CANONICAL CORRELATIONS BETWEEN CYCLE AND GRAIN PRODUCTION TRAITS IN LIMA BEAN[1]

GUILHERME ALEXANDRE LUZ DA COSTA[2]*, ARTUR MENDES MEDEIROS[3], VERÔNICA BRITO DA SILVA[4],
ÂNGELA CELIS DE ALMEIDA LOPES[4], WILLAME DOS SANTOS CANDIDO[5],
REGINA LUCIA FERREIRA GOMES[4]

ABSTRACT - Lima bean is an important crop in Northeastern Brazil and a source of food and income for farmers in the region. However, there have been few genetic studies on this species, which limits the knowledge available for use in breeding programs. The objective of this study was to estimate the relationship between cycle and production traits using canonical correlation analysis and to identify traits that can be used for the indirect selection of lima bean. The experiment was conducted at the Department of Plant Science of the Universidade Federal do Piauí, Teresina, in a randomized block design with five replications, in which 11 agronomic traits from six lima bean populations in the F_3 generation originating from biparental crosses were evaluated in 2019. The data were subjected to canonical correlation analysis using the virtual environment R. The results showed that only the first coefficient of the canonical pair was significant (r = 0.5531) by the qui-square test, suggesting that the studied groups were not independent, as the cycle traits showed coefficients of large magnitudes in the relationship between groups. The canonical correlation results suggested that there is a linear association between cycle and production traits in lima beans, in which days to flowering, days to maturation, pod length, seed width, and seed thickness contribute the most to the association between groups.

Keywords: *Phaseolus lunatus*. Multivariate analysis. Indirect selection.

CORRELAÇÕES CANÔNICAS ENTRE CARACTERES RELACIONADOS AO CICLO E À PRODUÇÃO DE GRÃOS EM FEIJÃO-FAVA

RESUMO - O feijão-fava é uma cultura importante no Nordeste do Brasil, sendo fonte de alimentação e renda para os agricultores da região. Estudos genéticos com a espécie são escassos, o que limita o conhecimento disponível para ser usado em programas de melhoramento. Assim, objetivou-se estimar a relação entre caracteres relacionados ao ciclo e à produção de grãos, por meio da análise de correlações canônicas, e identificar caracteres que possam ser utilizados na seleção indireta em feijão-fava. O experimento foi instalado no Departamento de Fitotecnia, do Centro de Ciências Agrárias, da Universidade Federal do Piauí, em Teresina - PI, no delineamento de blocos cazualizados, com cinco repetições, no qual foram avaliados 11 caracteres agronômicos em seis populações F_3 de feijão-fava, geradas de cruzamentos bi parentais, no ano de 2019. Os dados obtidos foram submetidos à análise de correlações canônicas, por meio do ambiente virtual R, mostrando que o primeiro coeficiente do par canônico foi significativo (r = 0.5531) pelo teste qui-quadrado, o que indica dependência entre os grupos estudados. Assim, existe associação linear entre os caracteres relacionados ao ciclo e à produção de grãos em feijão-fava, sendo que número de dias até o início da floração, número de dias para maturação, comprimento da vagem, largura da semente e espessura da semente são os que mais contribuem para a associação entre grupos.

Palavras-chave: *Phaseolus lunatus*. Análise multivariada. Seleção indireta.

INTRODUCTION

The Lima Bean *(Phaseolus lunatus* L.), belonging to the *Fabacea* family, is one of the five domesticated species of the *Phaseolus* genus and the second most cultivated species within the genus after the common bean (*P. vulgaris* L.) (PENHA et al., 2016; BITOCCHI et al., 2017). It is grown in tropical and subtropical regions and is distributed in countries such as Mexico, Guatemala, Ecuador, Peru, Colombia, Spain, Nigeria, Indonesia, and Philippines, among others (BRIA; SUHARYANTO; PURNOMO, 2019).

The crop is an important source of proteins and carbohydrates used in human and animal food and is also used as green manure or ground cover (PEGADO et al., 2008; BARREIRO NETO et al., 2017). It presents favorable traits for growth in dry areas and is moderately tolerant to salinity (BARREIRO NETO et al., 2015; ARTEAGA et al., 2018).

In Brazil, the cultivation of the Lima bean presents itself as an alternative source of food and income, mainly in the northeast region of the country, with a national production of 16.6 thousand tons in 2020, with the states of Ceará and Paraíba being the largest producers (IBGE, 2021).

Despite this, research studies on *Phaseolus lunatus* are scarce, which limits the planning and cultivation process of the crop (BRITO et al., 2020). This study is a pioneer in plant breeding studies of the Lima Bean Breeding Program of the Federal University of Piauí, where are being advanced segregating generations of populations developed from biparental crosses between genotypes from Argentina, Brazil, and United States.

Knowledge of phenotypic correlations can help in the process of selection of superior genotypes, as it enables the analysis of multiple traits, which facilitates the selection of the most adequate ideotype (COIMBRA et al., 2000). Canonical correlation analysis is a multivariate analysis method that allows the grouping of traits and the observation of linear multidimensional relationships between sets of traits, maximizing the correlation between groups, so that the observed associations allow the indirect selection of traits (BRUM et al., 2011; CRUZ; REGAZZI; CARNEIRO, 2012).

Canonical correlation analysis has earlier been used to estimate the relationships between groups of agronomic traits in different crops, such as common bean (COIMBRA et al., 1998; COIMBRA et al., 2000), sugarcane (SILVA et al., 2007), papaya (BRUM et al., 2011), wheat (CARVALHO et al., 2015), maize (ALVES et al., 2016), cotton (TEODORO et al., 2018), forage cactus (SILVA et al., 2020), and grapevine (CARGNIN, 2019), among others. Studies using this multivariate approach of canonical correlations related to Lima bean were not found in the existing literature; therefore, to the best of our knowledge, this study is a pioneering work in the use of this methodology for this species.

In this regard, this study aimed to estimate the relationship between traits related to the cycle and grain production through canonical correlation analysis and to identify agronomic traits that can be used in the indirect selection of Lima bean.

MATERIAL AND METHODS

Data were collected from experiments carried out in the Department of Plant Science of the Center for Agricultural Sciences at Universidade Federal do Piauí in Teresina, PI, at 05º05′21″ S, 42º48′07″ W, and 72 m altitude from March to August 2019. The climate in the region is of the Aw type, according to the Köppen classification, the soil used was classified as silty loam, with an average temperature of 28.07° C, average relative humidity of 67.40% and average precipitation of 1,588 mm in the year 2019 (BRASIL, 2021).

Six Lima bean populations in F_3 (Table 1) obtained from the advance of the F_2 generation by the bulk method in the Lima Bean Breeding Program at Universidade Federal do Piauí were evaluated. The experiment was set up in a randomized block design, with five replications, and the part of the plot used for data collection consisted of 10 plants, with a spacing of 0.5 m between plants and 1.0 m between rows. All plants in the plot were evaluated.

Table 1. Identification of the Lima bean (*Phaseolus lunatus* L.) populations obtained from biparental crossings and their respective parents, sourced from the Active Germplasm Bank of *Phaseolus*.

Population	Female Parent	Male Parent
H01	G25236	UC 92
H25	UFPI 628	G25236
H46	UFPI 728	UC HASKELL
H50	UFPI 628	UFPI 728
H81	UC 92	UFPI 628
H86	UFPI 728	UFPI 628
H94	UC 92	UFPI 728

Populations were obtained from biparental crosses between parents from different countries. Accession G25236, originating from Argentina, was made available by the International Center for Tropical Agriculture (CIAT), in Cali, Colombia; accessions UC 92 and UC HASKELL, from the United States, were made available by the University of California (UC) in Davis, CA, United States; UFPI 628 and UFPI 728, from Brazil, were made available by the Federal University of Piauí (UFPI), Campos Ministro Petrônio Portela, in Teresina, PI. The H50 and H86 populations were grouped into a single population as they originated from a reciprocal cross.

Cultural treatments and fertilization were performed according to soil analysis and recommendations for irrigated systems (LOPES; GOMES; ARAÚJO, 2010). Pests and disease control were preventively performed using pesticides.

The traits were evaluated according to the following descriptors for *Phaseolus lunatus* L. (BIOVERSITY INTERNATIONAL, 2007): number of days to the beginning of flowering (NDF), number of days to maturity (NDM), number of pods per plant (NPP), pod length (PL, mm), pod width (PW, mm), pod thickness (PT, mm), number of locules per pod (NLP), number of seeds per pod (NSP), seed length (SL, mm), seed width (SW, mm), and seed thickness (ST, mm).

The traits used in the canonical correlation analysis were divided into two groups: cycle traits, NDF and NDM (group 1), and production traits, NPP, PL, PW, PT, NLP, NSP, SL, SW, and ST (group 2).

The phenotypic correlation matrices were submitted to multicollinearity diagnosis based on condition number (CN) and the variance inflation factor (VIF) using the metan v1.15 package (OLIVOTO; LÚCIO, 2020) in the R statistical program (R CORE TEAM, 2021).

Canonical correlation analysis was performed according to the methodology proposed by Cruz, Regazzi, and Carneiro (2012) described below, with two groups of traits, X and Y, defined as follows: $X'=[x_1 x_2 \dots x_p]$ (1) is the vector of observations of p traits constituting group 1, and $Y'=[y_1 y_2 \dots y_q]$ (2) is the vector of observations of q traits constituting group 2. If each X_1 and Y_1, represent one of the linear combinations of the traits from groups 1 and 2, it follows that $X_1=a_1 x_1+a_2 x_2+\dots+a_p x_p$ (3) and $Y_1=b_1 y_1+b_2 y_2+\dots+b_q y_q$ (4), in which $a'=[a_1 a_2 \dots a_p]$ (5) is the vector $1 \times p$ of the trait weights of group 1, and $b'=[b_1 b_2 \dots b_q]$ (6) is the vector $1 \times q$ of the trait weights of group 2.

In this regard, the first canonical correlation was defined as the one that maximized the relation between X_1 and Y_1. Thus, the functions X_1 and Y_1 are the first pair associated with the canonical correlation, according to the expression below (7):

$$r=\frac{C\hat{o}v(X_1,Y_1)}{\widehat{V}(X_1).\widehat{V}(Y_1)}$$

In which,

$$C\hat{o}v(X_1,Y_1)=a'S_{12}b \qquad (8)$$

$$\widehat{V}(X_1)=a'S_{11}a \qquad (9)$$

$$\widehat{V}(Y_1)=b'S_{22}b \qquad (10)$$

In this case, S_{11} is the $p \times p$ covariance matrix between the traits from group 1, S_{22} is the $q \times q$ covariance matrix between the traits from group 2, and S_{12} is the $p \times q$ covariance matrix between the traits from groups 1 and 2. Therefore, the first canonical correlation (r^1) between the linear combinations of the traits from groups 1 and 2 is $r=(\lambda l)^{0.5}$ (11), where λ_1 is the biggest eigenvalue from the matrix $R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}$, which is squared and presents an asymmetric order p. The first canonical factor is given by $X_1=a'X$ (12) and $Y_1=b'Y$ (13), in which a is the eigenvector associated with the first eigenvalue $R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}$ and eigenvector b is associated with the first eigenvalue $R_{22}^{-1}R_{21}R_{11}^{-1}R_{12}$. The other correlations and canonical factors are estimated using the eigenvalues and eigenvectors corresponding to the estimation of the correlation order.

The significance of the canonical correlations estimated between the groups 1 and 2 was evaluated by the qui-square test at 1% probability, with the analyses being carried out using the candisc package (FRIENDLY; FOX, 2021) in the R program.

RESULTS AND DISCUSSION

Phenotypic correlations varied in magnitude from −0.38 (SL with NDF and NSP) to 0.90 (SL and SW) (Figure 1) and were significant between traits related to cycle (0.51 between NDF and NDM), between production traits such as NPP with NLP, NSP, ST, and PL (0.32, 0.43, 0.24, and −0.17, respectively), and between traits from the different groups. Knowledge of the correlations and their magnitudes can assist in indirect selection; however, it is important to highlight that phenotypic correlations include genetic and environmental values, as only the genetic values are passed on to the next generation (SILVA et al., 2007).
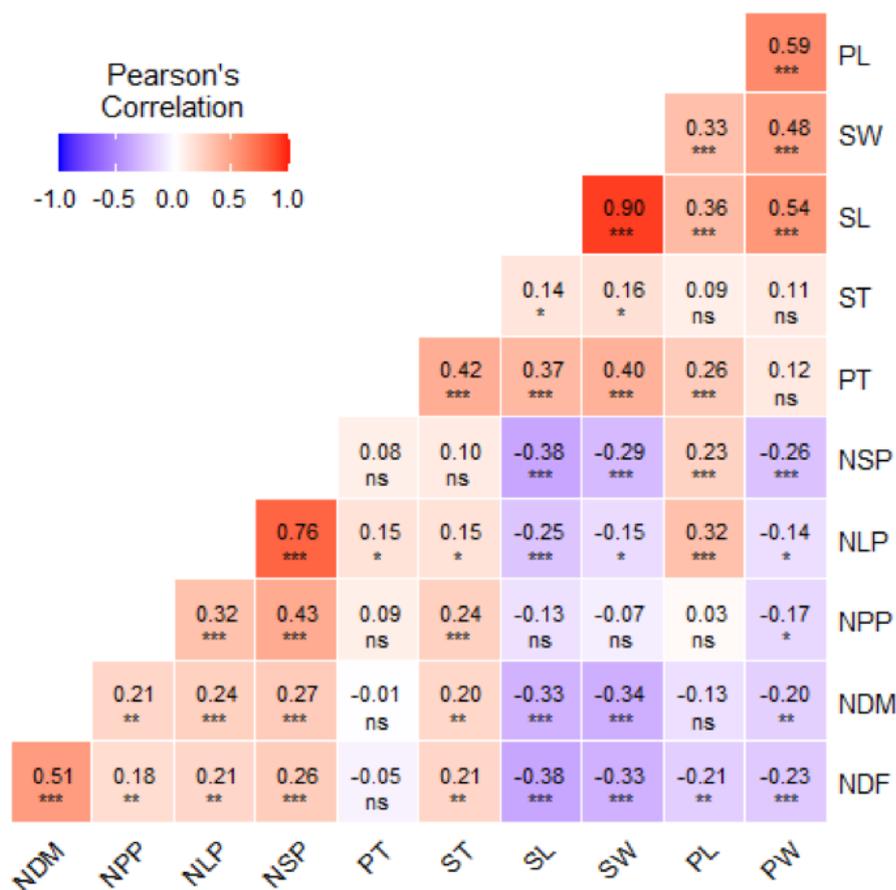
Figure 1. Estimates of Pearson's phenotypic correlation between 11 traits evaluated in six Lima bean populations obtained from biparental crosses in the Lima Bean Breeding. NDF: number of days to the beginning of flowering, NDM: number of days to maturity, NPP: number of pods per plant, PL: pod length, PW: pod width, PT: pod thickness, NLP: number of locules per pod, NSP: number of seeds per pod, SL: seed length, SW: seed width, ST: seed thickness. *significant at 5%, ** significant at 1%, and *** significant at 0.1%, ns – not significant.

The positive and significant phenotypic correlation value between the number of days to the beginning of flowering and the number of days to maturity (0.51) found in the lima bean populations was larger than that observed by Corrêa et al. (2015) in cowpea (0.44), implying a stronger association. These traits correlated positively with the number of seeds per pod and negatively with seed length and width, suggesting that plants with longer cycles tend to produce a larger number of smaller-sized seeds.

The pod length was positively and significantly correlated with pod width (0.59) and pod thickness (0.26), and these traits were positively correlated with seed length (0.36, 0.54, and 0.37, respectively) and seed width (0.33, 0.48, and 0.40, respectively). This shows that plants with longer, larger, and thicker pods tend to produce larger seeds, which is an important finding for the improvement of the species.

The correlation coefficients were positive and significant between the number of pods per plant and number of locules per pod (0.32), and between the number of pods per plant and number of seeds per pod (0.43), which was in contrast with those

obtained by Assunção Filho et al. (2022), who did not find significant correlations between these traits. For pod length, positive and significant correlations were estimated with the number of locules per pod (0.32) and number of seeds per pod (0.23), which was in line with the findings of Assunção Filho et al. (2022) that there were significant correlations between the same traits (0.43 and 0.37, respectively).

The number of locules per pod and number of seeds per pod, important grain production components, presented a positive, significant, and high magnitude correlation (0.76), indicating that pods with more locules present a larger quantity of seeds. The number of seeds per pod correlated negatively with pod width (−0.26), seed length (−0.38), and seed width (−0.29), suggesting a tendency for greater production of seeds in seeds of smaller size.

The evaluated traits presented the value for condition number (CN) of 42.7 and values of inflation of variance (VIF) of 1.71 to NDM, 1.52 to NDM, 1.53 to NPP, 2.53 to PL, 2.38 to PW, 1.59 to PT, 2.69 to NLP, 3.14 to NSP, 6.96 to SL, 6.01 to SW and 1.46 to ST, indicating weak

multicollinearity (CRUZ; REGAZZI; CARNEIRO, 2012; MONTGOMERY; PECK; VINIGIN, 2012) indicating that none of the traits needed to be excluded from the analysis.

The canonical correlation analysis allows to test if linear dependency exists between the groups analyzed (ALVES et al., 2016). In the present study, the first coefficient of the canonical pair (r = 0.5531) was significant by the qui-square test (Table 2), suggesting association between the evaluated groups.

Table 2. Canonical correlation coefficients and estimated canonical pairs between traits related to the cycle (group 1) and grain production (group 2) evaluated in six F_3 lima bean populations derived from biparental crosses in the Lima Bean Breeding Program.

Traits	Canonical Pairs	
	1°	2°
	Group 1	
NDF (days)	−0.6139	0.9917
NDM (days)	−0.5342	−1.0367
	Group 2	
NPP	−0.1118	−0.3501
PL (mm)	0.4164	−0.2442
PW (mm)	−0.1760	0.0508
PT (mm)	−0.0220	−0.3064
NLP	−0.2519	−0.5119
NSP	−0.1929	0.3762
SL (mm)	0.2258	−1.8854
SW (mm)	0.4215	1.9977
ST (mm)	−0.4575	0.2043
r	0.5531***	0.1763^{ns}
Degrees of Freedom	18	8
X^2	8.13×10^{-10}	0.6047

NDF: number of days to the beginning of flowering; NDM: number of days to maturity; NPP: number of pods per plant; PL: pod length; PW: pod width; PT: pod thickness; NLP: number of locules per pod; NSP: number of seeds per pod; SL: seed length; SW: seed width; ST: seed thickness. ***significant at 0.01 of probability by the qui-square test (X^2), ns – not significant.

Considering the first canonical pair, the traits that contributed the most to the association related to the cycle and grain production in the populations of lima bean studied were NDF, NDM, PL, NLP, SL, SW and ST, particularly NDF and NDM. In studies carried out by Coimbra et al. (1998) and Coimbra et al. (2000) with common bean, positive and significative canonical correlations were found between primary and secondary traits related to yield.

The first canonical pair associates plants with shorter days to flowering and maturity with longer pods, and longer and larger seeds. Thus, a longer cycle causes a decrease in pod length, seed length, and seed width. Therefore, selection should be carried out in a way that decreases the number days in the cycle (NDF and NDM), causing an increase in the production traits relevant to the market, such as those related to pods and seeds. In addition, larger pods are desirable in the breeding of the species because they help in the harvest, which is carried out manually by farmers (SILVA; NEVES, 2011).

The NDF, NDM, and ST were the traits that were correlated in the same manner. Thus, an increase in the number of days to the beginning of flowering or in the number of days to maturation would cause an increase in seed thickness, which indicates that these traits can be used in the indirect selection of plants with traits that are desirable for the consumer.

The number of pods per plant, pod width, pod thickness, and number of seeds per pod presented low-magnitude values in the association between the groups. Thus, an increase in NDF and NDM would result in a low degree of change in these traits. In relation to NPP, a decrease in the trait is not desirable because pod production is an important trait that should be considered in the selection.

CONCLUSION

Linear dependency exists between the traits related to the cycle and grain production in lima bean. The number of days to the beginning of flowering, the number of days to maturity, pod length, seed width, and seed thickness are the traits that contributed the most to the association between the cycle and grain production in lima bean.

In the studied populations, a decrease in the number of days to the beginning of flowering and the number of days to maturity resulted in longer pods and longer and wider seeds. The number of days to the beginning of flowering and the number of days to maturity are the traits that can be used in the indirect selection of grain yield components in lima beans.

ACKNOWLEDGMENTS

REFERENCES

ALVES, B. M. et al. Correlações Canônicas entre Caracteres Agronômicos e Nutricionais Proteicos e Energéticos em Genótipos de Milho. **Revista Brasileira de Milho e Sorgo**, 15: 171-185, 2016.

ARTEAGA, S. et al. Screening for Salt Tolerance in Four Local Varieties of *Phaseolus lunatus* from Spain. **Agriculture**, 8: 201-214, 2018.

ASSUNÇÃO FILHO, J. R. et al. Selection of superior genotypes of lima bean landraces by multivariate approach. **Revista Caatinga**, 35: 87-95, 2022.

BARREIRO NETO, M. et al. Características morfológicas e produtivas em acessos de feijão-fava consorciados. **Tecnologia & Ciência Agropecuária**, 9: 23-27, 2015.

BARREIRO NETO, M. et al. Valoração de custos e rentabilidade econômica de sistemas de produção de feijão fava de crescimento determinado na Mata Paraibana. **Tecnologia & Ciência Agropecuária**, 11: 75-83, 2017.

BITOCCHI, E. et al. Beans (*Phaseolus* ssp.) as a model for understanding crop evolution. **Frontiers in Plant Science**, 8: 1-21, 2017.

BIOVERSITY INTERNATIONAL. **Guidelines for the development of crop descriptor lists**. Bioversity Technical Bulletin Series. Rome: Bioversity International. 2007. 72 p.

BRASIL. INMET - Instituto Nacional de Meteorologia. Ministério da Agricultura, Pecuária e Abastecimento. **Gráficos Anuais de Estações Automáticas**. Disponível em: <https://tempo.inmet.gov.br>. Acesso em: 25 jun. 2021.

BRIA, E. J.; SUHARYANTO, E.; PURNOMO. Variability and Intra-Specific Classification of Lima Bean (*Phaseolus lunatus* L.) from Timor Island based on Morphological Characters. **Journal of Tropical Biodiversity and Biotechnology**, 4: 62-71, 2019.

BRITO, M. V. et al. Univariate and multivariate approaches in the characterization of lima bean genotypes. **Revista Caatinga**, 33: 571-578, 2020.

BRUM, B. et al. Correlações canônicas entre variáveis de semente, plântula, planta e produção de grãos em mamoneira. **Ciência Rural**, 41: 404-411, 2011.

CARGNIN, A. Canonical correlations among grapevine agronomic and processing characteristics. **Acta Scientiarum. Agronomy**, 41: 1-6, 2019.

CARVALHO, I. R. et al. Correlações canônicas entre caracteres morfológicos e componentes de produção em trigo de duplo propósito. **Pesquisa Agropecuária Brasileira**, 50: 690-697, 2015.

COIMBRA, J. L. M. et al. Coeficiente de trilha, correlações canônicas e divergência genética: II. Entre caracteres primários e fitossanitários em genótipos de feijão preto. **Pesquisa Agropecuária Gaúcha**, 4: 195-201, 1998.

COIMBRA, J. L. M. et al. Correlações Canônicas: II - Análise do rendimento de grãos de feijão e seus componentes. **Ciência Rural**, 30: 31-35, 2000.

CORREA, A. M. et al. Variabilidade genética e correlações entre caracteres de feijão-caupi. **Revista Agro@mbiente On-line**, 9: 42-47, 2015.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**, 3. ed. Viçosa, MG: UFV, 2012, 480 p.

FRIENDLY, M.; FOX, J. **Candisc**: visualizing generalized canonical discriminant and canonical correlation analysis. Disponível em: <https://CRAN.R-project.org/package=candisc>. Acesso em: 29 jul. 2021.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Produção Agrícola Municipal**. 2021. Disponível em: <https://sidra.ibge.gov.br/tabela/1612>. Acesso em: 10 jan. 2022.

LOPES, A. C. A.; GOMES, R. L. F. ARAÚJO, A. S. F. **A cultura do feijão-fava na Região Meio-Norte do Brasil**. 1. ed. Teresina, PI: EDUFPI, 2010. 272 p.

MONTGOMERY, D. C.; PECK, E. A.; VINIGIN, G. G. **Introduction to linear regression analysis**. 5 ed. New York: John Wiley & Sons, 2012. 642 p.

OLIVOTO, T.; LÚCIO, A. D. Metan: an r package for multi :environment trial analysis. **Methods in Ecology and Evolution**, 11: 783-789, 2020.

PEGADO, C. M. A. et al. Decomposição superficial e subsuperficial de folhas de fava (*Phaseolus lunatus* L.) na região do Brejo da Paraíba, Brasil. **Revista Caatinga**, 21:218-223, 2008.

PENHA, J. S. et al. Estimation of natural outcrossing rate and genetic diversity in Lima bean (*Phaseolus lunatus* L. var. *lunatus*) from Brazil using SSR markers: implications for conservation and breeding. **Genetic Resources and Crop Evolution**, 64: 1355-1364, 2016.

R CORE TEAM. **R: A language and environment for statistical computing**. Disponível em: <https://www.R- project.org>. Acesso em: 29 jul. 2021.

SILVA, A. S. et al. Análise multivariada da palma forrageira: características morfoprodutivas sob correlações canônicas. **Agrarian**, 13: 100-106, 2020.

SILVA, J. A. L.; NEVES, J. A. Produção de feijão-caupi semi-prostrado em cultivos de sequeiro e irrigado. **Revista Brasileira de Ciências Agrárias**, 6: 29-36, 2011.

SILVA, J. W. et al. Correlações canônicas de características agroindustriais em cana-de- açúcar. **Acta Scientiarum. Agronomy**, 29: 345-349, 2007.

TEODORO, P. E. et al. Interrelations between agronomic and technological fiber traits in upland cotton. **Acta Scientiarum. Agronomy**, 40: 1-7, 2018.