

International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
NAAS Rating (2026): 5.29
IJABR 2026; SP-10(8): 2320-2327
www.biochemjournal.com
Received: 17-05-2026
Accepted: 22-06-2026

HH Maru

Ph.D. Scholar, Department of
Genetics and Plant Breeding,
College of Agriculture, Junagadh
Agricultural University,
Junagadh, Gujarat, India

Dr. RB Madariya

Professor and Head, Department
of Genetics and Plant Breeding,
College of Agriculture, Junagadh
Agricultural University, Mota
Bhandariya, Gujarat, India

Dr. GK Sapara

Assistant Research Scientist,
Main Oilseeds Research Station,
Junagadh Agricultural
University, Junagadh, Gujarat,
India

Dr. CJ Rajani

Assistant Research Scientist,
Main Oilseeds Research Station,
Junagadh Agricultural
University, Junagadh, Gujarat,
India

Dr. CM Godhani

Assistant Research Scientist,
Oilseeds Research Station,
Sardarkrushinagar Dantiwada
Agricultural University,
Dantiwada, Gujarat, India

RM Jambukiya

Ph.D. Scholar (Agri.),
Department of Genetics and
Plant Breeding, College of
Agriculture, Junagadh
Agricultural University,
Junagadh, Gujarat, India

DR Vaghasiya

Ph.D. Scholar (Agri.),
Department of Genetics and
Plant Breeding, College of
Agriculture, Junagadh
Agricultural University,
Junagadh, Gujarat, India

Corresponding Author:**HH Maru**

Ph.D. Scholar, Department of
Genetics and Plant Breeding,
College of Agriculture,
Junagadh Agricultural
University, Junagadh,
Gujarat, India

PCA and D² analysis in seventy genotypes of groundnut (*Arachis hypogaea* L.)

HH Maru, RB Madariya, GK Sapara, CJ Rajani, CM Godhani, RM Jambukiya and DR Vaghasiya

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i8Sz.9745>

Abstract

Seventy groundnut accessions sourced from Main Oilseeds Research Station, Junagadh Agricultural University during Summer-2026 were evaluated utilizing Randomized Block Design with three replications. Individual genotypes were sown in solitary rows measuring 2.0 m length with 30 cm inter-row spacing and 10 cm inter plant spacing. Fourteen quantitative traits were assessed from five plants per genotype per replication, with oil content determined utilizing Nuclear Magnetic Resonance methodology. Principal Component Analysis utilizing singular value decomposition methodology extracted three principal components explaining 69.71% aggregate variance, with PC1 predominantly representing productivity and temporal advancement mechanisms. Mahalanobis D² metric partitioned germplasm into twelve genetically discrete assemblages through Tocher's clustering approach. Harvest index and kernel yield per plant constituted paramount divergence contributors, accounting for 33.74% and 13.75% total variance, respectively. Cluster means exhibited considerable variation across measured characteristics, substantiating substantial genetic material availability for advancement initiatives. Analysis revealed pronounced heterogeneity in cluster magnitude, with majority germplasm demonstrating genetic similarity while subset accessions harboured unique allelic combinations. Genotypes manifesting favourable yield characteristics concurrent with superior quality parameters merit prioritization as parents for hybridization programs targeting heterotic vigour exploitation. Principal Component Analysis and divergence analysis demonstrated dependability as effective instruments for systematic heritable variability assessment and parental combination identification within groundnut improvement programs.

Keywords: Groundnut, Mahalanobis D², Principal Component Analysis, Heterotic Vigour

Introduction

Arachis hypogaea L. represents a significant economic oilseed crop on the international stage, positioning itself as the 13th most important food crop globally. The crop functions as humanity's fourth primary oil resource and third major supplier of plant based proteins. From a botanical perspective, this species constitutes an annual herb within the Fabaceae (Leguminosae) family. The scientific nomenclature originates from Greek terminology, where *Arachis* signifies a pulse crop and *hypogaea* translates to "subterranean," highlighting the distinctive trait wherein pod maturation takes place underground. Chromosomally, groundnut exhibits an allotetraploid configuration ($2n = 4x = 40$) with a fundamental chromosome value of $x = 10$ (Stalker, 1991) [17]. The plant demonstrates cleistogamous pollination mechanisms, which results in predominantly self-fertilizing characteristics (Knauff and Wynne, 1995) [9].

Diverse geographical regions assign different terminologies to this legume. North American regions favor the "peanut," though numerous alternative appellations exist including African nut, Chinese nut, Manila nut, Kipper nut, Hawk's nut, Jar nut, Earth chestnut, Monkey nut, Goober pea, Ground pea and Ground bean (Johnson, 1964) [7]. Within India, the crop receives the traditional name mungphali. Its substantial commercial importance has earned it the prestige title "King of Oilseeds."

This legume thrives across tropical, subtropical and warm temperate zones between approximately 40° N and 40° S latitude bands. Historical evidence from phylogenetic and geographic analyses indicates the species originated in southern Bolivia along with northern Argentina (Kochert *et al.* 1996) [10]. The genus *Arachis* demonstrates its highest genetic variation across South America, particularly throughout western Brazil, Bolivia,

Paraguay and northern Argentina, where wild relatives and ancestral variants are naturally distributed.

The country's low groundnut production is mostly attributable to a lack of adoption of better varieties and uneven performance in a variety of locations, notably rainfed situations. Therefore, the creation of high-yielding cultivars with wide adaptability, fertilizer responsiveness and resistance to biotic and abiotic challenges, together with short seed dormancy, is crucial to increase output. It also has many health benefits including weight gain control (Petersen *et al.*, 2022) ^[13], prevention of cardiovascular diseases (Feldman, 2002) ^[4], protection against Alzheimer's disease (Vaishnavi *et al.*, 2023) ^[18] and cancer inhibition (Wang *et al.*, 2021) ^[19]. Groundnuts contain 21 - 36.4% protein, 18% carbohydrates, 36-54% fat and 100-120 grams of fibre per kg of kernels (Bonku and Yu, 2020) ^[20].

Assessing genetic variability among genotypes is essential in plant improvement programs and both Principal Component Analysis (PCA) and divergence analysis serve this purpose effectively. While PCA reduces multiple inter-correlated traits into a few independent components to identify major contributors to variability, divergence analysis based on the Mahalanobis D² statistic groups genotypes according to their genetic distance from one another. Genetic diversity is of particular importance, as hybrids from genetically dissimilar parents generally exhibit greater heterotic vigor and broader variability in segregating generations than those from closely related parents (Arunachalam, 1981) ^[1]. Together, PCA (Hotelling, 1933) ^[6] and the D² statistic (Mahalanobis, 1936) ^[11] remain

fundamental tools for identifying diverse parental combinations in plant breeding programs.

2. Materials and Methods

The research utilized 70 peanut accessions sourced from the Main Oilseed Research Station at Junagadh Agricultural University, Junagadh procured during the Summer-2026. A Randomized Block Design framework incorporating three replications formed the basis of this experimental setup. Individual genotypes were sown in solitary rows measuring 2.0 m in length with 30 cm intervals between rows and 10 cm spacing between individual plants. Standard agronomic protocols and protective measures against pests and diseases were implemented throughout the cultivation period.

Data collection focused on pod yield per plant alongside various associated parameters including days to 50% flowering, days to maturity, plant height (cm), number of primary branches per plant, number of mature pods per plant, 100-pod weight (g), kernel yield per plant (g), 100-kernel weight (g), shelling percentage (%), sound mature kernel (%), biological yield per plant (g), harvest index (%), oil content (%) and subsequent correlation analysis along with path coefficient investigation. Oil content was determined using Nuclear Magnetic Resonance methodology. Five plants from each genotype per replication were assessed for trait evaluation, with the exception of days to 50% flowering and days to maturity which were measured at the plot level rather than individual plant basis.

$$\text{Shelling percentage (\%)} = \frac{\text{Weight of kernels (g)}}{\text{Weight of pod samples (g)}} \times 100$$

$$\text{Sound mature kernel (SMK\%)} = \frac{\text{Number of well mature kernels}}{\text{Total number of kernels}} \times 100$$

$$\text{Harvest index (\%)} = \frac{\text{Pod yield per plant (g)}}{\text{Biological yield per plant (g)}} \times 100$$

Principal Component Analysis (PCA) was performed to assess the pattern of variation among the genotypes based on the fourteen quantitative traits studied, following the method described by Hotelling (1933) ^[6], using the singular value decomposition approach in R-Studio software. Genetic divergence among the genotypes was estimated using the Mahalanobis D² statistic (Mahalanobis, 1936) ^[11] and the genotypes were subsequently grouped into clusters following Tocher's method as described by Rao (1952).

3. Results and Discussion

3.1 Principal Component Analysis

Principal Component Analysis was first conceptualized by Pearson (1901), who introduced the idea of fitting lines and planes of closest fit to multivariate data and was later developed into its modern statistical form by Hotelling (1933) ^[6], who coined the term "principal components" and formulated the technique as a means of transforming correlated variables into a smaller set of uncorrelated components that successively account for the maximum variance in the data.

3.1.1 Inference from the Variance Distribution in Principal Components: Principal Component Analysis

(PCA) was performed to assess the pattern of variation among the fourteen quantitative traits - pod yield per plant (PYPP), days to 50% flowering (DF), days to maturity (DM), plant height (PH), number of primary branches per plant (PBPP), number of mature pods per plant (MPPP), 100-pod weight (100PW), kernel yield per plant (KYPP), 100-kernel weight (100KW), shelling percentage (SP), sound mature kernel (SMK), biological yield per plant (BYPP), harvest index (HI) and oil content (OC) and to identify the components contributing most substantially to the observed multivariate variability. The analysis was performed using the singular value decomposition (SVD) approach, preferred over the classical eigen-decomposition of the covariance/correlation matrix owing to its greater numerical stability and computational efficiency when handling datasets comprising multiple inter-correlated variables. A total of fourteen principal components were extracted, corresponding to the number of original variables analysed, since PCA yields as many components as there are input variables.

Of these fourteen components, only three exhibited eigenvalues greater than unity. Component retention was based on Kaiser's criterion (Kaiser, 1960) ^[8], according to which only principal components with an eigenvalue

exceeding 1.0 are considered to represent meaningful, non-trivial sources of variation, as such components explain a greater proportion of variance than any single original variable. Accordingly, PC1, PC2 and PC3 were retained for further interpretation, while the remaining eleven components, each accounting for comparatively negligible variance, were excluded from subsequent analysis.

Among the retained components, PC1 explained the highest proportion of total variance (37.54%), corresponding to an eigenvalue of 5.2441, indicating that this axis captured the principal direction of variability among the genotypes evaluated. PC2 and PC3 accounted for an additional 20.72% and 11.45% of the total variance, respectively. Collectively, the first three principal components explained approximately 69.71% of the cumulative variance, indicating that a substantial share of the multivariate variability among the studied traits could be adequately summarized along these three axes, thereby supporting the use of dimensionality reduction for interpreting trait associations within the dataset (Table 1 and Figure 1).

3.1.2 Inference from Correlation between Principal Components and Variables

The trait-loading (correlation) coefficients of the fourteen variables with the first three principal components were examined to determine the relative contribution of each trait to the observed variation pattern and to elucidate the biological basis of each component (Figure 2 and 3).

PC01 which accounted for the largest share of total variance, showed strong positive correlations with harvest index (HI; 0.3701), 100-kernel weight (100KW; 0.3210), shelling percentage (SP; 0.3205), kernel yield per plant (KYPP; 0.2906), pod yield per plant (PYPP; 0.2485), sound mature kernel (SMK; 0.2414), 100-pod weight (100PW; 0.2223), oil content (OC; 0.1292) and mature pods per plant (MPPP; 0.0666). In contrast, negative correlations were recorded with plant height (PH; -0.185), days to maturity (DM; -0.2048), primary branches per plant (PBPP; -0.2862),

biological yield per plant (BYPP; -0.2971) and days to 50% flowering (DF; -0.3686). This pattern suggests that PC1 represents predominantly a yield and earliness associated axis of variation, wherein genotypes exhibiting higher values for yield contributing and kernel quality traits tend to flower and mature earlier, are relatively shorter in stature and invest comparatively less in vegetative biomass and branching.

PC02 exhibited positive loadings for 100-pod weight (100PW; 0.2832), 100-kernel weight (100KW; 0.211) and shelling percentage (SP; 0.0221), while showing negative correlations with days to 50% flowering (DF; -0.0501), sound mature kernel (SMK; -0.0876), harvest index (HI; -0.0912), days to maturity (DM; -0.1624), biological yield per plant (BYPP; -0.1941), plant height (PH; -0.207), primary branches per plant (PBPP; -0.2128), oil content (OC; -0.2531), kernel yield per plant (KYPP; -0.4011), pod yield per plant (PYPP; -0.4262) and mature pods per plant (MPPP; -0.5455). The pronounced negative loadings of MPPP, PYPP and KYPP indicate that PC2 largely reflects variation inversely associated with pod number and yield related parameters, distinguishing genotypes primarily on the basis of test weight (pod and kernel size) characteristics relative to overall yield output.

PC03 showed positive associations with harvest index (HI; 0.2954), mature pods per plant (MPPP; 0.2282) and days to 50% flowering (DF; 0.0380), while displaying negative correlations with sound mature kernel (SMK; -0.0465), days to maturity (DM; -0.0609), kernel yield per plant (KYPP; -0.1090), pod yield per plant (PYPP; -0.1121), shelling percentage (SP; -0.2136), oil content (OC; -0.2265), plant height (PH; -0.292), primary branches per plant (PBPP; -0.3264), 100-kernel weight (100KW; -0.3599), biological yield per plant (BYPP; -0.4126) and 100-pod weight (100PW; -0.4902). The strongly negative loadings of 100PW and BYPP suggest that PC3 primarily differentiates genotypes based on an inverse relationship between pod test-weight traits and overall biomass accumulation.

Table 1: Eigenvalues and variance explained by principal components for yield and yield-attributing traits in groundnut

Principal Component	Eigenvalue (SD ²)	Proportion of Variance	Cumulative Proportion
PC01	5.244	37.544	37.544
PC02	2.890	20.715	58.259
PC03	1.602	11.447	69.707
PC04	0.909	6.496	76.203
PC05	0.818	5.849	82.052
PC06	0.784	5.601	87.653
PC07	0.697	4.983	92.636
PC08	0.364	2.606	95.242
PC09	0.314	2.247	97.489
PC10	0.215	1.540	99.030
PC11	0.103	0.736	99.766
PC12	0.022	0.161	99.926
PC13	0.008	0.064	99.990
PC14	0.001	0.010	100.000

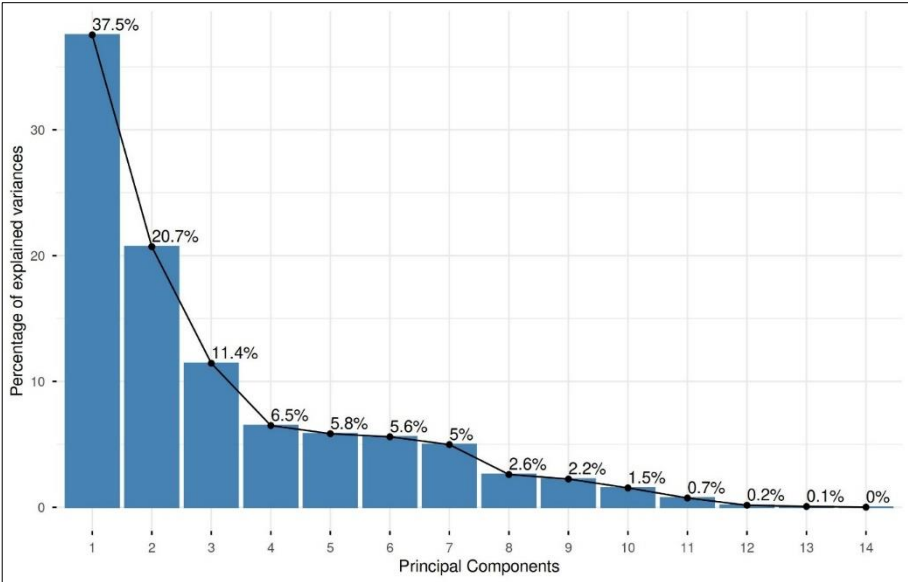


Fig 1: Scree plot showing the percentage of variance explained by each of the fourteen principal components

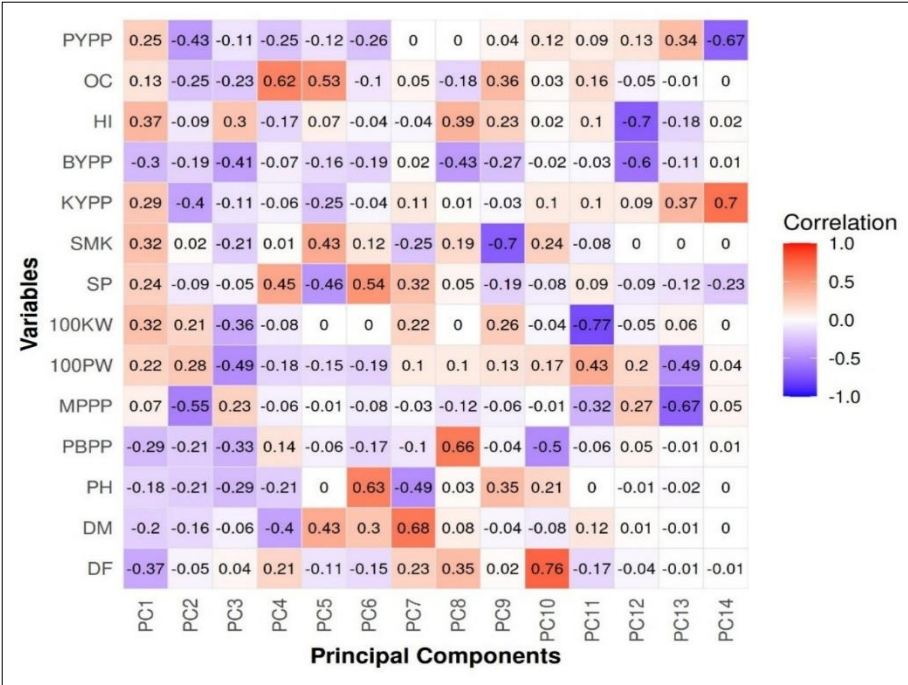


Fig 2: Correlation heatmap depicting the relationship between principal components and quantitative traits associated with pod yield

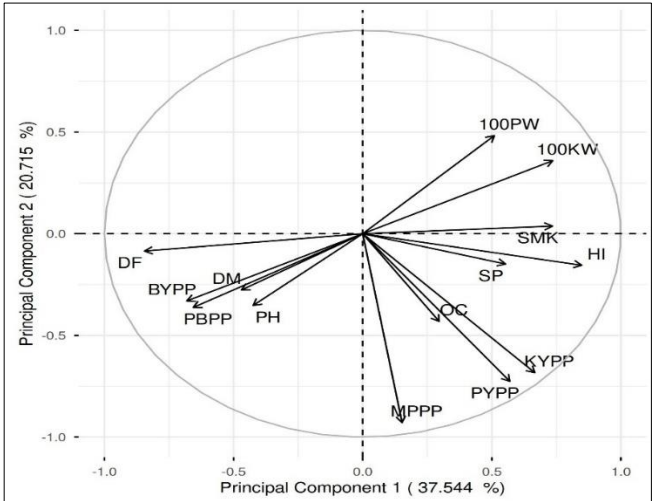


Fig 3: Correlation circle showing loadings of quantitative traits on Principal Component 1 and Principal Component 2

Vector direction indicates the trait's correlation with PC1 and PC2. A vector pointing right correlates positively with PC1; one pointing left correlates negatively. Similarly, upward vectors correlate positively with PC2 and downward vectors correlate negatively; Vector length reflects how well that trait is represented by these two components - vectors reaching close to the outer circle (radius = 1) are well explained by PC1- PC2, while shorter vectors indicate the trait's variation is better captured by other components (PC3 onward); Angle between two vectors indicates the correlation between those traits: a small angle (vectors pointing the same direction) suggests strong positive correlation, a $\sim 90^\circ$ angle suggests little to no correlation and vectors pointing in opposite directions ($\sim 180^\circ$) suggest strong negative correlation.)

3.2 Clustering of the genetic material using Tocher's method and Mahalanobis D² distance

Genetic divergence among the genotypes was assessed using Mahalanobis D² statistic (Mahalanobis, 1936) ^[11], which provides a quantitative measure of the genetic distance between genotypes based on multiple traits simultaneously, while accounting for the correlation existing among them. Based on the D² values, the genotypes were grouped into distinct clusters following Tocher's method (Rao, 1952), wherein genotypes exhibiting the smallest intra-cluster distances are grouped together such that the average intra-cluster D² value remains lower than the average inter-cluster D² value. This clustering approach enables the identification of genetically similar genotypes within a cluster and genetically divergent genotypes across clusters, thereby facilitating the selection of diverse parents for hybridization programs aimed at generating desirable transgressive segregants and exploiting heterosis. The distribution of genotypes into clusters, along with cluster means and percent contribution of individual traits toward total divergence, is presented in the sections that follow.

3.2.1 Inference Regarding the Distribution of Genotypes into Clusters

Genetic divergence analysis grouped the genotypes into twelve distinct clusters, revealing marked heterogeneity in cluster size and reflecting an uneven pattern of genetic diversity among the genotypes evaluated. Cluster 1 emerged as the largest, comprising 38 genotypes, followed by Cluster 2 with 11 genotypes, indicating a relatively high degree of genetic similarity among the majority of genotypes included in these two clusters. In contrast, Clusters 7, 8, 9, 10, 11 and 12 were each represented by a solitary genotype, suggesting that these genotypes possess a distinct genetic constitution sufficiently divergent from the rest of the germplasm to warrant separate clustering. Such disproportionate clustering, wherein a small number of large clusters coexist with several singleton clusters, is a commonly observed phenomenon in genetic divergence studies and indicates that while considerable genetic similarity exists among a majority of the genotypes, a subset of accessions harbors unique allelic combinations that set them apart, potentially representing valuable sources of variability for future hybridization programs.

3.2.2 Inference Regarding the Contribution of Traits to Genetic Divergence

The percentage contribution of individual traits toward total genetic divergence was estimated following the method of Singh (1981). Among the fourteen traits studied, harvest

index (HI) contributed maximally to the observed genetic divergence, accounting for 33.74% of the total variation, followed by kernel yield per plant (KYPP), which contributed 13.74%. The substantial contribution of these two traits indicates that they were the principal drivers of genetic differentiation among the genotypes and consequently, greater emphasis should be placed on these traits while selecting divergent parents for hybridization programs aimed at exploiting heterosis and generating desirable transgressive segregants.

3.2.3 Inference Regarding Cluster Means for Component Traits

Cluster mean values for the fourteen traits exhibited considerable variation across the twelve clusters, further substantiating the extent of genetic diversity captured through clustering.

For days to 50% flowering (DF), cluster means ranged from 32.66 days (Cluster 12) to 43.66 days (Cluster 9), indicating the presence of both early- and late-flowering genotype groups. Days to maturity (DM) followed a similar trend, ranging from 113.66 days (Cluster 5) to 123.00 days (Cluster 9). Plant height (PH) was maximum in Cluster 9 (25.17 cm) and minimum in Cluster 11 (15.02 cm), while primary branches per plant (PBPP) ranged from 5.10 (Cluster 7) to 8.133 (Cluster 9).

For mature pods per plant (MPPP), the highest mean was recorded in Cluster 6 (24.50) and the lowest in Cluster 9 (11.40). 100-pod weight (100PW) varied from 64.05 g (Cluster 3) to 119.98 g (Cluster 10), whereas 100-kernel weight (100KW) ranged from 24.82 g (Cluster 9) to 52.20 g (Cluster 12). Shelling percentage (SP) ranged between 56.59% (Cluster 8) and 69.10% (Cluster 12) and sound mature kernel (SMK) percentage varied from 57.21% (Cluster 3) to 88.05% (Cluster 5).

With respect to yield components, kernel yield per plant (KYPP) ranged from 5.33 g (Cluster 9) to 13.0 g (Cluster 6), while biological yield per plant (BYPP) varied from 20.86 g (Cluster 11) to 48.19 g (Cluster 4). Harvest index (HI) exhibited its maximum value in Cluster 11 (85.89%) and minimum in Cluster 9 (22.11%). Oil content (OC) ranged from 48.90% (Cluster 7) to 54.42% (Cluster 12) and pod yield per plant (PYPP), the primary trait of interest, ranged from 8.46 g (Cluster 9) to 21.50 g (Cluster 8).

Notably, Cluster 9 consistently recorded extreme values across multiple traits being the highest for DF, DM and PH, yet the lowest for MPPP, 100KW, KYPP, HI and PYPP suggesting that genotypes within this cluster are late-maturing, tall-statured, but poor performers for yield and yield-contributing traits and may therefore be of limited direct use in yield-improvement breeding programs unless targeted for specific traits.

Similar results were also reported by several researchers in groundnut. Habite (2023) ^[5] reported significant genetic variation among 64 groundnut genotypes evaluated at Pawe, Northwestern Ethiopia, with cluster analysis grouping genotypes into five divergent clusters and the first four principal components accounting for 70.03% of total variation, findings broadly consistent with the pattern of genetic divergence and variance distribution observed in the present study. Dama *et al.* (2026) ^[3] observed PC1 explaining 29.77% of total variance, governed predominantly by kernel yield, oil yield, hundred kernel weight, biological yield and harvest index in agreement with the yield-associated loading pattern observed for PC1 in the

present study. Likewise, Rajput *et al.* (2024) ^[14] reported hundred pod weight and hundred kernel weight as the major contributing traits in PCA, corroborating the strong positive loadings of these traits on PC1 and PC2 herein. Yami *et al.* (2025) ^[20] similarly recorded four principal components with eigenvalues greater than unity, cumulatively explaining 77.92% of total variation, a pattern consistent with the

substantial cumulative variance (69.71%) captured by the first three components in the present investigation. These consistent findings across independent studies strengthen the reliability of PCA and Mahalanobis D² statistic as effective tools for assessing genetic diversity and identifying superior yield-contributing traits in groundnut improvement programs.

Table 2: Cluster-wise distribution of genotypes based on genetic divergence

Cluster	Total Genotype	Genotype
Cluster 1	38	GG-2, GG-7, JB-202511, JB-202512, JB-202513, JB-202514, JB-202516, JB-202517, JB-202518, JB-202519, JB-202521, JB-202523, JB-202524, JB-202527, JB-202528, JB-202529, JB-202534, JB-202536, JB-202539, JB-202540, JB-202541, JB-202542, JB-202543, JB-202544, JB-202546, JB-202547, JB-202548, JB-202549, JB-202550, JB-202551, JB-202552, JB-202553, JB-202554, JB-202555, JB-202556, JB-202561, JB-202562, JB-20259
Cluster 2	11	GG-5, JB-20252, JB-202522, JB-202525, JB-202526, JB-202531, JB-202535, JB-202538, JB-20254, JB-202545, JB-202557
Cluster 3	4	GG-40, JB-202510, JB-202515, JB-202558
Cluster 4	4	Girnar-4, JB-20251, JB-20255, JB-20256
Cluster 5	5	GG-37, JB-202520, JB-202530, JB-202559, TG-37A
Cluster 6	2	JB-202533, JB-202537
Cluster 7	1	JB-20253
Cluster 8	1	JB-20257
Cluster 9	1	JB-20258
Cluster 10	1	JB-202532
Cluster 11	1	JB-202560
Cluster 12	1	GG-34

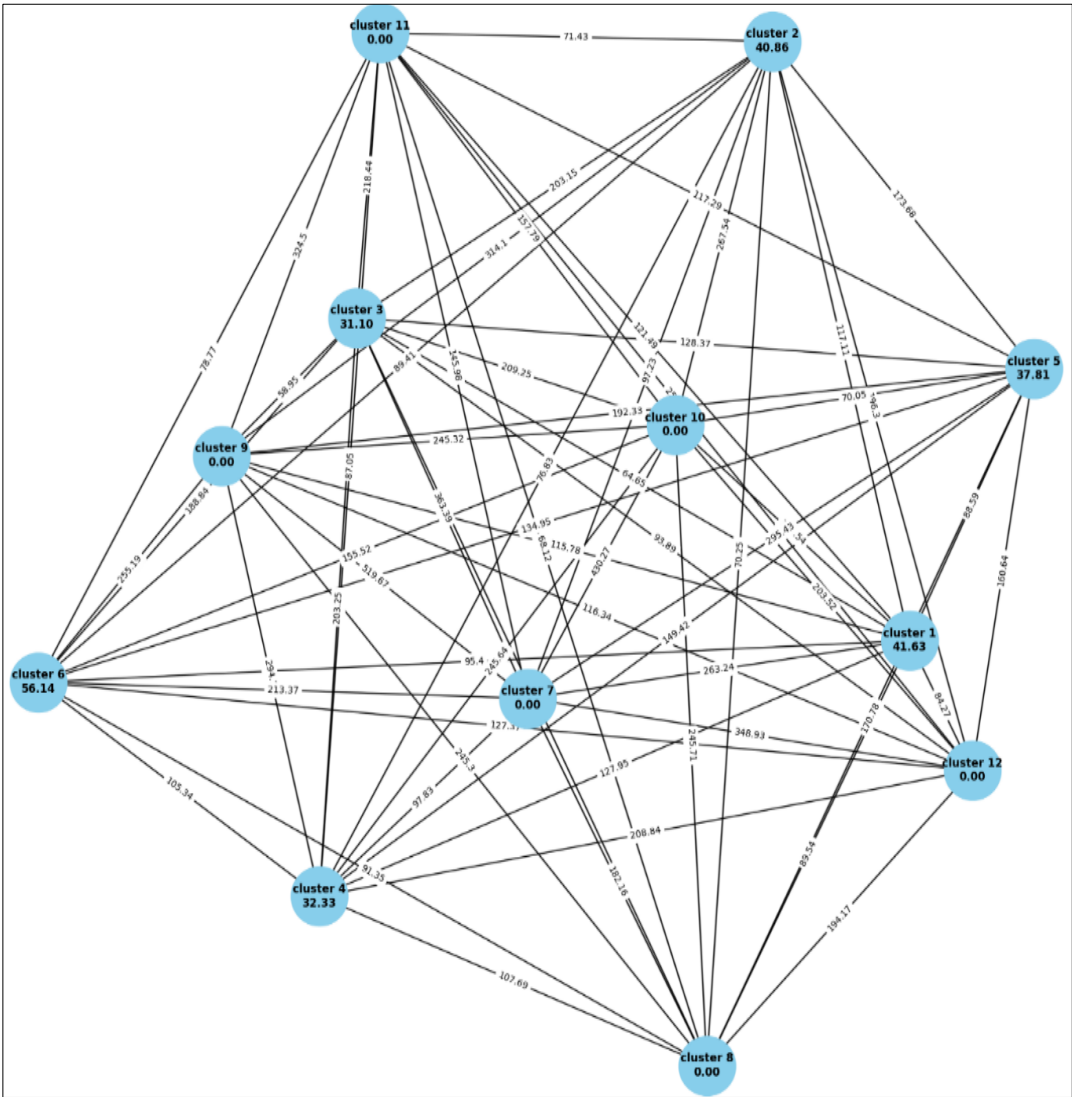


Fig 4: Cluster diagram depicting genetic divergence among clusters based on Mahalanobis D² distances

Table 3: Intra and inter cluster D² distances among clusters based on Tocher's method

Cluster Number	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
Cluster 1	41.63	117.11	64.65	127.95	88.59	95.40	263.24	89.54	115.78	151.54	121.49	84.27
Cluster 2	117.11	40.86	203.15	76.83	173.68	89.41	97.23	70.25	314.10	267.54	71.43	196.30
Cluster 3	64.65	203.15	31.10	203.25	128.37	188.84	363.39	149.41	58.95	209.25	218.44	93.89
Cluster 4	127.95	76.83	203.25	32.33	149.42	105.34	97.83	107.69	294.18	245.64	87.05	208.84
Cluster 5	88.59	173.68	128.37	149.42	37.81	134.95	295.43	170.78	192.33	70.05	117.29	160.64
Cluster 6	95.40	89.41	188.84	105.34	134.95	56.14	213.37	91.35	255.19	155.52	78.77	127.37
Cluster 7	263.24	97.23	363.39	97.83	295.43	213.37	0.00	182.16	519.67	430.27	145.98	348.93
Cluster 8	89.54	70.25	149.41	107.69	170.78	91.35	182.16	0.00	245.30	245.71	68.12	194.17
Cluster 9	115.78	314.10	58.95	294.18	192.33	255.19	519.67	245.30	0.00	245.32	324.50	116.34
Cluster 10	151.54	267.54	209.25	245.64	70.05	155.52	430.27	245.71	245.32	0.00	157.79	203.52
Cluster 11	121.49	71.43	218.44	87.05	117.29	78.77	145.98	68.12	324.50	157.79	0.00	258.03
Cluster 12	84.27	196.30	93.89	208.84	160.64	127.37	348.93	194.17	116.34	203.52	258.03	0.00

Conclusion

Principal Component Analysis and Mahalanobis D² divergence analysis revealed substantial genetic heterogeneity among seventy groundnut accessions. Three principal components explained 69.71% aggregate variance, with PC1 representing productivity and earliness mechanisms. Twelve clusters were identified through Mahalanobis D² metric, demonstrating pronounced variation in cluster magnitude. Harvest index and kernel yield per plant emerged as paramount divergence contributors. Cluster means exhibited considerable variation across measured traits, indicating substantial genetic material for advancement initiatives. Genotypes demonstrating favourable yield characteristics merit prioritization as parents for hybridization programs. Principal Component Analysis and divergence analysis confirmed their utility as effective tools for assessing heritable variability and identifying superior parental combinations in groundnut improvement programs.

Acknowledgments

I wish to express my profound gratitude to Dr. R. B. Madariya, serving as the major advisor, along with the members of my advisory committee for rendering valuable scholarly direction throughout this investigative endeavour. I further extend sincere appreciation to Junagadh Agricultural University for facilitating the provision of essential facilities and infrastructure that contributed significantly to the successful completion of this research undertaking.

Conflict of Interest: None.

References

1. Arunachalam V. Genetic distance in plant breeding. *Indian J Genet.* 1981;41:226-236.
2. Bonku R, Yu J. Health aspects of peanuts as an outcome of its chemical composition. *Food Sci Hum Wellness.* 2020;9(1):21-30.
3. Dama DB, Tesfamariam SA, Hassen AM. Multivariate analysis for agro-morphological and quality traits in groundnut (*Arachis hypogaea* L.) genotypes in Eastern Ethiopia. *PLoS One.* 2026;21(4):e0347850.
4. Feldman EB. The scientific evidence for a beneficial health relationship between walnuts and coronary heart disease. *J Nutr.* 2002;132(5):1062-1101.
5. Habite DN. Principal component analysis and genetic divergence studies for yield and yield-related trait of groundnut (*Arachis hypogaea* L.) genotypes at Pawe North Western Ethiopia. *East Afr Sch J Agric Life Sci.* 2023;6(8):154-160.
6. Hotelling H. Analysis of a complex of statistical variables into principal components. *J Educ Psychol.* 1933;24(6):417-441.
7. Johnson FR. The peanut story. North Carolina, USA; 1964.
8. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas.* 1960;20(1):141-151.
9. Knauff DA, Wynne JC. Peanut breeding and genetics. *Adv Agron.* 1995;55:393-445.
10. Kochert G, Stalker HT, Gimenes M, Galgaro L, Lopes CR, Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am J Bot.* 1996;83:1282-1291.
11. Mahalanobis PC. On the generalized distance in statistics. *Proc Natl Inst Sci India.* 1936;12:49-55.
12. Pearson K. On lines and planes of closest fit to systems of points in space. *Lond Edinb Dublin Philos Mag J Sci.* 1901;2(11):559-572.
13. Petersen KS, Murphy J, Whitbread J, Clifton PM, Keogh JB. The effect of a peanut-enriched weight loss diet compared to a low fat weight loss diet on body weight, blood pressure and glycemic control: a randomized controlled trial. *Nutrients.* 2022;14(14):2986.
14. Rajput LS, Sikarwar RS, Ingle SV, Tripathi MK, Prajapati PK, Yadav RK. Principal component analysis for groundnut yield and seed attributes (*Arachis hypogaea* L.). *Int J Res Agron.* 2024;7(5S):781-784.
15. Rao CR. Advanced statistical methods in biometric research. New York: John Wiley and Sons Inc.; 1952. p. 236-272.
16. Singh D. The relative importance of characters affecting genetic divergence. *Indian J Genet Plant Breed.* 1981;41(2):237-245.
17. Stalker HT. A new species in section *Arachis* of peanuts with AD genome. *Am J Bot.* 1991;78(5):630-637.
18. Vaishnavi G, Justin Thenmozhi A. Nuts and their potential role in Alzheimer's disease. In: *Nutraceuticals*

- for Alzheimer's disease: a promising therapeutic approach. Springer Science Reviews. 2023;9:125-149.
19. Wang C, Wang N, Li N, Yu Q, Wang F. Combined effects of resveratrol and vitamin E from peanut seeds and sprouts on colorectal cancer cells. Front Pharmacol. 2021;12:760919.
 20. Yami AS. Assessment of genetic variability for yield and yield-contributing traits in groundnut (*Arachis hypogaea* L.) genotypes. J Food Qual. 2025;3370389.