

**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS – UFGD**

**DIVERGÊNCIA GENÉTICA ENTRE LINHAGENS  
PARCIALMENTE ENDOGÂMICAS DE MILHO PIPOCA POR  
MÉTODOS MULTIVARIADOS E REDES NEURAS ARTIFICIAIS**

**PRISCILA CARVALHO DA SILVA**

**DOURADOS  
MATO GROSSO DO SUL  
2018**

**DIVERGÊNCIA GENÉTICA ENTRE LINHAGENS  
PARCIALMENTE ENDOGÂMICAS DE MILHO PIPOCA POR  
MÉTODOS MULTIVARIADOS E REDES NEURAIS ARTIFICIAIS**

**PRISCILA CARVALHO DA SILVA**  
Biotecnologista

Orientador: Prof<sup>o</sup>. Dr<sup>o</sup> Manoel Carlos Gonçalves

Co-orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Liliam Silvia Candido

Dissertação apresentada à Universidade Federal da Grande Dourados, como parte das exigências do Programa de Pós-Graduação em Agronomia - Produção Vegetal, para obtenção do título de Mestre.

DOURADOS  
MATO GROSSO DO SUL  
2018

**Dados Internacionais de Catalogação na Publicação (CIP).**

S586d Silva, Priscila Carvalho Da

Divergência genética entre linhagens parcialmente endogâmicas de milho pipoca por métodos multivariados e redes neurais artificiais / Priscila Carvalho Da Silva – Dourados, MS: UFGD, 2018.

86f. : il. ; 30 cm.

Orientador: Manoel Carlos Gonçalves

Co-orientadora: Liliam Silvia Candido

Dissertação (Mestrado em Agronomia) - Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados.

1. Zea Mays L. everta. 2. dissimilaridade genética.  
3. métodos de agrupamento. 4. mapas auto-organizáveis de Kohonen. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

**©Direitos reservados. Permitido a reprodução parcial desde que citada a fonte.**

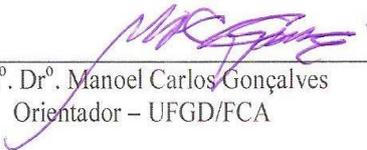
**DIVERGÊNCIA GENÉTICA ENTRE LINHAGENS PARCIALMENTE  
ENDOGÂMICAS DE MILHO PIPOCA POR MÉTODOS MULTIVARIADOS E  
REDES NEURAIS ARTIFICIAIS**

Por:

**Priscila Carvalho da Silva**

Dissertação apresentada como parte dos requisitos exigidos para obtenção do título de  
Mestre em Agronomia

Aprovado em: 08/02/2018



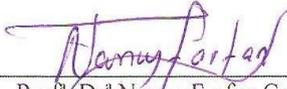
---

Prof. Dr.º Manoel Carlos Gonçalves  
Orientador – UFGD/FCA



---

Dr.º Bruno do Amaral Crispim  
Membro da Banca – UFGD/FCBA



---

Prof.ª Dr.ª Nancy Farfan Carrasco  
Membro da Banca – IFMS

*Nenhum obstáculo é tão grande, se sua fé e vontade de vencer for maior!*

## **Dedico**

*À Deus, a Jesus e a nossa senhora Aparecida.*

*Aos meus pais, Sueli Nascimento Carvalho Silva e José Carlos da Silva, por terem dedicado suas vidas a me criarem e nunca medirem esforços em meu favor, em todos os momentos, para que eu pudesse concluir mais essa fase. Nunca terei palavras para agradecer tamanho amor, zelo, incentivo, esforço e dedicação.*

*Aos meus avós Tereza Maria Nascimento Carvalho e Sebastião Carvalho, pelo exemplo de carinho, humildade e generosidade.*

*A todos de minha família, o apoio e carinho de vocês foi essencial nessa caminhada.*

*Mãe, obrigada por sempre zelar e interceder por mim em suas orações!*

## AGRADECIMENTOS

À Deus, em primeiro lugar, pela oportunidade, por iluminar e abençoar essa caminhada, por seu amor incondicional, por me dar forças e serenidade em todos os momentos que mais precisei.

Aos meus pais, por me apoiarem incondicionalmente, por suas lutas, suas histórias de vida, por respeitarem minhas decisões, por serem o meu maior exemplo de amor, respeito e perseverança.

A Universidade Federal da Grande Dourados, ao programa de pós graduação em Produção Vegetal - FCA. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro concedido.

A todos os professores que compartilharam de seus conhecimentos e contribuíram para minha formação profissional.

Em especial ao meu orientador professor Dr. Manoel Carlos Gonçalves, pela confiança, paciência, ensinamentos, apoio e dedicação. Deixo o meu muito obrigada!

A professora Liliam Silvia Candido, por me acolher desde o primeiro ano que ingressei na universidade, por todos os ensinamentos profissionais e pessoais, sendo uma profissional, professora e amiga incrível. Pessoa na qual pra sempre vou me espelhar. Obrigada por tudo.

Ao professor Cosme Damião Cruz pelo apoio e tempo disponibilizado para me auxiliar nas análises estatísticas deste trabalho.

Aos membros da banca Prof<sup>a</sup>. Dr<sup>a</sup>. Nancy Farfan Carrasco e Dr<sup>o</sup>. Bruno do Amaral Crispim pela disponibilidade, correções, sugestões e contribuições.

A todos do grupo de pesquisa de Melhoramento e Biotecnologia Vegetal (GMBV), estes que se tornaram uma família, suas contribuições foram fundamentais para a realização deste trabalho. Aos funcionários da FAECA, em especial ao seu Jesus, por todo o carinho, dedicação e prestatividade.

Ao meu amigo Vander Berres, por toda a amizade, por todo o auxílio, paciência e compreensão em nossos trabalhos de grupo e experimentos, com certeza sua mão amiga fez o caminho ser mais leve.

A Renata Marchiori, um presentinho em minha vida, com um amor imenso, que mesmo com a distância sempre esteve presente com suas palavras amigas, carinho e apoio. A Caroline Alves Gabbi, por ser uma pessoa com um caráter e um coração sem tamanho, por todo o carinho e apoio.

As minhas amigas de longa data, Simoni Santos e Lorraine Venâncio, pelo apoio, amor e amizade incondicional. A minha amiga de pós graduação Thamiris Barbizian, por todo o apoio, motivação e companhia. Ao meu amigo Mailson Viera Jesus, por todo o carinho e prestatividade.

A todas as pessoas que direta ou indiretamente contribuíram para realização deste trabalho de dissertação, e em meu mestrado.

Deixo aqui registrada minha gratidão a todos vocês.

## SUMÁRIO

	PÁGINA
LISTA DE FIGURAS.....	I
LISTA DE TABELAS.....	III
RESUMO.....	V
ABSTRACT.....	VI
<b>1. INTRODUÇÃO</b> .....	1
<b>2. OBJETIVOS</b> .....	3
<b>3. REVISÃO DE LITERATURA</b> .....	4
3.1 Aspectos evolutivos, morfológicos e socioeconômicos da cultura do milho pipoca.....	4
3.2 Melhoramento genético da cultura do milho pipoca.....	6
3.3 Descritores morfológicos e agronômicos do milho pipoca.....	8
3.4 Análise de divergência genética.....	12
3.5 Medidas de dissimilaridade e métodos para a estimação da divergência genética.....	14
3.5.1 Técnicas de agrupamento.....	15
3.5.1.1 Variáveis canônicas.....	16
3.5.2 Redes neurais artificiais.....	17
3.6 Importância relativa dos caracteres.....	21
<b>4. MATERIAL E MÉTODOS</b> .....	22
4.1 Implantação dos experimentos.....	22
4.2 Avaliações das linhagens.....	24
4.3 Análises estatísticas dos dados.....	25
4.3.1 Análises multivariadas.....	27
4.3.1.1 Distância generalizada de Mahalanobis.....	27
4.3.1.2 Método de otimização de Tocher.....	27
4.3.1.3 Método hierárquico de ligação média não ponderada entre grupos (UPGMA) e vizinho mais distante.....	28
4.3.1. Dispersão gráfica de variáveis canônicas.....	29
4.3.2 Redes neurais artificiais.....	30
4.3.3 Importância relativa dos caracteres.....	32
<b>5. RESULTADOS E DISCUSSÃO</b> .....	33

5.1. Experimento linhagens $S_2$ , safra I (2014/2015).....	33
5.1.1 Análises univariadas.....	33
5.1.2. Análises multivariadas.....	34
5.1.2.1 Distância generalizada de Mahalanobis.....	34
5.1.2.3 Análise de agrupamentos pelos métodos de Tocher, ligação média não ponderada entre grupos (UPGMA) e vizinho mais distante.....	37
5.1.2.4 Dispersão gráfica de variáveis canônicas.....	42
5.1.3 Redes neurais artificiais.....	44
5.1.4 Importância relativa dos caracteres.....	48
5.1.5 Considerações experimento linhagens $S_2$ (safra I).....	50
5.2. Experimento linhagens $S_3$ , safra II (2016/2017).....	51
5.2.1. Análises univariadas.....	51
5.2.2 Análises multivariadas.....	52
5.2.2.1 Distância generalizada de Mahalanobis.....	52
5.2.2.2 Análise de agrupamentos pelos métodos de Tocher, ligação média não ponderada entre grupos (UPGMA) e vizinho mais distante.....	55
5.2.2.3 Dispersão gráfica de variáveis canônicas.....	59
5.2.3 Redes neurais artificiais.....	61
5.2.4 Importância relativa dos caracteres.....	65
5.2.5 Considerações experimento linhagens $S_3$ (safra II).....	67
5.3 Considerações finais.....	68
<b>6. CONCLUSÕES.....</b>	<b>71</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>72</b>
<b>8. ANEXOS.....</b>	<b>84</b>

## LISTA DE FIGURAS

	PÁGINA
<b>Figura 1.</b> Modelo da estrutura de camadas de uma rede neural artificial.....	18
<b>Figura 2.</b> Modelo da estrutura de camadas de um mapa auto-organizável de Kohonen.....	20
<b>Figura 3.</b> Climograma representativo dos dados meteorológicos coletados em campo segundo a estação climatológica da Embrapa Agropecuária Oeste (CPAO) de Dourados/MS, para a safra I nos períodos de setembro à janeiro de 2014/2015.....	22
<b>Figura 4.</b> Climograma representativo dos dados meteorológicos coletados em campo segundo a estação climatológica da Embrapa Agropecuária Oeste (CPAO) de Dourados/MS, para a safra II nos períodos de novembro à março de 2016/2017.....	23
<b>Figura 5.</b> Arquitetura da rede neural artificial Kohonen composta por n acessos e m caracteres na camada de entrada, e por k neurônios na camada competitiva.....	31
<b>Figura 6.</b> Dendrograma representativo da dissimilaridade genética entre 20 linhagens S <sub>2</sub> de milho pipoca, obtidas pela ligação média não ponderada entre grupos (UPGMA), a partir da medida de dissimilaridade generalizada de Mahalanobis.....	40
<b>Figura 7.</b> Dendrograma representativo da dissimilaridade genética entre 20 linhagens S <sub>2</sub> de milho pipoca, obtidas pela ligação completa dos vizinhos mais distantes, a partir da medida de dissimilaridade generalizada de Mahalanobis.....	41
<b>Figura 8.</b> Dispersão de escores de 20 linhagens S <sub>2</sub> de milho em relação a três variáveis canônicas (VC1, VC2 e VC3), tendo como base em 15 caracteres morfoagronômicos.....	44
<b>Figura 9.</b> Mapa auto-organizável de Kohonen para nove classes de classificação, formadas por meio de rede neural artificial.....	45
<b>Figura 10.</b> Mapa auto-organizável de Kohonen, classificação das observações dos agrupamentos das linhagens nas classes formadas por meio de rede neural artificial.....	46
<b>Figura 11.</b> Contribuição relativa de quinze caracteres morfoagronômicos de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP), avaliados para divergência (D <sup>2</sup> ) entre 20 linhagens S <sub>2</sub> de milho pipoca, na safra de 2014/2015, em Dourados/MS.....	49

<b>Figura 12.</b> Dendrograma representativo da dissimilaridade genética entre 20 linhagens S <sub>3</sub> de milho pipoca, obtidas pela ligação média não ponderada entre grupos (UPGMA), a partir da medida de dissimilaridade generalizada de Mahalanobis.....	58
<b>Figura 13.</b> Dendrograma representativo da dissimilaridade genética entre 20 linhagens S <sub>3</sub> de milho pipoca, obtidas pela ligação completa do vizinho mais distante, a partir da medida de dissimilaridade generalizada de Mahalanobis.....	58
<b>Figura 14.</b> Dispersão de escores de 20 linhagens S <sub>3</sub> de milho em relação a três variáveis canônicas (VC1, VC2 e VC3), tendo como base em 15 caracteres morfoagronômicos.....	60
<b>Figura 15.</b> Mapa auto-organizável de Kohonen para nove classes de classificação, formadas por meio de rede neural artificial.....	62
<b>Figura 16.</b> Mapa auto-organizável de Kohonen, classificação das observações dos agrupamentos das linhagens nas classes formadas por meio de rede neural artificial.....	63
<b>Figura 17.</b> Contribuição relativa de quinze caracteres morfoagronômicos de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP), avaliados para divergência (D <sup>2</sup> ) entre 20 linhagens S <sub>3</sub> de milho pipoca, na safra de 2016/2017, em Dourados/MS.....	66

## LISTA DE TABELAS

PÁGINA

<b>Tabela 1.</b> Resumo das análises de variância para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S <sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.....	33
<b>Tabela 2.</b> Estimativas da dissimilaridade (D <sup>2</sup> ) para as 20 combinações de linhagens mais similares e mais divergentes, respectivamente, com base na avaliação de 15 caracteres morfoagronômicos quantitativos, em 20 linhagens S <sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.....	36
<b>Tabela 3.</b> Agrupamento pelo método de otimização de Tocher, com base na distância generalizada de Mahalanobis de 20 linhagens S <sub>2</sub> de milho pipoca, na safra de 2014/2015, em Dourados/MS.....	37
<b>Tabela 4.</b> Distâncias médias intra e intergrupos estimadas pelo método de otimização de Tocher, com base na dissimilaridade entre 20 linhagens S <sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.....	39
<b>Tabela 5.</b> Estimativas das variâncias (autovalores), variâncias percentuais acumuladas das variáveis canônicas, visando estimar a dissimilaridade genética entre 20 linhagens S <sub>2</sub> de milho milho pipoca, na safra de 2014/2015, em Dourados/MS.....	42
<b>Tabela 6.</b> Resumo das análises de variância para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S <sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados/MS.....	52
<b>Tabela 7.</b> Estimativas da dissimilaridade (D <sup>2</sup> ) para as 20 combinações de linhagens mais similares e mais divergentes, respectivamente, com base na avaliação de 15 caracteres morfoagronômicos quantitativos, em 20 linhagens S <sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados/MS.....	54
<b>Tabela 8.</b> Agrupamento pelo método de otimização de Tocher, com base na distância generalizada de Mahalanobis de 20 linhagens S <sub>3</sub> de milho pipoca, na safra de 2016/2017, em Dourados/MS.....	55
<b>Tabela 9.</b> Distâncias médias intra e intergrupos estimadas pelo método de otimização de Tocher, com base na dissimilaridade entre 20 linhagens S <sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados/MS.....	56

<b>Tabela 10.</b> Estimativas das variâncias (autovalores), variâncias percentuais acumuladas das variáveis canônicas, visando estimar a dissimilaridade genética entre 20 linhagens S <sub>3</sub> de milho pipoca, na safra de 2016/2017, em Dourados/MS.....	60
---	----

SILVA, P.C. **DIVERGÊNCIA GENÉTICA ENTRE LINHAGENS PARCIALMENTE ENDOGÂMICAS DE MILHO PIPOCA POR MÉTODOS MULTIVARIADOS E REDES NEURAI ARTIFICIAIS**, 2018. Dissertação (Mestrado em Agronomia) - Universidade Federal da Grande Dourados, Dourados-MS.

## RESUMO

No processo de seleção de genitores para cruzamentos dialélicos, visando à produção de híbridos, a avaliação da divergência genética é muito utilizada e pode ser mensurada por diferentes metodologias. As análises multivariadas são amplamente aplicadas para a avaliação da dissimilaridade genética, destacando-se os métodos de agrupamentos de Tocher, ligação média não ponderada entre grupos (UPGMA) e variáveis canônicas. Recentemente, tem-se empregado as técnicas de redes neurais artificiais, utilizando o modelo de mapa de Kohonen visando à organização da diversidade genética. Este trabalho foi realizado com o objetivo de estimar a divergência genética entre 20 linhagens parcialmente endogâmicas de milho pipoca, por métodos multivariados e redes neurais artificiais, visando à identificação de linhagens promissoras para cruzamentos dialélicos. Os experimentos foram conduzidos, em área de campo, na fazenda experimental da Universidade Federal da Grande Dourados, em Dourados/MS. Na safra I (ano agrícola 2014/2015) e safra II (ano agrícola 2016/2017). Vinte linhagens  $S_2$  (safra I) e  $S_3$  (safra II) de milho pipoca foram instaladas no delineamento em blocos casualizados, com três repetições. Foram avaliados os caracteres: dias para o florescimento feminino (FF) e masculino (FM), altura de plantas (AP), e da inserção da primeira espiga (AE), diâmetro de colmo (DC), prolificidade (PROLIF), plantas quebradas e acamadas (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espigas (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de cem grãos (M100), produtividade de grãos (PROD), capacidade de expansão dos grãos (CE) e volume de pipoca por hectare (VP). Foram realizadas as análises estatísticas multivariadas para a importância relativa dos caracteres, métodos de agrupamentos de Tocher, UPGMA, vizinho mais distante, variáveis canônicas e a técnica de redes neurais artificiais com o mapa de Kohonen. As escolhas das linhagens foram baseadas de forma geral pelas classificações dos métodos e contribuição dos caracteres. As linhagens  $S_2$  mais divergentes e potenciais identificadas foram as linhagens 9 (com menor AP, PROD e VP), 12 (com maior NFG e NEAP), 15 (com maior PROD, VP, CE e PROLIF) e 6 (com maior DC e menor FF e FM); e as linhagens  $S_3$  foram a 2 (menores CE e VP), 3 (maior CE), 10 (maiores PROD, VP, M100 e NEAP) e 8 (maior DE, FF e FM). Os caracteres que apresentaram a maior contribuição relativa para a divergência e diferenciação das linhagens, para as linhagens  $S_2$  foram a PROD, VP, M100 e CE; e para as linhagens  $S_3$  os caracteres CE, VP, AP e NEAP. Entre os métodos multivariados, as variáveis canônicas e o UPGMA, apresentaram a maior semelhança entre si para os agrupamentos. Entre os agrupamentos dos métodos multivariados e de redes neurais artificiais dos mapas de Kohonen, o maior relacionamento ocorreu entre os resultados apresentados pelas variáveis canônicas. A composição genética distinta das linhagens em  $S_2$  e  $S_3$  e as diferenças ambientais dos anos agrícolas avaliados, influenciaram nas diferenças obtidas em relação à divergência genética nos dois experimentos.

**Palavras-Chaves:** *Zea Mays* L. everta; dissimilaridade genética; métodos de agrupamento; mapas auto-organizáveis de Kohonen.

SILVA, P.C. **PARTIALLY ENDOGAMIC POPCORN LINEAGES GENETIC DIVERGENCE BY MULTIVARIATE METHODS AND ARTIFICIAL NEURAL NETWORKS**, 2018. Dissertação (Mestrado em Agronomia) - Universidade Federal da Grande Dourados, Dourados-MS.

### ABSTRACT

In the selecting genitors process for diallel crosses, aiming to the production of hybrids, the evaluation of genetic divergence is widely used and can be measured by different methodologies. The multivariate analyzes is widely applied for the evaluation of genetic dissimilarity, highlighting the Tocher clustering method, unweighted mean linkage between groups (UPGMA) and canonical variables. Recently, the techniques of artificial neural networks have been used, using the map model of Kohonen aiming at the organization of genetic diversity. This work was carried out with the objective of estimating the genetic divergence among 20 partially endogenous popcorn strains by multivariate methods and artificial neural networks, aiming the identification of promising lines for diallel crosses. The experiments were conducted, on a field area, at Big Dourados Federal University experimental farm, in Dourados - Ms. In harvest I (agricultural year 2014/2015) and harvest II (agricultural year 2016/2017). Twenty popcorn  $S_2$  lineages (crop I) and  $S_3$  (crop II) were installed in the randomized block design, with three replications. The following characteristics were evaluated: days for female and male flowering (FF; FM), plant height (AP), and first ear insertion (AE), stem diameter (DC), prolificacy (PROLIF), broken plants (AQ), number of spikes autofecunded per plot (NEAP), ear length (CPE), spike diameter (DE), number of grains rows (NFG), hundred grain mass (M100), grain yield (PROD), grain expansion capacity (CE) and popcorn volume per hectare (PV). Multivariate statistical analyzes were performed for the relative importance of the characters, grouping methods of Tocher, UPGMA, more distant neighbor, canonical variables and the technique of artificial neural networks with the Kohonen map. Lineage choices were generally based on classifications of the methods and characters contribution. The most divergent and potential strains identified in  $S_2$  lineages were 9 (with lower AP, PROD and VP), 12 (with higher NFG and NEAP), 15 (with higher PROD, VP, CE and PROLIF) and 6 (with higher DC and lower FF and FM); and in the  $S_3$  lineages 2 (lowest CE and VP), 3 (highest CE), 10 (highest PROD, VP, M100 and NEAP) and 8 (higher DE, FF and FM). The characters that presented the greatest relative contribution to the divergence and differentiation of the lineages, for the  $S_2$  lineages were the PROD, VP, M100 and CE; and for the  $S_3$  lineages the characters CE, VP, AP and NEAP. Among the multivariate methods, the canonical variables and the UPGMA presented the greatest similarity among groups. Among the groupings of multivariate methods and artificial neural networks of Kohonen maps, the greatest relationship occurred among the results presented by canonical variables. The genetic distinct composition from the  $S_2$  and  $S_3$  lineages and the environmental differences of the evaluated agricultural years influenced the differences obtained in relation to the genetic divergence in the two experiments.

**Keywords:** Zea Mays L. everta; genetic dissimilarity; cluster methods; self-organizing maps of Kohonen.

## 1. INTRODUÇÃO

O milho pipoca é uma cultura de alto valor agregado, mas em diversas regiões do país, sobretudo no estado de Mato Grosso do Sul, pouco se tem realizado o cultivo. O monopólio estabelecido pelas empresas empacotadoras dificultam o acesso dos produtores com menores condições tecnológicas à cultura, aliado a falta de um maior número de cultivares adaptadas as condições climáticas e abióticas da região.

O desenvolvimento de programas de melhoramento de milho pipoca, que visem o desenvolvimento de populações melhoradas e/ou híbridos adaptados às condições brasileiras, podem contribuir para diminuir a dependência da importação dos grãos, além de colaborar com o desenvolvimento agrícola.

No melhoramento do milho pipoca, o principal atributo de interesse, além da produtividade e de boas características agrônômicas é que os genótipos apresentem um alto índice de capacidade de expansão dos grãos (Moterle et al., 2012).

No desenvolvimento de cultivares alógamas do tipo híbrido, podem ser realizados cruzamentos entre linhagens parcialmente endogâmicas, ou seja, que não possuem *locus* completamente homozigóticos. A vantagem dos híbridos advindos de linhagens parcialmente endogâmicas se dá pela obtenção em um menor tempo e conseqüentemente menor custo agregado.

A avaliação da divergência genética que permite o conhecimento da diversidade genética entre as linhagens da população de trabalho, facilita a identificação de genitores com maior potencial heterótico para os cruzamentos dialélicos (Rotili et al., 2012). Os esforços devem ser concentrados nas combinações entre os genótipos mais divergentes e promissores, aumentando as chances de obtenção de linhagens superiores nas gerações segregantes.

Para a determinação da divergência genética, podem ser utilizados métodos biométricos, para estimar a dissimilaridade, por meio de técnicas de análise multivariadas. Dentre os métodos destacam-se, as variáveis canônicas e os métodos de agrupamento, como o método de otimização de Tocher e o hierárquico de ligação média não ponderada entre grupos (UPGMA). A dissimilaridade pode ser quantificada por meio de variáveis quantitativas por intermédio da distância Euclidiana e, principalmente da distância generalizada de Mahalanobis (Cruz et al., 2011).

Recentemente, também tem-se empregado o uso de técnicas de redes neurais artificiais para realizar a classificação de genótipos. Redes neurais artificiais (RNA's) são técnicas computacionais que apresentam um modelo inspirado na estrutura neural

do cérebro humano e que adquirem conhecimento por meio de experiência (Haykin, 2008).

No melhoramento genético de plantas, entre os modelos de RNA, pode-se destacar para uso o modelo de rede do Mapas auto-organizáveis de Kohonen (SOM), que é indicado para agrupamento de genótipos e reconhecimento de padrões. Possui a vantagem de não exigir do pesquisador a aplicação de um conhecimento prévio (origem e organização) dos genótipos em estudo, e poder fornecer subsídios para a atualização ou mesmo formulação de novas teorias acerca do problema (Barbosa et al., 2011), o que possibilitou seu emprego no trabalho em questão.

Nesse sentido, o investimento em estudo de metodologias que possibilitem a identificação de linhagens divergentes e promissoras, para potenciais cruzamentos dialélicos, torna-se de fundamental importância para o melhoramento genético da cultura do milho pipoca, podendo auxiliar, a obtenção de novas cultivares, adaptadas às diferentes condições edafoclimáticas do Brasil, e particularmente para a região Sul do estado de Mato Grosso do Sul. A maior disponibilidade de cultivares pode propiciar aos produtores rurais locais a possibilidade de agregar valor com a comercialização do produto final gerado, e conseqüentemente contribuir para o aumento da economia estadual.

## **2. OBJETIVOS**

### **2.1 GERAL**

Estimar a divergência genética entre 20 linhagens parcialmente endogâmicas de milho pipoca, por métodos multivariados e redes neurais artificiais, visando à identificação de linhagens promissoras para cruzamentos dialélicos.

### **2.2 ESPECÍFICOS**

Avaliar a dissimilaridade genética entre as 20 linhagens parcialmente endogâmicas de milho pipoca em diferentes ciclos de autofecundação ( $S_2$  e  $S_3$ ).

Avaliar o relacionamento dos agrupamentos, entre os métodos de análise multivariada de otimização de Tocher, hierárquico ligação média não ponderada entre grupos (UPGMA) e vizinho mais distante, dispersão gráfica de variáveis canônicas e redes neurais artificiais por meio dos mapas auto-organizáveis de Kohonen.

Identificar entre os caracteres avaliados, aqueles que mais contribuíram para a diferenciação das linhagens avaliadas.

### 3. REVISÃO DE LITERATURA

#### 3.1 Aspectos evolutivos, morfológicos e socioeconômicos da cultura do milho pipoca

O milho pipoca [*Zea mays* L. var. *everta* (Sturtey) L. H. Bailey] é originário da subfamília Panicoide, tribo Maydeae, gênero *Zea* e espécie *Zea mays* L., de forma semelhante ao milho comum. Estudos apontam o milho pipoca como sendo o precursor da espécie, pois, há relatos de que a evolução teria ocorrido no sentido do milho pipoca para o milho comum (Paterniani et al., 2000). Ambos têm em registro, como principal e mais antigo ancestral o *teosinte*, o qual produz uma espiga pequena com grãos que quando expostos ao calor estouravam, dando o aspecto da pipoca (Paterniani et al., 2000; Zinsly e Machado, 1987).

Em relação à arquitetura e à fisiologia das plantas, quando comparadas ao milho comum, geralmente as cultivares de milho pipoca possuem menor vigor, um sistema radicular menos desenvolvido, o que torna a planta mais suscetível, ao acamamento e à seca (Hallauer, 2010). Apresentam maior prolificidade de espigas e grãos menores (Zinsly e Machado, 1987). Entretanto, a principal diferença do milho pipoca em relação ao milho comum consiste na capacidade do grão de se expandir dando origem à pipoca (Brunson, 1937; Hallauer e Carena, 2009). A expansão do grão ocorre quando este é submetido à pressão de 135 psi, decorrente do aquecimento em torno de 180 °C (Hoseney et al., 1983).

O milho pipoca apresenta alta demanda do produto final no mercado e alto valor agregado na semente. O valor comercial da saca de 60 kg do milho comum gira em torno de R\$ 20 reais, o milho pipoca tem o preço médio saca de 30 kg a R\$ 105 reais (CEASA, 2017). Atualmente, os campos de produção da cultura vêm aumentando gradativamente no Brasil. Isso deve-se ao maior consumo pelo produto *in natura* ou na forma de produtos industrializados (Moterle et al., 2012).

Cerca de 60% da produção brasileira de milho pipoca está concentrada no estado de Mato Grosso (IBGE, 2017). Os municípios de Campo Novo dos Parecis-MT e Nova Prata-RS são atualmente os maiores produtores de milho pipoca do país, regiões onde atuam as grandes empresas empacotadoras. Contudo, a produção nacional está muito aquém da demanda.

O consumo nacional do alimento esta em torno de 80 mil toneladas, no entanto, ainda se observa alta dependência estrangeira, uma vez que 75% desse mercado corresponde a sementes do milho pipoca americano, importado principalmente da

Argentina, para suprir a demanda do produto pelas empresas empacotadoras (Carvalho et al., 2015; Sawazaki, 2010).

A dependência do mercado estrangeiro e a falta de cultivares adaptadas às diferentes condições edafoclimáticas do Brasil, são um dos maiores problemas enfrentados pelos produtores de milho pipoca. Contudo, a escassez de cultivares no mercado que reúnam características agronômicas favoráveis e com alto índice de capacidade de expansão é considerado o principal entrave para o crescimento do cultivo de milho pipoca no país (Rangel et al. 2011; Moterle et al. 2012).

O Centro-Oeste brasileiro tem a economia baseada no setor agrícola, e Mato Grosso do Sul é o estado de maior crescimento econômico da região, principalmente considerando o pequeno e médio produtor. Em se tratando da região, tem-se uma alta demanda pelo produto, e as condições climáticas dão a possibilidade de obtenção de boas produtividade, assim, o cultivo do milho pipoca torna-se uma opção viável para aumentar a renda dos agricultores, principalmente de pequeno e médio porte.

Segundo o Registro Nacional de Cultivares do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), no Brasil existem oitenta e nove cultivares registradas (MAPA, 2017). Dentre estas, apenas dezesseis estão disponíveis para os agricultores não conveniados as empresas empacotadoras, sendo oito disponibilizadas pela UENF (Universidade Estadual Norte Fluminense) e seis pelo IAC (Instituto Agrônomo de Campinas). A grande maioria dos registros pertence a empresas privadas, como exemplo, a Yoki Alimentos S/A e Pipolino Indústria Comércio LTDA.

As empresas empacotadoras, portadoras dos registros, estabelecem um contrato com os produtores para o fornecimento das sementes, geralmente provenientes de variedades e híbridos de origem americana, onde posteriormente, se realiza a compra parcial ou total da produção, gerando um monopólio no processo (Oliveira, 2016). Desses fatores, resulta a atual falta de cultivares comerciais disponíveis para os agricultores que não possuem contrato de compra e venda no mercado junto às empresas empacotadoras, acarretando uma problemática, no que se refere à comercialização e cultivo do milho pipoca no Brasil.

Desse modo, o investimento em programas de melhoramento de milho pipoca, para o desenvolvimento de novas cultivares adaptadas as regiões do país, garantindo alta produtividade e qualidade, torna-se indispensáveis para estimular e aumentar sua produção. Diminuindo assim, a dependência da importação da pipoca e

consequentemente aumentando economia e a rentabilidade da cultura para o agricultor (Leonello et al., 2009; Vilarinho et al., 2002).

### **3.2 Melhoramento genético da cultura do milho pipoca**

No melhoramento de milho pipoca, deve-se levar em consideração os interesses dos agricultores e dos consumidores (Brunson e Smith, 1945). Ao agricultor, interessa a elevada produtividade e bons caracteres agrônômicos. Ao consumidor, a alta capacidade de expansão (CE), que confere à pipoca melhor textura e maciez. As principais características das plantas e grãos para o melhoramento genético são a alta produtividade, o baixo acamamento e quebramento das plantas, a alta resistência às doenças e às pragas, a alta capacidade de expansão e boas características organolépticas como maciez, sabor, aroma e cor da flor da pipoca (Alexander e Creech, 1977).

O conhecimento das associações existentes entre os principais caracteres de interesse, contribui para obtenção de ganhos e para o sucesso da seleção dos genótipos superiores. A capacidade de expansão e a produtividade de grãos, apresentam frequentemente correlação fenotípica e genética negativa, ou seja, à medida que se obtém ganhos com a seleção baseada em um dos caracteres, existe a possibilidade de perda no outro, dificultando a obtenção de genótipos com ambas características favoráveis (Amaral Junior et al., 2016; Cabral et al., 2016). Deste modo, os programas de melhoramento de milho pipoca têm como principal desafio desenvolver cultivares que associem boa produtividade de grãos com alta capacidade de expansão e qualidade em seu produto final.

No Brasil, o melhoramento do milho pipoca obteve maior impulso a partir de 1990, embora o primeiro programa de milho híbrido tenha se iniciado em 1932, no Instituto Agrônomo de Campinas, sendo que em 1941 foi lançada a primeira variedade nacional, que se originou de ciclos de seleção massal na população-base *South American Mushroom* (SAM), derivada de *South American*, proveniente dos Estados Unidos da América (E.U.A). Mas, apenas em 1988 é que ocorreu o lançamento do primeiro híbrido simples, designado IAC 112 (Sawazaki, 2001). Porém, existem relatos de que o primeiro lançamento de híbrido no país ocorreu mesmo em 1941, pela empresa privada Pioneer, que denominou a cultivar de Zélia. Sendo estes, os primeiros lançamentos de cultivares provenientes de programas de melhoramento de milho pipoca no Brasil, servindo de precursores para outros programas (Silva, 2009).

As estratégias mais utilizadas para o melhoramento de milho pipoca são as mesmas empregadas ao milho comum, tendo como base o método de seleção recorrente, quando se visa a produção de variedades e a obtenção de linhagens endogâmicas para a produção de híbridos.

A seleção recorrente é um sistema designado para aumentar gradativamente a frequência de alelos desejáveis para caracteres quantitativos, por meio de repetidos ciclos de seleção, até atingir níveis satisfatórios de alelos favoráveis, sem reduzir a variabilidade genética da população. O ciclo é realizado em quatro etapas básicas: obtenção de progênies, avaliação, seleção e recombinação destas, para se conduzir outro ciclo de seleção. A seleção recorrente é mais utilizada e apropriada para trabalhos com objetivos em longo prazo, onde população obtida pode ser utilizada como cultivar ou mesmo como fonte de linhagens (Borém et al, 2017).

No desenvolvimento de linhagens elites para a produção de híbridos, vários métodos podem ser utilizados. A principal metodologia, conhecida como padrão, é a utilização de sucessivas autofecundações da população ou da variedade original (Miranda filho e Viégas, 1987).

Em cada geração de autofecundação o número de locos em heterozigose é reduzido, em que na primeira geração de autofecundação ( $S_1$ ) é alcançado 50% de homozigose, na segunda ( $S_2$ ) 75%, na terceira ( $S_3$ ) 87,5%, na quarta ( $S_4$ ) 93,75%, e assim sucessivamente até a sétima geração ( $S_7$ ) atingindo 99,22% de locos homozigotos (Nass et al., 2001). Quando se eleva o índice de homozigose, as linhagens tendem a perder o vigor, os alelos de maior herdabilidade são expressos e os caracteres fixados, evidenciando a diferença existente entre os genótipos. Geralmente após a sétima geração ( $S_7$ ) se realiza o cruzamentos entre as linhagens, para restabelecimento da heterose e obtenção dos híbridos de maior vigor.

Entretanto, no melhoramento da cultura do milho, uma outra concepção de híbrido pode ser utilizada, o qual é oriundo de linhagens parcialmente endogâmicas (híbridos intermediários) e os de  $F_2$  (híbridos intervarietais). Os híbridos de linhagens parcialmente endogâmicas são formados por linhagens de  $S_3$  e  $S_4$ , ou seja, oriundas de três ou quatro autofecundações sucessivas de uma população (Amorim e Souza, 2005). Esses híbridos são de menor custo de obtenção, pois necessitam, de menor número de autofecundações sucessivas, chegando mais rápido ao mercado.

No processo de seleção de genitores para cruzamentos, a avaliação da divergência genética é fundamental. Esse enfoque visa selecionar os genótipos mais

promissores, diminuindo os custos e o tempo necessário para a realização de várias combinações híbridas, muitas vezes desnecessárias. Os esforços são concentrados nas combinações entre os genótipos mais divergentes, com o intuito de captar a heterose ou vigor híbrido (Rinaldi et al., 2007).

A exploração da heterose é uma das maiores contribuições práticas da genética ao melhoramento do milho via hibridização. A heterose é definida como o comportamento superior de combinações híbridas em relação a média de seus parentais. Considerando-se que a heterose para um loco com frequência  $p_1$  em uma população e frequência  $p_2$  na população contrastante é dada por  $h = (p_1 - p_2)^2 d$ , em que  $d$  é uma medida de dominância (Miranda filho e Viégas, 1987). Quanto maior a divergência na população em estudo, maiores as chances de diferenças entre as frequências gênicas em locos contidos com dominância, e conseqüentemente maior a heterose do caráter estudado (Paterniani et al., 2010). A identificação de linhagens divergentes é um fator essencial para o sucesso de um programa de melhoramento de milho pipoca que visa à produção de híbridos superiores.

### **3.3 Descritores morfológicos e agronômicos do milho pipoca**

Em programas de melhoramento genético a avaliação da população de trabalho é fundamental para a seleção dos genótipos com atributos desejados, como exemplos, resistência a doenças e pragas, alta produtividade e capacidade de expansão, no caso do milho pipoca. Essa avaliação é realizada por meio dos descritores, com o intuito de caracterizar morfológica e agronomicamente a performance dos genótipos, para que estes, possam ser utilizados nos programas.

Na cultura do milho pipoca, deve-se levar em consideração os caracteres de interesse agronômicos, visando a produtividade, bem como os parâmetros de qualidade que estão relacionados diretamente a palatabilidade e aceitação da pipoca pelo consumidor. De forma que o conhecimento desses caracteres e suas associações são essenciais.

Em relação à arquitetura e à fisiologia das plantas, geralmente as cultivares de milho pipoca são mais prolíficas, com menor número de folhas, precocidade na maturação e na secagem dos grãos, apresentam colmo mais fino, além de possuírem um sistema radicular menos desenvolvido, o que torna a planta mais suscetível, também, ao acamamento e à seca (Fantin et al., 1991; Hallauer, 2010). Possivelmente, isso ocorre devido aos caracteres agronômicos do milho pipoca se correlacionem, negativamente,

com os caracteres de qualidade da pipoca (Daros et al., 2004; Rangel et al., 2011). Assim, pode-se pressupor que quando se seleciona genitores para maior capacidade de expansão, os genitores que apresentam tais características possivelmente trará consigo características indesejáveis, tais como: baixo vigor, baixa produtividade e maior suscetibilidade a fatores edafoclimáticos (Silva, 2016).

O volume de pipoca expandida por hectare é um caráter dado pelo efeito multiplicativo entre a produtividade média de grãos e capacidade de expansão. Possibilita assim, a seleção de genótipos com boa produtividade e capacidade de expansão, minimizando a problemática da correlação negativa existentes entre os principais caracteres desejados para o melhoramento da cultura (Amaral Júnior et al., 2016).

Em relação aos grãos do milho pipoca, quando comparados ao milho comum, são menores (0,5 a 1,0 cm) e mais duros, podendo ser, redondo, chato, ou pontiagudo e apresentam maior proporção de endosperma duro em relação ao endosperma macio, características essenciais da pipoca (Ziegler e Ashman, 1994). Quanto à coloração, podem ser rosa, vermelho, roxo, preto, azul entre outras, sendo mais comuns as cores branca e amarela (Zinsly e Machado, 1987).

O grão do milho pipoca é composto por embrião, aleuroma, endosperma e pericarpo, sendo os dois últimos componentes os mais importantes, pois influenciam diretamente na qualidade do grão e na capacidade deste expandir. Os grãos possuem uma estrutura mais arredondada, com um pericarpo cerca de quatro vezes mais espesso que o milho comum (Sawasaki, 2001). O endosperma do milho pipoca é em maior parte vítreo, com grãos de amido concentrados, circundados por uma matriz proteica dura e elástica, sem espaço entre as estruturas, enquanto que, o endosperma farináceo, constitui uma pequena porção na parte central do grão, onde os grânulos de amidos estão dispersos, formando espaços vazios entre as estruturas, que são ocupados pela água (Paraginski, 2014; Paes, 2006). Quando aquecido em temperatura em torno de 180°C, o grãos tem a capacidade de expandir, o óleo e a umidade exercem pressão sobre o pericarpo, até que se rompa, expondo o endosperma que é gelatinizado, dando o expecto da pipoca (Zinsly e Machado, 1987).

A capacidade de expansão (CE) é calculada pela relação entre o volume de pipoca obtido e o volume de grãos utilizados (Ruffato et al., 2000). A expansão é a principal característica considerada na avaliação da qualidade do milho pipoca, tendo

relação direta com a maciez da pipoca e com o número de piruás (Ruffato et al., 2000; Seifert et al., 2006).

Alguns fatores podem afetar a capacidade de expansão, sendo eles, umidade dos grãos, ideal entre 13 a 14%; danos no pericarpo e endosperma, ocasionados por danos mecânicos durante a colheita e armazenamento; grãos imaturos e ataque de fungos nos grãos (Ziegler e Ashman, 1994).

O tipo de material desejado pela indústria empacotadora apresenta as características de grãos tipo pérola, de tamanho médio a grande, de cor mais alaranjada brilhante, com pouca ponta e embrião pequeno. Para embalagens de saquinho, requer um mínimo de capacidade de expansão variando de 30 a 37 mL/g, dependendo da empresa; para microondas, exige-se um mínimo de 40 mL/g (Sawazaki, 2010).

O milho traz em sua composição vitaminas do complexo B, vitamina A, proteínas, gorduras, carboidratos, cálcio, ferro, fósforo e amido, além de ser rico em fibras. Os grãos de milho pipoca apresentam em média 10,4 % de proteína, 3,9% de gorduras e 56,5% de carboidratos e 22,9% de fibra, com valor energético de 302,9 kcal/100gramas (CIB, 2017; Germani et al., 1997).

Diante disso, a avaliação de caracteres referentes a qualidade é essencial para o sucesso de seleção, bem como a estimativa conjunta destes com os de produtividade. Alguns dos principais descritores morfoagronômicos, avaliados na cultura do milho pipoca, estão apresentados a seguir.

A determinação do número de dias para o florescimento feminino e masculino é importante para a definição do ciclo de maturação, classificado em plantas de ciclos curtos e tardios. A definição do ciclo é interessante para recomendação de cultivares conforme a região de cultivo. Em áreas de plantio no período de segunda safra em localidades que costumemente sofrem com o frio e a geada, é conveniente a indicação de cultivares com ciclos mais curtos, para que as plantas possam ter um escape do período crítico, na fase de enchimento dos grãos, como o caso da região Sul de Mato Grosso do Sul.

A altura da planta, da inserção da primeira espiga e o diâmetro de colmo são caracteres associados a arquitetura da planta. A mensuração de altura de planta e altura de espiga é utilizada para a determinação do porte da planta, plantas muito altas e com espigas posicionadas muito próximas a folha bandeira, tendem a apresentar maior acamamento e quebramento, devido ao peso e a fatores abióticos, como, vento e chuva. Plantas com elevado porte também apresentam dificuldades para a colheita mecanizada.

Porém, plantas muito baixas tendem a apresentar uma eficiência fotossintética menor, devido a menor interceptação da radiação fotossinteticamente ativa (Marchão et al., 2005).

O diâmetro do colmo, pode servir como indicativo de sustentação, pois colmos mais espessos tendem a apresentar maior resistência a quebras, principalmente referentes ao peso devido ao porte das plantas, bem como condições ambientais. Colmos mais espessos, também tendem a apresentar maior eficácia na distribuição de fotoassimilados.

A determinação do número de plantas quebradas e acamadas em relação ao estande final, avalia o número de plantas que sofreram danos em sua estrutura de colmo, mensurando a suscetibilidade do genótipo em relação aos fatores ambientais como pluviosidade, ventos, etc. Plantas que sofrem acamamento, podem se recuperar e completar o ciclo, porém plantas quebradas não se recuperam, inviabilizando sua colheita, o que pode inferir sobre a produtividade.

A prolificidade avalia a quantidade de espigas que as plantas podem produzir, servindo de indicativo da quantidade média de espigas que um genótipo vai apresentar, característica que pode inferir sobre a produtividade dos grãos. Plantas de milho pipoca tendem a ter uma prolificidade maior ao ser comparada com o milho comum. Para o melhoramento plantas prolíficas são especialmente interessantes, quando se trabalha com técnicas, como a obtenção de progênies de irmãos completos em que é necessário a polinização de ao menos duas espigas por planta, ou mesmo quando se realiza a autofecundação, caso ocorra a perda da primeira “boneca”, pode-se utilizar as demais para o trabalho.

O comprimento, diâmetro da espiga e o número de fileiras de grãos, são caracteres que estão associados a produtividade. Espigas com maior comprimento e número de fileiras, tendem a apresentar relativamente maior quantidades de grãos. Espigas mais espessas, tendem a maior resistência ao quebramento do sabugo nos processamentos pós colheitas.

A produtividade ou rendimento, se refere à quantidade de grãos que foram produzidos por unidade de área (plantas por área). Sendo um caráter de grande importância, principalmente para o produtor. A mensuração da massa de grãos, inferem sobre o peso real, em que sementes mais pesadas tendem a apresentar maior concentração de amido e proporcionar uma maior produtividade em kg, caracterizando ganho na produtividade e lucros.

Diante os atributos desejados, a seleção dos genótipos deve ser baseada em uma análise conjunta considerando os caracteres de interesse da cultura, pois quase sempre, estes estão associados, de maneira positiva ou negativa. Os métodos preditivos de análises multivariadas mostram-se de grande importância para a análise dos caracteres, pois esta, tem como objetivo avaliar as variáveis aleatórias, relacionadas entre si, de forma conjunta, onde cada uma possui o mesmo grau de importância, fornecendo coeficientes de distância genética entre os genótipos, proporcionando grande contribuição ao melhoramento genético (Cruz et al., 2014).

### **3.4 Análise de divergência genética**

A escolha de genitores para o desenvolvimento de populações, capazes de propiciar o ganhos de seleção desejados, ou ainda para direcionar cruzamentos dialélicos constitui uma das etapas mais importantes de um programa de melhoramento, em que o melhorista irá necessitar dispor do conhecimento da diversidade genética da população de trabalho, para facilitar a identificação dos genótipos promissores para a definição da estratégia e continuação do programa.

Existem duas maneiras de se inferir sobre a divergência genética. Quantitativamente, por meio das análises dialélicas, em que é necessário a realização dos cruzamentos entre os genitores para a posterior avaliação dos híbridos resultantes. Embora, amplamente utilizada, tem como desvantagem o grande trabalho envolvido na obtenção e sobretudo, nas avaliações, demandando mais tempo. A divergência genética preditiva é fundamentada em estatística multivariada baseada em diferenças morfológicas, fisiológicas ou moleculares, quantificadas em medidas de dissimilaridade. A inferência preditiva leva em consideração a magnitude das variâncias e covariâncias residuais existentes entre os caracteres mensurados, fornecendo coeficientes de distância genética entre os genótipos, úteis para a predição preliminar de cruzamentos que otimizem a heterose (Rao, 1952, Cruz et al., 2014).

Neste conceito, parte-se da hipótese de que quanto maior a distância genética entre dois genótipos, maiores são as chances de combinações mais promissoras, tornando-se interessante para diminuir o tempo e os cruzamentos desnecessários (Hallauer et al., 2010). Esses estudos auxiliam a identificação de possíveis duplicatas, e fornecem parâmetros para escolha de genitores, que ao serem cruzados, possibilitem maior efeito heterótico na progênie (ou híbrido), isto é, aumentam as chances de

obtenção de genótipos superiores para as próximas gerações. Tais estimativas são de grande utilidade nos programas de melhoramento (Cruz et al., 2014).

Na seleção de genitores para a obtenção de novos híbridos, é recomendado que sejam observados sempre um conjunto de caracteres de interesse, ao invés de caracteres individuais. Os genitores também devem ser complementares para o maior número possível de características de importância agrônômica (Simon, 2012). Assim, o estudo de caracteres morfoagronômicos das plantas cultivadas, por meio de técnicas de análise multivariadas, torna-se fundamental para se conhecer a divergência genética do conjunto de germoplasma disponível para fins de utilização (Gonçalves et al., 2014).

A determinação da divergência genética, com o uso da análise multivariada, permite que diversos caracteres avaliados possam ser utilizados simultaneamente, possibilitando unificar as informações, desta forma cada genótipo será representado por um único valor referente às suas características analisadas, tornando-se bastante vantajosa, principalmente se o objetivo é o melhoramento de caracteres quantitativos, como exemplo, a produção do milho pipoca (Missio et al., 2007; Moura et al., 1999).

Por meio da análise de divergência, espera-se encontrar alto grau de similaridade dentre os genótipos contidos em um mesmo grupo, e alta dissimilaridade entre os grupos. Dessa forma, deve-se evitar cruzamentos dentro de um mesmo grupo (Cruz et al., 2011).

Diversos autores recomendam a utilização de genótipos parentais com a maior divergência possível para maximizar a heterose manifestada nos híbridos, aumentar a probabilidade de ocorrência de segregantes superiores em gerações avançadas e ampliar a base genética (Borém et al., 2017; Cruz et al., 2011; Cruz et al., 2014; Mingoti, 2007; Mohammadi e Prasana, 2003). Portanto, essas informações devem ser utilizadas para recomendação de cruzamentos e combinação das informações contidas na unidade experimental.

A utilização de técnicas multivariadas para estimar a divergência genética, tem se tornado comum entre os melhoristas de plantas, em diferentes espécies, como é o caso do melhoramento do sorgo sacarino (Rocha et al., 2014), pimenta (Pessoa et al., 2015); cana-de-açúcar (Dutra Filho et al., 2011; Souza, J. et al., 2013); soja (Santos et al., 2011; Lopes et al., 2014), milho (Rigon et al., 2015; Silva et al., 2016), milho pipoca (Arnhold et al., 2010); e do feijão (Correa e Gonçalves 2012; Sofi et al., 2014).

Diferentes metodologias de análise multivariada têm sido utilizadas para se estimar a divergência genética, destacando-se os métodos de agrupamentos (Cruz et al., 2014).

### **3.5 Medidas de dissimilaridade e métodos para a estimação da divergência genética**

Os estudos de diversidade genética podem ser baseados em diversos métodos multivariados como, por exemplo, por variáveis canônicas, métodos de agrupamentos, além dos métodos por inteligência computacional. As análises de agrupamento têm o objetivo de reunir as observações em vários grupos, apresentando homogeneidade dentro do grupo e heterogeneidade entre os grupos formados. As metodologias baseadas nos modelos de agrupamentos dependem, fundamentalmente, de medidas de dissimilaridade estimadas previamente, dentre elas, destacam-se a distância euclidiana e a generalizada de Mahalanobis, bem como por algoritmos como o de Gower, dentre outras (Cruz et al., 2011).

As medidas de dissimilaridade exercem grande importância em estudos de diversidade, pois visam a seleção de genitores com potencial para técnicas de hibridização. Genitores com menor distância genética ou geneticamente próximos tendem a não apresentar heterose, que é o fenômeno em que a geração  $F_1$  do cruzamento entre dois genitores apresenta desempenho superior a média dos mesmos. Desse modo, as medidas de dissimilaridade realizam a assimilação de genitores que possam apresentar características genéticas que proporcionam, na  $F_1$ , maior heterose por meio dos genes complementares, e que sejam repassados em posteriores gerações (Cruz et al., 2011; Vittorazzi, 2016).

A técnica de Gower permite a análise simultânea de dados quantitativos e qualitativos. Este método permite que valores da matriz de distância fiquem compreendidos entre 0 e 1, sendo necessário a padronização das variáveis quantitativas e qualitativas (Santos, 2010).

A distância Euclidiana é uma distância geométrica (física), ao qual quanto maior o valor observado, mais dissimilares serão os indivíduos de estudo. É recomendável a padronização ou normalização das variáveis quando os dados não estão no mesmo padrão de medidas (Statística, 2017).

A distância de Mahalanobis ( $D^2$ ) é baseada nas correlações, entre os conjuntos de dados dos caracteres analisados por meio da matriz de variâncias e covariâncias

residuais, no qual quanto maior o valor observado, mais similares serão os indivíduos de estudos. Sua escala é invariante, ou seja, não depende da escala de medida. Preferencialmente é empregada em experimentos com repetições (Cruz et al., 2014). Segundo Cruz et al. (2014), a técnica de distância de Mahalanobis permite quantificar a importância relativa de caracteres para a diversidade genética, por meio da avaliação da contribuição destes para os valores de  $D^2$ .

Essas medidas são amplamente utilizadas em análise de agrupamento e outras técnicas de classificação. Dentre as técnicas mais utilizadas, estão as variáveis canônicas e os métodos de agrupamento, destacando os métodos de otimização e os hierárquicos (Cruz et al., 2014).

### **3.5.1 Técnicas de agrupamento**

A análise de agrupamentos (ou análise de cluster), tem por finalidade reunir os indivíduos em grupos, de forma que exista a máxima homogeneidade dentro do grupo e a máxima heterogeneidade entre os grupos. Ao qual permite uma avaliação da similaridade (ou dissimilaridade) entre elementos, levando em consideração várias características concomitantemente (Sartorio, 2008). As técnicas de agrupamentos podem ser divididas em métodos otimização e hierárquico, destacando-se o método de otimização de Tocher e o hierárquico de ligação média não ponderada entre grupos (UPGMA), que são amplamente utilizados (Cruz et al., 2014).

No procedimento de otimização, ocorre à partição do conjunto de indivíduos em subgrupos não vazios e mutuamente exclusivos, permitindo a realocação dos objetos, por meio da maximização ou minimização de alguma medida pre estabelecida, ou seja, objetiva-se alcançar uma partição que otimize alguma medida pré-definida. Citam-se, os métodos de Tocher e Tocher modificado que correspondem aos mais utilizados na análise de divergência genética entre acessos (Cruz et al., 2014).

O método de Tocher requer a obtenção da matriz de dissimilaridade, sobre a qual é identificado o par de indivíduos mais similares. Estes formarão o grupo inicial e a partir daí é avaliada a possibilidade de inclusão de novos indivíduos, adotando-se que a distância média intragrupo deve ser menor que a distância média intergrupo. Sendo que, a entrada de um indivíduo em um grupo sempre aumenta o valor médio da distância dentro do grupo (Rao, 1952; Cruz et al., 2014). Em contrapartida, no Tocher modificado o processo de agrupamento é sequencial e não simultâneo, não existindo influência dos indivíduos já agrupados (Vasconcelos et al., 2007).

Nos métodos hierárquicos, os genótipos são agrupados por um processo que se repete em vários níveis até o estabelecimento de um dendrograma de alto conteúdo informativo. Dentre os métodos que vêm sendo comumente utilizados pode-se destacar a ligação média não ponderada entre grupos (UPGMA) (Cruz et al., 2014).

O UPGMA é um método de agrupamento não-ponderado, que utiliza a média das distâncias entre todos os pares de genótipos para formação de cada grupo, evitando caracterizar a dissimilaridade por valores extremos (máximo e mínimo). O algoritmo UPGMA constrói uma árvore com raiz (dendrograma), gerados por coeficientes, onde concede peso igual a cada indivíduo do grupo, calculando a similaridade média de um indivíduo que pretende ser juntado ao grupo já existente. A distribuição dos indivíduos no dendrograma não segue um critério de formação de grupos mutuamente exclusivos, já que a principal característica desse método é a formação de ramificações ou dendrograma (Legendre e Legendre, 1998; Cruz et al. 2014; Oliveira, 2015).

Outro método de agrupamento hierárquico utilizado é o do vizinho mais distante, onde a distância entre grupos será definida como a distância entre os pares de acessos mais distantes, ou seja, formam-se todos os pares com um membro de cada grupo. A distância entre os grupos é definida pelo par que possui maior distância (Bussab et al, 1990).

### **3.5.1.1 Variáveis canônicas**

A técnica multivariada de variáveis canônicas é um processo alternativo para a avaliação do grau de similaridade genética entre genitores, que leva em consideração tanto a matriz de covariância residual quanto a de covariância fenotípica entre os caracteres avaliados. Esta técnica apresenta o princípio do processo de agrupamento com base na distância de Mahalanobis ( $D^2$ ) (Cruz et al., 2014).

As variáveis canônicas, quando utilizada em estudos de divergência genética, tem como propósito possibilitar a identificação de indivíduos com maior semelhança, em gráficos. A genética é avaliada por meio da dispersão gráfica dos escores relativos aos genótipos estudados em eixos cartesianos, que geram os vetores canônicos, contendo a concentração da variabilidade disponível (Cruz et al., 2011).

A análise baseada em variáveis canônicas foi relatada primeiramente por Rao (1952). Sua utilização permite avaliar a importância relativa dos caracteres e promover o descarte daqueles que tem pouca contribuição na discriminação do material avaliado, reduzindo, assim, mão-de-obra, tempo e custo despendido na análise e interpretação dos

dados experimentais; bem como, o propósito de agrupar indivíduos que possuam maior grau de similaridade, por meio de dispersões gráficas bi ou tridimensionais, ou seja, com a representação de dois ou três vetores canônicos. Em geral, para a interpretação satisfatória da variabilidade manifestada entre genótipos, os pesquisadores têm optado pela representação gráfica, quando os dois primeiros vetores canônicos estiverem próximo ou acima de 80% da variabilidade total disponível. Esta técnica baseia-se em informações entre e dentro de genótipos, havendo a necessidade de experimentos, com repetições (Cruz et al., 2014).

Camussi et al. (1985) relataram que as transformações por variáveis canônicas, permitem a otimização da visualização das diferenças existentes entre as populações, reduzindo as dimensões, preservando as informações que caracterizam a maior contribuição para a divergência.

A técnica de variáveis canônicas vem sendo empregada em diversos trabalhos, com milho (Colombo et al., 2014; Prazeres e Coelho, 2016); girassol (Vogt et al., 2012); feijão (Cabral et al., 2011; Correa e Gonçalves, 2012); soja (Santos et al., 2011); cana-de-açúcar (Souza et al., 2011); milho pipoca (Arnhold et al., 2010; Miranda et al., 2003), dentre outros.

### **3.5.2 Redes neurais artificiais**

Redes neurais artificiais (RNA's) são técnicas computacionais que apresentam um modelo inspirado na estrutura neural do cérebro humano e que adquirem conhecimento por meio de experiência. As redes neurais artificiais possuem como características a tolerância a falhas, adaptabilidade a novas condições, resolução de problemas com base no conhecimento passado e reconhecimento de padrões (Haykin, 2008).

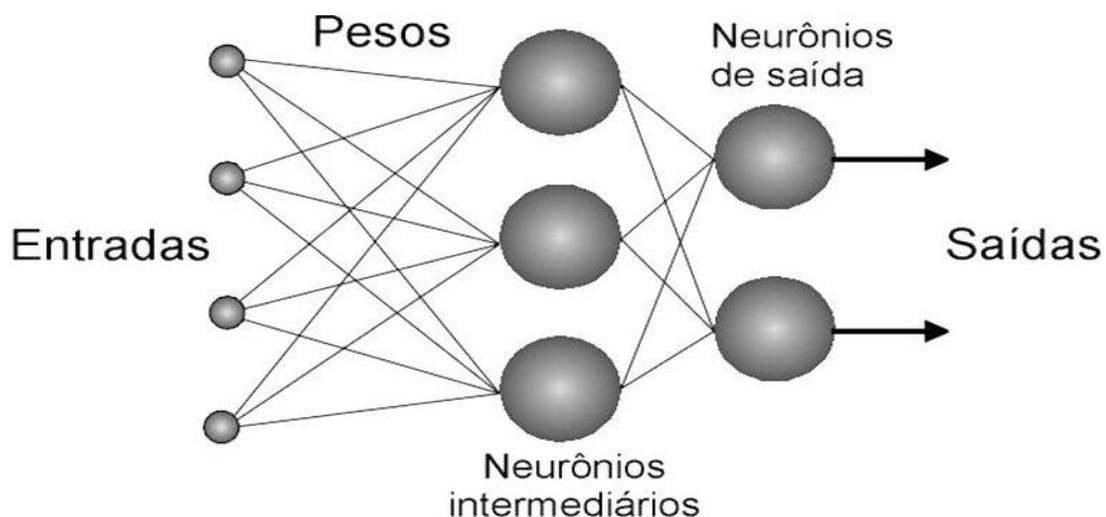
A solução de problemas por meio das redes neurais artificiais é bastante atrativa, já que a forma como estes são apresentados internamente pela rede cria a possibilidade de um desempenho superior ao dos modelos convencionais (Braga et al., 2011).

Uma das vantagens desta técnica está em sua estrutura não linear, que permite captar propriedades mais complexas dos dados (Galvão et al., 1999). Também destaca a sua capacidade de não requerer informação detalhadas sobre os processos físicos do sistema a ser modelado (Sudheer et al., 2003). Como método de classificação, as RNA's apresentam certas vantagens, como a de ser não paramétricas (Kavzoglu e Mather, 2003) e serem tolerantes a perda de dados (Bishop, 1995). Além disso, as RNA's se

mostraram eficazes quando se deseja prever algo, reconhecer padrões ou estabelecer agrupamentos (Haykin, 2008).

As redes neurais são compostas por unidades de processamentos simples que calculam determinadas funções matemáticas. Na programação matemática ou otimização, a modelagem é feita pelo uso de relações lineares e/ou não-lineares. A modelagem em uma rede neural artificial é baseada em dados. Não existe a criação de restrições (Benite, 2003).

De forma geral, a estrutura de uma rede neural artificial (RNA) pode ser dividida em três camadas (Figura 1). A primeira delas é a camada de entrada, na qual os valores das variáveis das amostras são informados à RNA. A segunda são as camadas intermediárias, que possuem neurônios capazes de extrair as características associadas ao sistema a ser inferido. A terceira camada também constituída de neurônios é responsável pela obtenção dos resultados, por meio do processamento das amostras realizado pelos neurônios das camadas anteriores (Silva et al., 2010). Assim, baseado nessa divisão, as arquiteturas de uma RNA apresentam uma estrutura base, porém apresentam particularidades próprias.



**Figura 1.** Modelo da estrutura de camadas de uma rede neural artificial. Fonte: ICMC-USP, 2009.

Existem diversos tipos de redes neurais e diferentes maneiras de classificá-las. Uma das formas mais importantes de classificação se dá pela forma de aprendizagem dos parâmetros, que pode ser supervisionada ou não supervisionada (Jang et al., 2012).

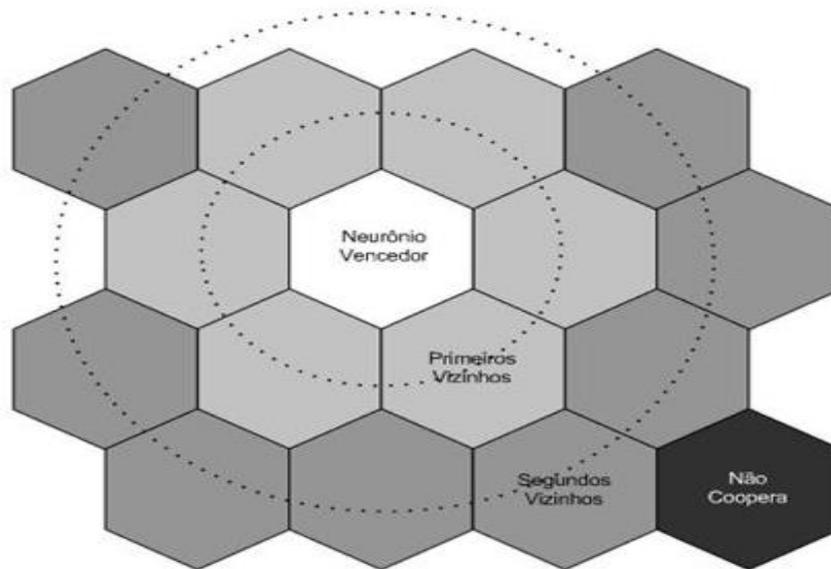
No aprendizado supervisionado a RNA utiliza um agente externo que assinala acertos e erros de acordo com o padrão de entrada e a resposta desejada (Braga et al., 2011). Desta forma, há então a necessidade de se disponibilizar amostras, com entradas

e saídas desejadas, representativas do processo a ser mapeado (Quintão, 2015). Durante este processo, a rede realiza um ajustamento dos pesos sinápticos e os limiares mediante as aplicações de ações comparativas, entre os elementos de processamento, executadas pelo próprio algoritmo de aprendizagem, que realiza o supervisionamento, até que o erro entre os padrões de saída gerados pela rede alcance um valor mínimo aceitável (Braga et al., 2011). Dentre algumas redes com métodos de aprendizado supervisionado, podemos citar, Perceptrons, Adaline e Madaline e a Redes de Função de Base Radial.

No aprendizado não supervisionado não existe um agente externo indicando a resposta desejada para os padrões de entrada. Utiliza-se, portanto, exemplos semelhantes para que a rede calcule as correlações entre eles e responda corretamente (Braga et al., 2011). Desta forma, a própria RNA deve se auto organizar em relação às particularidades existentes entre os elementos do conjunto total de amostras (Quintão, 2015). A rede irá analisar os conjuntos de dados apresentados, determinar algumas propriedades dos conjuntos de dados e aprender a refletir estas propriedades na sua saída. A rede utiliza padrões, regularidades e correlações para agrupar os conjuntos de dados em classes. As propriedades que a rede vai aprender sobre os dados podem variar em função do tipo de arquitetura utilizada e da lei de aprendizagem (Braga et al., 2011). Algumas das redes com métodos de aprendizado não supervisionado, são os Mapas auto-organizável de Kohonen, Redes de Hopfield e Memória Associativa Bidirecional.

Podemos destacar para uso no melhoramento genético de plantas o modelo introduzido por Kohonen (1982), conhecido como rede de Kohonen, ou Mapas auto-organizáveis de Kohonen (SOM), que é indicado para agrupamento de acessos e reconhecimento de padrões. Possui a vantagem de não exigir do pesquisador a aplicação de um conhecimento prévio (origem e organização) dos materiais em estudo, e pode fornecer subsídios para a atualização ou mesmo formulação de novas teorias acerca do problema (Barbosa, 2011).

O modelo de Kohonen é uma rede de duas camadas que pode organizar os dados de entrada em um mapa topológico a partir de um início aleatório (Figura 2). A rede combina uma camada de entrada com uma camada competitiva de unidades processadoras e é treinada pelo algoritmo de aprendizagem não supervisionado (Haykin, 2008).



**Figura 2.** Modelo da estrutura de camadas de um mapa auto-organizável de Kohonen. Fonte: ICMC – USP, 2009.

Os padrões são classificados pelas unidades que eles ativam na camada competitiva, as similaridades entre os padrões são mapeadas em relacionamento de proximidades sobre a grade da camada competitiva. Depois que o treino está completo, os relacionamentos entre os padrões são observados na camada competitiva (Kovacs, 1997).

Com a inicialização do processamento dos dados, tem-se dois procedimentos essenciais envolvidos na implementação do modelo de Kohonen (Haykin, 2008):

a) **Competição:** para cada padrão de entrada, os neurônios da grade calculam seus respectivos valores de uma função discriminante. Essa função discriminante fornece a base para a competição entre os neurônios. O neurônio particular com o maior valor da função discriminante é declarado vencedor da competição.

b) **Adaptação Sináptica:** este mecanismo permite que os neurônios excitados aumentem seus valores individuais da função discriminante, em relação ao padrão de entrada, por meio de ajuste dos seus pesos sinápticos.

O uso das técnicas de redes neurais artificiais na agricultura, tem-se mostrado uma tecnologia promissora, pois possibilitam, por exemplo, a previsão de produção das culturas (Kaul et al., 2005; Ji et al., 2007), comportamento de doenças e pragas (Batchelor et al., 1997), retenção de água no solo (Schaap e Bouten, 1996), dentre outros. No melhoramento genético, tem sido empregadas em estudos de previsão de valor genético (Silva et al., 2014), análises de adaptabilidade e estabilidade (Barroso et al., 2013) e diversidade genética, destacando-se o modelo de mapas auto-organizáveis

de Kohonen, visando agrupamentos e identificações de genótipos (Barbosa et al., 2011).

Dentre os trabalhos realizados no melhoramento vegetal, podemos citar a classificação e diferenciação de bananeira (Oliveira et al., 2013), seleção de genótipos de algodoeiro (Junior, 2017), eficiência da estimação da área foliar de couve (Azevedo et al., 2017), seleção de caracteres complexos em alfafa (Santos, 2017), classificação de tipologias florestais (Gonçalves, et al., 2016), dentre outros.

### **3.6 Importância relativa dos caracteres**

A importância relativa dos caracteres estimada pelo método de Singh (1981) avalia a importância dos caracteres na determinação dos valores da distância entre pares de acessos. Para identificar a variável de menor importância pelo método de Singh ou das variáveis canônicas, dentre aquelas contidas na análise, realiza-se então novas análises, excluindo-se, passo a passo, dentro do conjunto de variáveis, aquelas de menor contribuição relativa sobre a divergência, sendo a de menor importância. Considera-se uma variável de passível de descarte quando a sua exclusão não altera o padrão de agrupamento anteriormente obtido (Cruz et al., 2014).

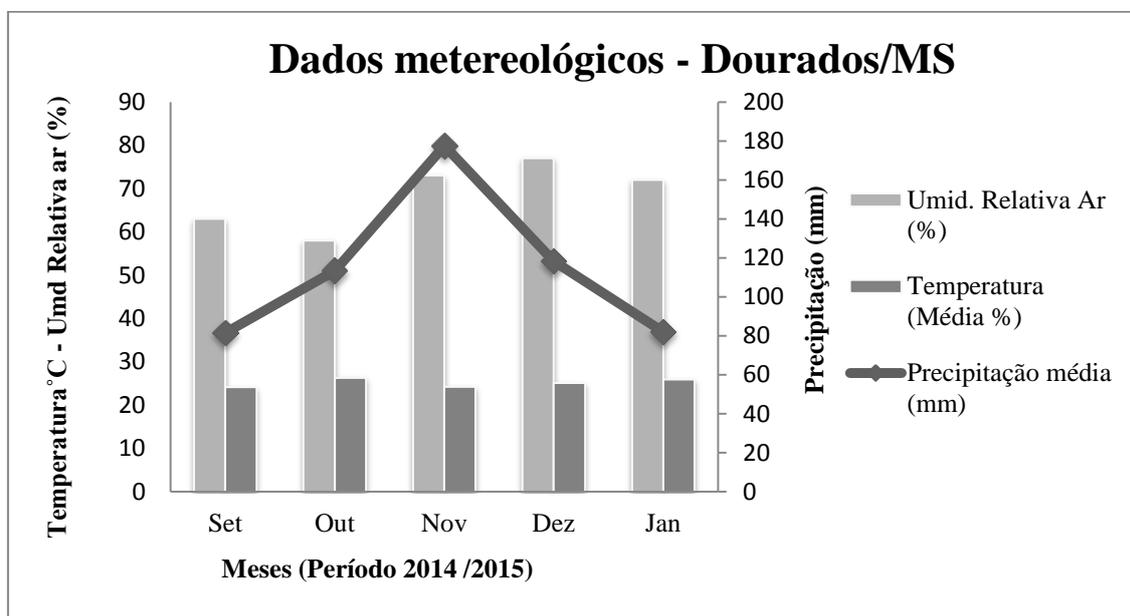
Sendo assim, o estudo da importância relativa de cada caráter, torna-se uma importante contribuição para a seleção, uma vez que indica caracteres que possam ser descartados, diminuindo a quantidade de dados e salientando quais descritores apresentaram maior contribuição dentro do estudo, como exemplo, para a divergência genética.

## 4. MATERIAL E MÉTODOS

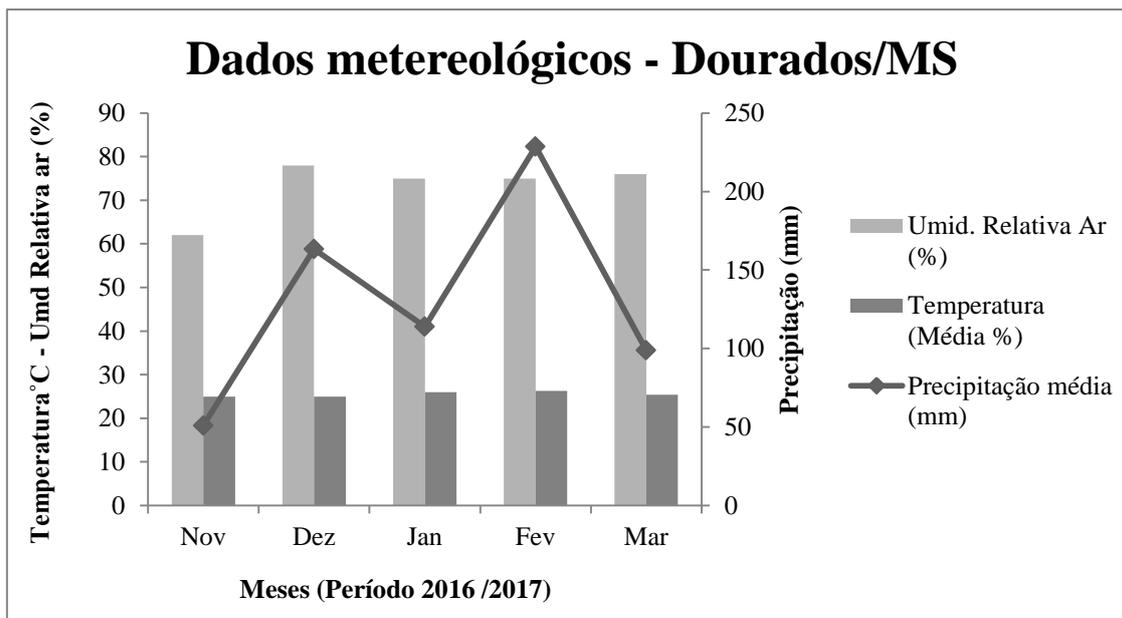
### 4.1 Implantação dos experimentos

Os experimentos foram conduzidos em período de primeira safra, nos anos agrícolas 2014/2015 e 2016/2017, em área da Fazenda Experimental de Ciências Agrárias (FAECA), da Universidade Federal da Grande Dourados (UFGD), localizada no município de Dourados/MS. A FAECA está localizada com as coordenadas de 22°13'15" de latitude Sul, 54°48'21" de longitude Oeste e 430 m de altitude. Sendo o solo do local do tipo latossolo vermelho distroférrico (LVdf). O clima da região de Dourados, segundo a classificação de Köppen é do tipo Cwa classe climática caracterizada como clima temperado úmido, com verões quentes e invernos secos, temperaturas máximas durante os meses de dezembro e janeiro e as temperaturas mínimas entre maio e agosto (Fietz e Fisch, 2008).

Informações meteorológicas das épocas de condução dos experimentos em campo foram obtidas da estação climatológica da Embrapa Agropecuária Oeste (CPAO) - Dourados/MS, no período de setembro à janeiro de 2014/2015 para a safra I (Figura 3) e novembro à março de 2016/2017 para a safra II (Figura 4).



**Figura 3.** Climograma representativo dos dados meteorológicos coletados em campo segundo a estação climatológica da Embrapa Agropecuária Oeste (CPAO) de Dourados/MS, para a safra I nos períodos de setembro à janeiro de 2014/2015.



**Figura 4.** Climograma representativo dos dados meteorológicos coletados em campo segundo a estação climatológica da Embrapa Agropecuária Oeste (CPAO) de Dourados/MS, para a safra II nos períodos de novembro à março de 2016/2017.

Antes da semeadura, as sementes estavam armazenadas no laboratório de Melhoramento Vegetal da Faculdade de Ciências Biológicas e Ambientais da UFGD. O solo foi preparado de forma convencional com uma gradagem. O primeiro experimento foi implantado em 24 de setembro de 2014 (safra 2014/2015), onde foram semeadas 20 linhagens  $S_2$  (2ª geração de autofecundação) de milho pipoca. O segundo experimento foi implantado em 19 de novembro de 2016 (safra 2016/2017) onde foram semeadas as mesmas 20 linhagens, mas em  $S_3$  (3ª geração de autofecundação). As sementes foram provenientes do banco de germoplasma do Grupo de Melhoramento e Biotecnologia Vegetal (GMBV) da UFGD, de origem crioula doadas por pequenos produtores da região da Grande Dourados.

O delineamento experimental adotado para os dois experimentos foi o de blocos casualizados, com três repetições. Cada parcela foi composta por duas linhas de 3,0 metros, com espaçamento de 0,90 metros entre linhas e 0,20 metros entre plantas. A adubação foi efetuada por meio da fórmula 8-20-20 em dose de 300 kg.ha<sup>-1</sup>. O desbaste foi realizado quando as plantas apresentavam de 3 a 4 folhas totalmente emergidas deixando-se uma planta por cova, para a obtenção de 5 plantas por metro linear.

Após os plantios, foi realizado o acompanhamento periódico das linhagens, verificando o desenvolvimento das plantas, proliferação de pragas e de vegetação espontânea, realizando o controle quando necessário por meio de aplicação de controle

químico para as pragas, e capina com enxada entre as linhas para controle da vegetação espontânea.

Nos dois experimentos, cerca de 60 dias após a semeadura, deu-se início aos trabalhos de autofecundação, em que cada planta teve a parte feminina, a espiga, coberta com um saco plástico antes do aparecimento dos estilo-estigmas, a fim de se evitar a polinização por pólen indesejado. No momento em que os pendões apresentavam um terço das anteras abertas e emitindo polén, foram cobertos com saco de papel Kraft. As espigas que não apresentaram o aparecimento de estilo-estigma, foram estimuladas com o auxílio de um canivete com um corte realizado em seu ápice, a fim de estimular a exposição do estilo-estigma para a posterior autopolinização.

No dia seguinte da cobertura de cada pendão realizava-se a autofecundação, onde os papéis Krafts eram retirados do pendão, com cuidado para que não ocorresse perda de pólen presente em seu interior. O saco plástico de proteção da espiga da mesma planta, era levantado e o pólen depositado sobre os estilo-estigmas, que encontravam-se com cerca de 1 centímetro de comprimento. Posteriormente, o papel Kraft utilizado para a polinização era preso ao redor do colmo da planta e fixado com o auxílio de um grampeador, afim de cobrir a espiga e continuar protegendo de pólen exógeno até o momento da colheita. As autofecundações terminaram com aproximadamente 75 dias após a semeadura.

#### **4.2 Avaliações das linhagens**

As características morfoagronômicas avaliadas em campo, foram o florescimento feminino (FF) data em que 50% das plantas da parcela apresentaram emissão de estilo-estigma; florescimento masculino (FM), data em que 50% das plantas da parcela apresentaram um terço de pendão emitindo pólen; altura média de plantas (AP), utilizando para as mensurações régua métrica, posicionada em nível do solo até a inserção da folha-bandeira, em cinco plantas competitivas da parcela, em cm; altura média de inserção da espiga (AE), utilizando para as mensurações régua métrica posicionada em nível do solo até a inserção da espiga superior, nas mesmas plantas em que foi mensurada a AP, em cm e prolificidade (PROLIF), número de espigas por planta de cinco plantas competitivas da parcela. Avaliações de estande final (STAND), plantas acamadas (PA) e quebradas (PQ), foram realizadas no momento da colheita. Colmo quebrado abaixo da inserção da espiga foi contabilizado como planta quebrada (PQ); colmo formando um ângulo maior que 20 graus com a

vertical foi contabilizado como acamado (PA). Assim, a variável AQ consistiu na soma do número de plantas quebradas e acamadas de cada parcela.

Posteriormente, foram realizadas as avaliações das características de pós-colheita, no laboratório de Melhoramento Vegetal da Faculdade de Ciências Biológicas e Ambientais da UFGD, sendo elas: número de espigas autofecundadas por parcela (NEAP), contabilizando o número de espiga autofecundadas que foram colhidas em cada parcela; comprimento de espiga (CPE), medida com régua métrica em centímetros, em cinco espigas aleatórias de cada parcela; diâmetro de espiga (DE), obtido com um paquímetro, onde o mesmo foi posicionado no terço médio de 5 espigas de cada parcela para padronizar a medida da circunferência; número de fileira de grãos médio de 5 cinco espigas aleatórias de cada parcela (NFG); produtividade de grãos, em  $\text{kg ha}^{-1}$  (PROD); massa de 100 grãos (M100), dado pela contagem dos grãos que posteriormente foram pesados em balança semi-analítica; capacidade de expansão (CE), utilizando duas amostras (repetições) de cada parcela de 30 gramas de grãos em forno de micro-ondas, por um tempo médio de 1 minuto e 30 segundos, o volume expandido foi quantificado em uma proveta graduada de 1.000 mL; e volume de pipoca por hectare, em  $\text{m}^3 \text{ ha}^{-1}$  (VP), obtido pelo produto da multiplicação dos valores de produtividade e capacidade de expansão para cada parcela.

Para a análise estatística, a produtividade foi corrigida para umidade uniforme de 13,0%, utilizando-se a metodologia da covariância média proposta por Vencovsky e Barriga (1992) por meio da expressão:

$$\text{PROD} = \frac{P(1-U)}{(1-0,13)}$$

Em que,

PROD = produtividade corrigida para a umidade ideal de 13%;

P = produtividade observada;

U = umidade observada.

Os dados coletados foram compilados e posteriormente submetidos às análises estatísticas.

### **4.3 Análises estatísticas dos dados**

Inicialmente foi realizado o teste da normalidade dos erros por meio do teste de Lilliefors e verificado esses pressupostos. Os dados de cada caráter foi submetido à

análise de variância individual para cada ambiente (ano agrícola), de acordo com o seguinte modelo (Cruz et al., 2014):

$$Y_{ij} = \mu + g_i + b_j + \varepsilon_{ij}$$

Em que:

$Y_{ij}$ : valor observado do i-ésimo genótipo no j-ésimo bloco;

$\mu$ : média geral do ensaio;

$g_i$ : efeito do genótipo i;

$b_j$ : efeito do bloco j;

$\varepsilon_{ij}$ : erro aleatório associado à observação  $Y_{ij}$

Os parâmetros genéticos foram obtidos para cada característica utilizando-se os seguintes estimadores:

Estimativa de variância genotípica ( $\hat{\sigma}_g^2$ ):

$$\hat{\sigma}_g^2 = \frac{QM_g - QM_r}{r}$$

Estimativa da variância fenotípica ( $\hat{\sigma}_f^2$ ):

$$\hat{\sigma}_f^2 = \frac{QM_f}{r}$$

Estimativa da variância ambiental

$$\hat{\sigma}_E^2 = \frac{QM_r}{r}$$

Estimativa do coeficiente de variação genético percentual:

$$CV_g = (\sqrt{\hat{\sigma}_g^2} / \hat{m}) \times 100$$

Estimativa do Índice de variação:

$$I_v = \frac{CV_g}{CV_e}$$

Estimativa da acurácia estatística percentual:

$$\hat{r}_{gg} = (\sqrt{h_x^2}) \times 100$$

Em que  $h_x^2$ : herdabilidade genética no sentido amplo

$$h_x^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_f^2}$$

Em que:

$QM_g$ : quadrado médio das linhagens

$QM_r$ : quadrado médio do resíduo;

$r$ : número de repetições;

$\hat{m}$ : média geral das linhagens.

### 4.3.1 Análises multivariadas

As análises multivariadas foram conduzidas utilizando-se os recursos computacionais do programa Genes (Cruz, 2016).

#### 4.3.1.1 Estimativa da distância generalizada de Mahalanobis

As estimativas da distância generalizada de Mahalanobis  $D^2$  foram obtidas por meio da expressão:

$$D_{ii'}^2 = \delta' \varphi^{-1} \delta$$

Em que:

$D_{ii'}^2$  : é a distância de Mahalanobis entre os genótipos  $i$  e  $i'$ ;

$\varphi$ : matriz da variância e covariâncias residuais;

$\delta$ :  $[d_1, d_2, \dots, d_v]$ , sendo  $d_j = Y_{ij} - Y_i'j$ ;

$Y_{ij}$ : é a média do  $i$ -ésimo genótipo em relação a  $j$ -ésima variável.

#### 5.3.1.2 Método de otimização de Tocher

O método de otimização de Tocher foi realizado a partir da matriz de distâncias de Mahalanobis, sobre a qual é identificado o par de linhagens mais similares. Essas linhagens formarão o grupo inicial (Cruz et al., 2014).

Para estabelecimento dos grupos a distância entre a linhagem  $k$  e o grupo formado pelas linhagens  $ij$  é dada por:

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

A inclusão, ou não, da linhagem  $k$  no grupo é, então, feita considerando:

$$\text{Se } \frac{d_{(grupos)k}}{n} \leq q, \text{ inclui-se a linhagem } k \text{ no grupo;}$$

Se  $\frac{d(\text{grupos})k}{n} > q$ , a linhagem k não é incluído no grupo.

Sendo n o número de linhagens que constitui o grupo original.

Para determinar a distância média intragrupo tem-se:

$$d_{ii1} = \frac{(2 \sum_{j>}^n \sum_{j'}^n d_{jj'})}{n(n-1)}$$

E para determinar a distância média intergrupo temos:

$$d_{ii1} = \frac{(\sum_{j=1}^{n_1} \sum_{j'=1}^{n_2} d_{jj'})}{n_1 n_2}$$

Em que  $n_1$  e  $n_2$  são o número de genótipos dentro dos grupos i e i', respectivamente.

A distância média dentro do grupo é a média das distâncias entre cada par de genitores que o constitui. Pelo critério adotado, esta distância é sempre menor que as distâncias médias intergrupos, que são obtidas pela média das distâncias entre pares de genótipos pertencentes aos diferentes grupos (Cruz et al., 2014).

#### 4.3.1.3 Método hierárquico de ligação média não ponderada entre grupos (UPGMA) e vizinho mais distante

O método de agrupamento hierárquico UPGMA é um método não ponderado, e promove o agrupamento dos genótipos avaliados aos pares por média aritmética, que consiste em atribuir pesos iguais a cada indivíduo do grupo, calculando a similaridade média de um indivíduo que pretende ser alocado ao grupo já existente (Cruz et al., 2011).

O método identifica a distância de um grupo em relação aos demais indivíduos (Cruz et al., 2014), em que, a distância entre uma linhagem k e um grupo formado pelas linhagens i e j é dada por:

$$d_{(ij)k} = \frac{d_{ij} + d_{jk}}{2}$$

Em que,  $d_{(ij)k}$  é dado pela média das distâncias dos pares de linhagens (i e k) e (j e k), gerando uma nova matriz com esses valores reduzindo a dimensionalidade passo a passo. Com a construção do dendrograma, novas distâncias serão estimadas, de forma

que, se uma linhagem k for incorporada a um grupo (ij), a distância deste novo grupo (ij.k) em relação a linhagem (l) ou a outro grupo (lm), é dada por:

$$d_{(ij)k} = \frac{d_{ij} + d_{jl} + d_{jk}}{2}$$

$$d_{(ij)k} = \frac{d_{ij} + d_{im} + d_{jl} + d_{jk}}{4}, \text{ e assim, sucessivamente.}$$

Para o método do vizinho mais distante, a expressão empregada é dada pela distância máxima:

$$d_{(ij)k} = \max (d_{ik}, d_{jk})$$

#### 4.3.1.4 Dispersão gráfica de variáveis canônicas

As variáveis originais (caracteres mensurados) foram transformadas em um novo conjunto, originando as variáveis canônicas, estabelecendo-se as seguintes propriedades (Cruz et al., 2014):

a) Se  $Y_{ij}$  é uma variável canônica, então,  $Y_j$  é uma combinação linear de  $X_j$  variáveis originais:

$$Y_{ij} = a_1 X_{i1} + a_2 X_{i2} + \dots + a_n X_{in}$$

b) Se  $Y_{ij}$  é outra variável canônica, então:

$$Y_{ij'} = b_1 X_{i1} + b_2 X_{i2} + \dots + b_n X_{in}$$

$$\sum_j \sum_{j'} a_j a_{j'} \sigma_{jj'} = \sum_j \sum_{j'} b_j b_{j'} \sigma_{jj'} = 1$$

$$\sum_j \sum_{j'} b_j b_{j'} \sigma_{jj'} = 0$$

Em que:

$\sigma_{jj'}$  é a covariância residual entre os caracteres j e j'.

c) Dentre todas as variáveis canônicas:

$$\text{Variância } Y_{i1} > \text{variância } Y_{i2} > \dots > \text{variância } Y_{in}$$

Segundo Rao (1952), as variâncias de cada variável canônica e os coeficientes de ponderação dos caracteres podem ser estimados pela solução dos sistemas de equações:

$$(T - \lambda_j E) \delta_j = \Phi$$

E a j-ésima variância, pela raiz característica de ordem correspondente, obtida pela solução de:

$$\det [T - \lambda_j E] = 0$$

Em que:

T = matriz de variâncias e covariâncias fenotípicas;

E = matriz de variâncias e covariâncias residuais;

$\delta_j$  = autovetor associado a cada estimativa das raízes características, cujos elementos são os coeficientes de ponderação; e

$\lambda_j$  = raízes características (autovalores) da matriz  $E^{-1}T$ .

#### 4.3.2 Redes neurais artificiais

Foi aplicada a técnica de redes neurais artificiais, por meio da integração entre os programas computacionais Genes (Cruz, 2016) e Matlab<sup>®</sup> (Matlab, 2009a). A metodologia baseada no modelo de Mapas auto-organizáveis de Kohonen (SOM) foi aplicada com a abordagem tradicional não supervisionada, de acordo com as características e necessidades do presente estudo.

Os parâmetros iniciais expressos foram:

Épocas : 2000, equivalente ao total de comparações realizadas;

Raio : 1

Grade - Linhas : 3

Grade - Colunas : 3

Topologia (hextop, gridtop, randtop) : 1

Distância (linkdist, dist, mandist) : 2

Em que:

As grades foram responsáveis pelas determinações da quantidade de classes, como base para serem formados os grupos, onde foram preestabelecidos 3 linhas x 3 colunas, resultando em 9 classes organizacionais.

A topologia representa os espaços topológicos, sendo considerado como uma extensão da geometria, ou seja, a figura gráfica a ser representada. A topologia definida foi hextop (1), para topologia de rede neural SOM com vizinhança hexagonal.

Distâncias utilizadas para configuração das redes neurais artificiais para o programa Matlab<sup>®</sup>:

Linkdist:

$$D_{ij} = 0, \text{ if } i = j$$

- = 1, if  $\sum((P_i - P_j)^2)^{0.5}$  is  $\leq 1$
- = 2, if  $k$  exists,  $D_{ik} = D_{kj} = 1$
- = 3, if  $k_1, k_2$  exist,  $D_{ik_1} = D_{k_1 k_2} = D_{k_2 j} = 1$ .
- =  $N$ , if  $k_1 \dots k_N$  exist,  $D_{ik_1} = D_{k_1 k_2} = \dots = D_{k_N j} = 1$
- =  $S$ , se nenhuma das condições acima se aplicar.

Dist:

Distância Euclidiana  

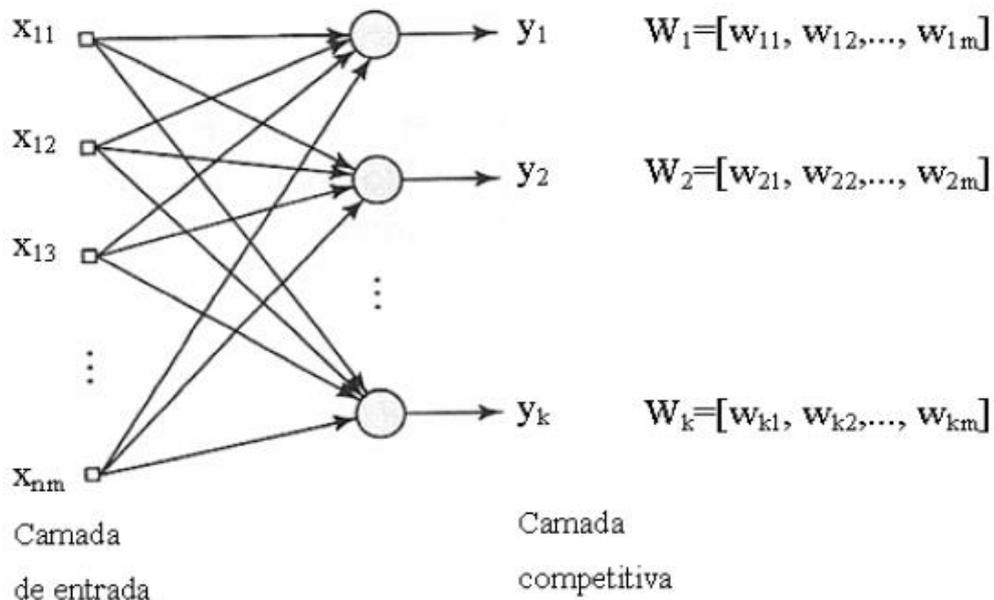
$$D = \sqrt{\sum (X - Y)^2}$$

Mandist:

Distância de Manhattan  

$$D = \sum |X - Y|$$

A arquitetura da rede neural baseada no modelo de Kohonen, de acordo com o contexto deste estudo, pode ser visualizada na Figura 5.



**Figura 5.** Arquitetura da rede neural artificial Kohonen composta por  $n$  acessos e  $m$  caracteres na camada de entrada, e por  $k$  neurônios na camada competitiva. Fonte: ICMC-USP, 2009.

A arquitetura é composta por uma matriz de entrada  $n \times m$ , sendo  $n$  linhas e  $m$  elementos de entrada ou caracteres, que juntos representam o vetor de entrada  $X$ , e por  $k$  neurônios de saída, representando as classes às quais as linhas podem ser agrupados.

A metodologia apresentada é baseada nos conceitos propostos por Haykin (2008). Dada uma entrada, apenas um neurônio de saída deverá ser ativado, registrando assim, a classe à qual a linhagem pertence. As classes devem agrupar linhagens com características semelhantes. Então, é realizada a classificação de vizinhança por relacionamento de similaridade de valores.

### 4.3.3 Importância relativa dos caracteres

A importância relativa dos caracteres foi calculada utilizando-se o método proposto por Singh (1981), que se baseia na partição do total das estimativas das distâncias  $D^2$  considerando todos os possíveis pares de indivíduos, para a contribuição referente a cada caráter.

Considerando-se que:

$$D_{ii'}^2 = \delta' \varphi^{-1} \delta = \sum_{j=i}^n \sum_{j'=i}^n \omega_{jj'} d_j d_{j'}$$

Em que  $\omega_{jj'}$  é o elemento da  $j$ -ésima linha e  $j'$ -ésima coluna da inversa da matriz de variâncias e covariâncias residuais.

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Experimento linhagens S<sub>2</sub>, safra I (2014/2015)

#### 5.1.1 Análises univariadas

Por meio da análise de variância (Tabela 1), para as linhagens S<sub>2</sub> na safra I (2014/2015), foi constatado para a fonte de variação linhagens efeito significativo a 1% de probabilidade pelo teste F, para as características AP, AE, AQ e VP, e significância de 5% de probabilidade para PROLIF, NEAP, CEP, NFG, M100, PROD e CE o que indica a existência de variabilidade entre as linhagens para essas características.

**Tabela 1.** Resumo das análises de variância para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S<sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.

Caráter	Quadrado Médio				
	Blocos (2) <sup>1</sup>	Linhagens (19) <sup>1</sup>	Resíduo (38) <sup>1</sup>	Média	Acurácia ( $\hat{f}_{gg}$ )
FF	1,85	11,23 <sup>ns</sup>	8,53	72,90	49,01
FM	4,01	8,41 <sup>ns</sup>	6,96	71,42	41,54
AP	0,49	316,96 <sup>**</sup>	113,51	160,77	80,11
AE	40,15	241,66 <sup>**</sup>	74,33	85,03	83,21
DC	92,25	8,10 <sup>ns</sup>	4,86	21,82	63,20
PROLIF	0,65	0,25 <sup>*</sup>	0,09	2,17	80,38
AQ	8,75	29,03 <sup>**</sup>	10,01	8,80	80,93
NEAP	28,46	18,29 <sup>*</sup>	8,36	9,28	73,68
CPE	1,33	3,41 <sup>*</sup>	1,42	13,21	76,22
DE	6,44	15,06 <sup>ns</sup>	12,19	30,67	43,81
NFG	0,40	1,92 <sup>*</sup>	0,98	12,44	69,84
M100	5,25	20,42 <sup>*</sup>	3,64	13,28	90,63
PROD	78923,40	121747,59 <sup>*</sup>	35379,28	608,19	84,22
CE	9,21	42,01 <sup>*</sup>	9,84	20,08	87,50
VP	14749305,28	60144649,91 <sup>**</sup>	19394031,61	12090,08	82,40

<sup>1</sup> Grau de liberdade. ns, \*\*, \*, Não significativo, Significativo a 1% e 5% de probabilidade respectivamente, pelo teste F.

Para FF, FM, DC e DE não foi verificado diferença estatística significativa entre as linhagens, indicando ausência de variabilidade entre as linhagens para essas características. Assim, 72,90; 71,42; 21,82 e 30,67 representam respectivamente a média das 20 linhagens.

As acurácias estimadas, em geral, foram boas e de elevada magnitude, todas com valores entre 63,20% á 90,63%, com exceção do FF, FM e DE, apontando, de maneira geral, boa precisão experimental, de acordo com os limites de classes estabelecidos em Cargnelutti Filho e Storck (2009).

A aplicação conveniente da estatística acurácia seletiva (AS) para avaliar a qualidade de um experimento é destacada no trabalho de Resende e Duarte (2007). Segundo Cargnelutti Filho e Storck (2009), a estatística de acurácia seletiva é considerada mais adequada, para avaliação da precisão experimental em ensaios de competição de cultivares de milho, do que o coeficiente de variação e que a diferença mínima significativa pelo teste de Tukey.

Trabalhos como de Pena et al. (2012), avaliando o desempenho agrônômico de linhagens de milho pipoca, encontraram valores de acurácia entre 70 a 90%, resultados que corroboraram com os obtidos no presente estudo.

## **5.1.2. Análises multivariadas**

### **5.1.2.1 Distância generalizada de Mahalanobis**

As medidas de dissimilaridade genética (Tabela 2), estimadas a partir da Distância Generalizada de Mahalanobis ( $D^2$ ) apresentaram magnitude de 9,22 a 107,09, indicando a presença de ampla variabilidade genética para as linhagens avaliadas. A combinação entre as linhagens 8 x 9 (107,09), seguida pela combinação entre 12 x 14 (94,38) apresentaram as estimativas mais elevadas da distância generalizada de Mahalanobis, sendo os pares mais divergentes. A menor divergência foi observada entre os pares 16 x 19 (9,22), e entre as linhagens 20 e 2 (9,75).

Segundo Oliboni et al. (2012) a divergência genética está relacionada com o grau de distância entre as linhagens no conjunto de caracteres genéticos que diferem entre as populações. Em que, na maioria dos casos a distância genética é positivamente correlacionada com a heterose.

Para as combinações entre as linhagens 8 x 14 (93,80), 14 x 15 (80,21), 1 x 9 (75,78), e 11 x 8 (73,62), também foram encontrados valores elevados de

dissimilaridade, indicando maior divergência genética entre essas linhagens. A formação destes grupos é de fundamental importância na escolha dos genitores, pois as novas combinações a serem estabelecidas devem ser baseadas na magnitude de suas dissimilaridades e no potencial *per se* dos genitores (Cruz et al., 2014).

Menores dissimilaridades também foram observadas entre as combinações 6 x 2 (10,06), 3 x 7 (12,23), 4 x 17 (12,48), 10 x 8 (13,10) e 18 x 17 (13, 15), resultando em linhagens geneticamente similares, isto indica, que possíveis cruzamentos entre esses pares, resultarão em progênies sem expressão heterótica.

Silva et al. (2016a), em um estudo de recombinação em progênies de meios-irmãos de milho-verde, observaram amplitudes das estimativas de dissimilaridade de Mahalanobis entre 1,51 a 109,66 na primeira safra sugerindo haver ampla variabilidade genética entre as linhagens avaliadas. A distancia de Mahalanobis é empregada em diversos trabalhos de diversidade, como de girassol (Rigon et. al., 2012) sorgo (Oliveira, 2015), milho pipoca (Melo et al., 2017), dentre outras culturas, sendo essencial para aplicação de metodologias para estimação da divergência genética por agrupamento.

**Tabela 2.** Estimativas da dissimilaridade ( $D^2$ ) para as 20 combinações de linhagens mais similares e mais divergentes, respectivamente, com base na avaliação de 15 caracteres morfoagronômicos quantitativos, em 20 linhagens  $S_2$  de milho pipoca na safra de 2014/2015, em Dourados/MS.

Linhagem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	20,86	18,01	41,01	16,83	25,48	12,41	37,03	75,98	32,66	33,01	55,88	30,48	49,05	55,59	27,47	43,73	39,56	35,61	21,90
2		-	19,88	25,85	17,50	10,06	10,64	39,84	30,78	21,69	23,83	29,29	15,05	35,00	34,82	12,13	13,79	20,70	14,43	9,75
3			-	24,49	16,56	24,15	12,23	50,92	69,12	49,32	41,34	57,15	32,91	55,93	55,24	16,67	27,80	36,55	29,23	25,43
4				-	48,84	22,15	26,46	70,70	40,84	59,26	47,10	28,69	43,21	67,22	63,28	22,50	12,48	14,27	24,32	23,80
5					-	18,38	18,41	59,79	69,48	45,24	31,19	70,67	34,69	27,74	59,93	19,63	45,88	52,29	37,63	24,36
6						-	21,50	67,96	27,41	44,04	26,51	31,03	31,28	31,40	51,00	17,85	22,52	19,35	31,36	11,29
7							-	42,70	50,91	31,49	25,93	44,64	22,57	46,12	50,47	15,41	23,11	31,44	23,53	16,27
8								-	107,09	13,10	73,62	58,89	30,28	93,80	69,52	63,69	53,05	68,86	52,92	40,72
9									-	64,77	39,84	43,08	45,78	53,65	60,21	38,42	23,81	32,46	44,91	35,52
10										-	46,22	46,46	24,66	53,80	60,52	44,89	35,37	49,59	37,42	24,77
11											-	62,42	28,52	34,38	63,31	31,21	36,94	48,59	42,63	36,42
12												-	38,97	94,38	67,31	50,31	24,66	25,77	37,70	30,59
13													-	62,64	27,69	30,18	23,78	43,20	29,70	22,18
14														-	80,21	40,36	54,81	54,71	57,91	42,03
15															-	35,55	43,83	41,39	32,53	55,60
16																-	17,56	24,60	9,22	25,00
17																	-	13,15	15,60	22,24
18																		-	23,36	28,75
19																			-	31,17
20																				-

### 5.1.2.3 Análises de agrupamentos pelos métodos de Tocher, ligação média não ponderada entre grupos UPGMA e vizinho mais distante

Utilizando o Método de otimização de Tocher, baseado na distância de Mahalanobis, foram obtidos oito grupos (Tabela 3). A formação desses grupos contribuirá na escolha de genitores, embasadas na magnitude de suas dissimilaridades. Assim, as linhagens que fazem parte do mesmo grupo apresentam maior similaridade genética entre si.

**Tabela 3.** Agrupamento pelo método de otimização de Tocher, com base na distância generalizada de Mahalanobis de 20 linhagens S<sub>2</sub> de milho pipoca, na safra de 2014/2015, em Dourados/MS.

Grupos	Linhagens										% Linhagens
I	16	19	2	17	7	6	20	3	4	18	50
II	8	10	13								15
III	1	5									10
IV	11										5
V	14										5
VI	15										5
VII	12										5
VIII	9										5

O grupo I foi o mais numeroso, contemplando 50% das linhagens de milho pipoca avaliadas, sendo, as linhagens 16, 19, 2, 17, 7, 6, 20, 3, 4, e 18. O grupo II apresentou 15% das linhagens (8, 10, 13). Assim, os dois primeiros grupos contemplaram 65% das linhagens avaliadas. Melo et al. (2017), em um estudo com milho pipoca avaliando progênies de meios-irmãos, observaram a formação de seis grupos pelo método de Tocher, com a maior concentração de genótipos também nos dois primeiros grupos, com 72,5% das linhagens em estudo.

Na busca de divergência genética entre as populações, uma alta porcentagem de linhagens similares, pode reduzir a possibilidade de obtenção de híbridos interpopulacionais. Neste seguimento, a escolha por linhagens de grupos que

apresentem maiores divergências pode resultar em cruzamentos mais contrastantes, possibilitando uma maior exploração da heterose (Rotili et al., 2012).

O grupo III apresentou 10% das linhagens, sendo elas, a 1 e a 5. Os grupos IV, V, VI, VII e VIII reuniram apenas uma linhagem em cada grupo, sendo elas, linhagens 1, 14, 15, 12 e 9, respectivamente, sugerindo que estas sejam as mais divergentes em relação as demais. A ocorrência de grupos com apenas uma linhagem indica uma ampla divergência genética, de modo que divergem mais em relação aos demais conjuntos formados, o que facilita a projeção dos trabalhos de melhoramento e possibilita a identificação de linhagens distintas para futuros cruzamentos (Rotili et al., 2012, Silva et al., 2015).

Dessa forma, de acordo com o agrupamento de Tocher as linhagens 1, 14, 15, 12 e 9, podem ser utilizadas em cruzamentos com qualquer genótipo dos outros grupos obtendo-se assim, híbridos potencias já que são resultantes de genitores contrastantes.

De Paula (2009), em um trabalho de caracterização e divergência genética de variedades crioulas de milho pipoca, utilizando método de Tocher baseado na distância de Mahalanobis, também observou a formação de grupos contendo apenas um genótipo. Resultado também corroborados por Melo et al. (2017), em seu estudo de estratégias de seleção entre progênes meios-irmãos de milho pipoca.

A partir do método de otimização de Tocher foi estimada a dissimilaridade intra e intergrupos (Tabela 4). A maior distância média contida intragrupo foi observada no grupo II (22,68), enquanto que a menor foi verificada no grupo III (16,83). As distâncias intragrupos foram inferiores a qualquer distância intergrupos, atendendo ao critério estabelecido para o método de otimização de Tocher, de acordo com Cruz et al. (2014).

As maiores distâncias intergrupo encontradas, foram entre os grupos V e VII (94,38), V e VI (80,21), II e VIII (72,55), e III e VIII (72,73), correspondendo aos grupos que possuem as maiores dissimilaridades, entre si, indicando provavelmente as melhores combinações para cruzamentos.

As menores distâncias intergrupo foram obtidas entre os grupos I e III (29,28), III e IV (32,10), IV e V (34,38), I e VII (35,98), I e IV (36,05). De acordo com Cruz et al. (2011) menores distâncias evidenciam que o intercruzamento das linhagens contidas nos respectivos grupos, podem não ser muito indicado para a obtenção de progênes superiores devido a similaridade genética, podendo apresentar baixa heterose, nas gerações segregantes.

**Tabela 4.** Distâncias médias intra e intergrupos estimadas pelo método de otimização de Tocher, com base na dissimilaridade entre 20 linhagens S<sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.

Grupos	I	II	III	IV	V	VI	VII	VIII
I	20,53	41,44	29,28	36,05	48,55	46,37	35,98	39,42
II		22,68	39,98	49,45	70,08	52,58	48,11	<b>72,55</b>
III			16,83	32,10	38,39	57,76	62,78	<b>72,73</b>
IV				-	34,38	63,31	62,42	39,84
V					-	<b>80,21</b>	<b>94,38</b>	53,65
VI						-	67,31	60,21
VII							-	43,08
VIII								-

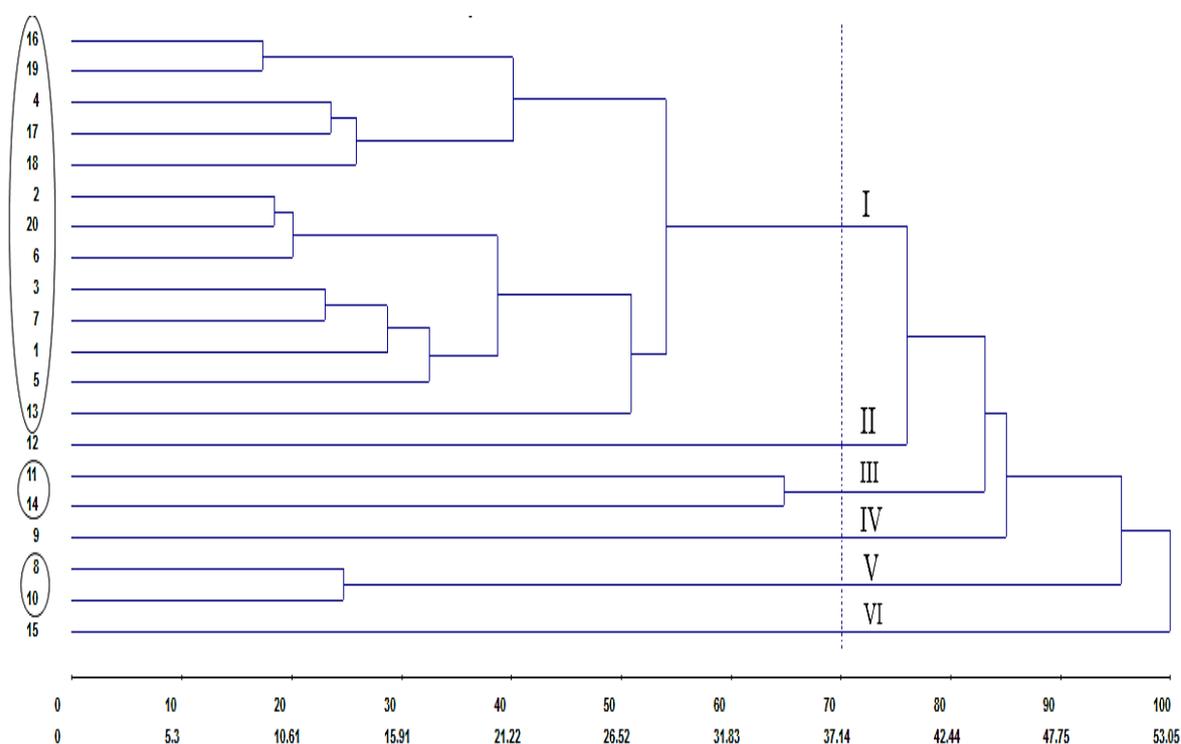
\*Em negrito estão destacadas as maiores distâncias médias intergrupos.

Em programas de melhoramento, recomenda-se a seleção de linhagens com a maior divergência possível, para maximizar a probabilidade de ocorrência de segregantes superiores e ampliar a base genética. Assim, deve-se levar em consideração as distâncias intergrupos, visando o cruzamento de indivíduos contidos entre os grupos mais distantes (Cruz et al., 2011). Outro fator que deve ser levado em consideração para a orientação nos cruzamentos, é o desempenho da linhagem, principalmente, quando são envolvidas as características de maior interesse da cultura, como a produtividade e a capacidade de expansão no milho pipoca.

Para o método de agrupamento UPGMA, o corte no dendograma foi realizado considerando 70% de variabilidade genética, o que possibilitou a divisão das linhagens em seis grupos distintos (Figura 6). Esse critério foi definido considerando a mudança abrupta de nível, conforme sugerido por Barroso e Artes (2003).

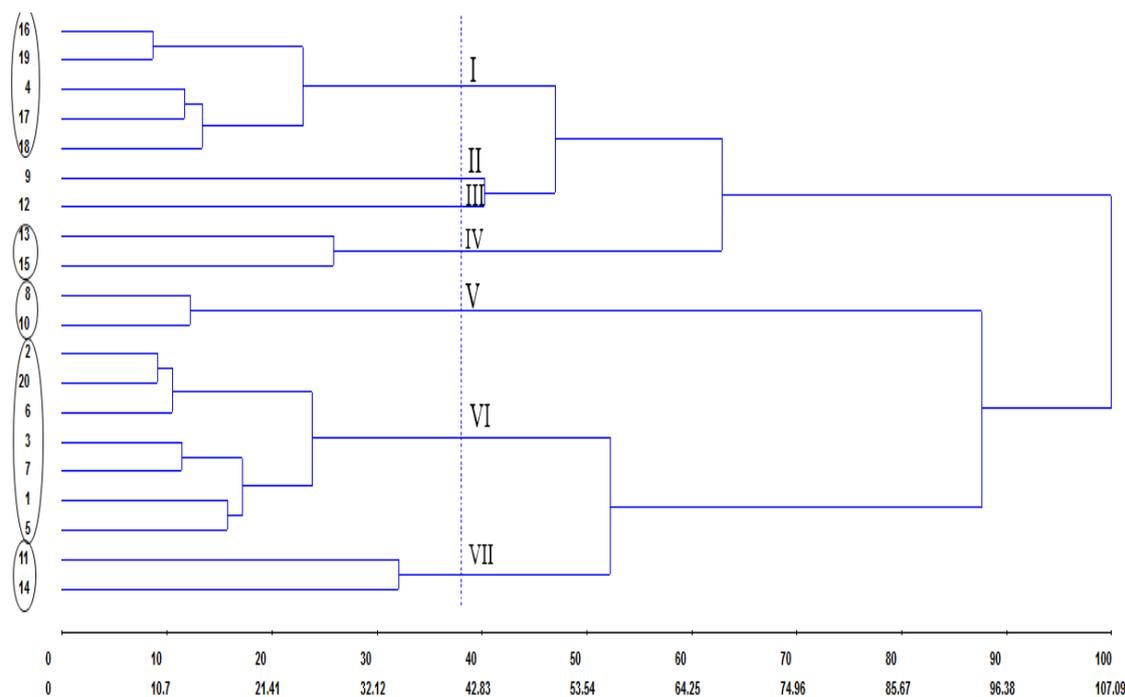
O grupo I foi formado por 13 das 20 linhagens analisadas, reunindo as linhagens 13, 5, 1, 7, 3, 6, 20, 2, 18, 17, 4, 19 e 16. O grupo II pela linhagem 12, que apresentou o maior número de fileira de grãos na espiga (NFG) e número de espigas autofecundadas por parcela (NEAP). O grupo III foi formado pelas linhagens 14 e 11. O grupo IV pela linhagem 9, que apresentou a menor altura de planta média. O grupo V pelas linhagens 8 e 10, que se destacaram por apresentarem a maior massa de cem grãos (M100) e produtividade (PROD). O grupo VI pela linhagem 15, tendo como principais caracteres a maior capacidade de expansão (CE), maior volume de pipoca (VP) e prolificidade

(PROLIF). Os valores médios dos caracteres para cada linhagem dos grupos podem ser observados no anexo 2.



**Figura 6.** Dendrograma representativo da dissimilaridade genética entre 20 linhagens  $S_2$  de milho pipoca, obtidas pela ligação média não ponderada entre grupos (UPGMA), a partir da medida de dissimilaridade generalizada de Mahalanobis.

Em relação ao dendrograma obtido pelo método do vizinho mais distante (Figura 7), com o corte realizado em 40% de variação, foram formados sete grupos, de boa concordância aos realizados pelo método de UPGMA. As linhagens 9 e 12 permaneceram em grupos isolados. Os agrupamentos também foram concordantes para as linhagens 14 e 11, linhagens 8 e 10. No entanto, a linhagem 15 que passou a constituir um grupo com a linhagem 13 (grupo IV vizinho mais distante). Pelo método UPGMA as linhagens 13, 18, 5, 17, 4, 3, 6, 7, 20, 2, 19, 1 e 16 foram reunidas em um grande grupo (grupo I), contudo pelo vizinho mais distante as linhagens 16, 19, 4, 17 e 18 passaram a formar um novo grupo (grupo I). Contudo, de maneira geral, é possível afirmar que os agrupamentos realizados pelo método de UPGMA foram consistentes aos encontrados pelo método do vizinho mais distante, inferindo sobre uma classificação eficiente em relação a dissimilaridade das linhagens.



**Figura 7.** Dendrograma representativo da dissimilaridade genética entre 20 linhagens  $S_2$  de milho pipoca, obtidas pela ligação completa dos vizinhos mais distantes, a partir da medida de dissimilaridade generalizada de Mahalanobis.

Silva et al. (2016) em um trabalho avaliando a dissimilaridade genética em progênies de meios-irmãos em milho, também constatou formação de grupos com indivíduos isolados, pelo método UPGMA. Resultados semelhantes, também foram relatados em trabalhos com soja (Lopes et al., 2014), sorgo sacarino (Oliveira, 2015), e milho pipoca (Faria, 2008; Miranda et al., 2003), demonstrando ser um método interessante para a identificação e seleção de linhagens com maiores divergência para posteriores cruzamentos.

Os métodos de Tocher e UPGMA apresentaram semelhanças nos agrupamentos das linhagens. As linhagens 12, 13 e 15 permaneceram em grupos isolados em ambos os métodos. As diferenças nos agrupamentos, foram em relação as linhagens 14 e 11 que formavam grupos dois grupos, cada um com uma das linhagens pelo método de Tocher, e foram incorporados em um único grupo (III) pelo método de UPGMA. As linhagens 1 e 5 que constituíram o grupo III pelo método de Tocher, já no método UPGMA foram alocadas no grupo I. A linhagem 13 que no método de Tocher foi reunida no grupo II com as linhagens 8 e 10, no método de UPGMA passou a ser agrupada com as demais linhagens do grupo I.

Os dois métodos foram condizentes em relação a divergência genética contida nas linhagens 12, 13 e 15, indicando que podem ser consideradas os genitores mais

promissores para serem utilizados em cruzamentos futuros, contribuindo para a restauração da heterose e obtenção de potenciais híbridos.

#### 5.1.2.4 Dispersão gráfica de variáveis canônicas

As estimativas dos autovalores ( $\lambda_i$ ) correspondentes as variáveis canônicas ( $VC_i$ ), estão apresentadas na Tabela 5. Dentre as 15 variáveis canônicas, observou-se que as três primeiras variáveis explicaram mais de 60% da variação total, sendo para a primeira 24,89%, a segunda 22,08% e para a terceira 13,58%. Para a descrição da divergência é desejável que as duas primeiras variáveis expliquem ao menos 80% da variação total (Cruz et al., 2011). Contudo, casos em que este limite não é atingido, a análise pode ser completada com a dispersão em relação a terceira e quarta variáveis canônicas (Cruz et al., 2014). Neste estudo, esta porcentagem (80%) só foi verificada na quinta variável canônica. Todavia, optou-se por manter as três primeiras variáveis canônicas, representando o valor de 60,55% da variação total, valor que já se pode ser representado em um espaço tridimensional para a interpretação geométrica, uma vez que atinge mais da metade da variação contida.

**Tabela 5.** Estimativas das variâncias (autovalores) e variâncias percentuais acumuladas das variáveis canônicas, visando estimar a dissimilaridade genética entre 20 linhagens  $S_2$  de milho pipoca, na safra de 2014/2015, em Dourados/MS.

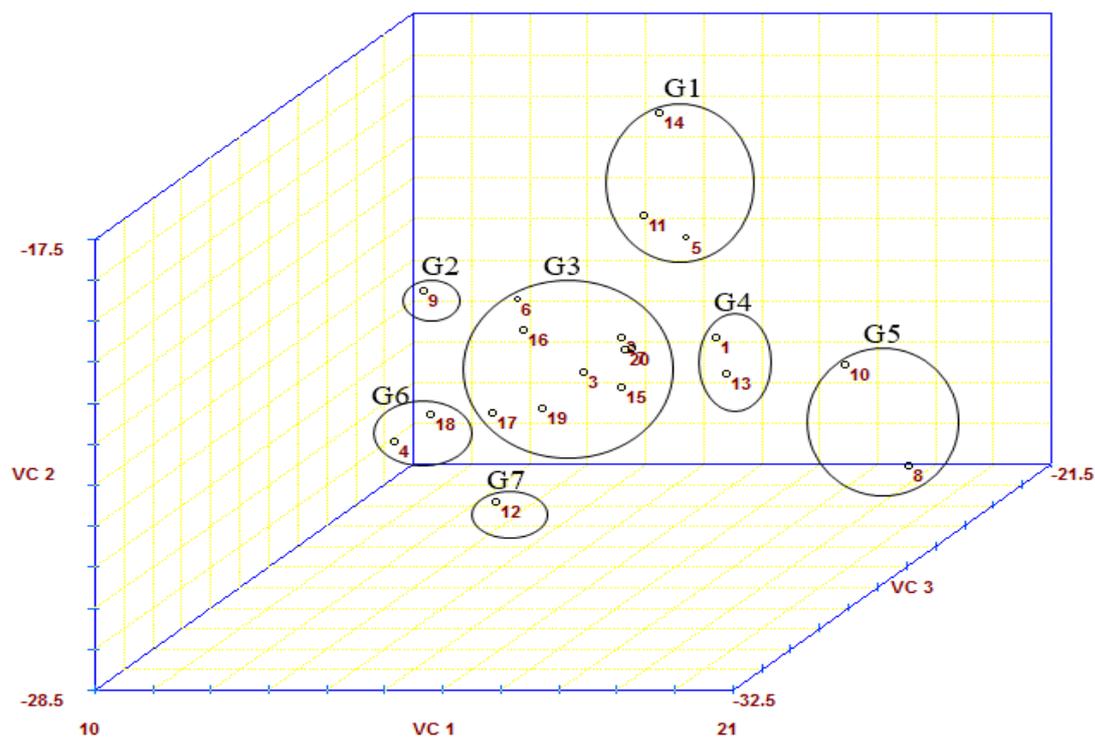
Variáveis Canônicas	Autovalores $\lambda_i$ (Variância)	Variância acumulada (%)
VC <sub>1</sub>	24,89	24,89
VC <sub>2</sub>	22,08	46,97
VC <sub>3</sub>	13,58	60,55
VC <sub>4</sub>	12,89	73,44
VC <sub>5</sub>	6,78	80,22
VC <sub>6</sub>	5,68	85,90
VC <sub>7</sub>	4,07	89,97
VC <sub>8</sub>	3,34	93,31
VC <sub>9</sub>	2,32	95,63
VC <sub>10</sub>	1,52	97,15
VC <sub>11</sub>	1,01	98,16
VC <sub>12</sub>	0,63	98,79
VC <sub>13</sub>	0,47	99,26
VC <sub>14</sub>	0,44	99,70
VC <sub>15</sub>	0,30	100,0

Teixeira (2011), em um trabalho de avaliação de acessos de açaizeiro, obteve resultados de 40,32% de variação na primeira variável canônica, necessitando utilizar a quarta variável canônica para atingir a porcentagem de 80%. Shimoya et. (2002) no estudo da divergência genética entre acessos capim-elefante, obteve cerca de 60% da variância total nas três primeiras variáveis canônicas, onde realizou a interpretação gráfica tendo boa percepção sobre os grupos, a variação de 80% ocorreu apenas na sétima variável canônica. Trabalhos estes que corroboram com os resultados encontrados no presente estudo.

A baixa variância nas primeiras variáveis canônicas, pode ser dada aos caracteres avaliados, que possivelmente não foram suficientes para discriminar a divergência genética contida entre as 20 linhagens de milho pipoca em estudo. Possivelmente, existem outros descritores que poderiam contribuir para o aumento da discrepância entre as linhagens, como a largura, espessura e comprimento de grãos (Coimbra et al., 2010) e o índice de área foliar (Simon, et al., 2012) para a cultura do milho, que demonstraram ser como caracteres de grande importância relativa para a divergência. Para o milho pipoca o número de folhas acima da espiga apresentou grande contribuição para a variância no trabalho realizado por Gonçalves (2016).

No gráfico de dispersão das linhagens, elaborado com base nas três primeiras variáveis canônicas (Figura 8), observa-se a formação de sete grupos, sendo o grupo I formado pelas linhagens 5, 11 e 14; grupo II linhagem 9; grupo III linhagens 6, 16, 20, 2, 7, 3, 15, 17 e 19, sendo o grupo mais numeroso; grupo IV linhagens 1 e 13; grupo V linhagens 10 e 8; grupo VI linhagens 4 e 18; e grupo VII linhagem 12.

As linhagens 9 e 12, mantiveram-se em grupos isolados por meio das variáveis canônicas, sendo concordante com os métodos de Tocher e de UPGMA. Para o agrupamento das linhagens 8 e 10 (grupo V), também houve concordância com o método UPGMA, e parcialmente com o método de Tocher, devido a presença da linhagem 13 no agrupamento (grupo II de tocher). As linhagens 11 e 14 que formaram um grupo pelo método de UPGMA (grupo III), pelas variáveis canônicas formaram o mesmo grupo com a incorporação da linhagem 5 formando o grupo I (Figura 8). As linhagens 4 e 18 que faziam parte dos grandes grupos nos métodos de Tocher e UPGMA, nas variáveis canônicas foram isolados e passaram a formar o grupo VI.

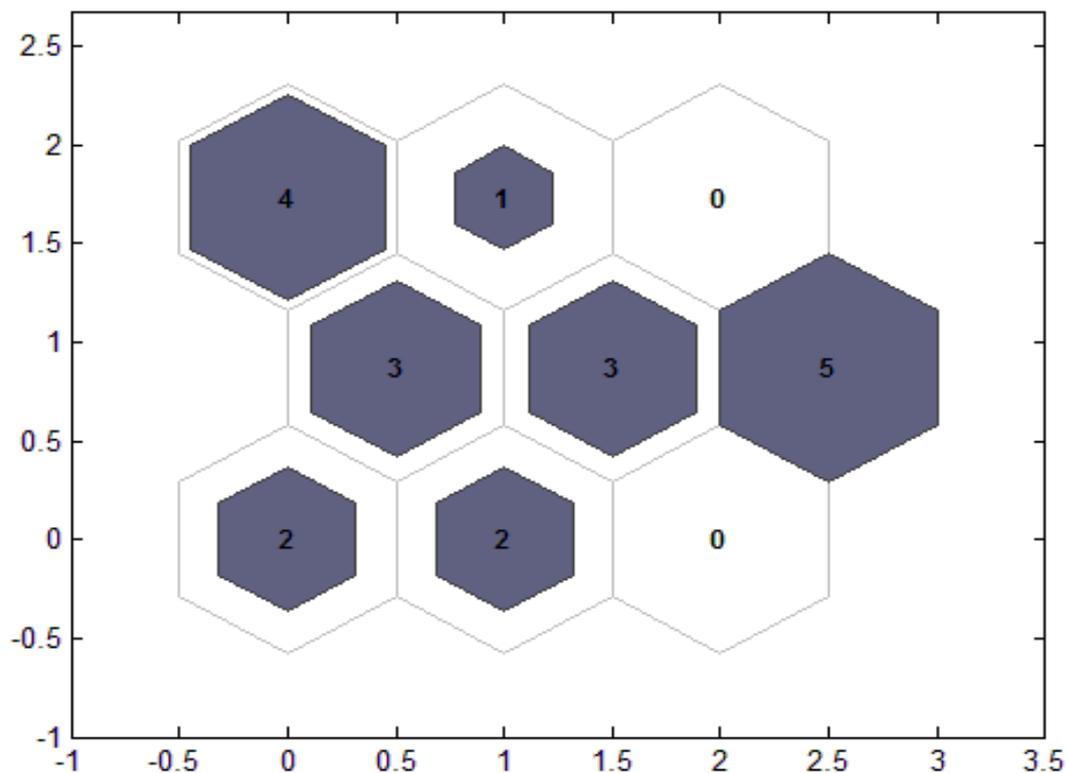


**Figura 8.** Dispersão de escores de 20 linhagens S<sub>2</sub> de milho em relação a três variáveis canônicas (VC1, VC2 e VC3), tendo como base em 15 caracteres morfoagronômicos.

Santos et al. (2011) realizou um trabalho de avaliação de divergência genética entre 48 acessos de soja, com o intuito de identificar as combinações mais promissoras para produzir recombinações superiores, utilizando os métodos de agrupamento de Tocher, UPGMA e variáveis canônicas. Por meio dos métodos, foi possível verificar que existe variabilidade entre os acessos testados, e a identificação para as hibridações de interesse. Os resultados dos métodos de agrupamento de Tocher, UPGMA e variáveis canônicas foram parcialmente concordantes entre si. O mesmo foi observado por Vogt et al. (2012) em relação aos métodos em um trabalho realizado de divergência com a cultura de girassol.

### 5.1.3 Redes neurais artificiais

Por meio da técnica de redes neurais artificiais, empregando o método de mapas de Kohonen, dos nove grupos preestabelecidos com 3 linhas e 3 grades para o comando, sete classes, formando sete grupos respectivamente (Figura 9).



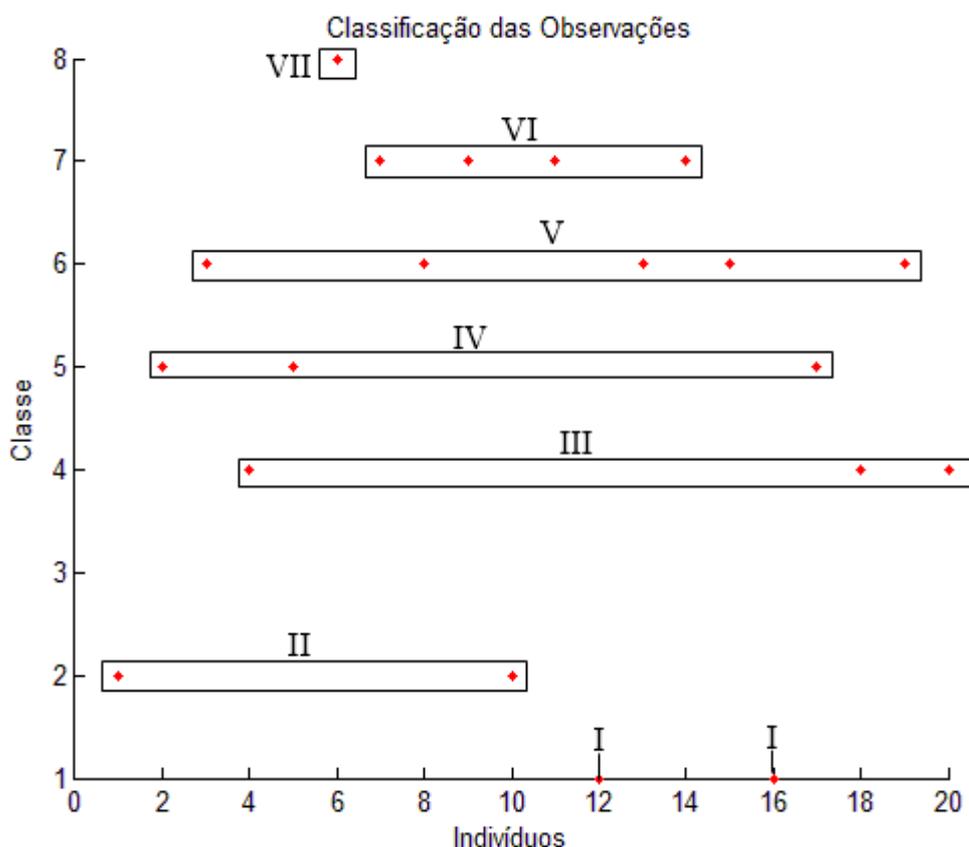
**Figura 9.** Mapa auto-organizável de Kohonen para nove classes de classificação, formadas por meio de rede neural artificial.

Pode-se observar que as linhagens foram organizadas em diferentes classes, que são representadas por meio dos hexágonos, o preenchimento da área dos hexágonos indica a concentração das linhagens contidas naquela classe, quanto maior a concentração de linhagens, maior a área preenchida, os grupos que possuem a maior quantidade de linhagens, é representado pelo preenchimento completo do hexágono. As classes onde as linhagens são alocadas no mapa, são consideradas regiões de classificações, onde as classes próximas, ou seja, que constituem divisas de espaços, são contidas de linhagens que possuem certa similaridade com a classe vizinha (primeiro vizinho), as classes mais divergentes constituem as regiões dos extremos (não cooperam) e as classes intermediárias ficam constituídas ao centro dos mapa (segundo vizinho). Desse modo, linhagens contidas nas classes dos extremos, podem ser consideradas as mais divergentes.

A classe 1 e 2 (linha 1 coluna 1, linha 1 coluna 2, respectivamente), foram constituída ambas por duas linhagens; para a classe 3 (linha 1 coluna 3), não ocorreu classificação, ou seja nenhuma linhagem foi incluída nesse grupo; para as classes 4 e 5 (linha 2 coluna 1, linha 2 coluna 2, respectivamente), foram contidas três linhagens; na classe 6 (linha 2 coluna 3), foram contidas cinco linhagens, sendo esta a classe com o

maior número de linhagens; para a classe 7 (linha 3 coluna 1), foram classificadas quatro linhagens; para a classe 8 (linha 3 coluna 2), apenas uma linhagem; e para a classe 9 (linha 3 coluna 3) não ocorreu a classificação de linhagens (Figura 9).

Na figura 10 podemos observar a classificação das 20 linhagens de milho pipoca nas classes e a constituição dos sete grupos preenchidos, por meio do mapa de Kohonen. O grupo I foi constituído com as linhagens 12 e 16, por possuírem similaridade na altura de planta (AP) e volume por pipoca (VP); o grupo II foi formado com as linhagens 1 e 10; o grupo III linhagens com as 4, 18 e 20, que possuem os valores de VP similares; o grupo IV com as linhagens 2, 5 e 17; grupo V linhagens 3, 8, 13, 15 e 19, que se caracterizam por possuírem as maiores produtividade e volume de pipoca (VP); grupo VI linhagens 7, 9, 11 e 14, que possuem os menores produtividades e menores VP; e o grupo VII formado pela linhagem 6, com maior DC e menor florescimentos. Os valores médios dos caracteres para cada linhagem dos grupos podem ser observados no anexo 2.



**Figura 10.** Mapa auto-organizável de Kohonen, classificação das observações dos agrupamentos das linhagens nas classes formadas por meio de rede neural artificial.

O número de linhagens variou em cada grupo, percebe-se que o grupo V foi o mais numeroso, possuindo 5 das 20, enquanto o grupo VII possuiu apenas uma. A diferença da quantidade de linhagens dentro de em cada grupo se deve a variação própria dos caracteres mensurados.

Barbosa et al. (2011), utilizou o método de Kohonen para o estudo da diversidade genética, visando agrupamentos e identificação de divergência entre 48 acessos de mamoeiro (*Carica papaya* L.). De acordo com autores, o método foi eficiente para a avaliação da divergência genética e seleção indivíduos com características superiores e favoráveis para a obtenção de novas cultivares de mamão, prevendo progresso genético. Clovis et al. (2015) utilizou o mapa de Kohonen, na cultura do milho para a análise de estabilidade, obtendo sucesso nos resultados com a utilização do método.

A técnica de redes neurais, também tem sido empregada em trabalhos para selecionar indivíduos dentro de famílias de cana-de-açúcar (Brasileiro et al., 2015); prever o volume de eucalipto (Bhering et al., 2015); identificação de linhagens de feijão-caupi semiprostrado com alta adaptabilidade e estabilidade fenotípicas (Teodoro et al., 2015); dentre outros.

Os grupos encontrados neste trabalho por meio da técnica, diferiram-se aos formados nos métodos de análise multivariada de otimização de Tocher, hierárquico UPGMA, vizinho mais distante e variáveis canônicas.

Por meio da rede dos mapa de Kohonen a linhagem 6, foi a única isolada em um grupo VII (classe 8), indicando que por esta metodologia é a mais divergente em relação as demais, a linhagem possui como características principais o menor ciclo de florescimento feminino e masculino e maior diâmetro de colmo.

Em relação aos cruzamentos intergrupos, a linhagem do grupo VII é a mais indicada para cruzamentos com as linhagens do grupos I, III e V (classe 6), devido a distância genética entre as mesmas. O grupo VI apresentou-se como o mais próximo ao grupo VII, indicando menor dissimilaridade genética, e intermediário aos grupos III e IV (classes 4 e 5), como pode ser observado na figura 9. Os grupos V e VI também se apresentam com alta dissimilaridades entre os caracteres, sendo as linhagens contidas, passíveis para cruzamentos.

Para os métodos de Tocher, variáveis canônicas, UPGMA e vizinho mais distante, as linhagens 9 e 12 ficaram isoladas cada uma em um grupo, enquanto para Kohonen a linhagem 6 foi a única isolada em um grupo. O método de Kohonen

concordou parcialmente com o agrupamento das linhagens 11 e 14 pelo método de UPGMA e variáveis canônicas, ao ficarem contidas dentro de um mesmo agrupamento. Por meio do método de Kohonen as linhagens 4, 18 e 20 formaram um grupo, concordando parcialmente com as variáveis canônicas que reuniu as linhagens 4 e 18 em um grupo, enquanto que pelos métodos de UPGMA e Tocher, as linhagens ficaram contidas em um grande grupo, o mesmo ocorrendo com as linhagens 2, 5 e 7.

De Melo (2015) em um estudo de divergência genética em acessos de melão utilizando redes neurais artificiais de Kohonen, relatou a diferença entre os grupos encontrados em seu trabalho, que se diferiram profundamente aos encontrados por Torres Filho (2008), com os mesmos dados utilizados, aplicando o método de análise multivariada, ressaltando as principais diferenças em relação ao método de Tocher.

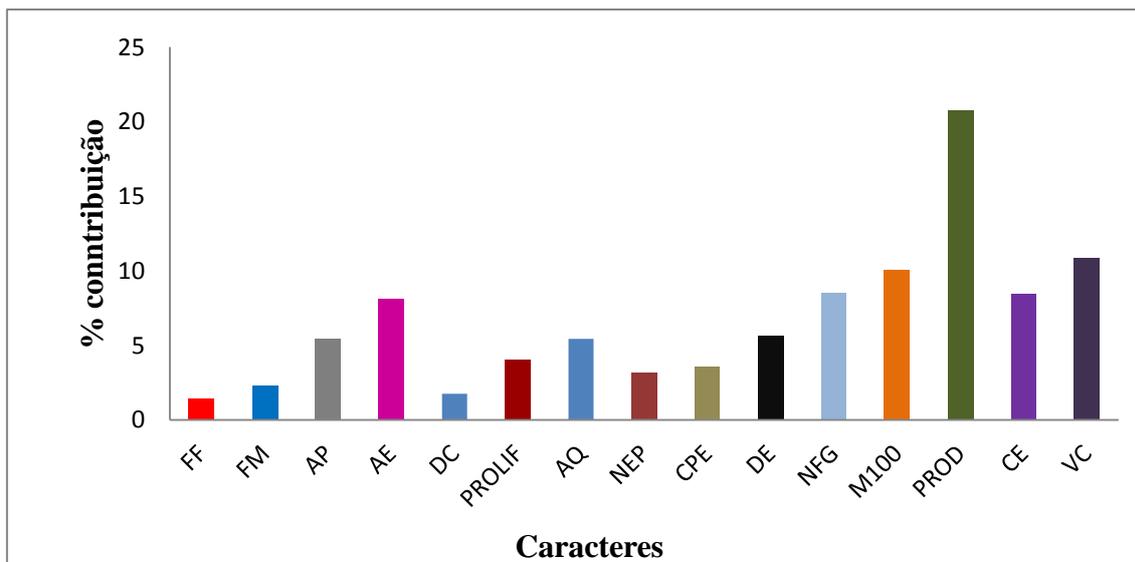
Esta diferença pode ser explicada devido as redes neurais artificiais possuírem uma estrutura não-linear, em que elas conseguem captar características mais profundas e complexas dos dados sejam quantitativos ou qualitativos, o que nem sempre é possível com a utilização das técnicas estatísticas tradicionais segundo Galvão et al. (1999).

#### **5.1.4 Importância relativa dos caracteres**

A análise para estimar a contribuição relativa de quinze caracteres morfoagronômicos para a expressão da divergência genética na safra I, foi estimada segundo o método de Singh (1981) baseado na distância Generalizada de Mahalanobis,

Os caracteres de PROD (20,80%), VP (10,90%) e M100 (10,10%) foram os que mais contribuíram para a divergência total entre as 20 linhagens S<sub>2</sub> de milho pipoca avaliados, conforme apresentado na figura 11. A CE (8,46%), também apresentou uma contribuição importante sobre a divergência.

São justificáveis os caracteres de produtividade proporcionarem a maior contribuição relativa, visto que a análise reuniu linhagens distintas, em que as variações foram expressamente contrastantes, em relação a produtividade. As médias variaram entre 1073,37 (linhagem 8) a 267,72 kg ha<sup>-1</sup> (linhagem 11), e para a massa de cem grãos ocorreram variações entre 19,55 (linhagem 8) a 8,54 gramas (linhagem 9) .



**Figura 11.** Contribuição relativa de quinze caracteres morfoagronômicos de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP), avaliados para divergência ( $D^2$ ) entre 20 linhagens  $S_2$  de milho pipoca, na safra de 2014/2015, em Dourados/MS.

A capacidade de expansão é uma característica importante para o milho pipoca, tendo em vista que é uma das características que mais contribuíram para a identificação do grupo ao qual as linhagens pertencem, variando entre 26,94 (linhagem 3) a 12,66 mL/g (linhagem 14). Quando se pretende identificar linhagens com características superiores e de melhor aceitação no mercado consumidor, deve se considerar as características de maior interesse pelo consumidor e pelo produtor, sendo para a cultura do milho pipoca, a produtividade e a capacidade de expansão.

Os caracteres florescimento feminino, florescimento masculino e diâmetro de colmo, foram sugeridos como passíveis de descarte, de acordo com o método de Singh, uma vez que foram pouco informativos para a avaliação da dissimilaridade genética entre as linhagens, já que às estimativas de contribuição relativa foram de pequena magnitude.

A importância relativa dos caracteres é de fundamental interesse dos pesquisadores, possibilitando concluir com confiabilidade sobre a viabilidade de empregar os descritores utilizados em estudos de divergência genética, reduzindo o número de características, e eliminando aquelas que contribuem pouco para o estudo (Cruz et al., 2011).

Rotili et al. (2012), em um trabalho de divergência genética com genótipos de milho, ao analisar a contribuição das características para a divergência, observou que a

característica que mais contribuiu para a divergência foi a produtividade (21,8%), concordando com os resultados encontrados neste trabalho. Arnhold et al. (2010), no estudo de seleção de linhagens S<sub>5</sub> de milho pipoca com base em desempenho e divergência genética, os caracteres de produtividade de grãos, massa de cem grãos e capacidade de expansão se apresentaram entre os de maiores índices de importância relativa sobre a divergência, corroborando com os resultados obtidos no presente estudo.

### **5.1.5 Considerações experimento linhagens S<sub>2</sub> (safra I)**

Por meio das metodologias empregadas, foi possível evidenciar a existência de variabilidade genética entre as linhagens S<sub>2</sub> avaliadas para o período da safra I (2014/2015). Os métodos de agrupamentos por técnicas multivariadas apresentaram semelhanças entre si para representar a divergência genética entre as linhagens avaliadas, sendo parcialmente concordantes nos agrupamentos. A maior semelhança entre o método de redes neurais pelo mapa de Kohonen em relação aos métodos multivariados, deu-se aos resultados apresentados por meio das variáveis canônicas e a maiores diferenças em relação ao método de Tocher..

Para os métodos de Tocher, variáveis canônicas, UPGMA e vizinho mais distante, as linhagens 9 e 12 ficaram em grupos isolados, enquanto a linhagem 6 foi alocado nos grandes grupos, nas referidas metodologias. O método de Kohonen concordou parcialmente com o agrupamento das linhagens 11 e 14 com o método de UPGMA, vizinho mais distante e variáveis canônicas, ao ficarem contidos dentro de um mesmo agrupamento. Por meio do método de Kohonen as linhagens 4, 18 e 20 formaram um grupo, concordando parcialmente com as variáveis canônicas que reuniu as duas linhagens 4 e 18 em um grupo, onde por UPGMA e o Tocher as linhagens ficaram contidos em um grande grupo.

Para a escolha dos genitores deve ser levado em consideração a sua dissimilaridade genética, bem como suas características biológicas. As linhagens mais divergentes e potenciais para serem indicadas como genitoras em futuros cruzamentos são: linhagem 9, que apresentou a menor altura de planta (AP), a menor produtividade (PROD) e volume de pipoca (VP); linhagem 12, que tem como característica o maior número de fileira de grãos (NFG) e número de espigas autofecundadas por parcelas (NEAP); linhagem 15, com a maior produtividade (PROD), volume por pipoca (VP), capacidade de expansão (CE) e prolificidade (PROLIF); e a linhagem 6 que apresenta

maior diâmetro de colmo (DC) e menores dias para o florescimento feminino (FF) e florescimento masculino (FM).

A indicação da linhagem 9 pode ser interessante, apesar dos valores médios baixos para PROD e VP, uma vez que nos cruzamentos divergentes existe a maior probabilidade de restabelecimento da heterozigose nos diversos locus gênicos. A escolha de genitores que possivelmente apresentem genes com locus homozigotos recessivos é uma maneira estratégica, para se obter o maior êxito nos cruzamentos com linhagens que apresentem a maioria de alelos dominantes e assim, obter o vigor híbrido e manifestação das características de interesse na descendência.

## **5.2. Experimento linhagens S<sub>3</sub>, safra II (2016/2017)**

### **5.2.1 Análises univariadas**

Verificou-se diferença estatística significativa entre as linhagens a 1% de probabilidade para as características FF, FM, AP, DC, AQ, NEAP e CE, e significância de 5% de probabilidade para PROLIF, NFG e VP, indicando a existência de variabilidade para essas características entre as linhagens avaliadas (Tabela 6). Para AE, CPE, M100 e PROD, não foi verificado efeito significativo entre as linhagens, indicando a inexistência de variabilidade entre as médias para essas características.

As acurácias estimadas, em geral, para cada caráter variaram acima de 64,49% a 96,77%, indicando elevada magnitude, exceto para CPE e DE, apontando uma boa precisão experimental, de acordo com os limites de classes estabelecidos em Cargnelutti Filho e Storck (2009). Quanto maior a acurácia seletiva, maior é a confiança na avaliação dos dados para o experimento.

**Tabela 6.** Resumo das análises de variância para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S<sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados/MS.

Caracteres	Quadrado Médio				
	Blocos (2) <sup>1</sup>	Linhagens (19) <sup>1</sup>	Resíduo (38) <sup>1</sup>	Média	Acurácia (f <sub>gg</sub> )
FF	37,26	22,00 <sup>**</sup>	7,88	77,28	80,11
FM	36,88	26,08 <sup>**</sup>	5,46	76,53	88,91
AP	1002,68	635,51 <sup>**</sup>	125,81	131,67	89,55
AE	650,46	151,55 <sup>ns</sup>	81,73	62,80	67,87
DC	107,33	6,76 <sup>**</sup>	2,48	16,99	79,58
PROLIF	2,37	0,21 <sup>*</sup>	0,09	2,07	74,18
AQ	9,65	6,54 <sup>**</sup>	2,42	5,85	79,35
NEAP	74,61	11,40 <sup>**</sup>	2,68	4,97	87,42
CPE	0,71	3,44 <sup>ns</sup>	2,66	9,561	47,50
DE	23,37	18,55 <sup>ns</sup>	16,34	24,70	34,53
NFG	11,27	3,82 <sup>*</sup>	1,89	11,11	71,00
M100	7,02	12,9594 <sup>ns</sup>	12,2956	13,19	71,55
PROD	17585,2395	13298,5623 <sup>ns</sup>	7767,5671	171,06	64,49
CE	2,81	34,5425 <sup>**</sup>	2,19	18,30	96,77
VP	6152573,5396	5219535,4921 <sup>*</sup>	2568607,1412	3112,63	71,26

<sup>1</sup> Grau de liberdade. <sup>ns</sup>, <sup>\*\*</sup>, <sup>\*</sup>, Não significativo, Significativo a 1% e 5% de probabilidade respectivamente, pelo teste F.

## 5.2.2 Análises Multivariadas

### 5.2.2.1 Distância generalizada de Mahalanobis

As medidas de dissimilaridade genética (D<sup>2</sup>), apresentaram magnitude de 9,71 a 156,72 (Tabela 7). As amplitudes dos valores das estimativas de dissimilaridade de Mahalanobis observadas sugerem haver ampla variabilidade genética, viabilizando desta forma, a escolha de genitores contrastantes para extração de linhagens endogâmicas ou utilização direta das mesmas (Simon et al., 2012). As combinações entre as linhagens 2 x 11 (156,72), e 3 x 2 (149,24), foram os pares mais divergentes,

apresentando as estimativas mais elevadas da distância generalizada de Mahalanobis. A menor divergência foi observada entre os pares 3 x 4 (9,71), e 8 x 10 (10,78).

Valores de elevada dissimilaridade, também foram encontrados para as combinações entre as linhagens 16 x 2 (145,40), 4 x 2 (125,40), 5 x 9 (115,64) e 7 x 9 (115,30), indicando maior divergência genética entre essas linhagens. Menores dissimilaridades também foram observadas entre as combinações 1 x 18 (11,50), 5 x 5 (13,30), 7 x 18 (13,63), e 19 x 17 (14, 39), indicando que estes cruzamentos devem ser evitados, uma vez que diminuem a chance de sucesso de reestabelecimento de heterose na descendência.

A distância de Mahalanobis é amplamente empregada em trabalhos para a avaliação da divergência, pois esta, mostra-se uma medida de cálculo de distância eficiente para a discriminação da dissimilaridade genética. Para a cultura do milho pipoca, tem-se encontrado na literatura, diversos trabalhos com a utilização desta medida para a estimativa da divergência genética, como, Miranda et al. (2003), De Paula (2009), Arnhold et al. (2010), Vittorazzi (2016), Melo et al. (2017), dentre outros.

**Tabela 7.** Estimativas da dissimilaridade ( $D^2$ ) para as 20 combinações de linhagens mais similares e mais divergentes, respectivamente, com base na avaliação de 15 caracteres morfoagronômicos quantitativos, em 20 linhagens S<sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados/MS.

Linhagem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	87,19	33,87	27,47	39,57	49,08	31,96	20,57	79,51	27,02	18,33	21,35	27,22	21,52	66,73	38,09	32,82	11,50	24,99	37,07
2		-	149,24	125,40	43,22	41,68	34,38	99,06	142,59	115,05	156,72	112,38	41,42	86,38	65,78	145,40	98,36	56,53	94,29	51,31
3			-	9,71	88,39	96,81	76,96	34,26	50,65	45,06	17,58	51,58	63,96	31,63	65,26	22,13	22,38	33,76	26,75	77,71
4				-	77,35	79,25	65,56	39,79	40,08	48,13	23,63	63,01	51,76	26,03	48,22	16,93	20,95	26,90	21,17	20,56
5					-	13,30	33,66	32,94	115,64	37,44	79,50	37,25	16,56	60,95	60,24	100,34	49,81	27,35	36,58	39,04
6						-	40,35	44,86	87,70	40,70	91,73	45,92	23,75	56,58	39,87	102,45	51,23	36,70	37,45	39,76
7							-	50,34	115,30	59,37	74,68	60,89	26,11	45,72	57,521	65,18	54,59	13,62	46,28	25,13
8								-	69,05	10,78	20,09	16,33	39,31	33,28	57,56	61,46	40,79	23,61	24,82	47,64
9									-	81,22	71,02	98,04	91,92	47,03	25,28	60,90	55,02	69,57	53,27	71,22
10										-	23,93	19,43	48,19	51,85	74,51	76,62	51,67	36,23	23,83	64,83
11											-	32,59	66,22	39,81	89,24	42,02	46,96	35,02	33,55	73,67
12												-	40,14	40,26	86,18	79,35	43,30	34,01	38,03	72,53
13													-	30,47	46,48	67,25	25,91	16,39	28,22	34,22
14														-	37,74	33,48	23,59	22,78	31,28	44,06
15															-	59,61	37,71	38,49	38,01	29,09
16																-	29,49	31,13	39,87	54,50
17																	-	21,06	14,39	53,19
18																		-	20,05	19,84
19																			-	38,21
20																				-

### 5.2.2.2 Análises de agrupamentos pelos métodos de Tocher, ligação média não ponderada entre grupos UPGMA e vizinho mais distante

Pelo método de Tocher o grupo I foi o mais numeroso (Tabela 8), contemplando 50% das linhagens de milho pipoca avaliadas, assim como nas linhagens S<sub>2</sub> na safra I. O grupo I foi constituído pelas linhagens 3, 4, 16, 17, 18, 19, 1, 14, 11 e 8, estes que apresentaram os maiores valores médios de capacidade de expansão, variando entre 18,49 e 23,50 mL/g. O grupo II apresentou 20% dos agrupamentos, formado pelas linhagens 5, 6, 13 e 7; o grupo III foi formado com as linhagens 10 e 12, com semelhantes para diâmetro médio de colmo (27,35 e 27,57mm). O grupo IV contemplou as linhagens 9 e 15, que apresentaram a mesma média de dias para o florescimento feminino (FF) (78 dias) e valores semelhantes para a massa de cem grãos (M100) com valores de 9,95 e 9,06 g respectivamente.

**Tabela 8.** Agrupamento pelo método de otimização de Tocher, com base na distância generalizada de Mahalanobis de 20 linhagens S<sub>3</sub> de milho pipoca, na safra de 2016/2017, em Dourados/MS.

Grupos	Linhagens										% linhagens
I	3	4	16	17	19	18	1	14	11	8	50
II	5	6	13	7							20
III	10	12									10
IV	9	15									10
V	20										5
VI	2										5

Os grupos V e VI apresentaram agrupamentos contendo apenas uma única linhagem, sendo a 20 e a 2, respectivamente, indicando que estas, apresentam a maior dissimilaridade genética em relação as demais, para eventuais cruzamentos, segundo o método proposto por Tocher. A linhagem 2 apresentou como a menor média de capacidade de expansão (CE) com 11,66 mL/g em relação as demais e a linhagem 20 apresentou um dos menores valores médios para plantas acamadas e quebradas (AQ) (1,0) além da menor quantidade de autofecundadas espigas por parcelas (NEAP) com 2,33 espigas, média.

Vasconcelos et al. (2007) ressaltaram que o método de Tocher apresenta a peculiaridade de reunir um maior número de linhagens nos primeiros grupos e geralmente indivíduos agrupados isoladamente nos últimos grupos. Nesse tipo de estudo, tal característica do método torna-se interessante, pois permite identificar indivíduos geneticamente dissimilares e não somente grupos.

Em outros trabalhos de divergência genética, por meio da utilização do método de otimização de Tocher, resultados semelhantes no padrão de agrupamento dos primeiros grupos, foram obtidos por Silva et al. (2016), para a dissimilaridade genética e definição de grupos de recombinação em progênies de meios-irmãos de milho verde; Rigon et al. (2015) entre híbridos de milho; Santos et al. (2015) com acessos de cupuaçu e Sulzbacher et al. (2017) na análise de divergência genética em feijão comum.

A partir do método de otimização de Tocher foi estimada a dissimilaridade intra e intergrupos (Tabela 9). A maior distância média contida intragrupo foi observada no grupo I (28,25), enquanto que a menor foi verificada no grupo III (19,43). As distâncias obtidas, atenderam ao critério estabelecido para o método de otimização de Tocher, onde os valores para as distâncias intragrupos foram inferiores a qualquer distância intergrupos (Cruz et al., 2014).

**Tabela 9.** Distâncias médias intra e intergrupos estimadas pelo método de otimização de Tocher, com base na dissimilaridade entre 20 linhagens S<sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados/MS.

Grupos	I	II	III	IV	V	VI
I	28,25	54,51	40,75	56,73	50,20	<b>109,86</b>
II		25,62	46,24	76,84	34,54	40,18
III			19,43	84,99	68,68	<b>113,72</b>
IV				25,28	50,16	<b>104,19</b>
V					-	51,31
VI						-

\*Em negrito estão destacadas as maiores distâncias médias intergrupos.

As maiores distâncias intergrupo encontradas foram entre os grupos III e VI (113,72), I e VI (109,86), e IV e VI (104,19), correspondendo aos grupos que possuem as maiores dissimilaridades genéticas, indicando provavelmente as combinações mais promissoras para cruzamentos. Segundo Barili et al. (2011), quanto maior o grau de divergência entre os genitores, maior será a variabilidade resultante na população segregante, e maior é a probabilidade de reagrupar os alelos em novas combinações favoráveis.

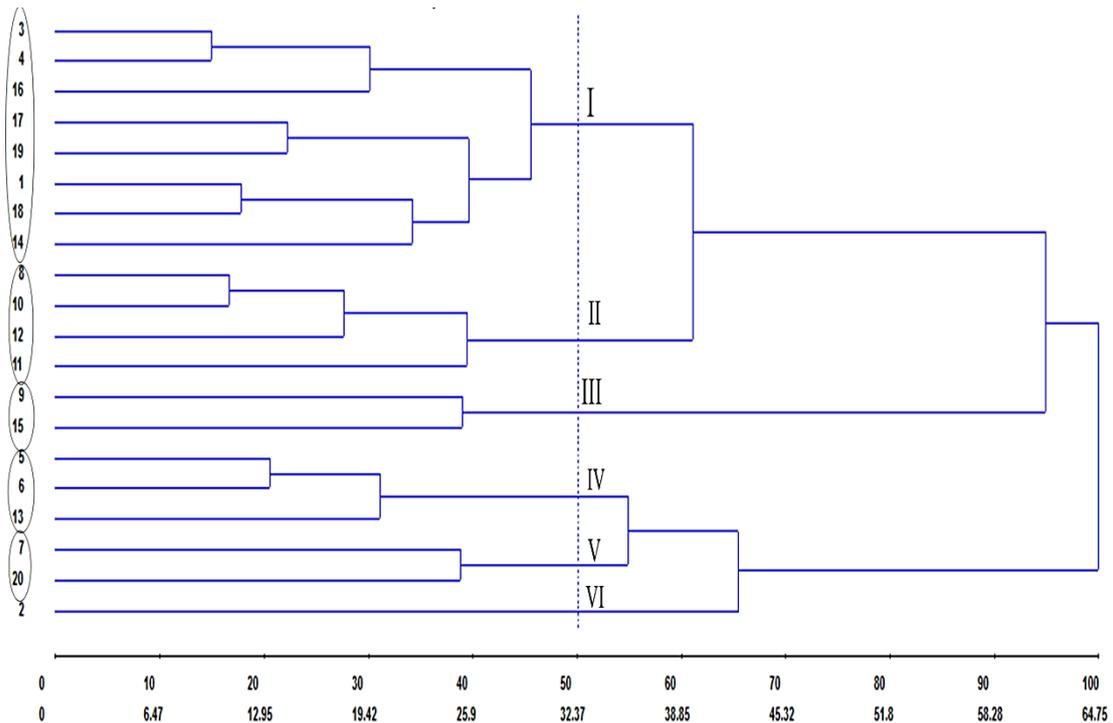
O grupo VI, contendo a linhagem 2, foi indicado para os três cruzamentos mais dissimilaridades. Isso indica que esta linhagem apresentou alta divergência genética em relação às demais, sendo interessante para utilização como genitora, pois possivelmente contribuirá para a maximização do vigor com o restabelecimento da heterose nos cruzamentos e ampliação da base genética.

Contudo, além de se levar em consideração as distâncias, é necessário que as linhagens destinadas aos cruzamentos associem elevado potencial para as características de interesse para a cultura do milho pipoca, como a produtividade e a capacidade de expansão. Atendido este requisito, existe grande possibilidade de seleção de genótipos transgressivos, devido à ocorrência de heterose e a ação de genes complementares (Falconer & Mackay, 1996; Carvalho et al., 2001). Sendo assim, torna-se interessante o cruzamento entre as linhagens dos grupos I e VI para a capacidade de expansão, uma vez que estes contêm os maiores e menor valor para esse caráter, respectivamente.

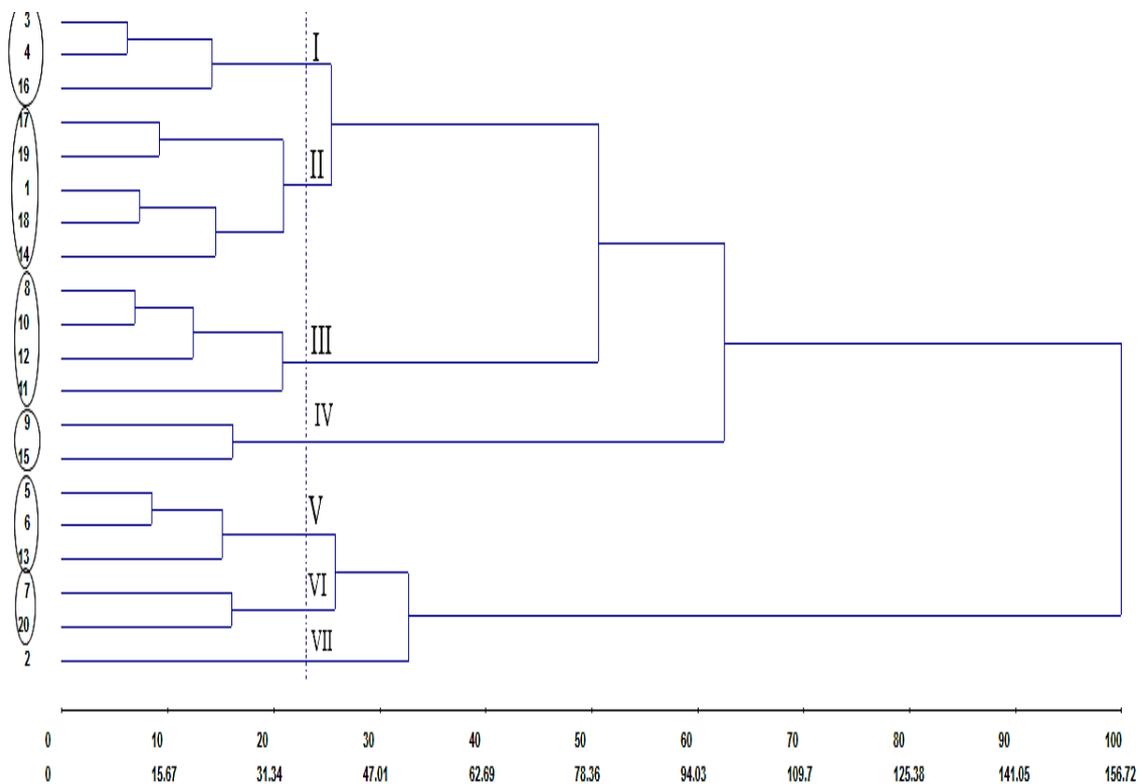
As menores distâncias intergrupo foram obtidas entre os grupos II e V (34,54), II e VI (40,18), e I e III (40,75) e II e III (46,24). O grupo II apresentou as menores distâncias em relação a três grupos, indicando possuir certa similaridade em relação aos caracteres, cruzamentos entre linhagens dos grupos com baixas distâncias genéticas, pode apresentar baixa heterose, não sendo indicados para cruzamentos.

Pelo método de agrupamento UPGMA, o corte no dendograma foi realizado considerando 50% variabilidade, que possibilitou a divisão das linhagens em seis grupos distintos (Figura 12). Esse critério foi definido considerando a mudança abrupta de nível, conforme sugerido por Barroso e Artes (2003). O grupo I foi formado por 8 das 20 linhagens analisadas, reunindo as linhagens 14, 18, 19, 17, 16, 4, 3, 1. O grupo II contou com as linhagens 8, 10, 11 e 12, apresentando o menor número de dias para florescimento masculino (FM). O grupo III pelas linhagens 15 e 9. O grupo IV 13, 6 e 5. O grupo V 7 e 20. O grupo VI com a linhagem 2, com menor capacidade de expansão (CE). Os valores médios dos caracteres para cada linhagem dos grupos podem ser observados no anexo 3.

Para o agrupamento pelo método do vizinho mais distante (Figura 13), com o corte realizado a cerca de 25% de variação, houve a formação de sete grupos, com grande concordância aos realizados pelo método de UPGMA. A linhagem 2 permaneceu isolada nesse agrupamento, assim como os agrupamentos realizados para as linhagens 7 e 20, linhagens 5, 6, 15, linhagens 9 e 15, linhagens 8, 20, 11 e 12 que também foram condizentes



**Figura 12.** Dendrograma representativo da dissimilaridade genética entre 20 linhagens  $S_3$  de milho pipoca, obtidas pela ligação média não ponderada entre grupos (UPGMA), a partir da medida de dissimilaridade generalizada de Mahalanobis.



**Figura 13.** Dendrograma representativo da dissimilaridade genética entre 20 linhagens  $S_3$  de milho pipoca, obtidas pela ligação completa do vizinho mais distante, a partir da medida de dissimilaridade generalizada de Mahalanobis.

A diferença entre as duas metodologias se deu em relação a formação do grande grupo, em que pelo UPGMA as linhagens 14, 18, 18, 17, 4, 3 e 1 foram reunidas em um mesmo grupo (grupo I), contudo pelo vizinho mais distante as linhagens 3, 4 e 16 passaram a formar um novo grupo (grupo I) e as linhagens 17,18, 19, 14 e 1 permaneceram reunidas em um mesmo agrupamento (grupo II).

Diante dos resultados observados, pode-se enfatizar que os agrupamentos realizados pelo método de UPGMA foram consistentes aos encontrados pelo método do vizinho mais distante, inferindo sobre uma classificação eficiente.

Os métodos de Tocher, UPGMA, também demonstraram semelhança no agrupamento das linhagens. A linhagem 2 foi isolada em ambos os métodos. Também, houve concordância em relação ao agrupamento das linhagens 9 e 15. As diferenças nos agrupamentos, foram em relação a linhagem 20 que formou um grupo isolado no método de Tocher e, pelo UPGMA a linhagem 20 foi agrupada com a linhagem 7 (grupo V). As linhagens 10 e 12 que constituíram o grupo III no método de Tocher, no método UPGMA foram alocados para o grupo II, juntamente com as linhagens 11 e 8.

### **5.2.2.3 Dispersão gráfica de variáveis canônicas**

Dentre as 15 variáveis canônicas, as três primeiras variáveis explicaram mais de 76% da variação total, sendo para a primeira 40,22%, a segunda 23,03% e para a terceira 13,10% (Tabela 10). Segundo Cruz et al. (2014), é desejável para a descrição da divergência que as duas primeiras variáveis expliquem pelo menos 80% da variação total, porém, em casos em que este limite não é atingido, a variação pode ser completada com a dispersão em relação a terceira e quarta variáveis canônicas.

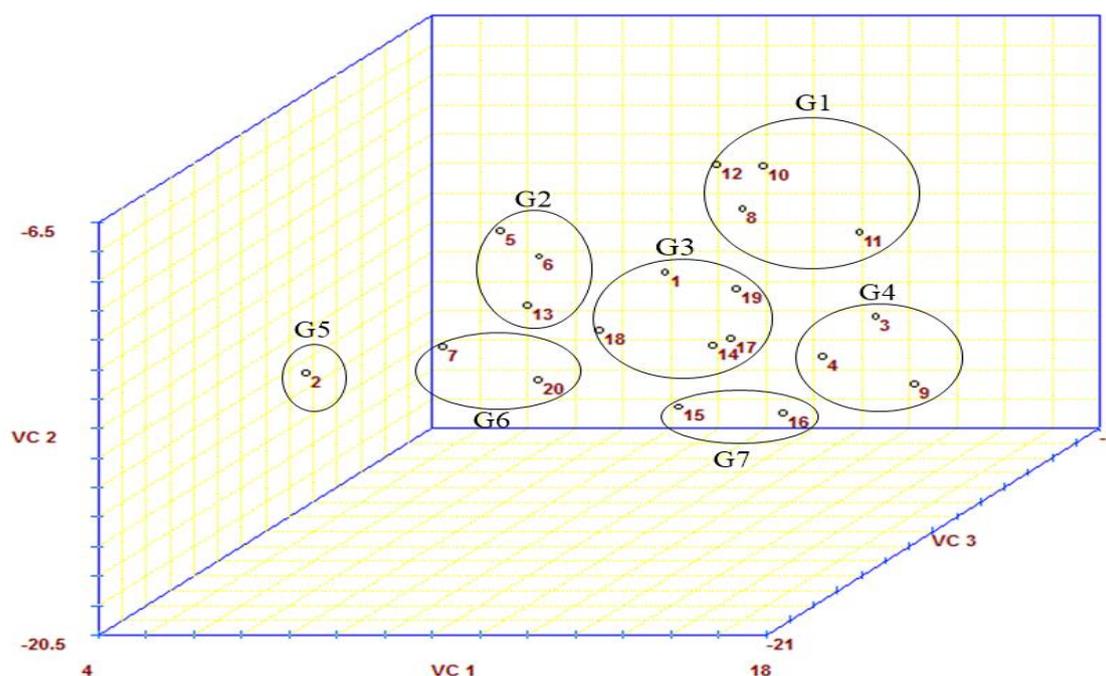
Neste trabalho, esta porcentagem (80%) foi verificada na quarta variável canônica. Contudo, optou-se por manter as três primeiras variáveis canônicas, representando 76,35% da variação total. Quando as três primeiras variáveis canônicas explicam mais de 70% da variação, os dados se adequam a uma representação gráfica tridimensional (Rêgo et al., 2003; Bento et al., 2007).

Santos et al. (2009) ao estudarem a divergência genética em *Capsicum spp* geração F<sub>2</sub>, relataram que as três primeiras variáveis explicaram cerca de 73% da variância total para as características quantitativas analisadas, ao qual foi gerado a dispersão gráfica em espaço tridimensional, resultados que corroboram com os encontrados no presente trabalho.

**Tabela 10.** Estimativas das variâncias (autovalores), variâncias percentuais acumuladas das variáveis canônicas, visando estimar a dissimilaridade genética entre 20 linhagens S<sub>3</sub> de milho pipoca, na safra de 2016/2017, em Dourados/MS.

Variáveis Canônicas	Autovalores $\lambda_i$ (Variância)	Variância acumulada (%)
VC <sub>1</sub>	40,22	40,22
VC <sub>2</sub>	23,03	63,25
VC <sub>3</sub>	13,10	76,35
VC <sub>4</sub>	6,96	83,31
VC <sub>5</sub>	5,15	88,46
VC <sub>6</sub>	3,55	92,01
VC <sub>7</sub>	2,23	94,24
VC <sub>8</sub>	2,00	96,24
VC <sub>9</sub>	1,37	97,61
VC <sub>10</sub>	1,04	98,65
VC <sub>11</sub>	0,68	99,33
VC <sub>12</sub>	0,38	99,71
VC <sub>13</sub>	0,18	99,89
VC <sub>14</sub>	0,08	99,97
VC <sub>15</sub>	0,03	100

No gráfico de dispersão das linhagens, elaborado com base nas três primeiras variáveis canônicas (Figura 14), observou-se a formação de sete grupos, sendo o grupo I formado pelas linhagens 12, 11, 10 e 8; grupo II linhagens 5, 6 e 13; grupo III linhagens 1, 14, 17, 18 e 19 sendo o grupo mais numeroso; grupo IV linhagens 4, 3 e 9; grupo V linhagem 2; grupo VI linhagens 7 e 20; e grupo VII linhagens 15 e 16.



**Figura 14.** Dispersão de escores de 20 linhagens S<sub>3</sub> de milho em relação a três variáveis canônicas (VC1, VC2 e VC3), tendo como base em 15 caracteres morfoagronômicos.

Segundo Cruz et al. (2011), cruzamentos entre genótipos de grupos distintos em relação as variáveis canônicas, resultarão em escolhas de genitores que apresentarão a maior exploração da heterose, e a distância de um grupo para o outro contribuirá ainda mais para este resultado.

A linhagem 2, manteve-se em grupo isolado por meio da metodologias das variáveis canônicas, Tocher, UPGMA e vizinho mais distante, sendo estas concordantes em relação ao mantê-lo em um grupo unitário. Para os agrupamentos das linhagens 7 e 20 (grupo VI), 12, 11, 10 e 8 (Grupo I), 5, 6 e 13 (grupo II), houve concordância em relação ao método de UPGMA com as variáveis canônicas. As linhagens 4 e 3 que faziam parte de grupos contendo várias linhagens pelos métodos de UPGMA (grupo I) e de Tocher (grupo I), pelas variáveis canônica formaram um novo grupo juntamente com a linhagem 9. As linhagens 15 e 9 que formavam um grupo nos métodos de Tocher e UPGMA, pela variáveis canônicas a linhagem 15 passou a ser reunido com a linhagem 16, e o 9 com as linhagens 4 e 3, como descrito anteriormente.

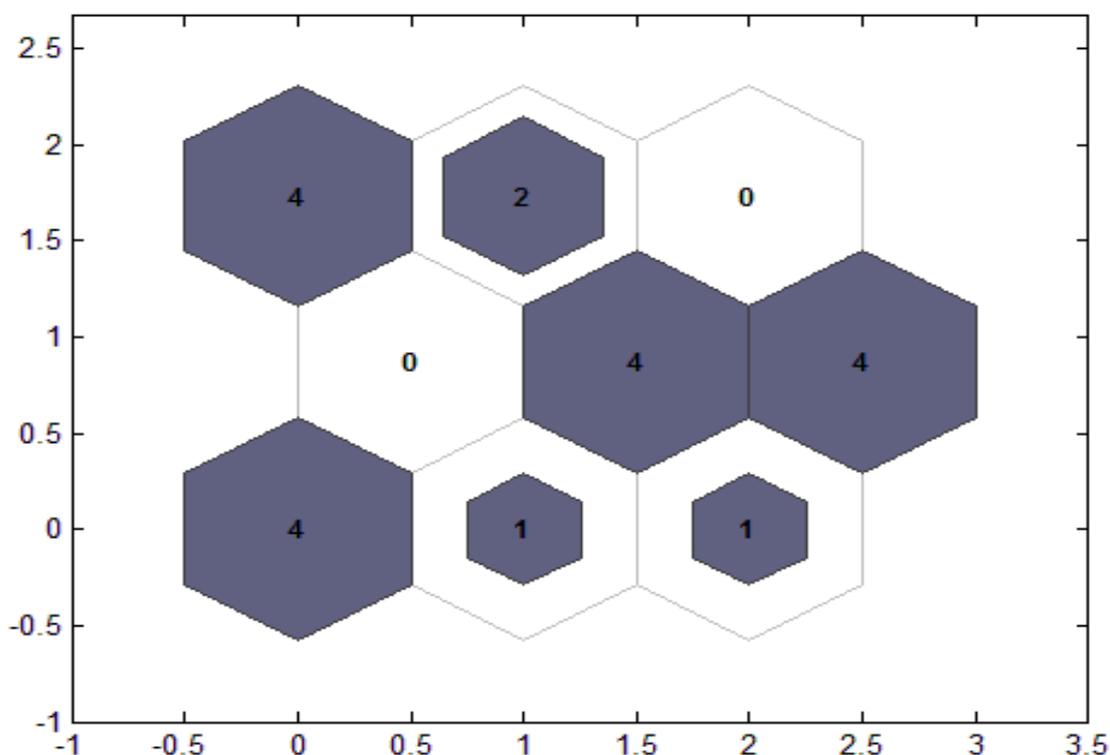
O método de variáveis canônicas apresentou maior semelhança nos resultados dos agrupamentos com o método de UPGMA. Este fato pode ser explicado em função do agrupamento de Tocher preconizar sempre as maiores distâncias entre grupos em relação à distância dentro dos grupos, o UPGMA apresenta uma formação mais complexa, referindo-se as menores distâncias, permitindo visualização de acessos mais similares dentro dos grupos (Bertan et al., 2006).

### **5.2.3 Redes Neurais Artificiais**

Por meio da técnica de redes neurais artificiais, empregando o método de Mapa auto-organizáveis de Kohonen, dos nove grupos pré estabelecidos com 3 linhas e 3 grades para o comando, as 20 linhagens foram classificadas e sete classes foram preenchidas, formando os sete grupos (Figura 15).

As linhagens foram organizadas em diferentes classes. A classe 1 (linha 1 coluna 1), foi constituída por quatro linhagens; para as classes 2 e 3 (linha 1 coluna 2, linha 1 coluna 3, respectivamente), foram alocados um linhagem em cada, destacando-se pela dissimilaridade em relação as demais classes; para a classe 4 (linha 2 coluna 1), não ocorreu a classificação de linhagens, formando um grupo sem linhagens alocadas; para as classes 5 e 6 (linha 2 coluna 2, linha 2 coluna 3, respectivamente), foram contidos quatro linhagens para cada classe; para a classe 7 (linha 3 coluna 1), foram classificadas

quatro linhagens; para a classe 8 (linha 3 coluna 2), duas linhagens; e para a classe 9 (linha 3 coluna 3) não foi verificado a presença de linhagens, como pode ser observado na Figura 15.



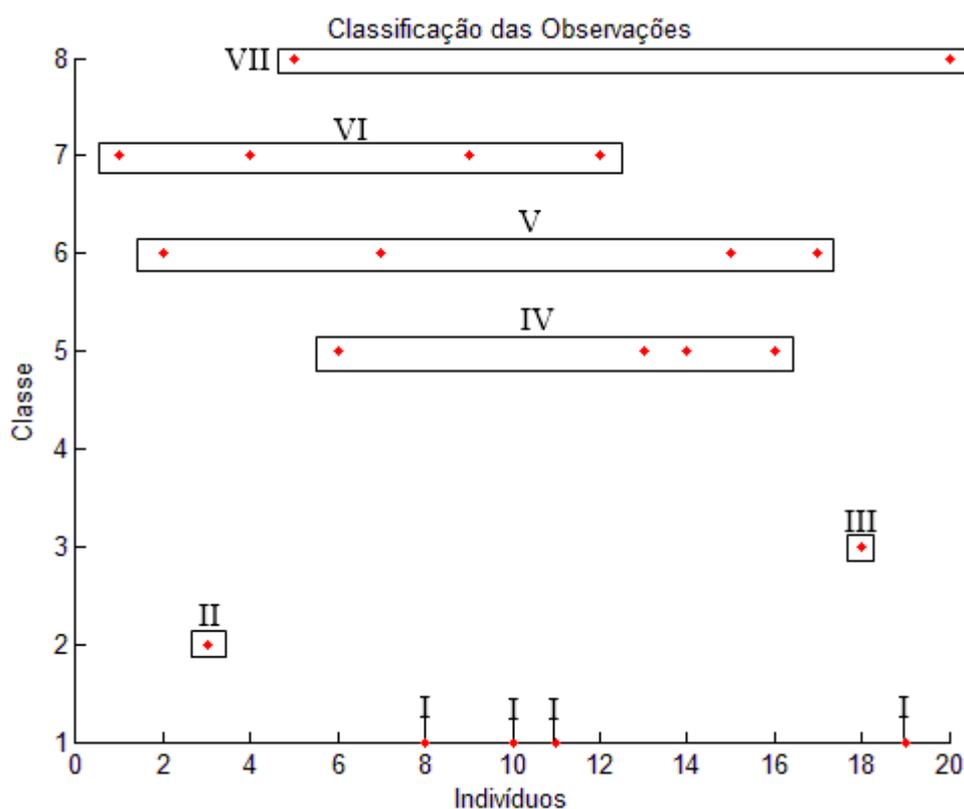
**Figura 15.** Mapa auto-organizável de Kohonen para nove classes de classificação, formadas por meio de rede neural artificial.

As classes onde as linhagens são classificadas no mapa, são consideradas regiões de organização da diversidade genética, onde as classes próximas representadas pelos hexágonos, o preenchimento da área dos hexágonos representa a quantidade de linhagens contidas naquela classe, quanto maior a concentração de linhagens, maior a área preenchida, os grupos que possuem a maior quantidade de linhagens, é representado pelo preenchimento completo do hexágono. As classes são contidas de indivíduos que possuem certa similaridade com a classe vizinha (primeiro vizinho), as classes mais divergentes constituem as regiões dos extremos (não cooperam) e as classes intermediárias ficam constituídas ao centro do mapa (segundo vizinho). Desse modo, linhagens contidas nas classes que não cooperam, apresentam maior dissimilaridade, podendo ser consideradas as mais distantes geneticamente.

Na figura 16 podemos observar a classificação das 20 linhagens de milho pipoca nas classes, e a constituição pelas linhagens dos sete grupos preenchidos, por meio dos mapas de Kohonen. O grupo I foi constituído com as linhagens, 8, 10, 11 e 19 por possuírem os maiores volume por pipoca (VP). O grupo II foi formado pela linhagem 3,

se diferenciando em relação as demais, por possuir a maior capacidade de expansão. O grupo III com a linhagem 18; o grupo IV com as linhagens 6, 13, 14 e 16, que possuem os valores de VP similares; grupo V com as linhagens 2, 7, 15 e 17; grupo VI com a linhagens 1 4, 9 e 12 que possuem semelhanças nos valores de VP; e o grupo VII formado pelas linhagens 5 e 20, sendo semelhantes em seus valores médios de VP.

O caráter de VP foi o responsável pela maioria das diferenciações e organizações entre as linhagens, pelo método de Kohonen. Os valores médios dos caracteres para cada linhagem dos grupos podem ser observados no anexo 3.



**Figura 16.** Mapa auto-organizável de Kohonen, classificação das observações dos agrupamentos das linhagens nas classes formadas por meio de rede neural artificial

As linhagens 3 e 18, foram alocadas em grupos isolados (grupos II e III, respectivamente), indicando serem as mais divergentes em relação as demais de acordo com o método de Kohonen. As linhagens pertencentes aos grupos I e III são as mais indicadas para cruzamentos com as linhagens dos grupos VII (classe 8) e grupo VI (classe 7) devido a distância organizacional entre os grupos no mapa, como pode ser observado na figura 14. Os grupos IV e V apresentam-se como o mais próximo ao grupo III, indicando menor dissimilaridade genética. O grupo II, tem como vizinhos

próximos os grupos I e IV. Os grupos II e III também estão alocados em classes de primeiros vizinhos, apresentando possíveis distâncias genéticas com grau certo grau de similaridade. Os grupos V e VI também se apresentam com alta dissimilaridades entre os caracteres, sendo passíveis para cruzamentos.

Campos (2012), utilizou o método de redes neurais com a metodologia de Kohonen, visando agrupamentos e identificações da análise da divergência em 138 genótipos de goiaba (*Psidium guajava L.*), ao qual foi possível identificar a presença de divergência genética entre as linhagens avaliadas, e que a técnica foi capaz de fornecer resultados para cruzamento entre genótipos de grupos que forneceriam as características de frutos com boas qualidades industriais e consumo *in natura*, sendo importantes para a utilização em programas de melhoramento da cultura.

A técnica de redes neurais tem sido empregada em diversos trabalhos de identificação, seleção de acessos e resolução de problemas. Azevedo (2016), utilizou a aplicação de redes neurais artificiais no processo de classificação de orquídeas do gênero *Cattleya*, buscando a identificação da classe de três espécies de orquídeas, onde constatou-se, que a utilização de redes neurais mostrou-se ideal para resolução da identificação e classificação no trabalho. Carneiro (2015), em um avaliou o potencial das redes neurais artificiais como ferramenta auxiliar no melhoramento da arquitetura de plantas do feijoeiro, gerando resultados confiáveis para predição da característica.

Os grupos encontrados neste trabalho por meio da rede de Kohonen, se diferiram em relação a alguns agrupamentos formados nos métodos de análise multivariada de otimização de Tocher, hierárquico UPGMA, vizinho mais distante e variáveis canônicas.

Para os métodos de Tocher, variáveis canônicas, UPGMA e vizinho mais distante a linhagem 2 foi agrupada separadamente, enquanto que as linhagens 3 e 18 foram alocadas em grandes grupos, formados por várias linhagens, pelos métodos de Tocher e UPGMA.

A classificação pelo método de Kohonen concordou parcialmente com o agrupamento das linhagens 8, 10 e 11 com os métodos de UPGMA, vizinho mais distante e variáveis canônicas, ao ficarem contidos dentro de um mesmo agrupamento. Exceto pela presença da linhagem 19 no lugar da 12. A linhagem 20 por Kohonen foi agrupada com a linhagem 5, enquanto que para UPGMA e variáveis canônicas o agrupamento foi realizado com a linhagem 7, e pelo método de Tocher a linhagem 20 ficou isolada. As linhagens 4 e 9 que formavam um grupo com a 3 (grupo IV) pelas

variáveis canônicas, no método de Kohonen passaram a constituir um grupo com as linhagens 1 e 2 (grupo VI), enquanto a 3 foi isolada. As linhagens 13 e 6, que faziam parte de um grupo com a linhagem 5 por UPGMA, vizinho mais distante e variáveis canônicas e parte de um grupo com 5, 6, 13 e 7 no método de Tocher, já por Kohonen o agrupamento se deu com as linhagens 14 e 16 (grupo IV).

Em geral, o método de mapa auto-organizáveis de Kohonen, apresentou maior semelhança ao agrupamento realizados por meio das variáveis canônicas e maiores diferenças em relação ao método de Tocher.

De Melo (2015) em um estudo de divergência genética em acessos de melão utilizando redes neurais artificiais de kohonen, relatou a diferença entre os grupos encontrados em seu trabalho, que se diferiram profundamente aos encontrados por Torres Filho (2008), com os mesmos dados, aplicando o métodos de análise multivariada, observaram que as principais diferenças ocorreram em relação ao método de Tocher, o que corrobora com os resultados encontrados no presente trabalho.

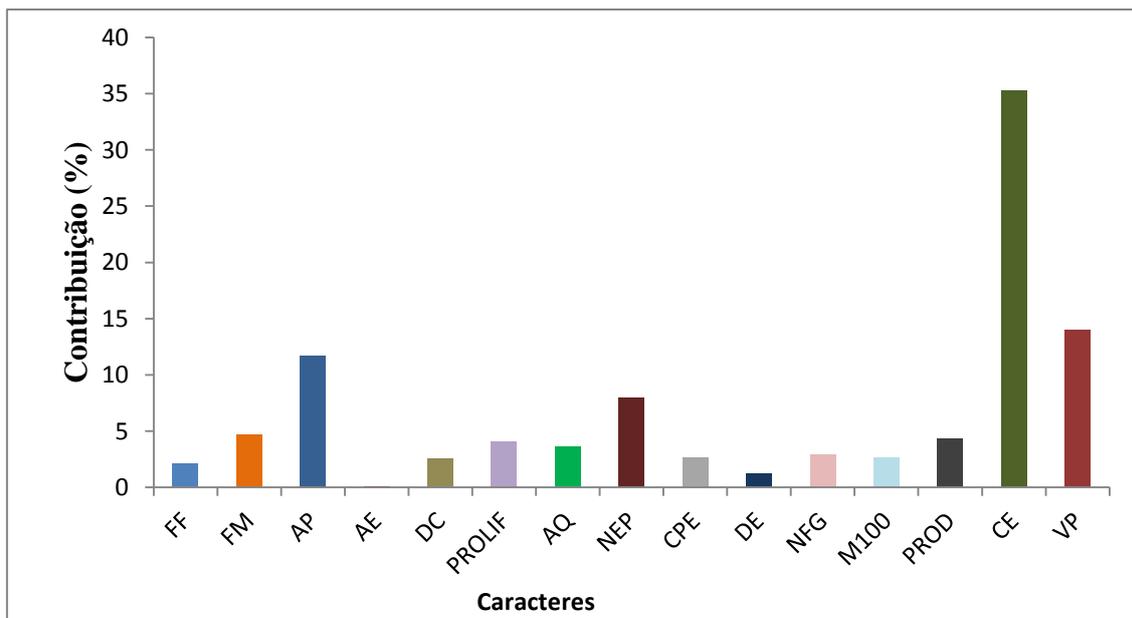
Segundo Galvão et al. (1999), esta diferença pode ser explicada em decorrência das redes neurais artificiais possuem uma estrutura não-linear e não paramétrica, ao qual conseguem captar características mais profundas e complexas dos dados, sejam quantitativos ou qualitativos, o que nem sempre é possível com a utilização das técnicas estatísticas tradicionais.

#### **5.2.4 Importância relativa dos caracteres**

A análise para estimar a contribuição relativa de quinze caracteres morfoagronômicos para a expressão da divergência genética, foi estimada segundo o método de Singh (1981) baseado na distância Generalizada de Mahalanobis, conforme apresentado na figura 17.

Os caracteres CE (35,28%), VP (14,04%) e AP (11,71%), apresentaram a maior contribuição para a divergência total entre as 20 linhagens  $S_3$  de milho pipoca avaliadas. O NEAP (7,95%) também apresentou uma contribuição relativamente importante sobre a divergência.

A explicação em relação a capacidade de expansão (CE) apresentar a maior contribuição relativa, pode ser dada em função que a análise reuniu linhagens distintas, onde as variações foram expressamente contrastantes, em relação a CE das médias que variaram de 11,66 mL/g (linhagem 2) a 23,50 mL/g (linhagem 3).



**Figura 17.** Contribuição relativa de quinze caracteres morfoagronômicos de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP), avaliados para divergência ( $D^2$ ) entre 20 linhagens  $S_3$  de milho pipoca, na safra de 2016/2017, em Dourados/MS.

A capacidade de expansão é um dos caracteres mais importante para o milho pipoca, pois está relacionada diretamente com a qualidade do produto final, onde quanto maior a CE, maior a palatabilidade da pipoca, refletindo em uma das principais características para a aceitação no mercado consumido.

O caráter volume de pipoca (VP), apresentou grande importância em relação aos agrupamentos das linhagens, uma vez que os grupos foram formados em relação a semelhança dos valores de VP apresentados, além de conter os dados referentes aos dois principais caracteres da cultura, a produtividade e a capacidade de expansão, apresentados de forma conjunta.

Em relação a contribuição relativa do número de espigas por parcela, as médias entre as linhagens apresentaram grandes variações, apresentando valores de 2,33 (linhagens 16 e 20) a até 10 espigas por parcela (linhagem 10).

De acordo com o método de Singh, os caracteres AE (0,10%) e DE (1,23%), apresentaram as menores contribuições para a divergência genética, sendo sugeridos como variáveis passíveis para descarte no presente estudo, pois seriam pouco informativas, devido às estimativas de contribuição relativa de pequenas magnitudes.

A importância relativa dos caracteres auxilia no descarte de variáveis, possibilitando uma melhor escolha de variáveis a serem consideradas na avaliação de divergência genética entre populações ou linhagens (Cruz et al., 2014).

Melo et al. (2017), no estudo de estratégias de seleção entre progênes meios-irmãos de milho pipoca, observou que a capacidade de expansão, foi o caráter com a maior contribuição relativa para a dissimilaridade genética. Arnhold et al. (2010), no estudo de seleção de linhagens de milho pipoca com base em desempenho e divergência genética, a capacidade de expansão também apresentou o maior índice de importância relativa sobre a divergência, corroborando com os resultados obtidos no presente trabalho.

### **5.2.5 Considerações experimento linhagens S<sub>3</sub> (safra II)**

Por meio das metodologias empregadas, foi possível evidenciar a existência de variabilidade genética entre as linhagens S<sub>3</sub> avaliadas para o período da safra II. Os métodos utilizados para os agrupamentos apresentaram semelhanças entre si, para representar a divergência genética entre as linhagens avaliadas, sendo parcialmente concordantes nos agrupamentos. Em relação a comparação entre os métodos de análise multivariada e as redes neurais dos mapas de Kohonen, a maior semelhança se deu entre os resultados apresentados por meio das variáveis canônicas, e maiores diferenças em relação ao método de Tocher.

Para os métodos de Tocher, variáveis canônicas, UPGMA e vizinho mais distante a L2 foi agrupada em grupo isolado, enquanto que as L3 e 18 foram alocadas em grandes grupos, pelos métodos de Tocher e UPGMA. A classificação pelo método de Kohonen concordou parcialmente com o agrupamento das L8, 10 e 11 com os métodos de UPGMA, vizinho mais distante e variáveis canônicas, ao ficarem contidos dentro de um mesmo agrupamento. Exceto pela presença da L19 no lugar da 12. A L20 por Kohonen foi agrupada com a L5, enquanto que para UPGMA, vizinho mais distante e variáveis canônicas o agrupamento foi realizado com a L7, e pelo método de Tocher a L20 ficou isolada.

As linhagens mais divergentes e potenciais para serem indicadas como genitoras em futuros cruzamentos são: linhagem 2, que apresenta a menor capacidade de expansão (CE), e volume de pipoca (VP); linhagem 3, tendo como característica a maior capacidade de expansão (CE); linhagem 10, com o maior produtividade (PROD), volume por pipoca (VP), massa de cem grãos (M100) e número de espiga por parcela

(NEAP); e a linhagem 8 que apresenta maior diâmetro de espiga (DE) e menores dias para o florescimento feminino (FF) e florescimento masculino (FM).

O restabelecimento da heterozigose nos locus gênicos é de fundamental importância, nesse sentido a escolha de genitores que apresentem caracteres contrastantes faz-se necessário, sendo assim, a escolha da linhagem 2 mostra-se uma estratégia eficiente para se obter maior êxito nos cruzamentos, ou seja, o vigor híbrido e manifestação das características de interesse nas progênes, o que justifica a indicação da linhagem 2 para os cruzamentos

### 5.3 Considerações finais

É possível constatar diferenças entre os experimentos realizados, principalmente em relação a dissimilaridade entre as linhagens. De acordo com a distancia generalizada de Mahalanobis a maior dissimilaridade para as linhagens  $S_2$  foi dada entre as linhagens 8 x 9 e 12 x 14, enquanto para as linhagens  $S_3$  a maior divergência ocorreram entre as linhagens 2 x 11 e 3 x 2.

Para as linhagens  $S_2$  (safra I), os genitores selecionados como os mais divergentes e promissores, foram as linhagens 9 apresentando a menor AP, menor PROD e VP; a linhagem 12, com o maior NFG e NEAP; a linhagem 15, com o maior PROD, VP, CE e PROLIF; e a linhagem 6 apresentando maior DC e menores dias para FF e FM.

Para as linhagens  $S_3$  (safra II) as linhagens selecionadas como genitores foram, as linhagens 2 que apresentou a menor CE e VP; linhagem 3 apresentando a maior CE; linhagem 10 com o maior PROD, VP, M100 e NEAP; e a linhagem 8 com maior DE e menores dias para FF e FM.

As características variaram de uma safra para outra, em relação a significância e a acurácia, indicando a possível influência do ambiente e da constituição genética em relação a variação das linhagens.

O avanço no ciclo de autofecundação resulta em alterações na sua constituição gênica das linhagens, resultando em diferenças em sua diversidade genética e expressões de suas características. A cada ciclo de autofecundação o número de *locus* em heterozigose é reduzido, onde na segunda geração de autofecundação ( $S_2$ ) obtém cerca 75% de homozigose, na terceira ( $S_3$ ) 87,5% (Nass et al., 2001). O avanço na geração aumenta em mais de 12% os níveis de homozigose, quando se eleva o índice, a

constituição genética das linhagens sofrem alterações e a divergência contida são diferenciadas, bem como as expressões de suas características.

Os caracteres mais importantes para discriminação das linhagens  $S_2$ , para a safra I, foram o PROD (20,80%), VP (10,90%) e M100 (10,10%) e a CE (8,46%). Já para as linhagens  $S_3$  na safra II, os caracteres CE (35,28%), VP (14,04%) e AP (11,71%) e o NEAP (7,95%), apresentaram a maior contribuição relativa para a divergência total entre as 20 linhagens  $S_3$  de milho pipoca avaliadas.

Contudo, em relação a safra II pode-se ressaltar o favorecimento da capacidade de expansão, onde mesmo com o avanço do ciclo de autofecundação, ocorreu pouca diferença em relação de uma safra para outra, ao qual para a safra II foi a principal característica de contribuição para a divergência genética nas 20 linhagens  $S_3$ , podendo indicar que entre as mesmas, tem-se linhagens que apresentam potencial para a característica, sendo de interesse para o melhoramento da cultura do milho pipoca.

Estresses abióticos, como temperatura e precipitação são fatores ambientais que interferem fortemente na produtividade e outras características do milho pipoca, que é uma cultura bastante exigente em água. O consumo de água durante todo o ciclo da cultura, na região de cerrado, varia de 550 mm a 650 mm e desde a emergência até a floração, a temperatura ideal para o desenvolvimento do milho é de 24 a 30°C (Fancelli e Dourado Neto, 2000; Vaz de Melo et al., 2012).

Para as linhagens  $S_2$  (safra I) no o período da semeadura (mês de setembro), ocorreu uma precipitação regular com média de 81,60 mm, fornecendo uma boa disponibilidade hídrica favorecendo a germinação e desenvolvimento para a cultura (Figura 3). Para as linhagens  $S_3$  (safra II) no período da data de semeadura (mês de novembro), ocorreu uma baixa precipitação com média de 51 mm (Figura 4) durante todo o mês, fase em que ocorreu a germinação e desenvolvimento das plântulas. A baixa disponibilidade hídrica pode ocasionar um estresse hídrico nas plântulas, afetando seu desenvolvimento, em alguns casos são ocasionados danos que influenciam sob as características em todo seu ciclo vegetativo. As plantas tendem a aumentar o gasto de energia para recuperação celular e passa a perder seu vigor de crescimento, podendo levar a plantas mais susceptíveis a fatores ambientais, como exemplo, ao vento, levando ao acamamento e o quebramento dos colmos, podendo afetar sobre a produtividade de grãos final.

Para a safra I, no mês de novembro, o índice médio de precipitação com 90 dias, foram de 118,2 mm (Figura 3). Para a safra II, no mês de fevereiro, onde as plantas

estavam com cerca de 90 dias, já com as espigas autofecundadas, ocorreu uma precipitação média com cerca de 228,6 mm (Figura 4). Para a o mês de fevereiro na safra II, o alto índice pluviométrico e o vento, levou a perdas dos papéis Krafts que realizam a proteção das espigas, acamamento das plantas e apodrecimento das espigas cobertas devida a alta umidade concentrada, afetando a qualidade e/ou ocasionando perdas dos materiais. Fatores ambientais que possivelmente podem ter contribuído sobre a diminuição da produtividade das plantas S<sub>3</sub> na safra II. Fatores, que não ocorreram no experimento para as linhagens S<sub>2</sub> conduzido na safra I.

Pode-se observar que ao decorrer de uma safra para outra, houve a redução dos valores médios para os caracteres, principalmente em relação a produtividade, que para as linhagens S<sub>2</sub> na safra I, foi o caráter que mais contribuiu para a divergência genética e para as linhagens S<sub>3</sub> safra II apresentou baixa contribuição, onde as linhagens passaram a apresentar queda de 437,10 kg ha<sup>-1</sup>, significando uma redução de cerca de 70% da produtividade, indicando que as linhagens avaliadas ao decorrer da elevação da homozigose nos *locus* gênicos, apresentaram a diminuição da expressão deste caráter, e/ou sofreram interferência devido as condições ambientais.

## 6. CONCLUSÕES

Por meio das metodologias empregadas, foi possível evidenciar a existência de variabilidade genética entre as 20 linhagens  $S_2$  e as 20 linhagens  $S_3$  de milho pipoca em estudo.

A composição genética distinta das linhagens em  $S_2$  e  $S_3$  e as diferenças ambientais dos anos agrícolas avaliados, influenciaram nas diferenças obtidas em relação à divergência genética nos dois experimentos.

Os métodos de agrupamentos apresentaram semelhanças para representar a divergência genética entre as linhagens avaliadas.

Entre os métodos multivariados, a dispersão gráfica de variáveis canônicas e o UPGMA, apresentaram a maior semelhança entre si para os agrupamentos.

Entre os agrupamentos dos métodos multivariados e o de redes neurais de Kohonen, o maior relacionamento ocorreu com os resultados apresentados pela dispersão gráfica de variáveis canônicas e as maiores diferenças em relação ao método de Tocher.

As linhagens  $S_2$  mais divergentes e potenciais identificadas foram a 9, 12, 15 e 6; e as linhagens  $S_3$  mais divergentes foram a 2, 3, 10 e 8, podendo ser indicadas como genitoras em futuros cruzamentos dialélicos.

Os caracteres que apresentaram a maior contribuição relativa para a divergência e diferenciação das linhagens, para as linhagens  $S_2$  foram a PROD, VP, M100 e CE; e para as linhagens  $S_3$  os caracteres CE, VP, AP e NEAP.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ALEXANDER, D. E.; CREECH, R. G. Breeding special industrial and nutritional types. In: SPRAGUE, G. F.; FUCCILLO, D. A. (Ed.). Corn and corn improvement. **Madison: American Society of Agronomy**, Madison, p. 363- 386, 1977.

AMARAL JÚNIOR, A. T.; SANTOS, A.; GERHARDT, I. F. S.; KUROSAWA, R. N. F.; MOREIRA, N. F.; PEREIRA, M. G.; GRABINA, G. A.; SILVA, F. H. L. Proposal of a super trait for the optimum selection of popcorn progênies based on path analysis. **Genetic and Molecular Research**, Ribeirão Preto, v.15, n.4, 2016.

AMORIM, E.P.; SOUZA, J.C. Híbridos de milho inter e intrapopulacionais obtidos a partir de populações S<sub>0</sub> de híbridos simples comerciais. **Bragantia**, Campinas, v.64, n.3, p.561-567, 2005.

AMARAL JUNIOR, A. T.; JÚNIOR, S. F.; RANGEL, R.; PENA, G.; RIBEIRO, R.; MORAIS, R.; SCHUELTER, A. A.T. Improvement of a popcorn population using selection indexes from a fourth cycle of recurrent selection program carried out in two different environments. **Genetics and Molecular Research**, Ribeirão Preto, v. 9, n. 1, p. 340-370, 2010.

ARNHOLD, E.; SILVA, R. G.; VIANA, J. M. S. Seleção de linhagens S<sub>5</sub> de milho pipoca com base em desempenho e divergência genética. **Acta Scientiarum Agronomy**, Maringá, v. 32, n. 2, p. 279-283, 2010.

AZEVEDO, A. M.; JÚNIOR, V. C. A.; JÚNIOR, A. S. S.; SANTOS, A. A.; CRUZ, C. D.; PEREIRA, S. L.; OLIVEIRA, A. J. M. Eficiência da estimação da área foliar de couve por meio de redes neurais artificiais. **Horticultura Brasileira**, Vitoria da Conquista, v. 35, n. 1, 2017

AZEVEDO, L.P. **Aplicação de redes neurais artificiais no processo de classificação de orquídeas do gênero *cattleya*** Monografia (Bacharelado em Sistemas de Informação) – Instituto Federal de Minas Gerais, Campus Sabará, p.48, 2016.

BARBOSA, C. D., VIANA, A. P., SILVA, S., QUINTAL, R., & PEREIRA, M. G. Artificial neural network analysis of genetic diversity in *Carica papaya* L. **Crop Breeding and Applied Biotechnology**, Viçosa, 11(3) p. 224–231, 2011.

BARILI, L. D.; VALE, N.M.; ARRUDA, B.; TOALDO, D.; ROCHA, F.; COIMBRA, J. L. M.; BERTOLDO, J. G.; Y GUIDOLIN, A. F. Escolha de genitores contrastantes para compor o bloco de cruzamentos de genótipos de feijão. **Revista Brasileira de Agrociência**, Pelotas, 17(3-4), p.303-310, 2011.

BARROSO, L. M. A., NASCIMENTO, M., NASCIMENTO, A. C. C., SILVA, F. F., & FERREIRA, R. DE P. Uso do método de eberhart e russell como informação a priori para aplicação de redes neurais artificiais e análise discriminante visando a classificação de genótipos de alfafa quanto à adaptabilidade e estabilidade. **Revista Brasileira de Biometria**, Jaboticabal, 31(2), 176–188, 2013.

BARROSO, L.P., ARTES, R. **Análise multivariada**. Lavras: UFLA, p. 151, 2003.

BATCHELOR, W. D., YANG, X. B., & TSCHANZ, A. T. Development of a neural network for soybean rust epidemics. **Transactions of the ASAE**, 40(1), 247– 252, 1997.

BENITE, M. **Aplicação de modelos de redes neurais na elaboração e análise de cenários macroeconômicos**. Dissertação (Mestrado em Engenharia), São Paulo, Escola de Engenharia de São Carlos – Universidade de São Paulo, p.123, 2003.

BENTO, CS; SUDRE, CP; RODRIGUES, R.; RIVA, EM; PEREIRA, MG. Descritores qualitativos e multicategóricos na estimativa da variabilidade fenotípica entre acessos de pimentas. **Scientia Agraria**, Curitiba, v. 8, n. 2, p. 149-156, 2007.

BERTAN I.; CARVALHO F.I.F.; OLIVEIRA A.C.; VIEIRA I.; HARTWIG E.A.; SILVA J.A.G.; SHIMIDT, D.A.M; VALÉRIO I.P.; BUSATO C.C.; RIBEIRO G. Comparação de métodos de agrupamento na representação da distância morfológica entre genótipos de trigo. **Revista Brasileira de Agrociência**, Pelotas, 12: 279-286. 2006.

BHERING, L.L.; CRUZ, C.D.; PEIXOTO, L. DE A.; ROSADO, A.M.; LAVIOLA, B.G.; NASCIMENTO, M. Application of neural networks to predict volume in eucalyptus. **Crop Breeding and Applied Biotechnology**, Viçosa, v.15, p.125-131, 2015.

BISHOP, C. M. **Neural networks for pattern recognition**. Oxford, New York: Oxford University Press, p.482, 1995.

BORÉM, A.; MIRANDA, G. V.; FRITSCHÉ-NETO, R. **Melhoramento de Plantas**. Viçosa, MG. Editora UFV, 7. ed., p. 543, 2017.

BRAGA, A. DE P.; CARVALHO, A. C. P; LUDEMIR, T. B. **Redes Neurais Artificiais - Teoria e Aplicações**. Rio de Janeiro: LTC, 2<sup>a</sup> ed., p. 251, 2011.

BRASILEIRO, B.P.; MARINHO, C.D.; COSTA, P.M. DE A.; CRUZ, C.D.; PETERNELLI, L.A.; BARBOSA, M.H.P. Selection in sugarcane families with artificial neural networks. **Crop Breeding and Applied Biotechnology**, Viçosa, v.15, p.72-78, 2015.

BRUNSON, A.M. Popcorn breeding. Washington: **Yearbook Agricultural**, v.1, p.395-404, 1937.

BRUNSON, A. M.; SMITH, G. M. Hybrid popcorn. **Journal of the American Society of Agronomy**, Madison, v. 37, p. 176-183, 1945.

BUSSAB, W. DE O.; MIAZAKI, E.S.; ANDRADE, D.F. DE. Introdução à análise de agrupamentos. In: Simpósio Nacional De Probabilidade E Estatística, 9., 1990, São Paulo. **Anais...** São Paulo: ABE, 1990. P.106, 1990.

CABRAL, P. D. S.; AMARAL JÚNIOR, A. T.; FREITAS, I. L. J.; RIBEIRO, R. M.; SILVA, T. R. C. Relação causa e efeito de caracteres quantitativos sobre a capacidade

de expansão do grão em milho pipoca. **Ciência Agrônômica**, Fortaleza, v. 47, n. 1, p. 108-117, 2016.

CABRAL, P. D. S.; SOARES, T. C. B.; LIMA, A. B. P.; SOARES, Y. J. B.; SILVA, J. A. Análise de trilha do rendimento de grãos de feijoeiro (*Phaseolus vulgaris* L.) e seus componentes. **Ciência Agrônômica**, Fortaleza, v.42, n.1, p.132-138, 2011.

CAMPOS, B. M. **Redes neurais artificiais e Ward-MLM aplicados à análise da divergência em goiaba (*Psidium guajava* L.)**. Mestrado (Mestrado em Genética e Melhoramento) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, UENF, Campos dos Goytacazes, Rio de Janeiro, p. 99, 2012.

CAMUSSI, A.; OTTAVIANO, E.; CALINSKI, T.; KACZMAREK, Z. Genetic distances based on quantitative traits. **Genetics**, 11: 945-62, 1985.

CARGNELUTTI FILHO, A.; STORCK, L. Medidas do grau de precisão experimental em ensaios de competição de cultivares de milho. **Revista Pesquisa Agropecuária Brasileira**, Brasília, 44: 111-117, 2009.

CARVALHO, C.; KIST, B. B.; SANTOS, C. E.; MULLER, I.; BELING, R. R. **Anuário brasileiro de milho**. Santa Cruz do Sul: Editora Gazeta Santa Cruz LTDA., p. 112, 2015.

CARVALHO, F.I.F; SILVA, S.A.; KUREK, A.J. **Estimativas e implicações da herdabilidade como estratégia de seleção**. Pelotas: Editora e gráfica da UFPel, p.99, 2001.

CEASA - CENTRAL DE ABASTECIMENTO. **Cotação de Preços**. Disponível em: <<http://www.ceasasc.com.br/cotacao-de-precos/?key=&p=2>> Acesso em: 12 set. 2017.

CIB - CONSELHO DE INFORMAÇÕES SOBRE BIOTECNOLOGIA. **Guia do milho**. Tecnologia do campo a mesa. São Paulo. Disponível em: <[http://www.cib.org.br/pdf/guia\\_do\\_milho\\_CIB.pdf](http://www.cib.org.br/pdf/guia_do_milho_CIB.pdf)> Acesso em: 09 out. 2017.

CLOVIS, L. R.; SCAPIM, C. A.; PINTO, R. J. Análise de estabilidade de híbridos de milho utilizando o mapa Auto-Organizável de Kohonen. In: 8º Congresso Brasileiro de Melhoramento de Plantas, 2015. **Anais...** cd-rom, Goiânia, GO., 2015.

COIMBRA, R. R.; MIRANDA, G. C.; CRUZ, C. D.; MELO, A. V.; ECKERT, F. R. Caracterização e divergência genética de populações de milho resgatadas do Sudeste de Minas Gerais. **Revista Ciência Agrônômica**, Ceará v. 41, n. 1, p. 159-166, 2010.

COLOMBO, G. A.; VAZ-DE-MELO, A.; TAUBINGER, M.; FARIA, E. A.; TAVARES, R. C. Fenotipagem de genótipos de milho em condições de estresse por alumínio e sua correlação com produtividade de grãos. **Revista Agrarian**, Dourados, V. 7, p. 60-71, 2014.

CORREA, A. M.; GONÇALVES, M. C. Divergência genética em genótipos de feijão comum cultivados em Mato Grosso do Sul. **Revista Ceres**, Viçosa, v.59, n.2, p.206-212, mar./abr. 2012.

CRUZ, C.D., CARNEIRO, P.C.S., REGAZZI, A.J. **Modelos biométricos aplicado ao melhoramento genético**. Viçosa: UFV, 3. ed., v. 2, p. 668, 2014.

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

CRUZ, C. D.; Genes Software – extended and integrated with the R, Matlab and Selegen **Acta Scientiarum**, Maringá, v.38, n.4, p.547-552, 2016.

DAROS, M.; AMARAL JÚNIOR, A.T.; PEREIRA, M.G.; SANTOS, F.S.; SCAPIM, C.A.; FREITAS JÚNIOR, S.P.; DAHER, R.F.; ÁVILA, M.R. Correlações entre caracteres agrônômicos em dois ciclos de seleção recorrente em milho pipoca. **Ciência Rural**, Santa Maria, v.34, p.1389-1394, 2004.

DE MELO, S. B. **Divergência Genética Em Acessos De Melão Utilizando Redes Neurais Artificiais**. Tese (Doutorado em Fitotecnia) Universidade Federal Rural do Semi-Árido, Mossoró, p. 72, 2015.

DE PAULA, T. O. M. **Adaptabilidade e estabilidade, divergência genética e otimização experimental em milho pipoca**. (Dissertação – Mestrado em Genética e Melhoramento de Plantas) Universidade Estadual do Norte Fluminense Darcy Ribeiro, p. 218, 2009.

DUTRA FILHO, J. A.; MELO, L. J. O. T.; 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 Agronômica**, Ceará, v. 42, p.185-192, 2011.

EMBRAPA - ESTAÇÃO CLIMATOLÓGICA EMBRAPA AGROPECUÁRIA OESTE (CPAO) DE DOURADOS-MS. **Guia e Clima**, 2017. Disponível em: <<http://www.cpaio.embrapa.br/clima/?lc=site/estatisticas/estatisticas-ver-dados>>. Acesso em: 20 out. 2017.

FALCONER, D.S.; MACKAY, T.F. **Introduction to quantitative genetics**. 4 ed. Londres: Editora Longman Group, p. 464, 1996.

FANCELLI, A.L.; DOURADO-NETO, D. **Produção de milho**. Guaíba: Agropecuária. p. 360, 2000.

FANTIN, G.M., SAWAZAKI, E., BARROS, B.C. Avaliação de genótipos de milho pipoca quanto à resistência a doenças e à qualidade da pipoca. **Summa Phytopathologica**, Botucatu, v.17, p. 91-104, 1991.

FARIA, V R.; VIANA, J M. S.; SOBREIRA, F. M.; SILVA, A. C. Seleção recorrente recíproca na obtenção de híbridos interpopulacionais de milho pipoca. **Pesquisa agropecuária brasileira**, Brasília, v.43 p.12, 2008.

FIETZ, R. C.; FISCH, G. F.; **O clima da região de Dourados, MS**. Documentos: Embrapa Agropecuária Oeste, ISSN 1679-043X, ed.2, p.32, 2008.

GALVÃO, C. O., VALENÇA, M. J. S., VIEIRA, V. P. P. B., DINIZ, L. S., LACERDA, E. G. M., CARVALHO, A. C. P. L. F, LUDERMIR, T. B. **Sistemas inteligentes: Aplicações a recursos hídricos e ciências ambientais**. Porto Alegre: UFRGS/ABRH, p. 246, 1999.

GERMANI, R.; PACHECO, C.A.P.; CARVALHO, C.W.P. Características físicas e químicas dos principais cultivares de milho pipoca planados no Brasil., **Arquivos de Biologia e Tecnologia**, Londrina, V.40(1),p.19-27, 1997.

GONÇALVES, D. L.; AMBROZIO, V. C.; BARELLI, M. A. A. Divergência genética de acessos tradicionais de feijoeiros através de características da semente. **Bioscience Journal**, Uberlândia, v. 30, n. 6, p. 1671- 1681, 2014.

GONÇALVES, G. M. B. **Caracterização e divergência genética de variedades crioulas de milho pipoca conservadas por agricultores do oeste de Santa Catarina**. Dissertação (mestrado em Recursos Genéticos Vegetais) - Universidade Federal de Santa Catarina, Centro de Ciências Agrárias, Florianópolis, p.82, 2016.

GONÇALVES, W. G. G.; RIBEIRO, H. M. C.; SÁ, J. A. S.; MORALES, G. P.; FILHO, H. R. F.; ALMEIDA, A. C. Classification of forest types using artificial neural networks and remote sensing data. **Revista Ambiente e Água**, Tabaté, v.11, n.3, 2016.

HALLAUER, A; CARENA, M.J. Maize Breeding. In: Cereals. ed.: Carena, M.J. **Springer**, Germany, p. 425, 2009.

HALLAUER, A. R., CARENA, J. M., FILHO, J. B. M. **Quantitative genetics in maize breeding**. Iowa: The Iowa State University Press, 663p 2010.

HAYKIN, S. **Neural Networks and Learning Machines**. 3rd ed. Pearson - Prentice Hall, Hamilton, p. 938, 2008.

HOSENEY, R.C., Zeleznak, K., Abdelrahman, A. Mechanism of Popcorn popping. **Jornal of Cereal Science**, London, V. 1, p. 43-52, 1983.

ICMC - Instituto de Ciências Matemáticas e de Computação (USP). **Redes Neurais Artificiais**. <<http://conteudo.icmc.usp.br/pessoas/andre/research/neural/>> Acesso em: 20 out. 2017

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Levantamento de safras**. Disponível em: <<http://www.ibge.gov.br/home/default.php>>. Acesso em: 12 out. 2017.

JANG, J.S. C.T. SUN, AND E. MIZUTANI. **Neuro-Fuzzy and Soft Computing – A computational Approach to Learning and Machine Intelligence**. PHI Learning Private Limited, New Delhi, p. 614, 2012.

JI, B., Y. SUN, S. YANG, AND J. WAN. Artificial neural networks for rice yield prediction in mountainous regions. **J. Agricultural Science**, Cambridge, v. 145(03), p.249, 2007.

JÚNIOR, E. G. S. **Seleção de genótipos de algodoeiro utilizando redes neurais artificiais**. Trabalho de Conclusão de Curso (Graduação em Agronomia) Universidade Federal de Uberlândia, UFU, Uberlândia, p. 20, 2017.

KAUL, M., R.L. HILL, AND C. WALTHALL. Artificial neural networks for corn and soybean yield prediction. **Agricultural Systems**, Essex, V. 85(1), p.1–18, 2005.

KAVZOGLU, T, MATHER, P. M. The use of backpropagation artificial neural networks in land cover classification. **International Journal of Remote Sensing**, Germany, v.24 (13): 4907-4938, 2003.

KOHONEN, T. Self-organized formation of topologically correct feature maps. **Biological Cybernetics**, Germany, 43: 59-69, 1983.

KOVACS, Z. L. **Redes neurais artificiais: fundamentos e aplicações**. São Paulo: Collegium Cognitio, p. 174, 1997.

LEGENDRE, P.; LEGENDRE, L. **Numerical Ecology**, 2nd English ed. Elsevier, 853 pp. 1998.

LEONELLO, L. A. F.; CAZETTA, D. A.; FORNASIERI FILHO, D. Características agronômicas e qualidade comercial de cultivares de milho pipoca em alta população. **Acta Scientiarum Agronomy**. Maringá, v. 31, n. 2, p. 215-220, 2009.

LOPES, L. A.; PELUZIO, J. M.; AFFÉRI, F. S.; CARVALHO, E. V. Variabilidade genética entre cultivares de soja, quanto ao rendimento de óleo, no estado do Tocantins. **Comunicata Scientiae**, Teresina, v.5 (3), p. 279-285, 2014.

MAPA - MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **Registros nacionais de cultivares**. Disponível em : <[http://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares\\_registradas.php](http://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares_registradas.php)>. Acesso em: 10 jan. 2018.

MARCHÃO, R. L.; BRASIL, E. M.; DUARTE, J. B.; GUIMARÃES, C. M.; GOMES, J.A. Densidade de plantas e características agronômicas de híbridos de milho sobespaçamento reduzido entre linhas. **Pesquisa agropecuária brasileira**, Brasília, v. 35, n.2, p. 93-101, 2005.

MELO, A. V.; COLOMBO, G. A.; VALE, J. C. V.; SANTANA, W. D.; FERNANDES, M. S. Estratégias de seleção entre progênies meios-irmãos de milho pipoca no cerrado Tocantinense **Brazilian Journal of Applied Technology for Agricultural Science**, Guarapuava-PR, v.10, n.1, p.41-50, 2017.

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

MIRANDA, G.V.; COIMBRA, R.R.; GODOY, C.L.; SOUZA, L.V.; GUIMARÃES, L.J.M.; MELO, A.V. Potencial de melhoramento e divergência genética de cultivares de milho pipoca. **Pesquisa Agropecuária Brasileira**, Brasília, v.38, n.6, p.681-688, 2003.

MIRANDA FILHO, J.B.; VIÉGAS, G.P. Milho híbrido. IN: PATERNIANI, E.; VIÉGAS, G.P. (Eds). **Melhoramento e produção do milho**. 2. ed. Campinas: Fundação Cargill, v. 1, p.277-326, 1987.

MISSIO, R. F.; MORAES, M. L. T.; DIAS, L. A. S. Efeito do desbaste seletivo sobre a divergência genética em progênies de *Pinus caribaea* Morelet var. *bahamensis*. **Scientia Forestalis**, São Paulo, v. 73,. p. 27-36. 2007.

MOHAMMADI, S.A., PRASANA, B.M. Analysis of genetic diversity in crop plants – salient statistical tools and considerations. **Crop Science**, Fitchburg, 43: 1235-1248, 2003.

MOTERLE LM, BRACCINI AL, SCAPIM, CA, BARTH PINTO RJ, GONÇALVES LSA, RODRIGUES R AND AMARAL JÚNIOR AT. Combining ability of popcorn lines for seed quality and agronomic traits. **Euphytica**, Wageningen, v.185, p. 337-347, 2012.

MOURA, W.M.; CASALI, V.W.D.; CRUZ, C.D.; Lima, P.C. Divergência genética em linhagens de pimentão em relação à eficiência nutricional de fósforo. **Pesquisa Agropecuária Brasileira**, Brasília, v.34, n.2, p.217-224. 1999.

NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. DE; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento**. Rondonópolis: Fundação MT, 1183 p. 2001.

OLIBONI, R.; FARIA, M. V.; NEUMANN, M.; BATTISTELLI, G. M.; TEGONI, R. G.; RESENDE, J. T. V. Genetic divergence among maize hybrids and correlations with heterosis and combining ability. **Acta Scientiarum**, Maringá, v. 34, n. 1, p. 37-44, 2012.

OLIVEIRA, A. C. DE L.; PASQUAL, M.; PIO, L. A. S.; LACERDA, W. S.; SILVA, O. S. Use of mathematical modeling (artificial neural networks) in classification of banana autotetraploid (*musa acuminata* colla). **Bioscience Journal**, Uberlândia, v. 29, n. 3, p. 617-622, 2013.

OLIVEIRA, G. H. F. **Capacidade combinatória e correlação em populações de milho pipoca**. Jaboticabal, (Doutorado Genética e Melhoramento de Plantas) - Universidade Estadual Paulista, p.50, 2016.

OLIVEIRA, T. C. **Divergência genética e correlação entre caracteres de genótipos de sorgo sacarino na região de Cáceres-MT**. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade do Estado de Mato Grosso. p.89, 2015.

PAES, M. C. D. **Aspectos físicos, químicos e tecnológicos do grão de milho**. Circular técnica 75. Ministério da Agricultura, Pecuária e Abastecimento, Sete Lagoas, MG; pg. 1-7, dez. 2006.

PARAGINSKI, R. T. **Efeitos da secagem e da incidência de defeitos na qualidade de grãos de milho, e do processamento na qualidade de grãos de milho pipoca**. Tese (Tese Doutorado) Programa de Pós-Graduação em Ciência e Tecnologia de Alimentos, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, p. 161, 2014.

PATERNIANI, E.; NASS, L.L.; SANTOS, M.X. O valor dos recursos genéticos de milho para o Brasil: uma abordagem história da utilização do germoplasma. In: UDRY, C.W.; DUARTE, W. (Org.) **Uma história brasileira do milho: o valor dos recursos genéticos**. Brasília: Paralelo 15, p.11-41, 2000.

PATERNIANI, M. E. A. G. Z.; CRISTIANI, S. B.; GUIMARÃES, P. S.; DONÁ, S.; GALLO, P. B.; DUARTE, A. P. Potencial produtivo e heterose de híbridos de populações F<sub>2</sub> de milho no estado de São Paulo. **Cadernos de Ciência & Tecnologia**, Brasília, v. 27, n. 1/3, p. 29-46, 2010.

PENA, G.F.; AMARAL JÚNIOR, A.T. DO; GONÇALVES, L.S.A.; CANDIDO, L.S.; VITTORAZZI, C.; RIBEIRO, R.M.; FREITAS JÚNIOR, S.P. Stability and adaptability of popcorn genotypes in the State of Rio de Janeiro, Brazil. **Genetics and Molecular Research**, Ribeirão Preto, v.11, p.3042-3050, 2012.

PESSOA, A. M. S.; BARROSO, P.A., RÊGO, E. R. DO; MEDEIROS, G. A.; BRUNO, R. L.A; RÊGO, M.M. Genetic divergence of physiological-quality traits of seeds in a population of peppers. **Genetics and Molecular Research**, Ribeirão Preto, v.14 n.4, p.12479-12488, 2015.

PRAZERES, C. S.; COELHO, C. M. M. Divergência genética e heterose relacionada à qualidade fisiológica em sementes de milho. **Bragantia**, Campinas v.75 n.4, p. 411-417, 2016.

QUINTÃO, V. Q. **Rede neural e lógica fuzzy aplicadas no melhoramento do feijoeiro**. Dissertação (Mestrado em Genética e Melhoramento). Universidade Federal de Viçosa, p.91, 2015.

RANGEL, M.M.;AMARAL JÚNIOR, A.T; GONÇALVES, L.S.A; FREITAS JÚNIOR S.P.; CANDIDO, L.S. Análise biométrica de ganhos por seleção em população de milho pipoca de quinto ciclo de seleção recorrente. **Revista Ciência Agronômica**, Ceará, 42: 473-48, 2011.

RAO, C. R. 1952. *Advanced statistical methods in biometric research*. New York: J. Wiley, p. 390, 1952.

RÊGO, E.R; RÊGO, M.M; CRUZ, C.D; FINGER, F.L; AMARAL, D.S.S.L. Genetic Diversity analysis of peppers: a comparison of discarding variables methods. **Crop Breeding and Applied Biotechnology**, Londrina, v. 3, n. 1, p. 19-26, 2003.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, Goiás, v. 37, n. 03, p. 182-194, 2007.

RINALDI, D.A.; PIPOLO, V.C.; GRAGE, A.C.; RUAS, C.F.; JUNIOR, N.S.F.; SOUZA, A.; SOUZA, S.G.H.; GARBUGLIO, D.D. Correlação entre heterose e divergência genética estimadas por cruzamentos dialélicos e marcadores moleculares RAPD em populações de milho pipoca. **Bragantia**, Campinas, 66:183-192, 2007.

RIGON, J. P. G.; CAPUANI, S.; CHERUBIN, M. R.; ROSA, G. M.; WASTOWSKI, A. D. Dissimilaridade genética de girassol por meio de caracteres quantitativos. **Ciência Rural**, Santa Maria, 42:1954-1959, 2012.

RIGON, J. P. G.; CAPUANI, S.; RIGON, C. A. G. Genetic divergence among maize hybrids by morphological descriptors. **Bragantia**, Campinas, v.74, n.2, p.156-160, 2015.

ROCHA, M. J.; LEITE, S. S. P.; LOMBARDI, G. M. R.; AIRTON, J.; PARRELA, R. C. A. Divergência genética entre genótipos de sorgo sacarino usando técnicas multivariadas. In: 9º Congresso Internacional de Bioenergia, 2014. **Anais...** cd-rom, São Paulo, SP., 2014.

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

RUFFATO, S.; CORREA, P.C.; MARTINS, J.H.; MANTOVANI, B.H.M; SILVA, J.N. Efeito das condições de colheita, pré-processamento e armazenamento na qualidade do milho pipoca. **Pesquisa Agropecuária Brasileira**, Brasília, 35:591-597, 2000.

SANTOS, D. B. **Análise multivariada para o agrupamento de genótipos de maracujazeiro com o uso de dados morfológicos, agronômicos e moleculares.** Dissertação (Mestrado em Ciências Agrárias). Universidade Federal do Recôncavo da Bahia. p. 58, 2010.

SANTOS ER, BARROS HB, FERRAZ E DE C, CAPONE A, SANTOS AF DOS & FIDELIS RR. Divergência entre genótipos de soja, cultivados em várzea irrigada. **Revista Ceres**, Viçosa, 58:755-764, 2011.

SANTOS, I. G. S.; **SELEÇÃO DE CARACTERES COMPLEXOS EM ALFAFA POR MEIO DE INTELIGÊNCIA COMPUTACIONAL.** Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Universidade Federal de Viçosa – UFV, Viçosa, p.91, 2017.

SANTOS, RMC; REGO, ER.; SAPUCAY, MJLC; SILVA, DF; BAIRRAL, MA; BATISTA, DS; RÊGO, MM. Avaliação de F2 de pimenteira ornamental. **Horticultura Brasileira**. Brasília: ABH, v. 27. p. 1679-1684, 2009.

SANTOS, V. S., MARTINS FILHO, S., ALVES, R. M., RESENDE, M. D. V. AND SILVA, F. F. Genetic divergence among cupuaçu accessions by multiscale bootstrap resampling. **Bragantia**, Campinas, 74, 169-175, 2015.

SARTORIO, S. D. **Aplicações de técnicas de análise multivariada em experimentos agropecuários usando o software R.** Dissertação (Mestrado em estatística e experimentação agrônômica) - Escola Superior de Agricultura Luiz de Queiroz - ESALQ, Piracicaba. p.131, 2008.

SAWAZAKI, E. Milho pipoca. In: congresso nacional de milho e sorgo. **Anais...** cd-rom, Goiânia: UFG, 2010.

- SAWAZAKI, E. A cultura do milho pipoca no Brasil. **O Agrônomo**, 53:11-13, 2001.
- SCAPIM, C. A.; AMARAL JÚNIOR, A. T.; VIEIRA, R. A.; MOTERLE, L. M.; TEIXEIRA, L. R.; VIGANÓ, J.; JÚNIOR, G. B. S. Novos compostos de milho pipoca para o Brasil. **Semina: Ciências Agrárias**, Londrina, v. 31, n. 2, p. 321-330, 2010.
- SCHAAP, M. G., & BOUTEN, W. Modeling water retention curves of sandy soils using neural networks. **Water Resources Research**, 32(10), 3033–3040, 1996.
- SEIFERT, A.L.; PIPOLO, C.V.; FERREIRA, J.M.; GERAGE, A.C. Análise combinatória de populações de milho pipoca em topcrosses. **Pesquisa Agropecuária Brasileira**, Brasília, 41:771-778, 2006.
- SHIMOYA, A.; CRUZ, C.D.; FERREIRA, R.P. Divergência genética entre acessos de um banco de germoplasma de capim-elefante. **Pesquisa Agropecuária Brasileira**, Brasília, v.37, p.971-980, 2002.
- SILVA, D. F. G.; DE JESUS COELHO, C.; ROMANEK, C.; GARDINGO, J. R.; DA SILVA, A. R.; GRACZYKI, B. L.; OLIVEIRA, E. A. T.; MATIELLO, R. R. Dissimilaridade Genética E Definição De Grupos De Recombinação Em Progênes De Meios-Irmãos De Milho-Verde. **Bragantia**, Campinas, 75 (4): 401-410, 2016.
- SILVA, G. N., TOMAZ, R. S., CASTRO, I. DE, ANNA, S., NASCIMENTO, M., & BHERING, L. L. Neural networks for predicting breeding values and genetic gains. **Scientia Agricola**, Piracicaba, v. 71(6), p. 494–498, 2014.
- SILVA, I.N. DA, D.H. Spatti, and R.A. Flauzino. **Redes neurais Artificiais para engenharias e ciências Aplicadas**. 1st ed. São Paulo: Artliber, p.399, 2010.
- SILVA, K.C.L., SILVA, K.P. DA, CARVALHO, E.V., ROTILI, E.A., AFFÉRI, F.S., PELUZIO, J.M. Divergência genética de genótipos de milho com e sem adubação nitrogenada em cobertura. **Revista Agro@mbiente On-line**, Boa Vista, v. 9, p.102-110, 2015.
- SILVA, T.R. DA C. **fito-sistema milho pipoca versus fósforo: análise genético-biométrica da eficiência nutricional de linhagens oriundas de grupos divergentes**. Dissertação (Mestrado em Produção Vegetal) Universidade Estadual do Norte Fluminense Darcy Ribeiro, 91p. 2016a.
- SILVA, V. Q. R. **Melhoramento de milho pipoca: capacidade combinatória de linhagens, parametrização heterótica e herança de características agrônomicas**. Dissertação (Mestrado em Produção Vegetal) Universidade Estadual do Norte Fluminense Darcy Ribeiro, p.198, 2009.
- SIMON, G. A.; KAMADA, T.; MOITEIRO, M. Divergência genética em milho de primeira e segunda safra. **Semina: Ciências Agrárias**, Londrina, v. 33, n. 2, p. 449-458, 2012.
- SINGH, D. The relative importance of characters affecting genetic divergence. **The Indian Journal of Genetics e Plant Breeding**, New Delhi, v.41, p.237-245, 1981.

SOFI, P. A.; ZARGAR, M. Y.; SHEIKH, F.A.; IRAM, S., SHAFI, T. Genetic variability and factor analysis in common bean (*Phaseolus vulgaris* L.) germplasm collection for yield related traits. **Electronic Journal of Plant Breeding**, Coimbatore, v.5, n.2, p.254-259, 2014.

SOUZA, J. R. DE; BOIÇA JÚNIOR, A. L.; PERECIN, D.; CARGNELUTTI FILHO, A.; COSTA, J. T. DA. Divergência genética de cultivares de cana-de-açúcar quanto à resistência a *Diatraea saccharalis*. **Semina: Ciências Agrárias**, Londrina, v.34, n.6, p.3367-3376, 2013.

SOUZA, A. E.R.; OLIVEIRA, F. J.; ANUNCIÇÃO FILHO, C. J.; MELO FILHO, P. A.; MELO, L. J. O. T.; TABOSA, J. N. Seleção de famílias RB visando à alta produtividade e precocidade na maturação em cana-de-açúcar. **Bragantia**, Campinas, v. 70, p.788-795, 2011.

STATISTICA. **Cluster Analysis**. Disponível em <<<http://www.statsoft.com/textbook/cluster-analysis/#d>>> Acessado em: 12 out. 2017.

SUDHEER, K. P., GOSAIN, A. K., RAMASASTRI, K. S. Estimating actual evapotranspiration from limited climatic data using neural computing technique. **Journal of Irrigation and Drainage Engineering**, Califórnia, v.129, p. 214-218, 2003.

SULZBACHER, L. J.; SILVA, V. P.; ZAGO, B. W.; CORRÊA, C. L.; DUARTE, A. V. M.; BARELLI, M. A. A. Análise da divergência genética através de caracteres agronômicos em genótipos de feijão comum. **Revista Espacios**, Caracas, v. 38, n.23, p. 26-38, 2017.

TEIXEIRA, D.H.L.; OLIVEIRA, M.S.P. DE.; PEDROSO, A.J.S. Estimativas de parâmetros genéticos e fenotípicos em progênies de meio-irmãos de açaizeiro para caracteres de produção de frutos em vários anos de colheita. In: congresso brasileiro de fruticultura, annual meeting of the interamerican society for tropical horticulture, 54., 2008, Vitória. **Anais...** Sociedade Brasileira de Fruticultura, p.1-5, Vitória, 2008.

TEODORO, P. E.; BARROSO, L. M. A.; NASCIMENTO, M.; TORRES, F. E.; SAGRILO, E.; SANTOS, A.; RIBEIRO, L. P. Redes neurais artificiais para identificar genótipos de feijão-caupi semiprostrado com alta adaptabilidade e estabilidade fenotípicas. **Pesquisa Agropecuária Brasileira** (Online), Brasília, v. 50, p. 1054- 1060, 2015.

TORRES FILHO, J. **Caracterização morfo-agronômica, seleção de descritores e associação entre a divergência genética e a heterose em meloeiro**. Tese (Doutorado em Fitotecnia) – Universidade Federal Rural do Semi-Árido, Mossoró, p. 150, 2008.

VASCONCELOS, E. S.; CRUZ, C. D.; BHERING, L. L.; RESENDE JÚNIOR, M. F. R. Método alternativo para análise de agrupamento. **Pesquisa Agropecuária Brasileira**, Brasília, v.42, n.10, p.1421-1428, 2007.

VAZ-DE-MELO, A.; SANTOS, L. D. T.; FINOTO, E. L.; DIAS, D. C. F. S.; ALVARENGA, E. M. Germinação e vigor de sementes de milho pipoca submetidas ao estresse térmico e hídrico. **Bioscience Journal**, Uberlândia, v.28, n.5, p.687-695, 2012.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, p. 496, 1992.

VILARINHO, A.A.; VIANA, J.M.S; CÂMARA T,M.M, SANTOS, J.F. Seleção de progênies endogâmicas S<sub>1</sub> e S<sub>2</sub> de melhoramento intrapopulacional de milho pipoca. **Acta Scientiarum**, Maringá, v.24, p. 1419-1425, 2002.

VITTORAZZI, C. **Grupos heteróticos em germoplasma de milho pipoca**. Tese (Doutorado em Genética e Melhoramento de Planta) Universidade Estadual do Norte Fluminense Darcy Ribeiro, p. 68, 2016.

VOGT, G. A.; BALBINOT JUNIOR, A. A.; SOUZA, A. M. Divergência fenotípica em genótipos de girassol. **Revista de Ciências Agroveterinárias**, Campinas, V.11, p.26-34, 2012.

ZIEGLER, K.E.; ASHMAN, B. Popcorn. In: HALLAUER, A.E. (ed.). **Specialty Corns**. London: CRC,. p. 189-223. 1994

ZINSLY, J.R.; MACHADO, J.A. Milho pipoca. In: PATERNIANI, E.; VIEGAS, G.P. (ed.). **Melhoramento e produção do milho**. Campinas: Fundação Cargill, p. 713-717, 1987.

## 8. ANEXOS

**Anexo 1.** Resumo das análises de variância conjunta (safras I 2014/2015 e safra II 2016/2017) para os caracteres de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP); quadrado médio dos blocos no ambiente (BI/amb), interação linhagens x ambiente (L x Amb).

Caract	Quadrado Médio						
	BI/Amb (4) <sup>1</sup>	Linhagens (19) <sup>1</sup>	Ambiente (1) <sup>1</sup>	L x Amb (19) <sup>1</sup>	Resíduo (76) <sup>1</sup>	Média	Ac ( $\bar{r}_{gg}$ )
FF	19,55	17,25 <sup>ns</sup>	576,40 <sup>****</sup>	15,98*	8,20	75,09	25,57
FM	20,44	17,53 <sup>ns</sup>	785,40 <sup>**</sup>	16,96 <sup>**</sup>	6,21	73,97	17,06
AP	501,59	803,24 <sup>**</sup>	25404,00 <sup>**</sup>	149,24 <sup>ns</sup>	119,66	146,21	90,22
AE	345,30	286,47*	14834,52 <sup>**</sup>	106,74 <sup>ns</sup>	78,03	73,91	79,17
DC	99,79	8,76 <sup>ns</sup>	698,99 <sup>ns</sup>	6,10 <sup>ns</sup>	3,67	19,40	54,83
PROF	1,51	0,33*	0,27 <sup>ns</sup>	0,14 <sup>ns</sup>	0,09	2,12	75,30
AQ	9,20	15,14 <sup>ns</sup>	1062,07 <sup>**</sup>	20,42 <sup>**</sup>	6,21	5,82	14,20
NEAP	51,54	19,15 <sup>ns</sup>	559,00*	10,53*	5,52	7,12	94,91
CPE	1,02	3,12 <sup>ns</sup>	399,67 <sup>**</sup>	3,72*	2,04	11,38	15,80
DE	14,90	20,70 <sup>ns</sup>	1068,39 <sup>**</sup>	12,94 <sup>ns</sup>	14,26	27,68	60,91
NFG	5,83	3,41 <sup>ns</sup>	53,14*	2,33 <sup>ns</sup>	1,44	11,77	56,13
M100	6,14	24,54*	0,21 <sup>ns</sup>	8,84 <sup>ns</sup>	7,97	13,23	79,93
PROD	48254,32	85068,77 <sup>ns</sup>	5732369,82 <sup>**</sup>	49977,38 <sup>**</sup>	21573,48	389,62	64,16
CE	6,01	48,45 <sup>ns</sup>	94,11*	28,10 <sup>**</sup>	6,01	19,18	64,69
VP	10450939,41	33730530,23 <sup>ns</sup>	2417849499,38 <sup>**</sup>	31633655,16*	17315839,49	7601,35	23,79

<sup>1</sup> Grau de liberdade. <sup>ns</sup>, \*\*, \*, Não significativo, Significativo a 1% e 5% de probabilidade respectivamente, pelo teste F.

**Anexo 2.** Agrupamento de médias pelo teste de Scott-Knott para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade de grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S<sub>2</sub> de milho pipoca na safra de 2014/2015, em Dourados-MS.

<b>Linhagens</b>	<b>FF</b>	<b>FM</b>	<b>M100</b>	<b>PROD</b>	<b>CE</b>	<b>VP</b>	<b>NFG</b>	<b>DE</b>
<b>1</b>	70,33 a	70,33 a	15,36 b	575,74 b	20,94a	11854,09 b	12,60 a	31,24 a
<b>2</b>	73,66 a	71,33 a	13,07 c	693,22 a	19,05 b	13115,20 a	12,55 a	31,71 a
<b>3</b>	71,33a	70,00 a	14,52 b	650,22 a	26,94 a	16866,47 a	12,72 a	33,03 a
<b>4</b>	72,33 a	71,00 a	11,56 c	500,03 b	22,00 a	10523,33 b	12,66 a	29,25 a
<b>5</b>	72,66 a	70,00 a	15,36 b	624,65 a	21,05 a	13305,20 a	11,50 a	29,40 a
<b>6</b>	69,33 a	68,66 a	11,88 c	520,04 b	17,72 b	8801,51 b	11,90 a	27,42 a
<b>7</b>	75,00 a	71,66 a	12,39 c	339,43 b	24,27 a	8184,59 b	12,66 a	34,27 a
<b>8</b>	77,33 a	76,00 a	19,55 a	1073,27 a	16,83 b	17997,34 a	13,06 a	34,55 a
<b>9</b>	71,66 a	71,00 a	8,54 c	305,63 b	16,66 b	4803,18 b	11,73 a	28,35 a
<b>10</b>	74,66 a	72,66 a	17,86 a	817,80 a	14,88 b	12160,50 b	12,75 a	35,04 a
<b>11</b>	70,00 a	69,00 a	9,47 c	267,72 b	20,16 b	5386,60 b	11,42 a	32,56 a
<b>12</b>	71,33 a	70,00 a	12,36 c	687,92 a	16,66 b	11490,12 b	14,54 a	29,74 a
<b>13</b>	73,66 a	71,33 a	13,47 c	819,62 a	22,16 a	18009,78 a	11,73 a	30,11 a
<b>14</b>	74,66 a	72,33 a	13,76 c	387,85 b	12,66 b	4920,02 b	11,53 a	30,58 a
<b>15</b>	73,33 a	73,33 a	10,90 c	850,23 a	24,61 a	20930,70 a	12,73 a	29,19 a
<b>16</b>	73,66 a	72,33 a	13,03 c	468,92 b	24,61 a	11596,46 b	11,61 a	28,10 a
<b>17</b>	74,00 a	72,00 a	12,63 c	625,54 a	22,33 a	13979,06 a	12,53 a	29,30 a
<b>18</b>	71,66 a	72,00 a	10,84 c	545,91 b	19,38 b	10606,82 b	13,44 a	28,14 a
<b>19</b>	75,00 a	73,00 a	14,16 b	728,11 a	22,55 a	16660,53 a	13,29 a	31,05 a
<b>20</b>	72,66 a	70,33 a	14,75 b	680,88 a	16,00 b	10610,13 b	11,73 a	30,21 a
<b>Linhagem</b>	<b>AQ</b>	<b>NEAP</b>	<b>AE</b>	<b>AP</b>	<b>PROLIF</b>	<b>CPE</b>	<b>DC</b>	
<b>1</b>	13,66 a	9,33 b	90,83 a	168,16 a	1,86 b	12,80 b	18,53 a	
<b>2</b>	8,33 b	9,00 b	89,00 a	163,83 a	2,06 b	13,04 b	21,27 a	
<b>3</b>	8,00 b	7,66 b	97,83 a	176,50 a	2,00 b	14,68 a	23,97 a	
<b>4</b>	7,33 b	9,00 b	88,16 a	170,16 a	2,43 a	14,56 a	23,54 a	
<b>5</b>	11,00 a	8,33 b	102,83 a	175,16 a	2,06 b	12,70 b	23,76 a	
<b>6</b>	11,33 a	9,33 b	93,00 a	164,66 a	2,00 b	12,98 b	24,31 a	
<b>7</b>	7,66 b	5,66 b	82,50 b	162,66 a	1,83 b	14,76 a	20,24 a	
<b>8</b>	11,00 a	12,33 a	86,50 a	171,33 a	2,00 b	12,15 b	20,15 a	
<b>9</b>	2,33 b	9,66 b	63,66 b	132,16 b	2,06 b	14,60 a	22,54 a	
<b>10</b>	8,33 b	9,66 b	75,00 b	155,00 a	2,00 b	12,77 b	19,76 a	
<b>11</b>	7,66 b	8,66 b	77,16 b	168,5 a	2,26 a	14,03 a	21,93 a	
<b>12</b>	9,00 b	12,33 a	73,83 b	156,00 a	1,66 b	11,83 b	21,93 a	
<b>13</b>	9,33 b	13,66 a	79,48 b	158,33 a	2,16 b	12,65 b	21,90 a	
<b>14</b>	10,33 a	5,66 b	90,41 a	161,71 a	2,56 a	14,00 a	23,50 a	
<b>15</b>	14,66 a	6,66 b	88,66 a	144,83 b	2,86 a	11,06 b	20,82 a	
<b>16</b>	5,66 b	7,00 b	88,50 a	156,18 a	2,40 a	12,71 b	22,72 a	
<b>17</b>	4,33 b	9,66 b	76,11 b	156,16 a	2,23 b	13,84 a	23,20 a	
<b>18</b>	12,33 a	6,33 b	85,83 a	158,66 a	2,36 a	12,79 b	21,41 a	
<b>19</b>	4,66 b	9,66 b	83,66 a	155,50 a	2,56 a	11,99 b	19,42 a	
<b>20</b>	9,00 b	12,00 a	87,66 a	159,83 a	1,93 b	14,23 a	21,43 a	

\* médias seguidas de mesma letra não diferem pelo teste de Scott-knott a 5% de probabilidade.

**Anexo 3.** Agrupamento de médias pelo teste de Scott-Knott para os componentes de florescimento feminino (FF), florescimento masculino (FM), altura de plantas (AP), altura de espigas (AE), diâmetro de colmo (DC), prolificidade (PROLIF), planta acamada e quebrada (AQ), número de espigas autofecundadas por parcela (NEAP), comprimento de espiga (CPE), diâmetro de espiga (DE), número de fileiras de grãos (NFG), massa de 100 grãos (M100), produtividade grãos (PROD), capacidade de expansão (CE) e Volume de pipoca (VP) em 20 linhagens S<sub>3</sub> de milho pipoca na safra de 2016/2017, em Dourados-MS.

<b>Linhagem</b>	<b>FF</b>	<b>FM</b>	<b>M100</b>	<b>PROD</b>	<b>CE</b>	<b>VP</b>	<b>NFG</b>	<b>DE</b>
1	76,00 b	74,33 b	16,38 a	177,15 a	19,33 b	3424,41 b	11,90 a	26,35 a
2	80,66 a	78,66 a	12,38 a	129,66 a	11,66 c	1511,93 b	10,23 a	23,22 a
3	81,66 a	80,66 a	13,64 a	178,66 a	23,50 a	4217,21 a	10,76 a	26,05 a
4	78,33 a	79,00 a	14,94 a	144,23 a	22,16 a	3185,77 b	10,69 a	23,83 a
5	75,00 b	74,33 b	13,99 a	230,32 a	13,16 c	2991,48 b	12,23 a	28,20 a
6	74,33 b	74,66 b	15,07 a	199,70 a	12,33 c	2479,04 b	12,26 a	25,79 a
7	79,33 a	77,66 a	11,64 a	71,95 a	16,83 b	1223,72 b	8,52 a	20,50 a
8	72,66 b	72,00 b	14,60 a	286,70 a	18,83 b	5407,90 a	12,06 a	28,54 a
9	78,00 a	78,00 a	9,95 a	161,89 a	20,50 b	3420,58 b	11,26 a	19,89 a
10	73,00 b	72,00 b	16,65 a	320,43 a	17,83 b	5769,54 a	11,80 a	27,35 a
11	74,66 b	73,66 b	14,75 a	215,17 a	22,83 a	4918,80 a	11,17 a	25,38 a
12	75,00 b	72,66 b	13,21 a	208,87 a	16,99 b	3474,48 a	13,11 a	27,57 a
13	77,00 b	76,33 b	11,72 a	177,92 a	14,83 c	2537,33 b	11,43 a	26,37 a
14	76,33 b	74,33 b	12,29 a	133,02 a	19,00 b	2527,88 b	11,04 a	23,34 a
15	78,00 a	78,66 a	9,06 a	94,79 a	17,33 b	1654,19 b	10,43 a	20,87 a
16	81,00 a	81,00 a	15,68 a	104,54 a	22,99 a	2404,10 b	9,00 a	25,34 a
17	81,33 a	81,33 a	11,03 a	85,12 a	19,49 b	1655,62 b	11,53 a	25,08 a
18	78,66 a	78,00 a	11,91 a	102,39 a	18,49 b	1886,30 b	9,77 a	22,91 a
19	78,33 a	77,66 a	13,08 a	237,76 a	19,99 b	4661,50 a	12,00 a	24,04 a
20	76,33 b	75,66 b	11,80 a	160,67 a	17,99 b	2900,76	10,88 a	23,37 a
<b>Linhagem</b>	<b>AQ</b>	<b>NEAP</b>	<b>AE</b>	<b>AP</b>	<b>PROLIF</b>	<b>CPE</b>	<b>DC</b>	
1	1,66 b	4,00 b	60,53 a	137,13 a	2,06 b	10,31a	16,14 b	
2	1,00 b	3,00 b	64,26 a	136,80 a	2,33 a	8,62 a	16,95 b	
3	3,33 a	5,33 b	61,93 a	137,33 a	2,06 b	8,17 a	16,39 b	
4	2,00 b	5,00 b	65,66 a	137,13 a	2,60 a	9,29 a	18,55 a	
5	2,00 b	6,00 b	72,46 a	144,26 a	2,00 b	9,55 a	19,23 a	
6	3,66 a	7,00 b	66,13 a	128,73 a	2,13 b	10,58 a	18,27 a	
7	1,66 b	2,66 c	74,93 a	150,40 a	1,66 b	9,75 a	15,08 b	
8	2,33 b	6,66 b	67,40 a	139,33 a	2,00 b	8,36 a	15,74 b	
9	3,00 a	6,33 b	43,60 a	88,46 c	2,60 a	8,05 a	16,65 b	
10	2,00 b	10,00 a	71,66 a	145,80 a	1,93 b	9,72 a	16,51 b	
11	0,66 b	6,00 b	59,33 a	138,80 a	2,06 b	8,70 a	15,93 b	
12	4,00 a	6,00 b	56,86 a	131,66 a	1,66 b	8,46 a	14,91 b	
13	3,33 a	5,00 b	57,93 a	137,80 a	2,26 a	11,09 a	18,62 a	
14	5,66 a	4,33 c	59,73 a	124,26 a	2,40 a	9,82 a	14,75 b	
15	4,00 a	4,33 c	58,40 a	106,60 b	2,26 a	9,13 a	17,36 b	
16	4,33 a	2,33 c	58,46 a	120,00 b	1,800 b	11,66 a	15,91 b	
17	5,66 a	3,66 c	59,86 a	131,80 a	2,00 b	9,35 a	19,10 a	
18	1,66 b	2,66 c	62,66 a	135,73 a	1,86 b	8,93 a	16,82 b	
19	4,00 a	6,66 a	72,53 a	143,20 a	1,93 b	10,01 a	19,86 a	
20	1,00 b	2,33 c	61,53 a	117,13 b	1,8 b	11,51 a	17,01 b	

\* médias seguidas de mesma letra não diferem pelo teste de Scott-knott a 5% de probabilidade.