RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 1: 41-45 (January 2026)

Multivariate Analysis in Groundnut (*Arachis hypogaea* L.) Germplasm Accessions for Yield Traits by Principle Component Analysis

K. Hariharan[1], K. Satheeskumar[2], R. Mahendran[1],
S. Thirugnanakumar[1], Mohammed Ashraf[3], J. Vanitha[1]

10.18805/IJARe.A-6395

ABSTRACT

Background: The present study was conducted using groundnut germplasm lines obtained from ICRISAT, Hyderabad to classify 93 groundnut germplasm lines along with 5 checks using Principal component analysis (PCA) for yield traits.

Methods: Using version 4.2.2 of R studio software, nine principle components had been extracted of which first six principal components showed over 93.3% of variation.

Result: Characters were shown in the first two components biplot were plant dry weight, 100 seed weight, 100 pod weight, 100 shell weight, fresh weight, kernel color, shelling percentage, days to 50% flowering and plant height. Nine PC scores that had for almost 100% of the variation were used to perform cluster analysis, which divided the 98 accessions into eight groups. This allows the selection of superior genotypes from each cluster for various yield traits in the groundnut crop improvement.

Key words: Clusters, Germplasm, Groundnut, Principal component analysis.

INTRODUCTION

Groundnut is an annual legume crop that is primarily produced for its easily digestible protein (22-30%) and high-quality edible oil (44-56%) in the seeds. It is distinguished by the high concentrations of linoleic and oleic acids (Yol *et al.*, 2017). It contains oil, protein, fiber, vitamins B complex, K, E and minerals (calcium, magnesium, phosphorus, Iron and zinc) and carbs (10-25%). The legume crop roots and haulm (above ground vegetative components) contribute organic matter and nitrogen (100-152 kg/ha) to the soil, improving soil fertility. In the feed and fertilizer industries, groundnut shell is also utilized as fuel, animal feed, cattle litter and filler material. (Janila *et al.*, 2016).

The development of superior hybrids with high yield in groundnut cultivation is very difficult for all the breeders. Hence, it is essential to identify the presence of genetic variability for selection in groundnut crop improvement to develop the superior varieties. The effective utilization of existing crop variability aims to identify and select superior genotypes with desirable traits from a diverse range of breeding materials. Principal component analysis is a method for estimating variability that reduces data dimensionality through covariance analysis among factors. PCA decreases the variable count by identifying new components that are linear combinations of the primary variables. The analysis assesses the net impact of each variable on the overall variance of the data set and subsequently extracts the maximum achievable variance from the provided material. This study involved principal component analysis of ninety eight groundnut accessions to estimate variability through genotype classification.

[1]Department of Genetics and Plant Breeding, SRM College of Agricultural Sciences, SRM Institute of Science and Technology, Baburayanpettai, Chengalpattu-603 201, Tamil Nadu, India.
[2]Department of Basic Sciences, SRM College of Agricultural Sciences, SRM Institute of Science and Technology, Baburayanpettai, Chengalpattu-603 201, Tamil Nadu, India.
[3]Department of Agronomy, SRM College of Agricultural Sciences, SRM Institute of Science and Technology, Baburayanpettai, Chengalpattu-603 201, Tamil Nadu, India.

Corresponding Author: J. Vanitha, Department of Genetics and Plant Breeding, SRM College of Agricultural Sciences, SRM Institute of Science and Technology, Baburayanpettai, Chengalpattu-603 201, Tamil Nadu, India. Email: vanithajrm@gmail.com

How to cite this article: Hariharan, K., Satheeskumar, K., Mahendran, R., Thirugnanakumar, S., Ashraf, M. and Vanitha, J. (2026). Multivariate Analysis in Groundnut (*Arachis hypogaea* L.) Germplasm Accessions for Yield Traits by Principle Component Analysis. *Indian Journal of Agricultural Research*. **60(1):** 41-45. doi: 10.18805/IJARe.A-6395.

Submitted: 02-04-2025 **Accepted:** 14-08-2025 **Online:** 01-09-2025

MATERIALS AND METHODS

Five check varieties (TMV 13, TMV8, TMV 6, VRI 12 and TMV7) and ninety-three test genotypes from ICRISAT, Hyderabad, were used in the experiment. During *Kharif* 2024, these genotypes were cultivated in an augmented design in SRM College of Agricultural Sciences. Four blocks were used to repeat each check. The distance between genotypes was 30 × 10 cm. Under irrigated conditions, all normal procedures were followed. And assessed for multivariate analysis across 9 characteristics in 98

accessions. Six plants were used at random for each entry in each replication to record the following traits *via* days to 50% flowering (DFF), plant height (PH), fresh weight (PFW), dry weight (PDW), 100 pod weight (HPW), 100 seed weight (HSW), kernel color (KC), shelling percentage (SP) and 100 shell weight (HSW). Principal component and cluster analysis was carried out using, R software version 4.2.2 was used.

RESULTS AND DISCUSSION

Eigen values and eigen vectors were computed using the principal component analysis of ninety-eight accessions using a correlation matrix. Table 1 presents the eigenvalues along with the cumulative percentage of variation indicated by the eigen vectors. In the original data units, the first six principle components contributed 93.3% to the total variation. PC1 may represent overall yield potential. PC2 could indicate quality or compositional traits (*e.g.*, kernel colour or seed size/weight). This separation could be helpful in genotype selection, where high PC1 and moderate PC2 scores identify high yielding varieties. PC1 appears to capture variation related to traits such as PFW, HPW, PDW, HSHW, since these have the highest positive loadings. PC2 reflect traits like seed weight and kernel characteristics (HSW, KC), with strong negative loadings. SP (shelling percentage) and KC (kernel colour) contribute negatively, possibly indicating contrast or trade-offs with yield components. These findings are in agreement with Vishnuprabha *et al.* (2024) and Sukrutha *et al.* (2023), who highlighted pod and kernel traits as primary contributors in groundnut diversity studies.

Table 1: Eigen analysis of correlation matrix.

Traits	PCA 1	PCA 2
DFF	0.16	0.18
PH	0.16	0.04
PFW	0.44	0.14
HPW	0.44	-0.28
HSW	0.26	-0.62
PDW	0.47	0.17
KC	-0.08	-0.65
HSHW	0.45	0.02
SP	-0.26	-0.17

PC- Principal component.

Fig 1 illustrates the scree graph, depicting the variation defined by the different principal components associated with cattle. The first principal components eigenvalue represented 36.7% of the total variance. The second to sixth principal components, which comprised 15.2%, 14.6%, 10.4%, 9.7%, 6.7% and 4% of the total, respectively. The scree plot supports the use of 2 to 3 principal components for further analysis, depending on the desired trade-off between simplicity and information retention. Beyond PC3, the marginal gain in explained variance is minimal, indicating those dimensions can likely be excluded without significant loss of important data structure.

The relative weight given to the variable in each of the nine primary components is represented by the eigenvector, which has been calculated so that the greatest element in each vector is set to one. The orientation of vector is depends on the value of variance and covariance of the principle component. The aim of principal component analysis is to develop new variables that account for 90% of the total variance in the dataset (Vishnuprabha *et al.*, 2025 and Sudhir *et al.*, 2010).

The study incorporates nine quantitative traits in the PCA analysis. Nine Principal Components (PCs) are utilized to decompose the entire dataset. Table 2 indicates that the first three principal components (PCs) possess eigenvalues greater than 1, collectively accounting for 66.5% of variability. PC1, possessing an eigenvalue of 1.81, accounted for 36.7% of the total variability. According to PC1, DFF (0.16) and PH (0.16) demonstrate the strongest positive correlation. Following this, the values are PDW (0.47), HSHW (0.45), PFW (0.44), HPW (0.44) and HSW (0.26). In contrast, PC1 exhibits a significant negative correlation with KC (0.08) and SP (0.26). Similarly it was reported by Dama, (2022) for the PC1 eigenvalue.

PC2 has an eigenvalue of 1.16 and represents 15.2% of the total variability, as shown in Table 2. PC2 exhibits a

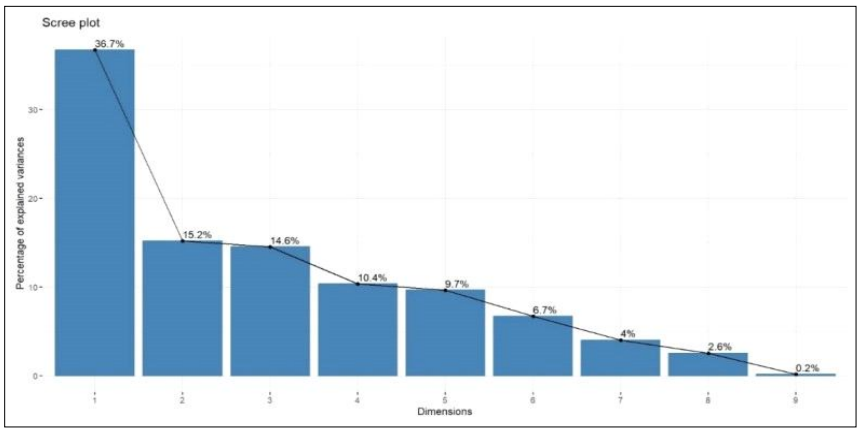


Fig 1: Cattle scree graph for variation explained by various principal components on groundnut germplasm.

strong positive correlation with DFF (0.18), followed by PH (0.04), PFW (0.14), PDW (0.17) and HSHW (0.02). There exists a strong negative correlation with attributes including HPW (-0.28), HSW (-0.62), KC (-0.65) and SP (-0.17), beside a weak negative correlation with PH (-0.086). Similarly it was reported by Mubai *et al.* (2020) and Danalakoti *et al.* (2024).

Biplot analysis

Fig 2 illustrates the biplot of PCA, classified into four quadrants, with Dimension 1 on the X-axis and Dimension 2 on the X-axis, accounting for the maximum variance. The biplot demonstrates the variance between traits and genotypes, with a colour indicating the degree of this variance. The green colour represents high variance, whereas yellow, brown and black denote moderate and low variance, respectively. Quadrant 1 (Q-1) is located directly above the centroid and exerts a positive effect on both Dimension 1 and Dimension 2. Seventeen genotypes, specifically 1, 75, 82, 54, 63, 25, 12, 4, 78, 60, 41, 88, 72, check 1, check 2, check 3 and check 5 in Q-1, exhibit a significant correlation with PDW, subsequently followed by HSHW, HSW and PH. Mekonen, (2022), Kohar *et al.* (2023) and Isler *et al.* (2023) reports similar findings by constructing a biplot with groundnut accessions, positioning the hundred seed, plant height and shelling percentage on the positive side of the graph.

Thirty-two accessions includes the genotypes 58, 51, 49, 43, 85, 23, 59, 86, 94, 52, 56, 39, 14, 23, 26, 3, 61, 70, 22, 81, 95, 44, 90, 67, 87, 31, 68, 38, 35, 24, 20 and 42 are classified in quadrant 2 (Q-2), located to the left higher than the centre. This positioning indicates a positive effect on Dim.2 and a negative effect on Dim.1, thus influencing the variance in traits such as SP and DFF. Genotypes 58 and 52 and the trait SP exhibit significant diversity and are suitable for selection purposes.

The third quadrant (Q-3) is located to the left of the centre and has a negative impact on both Dimension 1 and Dimension 2. Twenty-five genotypes are categorized within this quadrant, including 77, 92, 27, 15, 69, 89, 36, 91, 18, 58, 84, 62, 8, 48, 6, 46, 17, 9, 2, 19, 66, 93, 10, 79 and check 4.

In quadrant 4 (Q-4), directly beneath the centroid, Dim.1 exhibits positive effects while Dim.2 shows negative effects. The variables PFW and HPW demonstrate significant correlations with 24 genotypes: 7, 28, 47, 37, 30, 64, 40, 50, 76, 5, 80, 13, 45, 57, 34, 32, 88, 55, 21, 83, 11, 71 and 73. The contribution of variance is substantial, with values of 73, 83 and 21 observed for genotypes and HPW in Q-4.

Table 2: Correlation matrix and its contribution of traits to principal components.

PCs	PCA 1	PCA 2	PCA 3	PCA 4	PCA 5	PCA 6	PCA 7	PCA 8	PCA 9
Eigen value	1.81	1.16	1.14	0.96	0.92	0.77	0.60	0.47	0.12
Proportion of variance (%)	36.70	15.20	14.60	10.40	9.70	6.70	4.00	2.60	0.20
Cumulative variance per cent	36.70	51.90	66.50	76.80	86.50	93.20	97.30	99.80	100.00

DFF- Days to fifty percentage flowering, PH-Plant Height, NPB- No of primary branches, NSB- No of secondary branches, LS- Leaf size, FMA- Flower on main axils, PFW- Plant fresh weight, HPW- 100 Pod weight, HSW-100 seed weight, PDW- Plant dry weight, KC-Kernel colour, KS- Kernel shape, HSHW- 100 shell weight, SP- Shelling percentage.

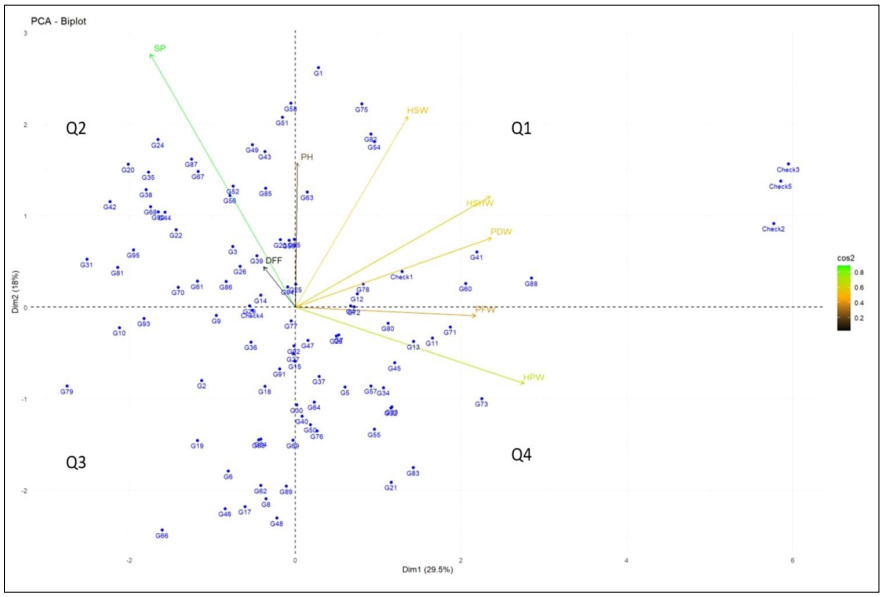


Fig 2: Biplot analysis of PC1 and PC2 in contribution of variance to traits and genotypes of study.

Mamun *et al.* (2022) and Sukrutha *et al.* (2023) conducted a principle component analysis revealing a significant association between hundred kernel weight, alongside plant height's contribution to total variability.

The biplot representation clearly discriminated the genotypes and revealed that PFW, PDW and NPPP are the most important variables in determining diversity. Hence, these traits can be used to select genetically diverse and superior genotypes. Similar findings were reported by Vishnuprabha *et al.* (2024) and Sukrutha *et al.* (2023), who also identified pod and kernel traits as major contributors to diversity.

Ward's hierarchical cluster analysis, utilizing nine principal component scores that accounted for over 90% of total variation, identified eight distinct clusters (Fig 3). The initial cluster consisted of four accessions, including one check variety and one novel germplasm line. The second cluster included 21 accessions, comprising 4 check varieties and 17 new germplasm lines. The third cluster includes twelve new varieties, while the fourth cluster consists of 29 new germplasm lines. The fifth cluster contains 14 new germplasm lines, while the sixth cluster consists of 13 new germplasm lines. The seventh

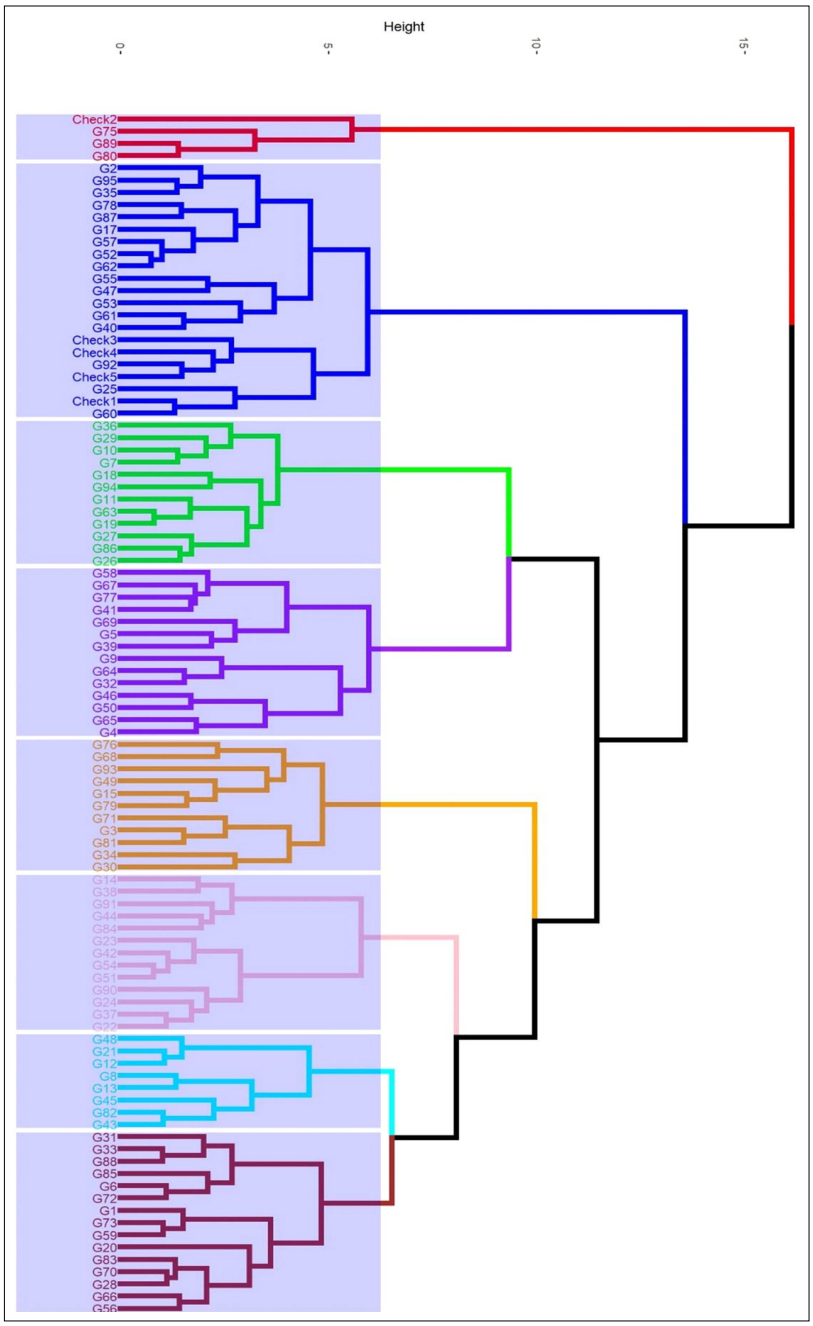


Fig 3: Dendrogram showing ninety eight accessions of groundnut based on the scores of principal components.

cluster comprises 8 accessions, while the eighth cluster contains 15 germplasm accessions. Similarly Upadhyaya *et al.* (2006) evaluated groundnut core collection and identified three clusters through principal component analysis. These results agree with earlier reports by Kumar *et al.* (2024) and Patel *et al.* (2024) who also emphasized the role of pod and yield traits in genetic diversity analysis.

CONCLUSION

The principal component analysis (PCA) conducted on 98 groundnut accessions using a correlation matrix has effectively unraveled the major sources of phenotypic variation and aided in genotype differentiation. The biplot analysis provided a clear visualization of trait-genotype associations, effectively classifying genotypes into four quadrants based on the contribution of traits to variation. Genotypes in Quadrants 1 and 4 showed strong associations with yield components, while those in Quadrant 2 were linked with shelling percentage and flowering duration. Genotypes in Quadrant 3 displayed low contribution to both dimensions, indicating less favorable trait combinations.

Ward's hierarchical clustering based on the principal components further grouped the accessions into eight distinct clusters, suggesting significant genetic diversity within the population. This clustering pattern supports the utility of PCA in identifying genetically diverse and agronomically superior genotypes. Overall, traits such as pod weight, hundred seed weight and plant height emerged as key contributors to variability and should be prioritized in groundnut breeding programs. The alignment of these findings with previous studies reinforces the robustness of PCA as a tool for genetic diversity analysis and genotype selection in groundnut improvement programs.

Conflict of interest

All authors declare that they have no conflict of interest.

REFERENCES

Dama, D.B. (2022). Genetic variability and association among agromorphological and quality traits in groundnut (*Arachis hypogaea* L.) genotypes in Eastern Ethiopia (Doctoral dissertation, Haramaya University).

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*. doi: 10.18805/ag.D-5805.

Isler, N., Sahin, C.B., Yildiz, R. and Yilmaz, M. (2023). Comparison of quality and yield components of peanut market types using PCA. *KSU. J. Agric Nat*. **26(3):** 610-618.

Janila, P., Pandey, M.K., Shasidhar, Y., Variath, M.T., Sriswathi, M., Khera, P. and Varshney, R.K. (2016). Molecular breeding for introgression of fatty acid desaturase mutant alleles (ahFAD2A and ahFAD2B) enhances oil quality in high and low oil containing peanut genotypes. *Plant Science*. **242:** 203-213.

Kohar, G.R. and Yadav, D.R. (2023). Evaluation of groundnut genotypes for yield and yield attributing traits through genotype by trait biplot analysis. *International Journal of Recent Advances in Multidisciplinary Topics*. **4(1):** 33-37.

Kumar, S., Sahi, V.P., Choudhary, S. and Singh, A.K. (2024). Assessment of the genetic diversity in groundnut (*Arachis hypogaea* L.) genotypes for yield and its attributing traits using D2 statistics. *Plant Cell Biotechnology and Molecular Biology*. **25:** 61-70.

Mamun, A.A., Islam, M.M.I., Adhikary, S.K. and Sultana, M.S. (2022). Resolution of Genetic Variability and Selection of Novel Genotypes in EMS Induced Rice Mutants Based on Quantitative Traits Through MGIDI. *International Journal of Agriculture and Biology*. 1560-8530. doi: 10.17957/IJAB/15.1957.

Mekonen, G.S. (2022). Phenotypic and Genotypic Variability of Bambara Groundnut (*Vigna subterranean* L.) Accessions at Tepi South west, Ethiopia (Doctoral dissertation, Jimma University)

Mubai, N., Sibiya, J., Mwololo, J., Musvosvi, C., Charlie, H., Munthali, W., Kori, P. (2020). Phenotypic correlation, path coefficient and multivariate analysis for yield and yield-associated traits in groundnut accessions. *Cogent Food and Agriculture*. **6(1)**. https://doi.org/10.1080/23311932.2020.1823591.

Patel, M.K., Tiwari, D., Sharma, V., Singh, D. and Kumar, S. (2024). Unraveling diversity and character association in sesame (*Sesamum indicum* L.) using different agro-morphological traits. *Environ Ecol*. **42(1):** 109-115.

Sudhir, K.I., Marappa, N. and Govindaraj, M. (2010). Classification of new germplasm and advanced breeding lines of groundnut (*Arachis hypogaea* L.) through principal component analysis. *Legume Research*. **33(4):** 242-248.

Sukrutha, B., Reddy, C.K.K., Madhuri, K.V.N., Reddy, C.B.R., Kumar, A.R.N., Vemireddy, L.N. and Akkareddy, S. (2023). Principal component analysis and path coefficient analysis for groundnut yield and seed quality attributes (*Arachis hypogaea* L.). *Legume Research*. **48(7):** 1096-1102. doi: 10.18805/LR-5075.

Upadhyaya, H.D., Gowda, C.L.L., Pundir, R.P.S., Reddy, V.G. and Singh, S. (2006). Development of core subset of finger millet germplasm using geographical origin and data on 14 quantitative traits. *Genetic Resources and Crop Evolution*. **53:** 679-685.

Vishnuprabha, R.S., Viswanathan, P.L., Manonmani, S., Rajendran, L. and Selvakumar, T. (2024). Development of early maturing varieties in groundnut (*Arachis hypogaea* L.). *Legume Research. An International Journal*. **47(11):** 1870-1874. doi: 10.18805/LR-4784.

Vishnuprabha, R.S., Viswanathan, P.L., Manonmani, S., Rajendran, L. and Selvakumar, T. (2025). Evaluation of variability and principal component analysis in segregating populations of groundnut (*Arachis hypogaea* L.). *Plant Science Today*. **12(1):** 1-10. https://doi.org/10.14719/pst.4025

Yol, E., Ustun, R., Golukcu, M. and Uzun, B. (2017). Oil content, oil yield and fatty acid profile of groundnut germplasm in mediterranean climates. *Journal of the American Oil Chemists Society*. **94(6):** 787-804.