



Analysis of Morphological Variation, Grouping and Path Coefficient Studies in a Set of Maize Inbred Lines Local to North East Hill Region of India

**Duddukur Rajasekhar ^{a*}, K. L. Naveenkumar ^a, Pramod Kumar Pandey ^a
and Devyani Sen ^a**

^a *School of Crop Improvement, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University (Imphal), Umiam, Meghalaya, India.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/IJPSS/2022/v34i1731042

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/86769>

Original Research Article

Received 14 February 2022

Accepted 28 April 2022

Published 30 April 2022

ABSTRACT

In the current study a total of 191 maize inbred lines were set up in an augmented randomized complete block design to study the genetic variation and interrelationship of different yield contributing parameters. Observations were recorded on a total of nine phenological and ear / kernel related traits. Significant results for ANOVA indicated presence of substantial variation while the first three axes of PCA could explain 75.54% of the variation present. Ear and flowering traits were orthogonal to each other and the highest variation in PC1 could be attributed to total grain weight. This was also reflected in the Agglomerative Hierarchical Clustering which grouped the inbreds into five clusters primarily on the basis of total grain weight which accounted for the highest variation between the clusters. Substantial heritability genetic advance as percentage of mean were observed for total grain weight indicating that selection for grain weight would be fruitful. Association and path coefficient studies after adjustments for multi-collinearity using k constant method (0.05) revealed that total grain weight was highly and positively correlated with ear weight, number of grain kernel rows and seed index. Therefore, simultaneous selection for these traits would be useful. Overall, our studies go on to show presence of sufficient genetic variability in the inbreds under study and hybridization between inbred lines grouped in different clusters have a high chance of producing heterotic hybrids.

Keywords: Maize; inbreds; genetic variation; correlation; path analysis.

1. INTRODUCTION

Successful crop improvement programmes depend on effective utilization of genetic diversity found in crop plants. Crop diversity in turn is the outcome of interventions both by the environment and local farmers [1]. Among cereals, maize possesses a high degree of genetic diversity having undergone extensive modification both since domestication and following migration to different parts of the world [2]. Today, maize is a multipurpose crop with cultivation spanning from approximately 50°N to 45°S latitudinally and broadly classified as temperate, subtropical and tropical germplasm [3, 4].

Hybrid maize development first began in the early 1900s in the USA with landmark contributions from maize breeders who successfully explored the concept of inbred-hybrid development in maize utilizing local germplasm. First proposed by Shull [5] while discussing the composition of open pollinated field maize, this concept focused on inbred development following successive rounds of selfing. Inbred lines in maize can be derived from a multitude of source populations including landraces and have been a rich resource for both fundamental and applied investigations [6].

A general consequence of inbreeding in maize is that it leads to a high degree of uniformity within the lines and high variability between the lines even for inbreds originating from the same parental stock. This variability is the result of fixation of allelic combinations at different loci [7]. Inbreds are highly influenced by environment and identification of stable inbreds is crucial for hybrid development programmes. Such programmes require that a large number of segregating families be critically evaluated in order to identify those with greater yield potential. Evaluation on the basis of cobs with higher grain weight is practical since it ensures availability of seeds at later stages of breeding programmes given that successive selfings lead to inbreeding depression in maize. However, the outcome of grain weight could be due to the direct/indirect influence of other yield contributing traits. Therefore, besides an understanding of the type of variation present, knowledge of the interrelationship between traits which influence the genetic potential of inbreds is also required. Use of path analysis first proposed by Wright [8]

is one such scientific method of analysis which allows the elucidation of the direct and indirect effects influencing the association shared among heritable components [9, 10].

Keeping in mind the necessity to understand factors governing variation and heritability, the current study was framed to evaluate a set of inbreds developed from germplasm local to North East India. An attempt to understand the factors contributing to variation in the experimental material using Analysis of Variance (ANOVA), Principal Component Analysis (PCA) and heritability was followed by path analysis studies. Since the inbreds developed were a collection of different parental stock, genetic diversity analysis using agglomerative hierarchical clustering method was also taken up.

2. MATERIALS AND METHODS

2.1 Plant Material

The current study included a total of 191 maize inbred lines. These inbreds had been developed following three rounds of sibbing followed by two rounds of selfing from a selection of landraces local to North Eastern Hill Region of India -an important secondary center of maize diversity in Asia [11].

2.2 Field Layout

The experiment was laid out in an augmented randomized complete block design consisting of two blocks (Fig. 1). Three commercial checks were repeated thrice at random within each block. The inbreds were sown in single rows with a row and plant to plant spacing of 60 x 30cm and the data on three plants within each row was considered for analysis. All the recommended intercultural practices were taken up and need based application of fertilizers and pesticides was done.

2.3 Observations Recorded

Observations on days to 50 percent tasseling (DT) and days to 50 percent silking (DS) were recorded for calculating the anthesis silking interval (ASI). Ear related traits -cob length (CL) (cm), cob diameter (CD) (cm), ear weight (EW) (g), number of kernel rows per cob (NGKR), 100-seed weight (SI) (g) and grain yield per plant (TGW) (g) were recorded for further statistical analysis.

Block-1					Block-2				
Border row									
Check-3						Check-1			Check-3
Inbred-1									
Inbred-2		Check-1		Check-2					
Inbred-3									
Inbred-4						Check-1			
Inbred-5									
			Check-3					Check-3	
	Check-2				Check-3	Check-2			
							Check-3		
									Inbred-187
		Check-1						Check-2	Inbred-188
						Check-1			Inbred-189
									Inbred-190
									Inbred-191
Check-2									Check-2
									Border row
Border row									

Fig. 1. Experimental plot layout of 191 inbred lines and three checks studied using an augmented block design

2.4 Statistical Analysis

Mean and variance were calculated. The analysis of variance (ANOVA), heritability, and genetic advance as percentage of mean were calculated using “augmentedRCBD.bulk” function of ‘augmentedRCBD’ package [12] in R software [13]. Principal Component analysis was done using XLSTAT software (Version 2016.02.28451). Agglomerative clustering of ward’s method was done by scaling the observation variables (mean of zero and standard deviation of 1) to avoid any bias. Function ‘hclust’ of R software using “clv” package [14] of R software was used to calculate cluster distances. The Pearson correlation matrix was computed using the “corr_coef” function and the correlation plot constructed using the “corr_plot” function of ‘metan’ package [15] in R software [13]. The correlation matrix generated was analysed for multicollinearity following which path analysis studies under collinearity (k constant method) was taken up using the GENES software [16].

3. RESULTS AND DISCUSSION

A study of the mean performance and distribution (Fig. 2) revealed that traits total grain weight, ear weight and number of grain kernel rows were

significantly and positively skewed with higher number of individuals concentrated around the lower values while mean values for days to flowering were significantly skewed towards the larger values. This is expected since inbreeding depression in a highly allogamous species like maize is known to reduce the reproductive fitness [7, 17, 18] and it would practically take more than one generation of open pollination to restore these lines to their original levels of heterozygosity [7, 19]. Exceptions were observed for cob diameter and seed index which despite being associated with reproductive fitness, were negatively skewed while values for cob length appeared to follow a normal distribution. Highest variance was observed for ear and total grain weight and seed index - all three being major yield contributing traits. The lowest variation was detected for number of grain kernel rows. ANOVA studies revealed significant variation at 0.05 significance level of significance for seven quantitative traits with flowering days, anthesis silking interval, cob length and seed index also significant at 0.01 significance level (Table 1). Significant F-values implied presence of variability between the lines under study. Further analysis using PCA revealed that the first three axes accounted for 75.54% of the total variation in the data. All ear related traits barring cob length loaded heavily on PC1. As revealed by

squared cosine values of the variables, total grain weight accounted for the highest percent contribution in this axis. Flowering traits were orthogonal to the ear related traits and loaded heavily on PC2. The highest loading for cob length was observed for PC3 (Table 2).

Agglomerative Hierarchical clustering using Ward's method divided the 191 inbred lines into five clusters (Fig. 3). The largest inter cluster distance was recorded for clusters I and III (5.63) with these clusters having grouped inbreds with highest and lowest respective values for the ear related traits. The lowest inter-cluster distance was observed between cluster IV and V (3.57) (Table 3). The highest contribution to variation in PC1 came from individuals grouped in Cluster I with most of the lines having recorded ear and total grain weight on the higher side. The genetic advance calculated as percentage of mean ranged from medium in case of cob diameter (15%) to high (40-55%) for the remaining ear related traits (Fig. 4). A high genetic advance implied that selection for these traits would have

a genetic basis despite being influenced by environment. With respect to number of grain kernel rows, the variance was highly non-significant and therefore genetic advance could not be calculated for this trait.

Besides knowledge of variation and heritability, an understanding of the degree of correlation among traits is important for formulating selection strategies in crop improvement. The Pearson correlation, regarded as the strength of the linear association between two traits [20, 21] is one of the most commonly used tools for estimating such associations [22-24]. Pearson's correlation studies revealed that total grain weight with maximum variability in PC1 and a high genetic advance was highly and positively correlated with cob diameter (0.55), ear weight (0.78), number of grain kernel rows (0.63) and seed index (0.74). Maize grain yield is the outcome of ear length, ear diameter and thousand-kernel weight which are known to be significantly correlated with each other and influence the final outcome [25].

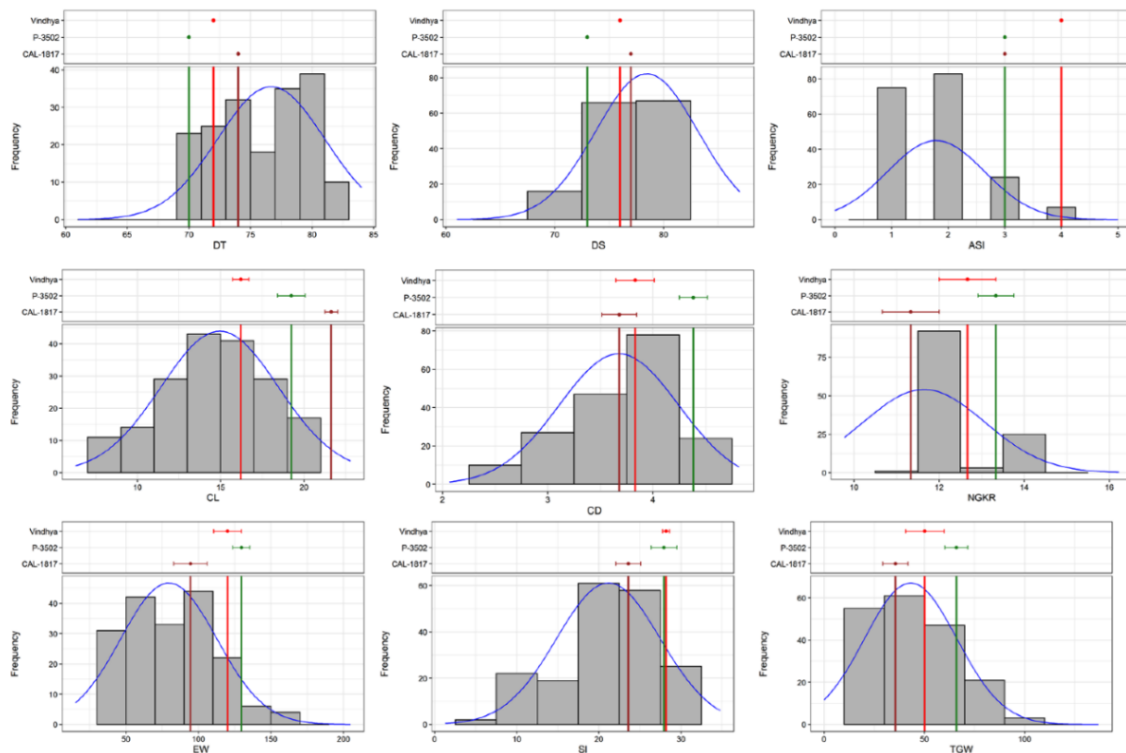


Fig. 2. Mean performance and variance of the nine yield contributing traits studied for the 191 inbred lines along with checks varieties Vindhya (red), P-3502 (green) and CAL-1817 (burgundy). The y-axis represents the frequency of the individuals for each of the parameters studied

DT: Days to 50 % tasseling; DS: Days to 50 % silking; ASI: Anthesis Silking Interval; CL: Cob length; CD: Cob diameter; EW: Ear weight; NGKR: Number of kernel rows per cob; SI: 100-seed weight; TGW: Total grain weight

Table 1. Mean square values for the nine yield contributing traits under study

Source of Variation	df	DT	DS	ASI	EW	CD	CL	NGKR	SI	TGW
Block (ignoring Treatments)	1	316**	293**	0.23**	2663*	0.1465	2.53	9.2	0.42	2467**
Treatment (eliminating Blocks)	193	19**	21**	0.92**	1201*	0.3353	13.28**	2.145	41**	518*
Treatment: Check	2	24**	26**	2**	1972*	0.82*	44.47**	6.222	40*	1388*
Treatment: Test and Test vs. Check	191	19**	21**	0.92**	1193*	0.3302	12.96**	2.102	41**	509.6
Residuals	14	0	0	0	525.7	0.1657	1.66	2.222	9.77	248.9

* $P \leq 0.05$; ** $P \leq 0.01$; DT: Days to 50 % tasseling; DS: Days to 50 % silking; ASI: Anthesis Silking Interval; CL: Cob length; CD: Cob diameter; EW: Ear weight; NGKR: Number of kernel rows per cob; SI: 100-seed weight; TGW: Total grain weight

Table 2. Principal Component Analysis for the nine yield contributing traits under study

Factor loadings:	Contribution of the variables (%):			Squared cosines of the variables:		
	PC1	PC2	PC3	PC1	PC2	PC3
Traits						
DT	0.04	0.94	0.00	0.04	37.55	0.00
DS	0.02	0.98	0.04	0.01	40.98	0.21
ASI	-0.11	0.65	0.24	0.36	18.29	6.70
EW	0.88	-0.03	0.09	22.25	0.04	0.94
CD	0.71	0.15	-0.03	14.42	1.02	0.10
CL	0.47	-0.21	0.82	6.40	1.90	78.94
NGKR	0.71	0.05	-0.25	14.38	0.13	7.55
SI	0.82	-0.04	-0.07	19.10	0.05	0.52
TGW	0.90	0.03	-0.21	23.03	0.04	5.05

DT: Days to 50 % tasseling; DS: Days to 50 % silking; ASI: Anthesis Silking Interval; CL: Cob length; CD: Cob diameter; EW: Ear weight; NGKR: Number of kernel rows per cob; SI: 100-seed weight; TGW: Total grain weight

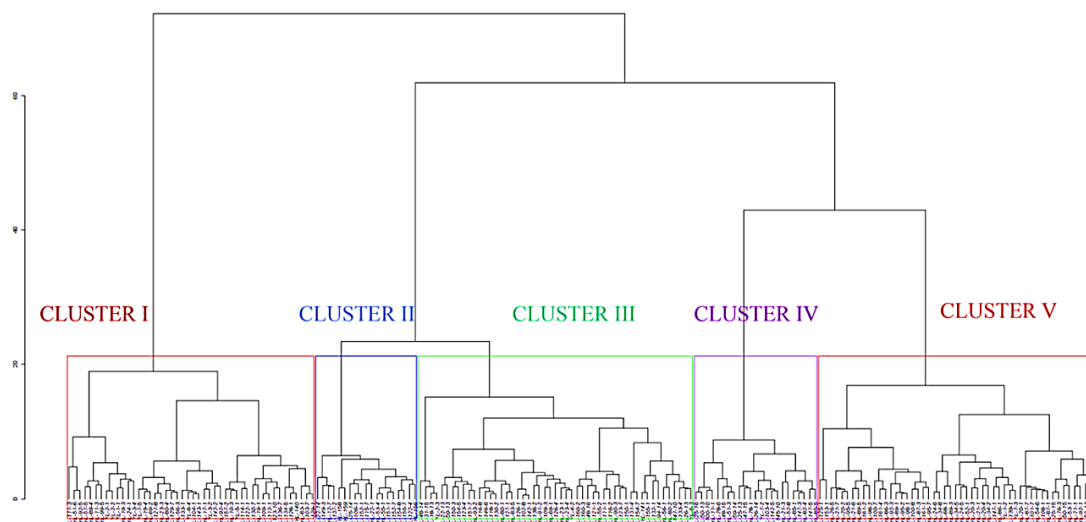


Fig. 3. Clusters based on Agglomerative Hierarchical Clustering with Euclidean Distances using Ward’s method. A total of five cluster were generated

Table 3. Inter and intra cluster distances of the five clusters generated based on the nine quantitative yield traits studied

	No. of inbreds	Cluster-1	Cluster-2	Cluster-3	Cluster-4	Cluster-5
Cluster-1	46	3.11	4.07	5.63	4.56	3.79
Cluster-2	52	4.07	3.23	4.18	4.25	4.30
Cluster-3	23	5.63	4.18	2.61	3.83	4.78
Cluster-4	51	4.56	4.25	3.83	3.06	3.57
Cluster-5	19	3.79	4.30	4.78	3.57	2.58

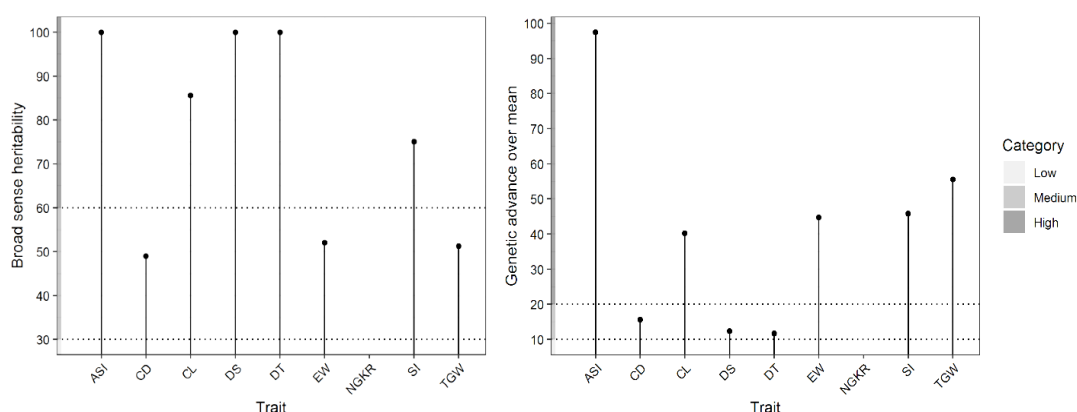


Fig. 4 Broad sense heritability (hBS) and genetic advance as percentage of mean (GAM) and for each of the nine parameters studied

ASI: Anthesis Silking Interval; CD: Cob diameter; CL: Cob length; DS: Days to 50 % silking; DT: Days to 50 % tasseling; EW: Ear weight; NGKR: Number of grain kernel rows; SI: 100-seed weight; TGW: Total grain weight

Use of path coefficients- a form of standardized partial regression analysis can further supplement the information gained from correlation studies by splitting correlation values into direct and indirect effects. Conventional path

analysis for the present study however indicated presence of multicollinearity in the estimation of path coefficients. The direct effects of days to tasseling (-0.30) and that to silking (0.35) were high in magnitude but opposite in direction. Multi-

Table 4. Estimates of direct effects on total grain weight and measures of multicollinearity diagnosis

Trait	Conventional path analysis		Path analysis under collinearity (K=0.05)	
	Direct effects	VIF	Direct effects	VIF
DT	-0.30	2483264605.5	0.013	10.38
DS	0.35	2933107990.6	0.008	12.02
ASI	-0.08	93472538.4	-0.023	1.64
EW	0.44	2.5	0.408	2.35
CD	0.07	1.6	0.084	1.58
CL	-0.12	1.3	-0.097	1.31
NGKR	0.25	1.4	0.248	1.46
SI	0.33	1.9	0.327	1.88

DT: Days to 50 % tasseling; DS: Days to 50 % silking; ASI: Anthesis Silking Interval; CL: Cob length; CD: Cob diameter; EW: Ear weight; NGKR: Number of kernel rows per cob; SI: 100-seed weight; TGW: Total grain weight

Table 5. Estimates of Pearson's correlation coefficients and of the direct and indirect path coefficients of the respective traits on total grain weight after adjustments for multicollinearity

	Correlation (r) coefficients with TGW	Direct effect	Indirect effect via							
			DT	DS	ASI	EW	CD	CL	NGKR	SI
DT	0.05	0.015	0	0.005	-0.008	0.008	0.011	0.012	0.016	-0.008
DS	0.03	0.005	0.015	0	-0.011	-0.001	0.010	0.013	0.012	-0.010
ASI	-0.08	-0.022	0.006	0.003	0	-0.047	0.002	0.010	-0.016	-0.018
EW	0.77**	0.407	0	0	0.003	0	0.046	-0.044	0.128	0.218
CD	0.55**	0.084	0.002	0.001	0	0.223	0	-0.023	0.101	0.161
CL	0.25**	-0.098	-0.002	-0.001	0.002	0.181	0.020	0	0.054	0.098
NGKR	0.63**	0.248	0.001	0	0.001	0.209	0.034	-0.021	0	0.146
SI	0.73**	0.328	0	0	0.001	0.270	0.041	-0.029	0.110	0

DT: Days to 50 % tasseling; DS: Days to 50 % silking; ASI: Anthesis Silking Interval; CL: Cob length; CD: Cob diameter; EW: Ear weight; NGKR: Number of kernel rows per cob; SI: 100-seed weight; TGW: Total grain weight

collinearity generates bias in the estimation of path coefficients, especially in traits with large VIF (Variance Inflation Factors) (Table 4). Very high variance inflation factors (VIF) were observed for traits DT, DS, and the ASI, respectively and therefore collinearity studies using the k constant approach ($k = 0.05$) was taken up. Path coefficients after multicollinearity adjustments revealed that traits ear weight (0.40), seed index (0.32) and number of grain kernel rows (0.24) in that order exhibited the largest direct and positive effect on total grain weight. With respect to ear weight, besides a high direct effect, indirect effect via seed index and cob diameter was also observed to influence the outcome of total grain weight (Table 5).

Figures in bold indicate the high direct effects by the respective explanatory variables on the dependent variable

Insights into the interrelationships between major yield parameters and their associated independent variables is necessary for effective selection especially in traits with low to medium heritability [26]. Path analysis under multicollinearity has been successfully used to understand associations between yield contributing traits after adjustments to address the negative impact of multicollinearity [27-30]. For the present study, ear weight, number of grain kernel rows and seed index appear to have a high direct effect on total grain yield. Therefore, simultaneous selection of these traits when selecting for higher grain weight would prove useful.

4. CONCLUSION

Based on studies on variance components, and genetic parameters the present study revealed presence of substantial variability in the inbreds especially for ear related traits. Accordingly, cluster analysis grouped the inbreds with contrasting ear and total grain weight into the most distant clusters. Hybridization between these contrasting lines from different clusters may be expected to yield heterosis and also produce transgressive segregation during population development. Path analysis after adjustments for multicollinearity revealed that simultaneous selection for ear weight, grain kernel rows and seed index need to be considered to enhance the heritability of total grain weight.

DISCLAIMER

The products used for this research have been developed by the authors. The research was not funded any producing company. Rather it was funded by personal efforts of the authors.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Harlan JR. Crops and man. American Society of Agronomy; 1975.
2. Aci MM, Lupini A, Mauceri A, Morsli A, Khelifi L, Sunseri F. Genetic variation and structure of maize populations from Saoura and Gourara oasis in Algerian Sahara. *BMC Genet.* 2018;19(1):51.
3. Leff B, Ramankutty N, Foley JA. Geographic distribution of major crops across the world. *Global Biogeochemical Cycles.* 2004;18(1).
4. Wegary D, Teklewold A, Prasanna BM, Ertiro BT, Alachiotis N, Negera D, Awas G, Abakemal D, Ogugo V, Gowda M, Semagn K. Molecular diversity and selective sweeps in maize inbred lines adapted to African highlands. *Scientific Reports.* 2019; 9(1):13490.
5. Shull GH. The composition of a field of maize. *J. Hered.* 1908(1):296-301.
6. Lee M. Inbred Lines of Maize and Their Molecular Markers, in *The Maize Handbook*, M. Freeling and V. Walbot, Editors. Springer New York: New York, NY. 1994;423-432.
7. Hallauer AR, Carena MJ, Miranda Filho JD. Quantitative genetics in maize breeding. New York: Springer Science & Business Media. 2010;6.
8. Wright S. Correlation and causation. *Journal of Agricultural Research*, 1921; 20(7):557-585.
9. Wooldredge J. Path Analysis, in *The Encyclopedia of Research Methods in Criminology and Criminal Justice.* 2021; 515-522.
10. Aalen OO, Frigessi A. What can Statistics Contribute to a Causal Understanding? *Scandinavian Journal of Statistics.* 2007;34(1):155-168.
11. Sharma S, Brahma P. Gene bank curators: towards implementation of the international

- treaty on plant genetic resources for food and agriculture by the Indian National Gene Bank, in Plant Genetic Resources and Food Security: Stakeholder Perspectives on the International Treaty on Plant Genetic Resources for Food and Agriculture, C. Frison, F. López, and J. Esquinas-Alcázar, Editors. FAO and Bioversity International: Oxon. 2011;183-196.
12. Aravind J, Mukesh S, Wankhede D, Kaur V. Augmented RCBD: Analysis of Augmented Randomised Complete Block Designs. R package version 0.1, 2020; 2.
 13. R Core Team, R: A language and environment for statistical computing. 1993, R Foundation for Statistical Computing: Vienna, Austria.
 14. Nieweglowski L. Cluster Validation Techniques; 2020.
 15. Olivoto T, Lúcio ADC. Metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*. 2020;11(6):783-789.
 16. Cruz CD. Genes Software-extended and integrated with the R, Matlab and Selegen. *Acta Scientiarum. Agronomy*. 2016;38:547-552.
 17. Dieterich Mabin ME, Brunet J, Riday H, Lehmann L. Self-Fertilization, Inbreeding, and Yield in Alfalfa Seed Production. *Frontiers in Plant Science*. 2021;12.
 18. Busbice TH. Inbreeding in Synthetic Varieties1. *Crop Sci.*, 1969; 9(5): p. cropsci1969.0011183X000900050026x.
 19. Bernardo R. Breeding for quantitative traits in plants. Stemma press Woodbury. 2002;1.
 20. Pearson K. Notes on the history of correlation. *Biometrika*, 1920;13(1):25-45.
 21. Profillidis VA, Botzoris GN. Chapter 5 - Statistical Methods for Transport Demand Modeling, in *Modeling of Transport Demand*, V.A. Profillidis and G.N. Botzoris, Editors. Elsevier. 2019;163-224.
 22. Magar BT, Acharya S, Gyawali B, Timilsena K, Upadhayaya J. Shrestha, Genetic variability and trait association in maize (*Zea mays* L.) varieties for growth and yield traits. *Heliyon*. 2021;7(9): e07939.
 23. Nikolić A, Ignjatović-Micić D, Dodig D, Anđelković V, Lazić-Jančić V. Identification of QTLs for Yield and Drought-Related Traits in Maize: Assessment of Their Causal Relationships. *Biotechnology & Biotechnological Equipment*. 2012;26(3): 2952-2960.
 24. Peng Y, Xiong D, Zhao L, Ouyang W, Wang S, Sun J, Zhang Q, Guan P, Xie L, Li W, Li G., Yan J, Li X. Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. *Nature Communications*. 2019;10(1):2632.
 25. Fan X, Chen H, Tan J, Xu C, Zhang Y, Luo L, Huang Y, Kang M. Combining abilities for yield and yield components in maize. *Maydica*. 2008;53(1):39.
 26. Mohammadi S, Prasanna B, Singh N. Sequential path model for determining interrelationships among grain yield and related characters in maize. *Crop Sci*. 2003;43(5):1690-1697.
 27. Alves BM, Cargnelutti Filho A. Genotypic correlation and path analysis in early and super-early maize genotypes. *Genet Mol Res*, 2017;16(2).
 28. Toebe, M. and A. Cargnelutti Filho, Multicollinearity in path analysis of maize (*Zea mays* L.). *Journal of Cereal Science*. 2013;57(3):453-462.
 29. De Carvalho CG, Borsato R, Cruz CD, Viana JM. Path analysis under multicollinearity in S0 x S0 maize hybrids. *Crop Breeding and Applied Biotechnology*. 2001;1(3).
 30. Crispim-Filho AJ, Dos Santos FP, Pinto JFN, Melo PGS, Dos Reis EF, Mendes-Resende MP. Dealing with multiple traits in maize: A new approach for selecting progenies. *Crop Sci*. 2020; 60(6):3151-3165.

© 2022 Rajasekhar et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
 The peer review history for this paper can be accessed here:
<https://www.sdiarticle5.com/review-history/86769>