

Potential of grain physical traits to the study of variability in maize

ABSTRACT

Aims: The identification of new traits that can be used as phenotypic markers as well as potential indirect selection tools is interesting for plant breeding. The objective of the study was to investigate the potential use of physical traits of grains to study phenotypic variability in maize.

Methodology: Ten maize varieties were evaluated, one commercial variety and nine local open-pollinated varieties. After harvesting, the following traits were evaluated: mass of 1000 grains, weight of ears, grain yield per hectare, estimated from plot production; real volume and apparent volume; real density and apparent density; sphericity and volumetric weight; obtained in samples of 50 grains of each of the studied varieties. A principal component analysis was performed on the data. First, those traits that showed a strong correlation with components capable of explaining a smaller portion of the total variation were excluded from the analysis. A new principal component analysis was performed with the remaining traits.

Results: The result revealed the possibility of excluding five of the ten analyzed variables: mass of 1000 grains, weight of ears, apparent volume, real density, and volumetric weight. Some possibilities of trait exclusion can be explained by correlation and method of estimate among them. The genotypes G2 and G3 showed great difference, mainly due the grain yield. The porosity and real volume contributed to the majority variation among genotypes.

Conclusion: Some physical grain traits showed potential for use in divergence studies. Apparent density can be used in indirect selection strategies for higher grain yield.

Keywords: maize breeding; phenotypic markers; indirect selection; principal component analysis.

1. INTRODUCTION

World grain production surpassed 1.5 billion tons in 2020 [1]. The constant population growth generates an increase in the demand for food, a factor that, added to climate change, has pointed to imminent risks to food security [2]. The need to adapt plants to grow in adverse conditions puts pressure on countries to invest in technologies to increase productivity [3].

In this scenario, plant breeding has contributed to increased productivity by developing plants that are more adapted to extreme climatic conditions, more efficient in the use of nutrients, and more resistant to diseases [4][5]. The development of new cultivars begins with the characterization of the germplasm to understand the potential use of each accession, as well as the genetic variation available for use in breeding programs [6].

For a precise characterization of the available accessions, the main challenge for plant breeders is to identify the genetic value of each plant and thus understand how much of the observed variations are due to genetic causes and how much of this variation is the result of environmental effects [7]. To more accurately access a greater volume of genetic information, breeders have used tools ranging from phenotype evaluation to the use of predictions based on molecular information and the estimation of epigenetic effects [8][9].

Despite the existence of increasingly sophisticated tools, plant breeders in countries with low investment in research have limitations in acquiring or maintaining equipment capable of generating genomic information. Still, even the use of more modern biotechnological tools in plant breeding requires accurate and abundant phenotypic evaluations to enhance their efficiency [8][10].

Thus, the identification of traits that can be used as phenotypic markers capable of accessing the variability of a larger set of genes is important to enhance studies of genetic divergence between accessions in different crops [11][12][13]. In addition to contributing to a better understanding of genetic variation between individuals, some secondary traits can also help genetic gains in traits of interest through indirect selection [14][15]. The objective of this study was to investigate the potential use of physical traits of grains to study phenotypic variability in maize varieties.

2. MATERIAL AND METHODS

We evaluated 10 maize varieties, one of which is a commercial cultivar and nine are open-pollinated varieties grown by local farmers in the southern region of Espírito Santo, Brazil. For grain production, a field experiment was installed in Jerônimo Monteiro, Espírito Santo province (20°49'48" S and 41°23'10" W) during the 2019-2020 harvest (September to January).

The experiment was conducted in randomized blocks with four replications. The experimental unit consisted of three rows 5 m in length with 20 plants by row. The spacing used was 0.25 m between plants and 0.80 m between rows. The plants received irrigation for 1 hour a day on alternate days. For fertilization, 80 kg ha^{-1} of phosphorus and 150 kg ha^{-1} of nitrogen was applied. Nitrogen fertilization was divided into two applications 30 and 45 days after planting, respectively.

After 110 days after sowing, all ears of the plots were harvested and threshed. The agronomic traits related to production were also evaluated: mass of 1000 healthy grains (W1000) with 13% moisture; ear weight (EW) in grams (g), and; grain yield estimated for 1 hectare (Yield) in kilograms (kg). The inclusion of these traits in this study aimed to infer their correlations with the physical traits of grains to ensure that these new traits are capable of generating new information. In other words, conventional production traits were used as a test to verify the potential of physical traits to explain variance among the genotypes.

From the measurements of 50 grains of each variety, the following were obtained: the apparent volume (A-vol) and real volume (R-vol) of the grains (mm^3); apparent density (A-density) and real density (R-density) of grains (kg m^{-3}); porosity (%) of samples (Porosity); the volumetric weight (VW) obtained from the surface by volume ratio and; the sphericity (Sph) of the grains (%). To obtain these traits, firstly were measured the length (L), the maximum diameter (Dmax) and the minimum diameter (Dmin) of the grains. The mass of 50

grains at 13% of moisture was measured in a balance with precision of 0.0001 grams. Then, the samples were dried in a forced ventilation oven at 60°C until reaching constant mass.

The A-vol was obtained from grains at 13% of moisture, after harvest, and estimate by differences from deposition of sample in a graduated cylinder with 0.1 mL of precision. To obtain the R-vol, the samples dried were added in the graduated cylinder now with the vegetal oil at 100mL. So, the sample of grains was added and, the new volume was observed.

The A-density of i samples was estimate by:

$$A_density_i = \frac{M(13\%)_i}{A-vol_i} / N_i$$

Where $M(13\%)$ is the mass of the sample i after harvest at 13% of moisture; and N is the number of grains in the " i " sample.

The R-density of i samples was estimate by:

$$A_density_i = \frac{M_i}{R-vol_i} / N_i$$

Where M is the mass of sample i after dry at constant mass; and N is the number of grains in the " i " sample.

The porosity of samples was obtained by relation among A-Vol and R-Vol not in grain unit but in the sample:

$$Porosity_i(\%) = 100 \cdot \left(1 - \left(\frac{A - vol}{R - vol}\right)\right)$$

To obtain the sphericity of the grains we use the means of sphericity of the grains estimate by equation suggested by [29]:

$$Sphericity_i(\%) = \left[\frac{(L \cdot Dmax \cdot Dmin)^{\frac{1}{3}}}{L} \right] \cdot 100$$

In the statistical proceeds, the principal component analysis was conducted for the ten characteristics evaluated [16]. From the eigenvalues of the components, the number of components that presented variance less than 0.7 were identified. Thus, the number of characteristics subject to deletion was defined as the number of components with a variance less than 0.7 (eigenvalues). The Pearson's correlation among traits and components were performed. The choice of variables to be excluded was defined based on the correlation values between variables and components. Those traits that showed stronger correlations with the less representative components (variance < 0.7) were excluded. This procedure ensures that traits that explain little of the total variance are removed from the analysis [17].

So, from the Yield, Porosity, Sph, A-Density and R-vol data, the standard means were used to proceed a new PCA - principal component analysis. The data was standard from mean/standard deviation of each trait. The total variance explained by each component were

obtained by length of each component (λ) and the eight values (%) were plotted in a barplot. The PCA analysis was performed in the R software on the functions of 'factoextra' package [18] and biplots were performed on the functions of the 'ggplo2' [19].

3. RESULTS AND DISCUSSION

The physical grain and yield characteristics of nine open-pollinated maize varieties and one commercial cultivar were evaluated for potential use in genetic divergence studies. The results reveal that of the ten components generated, the first five components, jointly, explained 94.36% of the total variation (Figure 1).

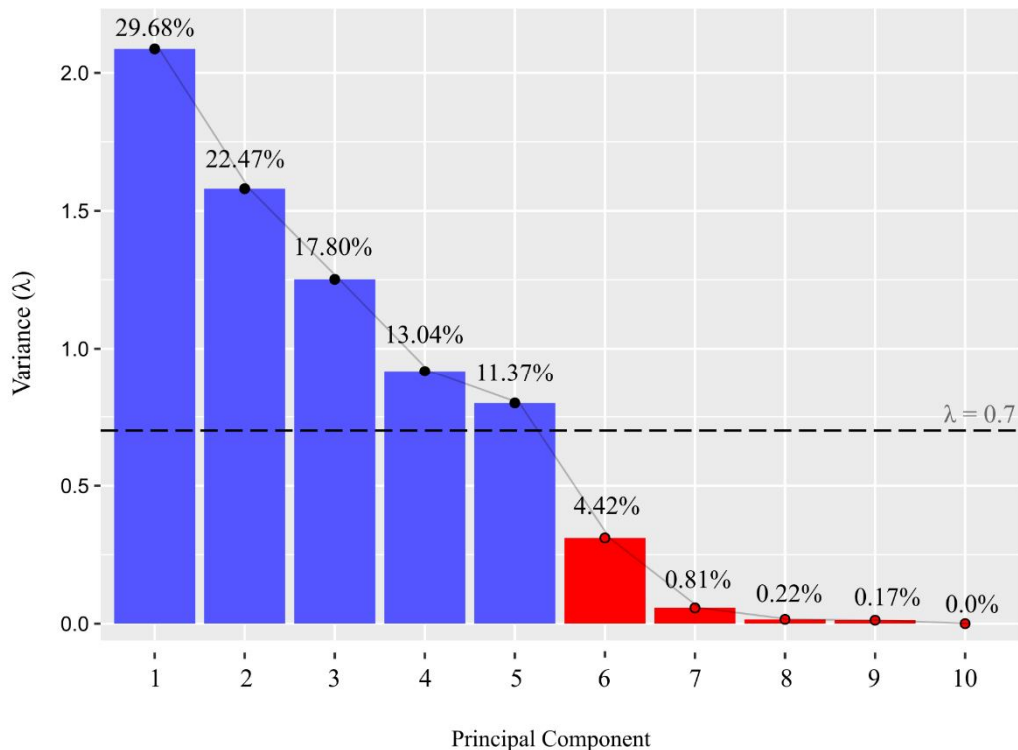


Figure 1. Total variance explained by each principal component for yield traits and physical aspects of maize grains.

Components 6, 7, 8, 9, and 10 accumulated variances of less than 0.7 which indicated that the contribution of these components to explain the variability between genotypes is very low. This value is defined by [20] whom he showed that this criterion guarantees a satisfactory result with the exclusion of highly correlated variables. This result also indicated that of the ten response variables evaluated, five can be discarded for variability studies among varieties.

The correlation of the variables with each component showed that the apparent volume (A-vol) and the volumetric weight of the grains (VW) presented the highest correlations with components 10 and 9, respectively (Table 1). These components together explained less than 0.2% of the variance between varieties.

Table 1. Correlation of agronomic and physical traits of grains and yield with the principal components generated from the evaluation of ten different maize varieties.

Traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
EW ^{5*}	-0.17	-0.07	0.682	-0.26	-0.26	-0.60	-0.04	0.02	-0.01	-0.01
Yield	-0.22	-0.02	0.55	-0.10	0.66	0.44	0.06	-0.02	-0.01	0.00
A-vol ¹	-0.39	0.33	-0.16	0.08	0.14	-0.22	0.35	0.17	-0.35	0.60
R-vol	-0.39	0.30	-0.12	0.23	0.14	-0.25	0.33	-0.37	0.25	-0.55
W1000 ⁴	0.01	0.62	0.01	0.10	0.19	-0.14	-0.69	0.24	0.10	-0.04
A-density	0.43	0.24	0.14	0.02	0.04	-0.04	0.51	0.58	0.35	-0.09
R-density ³	0.42	0.30	0.10	-0.11	0.04	-0.04	0.08	-0.64	0.30	0.44
Porosity	0.12	-0.25	0.25	0.90	0.04	-0.13	-0.07	-0.03	0.03	0.16
VW ²	0.43	0.24	0.14	0.05	0.06	-0.05	0.12	-0.16	-0.76	-0.32
Sph	0.22	-0.37	-0.29	-0.16	0.64	-0.54	-0.06	0.03	0.02	-0.01

**Numbers indicate the order of exclusion of variables due to their correlation with components that are not very explanatory (in bold).*

Also, the traits real grain density (R-density), thousand-grain weight (W1000), and ear weight (EW) showed higher correlations with components 8, 7, and 6, respectively. Due to the greater correlation of these traits with components that do not effectively contribute to the distinction of genotypes, they were excluded from the analysis and considered inefficient for the study of variability between accessions.

The exclusion of the EW and W1000 variables was expected since these characteristics tend to have a high correlation with yield[21]. This correlation is because productivity represents an estimate calculated from the weight of grains. The yield of greater correlation with the first components is since features with greater variation are the most correlated with the first components, which explains most of the variation. The greater yield variation between EW and W1000 occurs because the yield is the value of the plot's grain weight extrapolated to one hectare, which further increases the deviations and differences between treatments [22].

The exclusion of the variables A-vol, VW, and R-density may be related to their greater correlation with R-vol and A-density. However, as we did not find studies on the physical characteristics of maize grains, future studies may contribute to elucidating the potential of these variables to explain the variation and also the degrees of correlation between these and other variables of the same nature. However, the possibility of excluding some variables, as pointed out in this study, is important for optimizing time and accuracy in the analysis of divergence between genotypes.

After excluding the variables, the new principal component analysis revealed that the first two components explained 64.1% of the variation between genotypes for the five traits of interest (Figure 2).

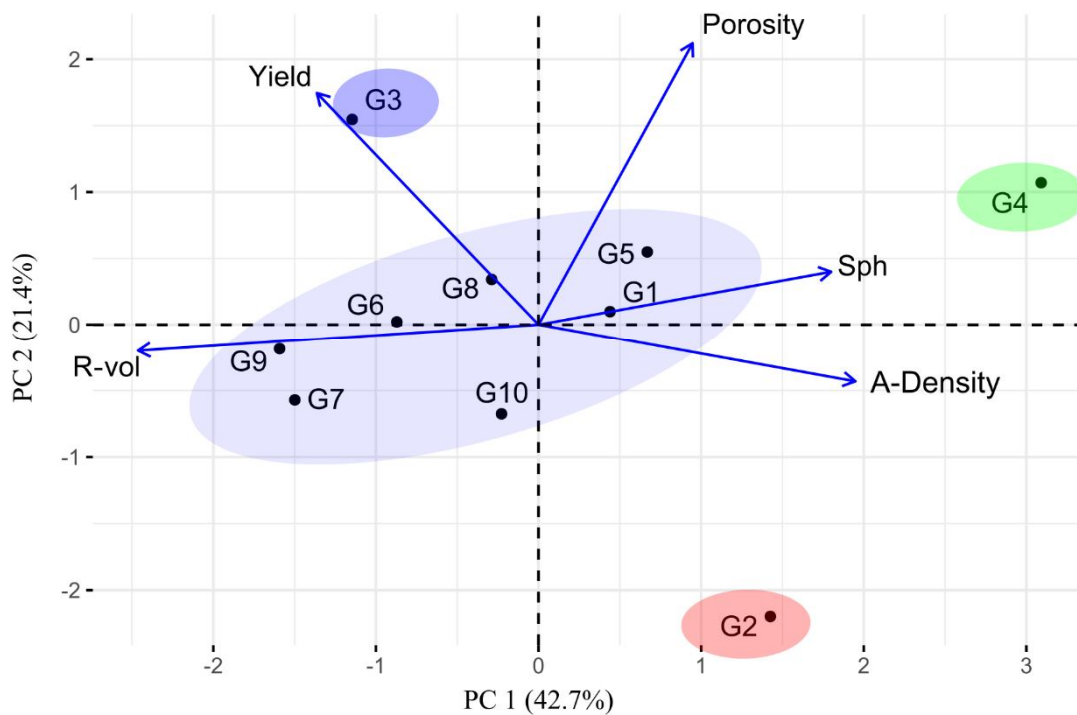


Figure 2. Biplot for the first two principal components generated from four physical traits of grains and grain yield among ten maize genotypes.

The analysis revealed the existence of negative correlations (angles greater than 90° between vectors) between A-Density and Yield and R-vol e characteristics; between R-vol and Porosity and Sph characteristics. Strong correlations, whether negative or positive, can help plant breeders in the indirect selection of traits [23] [24]. Our results show that negative selection for A-density can result in yield gains. Yield is a common feature in maize evaluations, and in these evaluations, the loss of part of the plots due to environmental factors not controlled during experiments is a recurrent event. In these cases, the plot's grain production is impaired, which can lead to inaccuracies in the grain yield estimate. So, our results indicate that indirect selection via A-density can be used in those plots where there was a minimum grain production, which can be evaluated in terms of their physical aspects.

The greatest variations between genotypes occurred for R-vol and Porosity, respectively (evident by the greater length of the vectors). Conversely, Sph was the trait with the least variation between genotypes. From the biological aspect, these correlations indicated that bulky grains tend to be less dense, that is, with less accumulation of dry matter, which directly impacts the reduction of productivity per hectare. The lower density of larger grains may be related to the inefficiency of plants to produce photoassimilates to fill the grains. Correlations of this nature are also observed in popcorn, for example, where the highest yield tends to be negatively correlated with the amount of dry matter in the kernels, producing popcorn of smaller volume after the expansion of the kernels [25] [26] [27].

For the performance of the genotypes, the commercial cultivar (G3) was the one that presented the highest productivity and, consequently, the lowest apparent density. This revealed that despite the efforts of local farmers to improve their varieties, commercial cultivars still maintain higher grain yields. Genotype 2 (G2) showed the lowest apparent

productivity and highest density. This genotype consists of a variety of white pericarp widely used by local farmers for the consumption of cooked ears, the so-called green corn, and also for the manufacture of flour [28]. Probably, the higher starch content of this genotype was responsible for ensuring lower porosity and consequently higher grain density. The G4 genotype, on the other hand, refers to a variety developed by local extension agencies, recommended for silage, due to the harder appearance of the grains. The more rounded and smaller grains of this genotype are typical of Flint-type varieties.

In an overview, the analysis showed that the physical traits of grains can be informative in the process of distinguishing maize genotypes. Although this study was conducted with only ten genotypes, it is important to highlight that nine of these are open-pollinated varieties. This indicates that the genetic variability sampled may have a degree of representativeness that makes the results important for the crop in general. Another important aspect is that the correlations also have a share of genetic effects, which is particular to each plant species and may be the object of attention by breeders of the respective species.

Although the PCA is a usual analysis in plant breeding, it use have been majority to infer genotype divergences and trait correlations [29][30]. However, there are many other functions to be explored from the PCA, as shown in this research. Inferences about the potential use of the traits can help breeders to optimize time and consequently reduce costs in breeding programs.

It is expected that more studies can be carried out with this type of trait so that maize breeders can access the existing variability between their accessions in an increasingly efficient way and thus make the best decisions for the genetic progress of the crop.

4. CONCLUSION

The physical traits: of real grain volume, porosity, apparent density, and sphericity can be used to enrich studies of phenotypic divergence between common maize genotypes in the characterization stages.

Corn breeders should pay attention to the trend that visibly larger grains have a lower density, which can lead to lower grain yield. Apparent grain density can be used for the indirect selection of more productive genotypes due to the negative correlation between these traits.

REFERENCES

1. FAO – Food and Agriculture Organization of United Nations. Food and Agriculture Data. Accessed 20 September 2022. Available: <https://www.fao.org/faostat/en/#data/QI>.
2. Leisner CP. Review: Climate change impacts and food security- focus on perennial cropping systems and nutritional value. *Plant Science*. 2020; 293: 110412.
3. Cole MB, Augustin MA, Robertson MJ, Manners JM. The science of food security. *npj Science of Food*. 2018; 2(14).
4. Lenaerts B, Collard BCY, Demont M. Review: Improving global food security through accelerated plant breeding. *Plant Science*. 2018; 287: 110207.

5. Qaim M. Role of new plant breeding technologies for food security and sustainable agricultural development. *Applied Economic Perspectives and Policy*. 2020; 42(2): 129-150.
6. Prasanha BM. Diversity in global maize germplasm: characterization and utilization. *Journal of Biosciences*. 2012; 37: 843-855.
7. Dudley JW, Moll RH. Interpretation and use of estimates of heritability and genetic variances in plant breeding. *Crop Science*. 1969; 9(3): 257-262.
8. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, Campos G, Burgueño J, González-Camacho J, Pérez-Elizalde S, Beyene Y, Dreisigacker S, Singh R, Zhang X, Gowda M, Roorkiwal M, Rutkoski J, Varshney RK. Genomic selection in plant breeding: Methods, Models and Perspectives. *Trends in Plant Science*. 2017; 22(11): 961-975.
9. Sobral M, Sampedro L. Phenotypic, epigenetic, and fitness diversity within plant genotypes. *Trends in Plant Science*. 2022; 27(9): 843-846.
10. Cursi DE, Hoffmann HP, Barbosa GVS, Bressiani JA, Gazaffi R, Chapola RG, Fernandes Junior AR, Balsalobre TWA, Diniz CA, Santos JM, Carneiro MS. History and current status of sugarcane breeding, germplasm development and molecular genetics in Brazil. *Sugar Tech*. 2022; 24(1): 112-133.
11. Ye H, Roorkiwal M, Valliyodan B, Zhou L, Chen P, Varshney RK, Nguyen HT. Genetic diversity of root system architecture in response to drought stress in grain legumes. *Journal of Experimental Botany*. 2018; 69(13): 3267-3277.
12. Mathew I, Shimelis H. Genetic analyses of root traits: implications for environmental adaptation and new variety development: A review. *Plant Breeding*. 2022; 141: 695-718.
13. Silva LOE, Schmidt R, Almeida RN, Feitoza RBB, Cunha M, Partelli FL. Morpho-agronomic and leaf anatomical traits in *Coffea canephora* genotypes. *Ciencia Rural*. 2023; 53(7): e20220005.
14. Ramchander S, Ushakumari R, Pillai MA. Association analysis of yield and yield contributing traits in Semi-Dwart and early mutants of rice. *Trends in Biosciences*. 2014; 7(11): 1147-1150.
15. Rigatti A, Pelegrin AJ, Meier C, Lunkes A, Klein L, Silva AF, Bellé EP, Silva ADB, Marchioro VS, Souza VQ. Combination capacity and association among traits of grain yield in wheat (*Triticum aestivum* L.): A review. *Journal of Agricultural Science*. 2018; 10(5): 179-187.
16. Abdi H, Williams, LJ. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4): 433-459. <https://doi.org/10.1002/wics.101>
17. Regazzi AJ. Análise multivariada: notas de aula. Viçosa: UFV, 2002 in Meira CT, Pereira IG, Farah MM, Pires AV, Garcia DA, Cruz VAR. Identification of morphofunctional traits in MangalarcaMarchador horse using principal component analysis. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*. 2013; 65(6): 1843-1848.

18. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2020; R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>

19. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

20. Jolliffe IT. Discarding Variable in a principal component analysis. Journal of the Royal Statistical Society. 1972; 21(2): 160-173.

21. Kamara MM, Rehan M, Ibrahim KM, Alsohim AS, Elshakawy MM, Kheir AMS, Hafez EM, El-Esawi MA. Genetic diversity and combining ability of white maize inbred lines under different plant densities. Plants. 2020; 9(9): 1140.

22. Tandzi LN, Mutengwa CS. Estimation of maize (*Zea mays* L.) yield per harvest area: appropriate methods. Agronomy. 2020; 10(29): 1:18.

23. Ertiro BT, Olsen M, Das B, Gowda M, Labuschagne M. Efficiency of indirect selection for grain yield in maize (*Zea mays* L.) under low nitrogen conditions through secondary traits under low nitrogen and grain yield under optimum conditions. Euphytica. 2020; 216: 134.

24. Singh A, Pandey J, Singh S, Singh RP, Singh RK. Correlation and path coefficient analysis for yield and yield attributing traits in advanced bread wheat (*Triticum aestivum* L.) lines. The Pharma Innovation Journal. 2021; 10(8): 482-488.

25. Dofing SM, D'Croze-Mason N, Thomas-Compton MA. Inheritance of expansion volume and yield in two popcorn x dent corn crosses. Crop Science. 1991; 31(3): 715-718.

26. Amaral Júnior AT, Santos A, Gerhardt IES, Kurosawa RNF, Moreira NF, Pereira MG, Gravina GA, Silva FH. Proposal of a super trait for the optimum selection of popcorn progenies based on path analysis. Genetics and Molecular Research. 2016; 15(4): gmr15049309.

27. Saito MA, Alves AV, Kuritza DP, Souza YP, Mioli MFSD, Amaral Júnior AT, Bento AC, Scapim CA, Pinto RJB. Influence of agronomic and kernel-related properties on popping expansion in popcorn. Agronomy Journal. 2021; 113(3): 2260-2272.

28. Souza ARR, Miranda GV, Pereira MG, Souza LV. Predicting the genetic gain in the Brazilian white maize landrace. Ciencia Rural. 2009; 39(1): 19-24.

29. Sheela KS, Robin S, Manonmani S. Principal component analysis for grain quality characters in rice germplasm. Electronic Journal of Plant Breeding. 2020; 11(01): 127-131.

30. Zayed E, Zeinab EG, Saad KI. Genetic diversity and principal component analysis (PCA) of Faba bean landraces based on yield-traits and protein sds-page. Journal of Global Agriculture and Ecology. 2022; 13(4): 1-16.