Andhra Pradesh J Agril. Sci : 10(4): 311-317, 2024

GENETIC DIVERSITY ANALYSIS OF 64 MAIZE INBRED LINES FOR YIELD TRAITS USING D^2 STATISTICS AND PRINCIPAL COMPONENT ANALYSIS

N. SUDHARSHAN, I. SUDHIR KUMAR*, M. SHANTHI PRIYA, P. MUNIRATHNAM AND M. REDDY SEKHAR

Department of Genetics and Plant Breeding, S.V. Agricultural College, ANGRAU, Tirupati-517 502.

Date of Receipt: 24-12-2024 Date of Acceptance: 29 -12-2024

ABSTRACT

To discriminate maize inbred lines based on yield traits using D^2 statistics and principal component analysis (PCA), sixty four genotypes available at Department of Genetics and Plant breeding, Agriculture Research Station, Peddapuram, Andhra Pradesh were tested during *Rabi*, 2023-24 under irrigation condition in Alpha lattice design with two replications. Analysis of variance revealed significant differences for 14 characters studied among the genotypes. The D^2 statistics displayed that cluster II, being largest group, comprises 37 maize genotypes followed by cluster I (15), III (10), IV (1) and V (1). The maximum intra-cluster distance was obtained for cluster I (99.05) while cluster IV and V showed null values as they were monogenetic clusters. Furthermore, maximum inter-cluster distance was recorded between cluster V and IV (2342.19) followed by clusters I and IV (2038.71) and clusters V and III (1391.95). D^2 statistics and principal component analysis (PCA) emphasized that days to 50 per cent anthesis, plant height, ear height, 100- kernel weight, shelling percentage, number of kernels row^{-1}, and number of kernel rows ear^{-1} and kernel yield plant^{-1} were major contributing traits for total genetic diversity. Research results indicated that the inbreds under study are highly diversified and hence, high heterotic hybrids would be resulted through the crossing of diverse lines clubbed under different clusters as parent in any breeding programme aimed to enhance grain yield.

KEYWORDS: Genetic diversity, Inbred lines, D^2 stastic, Yield traits, Principal component analysis.

INTRODUCTION

Maize (*Zea mays* L.) is a diploid plant species with a chromosome number of 20 (2n=20), ranking as one of the world's most important cereal crops, alongside rice and wheat. It belongs to the family Poaceae, the subfamily Panicoideae, and the tribe Maydeae. Archaeological and molecular evidence indicates that modern maize was domesticated from annual teosinte (*Zea mays* ssp. *Parviglumis*) in southern Mexico between 6,600 and 9,000 years ago (IIMR, 2023). Known as the "Queen of Cereals," maize boasts the highest yield potential among all cereals. It is a versatile crop with wide genetic variability, capable of thriving in tropical, subtropical, and temperate agro-climatic conditions worldwide.

Global maize production reached approximately 1.21 billion tonnes in 2023. The United States led the production with an estimated production of 384 mt, making it the top producer worldwide. Other major maize-producing countries are China, Brazil, and Argentina. India has been among the top 10 maize producers in the world since 1961, and currently ranks 5th (FAO 2023) and 14th largest exporter of Maize in the world (UN-COMTRADE data, 2022), contributing around 38.09 mt in 2023. Andhra Pradesh produced approximately 4.14 mt of maize, with an area under cultivation of around 2.6 lakh hectares. This accounted for about 11.5% of India's total maize production, which was around 33 m t (APSSDC, 2023).

The hybrid seed industry utilizes maize inbred lines selectively, depending on the availability of lines from diverse source populations and gene pools, while evaluating their individual performance and hybrid potential. Comprehensive understanding of genetic diversity among genotypes is critical for selecting parents in hybridization programs, particularly in cross-pollinated crops like maize.

Grain yield, a complex quantitative trait, results from the interplay of key processes such as photosynthesis, transpiration, and food storage (Naushad *et al.*, 2007), alongside the coordinated contribution of traits like the number of kernels per plant, number of kernel rows per ear, number of kernels per row, and 100-kernel weight (Viola *et al.*, 2003). Significant genetic variability, which reflects heritable differences among cultivars, is essential to ensure the effectiveness and longevity of plant breeding programs. Advanced biometrical techniques, such as multivariate analysis (Rao, 1952) using Mahalanobis' (1936) D^2 statistics and principal component analysis (PCA) (Pearson, 1901), have made it possible to quantify and assess the genetic diversity within germplasm. This study focused on distinguishing maize inbred lines based on yield and yield contributing traits using D^2 statistics and PCA.

Observations on various parameters were recorded for three plants selected at random from each entry in each replication except for days to 50% anthesis, days to 50% silking, anthesis silking interval were recorded on plot basis.

Statistical analysis

---

*Corresponding author, E-mail: i.sudhirkuma@angrau.ac.in

MATERIAL AND METHODS

The study was conducted at Agricultural Research Station, Peddapuram, Andhra Pradesh during *Rabi* 2023-24 under irrigated conditions. The experiment was laid out in Alpha Lattice design. The experimental material used in the present study consisted of 64 maize inbred lines developed at ARS, Peddapuram. The entire experimental area was divided into two replications, each comprising (64 of 8) = 8 blocks. Within each block, eight inbred lines were allocated. Each inbred line was planted in two rows, each spanning 4 meters in length, with a spacing of 60cm between rows and 20cm within the rows.

Statistical analysis of genetic divergence was computed by using D^2 statistics of Mahalanobis (1936), and clustering of genotypes was done according to Tocher's optimization method as described by Rao (1952). Relative contribution of characters towards genetic divergence was calculated as per the guideline of Singh and Chaudhary (1985). The principal component analysis was done as described by Pearson (1901).

Table 1. List of genotypes studied

S. No.	Inbred line	T. No.	Inbred line	S. No.	Inbred line	T. No.	Inbred line
1	PL 23039	**17**	PL 23055	**33**	PL 23071	**49**	PL 23087
2	PL 23040	**18**	PL 23056	**34**	PL 23072	**50**	PL 23088
3	PL 23041	**19**	PL 23057	**35**	PL 23073	**51**	PL 23089
4	PL 23042	**20**	PL 23058	**36**	PL 23074	**52**	PL 23090
5	PL 23043	**21**	PL 23059	**37**	PL 23075	**53**	PL 23091
6	PL 23044	**22**	PL 23060	**38**	PL 23076	**54**	PL 23092
7	PL 23045	**23**	PL 23061	**39**	PL 23077	**55**	PL 23093
8	PL 23046	**24**	PL 23062	**40**	PL 23078	**56**	PL 23094
9	PL 23047	**25**	PL 23063	**41**	PL 23079	**57**	PL 23095
10	PL 23048	**26**	PL 23064	**42**	PL 23080	**58**	PL 23096
11	PL 23049	**27**	PL 23065	**43**	PL 23081	**59**	PL 23097
12	PL 23050	**28**	PL 23066	**44**	PL 23082	**60**	PL 23098
13	PL 23051	**29**	PL 23067	**45**	PL 23083	**61**	PL 23099
14	PL 23052	**30**	PL 23068	**46**	PL 23084	**62**	PL 23100
15	PL 23053	**31**	PL 23069	**47**	PL 23085	**63**	PL 23101
16	PL 23054	**32**	PL 23070	**48**	PL 23086	**64**	PL 23102

RESULTS AND DISCUSSION

Cluster composition

The 64 inbred lines were grouped into five clusters by using Tocher's method. The distribution of inbred lines into five clusters is presented in Table 2. Cluster I had 15 inbred lines while, cluster II had highest number of inbred lines i.e., 37, cluster III had 10 inbred lines. The clusters IV and V had single inbred line i.e., unitary clusters. Jaishreepriyanka *et al.* (2020) reported similar grouping pattern of maize inbred lines in their studies.

Intra and Inter-cluster average distance

The intra- and inter-cluster distances for the five clusters are summarized in Table 3 and visualised in Figure 2. Intra-cluster average D^2 values ranged from 0 to 99.05, with cluster II exhibiting the highest intra-cluster distance (99.05) and cluster I showing the lowest (51.91). Clusters IV and V had zero intra-cluster distances, as each contained a single inbred line, indicating their uniqueness and potential utility in breeding programs. Among inter-cluster distances, the lowest D^2 value was

yield plant^{-1} contributed the maximum (14.53%) to the diversity by taking the first rank in 291 times, followed by shelling percentage (12.32%) with 246 times ranked first, number of kernels rows^{-1} (9.25%) with 185 times ranked first and number of kernel rows ear^{-1} contributed 8.54% with 170 times ranked first. 100-Kernel weight (8.54%) with 171 times ranked first. Other characters *viz*., anthesis- silking interval, ear height, tassel height, days to maturity, ear girth, ear length, plant height, days to 50% silking, days to 50% anthesis contributed 7.87, 6.58, 5.88, 5.87, 4.87, 4.8, 3.87, 3.76 and 3.32 per cent, respectively to the total genetic diversity (Table 5). Similar kind of results were reported by Amin *et al.* (2013), Maruthi *et al.* (2015) and Lone *et al.* (2017).

In this study, Mahalanobis D^2 statistics and the clustering pattern showed that Cluster II had the highest observed between clusters I and V (143.11), followed by clusters IV and III (175.83). The highest inter-cluster distance was recorded between clusters V and IV (2342.19), followed by clusters I and IV (2038.71).

Relative contribution of individual characters towards divergence

The characters appeared in the first rank contribute more towards diversity. Among all the characters, kernel

intra-cluster distance (09.05), indicating significant variability and potential for within-cluster selection. Clusters IV and V, containing only one inbred line each, were unique and may harbor rare alleles beneficial for breeding programs. The highest inter-cluster distance was recorded between Clusters IV and V (2342.19), followed by Clusters I and IV (2038.71), reflecting their extreme genetic divergence. Inbred lines from clusters with high inter-cluster distances are likely to produce

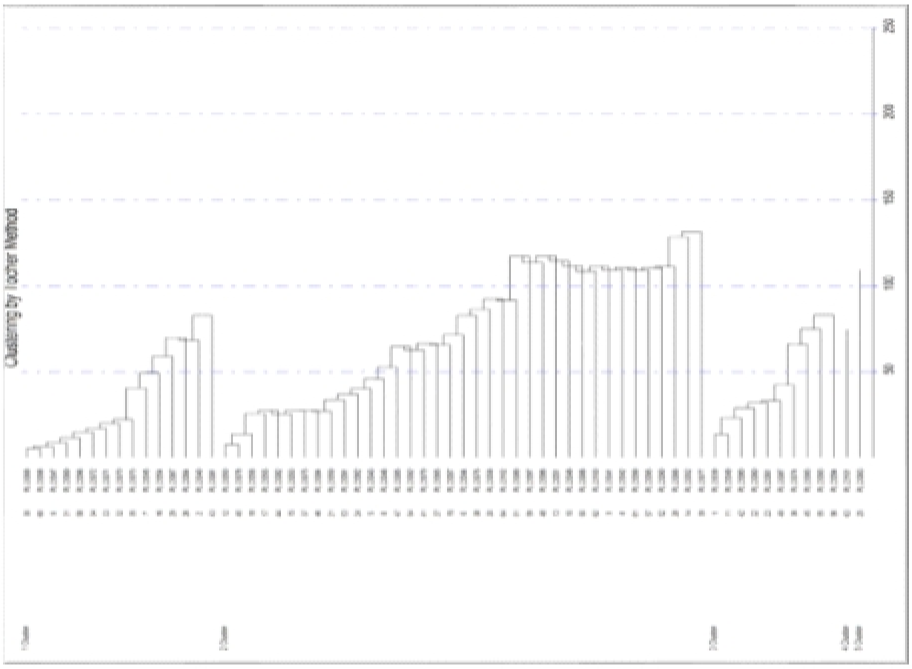
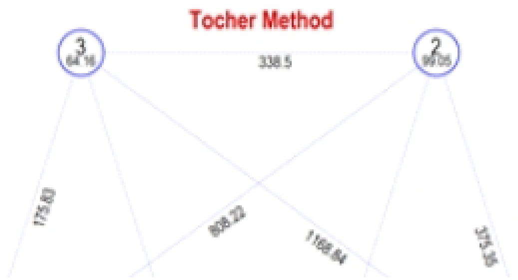


Fig 1. Dendrogram showing relationship based on Mahalanobis D^2 values in 5 clusters among 64 inbred lines of maize (*Zea mays* L.).

Table 2 Clustering pattern by Tocher's method

Cluster Group	No. of Genotypes	List of Genotypes
I Cluster	15	PL23068, PL23098, PL23047, PL23069, PL23096, PL23072, PL23071, PL23070, PL23073, PL23045, PL23054, PL23067, PL23064, PL23040 and PL23081
II Cluster	37	PL23050, PL23078, PL23056, PL23055, PL23082, PL23053, PL23075, PL23084, PL23059, PL23091, PL23062, PL23043, PL23046, PL23085, PL23092, PL23079, PL23065, PL23057, PL23044, PL23076, PL23058, PL23102, PL23089, PL23097, PL23086, PL23051, PL23048, PL23088, PL23100, PL23041, PL23042, PL23099, PL23095, PL23090, PL23066, PL23052 and PL23077
III Cluster	10	PL23039, PL23049, PL23080, PL23060, PL23061, PL23087, PL23074, PL23083, PL23093 and PL23094.
IV Cluster	1	PL23101
V Cluster	1	PL23063



PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) simplifies large datasets, like the one with 14 variables in this case, into a smaller set of variables called principal components (in this study, six components). PCA is a straightforward eigenvector-based multivariate analysis, typically performed on a symmetric covariance matrix.

Fig 2. Intra and inter-cluster distances in five clusters based on Tocher's method in maize.

hybrids with maximum heterosis. Key traits contributing to genetic diversity included kernel yield plant^{-1} (14.53%), shelling percentage (12.32%), and number of kernels row^{-1} (9.25%), making lines excelling in these traits ideal for hybridization programs.

The process involves eigenvalue decomposition of the covariance matrix or singular value decomposition of the data matrix to extract the principal components.

The first principal component (PC1) explained 38.35 per cent of the variance, mainly driven by anthesis-silking interval, days to 50 per cent silking, and kernel traits. The second component (PC2) accounted for 12.80 per cent of the variance, with kernel yield plant^{-1} and 100-kernel weight as key factors. PC3 (11.23%) was influenced by tassel height and ear length, while PC4 (10.38%) highlighted 100-kernel weight and days to 50 per cent anthesis. PC5 (8.11%) was shaped by days to 50% anthesis and ear height, and PC6 (6.19%) by the number of kernel rows^{-1} and days to 50% anthesis.

The PCA scores for 64 maize genotypes across the first three principal components were calculated and presented as three axes (X, Y, Z), with squared distances of each genotype from these axes. The scores were plotted in biplot. Figures 3. Genotypes positioned at the extreme

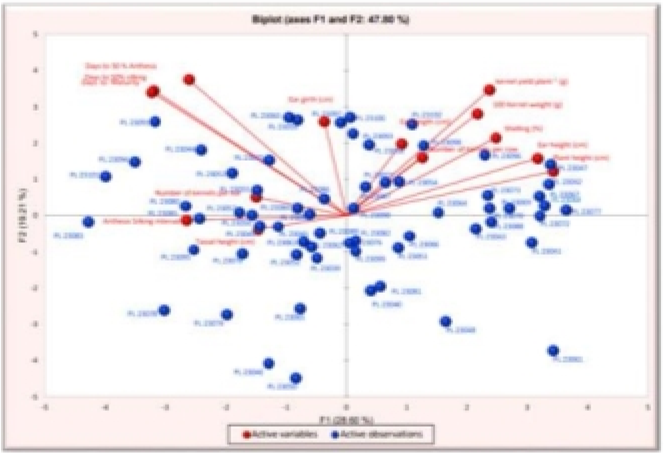


Fig 3. Biplot showing scattering of 64 genotypes of maize (*Zea mays* L.) based on PCA score.

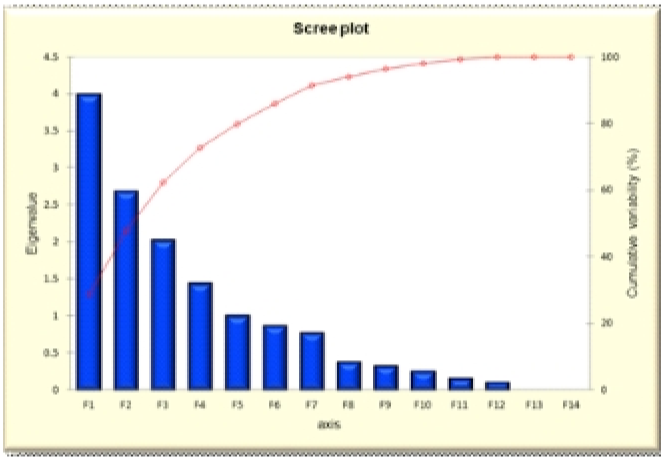


Fig 3. Scree plot showing the eigen values of principal components in PCA of maize (*Zea mays* L.).

Table 5. The relative contribution of 14 characters to genetic diversity in 64 maize inbred lines

S. No.	Source	Contribution %	Times ranked 1st
1	Days to 50% anthesis	3.32	66
2	Days to 50% silking	3.76	75
3	Anthesis-silking Interval	7.87	157
4	Plant height (cm)	3.87	77
5	Tassel height (cm)	5.88	118
6	Ear height (cm)	6.58	132
7	Ear length (cm)	4.8	96
8	Ear girth (cm)	4.87	97
9	Days to maturity	5.87	117
10	Number of kernels rows ear^{-1}	8.54	171
11	Number of kernels row^{-1}	9.25	185
12	100-Kernel weight (gm)	8.54	171

13	Shelling (%)	12.32	246
14	Kernel yield plant^{-1} (g)	14.53	291

positive side on both axes, such as PL23101, PL23094, PL23083, PL23093, PL23049, and PL23060 along PCA I, and PL23093, PL23094, PL23102, PL23089, and PL23101 along PCA II, were identified as superior. In the biplot, genotypes like PL23091, PL23094, PL23101, PL23083 located far from the center which may be useful for hybridization programs.

The scree plot graph, constructed using eigen values and component numbers (Table 6), shows that from the 6th component onward, the line becomes nearly flat, indicating that each successive component contributes less to the total variance. In plant breeding, only those principal components with eigenvalues greater than 1 are typically retained, as they explain a significant amount of variance. In this study, components like days to 50 per cent anthesis (4.004), days to 50 per cent silking (2.689), anthesis-silking interval (2.032), plant height (1.438), and tassel height (1.002) have eigenvalues greater than 1. Components with eigenvalues less than 1, including traits such as ear height, ear length, ear girth, days to maturity, and others, account for less variance and are considered less useful. The scree plot of these values is shown in Figure 4.

Table 6. Eigen values, proportion of total variance represented by first six principal components

Sl. No.	Canonical Roots Analysis (P.C.A.)	1 Vector	2 Vector	3 Vector	4 Vector	5 Vector	6 Vector
	Eigen Value (Root)	5.369	1.791	1.572	1.454	1.135	0.867
	% Vac. Exp.	38.353	12.796	11.229	10.384	8.109	6.192
	Cum. Van Exp.	38.353	51.149	62.379	72.763	80.872	87.064
1.	Days to 50% anthesis	0.047	0.305	0.308	0.242	0.520	0.415
2.	Days to 5 0% silking	0.393	0.201	0.150	-0.082	0.142	0.046
3.	Anthesis silking interval	0.406	0.126	0.071	-0.158	0.000	-0.072
4.	Plant height	-0.121	0.255	-0.404	-0.477	0.115	-0.132
5.	Tassel height	-0.200	-0.029	0.576	-0.158	-0.049	0.075
6.	Ear height	-0.045	0.398	0.015	-0.143	0.515	-0.307
7.	Ear length	-0.1S9	0.036	0.535	-0.037	-0.162	-0.535
8.	Ear girth	-0.302	0.200	0.113	-0.200	-0.244	0.364
9	Days to Maturity	-0.388	0.089	-0.191	-0.183	0.080	0.180
10.	Number of kernel rows ear^{-1}	0.331	-0.050	0.125	-0.248	-0.131	0.429
11.	Number of Kernels row^{-1}	0.258	-0.028	0.124	-0.549	-0.003	-0.144
12.	100 Kernel weight	0.152	0.473	-0.096	0.441	-0.227	-0.183
13.	Shelling percentage	-0.381	0.099	0.086	-0.015	0.149	0.051
14.	kernel yield plant^{-1}	-0.017	0.584	0.005	-0.070	-0.496	0.103

The application of Mahalanobis D^2 statistics and Principal Component Analysis (PCA) in this study effectively identified significant genetic diversity among 64 maize inbred lines. Cluster analysis revealed substantial intra- and inter-cluster variation, with clusters IV and V showing unique potential for breeding plant height and ear height as primary contributors to genetic diversity. Genotypes like PL23101 and PL23094 on extreme positive axes of PCA were superior, while combining clusters with high D^2 and PCA identified traits that ensures generation of heterotic hybrids with significant yield improvement. Selecting complementary traits, such as pairing high kernel yield with desirable morphological traits, can enhance hybrid performance and accelerate breeding for improved maize varieties.

LITERATURE CITED

Amin, M.N., Amiruzzaman, M., Ahmed, A., Rohman, M.M and Ali, M.R. 2013. The genetic diversity of white inbred lines of quality protein maize (QPM). *Eco-Friendly Agricultural Journal*. 6(9): 13-17.

Beck, D.L., Vasal, S.K and Crossa, J. 1990.Heterosis

due to their genetic divergence. The PCA reduced the complexity of 14 morphological traits into six principal components, explaining over 87% of the total variability, and identified key traits such as kernel yield plant^{-1},

and combining ability of CIMMYT's tropical early and intermediate maturity maize (*Zea mays* L.) germplasm. *Maydica.* 35 : 279-285.

DESAP. 2022. Annual report. Directorate of Economics and Statistics Andhra Pradesh. https://des.ap.gov.in.

FAO .2023 .Database .Foo dandAgricultureOrganization, United Nations. www.fao.org/publications.

Jaishreepriyanka, R., Sinha, S., Mandal, S.S., Singh, B and Rashmi, K. (2020). Diversity Analysis of Maize Inbred Lines. *International Journal of Current Microbiology and Applied Sciences*. 9(10), 2765-2773.

Lone, R.A. 2017. Genetic diversity in maize (*Zea mays* L.) inbred lines from Kashmir. *International Journal of Pure Applicational Bioscience*. 5 (1): 229-235.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proceedding of National Institute of Science*. 49-55.

Maruthi, R.T and Rani, K.J. 2015. Genetic variability, heritability, and genetic advance estimates in maize (*Zea mays* L.) inbred lines. *Journal of Applied and Natural Science*. 7 (1): 149-154.

Moll, R.H., Lonquist, J.H., Vele zFortuno, J. an dJohnson, E.C. 1965.The relationship of heterosis and genetic divergence in maize. *Genetics*. 52:139-144.

Naushad, A., Turi, S., Shah, S., Ali, S., Rahman, H., Ali, T and Sajjad, M. 2007. Genetic variability for yield parameters in maize (*Zea mays* L.) genotypes. *Journal of Agricultural and Biological Science*. 2(4-5): 1-3.

Panse ,V.G .an dSukhatme ,P.V .1967 .Statistica lmethods for agricultural workers. *ICAR Publications*.

Pearson, K. 1901.On lines planes of closest fit to systems of points in space. *Philosophical Magazine*. 2: 559–572.

Rao, C. R. 1952. Advanced statistical method in biometric research. *John Wiley and Sons.*

Singh, R.K. and Chaudhary, B.D. 1985. Biometrical methods in quantitative genetic analysis. *Kalyani Publishers*.

UN-COMTRADE. 2022. Database. The United Nations commodity trade statistics database. https://comtradeplus.un.org/Publication/ITSY

Viola, G., Ganesh, M., Reddy, S.S and Kumar, C.V.S. (2003). Study on heritability and genetic advances in elite baby corn (*Zea mays* L.) lines. *Progressive Agriculture*. 3(2): 127-128.