, correlação, melhoramento de plantas

MULTIVARIATE ANALYSIS FOR CHARACTERIZATION AND DIVERGENCE OF GENOTYPES AND CORRELATION IN MAIZE CHARACTERS

ABSTRACT - The main objectives of the research were: to allocate corn genotypes in divergent groups by the multivariate approach and to correlate agronomic traits with traits related to maize seed quality. For allocation in heterotic groups, two multivariate methods, K-means method and Tocher method were used. Based on the evaluation of agronomic characters (plant height, relative position of the spike, lodging, breaking, prolificacy and grain yield), 229 partially inbred lines (S3) were evaluated, evaluated in three harvests. Already for the correlation between characters were used the methods Bayesian Networks and simple linear correlation of Pearson. The evaluated traits of 31 corn genotypes were agronomic characteristics (plant height, relative position of the spike, lodging, breaking, prolificacy and grain yield) evaluated in three harvests, and seed characteristics (first germination count, germination endpoint, first emergency count, final emergency and germination speed index). Statistical analyzes were performed using the SAS software, Selegen and RStudio. In the analysis of joint variance it was verified a significant effect for harvest in all characters analyzed, characterizing the influence of the crops and indicating difference between the evaluation environments. It was possible to allocate the maize genotypes to heterotic groups from the clustering methods. Pearson's linear correlation ranged from -0.511 (between breakage and productivity) to 0.987 (between first emergency count and final emergency). From the Directed Acyclic Graph it is possible to observe the correlation between agronomic characters and characters related to the quality of maize seeds. Thus, maize genotypes were allocated for divergence from the multivariate K-medium and Tocher cluster analyzes. It was possible to identify seven divergent groups by the K-means method and 21 by the Tocher method. Also, there is little correlation between agronomic characters and traits related to maize seed quality.

Key words – correlation, clustering, plant breeding.

CAPÍTULO 1 – Considerações gerais

1 INTRODUÇÃO

O Brasil destaca-se como um dos maiores produtores de milho (*Zea mays*) do mundo, sendo superado apenas pelos Estados Unidos e a China (FAOSTAT, 2017). O milho, por sua vez, destaca-se economicamente como importante cultura, por ser usado como ingrediente em uma inúmeras lista de produtos manufaturados e empregado na nutrição da população mundial (HALLAUER et al., 2010). Devido a isso, a demanda por genótipos de qualidade é alta, dependendo cada vez mais, de programas de melhoramento de milho para suprir esta necessidade.

O sucesso de programas de melhoramento, depende, em grande parte, da escolha adequada de genitores, os quais originarão as populações segregantes. Com isso, utilizar genitores com alta frequência de alelos favoráveis, aumenta a probabilidade de se extrair genótipos promissores para cruzamentos favoráveis.

Assim, o ponto de partida para garantir o sucesso de programas de melhoramento é a variabilidade genética, que é maximizada pelo cruzamento de diferentes genótipos (ERTIRO et al., 2013).

A variabilidade existente entre os genótipos é uma necessidade para obter ganhos de seleção nos cruzamentos de grupos geneticamente divergentes, que apresentam atributos de interesse (CRUZ et al., 2014).

Quanto mais divergentes forem os genitores, maior é a variabilidade resultante na população segregante e maior é a probabilidade de reagrupar alelos de interesse em novas combinações favoráveis (BARBIERI et al., 2005), sendo de suma importância também a complementariedade entre os genitores para obtenção de genótipos superiores.

Esses genitores são facilmente identificados por meio de métodos para obtenção de grupos de genótipos com características semelhantes e divergentes, dentro e entre grupos, respectivamente.

Além da variabilidade, outro aspecto importante nos programas de melhoramento é a relação entre caracteres, pois possibilita a obtenção de ganhos para caracteres de

interesse por meio da manipulação de outras características correlacionadas e de fácil observação e mensuração.

De acordo com Lopes et al. (2007), o conhecimento das correlações entre componentes da produção, em especial a produtividade de grãos, possibilita a seleção indireta de um caráter quantitativa, de difícil ganho de seleção, por meio da seleção de outro caráter diretamente a ela correlacionada de maior ganho genético ou de fácil seleção visual.

Diante do exposto, o objetivo do presente trabalho consisti em alocar genótipos de milho em grupos divergentes pela abordagem multivariada e correlacionar caracteres de campo com caracteres de laboratório de sementes de milho.

2 REVISÃO DE LITERATURA

2.1 GRUPOS HETERÓTICOS

O estudo da diversidade genética é importante dentro de um programa de melhoramento genético de milho para seleção de genitores. O conhecimento da divergência genética fornece parâmetros para a correta escolha de genitores que, ao serem cruzados, maximizam as chances de obtenção de genótipos superiores em gerações segregantes (ROTILI et al., 2012). Quanto mais divergentes forem os genitores, maior é a variabilidade resultante na população segregante, e maior é a probabilidade de reagrupar alelos em novas combinações favoráveis (BARBIERI et al., 2005). Sendo a variabilidade, existente entre os genótipos, uma estratégia para obter ganhos de seleção nos cruzamentos de grupos geneticamente divergentes, que apresentam atributos de interesse (CRUZ et al., 2014).

Grupo heterótico pode ser definido como uma coleção de germoplasma que, quando cruzado com germoplasma externo ao seu grupo, tende a exibir, um maior grau de heterose, que quando cruzado com membros do seu próprio grupo (LEE, 1995).

A análise de agrupamento é uma técnica estatística multivariada que permite resolver o problema de, dado um conjunto de n unidades amostrais (indivíduos, tratamentos, genótipos, etc.), os quais são medidos segundo p caracteres

(características), obtendo um esquema que possibilite reunir as unidades em um número de grupos de tal forma que exista homogeneidade dentro e heterogeneidade entre os grupos (REGAZZI, 2006). Para se obter os grupos, utiliza-se a matriz de similaridade ou dissimilaridade entre os elementos observados, como uma estimativa da distância entre esses elementos.

Segundo Melani e Carena (2005), a precisão na seleção dos genitores e na identificação dos grupos heteróticos, corresponde ao sucesso do desenvolvimento de linhagens e cultivares superiores de milho.

2.2 ANÁLISE MULTIVARIADA

A estatística multivariada corresponde a um conjunto de métodos e técnicas estatísticas que utilizam, simultaneamente, todas as informações das caracteres, na interpretação teórica do conjunto de dados obtidos, levando em consideração as correlações existentes entre as mesmas (HAIR et al., 2009).

Os métodos de análise de dados multivariados permitem um estudo global das caracteres avaliadas, evidenciando as ligações, semelhanças ou diferenças entre elas, com menor perda possível de informação (HAIR et al., 2009).

A análise multivariada pode ser dividida em dois grupos. O primeiro, consiste em técnicas exploratórias de sintetização da estrutura de variabilidade dos dados. Por sua vez, o segundo, consiste em técnicas de inferência estatística. Destacam-se no primeiro grupo, métodos como a análise de componentes principais, análise fatorial, análise de correlações canônicas, análise de agrupamento, análise discriminante e análise de correspondência. No segundo grupo, métodos de estimação de parâmetros, testes de hipóteses, análise de variância, covariância e regressão multivariada (MINGOTI, 2007).

Nesta análise as caracteres utilizadas devem ser aleatórias e inter-relacionadas para que seus diferentes efeitos sejam interpretados conjuntamente. Deste modo, as metodologias multivariadas classificam as caracteres em dependentes e independentes (MANLY, 2008). Em estudos de dependência o caráter identificada como dependente é explicada por outras caracteres, conhecidas como caracteres independentes, utilizando-

se métodos multivariados, como análise discriminante, análise conjunta, análise de correlação canônica, análise multivariada de variância e regressão múltipla (HAIR et al., 2009). Em contrapartida, estudos de interdependência envolvem uma análise simultânea de todas as caracteres no conjunto, sem definir se o caráter é independente ou dependente. Estes estudos são feitos pela análise de agrupamentos, análise de componentes principais, análises de correspondência e análise fatorial (HAIR et al., 2009).

Nesse contexto, as técnicas multivariadas podem ser de grande importância no estudo da diversidade genética e na seleção de genótipos superiores nas gerações segregantes por serem capazes de unir informações de um conjunto de diversos caracteres de interesse (CRUZ et al., 2014).

A utilização de técnicas multivariadas para estimar a divergência genética tem sido empregada em vários trabalhos e em diversas culturas, tais como: arroz (BENITEZ et al., 2011); cana-de-açúcar (DUTRA FILHO, 2011; milho (COIMBRA et al., 2010); girassol (VOGT; BALBINOT JUNIOR; SOUZA, 2010); eucalipto (CASTRO et al., 2013), soja (PELUZIO et al., 2009; SANTOS et al., 2011; ALMEIDA; PELUZIO; AFFÉRI, 2011 e SANTOS et al., 2013), entre outras.

2.2.1 ANÁLISE DE AGRUPAMENTO

A análise de agrupamento ou *cluster analysis* descreve diversos métodos numéricos, cujo objetivo é classificar as observações representadas em um espaço p-dimensional em um número conveniente de grupos a partir de um critério de similaridade (LANDIM, 2011).

Jain (2010) descreve a análise de agrupamento como o estudo formal dos métodos e algoritmos para agrupamento ou particionamento do conjunto de dados, segundo um atributo, padrão, característica ou medida.

Os métodos de agrupamento têm por finalidade separar um grupo original de observações em vários subgrupos, de forma a obter homogeneidade dentro do grupo e heterogeneidade entre os subgrupos. Este método depende de medidas de

dissimilaridade provenientes de caracteres quantitativas e qualitativas (CROSSA; FRANCO, 2004; ROCHA, 2010).

Estes métodos diferem pelo tipo de resultado a ser fornecido e pelas diferentes formas de definir a proximidade entre os indivíduos diante da formação de um grupo ou entre dois grupos (CRUZ et al., 2014). Dentre os métodos destacam-se os métodos hierárquicos, não-hierárquicos e de otimização.

Nos métodos hierárquicos, os genótipos são agrupados por um processo que se repete em vários níveis, até que seja estabelecido o dendrograma. Neste caso, o maior interesse está no dendrograma e suas ramificações e não no número ótimo de grupos. As delimitações podem ser estabelecidas por um exame visual, em que se avaliam pontos de alta mudança de nível, tomando-os em geral como delimitadores do número de genótipos para determinado grupo (CRUZ et al., 2014). O método não-hierárquico difere do hierárquico na quantidade de classes formadas, a qual é determinada antes da realização da análise, com o objetivo de minimizar a variabilidade dentro do conjunto e maximizar a variabilidade entre os conjuntos (PITTA et al., 2010). O método de agrupamento por otimização constitui em um método de agrupamento simultâneo, o qual se realiza a separação dos indivíduos de uma só vez (CRUZ et al., 2014).

2.2.1.1 MÉTODOS DE AGRUPAMENTO

- **K- MÉDIAS (*K-MEANS*)**

O método de agrupamento K-médias foi proposto pela primeira vez na década de 1950. Este método foi desenvolvido por vários pesquisadores de forma independente em diversas áreas da ciência. Jain (2010) cita que os pesquisadores responsáveis pelo K-médias foram Steinhaus (1956), Ball e Hall (1965), MacQueen (1967) e Loyd (1982).

O K-médias é um método não-hierárquico, bastante utilizado, que procura alocar os elementos de dados em grupos previamente definidos. Este método minimiza a soma dos quadrados residuais dentro de cada grupo formado, aumentando a homogeneidade dentro dele ao mesmo tempo em que aumenta a diferença entre eles (ALENCAR et al., 2013).

Na análise de agrupamentos pelo método K-médias utiliza-se um número de grupos/clusters previamente determinado para o cálculo dos pontos que representam os “centros” destes clusters.

O centróide pode ser a média ou a mediana de um grupo de pontos, sendo como objetivo do K-médias, minimizar a distância entre cada ponto e o seu respectivo centróide (HAIR et al., 2009).

Segundo Mingoti (2007), basicamente, cada elemento amostral é alocado àquele cluster cujo centróide (vetor das médias amostral) é o mais próximo vetor de valores observados para o respectivo elemento. Originalmente, o método foi composto por quatro passos: (1) primeiramente escolhe-se k centróides, chamados de “sementes” para inicializar o processo de partição; (2) cada elemento do conjunto de dados é, então, comparado com cada centróide inicial, por meio de uma medida de distância que, em geral, é a distância Euclidiana. O elemento é alocado ao grupo cuja distância é menor; (3) depois de aplicar o passo 2 para cada um dos n elementos amostrais, recalcula-se os valores dos centróides para cada novo grupo formado, e repete-se o passo 2, considerando os centróides desses novos grupos; (4) os passos 2 e 3 devem ser repetidos até que todos os elementos amostrais estejam bem “alocados” em seus grupos, isto é, até que nenhuma realocação de elementos seja necessária.

- **TOCHER**

O método de agrupamento por otimização ou método de Tocher, pela de distância generalizada de Mahalanobis (D^2), como medida de dissimilaridade, constitui um método de agrupamento simultâneo, que se realiza a separação dos indivíduos de uma só vez (CRUZ et al., 2003).

Este método requer a obtenção da matriz de dissimilaridade, sobre a qual é identificado o par de indivíduos mais similares. Esses indivíduos formarão o grupo inicial. A partir daí é avaliada a possibilidade de inclusão de novos indivíduos, adotando-se o critério de que a distância média intragrupo deve ser menor que a distância média intergrupo (CRUZ et al., 2011).

Neste método, a entrada de um indivíduo em um grupo sempre aumenta o valor médio da distância dentro do grupo. A inclusão, ou não, do indivíduo k no grupo é, então, feita considerando:

Se $d_{(\text{grupo})k}/n \leq \Theta$, inclui-se o indivíduo k no grupo;

Se $d_{(\text{grupo})k}/n > \Theta$, o indivíduo k não é incluído no grupo.

Sendo n o número de indivíduos que constitui o grupo original e Θ é a maior dentre as menores distâncias envolvendo cada acesso (critério de agrupamento).

Nesse caso, a distância entre o indivíduo k e o grupo formado pelos indivíduos ij é dada por:

$$d_{(ij)k} = d_{ij} + d_{jk}.$$

Este método apresenta uma inconveniência no agrupamento dos genótipos com maior dissimilaridade: na maioria dos casos, cada genótipo forma um grupo específico (um grupo de apenas um genótipo) em virtude de este agrupamento ser influenciado pela distância dos genótipos já agrupados (VASCONCELOS et al., 2007).

2.3 ANÁLISE DE CORRELAÇÃO

Em situações experimentais e não-experimentais é comum haver uma grande quantidade de informações, na forma de caracteres, que apresentam alguma relação entre si. Com o intuito de melhor compreender estas caracteres, busca-se a expressão da relação funcional entre elas. De modo particular, busca-se uma função matemática que indique como as caracteres estão inter-relacionadas. Adicionalmente, pode ser de interesse indicar, por meio de uma medida quantitativa, o grau de relacionamento entre as caracteres. A técnica de análise usada para o estabelecimento do grau de associação entre as caracteres é chamada análise de correlação (OSTLE; MALONE, 1988).

A análise de correlação tem como objetivo a avaliação do grau de associação entre duas caracteres, X e Y , ou seja, mede a intensidade e a direção da relação linear ou não-linear entre caracteres (CHARNET et al., 2008).

Não existe a distinção entre o caráter independente (X) e o caráter dependente (Y), ou seja, o grau de variação conjunta entre X e Y é igual ao grau de variação conjunta

entre Y e X. Neste sentido, o conhecimento do coeficiente de correlação é importante pois possibilita ao melhorista saber como a seleção para um carácter está relacionado a expressão de outros caracteres (FREIRE FILHO, 1988).

A percepção e a interpretação da correlação entre caracteres são as mais importantes ferramentas para o planejamento de ensaios experimentais e para a resolução de modelos matemáticos. A correlação descreve a associação entre caracteres, não definindo uma relação de causa e efeito (DORIA FILHO, 1999).

2.3.1 ANÁLISE DE CORRELAÇÃO LINEAR SIMPLES (COEFICIENTE DE PEARSON)

Em pesquisas, é crescente a importância de verificar se existe relação entre duas ou mais caracteres, isto é, saber se as alterações sofridas por uma das caracteres são acompanhadas por alterações nas outras.

O termo correlação significa relação em dois sentidos e é usado em estatística para designar a força que mantém unidos dois conjuntos de valores. O estudo da correlação verifica a existência e o grau de relação entre as caracteres.

Entende-se por correlação estatística, o grau de associação entre caracteres aleatórias. A dependência de duas caracteres X e Y é dada pelo coeficiente de correlação amostral, conhecido também por coeficiente r-de-Pearson ou coeficiente de correlação.

O coeficiente de correlação linear de Pearson mede a força e a direção da correlação (positiva ou negativa) entre duas caracteres aleatórias X e Y. O seu valor situa-se no intervalo entre -1 e 1. Quanto mais o valor absoluto do coeficiente de correlação se aproxima de 1 mais forte é a correlação entre as caracteres.

Uma vez caracterizada esta relação, procura-se descrevê-la sob forma matemática, pela função:

$$r_{XY} = \text{Cov}(X, Y) / \sqrt{V(X) \cdot V(Y)} = \sigma_{X,Y} / \sigma_X \cdot \sigma_Y$$

Em que: $\sigma_{X,Y}$ é a covariância entre as caracteres aleatórias X e Y; σ_X é o desvio padrão do carácter aleatória X; e σ_Y é o desvio padrão do carácter aleatória Y. A partir de r_{XY} podemos tirar conclusões sobre a direção e intensidade da relação existente entre as caracteres X e Y.

2.3.2 ANÁLISE BAYESIANA

As redes bayesianas são métodos que podem gerar informações precisas para o estudo da correlação entre as características. A representação gráfica das redes bayesianas pode facilitar a compreensão da relação existente entre um conjunto de características, principalmente devido a propriedade da relação de dependência condicional, representada pela distribuição conjunta dos dados (MARQUES; DUTRA, 2002). Além disso, é uma maneira intuitiva de entender as relações diretas e indiretas entre caracteres (FELIPE et al., 2015).

O Directed Acyclic Graph é a denominação dada à representação gráfica produzida por essa modelagem estatística, na qual são apresentados um conjunto de círculos que representam as caracteres, conectadas entre si diretamente por meio de setas (MARGARITIS, 2002).

O método bayesiano tem como princípio descrever a incerteza em torno de um parâmetro, utilizando como medida de incerteza a probabilidade desse parâmetro assumir determinados valores (FARIA et al., 2007).

O modelo bayesiano tem como filosofia o fato de tudo o que é desconhecido é incerto e, portanto, toda a incerteza deve ser quantificada em termos de probabilidade (PAULINO et al., 2003). Como pretende-se identificar os melhores genótipos entre aqueles em avaliação, é interessante considerar os seus efeitos como aleatórios, pois essa abordagem permite uma melhor predição dos verdadeiros valores genotípicos e, conseqüentemente, um melhor ranqueamento dos genótipos (SMITH et al., 2005).

Na rede bayesiana, as associações são dispostas de tal maneira que pode-se observar a ordem de importância em termos de predição e dependência condicional da associação direta entre as características avaliadas (FELIPE et al. 2015).

Segundo Resende (2000), a inferência à análise Bayesiana permite obter estimativas mais precisas de componentes de variância, parâmetros genéticos, valores genéticos e ganhos genéticos. Isso faz com que esse modelo se torne interessante para o propósito de predição e seleção de caracteres no contexto do melhoramento genético.

3 REFERENCIAL BIBLIOGRÁFICO

AHMAD, S. Q., KHAN, S. E AHMAD, G. (2011). Genetic diversity analysis for yield and other parameters in maize (*Zea mays*L.) genotypes. **Asian Journal of Agricultural Sciences**, 3, 385-388.

ALENCAR, B. J. de; BARROSO, L. C.; ABREU, J. F. de. Análise Multivariada de Dados no Tratamento da Informação Espacial: uma abordagem com análise de Agrupamentos. **Revista Iberoamericana de Sistemas, Cibernética e Informática**, Buenos Aires, v. 10, n. 2, p. 6-12, 2013.

ALMEIDA, R. D.; PELUZIO, J. M.; AFFÉRI, F. S. Divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do Estado Tocantins. **Revista Ciência Agronômica**, v.42, n.01, p.108-115, 2011.

BARBIERI, R.L.; LEITE, D.L.; CHOER, E. E SINIGAGLIA, C. (2005) – Divergência genética entre populações de cebola com base em marcadores morfológicos. **Ciência Rural**, vol. 35, n. 2, p. 303-308.

BENITEZ, L.C.; RODRIGUES, I.C.S.; ARGE, W.P.; RIBEIRO, M.V.; BRAGA, E.J.B. Análise multivariada da divergência genética de genótipos de arroz sob estresse salino durante a fase vegetativa. **Revista Ciência Agronômica**. v.42, n.02, p.409-416, 2011.

CASTRO, A.F.N. M.; CASTRO, R.V.O.; CARNEIRO, A.C.O.; LIMA, J.E.; SANTOS, R.C.; ALVES, I.C.N. Análise multivariada para seleção de clones de eucalipto destinados à produção de carvão vegetal. **Pesquisa Agropecuária Brasileira**, v.48, n.6, p.627-635, 2013.

CHARNET, R. et al. **Análise de modelos de regressão linear**. Campinas: Unicamp, 2008. 357 p.

COIMBRA, R.R.; MIRANDA, G.V.; CRUZ, C.D.; ECKERT, F.R. Caracterização e divergência genética de populações de milho resgatadas do Sudeste de Minas Gerais. **Revista Ciência Agronômica**, v.41, n.01, p.159-166, 2010.

CROSSA, J.; FRANCO, J. Statistical methods for classifying genotypes. **Euphytica**, v. 137, n. 1, p. 19-37, 2004.

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos Biométricos Aplicados ao Melhoramento Genético**. 2.ed. Viçosa: UFV, 2003. 585 p.

CRUZ, C. D.; FERREIRA, F. M.; PESSONI, L. A. **Biometria aplicada ao estudo da diversidade genética**. Visconde do Rio Branco: Suprema, 2011. 620p.

CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **Modelos Biométricos aplicados ao melhoramento genético**. 3. ed. Viçosa: UFV, v. 2, 2014. 688 p.

- DORIA FILHO, U. Introdução à bioestatística. São Paulo: Negócio Editora, 1999. 152p.
- DUTRA FILHO, J.A.; MELO, L.J.O.; RESENDE, L.V.; ANUNCIÇÃO FILHO, C.J.; BASTOS, G.Q. Aplicação de técnicas multivariadas no estudo da divergência genética em cana-de-açúcar. **Revista Ciência Agrônômica**, v.42, n.01, p.185- 192, 2011.
- ERTIRO, B. T., TWUMASI-AFRIYIE, S., BLÜMMEL, M., FRIESEN, D., NEGERA, D., WORKU, M., ABAKEMAL, D. AND KITENGE, K. (2013). Genetic variability of maize stover quality and the potential for genetic improvement of fodder value. **Field Crops Research**, 153, 79-85. <http://dx.doi.org/10.1016/j.fcr.2012.12.019>.
- FAOSTAT (2017) – Production: crops/maize. <http://faostat.fao.org/> [cit. 2014.02.01]. Food and Agriculture Organization of the United Nations Statistics.
- FARIA, C. U. de; MAGNABOSCO, C. U.; REYES, A. L.; LÔBO, R. B.; BEZERRA, L. A. F. Inferência bayesiana e sua aplicação na avaliação genética de bovinos da raça nelore: revisão bibliográfica. **Ciência Animal Brasileira**, v. 8, n. 1, p. 75-86, jan./mar. 2007.
- FELIPE, V. P.; Silva, M. A.; Valente, B. D.; Rosa, G. J. Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. **Poultry science**, v. 94, n. 4, p. 772-780, 2015.
- FREIRE FILHO, F. R. Genética no feijão-caupi. In: ARAÚJO, J. P. P.; WATT, E. E. (Ed.). **O feijão-caupi no Brasil**. Brasília: IITA/EMBRAPA-CNPAF, 1988. p. 159-229.
- HAIR, J. F.; BLACK, W.; BABIN, B.; ANDERSON, R. E.; TATHAM, R. L. **Análise Multivariada de dados**. Editora Bookman, Porto Alegre, 6ª ed., 2009. 688 p.
- HALLAUER, A. R.; CARENA, J. M.; MIRANDA FILHO, J. B. de. Quantitative genetics in maize breeding. New York: Springer, 2010. 500 p.
- JAIN, A. K. **Data clustering: 50 years beyond K-means**. Pattern Recognition Letters, n. 31, p. 651-666, 2010.
- LANDIM, P. M. B. Análise estatística de dados geológicos multivariados. São Paulo: Oficina de Textos, 2011. 208 p.
- LEE, M. (1995). DNA markers and plant breeding programs. **Advances in Agronomy**, 55: 265-344.
- LIRA, S. A. **Análise de correlação: análise teórica e de construção dos coeficientes com aplicações**. 2004. 196 p. Dissertação (Mestrado e Métodos Numéricos em Engenharia dos Setores de Ciências Exatas e de Tecnologia) – Universidade Federal do Paraná, Curitiba, 2004.

LOPES, S. J.; LUCIO, A. D.; STORCK, L.; DAMO, H. P.; BRUM, B.; SANTOS, V. J. D. Relações de causa e efeito em espigas de milho relacionadas aos tipos de híbridos. **Ciência Rural**, v. 37, n. 6, p. 1536-1542, 2007.

MANLY, B. F. J. **Métodos estatísticos multivariados: uma introdução**. 3° ed. Porto Alegre: Bookman, 2008. 229 p.

MARGARITIS, D. **Learning Bayesian network model structure from data**. 2002. 126 f. (PhD)- School of Computer Science, Carnegie Mellon University, Pittsburgh, 2003.

MARQUES, R. L.; DUTRA, I. N. Ê. S. **Redes Bayesianas: o que são, para que servem, algoritmos e exemplos de aplicações**. Rio de Janeiro: Coppe Sistemas; UFRJ, 2002.

MELANI, M.D., CARENA, M.J. (2005). Alternative maize heterotic patterns for the northern Corn Belt. **Crop Science**, 45: 2186-2194.

MINGOTI, A.S. (2007). **Análise de dados através de métodos de estatística multivariada**. Belo Horizonte, Editora UFMG, 295p.

OSTLE, B., MALONE, L.C. **Statistics in research**. 4. ed. Ames: Iowa State University Press, 1988. 664p.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003. 446 p.

PELUZIO, J.M.; VAZ DE MELO, A.; AFFÉRI, F.S.; SILVA, R.R.; BARROS, H.B.; NASCIMENTOS, I.R.; FIDELIS, R.R. Variabilidade genética entre cultivares de soja, sob diferentes condições edafoclimáticas. **Revista Ciência Agronômica**, v.42, n.01, p.108-115, 2009.

PITTA, R. M.; BOIÇA JR, A. L.; DE JESUS, F. G.; TAGLIARI, S. R. Seleção de genótipos resistentes de amendoineiro a *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae) com base em análises multivariadas. **Neotropical Entomology**, Londrina, v. 39, p. 260-265, 2010.

REGAZZI, A.J. **Análise Multivariada**, Notas de Aula EST746, Departamento de Estatística da Universidade Federal de Viçosa. V.2. 2006.

RESENDE, M.D.V. de. **Inferência Bayesiana e simulação estocástica (amostragem de Gibbs) na estimação de componentes de variância e de valores genéticos em plantas perenes**. Colombo: Embrapa Florestas, 2000. 68p. (Embrapa Florestas. Documentos, 46).

ROCHA, M.C; GONÇALVES, L.S.A.; RODRIGUES, R.; SILVA, P.R.A.; CARMO, M.G.F.; ABBOD, A.C.S. Uso do algoritmo de Gower na determinação da divergência genética

entre acessos de tomateiro do grupo cereja. **Acta Scientiarum. Agronomy**, v. 32, n. 3, p. 423-431, 2010.

ROTILI, E. A.; CANCELLIER, L. L.; DOTTO, M. A.; PELUZIO, J. M.; CARVALHO, E. V. Divergência genética em genótipos de milho, no Estado do Tocantins. **Revista Ciência Agronômica**, v.43, n.3, p.516-521, 2012. DOI: 10.1590/S1806-66902012000300014.

SANTOS, E. R.; SANTOS, A. F.; CAPONE, A.; SANTOS, W. R.; MOURA, S. G.; BARROS, H. B. Genetic dissimilarity between soybean genotypes cultivated in lowland irrigated inter-cropping. **Journal of Biotechnology and Biodiversity**, v.4, n.3: p.222-231, 2013.

SANTOS, E.R.S.; BARROS, H.B.; FERRAZ, E.C.; CELLA, A.J.S.; CAPONE, A.; SANTOS, A.F.; FIDELIS, R.R. Divergência entre genótipos de soja, cultivados em várzea irrigada. **Revista Ceres**, v.58, n.6, 2011.

SMITH, A. B.; CULLIS, B. R.; THOMPSON, R. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. **Journal of Agricultural Science**, v. 143, n. 6, p. 449-462, 2005.

VASCONCELOS, E.S. de; CRUZ, C.D.; BHERING, L.L. FERREIRA, A. Estratégias de amostragem e estabelecimento de coleções nucleares. **Pesquisa Agropecuária Brasileira**, v.42, p.507-514, 2007.

VOGT, G.A.; BALBINOT JÚNIOR, A.A.; SOUZA, A.M. Divergência genética entre cultivares de girassol no planalto norte catarinense. **Scientia Agraria**, v.11, n.04, p.307-315, 2010.

1 **CAPÍTULO 2 – Abordagem multivariada na alocação de genótipos de milho em**
 2 **grupos divergentes**

3 (Revista Ciência Agronômica)

4 **Abordagem multivariada na alocação de genótipos de milho em grupos**
 5 **divergentes**

6 **Multivariate approach in the allocation of maize genotypes in divergent groups**

7 **RESUMO** - O objetivo com esta pesquisa constitui em estimar a capacidade geral
 8 de combinação de genótipos de milho e a alocação dos mesmos em grupos divergentes
 9 pela abordagem multivariada. Foram avaliados 229 híbridos *topcrosses* em três épocas
 10 de semeadura. Foi utilizado o delineamento de blocos casualizados com duas repetições,
 11 avaliando-se os seguintes caracteres agronômicos: altura de planta, posição relativa da
 12 espiga, acamamento, quebramento, produtividade de grãos e prolificidade. Foram
 13 estimadas as capacidades gerais de combinação dos genótipos de milho e realizada
 14 análise de agrupamento dos mesmos pelos métodos não-hierárquico K-médias e de
 15 otimização Tocher. As análises estatísticas foram realizadas utilizando-se o programa
 16 computacional SAS e Selegen. Na análise de variância conjunta foi constatado efeito
 17 significativo para safra em todas caracteres analisadas, caracterizando a influência das
 18 safras e indicando diferença entre os ambientes de avaliação. A amplitude dos valores
 19 da capacidade geral de combinação para produtividade de grãos foi de 1.655,2 kg.ha⁻¹ a
 20 -1.906,0 kg.ha⁻¹, indicando variabilidade na frequência de alelos favoráveis para essa
 21 característica entre os genótipos avaliados. Foi possível alocar os genótipos de milho em
 22 grupos heteróticos a partir dos métodos de agrupamento. Assim, foi possível estimar a
 23 capacidade geral de combinação dos genótipos de milho. Foi possível também selecionar
 24 genótipos com capacidade geral de combinação favorável nas três safras, como os
 25 genótipos 3, 12, 16 e 54, e também genótipos específicos, com elevada capacidade geral
 26 de combinação, para a safra 15/16 e safrinha 16, genótipos 3 e 16, e para a safra 16/17,
 27 genótipos 40 e 113. Os genótipos de milho foram alocados quanto à divergência a partir
 28 das análises multivariadas de agrupamento K-médias e Tocher. Foi possível identificar
 29 sete grupos divergentes pelo método K-médias e 21 pelo método Tocher.

30
 31 **Palavras chave:** agrupamento, melhoramento de plantas, *Zea mays*.

32
 33 **ABSTRACT** - The objective of this research is to estimate the general ability of combining
 34 maize genotypes and their allocation to divergent groups by the multivariate approach.
 35 Were evaluated 229 topcross hybrids in three sowing seasons. A randomized block design
 36 with two replications was used, evaluating the following agronomic characters: plant
 37 height, relative position of the cob, lodging, breaking, grain yield and prolificacy. The
 38 general combining capacities of the maize genotypes were evaluated and the grouping
 39 analysis of the same were performed by the non-hierarchical *K-means* and Tocher
 40 optimization methods. Statistical analyzes were performed using the SAS and Selegen

41 software. In the analysis of joint variance, a significant effect for harvest in all analyzed
42 variables, characterizing the influence of the crops and indicating difference between the
43 evaluation environments. The amplitude of the general combining capacities for grain yield
44 was of 1.655,2 kg.ha⁻¹ a -1.906,0 kg.ha⁻¹, indicating variability in the frequency of favorable
45 alleles for this characteristic among the evaluated genotypes. It was possible to allocate
46 the maize genotypes to heterotic groups from the clustering methods. Thus, it was
47 possible to estimate the general combining ability of maize genotypes. It was also possible
48 to select genotypes with favorable overall combining ability in the three harvests, such as
49 genotypes 3, 12, 16 and 54, as well as specific genotypes with high overall combining
50 ability for the 15/16 and 16th crops, genotypes 3 and 16, and for the 16/17 crop, genotypes
51 40 and 113. Corn genotypes were allocated for divergence from the multivariate K-
52 medium and Tocher cluster analyzes. It was possible to identify seven divergent groups
53 by the K-means method and 21 by the Tocher method.

54

55 **Key words:** clustering, plant breeding, *Zea mays*.

56

INTRODUÇÃO

57

Os programas de melhoramento de plantas baseados em combinações híbridas
58 devem concentrar seus esforços na etapa de identificação de linhagens superiores e na
59 capacidade destas em transmitir características desejáveis aos híbridos (HALLAUER et
60 al., 2010). O grande desafio na avaliação de linhagens é a quantidade de cruzamentos
61 possíveis a serem avaliados em campo. Davis (1927) propôs o método *topcross*, no qual
62 as linhagens obtidas são cruzadas com um testador comum e essas progênies são
63 avaliadas em experimentos comparativos com repetições. Na tentativa de minimizar esta
64 dificuldade, o método de *topcross* foi proposto para a otimização deste processo e
65 redução de tempo e recursos (ELIAS; CARVALHO; ANDRÉ, 2000; SCAPIM et al., 2008).
66 Algumas vantagens desse método são a avaliação e identificação de híbridos superiores
67 e, também, a separação das linhagens em grupos heteróticos (GUIMARÃES et al. 2012).

68

O cruzamento *topcross* possibilita estimar o efeito da capacidade geral de
69 combinação (CGC) para selecionar os genitores mais promissores na obtenção de

70 híbridos (CARDOSO, 2006). Por esse método é possível avaliar as linhagens com melhor
71 comportamento médio dentro do conjunto de genitores testados (CRUZ et al., 2012).

72 As técnicas de análise de agrupamento têm por objetivo dividir um grupo original
73 de observações em vários grupos homogêneos, semelhantes dentro de cada grupo,
74 segundo algum critério de similaridade ou dissimilaridade (CRUZ, 1990). Entre os
75 métodos de agrupamento utilizados no melhoramento de plantas, citam-se os
76 hierárquicos, os não-hierárquicos e os de otimização. Nos métodos hierárquicos, os
77 genótipos são agrupados por um processo que se repete em vários níveis, até que seja
78 estabelecido o dendrograma. As delimitações podem ser estabelecidas por um exame
79 visual do dendrograma, em que se avaliam pontos de mudança brusca de nível, tomando-
80 os em geral como delimitadores do número de genótipos para determinado grupo (CRUZ;
81 CARNEIRO, 2003). Os métodos não-hierárquicos diferem dos hierárquicos na
82 quantidade de classes formadas, a qual é determinada antes da realização da análise
83 (PITTA et al., 2010). Um método não-hierárquico bastante utilizado é o K-médias (*K-*
84 *means*), que procura alocar os elementos de dados em grupos previamente definidos. É
85 um método que minimiza a soma dos quadrados residuais dentro de cada grupo formado,
86 aumentando a homogeneidade dentro deles ao mesmo tempo em que aumenta a
87 diferença entre eles (ALENCAR et al., 2013). Por sua vez, nos métodos de otimização,
88 destaca-se o algoritmo de Tocher, que constitui um método de agrupamento simultâneo,
89 o qual realiza a separação dos indivíduos de uma única vez.

90 Dadas as considerações apresentadas, o objetivo com esta pesquisa constitui em
91 estimar a capacidade geral de combinação de genótipos de milho e a alocação dos
92 mesmos em grupos divergentes pela abordagem multivariada.

MATERIAL E MÉTODOS

93

94

95

96

97

98

99

100

101

102

Foram obtidos 229 híbridos *topcrosses*, oriundos do cruzamento de 229 linhagens parcialmente endogâmicas (S_3) de milho com um testador de base genética ampla. Os cruzamentos *topcrosses* foram realizados no ano agrícola de 2014/2015, em lote isolado de despendoamento, utilizando-se a proporção de 4 linhas femininas (formadas com cada uma das linhagens parcialmente endogâmicas), para 2 linhas masculinas (formadas pelo testador). O despendoamento das plantas fêmeas foi realizado manualmente antes que estas liberassem grãos de pólen e, após a maturidade fisiológica dos grãos, foi realizada a colheita manual das linhas fêmeas, sendo as sementes armazenadas em câmara fria, logo após a debulha manual e identificação dos cruzamentos.

103

104

105

106

107

Os experimentos para avaliação dos *topcrosses* foram conduzidos em campo em três ambientes distintos: primeira safra 2015/2016 (safra 15/16), segunda safra 2016 (safrinha 16) e primeira safra 2016/2017 (safra 16/17), no município de Jaboticabal-SP (Latitude 21° 14' 05" S, longitude 48° 17' 09" W e altitude de 614 m acima do nível do mar).

108

109

110

111

112

113

114

115

Os experimentos foram dispostos no delineamento de blocos casualizados com duas repetições e 229 tratamentos. Cada parcela foi constituída por duas linhas de cinco metros de comprimento, com espaçamento de 0,5 m entre linhas e 0,33 m entre plantas, representando 64.000 plantas.ha⁻¹. Os caracteres avaliados nos experimentos foram: altura de planta (AP, obtida do solo até a inserção da folha bandeira, em cm), posição relativa da espiga (PRE, razão entre altura de planta e altura da inserção da espiga), acamamento (AC, em número de plantas/parcela inclinadas abaixo de 45° em relação ao solo, transformado para $\sqrt{x+0,5}$), quebraamento (QUE, em número de plantas/parcela com

116 quebramento do colmo abaixo da espiga principal, transformado para $\sqrt{x+0,5}$,
117 produtividade de grãos (PG, em $\text{kg}\cdot\text{ha}^{-1}$ corrigida para umidade 13% e estande ideal por
118 covariância) e prolificidade (PRO, razão entre número de espigas e estande final).

119 Com os dados coletados nos experimentos realizou-se a análise de variância por
120 ambiente e, posteriormente, a análise conjunta de acordo com Ramalho et al. (2000).

121 Foi estimar a capacidade geral de combinação (CGC), considerando-se cada
122 ambiente e as médias dos três ambientes de avaliação, utilizando o seguinte modelo
123 estatístico-genético:

$$124 \quad Y_i = m + g_i + e_i$$

125 sendo, Y_i = média do *topcross* entre o genótipo i e o testador; m = média geral dos
126 *topcrosses*; g_i = efeito da capacidade geral de combinação do genótipo i ; e_i = erro
127 experimental médio.

128 A g_i foi obtida de acordo com a seguinte expressão:

129 $g_i = c_i - m$, sendo ; c_i = média geral do i -ésimo *topcross*; m = média geral dos
130 *topcrosses*.

131 A CGC foi estimada apenas para a produtividade de grãos, por ser o caráter de
132 maior importância agrônômica e econômica.

133 Para alocação dos genótipos avaliados em grupos heteróticos foram utilizadas as
134 análises multivariadas de agrupamentos não-hierárquica K-médias (*K-means*) e de
135 otimização de Tocher, seguindo o critério de que a variância entre os genótipos não
136 poderia ser maior dentro dos grupos do que entre os grupos.

137 Foi realizado o agrupamento prévio por distância euclidiana pelo método hierárquico
138 das médias aritméticas não-ponderadas UPGMA (*Unweighted pair group method using*

139 *arithmetic average*) para os 229 genótipos de milho, formando sete grupos, sendo este
140 agrupamento realizado a partir do ponto de corte do dendrograma onde foram observadas
141 mudanças bruscas de níveis. Esta pressuposição foi utilizada para o método K-médias.

142 As análises estatísticas foram realizadas utilizando-se o programa computacional
143 SAS (SAS INSTITUTE, 2011) e Selegen (RESENDE, 2015).

144 **RESULTADOS E DISCUSSÃO**

145 Na análise de variância conjunta (Tabela 1) foi constatado efeito significativo para
146 ambiente em todos caracteres analisados, caracterizando a influência destes e indicando
147 diferença entre os ambientes de avaliação. Também foi detectado efeito significativo para
148 genótipo em relação a todas as caracteres, indicando que pelo menos um genótipo difere
149 dos demais. Considerando a interação dos genótipos com os ambientes, verifica-se que
150 os genótipos não apresenta comportamento diferente em função do ambiente para PRE,
151 PRO e PG.

152 Na análise conjunta os coeficientes de variação (CVs) variaram de 4,02 % (AP) a
153 38,49 % (AC), apresentando valores dentro dos padrões normais de experimentação em
154 milho, de acordo com Fritsche-Neto et al. (2012), confirmando a precisão experimental e
155 a acurácia na estimativa das médias, em todos os ambientes, com exceção das
156 caracteres AC e QUE, por apresentarem maior influência ambiental. Lima et al. (2006)
157 relatam que a magnitude do coeficiente para AC e QUE pode estar relacionada com
158 médias baixas, obtidas por valores reduzidos ou até mesmo nulos apresentados pela
159 característica. Gomes et al. (2010) relatam que a avaliação por contagem de plantas é
160 um método que pode ainda contribuir para aumentar o coeficiente de variação para estes
161 caracteres. Contudo, trabalhos com relação a estes caracteres indicam que os fatores

162 mais influentes são as rajadas de vento e chuvas, que apresentam inconstância de
163 ocorrências e, portanto, também contribuem para os elevados valores de CV% obtidos
164 (EASSON; WHITE; PICKLES, 1993; BUENO et al., 2011).

165 Nas análises de variância individuais (Tabela 1), houve efeito significativo dos
166 genótipos para AP, PRE, AC e QUE, na safrinha 16. Já para as safras 15/16 e 16/17,
167 houve efeito significativo dos genótipos para as caracteres AP e PRE, e AP, PRE e QUE,
168 respectivamente. Por sua vez, para as caracteres PRO e PG não foi observada diferença
169 significativa entre os genótipos, em nenhuma safra.

170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197

198 Tabela 1 – Resumo das análises de variância, individuais e conjunta, com os
 199 quadrados médios e significâncias para caracteres agronômicos em genótipos de
 200 milho.

FV	GL	AP ⁺⁺	PRE	AC ⁺	QUE ⁺⁺	PRO	PG ⁺
Safrinha 15/16							
Genótipo	228	0,13 ^{**}	0,0013 ^{**}	0,005 ^{ns}	0,01 ^{ns}	0,07 ^{ns}	7,08 ^{ns}
Repetição	1	27,30 ^{**}	0,0193 ^{**}	0,004 ^{ns}	0,39 ^{**}	4,82 ^{**}	1213,19 ^{**}
Erro	228	0,06	0,0008	0,004	0,01	0,06	8,22
Média		197,17	0,58	0,06	0,19	1,30	7425,79
Máximo		219,35	0,67	2,32	2,11	2,35	9491,7
Mínimo		172,20	0,51	0,71	0,71	0,97	5436,2
CV(%)		3,85	4,80	27,19	37,75	18,77	12,21
Safrinha 16							
Genótipo	228	0,14 ^{**}	0,0011 ^{**}	0,06 ^{**}	0,06 ^{**}	0,03 ^{ns}	11,75 ^{ns}
Repetição	1	0,91 ^{**}	0,0005 ^{ns}	2,53 ^{**}	0,001 ^{ns}	0,08 ^{ns}	2940,67 ^{**}
Erro	228	0,09	0,0007	0,04	0,03	0,03	10,99
Média		218,90	0,58	1,32	3,92	0,94	7425,79
Máximo		236,30	0,66	3,08	4,12	1,54	9658,0
Mínimo		193,80	0,51	0,71	0,96	0,30	5030,0
CV(%)		4,26	4,69	37,57	22,02	18,71	14,12
Safrinha 16/17							
Genótipo	228	0,15 ^{**}	0,0009 ^{**}	0,01 ^{ns}	0,04 [*]	0,03 ^{ns}	13,91 ^{ns}
Repetição	1	1,08 ^{**}	0,0022 ^{ns}	0,001 ^{ns}	0,14 [*]	1,54 ^{**}	315,06 ^{**}
Erro	228	0,08	0,0006	0,01	0,03	0,03	14,66
Média		231,10	0,61	0,14	1,08	1,17	7725,08
Máximo		253,45	0,65	1,97	2,97	1,74	10635
Mínimo		204,05	0,43	0,71	0,71	0,89	5487
CV(%)		3,90	4,20	36,54	37,07	15,09	15,67
Conjunta							
Genótipo (G)	228	0,23 ^{**}	0,0017 ^{**}	0,02 ^{**}	0,04 ^{**}	0,05 ^{**}	20,83 ^{**}
Ambiente (A)	2	135,83 ^{**}	0,0930 ^{**}	11,11 ^{**}	27,70 ^{**}	15,13 ^{**}	137,34 ^{**}
Repetição	3	9,76 ^{**}	0,0073 ^{**}	0,84 ^{**}	0,18 ^{**}	2,15 ^{**}	1489,64 ^{**}
G*A	456	0,09 ^{**}	0,0008 ^{ns}	0,02 ^{**}	0,04 ^{**}	0,04 ^{ns}	5,95 ^{ns}
Erro	684	0,07	0,0007	0,02	0,02	0,04	11,29
Média		215,72	0,59	0,35	1,32	1,14	7525,55
Máximo		230,2	0,65	1,90	2,41	1,47	9180,77
Mínimo		194,8	0,51	0,71	0,94	0,90	5619,55
CV (%)		4,02	4,56	38,49	30,24	17,71	14,12

201 Coeficientes de variação em % (CV), altura de planta (AP, em cm), posição relativa da espiga (PRE, em
 202 cm), acamamento (AC, número de plantas/parcela), quebramento (QUE, número de plantas/parcela),
 203 prolificidade (PRO, número de espigas/estande final) e produtividade de grãos (PG, kg/ha). ^{ns}, * e ^{**}, não
 204 significativo, significativo a 5% e significativo a 1% pelo teste F, respectivamente. +: dividido por 100.000,
 205 ++: dividido por 1.000, +++: dividido por 10. Máximo – valor máximo dos caracteres para os genótipos de
 206 milho. Mínimo – valor mínimo dos caracteres para os genótipos de milho.

207

208 A amplitude dos valores da CGC (Tabela 2), com base na média dos três
209 ambientes, para PG foi de 1.655,2 kg.ha⁻¹ (genótipo 3) a -1.906,0 kg.ha⁻¹ (genótipo 26).
210 Os genótipos 3, 12, 14, 16, 18, 49, 54, 113, 167, 168, 205, 211 e 236 destacaram-se com
211 CGC superior a 900 kg.ha⁻¹, na média das três safras. Observa-se para os genótipos 3,
212 12, 16 e 54 elevados valores de CGC nas três safras, sendo menos influenciados pelo
213 ambiente, indicando a estabilidade destes genótipos. Nos genótipos 5, 48, 151 e 186
214 ocorreu forte influência do ambiente, com elevada CGC em duas safras, porém valores
215 discrepantes e negativos em outra safra. Nos genótipos 14, 33, 39, 49 e 57 observou-se
216 forte influência do ambiente, porém, sem valores negativos. Os genótipos que
217 destacaram-se em cada safra foram: safra 15/16, genótipos 3 e 16 com CGC de 2.065,9
218 kg.ha⁻¹ e 1.692,5 kg.ha⁻¹, respectivamente; safrinha 16, genótipos 3 e 16, com CGC de
219 2.231,7 kg.ha⁻¹ e 1.616,3 kg.ha⁻¹, respectivamente; e na safra 16/17, os genótipos 40 e
220 113, com CGC de 1.889,5 kg.ha⁻¹ e 1.661,0 kg.ha⁻¹. A CGC é um indicador importante
221 em programas de melhoramento de milho visando a obtenção de híbridos, pois a partir
222 desta análise é possível identificar melhores genótipos em relação aos alelos favoráveis
223 para a característica considerada, dentre os demais genótipos avaliados.

224 Em trabalho com 47 genótipos de milho, foi estimado a capacidade de
225 combinação, observando que um genótipo destacou-se por apresentar alta CGC para PG
226 em quatro ambientes, sendo apropriado para melhoramento de populações por sua
227 estabilidade (Bolson et al., 2016). Por vez, Clovis et al. (2015), realizaram trabalho para
228 identificar linhagens adaptadas à condição de safrinha por meio da análise da capacidade
229 de combinação de 50 linhagens S₃ de milho. Destas linhagens apenas uma apresentou

- 230 alta capacidade geral de combinação para produtividade de grãos nos três ambientes
231 estudados.

232 Tabela 2 - Estimativas dos efeitos da capacidade geral de combinação (CGC) em kg.ha⁻¹, para produtividade de grãos
 233 (PG), de 229 genótipos (G) de milho em três ambientes (1^a – Safra 2015/2016, 2^a – Safrinha 2016, 3^a – Safra 2016/2017,
 234 M – Média da CGC nas três safras).

G	1 ^a	2 ^a	3 ^a	M	G	1 ^a	2 ^a	3 ^a	M	G	1 ^a	2 ^a	3 ^a	M
1	-787,88	-318,26	-208,88	-438,34	44	389,66	240,37	-187,06	147,65	91	-458,29	-560,14	546,54	-157,30
2	113,15	653,01	-274,34	163,94	45	137,90	213,58	396,99	249,49	92	-260,20	-125,69	803,10	139,07
3	2065,87	2231,74	668,03	1655,21	46	-105,05	-335,98	775,17	111,38	93	-667,45	-974,32	287,45	-451,44
4	-562,63	-1029,92	-858,35	-816,97	47	295,62	593,41	699,32	529,45	94	338,48	107,55	32,64	159,56
5	895,53	1143,03	-186,80	617,25	48	904,21	1500,06	-447,79	652,16	95	89,70	696,95	-271,48	171,73
6	-344,54	-105,59	-652,88	-367,67	49	882,84	598,77	1482,72	988,11	96	-75,49	-239,03	-1404,29	-572,93
7	216,24	668,76	158,30	347,76	50	-445,20	-996,99	157,35	-428,28	97	-346,28	-138,67	25,00	-153,32
8	266,76	871,16	-438,20	233,24	51	280,60	423,67	-62,89	213,79	98	-684,17	-1227,41	-227,65	-713,08
9	510,28	805,22	851,31	722,27	52	309,38	685,97	-76,44	306,30	99	226,25	135,80	778,01	380,02
10	-1125,57	-827,78	-826,10	-926,49	53	-30,80	438,82	-137,64	90,13	100	107,14	-126,64	-1180,60	-400,04
12	1078,48	1466,46	1263,66	1269,53	54	648,39	1165,45	1396,36	1070,07	101	537,68	762,38	188,47	496,18
13	-558,87	39,83	-795,02	-438,02	56	141,02	222,40	-490,51	-42,36	102	-665,85	-1740,66	-571,53	-992,68
14	369,61	979,71	1411,87	920,39	57	944,07	1388,04	94,94	809,02	103	550,76	708,08	692,36	650,40
15	260,68	496,78	-6,83	250,21	59	-69,32	585,37	424,17	313,41	104	379,63	-470,21	-1139,09	-409,89
16	1692,51	1616,31	1499,55	1602,79	60	-90,96	-496,57	374,06	-71,16	105	-178,35	-344,74	-2117,20	-880,10
17	705,82	860,29	-115,32	483,60	61	-30,10	93,01	164,01	75,64	106	-52,74	169,11	60,67	59,01
18	504,57	1038,73	1279,31	940,87	62	33,07	120,15	449,98	201,07	108	190,62	272,00	-779,58	-105,65
19	-74,47	231,87	-49,60	35,93	65	-324,01	201,60	-200,92	-107,78	109	-512,85	-1123,48	1362,69	-91,21
20	1028,12	1115,19	11,89	718,40	66	684,85	608,64	448,92	580,80	110	-61,65	-433,05	-300,07	-264,92
21	-402,88	-366,09	1268,76	166,60	68	429,60	952,36	583,13	655,03	111	-232,73	-823,41	786,18	-89,99
22	-338,25	-467,59	-19,59	-275,14	69	416,92	650,17	-635,74	143,78	112	88,43	304,59	344,03	245,69
23	451,39	527,07	-744,43	78,01	70	39,68	8,06	-957,23	-303,16	113	940,48	869,98	1661,06	1157,18
24	-89,77	185,22	-703,64	-202,73	71	-685,67	-694,48	-976,52	-785,56	114	292,51	371,04	-1102,19	-146,21
26	-1989,60	-1627,27	-2101,14	-1906,01	72	-186,14	-490,16	81,65	-198,22	115	-14,27	-650,54	104,26	-186,85
27	94,04	37,78	-1010,45	-292,88	73	-672,75	-1297,62	-286,42	-752,27	116	176,92	-585,59	592,15	61,16
28	-778,52	-711,39	-916,44	-802,12	74	180,42	370,93	-57,13	164,74	118	-796,06	-1015,59	-270,88	-694,18
29	-169,29	-20,52	-667,82	-285,88	75	-235,75	-264,51	465,10	-11,72	119	-316,56	68,57	264,63	5,55
30	23,05	557,21	-1355,34	-258,36	77	-551,94	-642,39	-823,77	-672,70	120	126,36	544,68	-450,88	73,39
31	-1115,98	-1344,06	1044,33	-471,91	78	-645,70	-252,02	-2218,47	-1038,73	121	476,07	540,34	-452,02	188,13
32	77,28	161,51	-25,52	71,09	79	207,88	44,33	-1266,55	-338,11	122	218,36	580,69	-396,12	134,31
33	177,23	401,93	1054,28	544,48	80	-892,01	-979,61	-1249,82	-1040,48	123	398,94	170,86	571,44	380,42
34	-632,42	-553,89	-701,91	-629,41	81	-1220,92	-1384,46	-98,48	-901,29	124	-339,71	185,90	340,35	62,18
35	264,19	27,56	571,01	287,59	82	-750,11	-907,95	-1031,58	-896,55	125	95,33	235,54	-273,06	19,27
36	-462,13	-301,96	1002,95	79,62	83	-239,64	-470,57	-945,12	-551,78	126	13,23	94,61	-128,52	-6,89
37	-676,48	-640,70	803,91	-171,09	84	-67,99	-149,90	-2237,59	-818,49	127	-1604,10	-2006,86	470,71	-1046,75
38	562,58	714,19	-131,44	381,78	85	320,19	772,71	401,90	498,27	128	251,64	751,60	996,91	666,72
39	738,45	1132,13	459,46	776,68	86	13,98	-424,82	-480,47	-297,11	129	309,47	201,91	321,32	277,57
40	-252,02	37,22	1889,53	558,24	87	-572,28	-429,21	924,36	-25,71	130	385,37	916,68	-1743,62	-147,19
41	-736,78	-886,07	374,67	-416,06	88	-1014,32	-2395,74	-1164,97	-1525,01	131	-32,34	189,51	-743,77	-195,53
42	-733,05	-598,54	-601,71	-644,43	89	-139,88	236,70	265,94	120,92	132	-120,21	-359,69	-854,87	-444,92
43	362,32	564,22	-40,81	295,24	90	405,64	247,79	582,51	411,98	133	500,24	584,56	986,71	690,54

235

236

G	1 ^a	2 ^a	3 ^a	M	G	1 ^a	2 ^a	3 ^a	M	G	1 ^a	2 ^a	3 ^a	M
135	551,54	942,38	242,45	578,79	175	7,24	-88,92	717,57	211,96	215	-897,06	-1586,47	-102,33	-861,95
136	415,39	859,36	262,72	512,49	176	799,05	1169,93	-446,63	507,45	216	-414,72	-333,34	75,55	-224,17
137	-298,09	-160,72	311,11	-49,23	177	-398,38	-514,49	-364,97	-425,95	217	336,49	-203,90	554,00	228,87
138	223,11	579,74	-1434,15	-210,43	179	-591,79	-294,00	-88,51	-324,77	218	-13,02	-851,46	78,30	-262,06
139	25,17	16,36	728,23	256,59	180	-429,79	-289,57	-1381,22	-700,19	219	1171,52	755,62	-231,03	552,04
140	44,23	496,75	-317,81	74,39	182	184,74	243,32	-0,32	142,58	220	445,90	-15,69	137,11	189,11
141	-1084,91	-733,98	-1461,27	-1093,38	183	-26,11	-662,39	1241,12	184,21	221	-983,60	-1729,00	66,59	-882
142	-461,11	-562,97	-854,65	-626,24	184	-427,59	156,85	-946,79	-405,84	222	-518,52	-755,15	-1414,02	-895,89
143	10,95	536,56	517,92	355,14	185	-1261,50	-1790,49	-405,56	-1152,52	223	193,18	128,38	-1302,70	-327,05
144	-221,17	-215,73	-786,15	-407,68	186	1096,85	936,16	-113,70	639,77	225	-619,19	-779,88	-830,74	-743,27
146	425,17	329,02	803,08	519,09	187	-955,79	-1892,97	-1348,09	-1398,95	226	-501,47	-501,74	961,72	-47,16
147	-1141,17	-477,93	175,44	-481,22	188	199,90	132,25	283,34	205,16	227	972,10	259,89	-137,76	364,75
149	-510,83	-221,58	906,42	58,00	189	-587,61	-660,97	-429,54	-559,37	228	411,41	712,06	654,45	592,64
150	-223,61	152,97	-1079,73	-383,46	190	-40,86	-417,97	442,37	-5,49	229	112,28	-299,03	424,16	79,14
151	912,32	1063,93	-141,34	611,64	192	783,59	710,24	-841,53	217,43	230	167,81	240,64	426,31	278,25
152	462,52	383,47	-553,24	97,58	193	409,84	564,30	-232,83	247,10	232	-1054,04	-1985,53	-444,44	-1161,34
153	114,08	321,68	-91,44	114,77	194	365,09	145,56	-747,21	-78,86	234	52,72	55,31	-89,25	6,26
154	481,87	1167,91	-961,25	229,51	195	475,15	314,46	898,73	562,78	235	-358,98	-685,80	-916,78	-653,85
155	-880,17	-1350,32	-765,67	-998,72	196	114,61	-99,22	521,32	178,90	236	991,08	897,77	1081,75	990,20
156	-283,31	-219,03	-698,13	-400,16	197	-207,72	-450,06	257,44	-133,45	237	-975,44	-1510,13	-491,01	-992,19
157	-43,17	35,36	884,78	292,32	198	178,71	639,78	-10,35,91	-72,47	238	-402,59	-473,09	715,45	-53,41
158	273,49	501,04	327,62	367,38	199	628,46	467,77	1216,93	771,05	239	-33,40	-711,42	174,65	-190,05
159	539,34	699,50	1212,85	817,23	200	-13,86	-239,09	-353,37	-202,10	240	-317,30	-300,46	-272,21	-296,66
160	-780,67	147,43	851,65	72,80	201	954,25	863,80	276,26	698,10	241	14,46	-449,99	-703,11	-379,54
161	-729,14	-603,17	817,88	-171,48	202	211,10	202,29	1410,57	607,99	242	-214,80	-48,93	-148,02	-137,25
163	-125,75	-98,52	-1988,34	-737,53	203	644,12	565,07	1315,03	841,41	243	417,56	186,63	-72,22	177,32
164	-304,50	-453,79	1013,63	85,11	204	410,59	72,37	644,00	375,65	244	-507,03	-583,24	29,12	-353,72
165	378,66	349,90	1157,27	628,61	205	927,65	1152,36	982,33	1020,78	245	-766,87	-775,68	-212,38	-584,98
166	-290,19	-824,88	-612,36	-575,81	206	505,38	271,60	-12,21	254,92	246	922,01	1242,60	-17,85	715,59
167	1161,62	933,54	714,86	936,67	207	-314,63	-773,37	1255,40	55,80	247	866,51	609,92	705,58	727,34
168	692,42	911,43	1175,15	926,33	208	147,17	315,89	-144,57	106,16	248	-494,60	185,74	1119,87	270,33
169	-880,04	-1896,01	129,72	-882,11	209	144,42	-246,94	413,10	103,53	249	194,14	413,14	554,17	387,15
170	228,50	973,38	799,57	667,15	210	672,30	540,11	419,78	544,06					
171	21,42	24,01	-807,88	-254,15	211	1331,83	1317,32	429,02	1026,06					
172	-858,83	-1515,06	-875,26	-1083,05	212	450,45	57356	1020,63	681,55					
173	-898,00	-839,42	-1938,27	-1225,23	213	96,95	-519,37	-219,72	-214,05					
174	-697,03	-154,32	580,35	-90,33	214	-333,35	-631,67	-185,51	-383,51					

197 De acordo com a análise de agrupamento pelo método K-médias os genótipos
198 avaliados foram alocados em sete grupos heteróticos, previamente determinados pelo
199 auxílio de um dendrograma, pré-estabelecido pelo método hierárquico UPGMA. A partir
200 do gráfico é possível observar o perfil da distribuição dos centróides de cada grupo em
201 relação a cada caráter (Tabela 3 e Figura 1).

202 O grupo 1, formado por 21 genótipos, destacou-se por apresentar maior valor para
203 o caráter AP e menor valor para o caráter PRE, sendo esta uma característica favorável
204 buscada em programas de melhoramento de milho (Tabela 3 e Figura 1). Plantas de
205 milho mais desenvolvidas, porém com a menor altura de inserção da espiga, favorecem
206 a colheita mecanizada (CORTEZ et al. 2009). Segundo Costa et al. (2012), a
207 homogeneidade dos valores médios de altura de inserção da espiga principal pode
208 favorecer a colheita mecanizada, com diminuição nas perdas de espigas não colhidas
209 pela plataforma da colhedora.

210 Em relação ao grupo 2, formado por 23 genótipos, foi possível observar maior valor
211 para o caráter PRE e menor valor para o caráter AP (Tabela 3 e Figura 1). De acordo
212 com Demétrio et al. (2008), a partir de 1970, os melhoristas passaram a se preocupar
213 com estudos sobre arquitetura da planta, baseados na premissa de que as plantas de
214 menor porte permitiriam uma semeadura mais adensada, com maior capacidade
215 fotossintética e, assim, maior rendimento de grãos.

216 O grupo 3, formado por 47 genótipos, apesar de ter apresentado baixo valor para
217 o caráter AP, apresentou alto valor para o caráter QUE, podendo justificar os menores
218 valores para o caráter PG, devido à quantidade de plantas quebradas (Tabela 3 e Figura
219 1).

220 Por sua vez, o grupo 4, formado por 25 genótipos, destacou-se por apresentar
221 maior valor para o caráter AC, sendo esta característica desfavorável e que resulta em
222 perdas de produtividade. Apesar disso, é possível observar também que este grupo
223 apresentou menor valor para o caráter QUE (Tabela 3 e Figura 1).

224 O grupo 5, formado por 34 genótipos, destacou-se por apresentar o menor valor
225 para o caráter PG. Isso pode ser justificado por este grupo ter apresentado alto valor para
226 o caráter AP, sendo que esta característica pode favorecer acamamento e quebramento
227 das plantas e por consequência a perda das espigas (Tabela 3 e Figura 1).

228 O grupo 6, formado por 39 genótipos, apresentou maior valor para caráter PG e
229 baixo valor para caráter AC. Menor número de plantas acamadas em uma parcela pode
230 favorecer a colheita e assim favorecer a produtividade desta (Tabela 3 e Figura 1).

231 O grupo 7, formado por 40 genótipos, apresentou maior valor para o caráter PRO
232 (Tabela 3 e Figura 1).

233 De acordo com Freddi et al. (2008) a análise de agrupamento K-médias permitiu
234 identificar três grupos de alta, média e baixa produtividade de plantas de milho, segundo
235 caracteres de solo, do sistema radicular e da parte aérea das plantas. Também, Silva et
236 al. (2015) a partir da análise de agrupamento pelo método de K-médias, conseguiram a
237 formação de três grupos para as médias padronizadas dos componentes produtivos de
238 feijão e dos atributos físicos do solo, sendo esta análise eficiente para verificar as
239 similaridades ou as diferenças na variabilidade da produtividade, a partir dos atributos do
240 solo.

241 A identificação dos grupos, realizada pelo método de agrupamento proposto por
242 Tocher, a partir das distâncias generalizadas de Mahalanobis (D^2), possibilitou a alocação

243 dos 229 genótipos de milho em 21 grupos, sendo que o maior deles (grupo 1) reuniu 147
244 genótipos (64,19%), e oito grupos reuniram apenas um genótipo, o que indica a
245 variabilidade no germoplasma estudado (Tabela 4). Vasconcelos et al. (2007)
246 argumentam que este método apresenta uma inconveniência no agrupamento de
247 genótipos com maior dissimilaridade. Na maioria dos casos, cada genótipo forma um
248 grupo específico (com apenas um genótipo) em virtude deste agrupamento ser
249 influenciado pela distância dos genótipos já agrupados. Isso ocorre devido ao fato de o
250 método utilizar um critério global de agrupamento, ou seja, baseia-se na maior distância
251 entre as menores distâncias encontradas na matriz de dissimilaridade durante todo o
252 processo.

253 Foram avaliadas 96 progênies de milho com o objetivo de estimar a divergência
254 genética entre estas. A divergência foi estimada usando a distância de Mahalanobis (D^2)
255 generalizada e as progênies foram agrupadas pelos métodos UPGMA e Tocher. O
256 método UPGMA foi mais sensível do que o de Tocher, uma vez que formou 11 grupos
257 geneticamente dissimilares em comparação com cinco grupos formados pelo método de
258 Tocher (Silva et al., 2016). Lopes et al. (2014), estudando a diversidade genética de 15
259 linhagens de milho doce, observaram que os genótipos foram divididos em grupos de
260 acordo com o método de otimização Tocher. Este método resultou na formação de oito
261 grupos.

262 A análise de agrupamento tem por finalidade reunir por algum critério de
263 classificação os genitores (ou qualquer outro tipo de unidade amostral) em vários grupos,
264 de tal forma que exista homogeneidade dentro do grupo e heterogeneidade entre grupos
265 (CRUZ; REGAZZI, 2004; BERTAN et al., 2006). Os grupos gerados pelos métodos K-

266 médias e Tocher facilitam a seleção de genótipos divergentes para obtenção de híbridos,
267 principalmente linhagens-elite, uma vez que permitem a seleção de genótipos
268 pertencentes a grupos heteróticos distintos a serem indicados para futuros cruzamentos,
269 facilitando a projeção dos trabalhos de melhoramento.

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305 Tabela 3 - Agrupamento de 229 genótipos de milho, quanto os caracteres altura de planta
 306 (AP), posição relativa da espiga (PRE), acamamento (AC), quebramento (QUE),
 307 produtividade de grãos (PG) e prolificidade (PRO), segundo o Método de K-médias.

Grupo	Nº	Genótipos
1	21	106, 108, 14, 157, 19, 202, 208, 21, 214, 217, 220, 23, 234, 246, 3, 35, 37, 44, 45, 46, 52
2	23	115, 121, 124, 135, 137, 153, 154, 164, 165, 179, 192, 203, 211, 229, 235, 65, 68, 79, 83, 85, 86, 87, 91
3	47	1, 100, 101, 102, 105, 110, 114, 125, 126, 127, 131, 132, 138, 144, 149, 150, 161, 163, 166, 174, 177, 183, 185, 187, 190, 206, 207, 209, 238, 4, 43, 59, 62, 69, 70, 71, 72, 80, 84, 88, 89, 92, 93, 94, 95, 96, 98
4	25	109, 111, 113, 122, 123, 128, 130, 133, 136, 143, 147, 159, 170, 186, 188, 199, 200, 205, 210, 212, 216, 230, 245, 75, 99
5	34	10, 104, 13, 142, 155, 160, 172, 189, 194, 198, 213, 215, 22, 221, 222, 226, 237, 239, 240, 242, 243, 244, 26, 27, 28, 31, 34, 41, 42, 61, 77, 78, 81, 82
6	39	103, 12, 139, 16, 167, 168, 17, 176, 18, 182, 2, 20, 201, 204, 219, 227, 228, 236, 247, 248, 30, 32, 33, 38, 39, 40, 47, 48, 49, 5, 51, 54, 57, 60, 66, 7, 8, 9, 90
7	40	112, 116, 118, 119, 120, 129, 140, 141, 146, 15, 151, 152, 156, 158, 169, 171, 173, 175, 180, 184, 193, 195, 196, 197, 218, 223, 225, 232, 24, 241, 249, 29, 36, 50, 53, 56, 6, 73, 74, 97

308

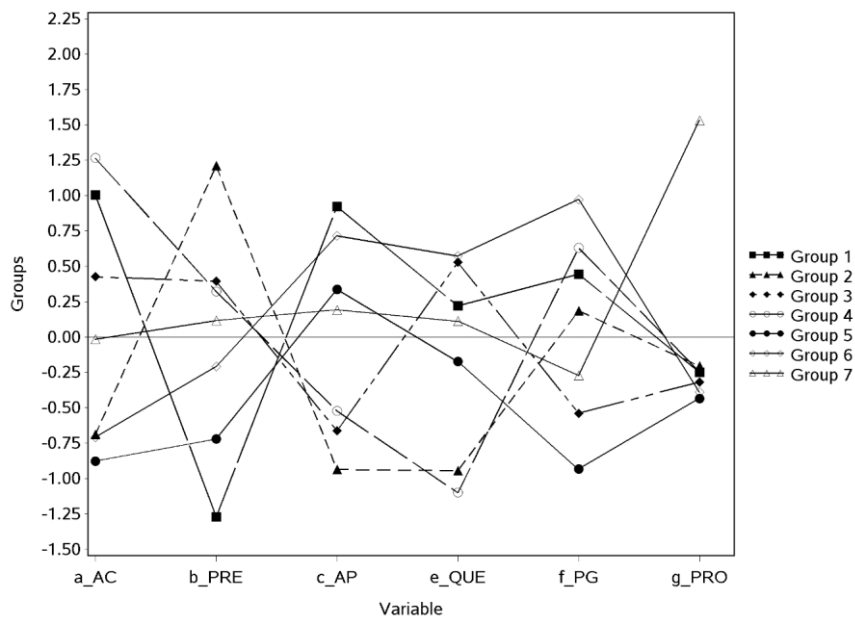
309

310

311

312

313



314

315 Figura 1 - Gráfico do perfil de 229 genótipos de milho da
 316 distribuição dos centróides dos grupos na análise de agrupamento
 317 pelo método K-médias, formados a partir dos caracteres: altura de
 318 planta (AP), posição relativa da espiga (PRE), acamamento (AC),
 319 quebramento (QUE), produtividade de grãos (PG) e prolificidade
 320 (PRO).

321

322

323

324

325

326

327

328

329

330

331 Tabela 4 - Agrupamento de 229 genótipos de milho, quanto os caracteres altura de
 332 planta (AP), posição relativa da espiga (PRE), acamamento (AC), quebramento
 333 (QUE), produtividade de grãos (PG) e prolificidade (PRO), segundo o Método de
 334 Tocher.

Grupo	Nº	Genótipos
1	147	1, 2, 4, 5, 6, 7, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 27, 28, 29, 30, 32, 33, 36, 37, 39, 39, 40, 41, 42, 48, 49, 51, 54, 55, 56, 57, 58, 59, 60, 62, 66, 68, 71, 73, 74, 75, 76, 78, 81, 82, 83, 84, 85, 87, 88, 89, 90, 92, 95, 96, 99, 100, 101, 102, 103, 105, 107, 111, 114, 115, 116, 118, 120, 121, 122, 123, 125, 126, 127, 128, 129, 131, 132, 133, 134, 136, 137, 139, 141, 143, 144, 145, 146, 147, 148, 149, 150, 152, 154, 157, 160, 162, 163, 164, 165, 166, 167, 170, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 185, 186, 187, 189, 190, 192, 196, 199, 202, 203, 207, 209, 210, 211, 212, 214, 215, 218, 219, 220, 221, 222, 223, 226, 228
2	8	63, 64, 65, 72, 86, 93, 97, 184
3	16	25, 35, 80, 117, 142, 158, 169, 197, 198, 201, 204, 205, 208, 213, 217, 224
4	7	13, 91, 104, 193, 194, 195, 216
5	10	9, 11, 15, 17, 31, 45, 46, 47, 52, 153
6	9	53, 67, 108, 109, 110, 119, 138, 168, 229
7	3	130, 155, 159
8	5	77, 106, 124, 140, 151
9	6	43, 44, 50, 98, 113, 188
10	4	94, 135, 171, 225
11	2	26, 70
12	2	61, 79
13	2	191, 200
14	1	3
15	1	34
16	1	24
17	1	206
18	1	112
19	1	156
20	1	227
21	1	161

335

336 Em geral, observa-se que os genótipos 3, 12, 16 e 54, apresentaram CGC
 337 favorável nas três safras. Na safra 15/16 e safrinha 16 destacaram-se os genótipos 3 e
 338 16, e na safra 16/17 os genótipos 40 e 113 (Tabela 2). Deste modo é importante salientar

339 que estes genótipos foram alocados de forma similar pelos dois métodos de
340 agrupamentos. Segundo o método de K-médias os genótipos 12, 16, 40 e 54 alocaram-
341 se no grupo 6, o genótipo 3 no grupo 1 e o genótipo 113 no grupo 4. Os três grupos
342 destacaram-se por apresentarem os valores mais altos para PG (Tabela 3 e Figura 1).
343 Pelo método de Tocher, os genótipos 12, 16, 40 e 54 alocaram-se no grupo 1, o genótipo
344 3 no grupo 14 e o genótipo 113 no grupo 9 (Tabela 4).

345 Comparando os resultados obtidos pela estimativa da CGC e pelos métodos de
346 agrupamento K-médias e Tocher, pode-se verificar a complementaridade entre os
347 mesmos. Observa-se que os genótipos 12, 16 e 54, que apresentaram CGC favorável
348 nas três safras, foram alocados no mesmo grupo (grupo 1 e 6 nos métodos Tocher e K-
349 médias, respectivamente), sendo este considerado o grupo com maior valor de PG
350 (Figura 1). É possível observar que os genótipos 40 e 113, que destacaram-se por
351 apresentar alta CGC na safra 16/17, foram alocados em grupos diferentes, porém os dois
352 grupos foram caracterizados por apresentarem alto valor de PG, sendo os grupos com
353 maior valor de PG (Figura 1). Comparando a estimativa da CGC, em relação a média das
354 três safras, com o método K-médias observa-se que os quatro genótipos que
355 apresentaram os maiores valores de estimativa de CGC (3, 16, 12 e 113) foram alocados
356 nos grupos 1, 6 e 4, respectivamente, sendo estes os grupos com alto valor de
357 produtividade (Figura 1). Os 3 genótipos com menores valores de CGC (26, 88 e 187)
358 foram alocados nos grupos 5 e 3, respectivamente, sendo estes os grupos com baixos
359 valores de PG.

360 Observa-se que a estimativa da CGC foi referente à PG dos genótipos e a alocação
361 nos grupos foi referente, além da PG, aos caracteres AP, PRE, AC, QUE, PG e PRO,

362 sendo estas de suma importância, por serem de interesse para os agricultores, além de
363 influenciarem diretamente na PG. Candido e Andrade (2008) identificaram correlações
364 genéticas altas e positivas entre altura de plantas e acamamento/quebramento e entre
365 altura de espiga e acamamento/quebramento, indicando que plantas de porte mais alto e
366 com maior altura de espiga apresentam maior incidência de acamamento e quebramento.
367 Assim, essa correlação é de grande interesse para os programas de melhoramento, pois
368 a seleção de plantas com menor incidência de acamamento pode resultar em ganhos de
369 produtividade.

370 Nas próximas etapas de programa de melhoramento deve-se buscar, dentro de
371 cada grupo, os genótipos com maior valor de CGC, para que sejam utilizados nos
372 cruzamentos e avaliações das próximas gerações, tornando-se possível selecionar
373 genótipos para futuros cruzamentos promissores, aumentando a chance de obter plantas
374 superiores em relação às características agronômicas desejáveis. Por exemplo, poderia
375 ser realizado cruzamento entre genótipos do grupo 2, com baixo valor para os caracteres
376 AP e QUE, e genótipos do grupo 6, com alto valor para caráter PG e baixo valor para
377 caráter AC, sendo possível obter genótipos mais produtivos, com menor porte e assim,
378 menor acamamento e quebramento de plantas.

379 A combinação dos resultados de CGC e análise multivariada de agrupamento é uma
380 estratégia interessante para o melhorista de milho, pois auxilia o direcionamento dos
381 cruzamentos, aumentando a probabilidade de obter genótipos superiores no programa.

382 **CONCLUSÕES**

383 1. Foi possível estimar a capacidade geral de combinação dos genótipos de
384 milho. Foi possível também selecionar genótipos com capacidade geral de

385 combinação favorável nas três safras, como os genótipos 3, 12, 16 e 54, e
 386 também genótipos específicos, com elevada capacidade geral de combinação,
 387 para a safra 15/16 e safrinha 16, genótipos 3 e 16, e para a safra 16/17,
 388 genótipos 40 e 113.

389 2. Os genótipos de milho foram alocados quanto à divergência a partir das
 390 análises multivariadas de agrupamento K-médias e Tocher. Foi possível
 391 identificar sete grupos divergentes pelo método K-médias e 21 pelo método
 392 Tocher.

393 REFERÊNCIAS

- 394 ALENCAR, B. J. de; BARROSO, L. C.; ABREU, J. F. de. Análise Multivariada de Dados
 395 no Tratamento da Informação Espacial: uma abordagem com análise de Agrupamentos.
 396 **Revista Iberoamericana de Sistemas, Cibernética e Informática**, v. 10, n. 2, p. 6-12,
 397 2013.
- 398 BERTAN, I. et al. Comparação de métodos de agrupamento na representação da
 399 distância morfológica entre genótipos de trigo. **Revista Brasileira de Agrociência**, v. 12,
 400 n. 03, p. 279-286, 2006.
- 401 BOLSON, E.; SCAPIM, C. A.; CLOVIS, L. R.; JUNIOR, A. T. A.; FREITAS, I. L. J.
 402 Capacidade combinatória de linhagens de milho avaliada por meio de testadores
 403 adaptados à safrinha. **Revista Ceres**, v. 63, n.4, p. 492-501, 2016.
- 404 BUENO, R. C.; CARVALHO, L. G.; VIANELLO, R. L.; MARQUES, J. J. G. S. M. Estudo
 405 de Rajadas de Vento e Direções Predominantes em Lavras, Minas Gerais, Por Meio da
 406 Distribuição Gama. **Ciência e Agrotecnologia**, v. 35, n. 4, p. 789-796, 2011.
- 407 CANDIDO, L. S.; ANDRADE, J. A. C. Breeding potential of maize composite Isanão VF1
 408 in small spacing in the second growing season. **Crop Breeding and Applied**
 409 **Biotechnology**, v. 8, p. 56-64, 2008.
- 410 CARDOSO AII. 2006. Dialelo entre linhagens de uma população de pepino tipo 'Caipira'.
 411 **Horticultura Brasileira** 24: 259-263.
- 412 CORTEZ, J. W.; FURLANI, C. E. A.; SILVA, R. P. Sistemas de adubação e consórcio de
 413 culturas intercalares e seus efeitos nas caracteres de colheita da cultura do milho.
 414 **Engenharia Agrícola**, v.29, n.2, p.277-287, 2009.

- 415 COSTA, N. R.; ANDREOTTI, M.; GAMEIRO, R. A.; PARIZ, C. M.; BUZETTI, S.; LOPES,
416 K. S. M. Adubação nitrogenada no consórcio de milho com duas espécies de braquiária
417 em sistema plantio direto. **Pesquisa agropecuária brasileira**, v.47, n.8, p.1038-1047,
418 2012.
- 419 Clovis LR, Scapim CA, Pinto RJB, Bolson E & Senhorinho HJC (2015) Avaliação de
420 linhagens S3 de milho por meio de testadores adaptados à Safrinha. **Revista Caatinga**,
421 28:109-120.
- 422 CRUZ, C. D. **Aplicações de algumas técnicas multivariadas no melhoramento de**
423 **plantas**. 1990. 188f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola
424 Superior de Agricultura Luiz de Queiróz, Universidade de São Paulo, Piracicaba, 1990.
- 425 CRUZ CD; REGAZZI AJ; CARNEIRO, PCS. 2012. **Modelos biométricos aplicados ao**
426 **melhoramento genético**. 4.ed. Viçosa: UFV. 514p.
- 427 CRUZ, C. D.; REGAZZI, A. J. Análise dialéctica. In.: **Modelos Biométricos Aplicados ao**
428 **Melhoramento Genético**. Ed. UFV. Viçosa. 2a ed. p. 131–285. 1997.
- 429 CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento**
430 **genético**. Viçosa: Editora UFV, 2003. 579p.
- 431 CRUZ C. D.; REGAZZI A. J. 2004. **Modelos biométricos aplicados ao melhoramento**
432 **genético**. Viçosa: UFV. 480 p.
- 433 DAVIS, R. L. **Report of the plant breeder**. Annals Reports Agricultural Experiment.
434 Puerto Rico, 1927. p. 14-15.
- 435 DEMÉTRIO, C. S.; FORNASIERI FILHO, D.; CAZETTA, J. O.; CAZETTA, D. A.
436 Desempenho de híbridos de milho submetido as diferentes espaçamentos e densidades
437 populacionais. **Pesquisa Agropecuária Brasileira**, v. 43, p.1691-1697, 2008.
- 438 EASSON, D. L.; WHITE, E. M.; PICKLES, S. J. The effects of weather, seed rate and
439 cultivar on lodging and yield in winter wheat. **Journal of Agricultural Science**, v. 121, n.
440 2, p.145-156, 1993.
- 441 ELIAS, H. T.; CARVALHO, S. P.; ANDRE. C. G. M. Comparação de testadores na
442 avaliação de famílias S2 de milho. **Pesquisa Agropecuária Brasileira**, v. 35, n. 06,
443 p.1135 1142, 2000.
- 444 FREDDI, O. S.; FERRAUDO, A. S.; CENTURION, J. F. Análise multivariada na
445 compactação de um latossolo vermelho cultivado com milho. **Revista Brasileira**
446 **Ciências Solo**, 32:953-961, 2008.

- 447 FRITSCHÉ-NETO, R.; VIEIRA, R. A.; SCAPIM, C. A.; MIRANDA, G. V.; REZENDE, L. M.
448 Updating the ranking of the coefficients of variation from maize experiments. **Acta**
449 **Scientiarum. Agronomy**, v. 34, n. 1, p. 99-101, 2012.
- 450 GOMES, L. S.; BRANDÃO, A. M.; BRITO, C. H.; MORAES, D. F.; LOPES, M. T. G.
451 Resistência ao acamamento de plantas e ao quebramento do colmo em milho tropical.
452 **Pesquisa Agropecuária Brasileira**, v. 45, n. 2, p. 140-145, 2010.
- 453 GUIMARÃES, L. J. M.; MIRANDA, G. V.; DELIMA, R. O.; MAIA, C.; OLIVEIRA, L. R. de;
454 SOUZA, L. V. de. Performance of testes with different genetic structure for evaluation of
455 maize inbred lines. **Ciência Rural**, . 42, n. 5, p. 770-776, 2012.
- 456 HALLAUER, A.R. Temperate maize and heterosis. In: COORS, J.G.; PANDEY, S. (ed.)
457 Genetics and Exploitation of Heterosis in Crops. Madison: **American Society of**
458 **Agronomy**, p.353–361. 1999.
- 459 HALLAUER, A. R.; CARENA, M. J.; MIRANDA FILHO, J. D. Quantitative genetics in maize
460 breeding. Iowa: **Springer**, 2010.
- 461 LIMA, M. L. A.; SOUZA JÚNIOR, C. L.; BENTO, D. A. V.; SOUZA, A. P.; CARLINI
462 GARCIA, L. A. Mapping QTL for grain yield and plant traits in a tropical maize population.
463 **Molecular Breeding**, v. 17, n. 3, p. 227-239, 2006.
- 464 MARCONDES, M. M.; FARIA, M. V.; MENDES, M. C.; GABRIEL, A.; NEIVERTH V.;
465 ZOCCHÉ, J. C. Breeding potential of S₄ maize lines in topcrosses for agronomic and
466 forage traits. **Acta Scientiarum. Agronomy**. v. 38, n. 3, p. 307-315, 2016.
- 467 OLIVEIRA, J. P.; CHAVES, L. C.; DUARTE, J. B.; BRASIL, E. M.; FERREIRA JÚNIOR,
468 L. T.; RIBEIRO, K. O. Teor de proteína no grão em população de milho de alta qualidade
469 protéica e seus cruzamentos. **Pesquisa Agropecuária Tropical**, v. 34, n. 1, p. 45–51.
470 2004.
- 471 PITTA, R. M.; BOIÇA JR, A. L.; JESUS, F.; TAGLIARI, S. R. A. Seleção de Genótipos
472 Resistentes de Amendoinzeiro a *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae)
473 com Base em Análises Multivariadas. **Neotropical Entomology**, Londrina, v. 39, n. 2, p.
474 260-265, 2010.
- 475 RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. de. **Experimentação em**
476 **genética e melhoramento de plantas**. Lavras: UFLA, 2000, 326 p.
- 477 RESENDE, M.D.V. Software Selegen-Reml/Blup. Embrapa Florestas, Colombo, 2015.
- 478 SAS INSTITUTE INC. SAS/STAT software 9.3. SAS Institute Inc. Cary, NC, USA, 2011.

- 479 SCAPIM, C. A. et al. Comparação de testadores na avaliação da capacidade de
480 combinação de famílias S2 de milho-pipoca. **Revista Brasileira de Milho e Sorgo**, v. 07,
481 n. 01, p.83-91, 2008.
- 482 SILVA, E. N. S.; MONTANARI, R.; PANOSSO, A. R.; CORREA, A. R.; TOMAZ, P. K.;
483 FERRAUDO, A. S. Variabilidade de atributos físicos e químicos do solo e produção de
484 feijoeiro cultivado em sistema de cultivo mínimo com irrigação. **Revista Brasileira**
485 **Ciências Solo**, 39:598-607, 2015.
- 486 SPRAGUE, G.F.; TATUM, L.A. General vs. specific combining ability in single crosses of
487 corn. **Journal of the American Society of Agronomy**, v.34, p.923-932, 1942.
- 488 VENCOVSKY, R. Herança quantitativa. In: PATERNIANI, E. (Ed.). **Melhoramento e**
489 **produção do milho no Brasil**. Campinas: Fundação Cargill, 1978. p. 122-195

1 **CAPÍTULO 3 – Correlação entre caracteres de milho**

2 (Revista Brasileira de Sementes)

3 **Correlation between traits of maize**

4 **ABSTRACT-** The objective of the research was to verify relationships between agronomic
 5 characters and characters related to maize seed quality. 30 corn topcross genotypes and
 6 one commercial control were evaluated in field and laboratory conditions. The agronomic
 7 traits evaluated in three harvests were: plant height, relative position of the spike, lodging,
 8 breaking, grain yield and prolificacy. Germination and emergency tests were carried out
 9 in the laboratory, and the first germination count, final germination, first emergency count,
 10 final emergency and emergency speed index readings were performed. Field experiments
 11 were arranged in a randomized block design with two replicates. For the laboratory
 12 experiments the design was completely randomized with four replicates. Simple linear
 13 correlation analysis (Pearson) and Bayesian networks were performed with statistical
 14 software RStudio. Pearson's linear correlation ranged from -0.511 (between breakage and
 15 productivity) to 0.987 (between first emergency count and final emergency). From the
 16 *Directed Acyclic Graph* it is possible to observe the correlation between agronomic
 17 characters and characters related to the quality of maize seeds. Thus, there is little
 18 correlation between agronomic characters and traits related to maize seed quality.

19 **Index terms** - *Zea mays*, Pearson's linear correlation, Bayesian networks.

20 **Correlação entre caracteres de milho**

21 **RESUMO** – O objetivo com a pesquisa foi verificar relações existentes entre caracteres
 22 agrônômicos e caracteres relacionados à qualidade de sementes de milho. Foram
 23 avaliados em condições de campo e de laboratório 30 genótipos *topcrosses* de milho e
 24 uma testemunha comercial. Os caracteres agrônômicos avaliados, em três safras, foram:
 25 altura de planta, posição relativa da espiga, acamamento, quebramento, produtividade
 26 de grãos e prolificidade. Em laboratório foram realizados os testes de germinação e
 27 emergência, sendo realizadas as leituras da primeira contagem de germinação,
 28 germinação final, primeira contagem de emergência, emergência final e índice de
 29 velocidade de emergência. Os experimentos em campo foram dispostos no delineamento
 30 em blocos casualizados com duas repetições. Para os experimentos em laboratório o
 31 delineamento foi inteiramente casualizado com quatro repetições. Foram realizadas as
 32 análises de correlação linear simples (Pearson) e redes bayesianas com auxílio do
 33 software estatístico RStudio. A correlação linear de Pearson variou de -0,511 (entre
 34 quebramento e produtividade) a 0,987 (entre primeira contagem de emergência e
 35 emergência final). A partir do *Directed Acyclic Graph* é possível observar a correlação
 36 entre caracteres agrônômicos e caracteres relacionados à qualidade de sementes de
 37 milho. Assim, observa-se pouca correlação existente entre os caracteres agrônômicos e
 38 os caracteres relacionados à qualidade de sementes de milho.

39

40 **Termos para indexação** - *Zea mays*, correlação linear de Pearson, Redes Bayesianas.

41

42

Introdução

43 Em programas de melhoramento genético de plantas desejam-se selecionar
44 genótipos favoráveis à produtividade usando diferentes caracteres e genótipos na
45 seleção.

46 A produtividade destaca-se como característica de suma importância para o
47 melhoramento de milho. No entanto, sabe-se que caracteres relacionados à qualidade de
48 sementes são imprescindíveis para a obtenção de estande de plantas homogêneo no
49 campo, resultando em elevada produtividade (Mambrin et al., 2015).

50 As sementes são o insumo básico da agricultura. É de suma importância o uso de
51 sementes de alta qualidade para o sucesso da lavoura, sendo estas a base para o início
52 de um programa de melhoramento de plantas. Assim, torna-se importante a seleção de
53 genótipos superiores para caracteres fisiológicas de sementes, sendo válido determinar
54 as inter-relações existentes entre caracteres agrônômicas com estes caracteres.

55 O conhecimento das relações existentes entre os caracteres de importância
56 econômica, estimadas pelas correlações, é de grande relevância para o melhoramento
57 genético, por fornecer informações que auxiliam no processo de seleção de genótipos
58 superiores. De acordo com Nogueira et al. (2012), as correlações fornecem informações
59 úteis, as quais podem facilitar o processo de seleção indireta.

60 Segundo Hallauer et al. (2010), os coeficientes de correlação simples são
61 importantes pois, a partir destes, é possível quantificar o grau de associação genético e
62 não-genético entre dois ou mais caracteres. Cruz et al. (2012) afirmaram que essas
63 associações possibilitam ganhos indiretos por seleção em caracteres correlacionados,
64 principalmente em seleção truncada em caracteres de baixa herdabilidade, aumentando
65 a eficiência na seleção.

66 Por outro lado, as redes bayesianas são metodologias que podem gerar
67 informações precisas para o estudo da correlação entre caracteres. A representação
68 gráfica das redes bayesianas pode facilitar a compreensão da relação existente entre um
69 conjunto de características, principalmente devido a propriedade da relação de

70 dependência condicional, representada pela distribuição conjunta de dados (Marques e
71 Dutra, 2002).

72 Dadas as considerações apresentadas, o objetivo com a pesquisa foi verificar
73 relações existentes entre caracteres agronômicos e caracteres relacionados à qualidade
74 de sementes de milho.

75 **Material e Métodos**

76 A pesquisa foi conduzida na área experimental da Fazenda de Ensino, Pesquisa
77 e Extensão (FEPE) e no Laboratório de Sementes do Departamento de Produção Vegetal
78 da Faculdade de Ciências Agrárias e Veterinária (FCAV/UNESP), no município de
79 Jaboticabal-SP (Latitude 21° 14' 05" S, longitude 48° 17' 09" W e altitude de 614 m acima
80 do nível do mar).

81 Foram avaliados em campo e laboratório 30 genótipos *topcrosses* de milho,
82 oriundos do cruzamento de 30 linhagens parcialmente endogâmicas (S_3) com um testador
83 de base genética ampla, e uma testemunha comercial, o híbrido DKB 390 VTPRO 2.

84 Os experimentos em campo foram dispostos no delineamento de blocos
85 casualizados com duas repetições e 31 tratamentos. Cada parcela foi constituída por
86 duas linhas de cinco metros de comprimento, com espaçamento de 0,5 m entre linhas e
87 0,33 m entre plantas, representando 64.000 plantas.ha⁻¹. Os caracteres agronômicos
88 avaliados foram: altura de planta (AP, obtida do solo até a inserção da folha bandeira, em
89 cm), posição relativa da espiga (PRE, razão entre altura de planta e altura da inserção da
90 espiga), acamamento (AC, em número de plantas/parcela inclinadas abaixo de 45° em
91 relação ao solo, transformados para $\sqrt{x+0,5}$), quebramento (QUE, em número de
92 plantas/parcela com quebramento do colmo abaixo da espiga principal, transformados
93 para $\sqrt{x+0,5}$), produtividade de grãos (PG, em kg.ha⁻¹ corrigida para umidade de 13%) e
94 prolificidade (PRO, razão entre número de espigas e estande final). A avaliação foi
95 realizada em três ambientes: primeira safra 2015/2016, segunda safra 2016 e primeira
96 safra 2016/2017. Para a realização das análises estatísticas e de correlação, foi utilizado
97 o valor da média das três safras.

98 Para a avaliação da qualidade das sementes, em laboratório, foram realizados os
99 testes de germinação e emergência. A germinação das sementes dos 30 genótipos

100 *topcrosses* e da testemunha foi conduzida com quatro repetições de 50 sementes. A
101 semeadura foi realizada entre papel umedecido com água destilada na proporção de 2,5
102 vezes o peso do papel seco. Os rolos foram mantidos em germinador regulado na
103 temperatura de 25°C. Foram realizadas a primeira contagem de germinação (G) e a
104 germinação final (GF) das plântulas normais aos 4 e 7 dias após a instalação do teste,
105 respectivamente. Os resultados foram expressos em porcentagem média.

106 A emergência em leito de areia foi conduzida em bandejas de plástico, sendo
107 semeadas 50 sementes/bandeja com quatro repetições por genótipo. O substrato foi
108 umedecido sempre que necessário e foram realizadas a primeira contagem de
109 emergência e a emergência final das plântulas até a estabilização da emergência, que
110 ocorreu 15 dias após a semeadura, cujos resultados foram expressos em porcentagem.

111 Para determinação do índice de velocidade de emergência, foram realizadas
112 contagens diárias do número de plântulas normais a partir da emergência da primeira,
113 sendo consideradas emergidas as plântulas com plúmulas visíveis, sendo o índice
114 calculado conforme Maguire (1962).

115 O delineamento para a avaliação da qualidade fisiológica de sementes em
116 laboratório foi o inteiramente casualizado, com quatro repetições.

117 Foram realizadas análises de variância para cada conjunto de caracteres
118 (agronômicos e relacionados à qualidade de sementes de milho), além da análise de
119 correlação entre os mesmos caracteres, por meio das análises de correlação linear
120 simples (Pearson) e redes bayesianas com auxílio do software estatístico RStudio.

121 **Resultados e Discussão**

122 Em relação à análise de variância para caracteres agronômicos, houve efeito
123 significativo dos genótipos para os caracteres AP, PRE, QUE e PG, indicando que os
124 genótipos diferem entre si para os mesmos. Já para a análise de variância dos caracteres
125 relacionados à qualidade de sementes, foi possível observar efeito significativo dos
126 genótipos para todos os caracteres. Tanto para os caracteres agronômicos como os de
127 sementes, os CVs apresentaram valores dentro dos padrões normais de experimentação
128 em milho, de acordo com Fritsche-Neto et al. (2012), confirmando a precisão experimental

129 e a acurácia na estimativa das médias, em ambas as avaliações, com exceção dos
130 caracteres AC, QUE e PRO, por apresentar maior influência ambiental (Tabela 1).

131 Os coeficientes de correlação linear de Pearson apresentaram elevada
132 variabilidade entre os caracteres de milho avaliados, com estimativas variando de -0,511
133 (entre QUE e PG) a 0,987 (entre E e EF) (Tabela 2). O coeficiente de correlação mede a
134 força e a direção da associação (positiva ou negativa) entre dois caracteres aleatórios. O
135 seu valor situa-se no intervalo entre -1 e 1 ($r_{xy} = -1$, correlação perfeita negativa; $-1 < r_{xy} < 0$,
136 correlação negativa; $0 < r_{xy} < 1$, correlação positiva; $r_{xy} = 1$, correlação perfeita positiva;
137 $0,2 < r_{xy} < 0,4$, correlação fraca; $0,4 < r_{xy} < 0,7$, correlação moderada; $0,7 < r_{xy} < 0,9$, correlação
138 forte), sendo que quanto mais o valor absoluto do coeficiente de correlação se aproxima
139 dos limites, mais forte é a correlação entre os caracteres.

140 Este resultado evidencia que a emergência final das sementes dos genótipos de
141 milho está diretamente associada com a primeira contagem de emergência, ou seja, é
142 possível pré-estabelecer a emergência final a partir da primeira contagem de emergência.
143 Observam-se correlação forte entre os testes de emergência (E e EF) e o IVE (0.880 e
144 0.865, respectivamente) (Tabela 2).

145 Em trabalho com seis lotes de sementes de milho híbrido, a partir da correlação
146 linear simples foi possível observar correlação alta e significativa com os testes de
147 emergência e índice de velocidade de emergência de plântulas em campo (Sbrussi e
148 Zucareli, 2015).

149 Observa-se correlação negativa entre os caracteres agrônômicos de milho (PRE
150 e AC) e o caráter PG (-0,181 e -0,099, respectivamente), demonstrando que estes
151 caracteres não estão diretamente relacionadas, ou seja, a produtividade de grãos de um
152 genótipo de milho não depende destes caracteres (Tabela 2).

153 Em relação aos caracteres relacionados à qualidade de sementes (G, EF e IVE)
154 com a PG observa-se correlação fraca ou moderada (0,475, 0,469 e 0,375,
155 respectivamente) (Tabela 2). Este resultado corrobora com estudo realizado por Cai et
156 al. (2014), onde verificaram a partir da correlação linear simples que os dados fisiológicos
157 de seis linhagens de milho não se correlacionaram com a produtividade de grãos. Fica

158 evidente a partir do maior valor de correlação de caracteres agronômicos e o caráter PG,
159 que a germinação final influenciará diretamente na produtividade de grãos (Tabela 2).

160 Os maiores valores de correlação, ou seja, correlação forte, ocorreram entre os
161 caracteres E x EF, e E x IVE, 0,987 e 0,880, respectivamente, evidenciando que a primeira
162 contagem de emergência de um campo influenciará diretamente na emergência final e no
163 índice de velocidade de emergência da parcela, por exemplo (Tabela 2).

164 As relações entre os caracteres avaliadas por meio das Redes bayesianas podem
165 ser observadas pelo *Directed Acyclic Graph* (Figura 1). As Redes bayesianas
166 representam uma análise de correlação mutuamente exclusiva, mostrando apenas o que
167 esta mais relacionado entre os caracteres estudados. Nas Redes Bayesianas, as
168 associações são dispostas de tal maneira que pode-se observar a ordem de importância
169 em termos de predição e dependência condicional da associação direta entre as
170 características avaliadas (Felipe et al., 2015).

171 É possível observar que houve correlação entre caracteres relacionados à
172 qualidade de sementes e os caracteres agronômicos (Figura 1).

173 Em relação aos caracteres de sementes, o caráter primeira contagem de
174 germinação esta diretamente relacionada com os caracteres germinação final e primeira
175 contagem de emergência (Figura 1). Sabe-se que esta relação existe pois os caracteres
176 GF e E são determinadas a partir do caráter G, ou seja, para ocorrer os caracteres GF e
177 E é preciso ocorrer o caráter G. O caráter primeira contagem de emergência (E) relaciona-
178 se diretamente com os caracteres emergência final e índice de velocidade de emergência,
179 sendo que estes caracteres são determinadas apenas quando o caráter E ocorre (Figura
180 1). Este caráter depende exclusivamente da primeira contagem de germinação para
181 ocorrer. Sendo assim, sem germinação não ocorre a emergência da semente em solo.
182 Segundo Yu et al. (2004), as redes bayesianas podem ser usadas para demonstração do
183 fenômeno biológico em dados experimentais.

184 Com relação aos caracteres agronômicos, observa-se correlação entre os
185 caracteres quebraamento, altura de planta e prolificidade com o caráter produtividade de
186 grãos. Logo, plantas prolíficas e quebradas influenciam diretamente na produtividade de
187 grãos (Figura 1).

188 Observa-se correlação entre dados agronômicos e dados de sementes. O caráter
 189 posição relativa da espiga esta diretamente relacionada com o caráter germinação final.
 190 O caráter acamamento correlaciona-se com o caráter índice de velocidade de
 191 emergência. Já o caráter germinação final está correlacionada diretamente com os
 192 caracteres produtividade de grãos e prolificidade, onde o número final de plantas de um
 193 stande influenciará diretamente no valor da produtividade (Figura 1).

194 De modo geral, os resultados observados nas Redes Bayesianas evidenciaram
 195 maior relação entre os caracteres agronômicos com os caracteres relacionados à
 196 qualidade de sementes do que os dados observados na análise de correlação linear de
 197 Pearson.

198 **Conclusões**

199 Observa-se pouca correlação existente entre os caracteres agronômicos e os
 200 caracteres relacionados à qualidade de sementes de milho.

201 **Referências**

- 202 CAI, Q.; YU, L.; YAO, W.; ZHANG, Y.; WANG, L.; CHEN, X.; DENG, J.; KANG, M.; FAN,
 203 X.. Correlation and combining ability analysis of physiological traits and some agronomic
 204 traits in maize. *Maydica* 59-2014.
- 205 CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. *Modelos biométricos aplicados ao*
 206 *melhoramento genético*. 4. ed. Viçosa: Editora UFV, 2012. 514p.
- 207 FELIPE, V. P.; SILVA, M. A.; VALENTE, B. D.; ROSA, G. J. Using multiple regression,
 208 Bayesian networks and artificial neural networks for prediction of total egg production in
 209 European quails based on earlier expressed phenotypes. *Poultry science*, v. 94, n. 4, p.
 210 772-780, 2015.
- 211 FRITSCHÉ-NETO, R.; VIEIRA, R. A.; SCAPIM, C. A.; MIRANDA, G. V.; REZENDE, L. M.
 212 Updating the ranking of the coefficients of variation from maize experiments. *Acta*
 213 *Scientiarum. Agornomy*, v. 34, n. 1, p. 99-101, 2012.
- 214 HALLAUER, A. R.; CARENA, M. J.; MIRANDA FILHO, J. B. Quantitative genetics in maize
 215 breeding. *Springer*, 2010. 580 p.
- 216 MAGUIRE, J.D. Speed of germination-aid in selection and evaluation for seedling
 217 emergence and vigor. *Crop Science*, v.2, n.1, p.176-177, 1962.

- 218 MAMBRIN, R. B.; HENNING, L. M. M.; HENNING, F. A.; BARKERT, K. A. Seleção de
219 linhagens de feijão com base no padrão e na qualidade de sementes. *Revista Caatinga*,
220 v. 28, n. 3, p. 147-156, 2015.
- 221 MARQUES, R. L.; DUTRA, I. N. Ê. S. Redes Bayesianas: o que são, para que servem,
222 algoritmos e exemplos de aplicações. Rio de Janeiro: *Coppe Sistemas*; UFRJ, 2002.
- 223 NOGUEIRA, A. P. O.; SEDIYAMA, T.; SOUSA, L. B. de; HAMAWAKI, O. T.; CRUZ, C.
224 D.; PEREIRA, D. G.; MATSUO, E. Análise de trilha e correlações entre caracteres em
225 soja cultivada em duas épocas de semeadura= Path analysis and correlations among
226 traits in soybean grown in two dates sowing. *Bioscience Journal*, v. 28, n. 6, 2012.
- 227 SBRUSSI, C. A. G., ZUCARELI, C.. Germination at high temperatures in evaluation of
228 maize vigor. *Ciência Rural*, v.45, n.10, out, 2015.
- 229 YU, J. et al. Advances to Bayesian network inference for generating causal networks
230 from observational biological data. *Bioinformatics*, v. 20, n. 18, p. 3594-3603, 2004.

ANEXO

Tabela 1 - Resumo das análises de variância individuais, com os quadrados médios e significâncias para caracteres agronômicos e relacionados à qualidade de sementes de genótipos de milho.

FV	GL	AP	PRE	AC	QUE	PG ⁺	PRO
Dados Agronômicos							
Genótipo	30	192.45**	0.0009*	0.13 ^{ns}	0.43*	23.86**	0.08 ^{ns}
Repetição	5	8209.75**	0.0062**	7.13**	13.46**	137.82**	2.05**
Erro	114	85.00	0.0005	0.14	0.27	8.70	0.09
Média		215.00	0.59	1.10	1.53	7488.01	1.21
Máximo		226.35	0.62	1.30	1.97	8551.6	1.47
Mínimo		203.77	0.57	0.80	0.94	6300.3	1.03
Testemunha		212.50	0.58	1.39	0.94	9103.9	1.02
CV(%)		4.29	3.94	34.13	34.27	12.45	24.96

FV	GL	G	GF	E	EF	IVE
Dados de Sementes						
Genótipo	30	359.25**	194.46**	178.72**	154.40**	2.84**
Repetição	3	20.65 ^{ns}	16.77 ^{ns}	56.56 ^{ns}	49.25 ^{ns}	0.59*
Erro	56	19.27	43.95	15.83	13.47	0.21
Média		88.37	90.26	90.94	91.97	10.90
Máximo		99	100	100	100	12.40
Mínimo		54	71	73.5	74.5	9.14
Testemunha		91	95	93	92.5	10.52
CV(%)		4.97	7.35	4.37	3.99	4.17

^{ns}, * e **, não significativo, significativo a 5% e significativo a 1%. + dividido por 100000.

Tabela 2 – Correlação entre os caracteres altura de planta (AP, em cm), posição relativa da espiga (PRE, em cm), acamamento (AC, número de plantas/parcela), quebramento (QUE, número de plantas/parcela), produtividade de grãos (PG, kg.ha⁻¹), prolificidade (PRO, número de espigas/estande final), primeira contagem de germinação (G, % de plântulas normais), germinação final (GF, % de plântulas normais), primeira contagem de emergência (E, % de plântulas visíveis), emergência final (EF, % de plântulas visíveis) e índice de velocidade de emergência (IVE)) de 31 genótipos de milho.

	AP	PRE	AC	QUE	PG	PRO	G	GF	E	EF	IVE
AP	-	-0,120	0,005	0,181	-0,421	0,134	-0,115	-0,188	-0,234	-0,217	-0,147
PRE		-	-0,177	-0,094	-0,181	0,051	-0,033	-0,348	-0,173	-0,209	-0,120
AC			-	-0,044	-0,099	-0,138	0,082	0,054	0,139	0,131	0,296
QUE				-	-0,511	0,396	-0,167	0,011	0,023	0,068	0,066
PG					-	-0,435	0,475	0,561	0,515	0,469	0,375
PRO						-	0,007	0,004	-0,009	0,022	0,080
G							-	0,827	0,815	0,778	0,748
GF								-	0,830	0,827	0,664
E									-	0,987	0,880
EF										-	0,865

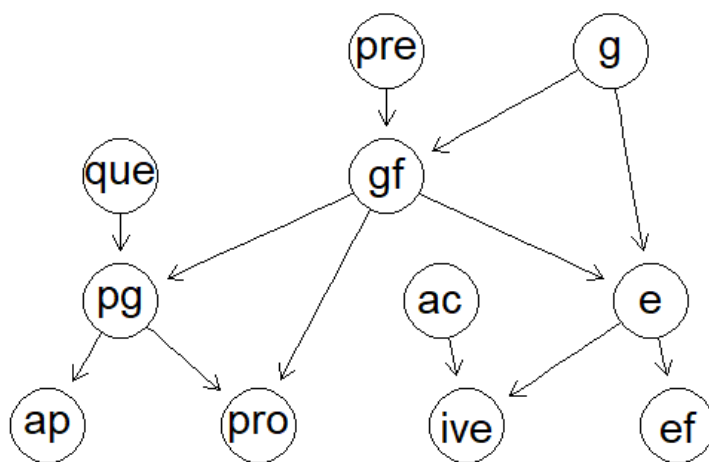


Figura 1 – Rede bayseana formada entre os caracteres de milho: altura de planta (AP), posição relativa da espiga (PRE), acamamento (AC), quebramento (QUE), produtividade de grãos (PG), prolificidade (PRO), primeira contagem de germinação (G), germinação final (GF), primeira contagem de emergência (E), emergência final (EF) e índice de velocidade de emergência (IVE).