***Original Research Article***

# Association Among Traits by Correlations and Path in Maize Genotype Selection

---

## ABSTRACT

**Aims:** In the state of Espírito Santo, family farmers have grown a number of maize varieties for decades, consisting of open-pollinated populations with valuable importance for the livelihood of smallholder farmers. As such, this study intended to analyze the cause and effect associations between agronomic traits for increasing yield in maize populations cultivated in Espírito Santo.

**Study design:** Randomized block design with three replicates.

**Place and Duration of Study:** The experiment was conducted in the Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo campus of Alegre, between October 2017 and October 2018.

**Methodology:** A total of 16 accessions of field maize from conservation works in eight different municipalities were evaluated, in a randomized block design with three replicates, totaling 48 experimental units. Genetic parameters, genotypic, phenotypic, and environmental correlation were estimated, and a path analysis was conducted.

**Results:** In general, the genotypes present genetic variability, showing a significant difference between the genotypes for all the traits analyzed by the F test (P=.05). Insertion height of the first cob, plant height, and final plant stand variables presented heritability (h2) higher than 80%, suggesting that superior genotypes can be achieved. The highest estimate of phenotypic correlation (rp) was found between the insertion height of the first cob with plant height 0.85 and total number of cobs with number of cobs per plant 0.85. For the estimates of genotypic correlation (rg), the highest was of 0.88 between the insertion height of the first cob and the plant height; number of grains per row and number of cobs per plant; total number of cobs and final plant stand.

**Conclusion:** The final plant stand and the mass of one thousand seeds are determinants to directly increase the grain yield.

*Keywords: Zea mays L., heritability, genetic parameters, simultaneous selection.*

## 1. INTRODUCTION

Maize (Zea mays L.) is one of the most grown cereals worldwide, for its high genetic diversity and wide adaptability [1]. Such a fact, associated with its nutritional characteristics, makes maize also one of the cereals of commercial importance in different countries [2]. Brazil is the third largest maize producer, meaning that the estimated national production for the first

2018/19 harvest is 27.4 million tons and 63.7 million tons for the second one, with a mean yield of 5.2 t ha-1 in the last harvest [3].

In the state of Espírito Santo, family farmers have grown a number of maize varieties for decades, consisting of open-pollinated populations with valuable importance for the livelihood of smallholder farmers. Estimated total contribution for the 2019 harvest is 38.0 thousand t, with yield (approximately 2.8 t ha-1) being one of the lowest in the country [3].

Open-pollinated populations present lower yield than that of cultivars due to their wide genetic basis [4];[5]. While such variability does not ensure high yield in the short term, it represents a source of alleles favorable for genetic response to adverse biotic and abiotic factors, conferring greater long-term yield stability to maize populations [6];[7];[8]. This is a knowledge that needs to be considered so that the conservation of germplasm of these maize populations is increasingly promoted [9].

With the purpose of exploring favorable alleles in these populations to obtain higher yield values, the study of agronomic traits has been of paramount importance in plant breeding, due to the possibility of identifying variability in germplasm and, especially, the chance to select superior accessions for morpho-agronomic traits of interest [10];[11]. In this regard, the association of knowledge on the correlations between agronomic traits allows the breeder to design strategies that ensure a higher probability of obtaining superior materials [12].

Correlation studies between traits can be conducted through simple correlations, combined with the study of the genetic parameters involved, given that the correlations between two traits can be of phenotypic, genotypic, or environmental nature, in which only genotypic correlations associate heritability [13]. Simple correlations, despite being useful, do not allow conclusions to be drawn about cause and effect relationships between them, i.e. they do not comprise the direct and indirect effects of traits on a basic variable. As an alternative, by means of a path analysis, it is possible to analyze these relationships between the variables, according to the Wright method [14].

In the state of Espírito Santo, both the Instituto Federal do Espírito Santo (Ifes) and the Instituto Capixaba de Pesquisa e Extensão Rural (Incaper) have been working on studies with maize populations grown in the state, gathering germplasm of these populations and working on breeding strategies so as to increase grain yield in the state [15];[16]. As such, studying correlations between agronomic traits in these maize populations will contribute to the optimization of the strategies used by breeders in the state.

UNDER PEER REVIEW

In this manner, the purpose of this study was to analyze the cause and effect associations between agronomic traits for increasing yield in populations of maize grown in Espírito Santo.

## 2. MATERIAL AND METHODS

The experiment was conducted in the Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo campus of Alegre in 2017. According to the Köppen international classification, the climate of the region is of “Cwa” type, i.e. tropical hot humid, with a cold and dry winter and a hot and rainy summer. The mean temperature of the region is 23.1º C, and the mean total precipitation is 1,341 mm [17].

In the 2017/2018 harvest season, 16 field maize genotypes from conservation work were evaluated in eight different municipalities in the state of Espírito Santo (Table 1). All accessions obtained are described as open-pollinated populations and have been cultivated for some years by smallholder farmers in their respective locations.

Table 1 - Accessions (populations) of maize evaluated in accordance with agronomic traits, in the municipality of Alegre, Espírito Santos state, Brazil, and respective municipalities where they were conservated.

| Acessions | GEBs | Origin |
|---|---|---|
| Aliança | - | Muqui-ES |
| Asa Branca | IFES Itapina | Itapina-Colatina-ES |
| Caiano | - | Linhares-ES |
| Caipira | - | Linhares-ES |
| Celina | - | Celina - Alegre-ES |
| BRS Cipotânea | IFES Itapina-ES | Itapina-Colatina-ES |
| BRS Diamantina | IFES Itapina-ES | Itapina-Colatina-ES |
| Emcapa 201 | INCAPER | Viana-ES |
| ES001 | IFES Itapina | Itapina-Colatina-ES |
| Fortaleza | - | Muqui-ES |
| Incaper Capixaba 203 | INCAPER | Viana-ES |
| MA008 | IFES Itapina-ES | Itapina-Colatina-ES |
| Palha Roxa | IFES Alegre-ES | Iúna-ES |
| Palha Roxa | IFES Alegre-ES | Muniz Freire-ES |
| Palha Roxa | - | Venda Nova do Imigrante-ES |
| Sertanejo | IFES Itapina-ES | Itapina-Colatina-ES |

GEBs= Germplasm banks; IFES= Instituto Federal do Espírito Santo; INCAPER= Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural.

UNDER PEER REVIEW

In order to implement the evaluation test of the genotypes in the field, it was opted for the randomized block design with three replicates, according to the following statistical model:

X_ij=m+t_i+b_j+e_ij

in which: m = represents the overall mean, ti = treatment effect (genotypes), bj = block effect, eij = effect of the experimental error.

The experiment comprised 48 experimental units, which were composed of three 4.0 m long lines spaced at 1.00 m between them. The evaluation of agronomic traits was performed on 16 plants within the central line of the plot.

During planting, 15 seeds per linear meter were uniformly distributed in furrow. At 21 days after sowing (DAS), thinning was performed to establish a population of 5 plants per linear meter of furrow, corresponding to a density of 50,000 plants ha-1 (adapted to Corrêa et al., 2014). The phytosanitary treatments followed the technical recommendations for the cultivation, [18]. The maize was naturally dried in the plant, without using desiccants, until it reached the water content of 13%. In February 2018, the harvest was done manually.

The agronomic traits evaluated were as follows: I – Plant height (HEI) – measured from the base of the stem to the apex of the tassel; II – Insertion height of the first cob (IHC) – measured from the base of the stem to the first cob; III – Total number of cobs (TNC); IV – final plant stand (FPS); V – Plant cob diameter (PCD): VI – Stripped cob length (SCL) – measured from the base to the apex of the cobs; VII – Stripped cob diameter (SCD) – measured in the middle of each cob; VIII – Number of rows of grain per cob (NRC); IX – Number of grains per row (NGR); X – Mass of one thousand seeds (MTS); XI – Yield (YIE); XII – Number of cobs per plant (NCP). All the traits under evaluation are descriptors established by Biodiversity International [19]

Aside from the conventional treatments to control the fall armyworm (Spodoptera frugiperda), leaf sprays were carried out using the Dipel WP biological insecticide, made from Bacillus thuringiensis. The quantity used was of 500g ha-1 and the applications were made with the aid of a manual knapsack sprayer with a 20-liter capacity. Mechanical weeding controlled the weeds at 30 DAS.

For statistical analyses, the phenotypic (rp), genotypic (rg), and environmental (re) correlation analyses and linear regression analysis were performed. For correlations, the following expressions were applied: phenotypic correlations:

UNDER PEER REVIEW

$$r_F = \frac{MPG_{XY}}{\sqrt{MSG_X\, MSG_Y}};$$

genotypic correlations:

$$r_G = \frac{(MPG_{XY} - PMR_{XY})/r}{\sqrt{\hat{\Phi}g_{(X)}\hat{\Phi}g_{(Y)}}} = \frac{\hat{\Phi}g_{(XY)}}{\sqrt{\hat{\Phi}g_{(X)}\hat{\Phi}g_{(Y)}}},$$

and environmental correlations:

$$r_p = \frac{MPG_{XY}}{\sqrt{MSR_X\, MSR_Y}}.$$

in which, $MPG_x$ = mean product between genotypes for traits X and Y;
$MPR_{xy}$ = mean product between residues for traits;
$MSG_x$ = mean square between genotypes for trait X;
$MSG_y$ = mean square between genotypes for trait Y;
$MSR_x$ = mean square between residues for trait X;
$MSR_y$ = mean square between residues for trait Y; $\hat{\Phi}g_{(XY)}$ = genotypic covariance estimator;
$\hat{\Phi}g_{(X)}, \hat{\Phi}g_{(Y)}$ = estimators of quadratic components associated with genotypic variabilities for traits X and Y, respectively.

The path analysis consisted of studying the direct and indirect effects of the above-mentioned explanatory independent variables (X) on grain yield, main dependent variable (Y). Considering Y to be a complex trait, resulting from the combined action of other traits, the following model can be defined:

$$Y = \beta_{1X1} + \beta_{2X2} + \cdots \beta_{nXn} + \varepsilon,$$

in which: $X_1, X_2, \dots, X_n$ are the explanatory variables,
and Y is the main variable (or dependent variable).

The direct and indirect effects of the explanatory variables are estimated on the main variable. Therefore, $r_{iy} = p_i + \sum_{j \neq 1}^{n} p_{ij}\, r_{ij}$ in which: correlation between the main variable (Y) and the i-th explanatory variable; $p_i$: direct effect of variable *i* on the main variable; and $p_j r_{ij}$: indirect effect of variable *i* by means of variable *j* on the main variable.

The significance of the genotypic correlation coefficient and the b1 of the regression were evaluated by the "t" test, and the bootstrap with 5000 simulations for phenotypic and environmental correlations was applied in accordance with [20].

In order to verify the collinearity between the traits, a multicollinearity test was conducted, in line with Montgomery and Peck cited by [21]. Subsequently, it was carried out the split of the simple correlation coefficients into direct and indirect effects provided by the trail analysis. For all analyses, the computational resources from the Genes program were applied [22].

UNDER PEER REVIEW

## 3. RESULTS AND DISCUSSION

All evaluated traits acted in accordance with the assumption of normality of error distribution (p = .05) by the Lilliefors test and homogeneity of residual variances (p = .05) by the Bartllet test. These results prove that, normally, the mathematical assumptions needed to conduct the analysis of variance and further studies were satisfied [23].

As a general rule, the genotypes exhibit genetic variability, showing a significant difference for all the traits under analysis, excluding the stem diameter of the plant, by the F test (P=.05) (Table 2). In this regard, [24] observed the variability of field maize, super sweet maize and teosinte maize populations, pointing out that the low plant genetic variability leads to lower genetic gain for breeding programs. The experimental coefficients of variation ($CV_e$%) showed magnitudes ranging from 5.03, for stripped cob diameter, to 25.14%, for plant stem diameter. In turn, the coefficient of genetic variation ($CV_{gi}$%) varied from 0.0 to 22.51 for the traits plant stem diameter and yield, respectively. These values are basically explained by the variability of the genetic material used in the analysis. Research with other cultures shows that high variability of genotypes makes the selection process of superior material easier [25]; [26]; [27]; [28].
As reported by [29], heritability values (h2) above 80% and variation index above unit 1.0 ensure satisfactory selection gains. For the insertion height of the first cob, plant height, and final plant stand, this situation was verified suggesting that, for these variables, superior genotypes can be obtained by means of simple selection methods, such as mass selection (Table 3).

The estimates of phenotypic (rp), genotypic (rg), and environmental (ra) correlation are depicted in (Table 2). Hence, among the 55 pairs of combinations for the 12 variables evaluated, 13 had significant (rp), 12 to 1 or 5% probability by the "t" test, and only one significant to 1% by the bootstrap method with 5000 simulations. There was a positive and high variation from 0.50 to 0.85 for 12 pairs. (rg) also presented 13 significant pairs, but 12 by the bootstrap method with 5000 simulations at 1 or 5% probability, and only one to 5% probability by the "t" test. For this one, there was a positive and high variation from 0.53 to 0.88 for 12 pairs. For (re), 23 combinations were significant using the bootstrap method with 5000 simulations at 1 or 5% probability, varying from 0.4 to 0.88 positive pairs (Table 3).

The highest estimate of (rp) was noticed between the insertion height of the first cob, with plant height 0.85, and the total number of cobs with number of cobs per plant 0.85. The lowest correlations (rp) were between stripped cob length with stripped cob diameter 0.50 and stripped cob length and mass of

UNDER PEER REVIEW

one thousand seeds 0.50, being significant estimates by the "t" test. These results corroborate the ones seen by [30], which indicate the variable cob height as the most influential in yield variation and of greater relevance in indirect selection for yield. Number of rows of grains per cob and mass of one thousand seeds showed significant and high value, though negative -0.55. The higher the number of rows, the lower the conversion of photoassimilates into starch [31].

Among the estimates of (rg), the highest one was 0.88 between the insertion height of the first cob and the plant height; number of grains per row and number of cobs per plant; total number of cobs; and final plant stand. Also among the positive estimates, the lowest one was found between the insertion height of the first cob and the final plant stand, with 0.53. There was only one significant and high pair, although with a negative value -0.57 between number of rows of grains per cob and mass of one thousand grains. The genotype selection with a larger diameter and mass of one thousand grains enables obtaining genotypes with a good yield [32]. Moreover, for (re), the highest estimate happened between total number of cobs and number of cobs per plant, with 0.88. Stripped cob length and stripped cob diameter 0.4 was the lowest positive and significant value. There was a significant and high pair, however with a negative value between the number of rows of grains per cob and mass of one thousand seeds -0.46, (Table 3).

The insertion height of the first cob indicated negative correlations (rg) with stripped cob length of -0.19, stripped cob diameter of -0.32, and number of rows of grains per cob of -0.54; and positive correlations with number of grains per row of 0.28, number of cobs per plant 0.47, mass of one thousand seeds of 0.09, and yield of 0.29, still not significant in all cases (Table 3). In accordance with [21] a non-significant or low magnitude correlation coefficient does not suggest lack of relationship between two variables, but absence of a linear relationship between them.

UNDER PEER REVIEW

Table 2. Estimates of genetic parameters for phytotechnical traits of 16 maize genotypes in the municipality of Alegre, Espirito Santo state, Brazil, 2017

| SV | DF | IHC | HEI | PCD | SCL | SCD | NRC | NGR | TNC | FPS | NCP | MTS | YIE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ------------ Mean squares ------------ | | | | | | | | | | | |
| Blocks | 2 | 0.01 | 0.00 | 0.59 | 1.49 | 2.26 | 0.36 | 19.20 | 134548.68 | 48868.35 | 0.01 | 1722.67 | 829.53 |
| Varieties | 15 | 0.10** | 0.15** | 13.15ns | 2.80* | 17.09** | 4.06** | 28.63* | 203435.78** | 93182.08** | 0.02** | 4422.62** | 4449.19** |
| Residue | 30 | **0.00** | 0.01 | 13.32 | 1.37 | 4.46 | 0.86 | 13.44 | 44870.97 | 14574.77 | 0.01 | 980.06 | 1075.33 |
| | | ------------ Genetic parameters ------------ | | | | | | | | | | | |
| $(\sigma^2_f)$ | | 0.03 | 0.05 | 4.38 | 0.93 | 5.69 | 1.35 | 9.54 | 67811.92 | 31060.69 | 0.00 | 1474.20 | 1483.06 |
| $(\sigma^2_e)$ | | 0.00 | 0.00 | 4.44 | 0.45 | 1.48 | 0.28 | 4.48 | 14956.99 | 4858.25 | 0.00 | 326.68 | 358.44 |
| $(\sigma^2_g)$ | | 0.03 | 0.04 | 0.0 | 0.47 | 4.20 | 1.06 | 5.06 | 52854.93 | 26202.43 | 0.00 | 1147.52 | 1124.62 |
| $(\hat{h}^2_g)$ | | 90.57 | 87.17 | 0.0 | 51.05 | 73.87 | 78.74 | 53.04 | 77.94 | 84.35 | 64.80 | 77.83 | 75.83 |
| $(CV_e\%)$ | | 9.30 | 6.11 | 25.14 | 7.50 | 5.03 | 7.41 | 11.42 | 16.00 | 7.59 | 12.35 | 11.14 | 22.01 |
| $(CV_{gi}\%)$ | | 16.66 | 9.20 | 0.0 | 4.42 | 4.88 | 8.24 | 7.00 | 17.37 | 10.17 | 9.67 | 12.05 | 22.51 |
| $(CV_{gi}/CV_e)$ | | 1.78 | 1.50 | 0.0 | 0.58 | 0.97 | 1.11 | 0.61 | 1.08 | 1.34 | 0.78 | 1.08 | 1.02 |
| $(r)$ | | 0.95 | 0.93 | 0.0 | 0.71 | 0.85 | 0.88 | 0.72 | 0.88 | 0.91 | 0.80 | 0.88 | 0.87 |
| Overall | | 1.05 | 2.30 | 14.51 | 15.61 | 41.98 | 12.54 | 32.10 | 41840.27 | 50289.35 | 0.82 | 281.01 | 4709.72 |

IHC - insertion height of the first cob; HEI – plant height; PCD - plant cob diameter; SCL - stripped cob length; SCD - stripped cob diameter; NRC - number of rows of grain per cob; NGR - number of grains per row; TNC – total number of cobs; FPS - final plant stand; NCP - number of cobs per plant; MTS - mass of one thousand seeds; YIE – yield.

Genotypic variance ($\sigma^2_g$), residual variance ($\sigma^2_e$), phenotypic variance ($\sigma^2_f$), heritability ($\hat{h}^2_g$), coefficient of genotypic

UNDER PEER REVIEW

variation ($CV_{gi}$%), coefficient of residual variation ($CV_e$%), coefficient of relative variation, considering ($CV_{gi}/CV_e$), coefficient of correlation $(r)$ and mean.

UNDER PEER REVIEW

The phenotypic and genotypic correlations between the number of rows of grains per cob and mass of one thousand seeds were predominantly significant and negative, however, of less than -0.60 magnitude (Table 3). The larger number of rows in a cob tends to reduce the grain size, thereby influencing the seed weight in a negative way. Cob volume and grain volume are those that most contribute to increasing the mean cob weight in production components, for super sweet maize populations [10].

The simultaneous selection of traits, such as number of grains per cob and grain weight per cob, is a difficult task for plant breeders, considering that the genes on which these traits are conditioned, have often negative correlations in genotypes [33]. The significance between the stripped cob diameter and the mass of one thousand seeds in the phenotypic and genotypic correlations, with values of 0.70 and 0.77, respectively, demonstrate that large cobs have a larger number of rows; nevertheless, the cob size is a relevant trait for a larger number of grains, contributing to a greater grain weight per cobs.

The variable insertion height of the first cob showed a positive phenotypic and genotypic correlation coefficient with regard to yield, with 0.31 and 0.29 respectively (Table 3). As stated by [34], when the correlation coefficient is positive, but the direct effect is negative, or when the value is insignificant, indirect effects cause the correlation. The variable stripped cob diameter also showed positive values by means of phenotypic and genotypic correlations, of 0.42 and 0.36, respectively, in relation to yield.

Table 3. Phenotypic ($r_p$), genotypic ($r_g$), and environmental ($r_e$) correlations among 12 phytotechnical traits of 16 maize accessions in the municipality of Alegre, Espirito Santo state, Brazil, 2017

| Variables | Correlation | HEI | SCL | SCD | NRC | NGR | NTE | TNC | NCP | MTS | YIE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| IHC | ($r_p$) | 0.85** | -0.05 | -0.25 | -0.45 | 0.25 | 0.55++ | 0.51* | 0.41 | 0.09 | 0.31 |
|  | ($r_g$) | 0.88++ | -0.19 | -0.32 | -0.54 | 0.28 | 0.58++ | 0.53+ | 0.47 | 0.09 | 0.29 |
|  | ($r_e$) | 0.62++ | 0.33++ | 0.06 | 0.02 | 0.29+ | 0.39++ | 0.38+ | 0.28+ | 0.08 | 0.47++ |
| HEI | ($r_p$) |  | 0.33 | -0.00 | -0.35 | 0.40 | 0.36 | 0.37 | 0.24 | 0.27 | 0.35 |
|  | ($r_g$) |  | 0.40 | -0.03 | -0.45 | 0.50 | 0.41 | 0.40 | 0.3 | 0.33 | 0.37 |
|  | ($r_e$) |  | 0.26 | 0.15 | 0.12 | 0.26 | 0.15 | 0.16 | 0.10 | 0.00 | 0.28+ |
| SCL | ($r_p$) |  |  | 0.50* | 0.11 | 0.33 | -0.09 | -0.15 | 0.01 | 0.50* | 0.32 |
|  | ($r_g$) |  |  | 0.59 | 0.22 | -0.07 | -0.20 | -0.26 | -0.05 | 0.59 | 0.20 |
|  | ($r_e$) |  |  | 0.4++ | -0.08 | 0.77++ | 0.09 | 0.06 | 0.11 | 0.39++ | 0.57++ |
| SCD | ($r_p$) |  |  |  | 0.02 | 0.19 | -0.05 | 0.13 | -0.18 | 0.70** | 0.42 |
|  | ($r_g$) |  |  |  | -0.08 | 0.24 | -0.14 | 0.08 | -0.30 | 0.77++ | 0.36 |

UNDER PEER REVIEW

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $(r_e)$ | 0.40++ | 0.12 | 0.21 | 0.31 | 0.08 | 0.47++ | 0.61++ |
| NRC | $(r_p)$ | | -0.26 | -0.28 | -0.13 | -0.33 | -0.55* | -0.19 |
| | $(r_g)$ | | -0.49 | -0.38 | -0.18 | -0.46 | -0.57* | -0.32 |
| | $(r_e)$ | | 0.15 | 0.05 | 0.09 | 0.00 | -0.46++ | 0.25 |
| NGR | $(r_p)$ | | | 0.48 | 0.30 | 0.51* | 0.19 | 0.67** |
| | $(r_g)$ | | | 0.79++ | 0.5 | 0.88++ | 0.35 | 0.85+ |
| | $(r_e)$ | | | -0.09 | -0.10 | -0.02 | -0.09 | 0.38+ |
| TNC | $(r_p)$ | | | | 0.84** | 0.85** | 0.04 | 0.81** |
| | $(r_g)$ | | | | 0.88++ | 0.86++ | 0.00 | 0.83++ |
| | $(r_e)$ | | | | 0.68++ | 0.88++ | 0.18 | 0.74++ |
| FPS | $(r_p)$ | | | | | 0.44 | 0.09 | 0.74 |
| | $(r_g)$ | | | | | 0.51 | 0.06 | 0.78++ |
| | $(r_e)$ | | | | | 0.26 | 0.23 | 0.58++ |
| NCP | $(r_p)$ | | | | | | 0.00 | 0.66** |
| | $(r_g)$ | | | | | | -0.02 | 0.69+ |
| | $(r_e)$ | | | | | | 0.09 | 0.61++ |
| MTS | $(r_p)$ | | | | | | | 0.43 |
| | $(r_g)$ | | | | | | | 0.43 |
| | $(r_e)$ | | | | | | | 0.42++ |

HEI – plant height; SCL - stripped cob length; SCD - stripped cob diameter; NRC - number of rows of grain per cob; NGR - number of grains per row; TNC – total number of cobs; FPS - final plant stand; NCP - number of cobs per plant; MTS - mass of one thousand seeds; YIE – yield. $^{*}$, $^{**}$ and ns = significative at 5 and 1% probability and no significative, respectively, by the t test. $^{+}$, $^{++}$ = significative at 1 and 5%, respectively, by the bootstrap method with 5000 simulations.

Before performing the path analysis, a multicollinearity analysis was conducted among the variables [21]. Genotypic correlation matrices were submitted to the diagnosis of multicollinearity on the basis of the number of conditions. The elimination of the variables stem diameter of the plant and number of cobs per plant was required in view of the severe multicollinearity, 107,826.62. For the remaining variables, insertion height of the first cob; plant height; stripped cob length; stripped cob diameter; number of rows of grains per cob; number of grains per row; total number of cobs; final plant stand; mass of one thousand seeds; and yield, the number of conditions was 742.19, which enabled classifying them from moderate to strong, not affecting the path analysis statistics. It was decided to use, in this study, only

UNDER PEER REVIEW

the variables with no collinearity, given that, in strong or severe multicollinearity, the variances associated with path coefficients are likely to reach high values [12].

Table 4 depicts the direct and indirect effects of the explanatory variables using grain yield as the main variable. The coefficient of determination ($R^2$) in the path analysis model displayed a value of 0.9236 and residual effects lower than 0.0143. As such, the model showed the cause and effect relationship between the explanatory variables and grain yield. The satisfactory use of path coefficients is directly linked to the composition of causal diagrams, which should be listed to the most important variables in the expression of the main variable [35] The diagram applied enabled to explain 92.36% ($R^2$) of the variation in grain yield (Table 4).

In accordance with [30], in five hybrid maize lines, the weight of 100 grains was the variable yield, which generated the greatest direct effect on grain yield, being the most indicated for indirect selection regarding yield. Greater direct effect and greater total correlation on grain yield point to a great contribution to increase yield [36]. The highest values of direct effects on grain yield were seen for the mass of one thousand seeds, with 0.8079, followed by the final plant stand, with 0.5850, and total number of cobs, with 0.4359, while the stripped cob diameter displayed negative direct effect and high magnitude, with -0.5173, and positive phenotypic correlation of mean magnitude, 0.42, with the indirect effect being considered on the mass of one thousand seeds, 0.6266, in the conditions in which the experiment was carried out (Table 4).

The direct effect between the variable insertion height of the first cob and yield was negative, -0.6416, while the phenotypic correlation was positive and with a mean magnitude of 0.31, which should take into account the indirect effect in the final plant stand 0.3150 to benefit from the gain in these two traits (Table 4). As stated by [23], the final plant stand is critical to increase the production of dry mass and commercial cobs, both directly and indirectly. This is not the case in [36], who reported a positive value for the trait insertion height of the cobs in the harvest with the yield, even though they had worked with precocious cycle maize cultivars. In this case, a direct selection on the causal factor effect may not be efficient to improve the trait yield. Also in agreement with [37], the plant height in the harvest can be considered for indirect selection, because of the positive linear relationship with yield.

The low correlation observed between the insertion height of the first cob and the stripped cob diameter with the grain yield was caused by the negative indirect effects via the stripped cob length, number of rows of grains per cob, plant height, and total number of cobs, with very low values similar to the ones reported by [23]. Conversely, when a direct effect displays positive results, indirect effects are responsible for the lack of correlation [38]. In this

UNDER PEER REVIEW

way, for the indirect selection of more productive genotypes, the trait yield should be related, with positive effects, through the desired trait. As mentioned by [11], the negative correlation between different phenotypic traits is assigned to different genes that are controlling these traits such is the existing negative correlation between grains per cob and grain thickness.

For indirect effects with negative values, the simultaneous selection in an indirect way leads to a low efficiency, [36]. In contrast, the positive values for the direct effects of primary components on grain yield are good predictors of genetic correlation according [39] and [40].

The indirect selection for increasing of the traits number of grains and number of rows is efficient in increasing the grain weight [33]. Indirect selection may result in faster genetic progress than direct selection of the desired trait, [41] In an experiment with open-pollinated varieties, [31] achieved a greater correlation between the number of grains per row and yield 0.586.

Table 4. Direct and indirect effects of agronomic variables of maize from 16 varieties of maize in the municipality of Alegre, Espirito Santo state, Brazil, 2017

| Variable | Effect | Via | Coefficients |
|---|---|---|---|
| IHC | Direct | YIE | -0.6416 |
| | Indirect | HEI | 0.1744 |
| | | SCL | -0.0072 |
| | | SCD | 0.1686 |
| | | NRC | -0.1190 |
| | | NGR | 0.0688 |
| | | TNC | 0.2558 |
| | | FPS | 0.3150 |
| | | MTS | 0.0796 |
| | Total | | 0.2946 |
| **HEI** | Direct | YIE | 0.1979 |
| | | **AIHC** | -0.5656 |
| | | **SCL** | 0.0153 |
| | | **SCD** | 0.0202 |
| | | **NRC** | -0.0999 |
| | | **NGR** | 0.1227 |

UNDER PEER REVIEW

| | | | |
|---|---|---|---|
| | | **TNC** | 0.1791 |
| | | **FPS** | 0.2390 |
| | | **MTS** | 0.2696 |
| | Total | | 0.3786 |
| SCL | Direct | YIE | 0.0382 |
| | Indirect | **IHC** | 0.1224 |
| | | **HEI** | 0.0796 |
| | | **SCD** | -0.3068 |
| | | **NRC** | 0.0489 |
| | | **NGR** | -0.0174 |
| | | **TNC** | -0.0885 |
| | | **FPS** | -0.1573 |
| | | **MTS** | 0.4822 |
| | Total | | 0.2015 |
| SCD | Direct | YIE | -0.5173 |
| | Indirect | **IHC** | 0.2091 |
| | | **HEI** | -0.0077 |
| | | **SCL** | 0.0226 |
| | | **NRC** | -0.0187 |
| | | **NGR** | 0.0588 |
| | | **TNC** | -0.0641 |
| | | **FPS** | 0.0508 |

UNDER PEER REVIEW

| | | | |
|---|---|---|---|
| | | **MTS** | 0.6266 |
| | Total | | 0.3602 |
| NCR | Direct | YIE | 0.2202 |
| | Indirect | **IHC** | 0.3468 |
| | | **HEI** | -0.0898 |
| | | **SCL** | 0.0084 |
| | | **SCD** | 0.0440 |
| | | **NGR** | -0.1189 |
| | | **NTE** | -0.1659 |
| | | **FPS** | -0.1090 |
| | | MTS | -0.4641 |
| | Total | | -0.3283 |
| NGR | Direct | YIE | 0.2417 |
| | Indirect | IHC | -0.1827 |
| | | HEI | 0.1004 |
| | | SCL | -0.0027 |
| | | SCD | -0.1260 |
| | | NRC | -0.1083 |
| | | TNC | 0.3478 |
| | | FPS | 0.2943 |
| | | MTS | 0.2894 |
| | Total | | 0.8539 |
| TNC | Direct | YIE | 0.4359 |
| | Indirect | IHC | -0.3765 |
| | | HEI | 0.0813 |
| | | SCL | -0.0077 |
| | | SCD | 0.0761 |
| | | NRC | -0.0838 |
| | | NGR | 0.1928 |
| | | FPS | 0.5153 |
| | | MTS | 0.0046 |
| | Total | | 0.838 |
| FPS | Direct | YIE | 0.5850 |

UNDER PEER REVIEW

|  |  |  |  |
|---|---|---|---|
|  | Indirect | IHC | -0.3455 |
|  |  | HEI | 0.0808 |
|  |  | SCL | -0.0102 |
|  |  | SCD | -0.0449 |
|  |  | NRC | -0.0410 |
|  |  | NGR | 0.1216 |
|  |  | TNC | 0.3840 |
|  |  | MTS | 0.0543 |
|  | Total |  | 0.7842 |
| MTS | Direct | YIE | 0.8079 |
|  | Indirect | IHC | -0.0632 |
|  |  | HEI | 0.0660 |
|  |  | SCL | 0.0228 |
|  |  | SCD | -0.4012 |
|  |  | NRC | -0.1265 |
|  |  | NGR | 0.0865 |
|  |  | TNC | 0.0024 |
|  |  | FPS | 0.0393 |
|  | Total |  | 0.4342 |
| Coefficient of Determination |  |  | 0.9632 |
| Effect of residual variable |  |  | 0.0143 |

IHC - insertion height of the first cob; HEI – plant height; SCL - stripped cob length; SCD - stripped cob diameter; ; NRC - number of rows of grain per cob; NGR - number of grains per row; TNC – total number of cobs; FPS – final plant stand; MTS - mass of one thousand seeds; YIE – yield.

## 4. CONCLUSION

The genetic variability identified in local breeds enables the selection of genotypes to be used in genetic breeding programs, given that the traits under evaluation showed a genotypic correlation of greater magnitude than the phenotypic one, indicating genetic effects whether of an additive, epistatic or dominance nature, exceeding the contribution of environmental origin.

The gain in the traits stripped cob diameter, which presented an indirect effect with mass of one thousand seeds, positive and with high magnitude, as well as the indirect effect of the variable insertion height of the first cob in the final plant stand, should be exploited. The final plant stand and the mass of one thousand seeds are determinants to directly increase the grain yield.

UNDER PEER REVIEW

## ACKNOWLEDGEMENTS

This work was conducted during a scholarship supported by CAPES/FAPERJ at the State University of Nothern Rio de Janeiro Darcy Ribeiro. Financed by CAPES and FAPERJ – Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil and Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro (FAPERJ).

## COMPETING INTERESTS

Authors have declared that no competing interests exist in this study.

## AUTHORS' CONTRIBUTIONS

'Author Richardson Sales Rocha' designed the study and performed the statistical analysis, 'Author Maxwel Rodrigues Nascimento'' managed the analyses of the study, 'Author José Tiago Barroso Chagas' and "Author Rafael Nunes De Almeida" managed the literature searches, "Author Paulo Ricardo Dos Santos" and "Author Camila Queiroz Da Silva Sanfim De Sant'anna" supporting at the experimental design and traits evaluations, "Author Derivaldo Pureza Da Cruz" and "Author Kleyton Danilo Da Silva Costa" suggests path analysis, "Author Geraldo De Amaral Gravina" and "Author Rogério Figueiredo Daher" review the manuscript and suggests some alterations. All authors read and approved the final manuscript."

## REFERENCES

1. Anilkumar, C., Lohithaswa, HC, Pavan, R. Assessment of genetic diversity in newly developed inbred lines of maize (Zea mays L.). Electron J Plant Breed. 2017; 8: 193.: https://www.researchgate.net/publication/316889104.

2. Araújo VS; Ekclund CRB; FB rabbit; RCV wedge; Lombardi CT; Aguiar RS. Crude protein content and maize fodder yield using residues from the mini - maize crop under no - tillage system. Brazilian Regative of Maize and Sorghum. 2010; 9: 266-276. English. http://dx.doi.org/10.18512/1980-6477/rbms.v9n3p266-276.

3. National Supply Company - CONAB. Follow up of the Brazilian grain harvest: agricultural monitoring 2018/2019 crop. Brasília: Conab. 2019; 6, (4). 75

UNDER PEER REVIEW

4. Andrade, JA da C e Filho, JB de M. Quantitative variation in the tropical maize population, ESALQ-PB1. Scientia Agrícola, 2008; 65, (2), 174-182. English. http://dx.doi.org/10.1590/S0103-90162008000200011.

5. Carpentieri-Pípolo, V, Souza, A de, Silva, DA da, Barreto, TP, Garbuglio, DD, Ferreira, JM. Evaluation of cultivars of criollo corn in a system of low technological level. Acta Scientiarum. Agronomy. Maringá, 2010; 32, (2), 229-233. English. http://dx.doi.org/10.4025/actasciagro()v32i2.430.

6. Nass LL, Valois ACC, Melo IS, Inglis MC Genetic Resources and breeding - plants. Rondonópolis: MT Foundation, 2001; 1183.

7. Machado, AT; Machado, CTT; Coelho, CHM; Nunes, JA. Management of corn genetic diversity and participatory improvement in agricultural communities in the states of Rio de Janeiro and Espirito Santo. Planaltina: Embrapa Cerrados, 2002; 22 (Embrapa Cerrados, Research and Development Bulletin, 32).

8. Abreu, L, Cansi, E, Juriatti, C. Evaluation of socioeconomic yield of Creole varieties and commercial hybrids of corn in the Chapecó microregion. Revista Brasileira Agroecologia, 2007; 2, (1), 1230-1233. English.

9. Parent, B, Tardieu, F. Can current crop models be used in the phenotyping for predicting the genetic variability of yield of plants subjected to drought or high temperature ?. J Exp Bot. 2014; 65: 6179-6189. https://doi.org/10.1093/jxb/eru223.

10. Entringer, Geovana Cremonini et al. Correlation and track analysis for corn production components superdoce.Ceres, 2015; 61, (3). English.

11. Nemati, A, Sedghi, M, Sharifi, RS, Seiedi, MN. Investigation of correlation between traits and path analysis of corn (Zea mays L.) grain yield at the climate of ardabil region (Northwest Iran). Not Bot Horti Agrobot Cluj-Napoca. 2009; 37: 194-198. http://dx.doi.org/10.15835/nbha3713120.

12. Coimbra, RR, Miranda, GV, Cruz, CD, Melo, AV de, Eckert, FR. Genetic characterization and divergence of corn populations rescued from the Southeast of Minas Gerais. Revista Agronômica, 2010; 41, (1), 159-166.

13. Nogueira, Ana Paula Oliveira et al. Track analysis and correlations between characters in soybean cultivated at two sowing times. Bioscience Journal. 2012; 28, (6). English.

14. Wright, S. Correlation and causation Journal of Agricultural Research. 1921; 20: 557-585.

UNDER PEER REVIEW

15. Silva, Olícia Larária et al. Study of the genetic variability of capixabas maize populations by means of molecular markers. Revista Univap, 2017; 22, (40), 385. English .. http://dx.doi.org/10.18066/revistaunivav22i40.900.

16. Valadares, Fernanda Vargas et al. Agronomic performance of complete sibs of corn. Revista Univap, 2017; 22, (40), 460. http://dx.doi.org/10.18066/revistaunivav22i40.1039.

17. Lima, Julião Soares de Souza Et Al. Temporal variability of monthly precipitation in alegre-es. Revista Agronômica, 2008; 39, (2). English.

18. Karam, Décio et al. Response of corn plants to the simulation of mechanical damage. Brazilian Journal of Corn and Sorghum. 2011; 9, (2), 201-211. http://dx.doi.org/10.18512/1980-6477/rbms.v9n2p201-211.

19. International Board For Plant Genetic Resources - (IBPGR). Descriptors for maize / Descriptors for maize / Descripteurs pour le maïs. 1991; 86.

20. Ferreira A; Cross CD; Vasconcelos ES; Birth M; Ribeiro MF; Silva MF. Use of non-parametric bootstrap for the evaluation of phenotypic, genotypic and environmental correlations. Acta Scientiarum. Agronomy. 2008; 30: 647-663. http://dx.doi.org/10.4025/actasciagro()v30i5.5966.

21. Cross CD; Regazzi AJ; Sheep PCS. Biometric models applied to genetic improvement. Viçosa: UF 514p; 2012.

22. Cross CD. (2013). GENES: a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum. Agronomy 35: 271-276. http://dx.doi.org/10.4025/actasciagro()v35i3.21251.

23. Birth, MR. et al. Biometric analysis in maize genotypes suitable for baby corn production in organic farming system. Brazilian Horticulture. 2018; 36, (3), 419-425. http://dx.doi.org/10.1590/s0102-053620180322.

24. Almeida, C, Amorim, EP, Barbosa Neto, JF, Cardoso Filho, JA, Sereno, MJCDM. Genetic variability in populations of sweet co

UNDER PEER REVIEW