The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
20(3): 328-333, 2025 www.thebioscan.com

Genetic characterization of chickpea genotypes using principal component analysis in diallel mating design

[+]Jainendra Pratap, [#]Shiva Nath*, Anupam Tripathi

Department of Genetics and Plant Breeding, ANDUA&T, Kumarganj-224229, Ayodhya (U.P.)

[#] Professor [+] Phd Scholar

DOI: 10.63001/tbs.2025.v20.i03.pp328-333

KEYWORDS
PCA analysis seed yield, chickpea

Received on:

28-05-2025

Accepted on:

19-06-2025

Published on:

22-07-2025

ABSTRACT

Experimental material consisted of 10 parents and 45 crosses, these crosses were developed by using half diallel mating design. 55 genotypes were evaluated in randomized block design with three replications. Crop geometry consisted of four rows per plot, each measuring 4 meters in length, with a 30 x 10 cm gap between each row and plant. Only three of the nine PCs had an Eigen value greater than 1.0 and a variability of 73.56%. While PC2 accounted 20.38% of the total variation and was strongly correlated with plant height, days to 50% flowering, number of primary branches per plant, and seed yield, PC1 contributed 38.83% of the total variation and was positively correlated with days to flowering and days to maturity. Primary branches per plant, secondary branches/plant, pod/plant, harvest index, seed output, days to 50% flowering, and days to maturity are all related to PC3, which accounted for an extra 14.33% of the overall variation. Using a hierarchical clustering technique, 55 genotypes were divided into 6 clusters. For the majority of the parameters examined, the current study showed that the chickpea germplasm exhibited significant genetic variation.

INTRODUCTION

After the common bean, the chickpea (*Cicer arietinum* L.) is the second-largest edible legume produced worldwide. India is the world's top producer of chickpeas (72.0%), followed by the United States, Australia, and Turkey (FAOSTAT, 2020). In India, chickpeas are often grown in the Central and Southern regions (dry and hot), the North West plains (wet and cool), the North East lowlands (humid/wet and cool), the North East Hills (wet and somewhat hot), and the North Hill (dry and cool). However, the increase of irrigated wheat production has resulted in a significant loss of chickpea land in some states, including Punjab, Haryana, Uttar Pradesh, and Bihar, while Madhya Pradesh, Andhra Pradesh, Maharashtra, and Karnataka have added more land for chickpea agriculture. (Arya *et al.*, 2019). With its great popularity and widespread use in India, chickpeas are the most significant rabi crop among all the pulses. More high-quality seed of enhanced cultivars is becoming accessible. to the farmers, and it is among the elements that have improved the chickpea yield in recent years. By choosing superior genotypes that are directly linked to seed yield, chickpea yield can be increased. These genotypes can then be used exclusively in breeding programs to increase grain yield. PCA is a common tool in contemporary data analysis because it provides an easy, non-parametric way to glean pertinent information from jumbled data sets. A number of (potentially) linked variables are converted into a (smaller) number of uncorrelated variables known as principle components by a mathematical process (Muniraja *et al.*, 2011). Recombination breeding must be used to combine the discovered component qualities from different parents, followed by the selection of transgressive segregants, in order to produce effective recombinants.

Breeding is made easier by dividing the different genotypes based on genetic diversity, with each grouping the genotypes according to the level of genetic divergence. One of the most accurate techniques for determining the relative contributions of various traits and constituents to overall variety, measuring the extent of divergence, and selecting genetically varied parental genotypes to produce desired recombinants is cluster analysis. (Vus *et al.*, 2020). Cluster analysis has been used frequently in breeding procedures to evaluate and compare the attributes of various crop genotypes. Cluster analysis was used in earlier research to work with tomato breeding material. (Evgenidis *et al.*, 2011), To analyze parental pair selection and genetic diversity in winter wheat accessions (Khodadadi *et al.*, 2011), Additionally, principal component analysis is used to evaluate the genetic diversity of wheat. (Mecha *et al.*, 2017). In order to help choose potential parental genotypes for hybridization, the latest study examined the genetic diversity of the chickpea genotypes using multivariate analysis. Through programs, such variety may eventually lead to the creation of new high-yielding chickpea genotypes

Material and Method

Present investigation was carried out at GPB farm ANDUAT Kumarganj, Ayodhya, to evaluate the superior chickpea genotype using principal component analysis and hierarchical cluster analysis using 10 parents(BG 3034, BG-362, JG-315, GNG-2207, BG-372, Rajash, JG-11,KWR- 108, GCP-105,PantG-186) and 45crosses procured from crossing of these parents in half diallele fashion during *Rabi* 2023-24.Genotypes assessed using three replications in a randomized block design(RBD). The genotypes were sown under late-sowing conditions during the first week of December. The experimental unit consisted of four rows per plot, each measuring 4 meters in length, with a 30 x 10 cm gap between each row and plant.Meteorological observations including weekly

mean maximum and minimum temperatures, rainfall, relative humidity, and sunshine hours were recorded throughout the crop growth periods during Rabi seasons of 2023-24 under timely (TS), late (LS), conditions. Recommended fertilizer doses (20 N + 40 P2O5 kg/ha) and standard agricultural practices were followed to ensure healthy crop growth for better characterization of parents and Crosses.Morphological triats days to 50% flowering (DFF),sto maturity (DM), No. of Primary branches (NPB),No. of Secondary branches (NSB), Plant Height (PH), Pods/ Plant (PPP),Seed/ Pod (SPP), Seed Index (SI) Biological yield (BI), Harvest Index (HI), Seed yield (SY) have been recorded from Five randomly selected plant.

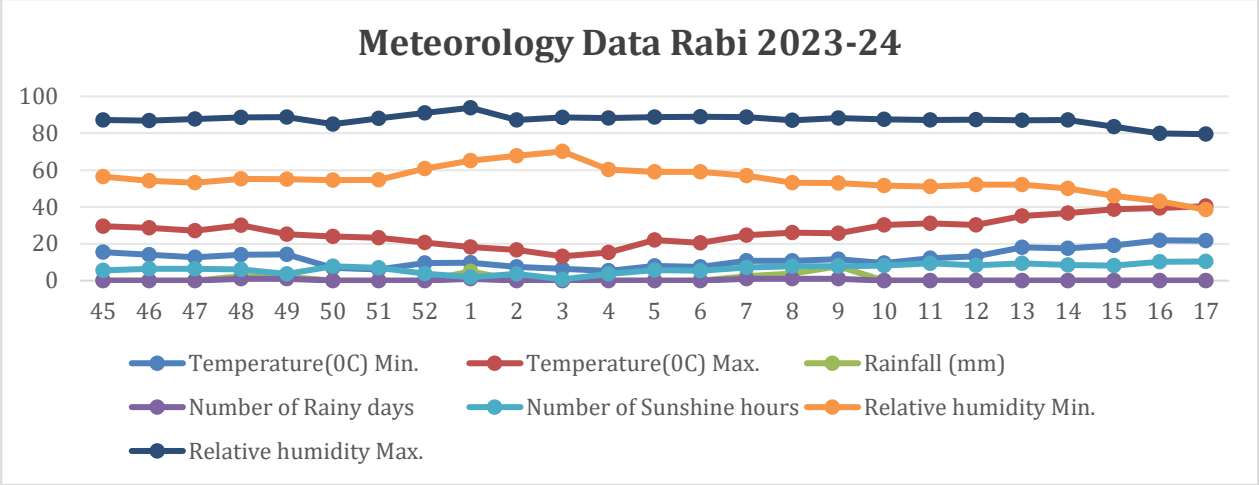

Fig1. Weekly Meteorological data (min and max temperature, rainfall, relative humidity (%), and sunshine hours)

Morphological traits days to 50% flowering (DFF), days to maturity (DM), No. of Primary branches (NPB),No. of Secondary branches (NSB), Plant Height (PH), Pods/ Plant (PPP),Seed/ Pod (SPP), Seed Index (SI) Biological yield (BI), Harvest Index (HI), Seed yield (SY) have been recorded from Five randomly selected plant.

Principal Component Analysis (PCA) was conducted using the correlation matrix approach in XLSTAT version 2024.2.2 to identify the major contributing traits and reduce the dimensionality of the dataset.

C1=b11 (X1)+b2 + …,b1p(Xp)

Where C_1 represents the subject's score on Principal Component 1, which is the first extracted component. The term b_1p denotes the regression coefficient (loading or weight) associated with the p^{th} observed variable used in the construction of Principal Component 1. X_p refers to the subject's actual score or value for the p^{th} variable. The principal component score is computed as a linear combination of standardized observed variables weighted by their respective coefficients.

Result and Discussion

Cluster analysis

Dendrogram of 10 parents and their 45 half diallel crosses was constructed using Ward clustering method (**Ward, 1963**) to evaluate the genotypic variability existing among the clusters **(Fig 2)**.

The intra- and inter-cluster distance analysis **Table 1** highlights the genetic divergence among six clusters derived from multivariate trait data. Cluster 1 exhibited the highest intra-cluster distance (17.65), indicating notable variability within its 43 genotypes, making it a valuable source for selection. Cluster 3 also showed moderate internal variability (23.38), while Clusters 2, 4, and 6 had no internal divergence due to being singleton clusters.

Table 1. Inter-cluster and Intra-cluster distance for 6 cluster in chick pea

	Intra-cluster distance		Inter cluster distance				
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Cluster 1	17.65		29.22	73.85	28.37	124.74	114.78
Cluster 2	0.00	29.22		110.30	75.59	167.34	193.26
Cluster 3	23.38	73.85	110.30		70.49	37.74	41.13
Cluster 4	0.00	28.37	75.59	70.49		124.69	72.21
Cluster 5	0.00	124.74	167.34	37.74	124.69		55.07
Cluster 6	0.00	114.78	193.26	41.13	72.21	55.07	

The maximum inter-cluster distance was observed between Cluster 2 and Cluster 6 (193.26), followed by Cluster 2 and Cluster 5 (167.34), suggesting that genotypes from these clusters are highly divergent and can be used in hybridization programs to generate transgressive segregants. A substantial divergence was also noted between Cluster 1 and Cluster 5 (124.74), and between Cluster 4 and Cluster 5 (124.69).

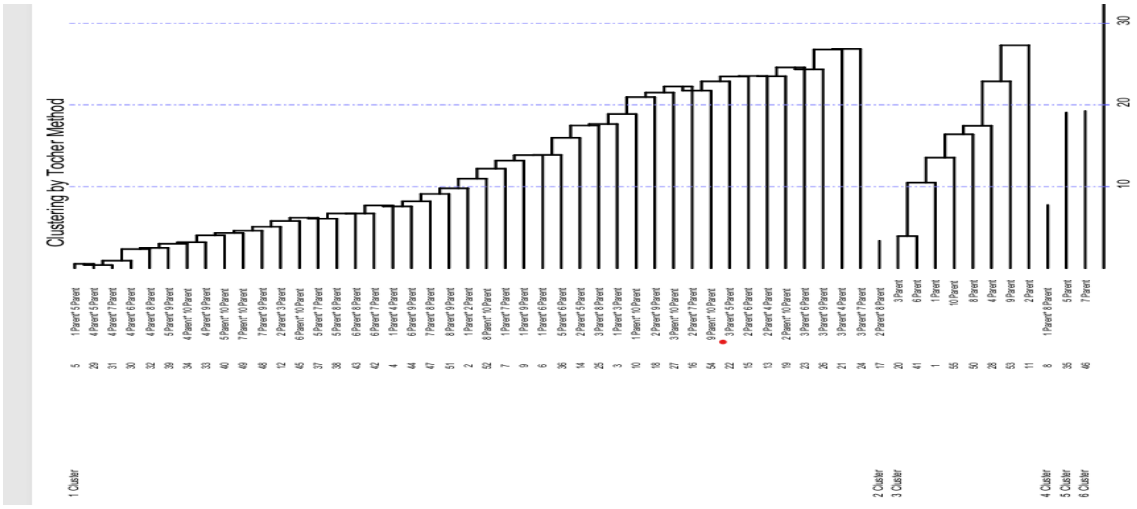

Fig2. Dendrogram showing relationship between parents and half-dialel crosses in field

The lowest inter-cluster distance was recorded between Cluster 1 and Cluster 4 (28.37), indicating genetic similarity and possibly limited benefit from crosses between these two. Cluster 3 appeared moderately divergent from most clusters, especially with Cluster 5 (37.74) and Cluster 6 (41.13), suggesting it harbors intermediate diversity.

This analysis underscores the importance of selecting parents from genetically distant mize heterosis and broadeclusters to maxin the genetic base for future breeding efforts.

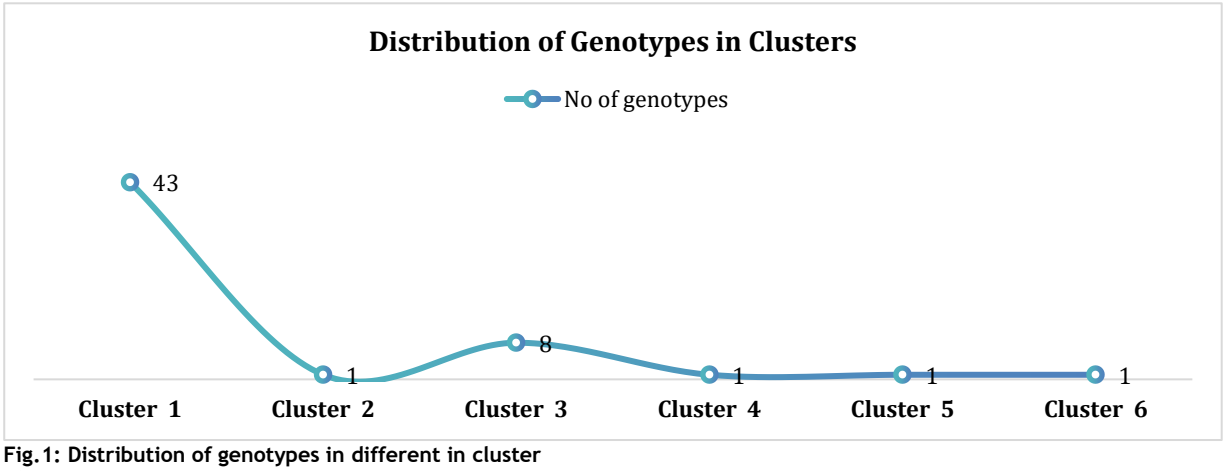

Fig.1: Distribution of genotypes in different in cluster

Cluster mean

The difference of cluster means for different characters can also endorsed an appreciable amount of diversity **(Table 4)**. A cluster analysis involving 55 genotypes grouped into six clusters revealed significant variation in agronomic traits. Cluster 1, the largest group with 43 genotypes, recorded early flowering (50.77 days), moderate maturity (95.53 days), and balanced performance in seed yield (20.99 g) and related traits. Cluster 2, though consisting of only one genotype, exhibited the highest biological yield (55.22 g) and seed yield (23.50 g), indicating high productivity potential.Cluster 3 was characterized by delayed flowering (65.58 days) and maturity (102.96 days) with comparatively lower seed yield (10.69 g), suggesting limited suitability for short-season environments. Cluster 4, another singleton, had the highest number of primary branches (3.47) and a relatively good harvest index (50.58%), making it desirable for branch-related yield traits.Cluster 5 showed the latest flowering (77.00 days), lowest biological yield (15.32 g), and poor seed yield (5.46 g), indicating it may be less promising. Cluster 6 had the least pods per plant (15.11) and seed yield (4.82 g), implying poor reproductive efficiency.

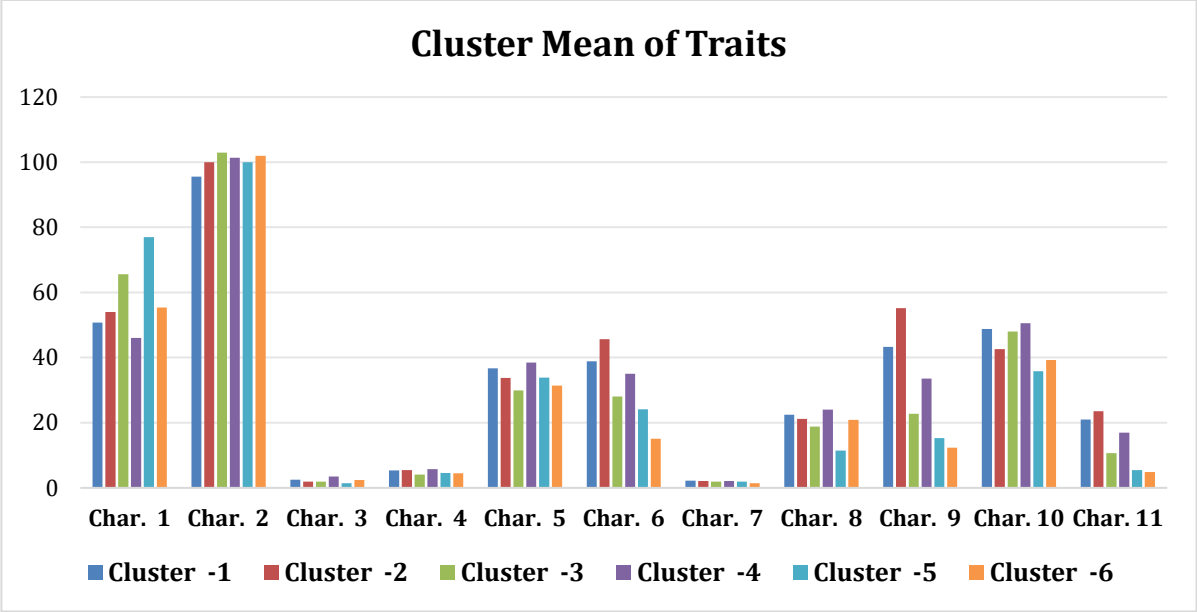

Fig 1: Mean values of cluster for different traits

Overall, clusters 1 and 2 appear promising for yield improvement, while clusters 4 offers superior branching traits.

Table 2 . Cluster means for different characters

Principal component analysis

Principal component analysis (PCA) given in **Table:3** revealed that the first three principal components (PCs) accounted for a cumulative variance of 73.56%, indicating a strong dimensional reduction. PC1 alone contributed the maximum variance (38.84%) with the highest eigenvalue (4.27), followed by PC2 (20.38%) and PC3 (14.34%). These components together effectively summarize the majority of variability in the multivariate dataset. Hence, PC1, PC2, and PC3 can be reliably used for genotype differentiation and multivariate trait interpretation in further analysi. Similar findings were also investigated by (Malik *et al.*, 2014). Eigen values helps to decide that how many variables to retain. The sum of the eigen values is generally equal to the number of variables (Abdi, 2007).

TABLE 3 . Eigen value and commulative variance of different principal component

	PC1	PC2	PC3
Eigene Value (Root)	4.2724	2.24223	1.57712
% Var. Exp.	38.83995	20.38393	14.33748
Cum. Var. Exp.	38.83995	59.22388	73.56136

Inference from correlation between Principal Components and Variables

The principal component (**Table-1)**PC1 had positive correlation with variables Days to 50% flowering (0.35825), Days to maturity (0.30547), while it has a negative correlation with variables biological yield (-0.43243), No. of secondary branch (-0.38721), , pods per plant (-0.37472), seed yield (-0.31781), plant height (-0.30246), 100 seed weight (-0.21885), harvest index (-0.17617), No. of Primary Branches (-0.13847), seeds per pod (-0.10205).The principal component PC2 had positive correlation with variables seed yield (0.27903), plant height (0.46533), Days to 50% flowering (0.05455) and No. of Primary Branches (0.55078) while negative correlation with variables seeds per pod (-0.05584), pods per plant (-0.12), Days to maturity (-0.12569), No. of secondary branch (-0.13837), 100 seed weight (-0.19244), biological yield (-0.20051), harvest index (-0.51922).The principal component PC3 had positive correlation with variables pods per plant (0.41626), harvest index (0.29138), days to maturity (0.28231), seed yield (0.24039), days to 50% flowering (0.20616), No. of primary branches (0.09121) and Secondary branches (0.08833), while negative correlation with variables plant height (-0.00884), biological yield (-0.01909), 100 seed weight (-0.45935), seeds per pod (-0.57816).**(Table 2)**

Table 4. Correlation of different variables with principal component

Canonical Roots Analysis (P. C. A.)	PC1	PC2	PC3
Days to 50% flowering	0.35825	0.05455	0.20616
Days to maturity	0.30547	-0.12569	0.28231
No. of Primary Branches p	-0.13847	0.55078	0.09121
SECONDARY BRANCH	-0.38721	-0.13837	0.08833
PLANT HEIGHT (cm)	-0.30246	0.46533	-0.00884
PODS PER PLANT	-0.37472	-0.12	0.41626
SEEDS PER POD	-0.10205	-0.05584	-0.57816

100 SEED WEIGHT	-0.21885	-0.19244	-0.45935
BIOLOGICAL YIELD	-0.43243	-0.20051	-0.01909
HARVEST INDEX	-0.17617	-0.51922	0.29138
SEED YIELD	-0.31781	0.27903	0.24039

CONCLUSION

The present study revealed significant genetic variation among the chickpea germplasm for most of the traits studied, indicating substantial scope for selection and improvement. The use of multivariate techniques such as cluster analysis and principal component analysis (PCA) proved effective in identifying genetic diversity and structuring the germplasm based on morphological traits.

Cluster analysis facilitated the identification of genetically diverse genotypes, with Cluster 1 exhibiting the highest intra-cluster distance and the greatest inter-cluster distance observed between Cluster 2 and Cluster 6. These findings emphasize the importance of selecting genetically distant parents to maximize heterosis and broaden the genetic base in future breeding programs.

PCA identified three principal components (PCs) with Eigenvalues greater than 1, accounting for 73.56% of the total variability. PC1 explained 38.83% of the variation and was positively associated with days to 50% flowering and days to maturity, suggesting a trade-off between early maturity and yield-related traits. PC2 contributed 20.38% of the total variation and showed strong positive correlations with plant height, number of primary branches per plant, days to 50% flowering, and seed yield. PC3 explained an additional 14.33% of the variation and was associated with yield-related traits such as pods per plant, harvest index, and seed yield.

Table 4 . cluster means for different characters

	No. of Genotypes	Days to 50% flowering	Days to maturity	No. of Primary Branches p	SECONDARY BRANCH	PLANT HEIGHT (cm)	PODS PER PLANT	SEEDS PER POD	100 SEED WEIGHT	BIOLOGICAL YIELD	HARVEST INDEX	SEED YIELD
Cluster 1	43	50.77	95.53	2.49	5.40	36.70	38.82	2.21	22.50	43.29	48.79	20.99
Cluster 2	1	54.00	100.00	1.87	5.47	33.80	45.63	2.13	21.18	55.22	42.56	23.50
Cluster 3	8	65.58	102.96	1.94	4.09	29.95	28.07	1.91	18.78	22.79	48.05	10.69
Cluster 4	1	46.00	101.33	3.47	5.73	38.43	35.03	2.07	24.01	33.56	50.58	16.96
Cluster 5	1	77.00	100.00	1.40	4.60	33.83	24.10	1.87	11.42	15.32	35.80	5.46
Cluster 6	1	55.33	102.00	2.40	4.47	31.40	15.11	1.47	20.87	12.34	39.23	4.82

REFERENCES

- , M., Dwivedi, S. and Chaturvedi, S.K. (2019). Management of biotic stresses in chickpea exploiting host plant resistance. International Journal of Agriculture, Environment and Biotechnology. 12: 141-149.
- Mishra, S.P., Sarkar, U., Taraphder, S., Datta, S., Swain, D., Saikhom, R., Panda, S. and Laishram, M., 2017. Multivariate statistical data analysis-principal component analysis (PCA). *International Journal of Livestock Research*, *7*(5), pp.60-78.
- Abdi, H., 2007. The eigen-decomposition: Eigenvalues and eigenvectors. Encyclopedia of measurement and statistics, pp.304-308.
- Ward, J.H. (1963). Hierarchical groupings to optimize an objective function. Journal of American Statistics Associations. 58: 236-244.
- FAOSTAT, (2020). Agriculture. Available from: http://faostat.fao.org.
- Muniraja, C., Satish, R.G., Raju, C. and Hart, M. (2011). Principal component analysis among genotypes of chickpea (Cicer arietinum L.). International Journal of Agricultural Sciences. 7: 382-386.
- Vus NA, Kobyzeva LN, Bezuglaya ON (2020). Determination of the breeding value of collection chickpea (Cicer arietinum L.) accessions by cluster analysis. Vavilov J. Genet. Breed. 24(3): 244-251.
- Evgenidis G, Traka-Mavrona E, Koutsika-Sotiriou M (2011). Principal component and cluster analysis as a tool in the assessment of tomato hybrids and cultivars. Int. J. Agron. 697879. doi. 10.1155/2011/697879.
- Mecha B, Alamerew S, Assefa A, Assefa E, Dutamo D (2017). Genetic diversity based on multivariate analysis for yield and its contributing characters in bread wheat (Triticum aestivum L.) genotypes. Agric. Res. Technol. 8(5): 555748.
- Khodadadi M, Fotokian MH, Miransari M (2011). Genetic diversity of wheat (Triticum aestivum L.) genotypes based on cluster and principal component analyses for breeding strategies. Aust. J. Crop Sci. 5(1): 17-24.
- Ghafoor, A., Sharif, A., Ahmad, Z., Zahid, M.A. and Rabbani, M.A. (2001). Genetic diversity in blackgram [Vigna mungo L. Hepper]. Field Crops Research. 69: 183-190.
- Malik, S.R., Shabbir, G., Zubur, M., Iqbal, S.M. and Ali, A. (2014). Genetic diversity analysis of morpho-genetic traits in Desi chickpea (Cicer arietinum L.).

International Journal of Agriculture and Biology. 16: 956-960.

- Jayalakshmi, V., Ronald, R.G. and Lakshmanna, K. (2014). Diversity analysis of chickpea germplasm in scarce rainfall zone of Andhra Pradesh, India. Legume Research. 37: 682 684.
- Vishnu, B., Jayalakshmi, V. and Sudha Rani, M. (2020). Genetic diversity studies among chickpea (Cicer arietinum L.) genotypes under rainfed and irrigated conditions for yield attributing and traits related to mechanical harvesting. Legume Research. 43: 190-194.