The Andhra Agric. J 72 (3): 278-284, 2025
doi :10.61657/aaj.2025.209

Principal component analysis and genetic divergence studies for yield and yield-contributing traits in groundnut (*Arachis hypogaea* L.)

K Mukhthambica, Lal Ahamed M, P Janila, K Yogendra, S Hari Kishan and G Rama Rao
Department of Genetics and Plant Breeding, Acharya N G Ranga Agricultural University,
Agricultural College, Bapatla-522101, Andhra Pradesh, India

ABSTRACT

Groundnut is a self-pollinated crop cultivated across diverse agro-climatic regions of world and serves as a source of income and livelihood for millions of farmers. To increase the productivity of groundnut in India, selection and breeding of high yielding cultivars of groundnut is need of the hour. Fifty groundnut genotypes were evaluated at ICRISAT, Hyderabad, during rainy season 2023 using alpha lattice design with three replications. The analysis of variance revealed a significant difference among the genotypes for all the recorded traits. Principal component analysis revealed that the PC1 and PC2 contributed around 80% of total variation in which pod yield/ plot and kernel yield/ plot contributed maximum and is considered as important yield trait for selecting parents in hybridization. In addition, the PCA biplot revealed that there exist positive associations among yield and yield-contributing traits such as pod yield, kernel yield and hundred kernel weight and positive association of hundred kernel weight with shelling percentage. The highest pod yield/ plot and kernel yield/ plot were expressed by the genotype, ICGV 15083 while, the highest shelling percentage was exhibited by JL 24 and hundred kernel weight by ICGV 00440. Cluster analysis the genotypes present in cluster II recorded high mean for pod yield/ plot and kernel yield/ plot while, cluster VIII showed high mean for shelling percentage and cluster I for hundred kernel weight. Therefore, the genotypes identified from PCA analysis and genotypes present in the cluster which recorded high mean values for the yield and yield-contributing traits, can be exploited in groundnut crop improvement programme by selecting them as parents in breeding for high yielding groundnut cultivars.

Keywords: *Cluster analysis, Groundnut, Kernel yield, PCA and Pod yield*

Groundnut (*Arachis hypogaea* L.), also known as peanut, is a self-pollinated crop belonging to the family Leguminosae, believed to have originated at Bolivia (Krapovickas, 1968). It is an allotetraploid with AABB genome, derived from its ancestral species, *A. duranensis* (AA) and *A. ipaensis* (BB) (Bertioli *et al.,* 2019). It is cultivated across diverse agro-climatic regions, including arid and semi-arid areas, where it serves as a source of income and livelihood for millions of farmers. Groundnut kernels are rich source of oil (40-60 %), protein (20-40%) and carbohydrate (10-20%) and is considered as “King of oilseeds”. Furthermore, groundnut kernels also contain mineral elements, vitamins, phytosterols, phenolics and antioxidants (Janila *et al.,* 2013 and Arya *et al.,* 2016). It is predominantly used as edible oil contributing to 25 per cent of vegetable oil. It is also extensively used as a confectionery in food, as well as for feed and fodder purposes. The haulm is used as cattle feed and comprised of protein (8-15%), carbohydrates (38-45%), minerals (9-17%) and lipids (1-3%) (Janila *et al.,* 2016). It is cultivated in more than 120 countries and plays a vital role in the global economy. The global production of groundnut was about 54.27 million metric tonnes from an area of 30.92 million hectares with the productivity of 1.76 tonnes/ ha (FAOSTAT, 2023). India is the second leading country after China in groundnut production, spreading over an area of 4.96 million hectares with a production of 10.3 million metric tons having productivity of 2.1 tonnes/ ha (FAOSTAT, 2023). However, the productivity of groundnut in India is more than the global average productivity but it is less than that of China’s productivity having 4 tonnes/ ha. Thus, there exists a need for increasing productivity of groundnut in India.

Significant variation has been reported among the groundnut genotypes for yield and yield-contributing traits (Yami and Abtew, 2025), which opens the opportunity for selecting high yielding genotypes. Breeding the groundnut genotypes for high yield could result in greater crop productivity. Therefore, the present study was planned to identify the high yielding genotypes.

For analysing huge data sets on yield and yield-contributing traits, Principal component analysis was performed to reduce a multivariate data and to evaluate the significant contribution of each principal component to the overall variance and also the variance of each recorded trait. Moreover, it also helps to know the associations of yield with yield-contributing traits of groundnut and promotes the identification of the best groundnut genotypes for yield and yield-contributing traits. While, the cluster analysis group the groundnut genotypes into clusters, where the genotypes with less genetic divergence/ closely related are grouped together. Cluster analysis also helps in identifying the best genotypes clusters for yield and yield-contributing traits.

MATERIAL AND METHODS

The present investigation was carried out with 47 groundnut genotypes along with three checks (Chico, ICGV03043 and ICGV91114), comprised of both Advanced breeding lines of ICRISAT and released varieties of groundnut (Table 1). The experiment was carried out at 17°50'N, 78°27'E, International Crops Research Institute for Semi-arid Tropics (ICRISAT), Patancheru, Hyderabad, India during Rainy 2023 (R-23) season in alfisols. These genotypes were evaluated using alpha lattice design in three replications and five blocks per replication. Seeds of each genotype were sown in four, 4m-long rows with 30 cm between rows and 10 cm between plants by practicing broad bed and furrow system. Standard agronomic management practices were followed and irrigation was given soon after planting and subsequently whenever required. Gypsum (500 kg ha^{-1}) was applied to the experimental field at flowering stage and protection was taken against insects and diseases as per the guidelines of groundnut production by Janila *et al.* (2018).

Data was collected on Pod yield/ plot (g) (PY), Kernel yield/ plot (g) (KY), Hundred kernel weight (HKW) and Shelling percentage (SH%). Hundred kernel weight (HKW) is the indicator of kernel size and shelling percentage indicates the proportion of the pod that is made up of kernels. The data on hundred kernel weight and shelling percentage was collected using CT scanning of groundnut pods. Computed Tomography (CT) scan is a non-invasive, rapid, and cost-effective method compared to conventional methods for determining the post-harvest data of rainy season (R-23) for shelling percentage and hundred kernel weight. The groundnut pods were scanned using the X-ray system CT portable160.90 ("2D X-ray images" system, X-Ray technology development centre, Fraunhofer Institute of Integrated Circuits, Germany). The resulting X-ray projections were pre-processed ("Image treatment" section) and X-ray transformation (XRT) of images was used to predict the groundnut pod characters, kernel weight and shell weight (Figure 1).

Data was subjected to analysis of variance (ANOVA) using SAS version9.4 Software. Analysis of variance was performed to assess the main and

Table 1. Details of experimental material

Entry no.	Designation	Entry no.	Designation
1	GG20	26	ICGV16688
2	GJG9	27	ICGV16690
3	GJGHPS-1	28	ICGV16697
4	GPBD4	29	ICGV16700
5	ICGV00308	30	ICGV171002
6	ICGV00350	31	ICGV171008
7	ICGV00351	32	ICGV171015
8	ICGV00362	33	ICGV171021
9	ICGV00371	34	ICGV171024
10	ICGV00440	35	ICGV171357
11	ICGV01258	36	ICGV181025
12	ICGV02022	37	ICGV181045
13	ICGV02144	38	ICGV181068
14	ICGV02266	39	ICGV181075
15	ICGV03169	40	ICGV191033
16	ICGV03207	41	ICGV191039
17	ICGV06040	42	ICGV201090
18	ICGV06420	43	ICGV93468
19	ICGV07220	44	ICGV99240
20	ICGV13189	45	ICGV99247
21	ICGV14421	46	JL24
22	ICGV15074	47	Sunoleic95-R
23	ICGV15083	48	Chico (check)
24	ICGV15090	49	ICGV03043(check)
25	ICGV16668	50	ICGV91114(check)

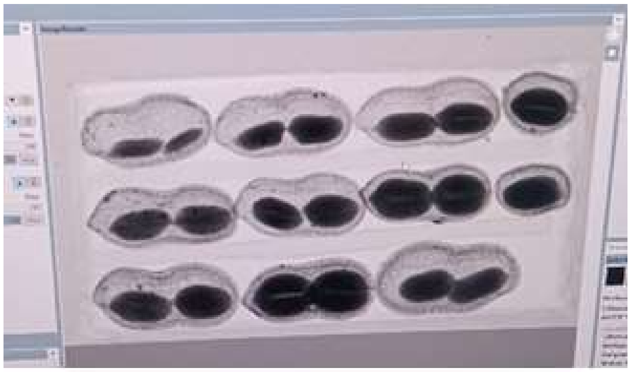

Fig. 1. CT scanned image of groundnut pods

interaction effects for environment, replication, block, and genotypes by considering these factors as random. The following statistical model, linear mixed model was used to estimate the ANOVA.

$$z_{ijk} = \mu + (r)_i + (r/b)_{ij} + g_k + \varepsilon_{ijk}$$

Where,

Z_{ijk} = Performance of the k^{th} genotype in the j^{th} block of i^{th} replication

μ = grand mean;

$(r)_i$ = random effect of replication *i*

$(r/b)_{ij}$ = random effect of block *j* nested with in replication *i*

g_k = random effect of genotype *k*

ε_{ijk} = random residual effect

BLUEs (Best Linear Unbiased Estimates) means of linear mixed model were estimated for each genotype. Principal Component Analysis (PCA) is a widely used dimensionality-reduction technique that condenses a large set of variables (multivariate data) into a smaller set while retaining most of the information contained in the original data. It is used to evaluate the significant contribution of each component to the overall variance. PCA also provides insights into the independent contribution of each trait to the total variance, with each eigen vector coefficient reflecting the amount of the original variable's contribution associated with each principal component. In this study, PCA was performed following the procedures outlined by Massey (1965) and Jolliffe (1986). Principal component biplots were constructed to determine association among traits and to identify the best genotypes for traits using R software.

Hierarchical cluster analysis was done based on average linkage method reported by Xu *et al.* (2021) using R software. Average linkage method of genetic distance D (*r, s*) was calculated by formula,

$$D_{average}\ (\mathrm{r,\ s}) = \frac{1}{Nr\ x\ Ns}\sum_{i=1}^{Nr}\sum_{j=1}^{Ns} d\ (\mathrm{xi, yj})$$

Where,

N_r and N_s are the number of points in cluster r and s, respectively, and d (x_i,yj) is the distance between a point x_i from cluster r and a point y_j from cluster s. The distance between two clusters is the average of the distance between all pairs of points, with one point from each cluster. This average linkage in hierarchical clustering was used to determine the genetic groupings of the test genotypes for multiple traits for its simultaneous selection.

RESULTS AND DISCUSSION

The success of a breeding programme largely depends on the extent of genetic variability present within the population. A broader range of variability enhances the chances of selecting desirable genotypes.

Analysis of variance (ANOVA)

The analysis of variance showed significant differences ($p < 0.05$) among the genotypes for all the recorded traits as presented in the Table 2. Analysis of variance for the traits of pod yield/ plot, kernel yield/ plot, hundred kernel weight and shelling percentage, expressed existence of high amount of variability in the material indicating that these genotypes can be selected for breeding the high yielding genotypes of groundnut. These conclusions were in accordance with the results reported by Rajanna *et al.* (2024), Suthur *et al.* (2023) and Yami and Abtew (2025) in groundnut.

Table 2. Analysis of variance (ANOVA) for yield and yield-contributing traits of groundnut

Source of variation	Df	Pod yield/ plot (g)	Kernel yield/ plot (g)	Shelling percentage	Hundred kernel weight (g)
Replications	2	12053	2674.4	0.005	0.016
Block within Replication	12	1191	1833.5	1.7*	1.5
Genotypes	49	145879**	45704**	8.31**	42.63**
Residual	74	49161	17234	10.45	9.9

** Significant at 1% level, * Significant at 5% level

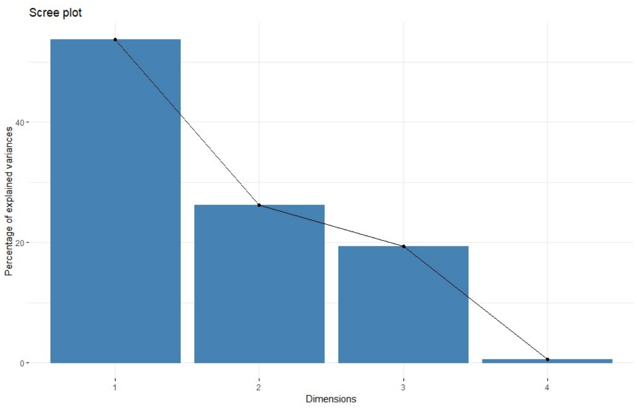

Fig. 2. Scree plot of groundnut genotypes between principal components and its proportion of variation

Principal component analysis

Table 3. Principal component scores, eigen values and variances for yield and yield-contributing traits of groundnut

Traits	PC1	PC2	PC3	PC4
Pod yield/ plot (g)	-0.66	0.08	-0.23	-0.71
Kernel yield/ plot (g)	-0.65	-0.05	-0.28	0.70
Shelling percentage	0.21	-0.80	-0.57	-0.09
Hundred kernel weight (g)	-0.30	-0.60	0.74	-0.03
Eigen values	2.15	1.05	0.78	0.02
Proportion of variation %	53.78	26.2	19.41	0.58
Cumulative variation %	53.78	80.01	99.42	100

Note: PC- Principal component

Principal component analysis for the assessed traits among 50 groundnut genotypes revealed two principal components (PC1 and PC2) with eigen values greater than one (Table 3, Figure 2). PC1 and PC2 together accounts for 80.01% of total variation. The principal component 1 (PC1) and principal component 2 (PC2) showed the variability of 53.78% and 26.23%, respectively, of the total variability. In PC1, pod yield/ plot and kernel yield/ plot showed high component scores while in PC2 and PC3, shelling percentage and hundred kernel weight expressed high component scores and in PC4 pod yield/ plot and kernel yield/ plot has high component scores (Figure 3). It indicates that pod yield/ plot and kernel yield/ plot contributes for much variability than other yield traits thus, it is considered as an important yield trait of groundnut for selecting parents in breeding for high yielding groundnut cultivars.

Smaller angles (<90^0) were present among dimension vectors of pod yield/ plot, kernel yield/ plot and hundred kernel weight and between shelling percentage and hundred kernel weight while, large angles (>90^0) were observed for shelling percentage with pod yield/plot and kernel yield/ plot. Smaller angles between dimension vectors revealed positive association among pod yield/ plot, kernel yield/ plot and hundred kernel weight and between shelling percentage and hundred kernel weight. Large angle expresses negative association of shelling percentage with pod yield/plot, kernel yield/ plot.

Genotypes that are best for a particular trait were plotted closer and furthest to the vector line. The highest pod yield/ plot and kernel yield/ plot was expressed by the genotype, ICGV 15083 (G23) while, the highest shelling percentage (SH%) was exhibited by JL 24 (G46) and hundred kernel weight (HKW) by ICGV 00440 (G10). It indicates that these identified genotypes showed high performance for the respective traits and can be selected in groundnut breeding programme for the genetic improvement of these traits. In addition, these genotypes may also contribute to increase in pod and kernel yields of groundnut as they are positively associated with pod and kernel yields of groundnut. Similar results were reported by Mubai *et al.* (2020), Abady *et al.* (2021), Sukrutha *et al.* (2023), Rajput *et al.* (2024), Danalakoti *et al.* (2025) and Yami and Abtew (2025) in groundnut.

Cluster analysis

Understanding the nature and extent of genetic divergence is essential for the effective planning of hybridization programmes. Studies on genetic divergence are particularly important for assessing variability among the genotypes, enabling breeders to identify and select suitable parents for breeding initiatives. In the present study, 50 genotypes were grouped into 8 clusters based on divergence analysis at the genetic distance of 150 (Table 4, Figure 4). Cluster IV was the largest with 14 genotypes. While clusters I, II, III, V, VI and VII has 4, 3, 9, 4, 6, 2 and 8 genotypes, respectively.

From the table (Table 5), the genotypes present in the cluster II recorded high mean for pod yield/ plot and kernel yield/ plot (1253.19g and 689.24g) while, cluster VIII showed high mean for shelling percentage (72.79%) and cluster I for hundred kernel weight (43.39g). In crop improvement programme, the genotypes present in cluster II can be selected as parents in breeding for high pod yield

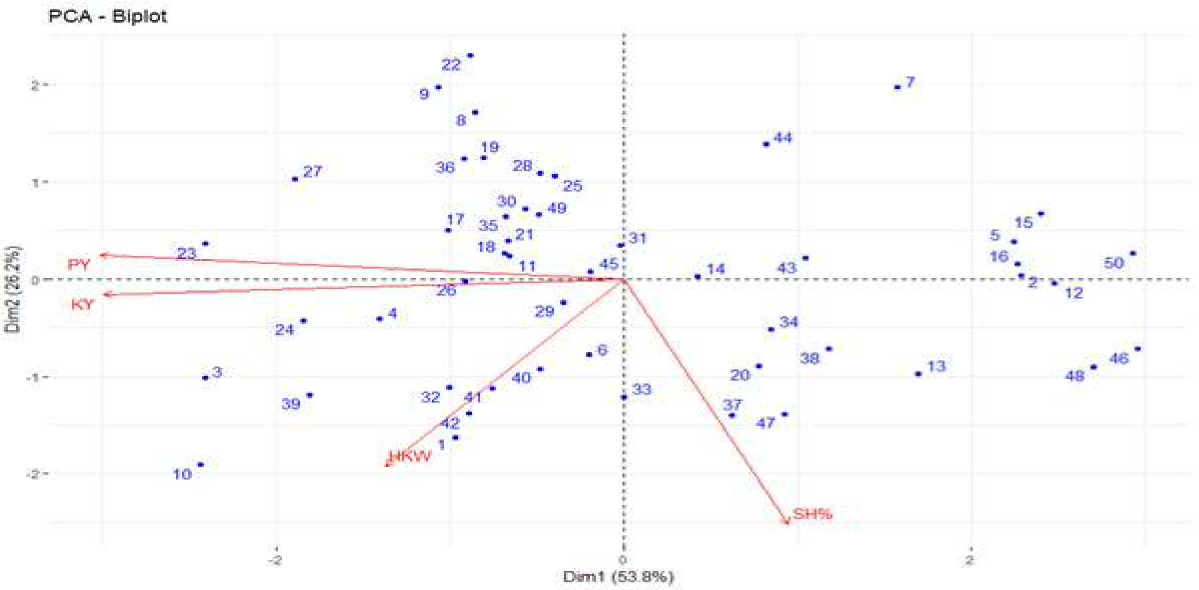

Fig. 3 Two dimensional biplot of Principal Component Analysis in groundnut

Table 4. Distribution of groundnut genotypes into different clusters

Name of Cluster	No.of Genotypes	Genotypes
I	4	GPBD 4, ICGV 15090, GJGHPS-1, ICGV 00440,
II	3	ICGV 16690, ICGV 15083, ICGV 181075
III	9	ICGV 99247, ICGV 171008, GG 20, ICGV 03043, ICGV 16668, ICGV 16700, ICGV 00350, ICGV 16697, ICGV 171002
IV	14	ICGV 00371, ICGV 181025, ICGV 06040, ICGV 16688, ICGV 191039, ICGV 06420, ICGV 171015, ICGV 201090, ICGV 14421, ICGV 191033, ICGV 00362, ICGV 07220, ICGV 01258, ICGV 15074
V	4	ICGV 171024, ICGV 181068, ICGV 181045, ICGV 93468
VI	6	ICGV 13189, ICGV 171021, Sunoleic95-R, ICGV 171357, ICGV 02266, ICGV 99240
VII	2	ICGV 00351, ICGV 03207
VIII	8	ICGV 03169, Chico, ICGV 91114, ICGV 02022, GJG9, ICGV 00308, ICGV 02144, JL 24

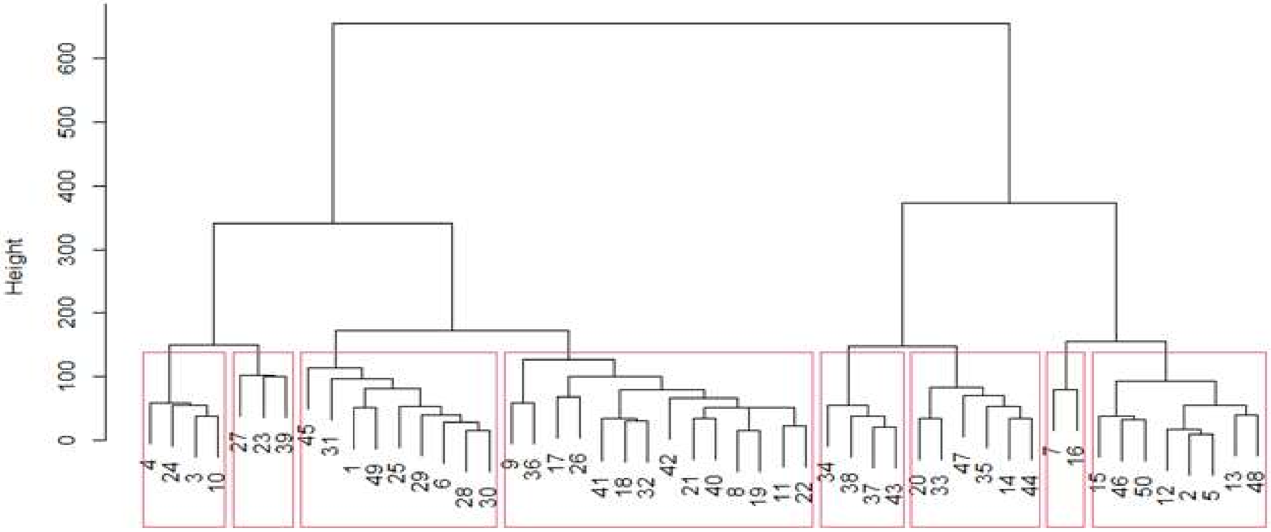

Fig. 4. Hierarchical clustering using average method showing grouping of 50 groundnut geno types based on the yield and yield-contributing traits

Table 5. Cluster mean values for yield and yield-contributing traits in groundnut genotypes

Character	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI	Cluster VII	Cluster VIII
Pod yield/ plot (g)	1118.75	1253.19	818.03	961.03	478.69	607.3	337.82	201.48
Kernel yield/ plot (g)	653.11	689.24	435.85	503.31	266.89	332.72	189.56	118.77
Shelling percentage	71.55	71.06	69.80	69.85	71.75	70.61	69.63	72.79
Hundred kernel weight (g)	43.39	34.89	36.95	34.68	39.22	38.54	27.88	31.98

and kernel yield and cluster VIII for high shelling percentage and cluster I for high hundred kernel yield. Moreover, selecting the genotypes from these clusters contribute for increasing pod and kernel yields of groundnut as they also positively contribute for increasing pod and kernel yield of groundnut. It is similar with the earlier reports of Mubai *et al.* (2020), Abadya *et al.* (2021), Roy *et al.* (2021), Yami and Abtew (2025) and Sonaniya *et al.* (2025) in groundnut.

CONCLUSION

Groundnut genotypes, ICGV 15083, JL 24 and ICGV 00440 and genotypes present in the cluster I, II and VIII which recorded high mean values for the yield and yield-contributing traits, can be exploited in groundnut crop improvement programme by selecting them as parents in breeding for high yielding groundnut cultivars.

LITERATURE CITED

Abadya S, Shimelis H, Pasupuleti J, Mashilo J, Chaudhari S and Manohar S S 2021. Assessment of the genetic diversity of groundnut (*Arachis hypogaea* L.) genotypes for kernel yield, oil and fodder quantity and quality under drought conditions. *Crop Science*. 61(3): 1926-1943.

Arya S S, Salve A R and Chauhan S 2016. Peanuts as functional food: A review. *Journal of Food Science and Technology.* 53: 31–41.

Bertioli D J, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, Leal-Bertioli S C M, Ren L, Farmer A D, Pandey M K and Samoluk S S 2019. The genome sequence of segmental allotetraploid peanut (*Arachis hypogaea* L.). *Nature Genetics*. 51(5): 877–884.

Danalakoti K, Dubey N, Darvhankar M, Avinashe H, Kumar D M, Sharadhi G P and Ghosh S 2024. Assessment of groundnut (*Arachis hypogaea* L.) cultivars on yield and yield contributing traits using principal component analysis. *Agricultural Science Digest-A Research Journal, Of.* 7 (4): 23-34.

FAOSTAT 2023. FAOSTAT production statistics, food and agriculture organization, Rome, Italy.

Janila P, Nigam S N, Pandey M K, Nagesh P and Varshney R K 2013. Groundnut improvement: Use of genetic and genomic tools. *Frontiers in Plant Science*. 4: 23.

Janila P, Pandey M K, Manohar S S, Variath M T, Nallathambi P, Nadaf H L, Sudini H and Varshney R K 2016. Foliar fungal disease resistant introgression lines of groundnut (*Arachis hypogaea* L.) record higher pod and haulm yield in multilocation testing. *Plant Breeding.* 135(3): 355-366.

Janila P, Manohar S S, Deshmukh D B, Chaudhari S, Papaiah V and Variath M. T 2018. *Standard Operating Procedures for Groundnut Breeding and Testing.* Documentation. ICRISAT.

Jolliffe I T 1986. *Principal Component Analysis.* 2nd edition. Springer-Verlag, New York Inc., 487.

Krapovickas A 1968. Origen, variabilidad y difusion del maní (*Arachis hypogaea*). *Actas y Memorias Congress International Americanistas (Buenos Aires)*. 2: 517–534.

Massey W F 1965. Principal components regression in exploratory statistical research. *Journal of theAmerican Statistical Association.* 60: 234–246.

Mubai N, Sibiya J, Mwololo J, Musvosvi C, Charlie H, Munthali W, Kachulu L and Okori P 2020. Phenotypic correlation, path coefficient and multivariate analysis for yield and yield-associated traits in groundnut accessions. *Cogent Food & Agriculture*. 6(1): 1823591.

Rajanna B, Divya K, Lokesh G Y, Shivakumar K and Sreedevi S C 2024. Genetic variability studies in groundnut (*Arachis hypogaea* L.) germplasm. *International Journal of Advanced Biochemistry Research.* 8(8S): 252-256. DOI: 10.33545/26174693.2024.v8.i8Sd.1778

Rajput L S, Sikarwar R S, Ingle S V, Tripathi M K, Prajapati P K and Yadav R K 2024. Principal component analysis for groundnut yield and seed attributes (*Arachis hypogaea* L.). *International Journal of Research in Agronomy*. 194-197.

Roy A, Lal A M, Babu J D B, Amaravathi Y, Viswanath K and Sreekanth B 2021. Studies on divergence in groundnut (*Arachis hypogaea* L.). *Asian Journal of Microbiology, Biotechnology and Environmental Science.* 23(3): 434-439.

Sonaniya P, Sikarwar R S, Sharma S, Solanki R, Jamnotia C and Shruti P 2025. Genetic diversity and clustering pattern in groundnut (*Arachis hypogaea* L.) for yield traits using D² analysis. *Journal of Experimental Agriculture International.* 47 (9): 381-391.

Sukrutha B, Reddy C K K, Madhuri K N and Reddy C B R 2023. Principal component analysis and path coefficient analysis for groundnut yield and seed quality attributes (*Arachis hypogaea* L.). *Legume Research.* 1(7): 1-8.

Suthur D M, Patel D G, Patel P R and Soni N V 2023. Studies on variability for yield and yield attributing traits in groundnut (*Arachis hypogaea* (L.)). *Biological Forum – An International Journal.* 15(11): 520-525.

Xu N, Finkelman R B, Dai S, Xu C and Peng M 2021. Average linkage hierarchical clustering algorithm for determining the relationships between elements in coal. *ACS omega.* 6(9): 6206-6217.

Yami A S and Abtew W G 2025. Assessment of genetic variability for yield and yield contributing traits in groundnut (*Arachis hypogaea* L.) genotypes. *Journal of Food Quality*. 2025(1): 3370389.

Received on 24.08.2025 and Accepted on 19.09.2025