ISSN: 1984-5529

Científica

Jaboticabal

v.48, n.4, p.357-362, 2020

http://dx.doi.org/10.15361/1984-5529.2020v48n4p357-362

Principal component analysis for selection of superior maize genotypes

Análise de componentes principais para seleção de genótipos superiores de milho

Eduardo Sávio Gomes CARNIMEO[1]; Luiz Eduardo Tilhaqui BERTASELLO[2];
Sophia Mangussi Franchi DUTRA[3]; Gustavo Vitti MÔRO[4]

[1] Graduando em Engenharia Agronômica, Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências Agrárias e Veterinárias, Câmpus de Jaboticabal, eduardocarnimeo@hotmail.com

[2] Doutorando em Agronomia – Genética e Melhoramento de Plantas, Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências Agrárias e Veterinárias, Câmpus de Jaboticabal, luiz.bertasello@unesp.br

[3] Doutora em Agronomia – Genética e Melhoramento de Plantas, Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências Agrárias e Veterinárias, Câmpus de Jaboticabal, sophiamfd@gmail.com

[4] Autor para correspondência - Professor Associado, Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências Agrárias e Veterinárias, Câmpus de Jaboticabal, Departamento de Ciências da Produção Agrícola, gv.moro@unesp.br

Recebido em: 25-05-2020; Aceito em: 21-07-2020

Abstract

Constant advances in studies on the behavior of maize genotypes and their interactions with the environment are of great importance for the best performance of the plant. This study verifies effects and causes of agronomic variables of maize hybrids on grain yields and performs the indirect selection of superior genotypes by principal component analysis (PCA). Two hundred and thirty maize genotypes were used, with two hundred and twenty--nine topcross hybrids (consisting of crossings of two hundred and twenty-nine partially inbred genotypes with a tester) and one check in a randomized block design with two repetitions. The genotypes were evaluated during the 2016 and 2016/2017 crops considering the agronomic variables plant height, ear insertion height, ear position, lodging, breakage, and grain yield. Data were submitted to analysis of variance and means were compared by the Scott-Knott test ($p<0.05$) with subsequent multivariate exploratory analysis by PCA. In the principal component analysis, components explained 52.07% and 55.69% of the variance contained in the original variables for the 2016 and 2016/2017 crops, respectively. The variable that was most significant in both crops was ear insertion height, allowing the indirect selection of more productive genotypes. Indirect selection of the most productive genotypes was also conducted through variables that contributed significantly in the principal component analysis. Thus, the use of multivariate exploratory analysis is efficient in the characterization and selection of maize genotypes evaluated in different crop seasons.

Additional keywords: first and second crop season; hybrids; principal components; *Zea mays*.

Resumo

Os constantes avanços de estudos no comportamento de genótipos de milho são de grande importância e suas interações com o ambiente, buscando o melhor desempenho da planta. Objetivou verificar por meio de análises exploratórias multivariada por componentes principais, o efeito e a causa que variáveis agronômicas de híbridos de milho exercem sobre a produtividade de grãos e realizar seleção indireta de genótipos superiores. Foram utilizados 230 genótipos de milho, dos quais 229 deles eram híbridos *topcrosses* (constituintes pelos cruzamentos de 229 genótipos parcialmente endogâmicos com um testador) e uma testemunha, em delineamento de blocos casualizados com duas repetições. Os genótipos foram avaliados durante as safras 2016 e 2016/2017, considerando as variáveis agronômicas altura de plantas, altura de inserção da espiga, posição relativa da espiga, acamamento, quebramento e produtividade de grãos. Os dados foram submetidos à análise de variância, as médias comparadas pelo teste de Scott-Knott ($p<0,05$) e foi realizada análise exploratória multivariada por componentes principais. Na análise de componentes principais, os componentes explicaram 52,07% e 55,69% da variância contida nas variáveis originais para a safra de 2016 e 2016/2017, respectivamente. A variável que mais contribuiu significativamente em ambas as safras foi a altura de inserção da espiga, permitindo a seleção indireta de genótipos mais produtivos. Realizou-se também seleção indireta dos genótipos mais produtivos por meio das variáveis que contribuíram significativamente na análise de componentes principais. Conclui-se, que o uso de análise exploratórias multivariada é eficiente na caracterização e seleção de genótipos de milho avaliados em diferentes safras agrícolas.

Palavras-chave adicionais: componentes principais; híbridos; primeira e segunda safra; *Zea mays*.

Introduction

In addition to ethanol and human food production, a large part of maize (*Zea mays*) crops is used for animal feed (cattle, swine, and poultry) (Galvão et al., 2014), being their main source of carbohydrates. It is known that breeding programs are essential for the selection of superior genotypes, ensuring increased yield among Brazilian crops (Alves et al., 2015).

Due to the economic importance of this crop, the demand for cultivars with superior agronomic characteristics and increased yield potential is growing (Seraguzi et al., 2016). Yield increases, in part, by the improvement of cultivation techniques (Besen et al., 2019) and by the use of promising genotypes, such as more adapted maize hybrids (Nardino et al., 2016). The marketing and choosing of seeds are increasingly selective in relation to maize hybrids due to requirements from farmers, and grain yield is one of the main variables taken into account when choosing the material to be cultivated (Nascimento et al., 2011; Serpa et al., 2012; Silva et al., 2015).

In maize breeding programs, increasing grain yield without losing nutritional quality and grain composition is the main premise (Alves et al., 2015). Therefore, the study of linear relationships between characters can provide important results, especially in the identification of those that can be used in indirect selection (Cargnelutti Filho et al., 2011). Thus, selection with only one character may cause concomitant changes in other plant characteristics.

Parental selection is one of the primary steps in plant breeding, in which favorable alleles must be concentrated for the characteristics of interest, allowing the achievement of superior individuals (Oliboni et al., 2013). As a result, tools that facilitate the collection of information on the base population under study have been increasingly used (Souza Neto et al., 2015).

Breeders have been using the topcross method proposed by Davis (1927), which evaluates the relative merit of many partially inbred strains crossbreeding with a tester genotype in common. Topcross is the crossing between selections, strains or clones with a male parent in common (Borém, 1998). Among procedures for measuring combinatorial capacity, topcross stands out mainly for being easy to perform and obtaining the combination parameters (Duarte et al., 2003). Thus, the use of the topcross technique allows us to evaluate and select strains, eliminating those with low agronomic performance, making the development of hybrids faster, cheaper, and more efficient.

Commercial hybrids are widely used in Brazil. According to Ferreira et al. (2010), Possatto Junior et al. (2017), and Gabriel et al. (2018), in order to confirm superior hybrids, they are tested and evaluated in various locations and environments, selecting hybrids with low presence of harmful alleles and abundance of favorable alleles for fixation and recommendation in different characteristics, such as resistance to pests and diseases, plant architecture, plant height, ear insertion height, vegetative cycle, and grain yield.

Conventional statistical analyzes are often insufficient to describe interactions between the maize agronomic variables studied (Samecima Junior, 2018). One of the solutions proposed is the use of multivariate statistics, as it is more capable of describing the relationships of interdependence between variables (Silva et al., 2016). Principal component analysis (PCA) is a multivariate exploratory modeling technique that defines important components based on a covariance structure (Hongyu et al., 2015). This technique is associated with reduced losses of data and information, and with a better view of all or almost all the variables studied to solve practical problems, which are often difficult to be solved when analyzed separately (Hongyu et al., 2015).

Some studies have applied multivariate techniques to evaluate agronomic variables with satisfactory results (Maia et al., 2006; Silva et al., 2016; Samecima Junior, 2018; Bertasello, 2018). Given the importance of these analyzes, the use of indirect selection can accelerate the obtaining and identification of promising genotypes, since the observed characters are qualitative and easy to measure, with early measurement and high heritability, besides being capable of correlating with other characters.

Therefore, this study verifies the effects and causes of agronomic variables of maize hybrids on grain yields and performs the indirect selection of superior genotypes using the multivariate exploratory modeling technique of principal component analysis.

Material and methods

The experiments were conducted during the 2016 and 2016/2017 crops at UNESP, Jaboticabal Campus, São Paulo State, Brazil (21°14’05’’ S latitude and 48°17’09’’ W longitude, altitude of 614 m). The climate corresponds to the Köppen climate classification category Aw, characterized by dry winter, with average temperature of 22.3 ºC and annual rainfall of 1444 mm. (Alvares et al., 2013). The soil of the experimental area is classified as Eutroferric Red Latosol (EMBRAPA, 2013).

The experimental design was a randomized block with two repetitions. Each plot consisted of two 5-m long rows, spaced at 0.5 meters between rows and 0.33 meters between plants, representing 62,500 plants per hectare. Sowing was performed with a plot seeder.

Two hundred and thirty maize genotypes were used, with two hundred and twenty-nine topcross hybrids (consisting of crossings of two hundred and twenty-nine partially inbred genotypes with a tester, all genotypes from the company Phoenix Agricola Ltd.) and one commercial check, the hybrid DKB 390 VT PRO2.

Base fertilization was performed with 350 kg ha^{-1} of NPK in the formula (8-28-16) at sowing. Topdressing was applied thirty days after sowing with 311 kg ha^{-1} of common urea (45% N) distributed in a continuous fillet at up to 10 cm from the plot, providing

140 kg ha^{-1} of N. One day before application, water was supplemented to the plants through sprinkler irrigation.

The genotypes were evaluated considering the following agronomic variables:

a) Plant height (PH): measured as the distance from the soil to the flag leaf insertion, measured using a ruler graduated every five centimeters. Eight plants were evaluated per plot, with results expressed in cm.

b) Ear insertion height (EH): measured as the distance from the soil to the insertion point of the main ear, measured using a ruler graduated every five centimeters. Eight plants were evaluated per plot, with results expressed in cm.

c) Ear relative position (ERP): Obtained by the ratio between ear insertion height and plant height.

d) Lodging (LO): at harvest, the number of plants with angle slope above 45° in relation to the vertical was evaluated.

e) Breakage (BRE): at harvest, the number of plants with breakage in the stem below the main ear was evaluated.

f) Grain yield (GY): obtained by threshing the harvested ears and weighing the grains of each plot, correcting the moisture to 13% and converting the values to tons per hectare (t ha^{-1}).

Plant lodging and breakage data were transformed into $\sqrt{x+0.5}$ for normality and later converted to percentage. The data of variables were submitted to analysis of variance (ANOVA) using the statistical software R version 3.6.0 (R Development Core Team, 2016) to verify the effects between treatments. Data normality and homogeneity of the variances were evaluated using the Shapiro-Wilk and Bartlett tests at 5% probability.

All data were submitted to the exploratory multivariate modeling technique of principal component analysis in both crop seasons. The software Statistica®, version 7.0 (Statsoft, 2008) was used for analysis.

Results and discussion

Genotypes affected all variables ($p<0.05$) except for grain yield ($p>0.05$). Lodging and breakage variables showed the highest variation coefficient among the variables evaluated, explaining variance during crops (Table 1).

Table 1 - Summary of the analysis of variance of the maize agronomic characters evaluated during the crops years 2016 and 2016/2017.

Crop season	Agronomic variables					
	PH	EH	ERP	LO	BRE	GY
			P-Value			
2016	0.0002*	0.003*	0.00004*	$1.9 \times 10^{-08*}$	0.0001*	0.31
2016-17	0.00001*	0.00001*	$9.8 \times 10^{-09*}$	0.00001*	0.002*	0.65
			CV(%)			
2016	5.47	4.69	4.12	75.47	50.34	14.12
2016-17	3.65	3.28	4.74	136.81	78.92	15.68

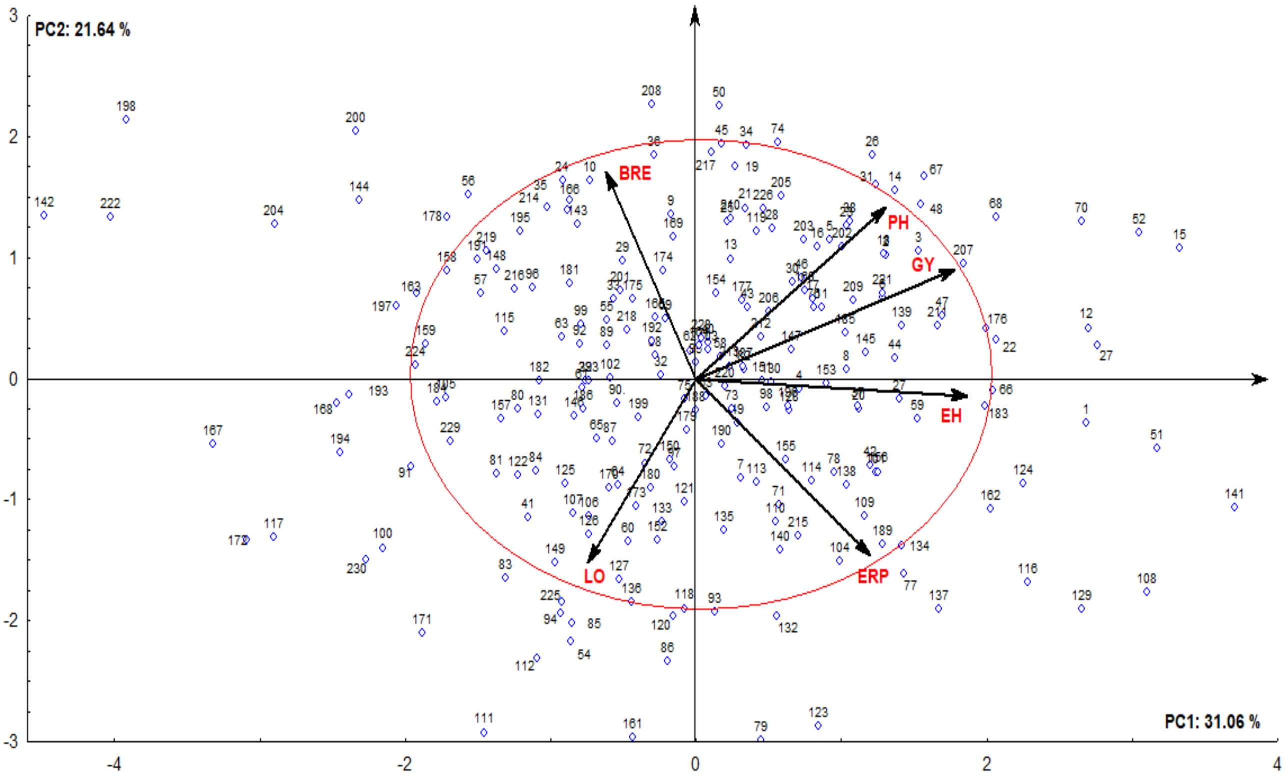
PH: Plant height. EH: Ear insertion height. ERP: Ear relative position. LO: Plant lodging. BRE: Plant breakage. GY: Grain yield. CV: coefficient of variation. *significant difference.

The total amount of information of the original variables retained by the two principal components in the 2016 crop was 52.07%. PC 1 showed 31.06% of variance and PC 2 showed 21.64% (Figure 1). In the principal component 1, the agronomic variable ear insertion height was more significant for being close to the axis. The agronomic variables plant height, grain yield, and ear relative position showed little relation or significance, since they were angled at more than 90º to the axis (Hongyu et al., 2015). In the principal component 2, the agronomic variables breakage and lodging were more significant because they were closer to the axis, although inversely proportional.

Ear insertion height was the variable that was more significant to explain the difference between the genotypes evaluated in PC1, explaining 31.06% of the variance between them in the 2016 crop (Figure 1). It is observed that genotypes 183, 59, and 27 differed for ear insertion height, possibly being the genotypes with the highest means. Genotypes 1, 51, 141, 124, 162, and 12, with lower means, stood out negatively for this agronomic variable.

The results corroborate those found by Silva et al. (2016) and Samecima Junior (2018), in which ear insertion height was the indirect selection with the most important contribution to the variability of maize genotypes evaluated by multivariate analysis.

The breakage and lodging variables were inversely proportional, as the breakage of the stem below the ear can topple or lodge maize plants, explaining 21.64% of the variance between the maize genotypes evaluated. Genotypes that differed the most for breakage were 10, 24, and 35, while genotypes that differed the most for lodging were 149, 60, 126, 127, and 136 (Figure 1).



PH: Plant height. EH: Ear height. BRE: Plant breakage. LO: Plant lodging. ERP: Ear relative position. GY: Grain yield.

Figure 1 - Biplot dispersion graph of 230 maize genotypes during the 2016 harvest.

Therefore, the most recommended genotypes, which are less susceptible to breakage, are 142, 222, and 198, and genotypes less susceptible to lodging were 111 and 161.

Thus, principal component analysis contributed significantly for a good indirect selection, identifying inferior and superior genotypes, which are capable of increased yield, reducing agricultural losses (Toebe & Cargnelutti Filho, 2013; Kavalco et al., 2014).

The amount of information of the original variables of the sum of principal components was 55.69%, with PC1 showing 31.55% of variance and PC2 showing 24.14%. In PC1, the agronomic variables plant height, ear insertion height, and ear relative position were more significant because they were close to the axis. In PC2, the lodging agronomic variable was more significant because it was closer to the axis (Figure 2).

Genotypes 21, 84, 126, and 200 differed for agronomic variables ear insertion height, plant height, and ear relative position (Figure 2), being possibly the most productive genotypes. Genotype 90 had the worst performance, with the lowest yield.

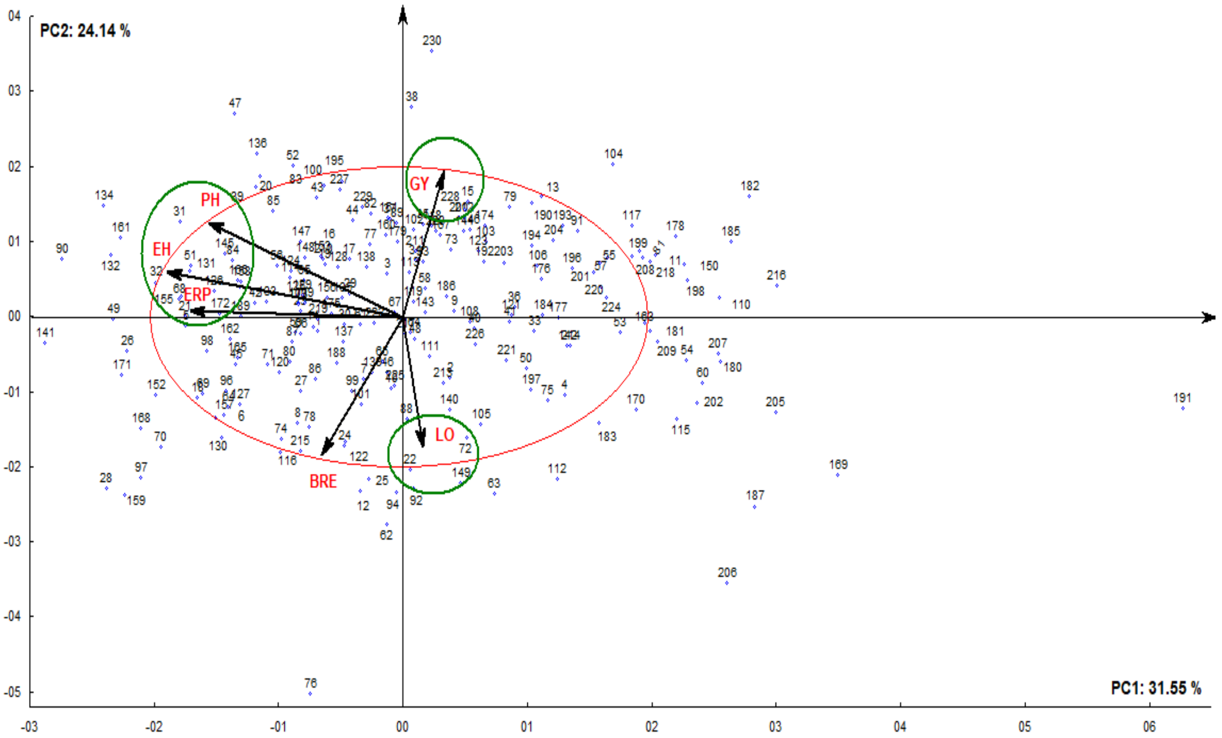
Therefore, although ear insertion height has a higher estimate of variance, this variable corresponds to indirect effects of other agronomic variables. This variance can be linked especially to indirect effects of plant height (Entringer et al., 2015). Thus, these variables are efficient for increasing ear insertion height.

Genotypes 22, 72, and 88 differed for the agronomic variable lodging (Figure 2), being the genotypes most susceptible to yield loss. Genotypes 62 and 76 stood out positively, being less susceptible to plant breakage, reducing yield loss.

Plant height has a great relationship with plant breakage and lodging in the field. When allied, these variables decrease stem resistance. Authors note that these variables are correlated, where the higher the plant height, the greater the breakage and lodging of maize genotypes (Vieira et al., 2017).

In this study, the relationship between ear insertion height and breakage and lodging was noticeable in the two crop seasons evaluated. Therefore, it is important to choose other agronomic variables to evaluate these causes and reduce yield losses, such as the measurement of the stem and the duration of the vegetative cycle of maize genotypes.

Thus, the selection of genotypes that correlate with the lodging and breakage characters is indicated to provide better resistance, since these are the variables that contributed and differed the most in the principal component analysis of this study (Figures 1 and 2). In general, this study helps in the recommendation and indirect selection of hybrids that are more resistant to breakage and lodging, increasing yields for maize grains.



PH: Plant height. EH: Ear height. BRE: Plant breakage. LO: Plant lodging. ERP: Ear relative position. GY: Grain yield.

Figure 2 - Biplot dispersion graph of 230 maize genotypes during the 2016/2017 harvest.

Conclusions

Multivariate analysis is a viable technique for obtaining and selecting maize genotypes. From the distribution in the principal component analysis, it is possible to indirectly select the most productive genotypes based on ear insertion height, plant height, ear relative position, lodging, and breaking. Indirect selection was not possible using grain yield, since this agronomic variable was not a discriminant, besides not showing any significant differences between the evaluated genotypes.

Acknowledgment

To the National Council for Scientific and Technological Development (CNPq) for granting financial support; to the Center of Studies in Genetics and Maize Improvement (NEGEMM) and to the Faculty of Agricultural and Veterinary Sciences of Jaboticabal-SP (UNESP/ FCAV).

References

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Koppen's climate classification map for Brazil. Meteor Zeits 22(6):711-728.

Alves BM, Cargnelutti Filho A, Burin C, Toebe M, Da Silva LP (2015) Divergência genética de milho transgênico em relação à produtividade de grãos e à qualidade nutricional. Ciência Rural 45(5):884-891.

Bertasello LET (2018) Efeito da adubação nitrogenada e inoculação com *Azospirillum brasilense* em caracteres agronômicos de genótipos de milho. Unesp-Fcav (Dissertação de Mestrado em Agronomia Genética e Melhoramento de Plantas) - Universidade Estadual Paulista, Jaboticabal.

Besen MR, Ribeiro RH, Figueroa LV, PIVA JT (2019) Produtividade do milho em resposta à inoculação com. Revista Brasileira de Milho e Sorgo 18(2):257-268.

Borém A (1998). Melhoramento de plantas. 2ª ed. UFV/Viçosa. 453p.

Cargnelutti Filho A, Lopes SJ, Toebe M, Silveira TR, Schwantes IA (2011) Tamanho de amostra para estimação do coeficiente de correlação de *Pearson* entre caracteres de *Crambe abyssinica*. Revista Ciência Agronômica 42(1):149-158.

Davis RL (1927) Report of the plant breeder. Annals Reports Agricultural Experiment. Puerto Rico, 14-15p.

Duarte IA, Ferreira JM, Nuss CN (2003) Screening potential of three maize *topcross* testers. Pesquisa Agropecuária Brasileira 38(3):365-372.

Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA (2013) Sistema brasileiro de classificação de solos. 3.ed. Brasília, 353p.

Entringer GC, Santos PHAD, Vettorazzi JCF, Cunha KSD, Pereira MG (2015) Correlação e análise de trilha para componentes de produção de milho super doce. Ceres 61(3):356-361.

Ferreira EA, Paterniani MEAGZ, Santos FMC (2010) Potencial de híbridos comerciais de milho para obtenção de linhagens em programas de melhoramento. Pesquisa Agropecuária Tropical 40(3):304-311.

Gabriel A, Faria MV, Mendes MC, Gava E, Zaluski WL, Júnior OP, Cescon LK (2018) Avaliação de linhagens S3 de milho em cruzamento *topcross* em dois espaçamentos entre linhas. Revista Brasileira de Milho e Sorgo 17(2):328-339.

Galvão JCC, Miranda GV, Trogello M, Fritsche-Neto R. (2014) Sete décadas de evolução do sistema produtivo da cultura do milho. Revista Ceres 61(2):819-828.

Hongyu K, Sandanielo VLM, Oliveira Junior GJ (2015) Análise de componentes principais: Resumo teórico, aplicação e interpretação. Engineering and Science 5(1):83-90.

Kavalco SAF, Figueiredo R, Groli EL, Zimmer CM, Baretta D, Tessmann EW, Júnior AM de M, Oliveira AC de (2014) Análise de trilha em genótipos de trigo submetidos ao estresse por encharcamento. Semina: Ciências Agrárias 35(4):1683-1695.

Maia MCC, Vello NA, Rocha MM, Pinheiro JB, Silva NF (2006) Adaptabilidade e estabilidade de linhagens experimentais de soja selecionadas para caracteres agronômicos através de método uni-multivariado. Bragantia 65(2):215-226.

Nardino M, Baretta D, Carvalho IR, Follmann DN, Konflanz VA, Souza VQ, Oliveira AC, Maia LC (2016) Correlações fenotípica, genética e de ambiente entre caracteres de milho híbrido da Região Sul do Brasil. Revista Brasileira de Biometria 34(3):379-394.

Nascimento FM, Bicudo SJ, Rodrigues JGL, Furtado MB, Campos S (2011) Produtividade de genótipos de milho em resposta à época de semeadura. Revista Ceres 58(2):193-201.

Oliboni R, Faria MV, Neumann M, Resende JTV, Battistelli GM, Tegoni RG, Oliboni DF (2013) Análise dialélica na avaliação do potencial de híbridos de milho para a geração de populações-base para obtenção de linhagens. Semina: Ciências Agrárias 34(1):7-17.

Possatto Junior O, Faria MV, Battistelli GM, Rossi ES, Marck DF, Silva CA (2017) Avaliação de linhagens S2 de milho em *topcrosses* com linhagem-elite testadora. Revista Brasileira de Milho e Sorgo 16(2):297-309.

R Development Core Team (2016) R: a language and environment for statistical computing.

Statsoft. Statistica 7.0 (2008). EUA Software. Tucksa: StatSoft.

Silva AG, Francischini R, Martins PDS (2015) Agronomic and economic performances of maize hybrids in off-season cultivation. Revista Agrarian 8(27):1-11.

Silva RA, Souza UO, Santos LGD, Melo NC, Vasconcelos RCD (2016) Características agronômicas de cultivares de milho verde submetidas a doses de Ribumin. Revista de Ciências Agrárias 39(3):395-403.

Samecima Junior EH (2018) Índice de vegetação por diferença normalizada e caracteres agronômicos em genótipos de milho. Unesp-Fcav (Dissertação de Mestrado em Agronomia Genética e Melhoramento de Plantas) – Universidade Estadual Paulista, Jaboticabal.

Seraguzi EF, Lima AR, Anselmo JL, Alvarez RCF (2016) Desempenho de híbridos de milho na região de Chapadão do Sul, MS. Tecnologia & Ciência Agropecuária 10(4):12-14.

Serpa M da Silva, Silva PRF, Sangoi L, Vieira VM, Marchesi DR (2012) Densidade de plantas em híbridos de milho semeados no final do inverno em ambientes irrigados e de sequeiro. Pesquisa Agropecuária Brasileira 47(4):541-549.

Souza Neto ILD, Pinto RJB, Scapim CA, Jobim CC, Figueiredo AST, Bignotto LS (2015) Análise dialélica e depressão endogâmica de híbridos forrageiros de milho para características agronômicas e de qualidade bromatológica. Bragantia 74(1):42-49.

Toebe M, Cargnelutti Filho A (2013) Não normalidade multivariada e multicolinearidade na análise de trilha em milho. Pesquisa Agropecuária Brasileira 48(5):466-477.

Vieira LC, Pedro Guerra M, Barbosa-Neto JF (2017) Análise preliminar de germoplasma de variedades crioulas de milho do Sul do Brasil. Revista Brasileira de Milho e Sorgo 15(3):557-571.