PLANT SCIENCE TODAY
Vol 12(4): 1-12
https://doi.org/10.14719/pst.11556

eISSN 2348-1900

RESEARCH ARTICLE

Principal component and cluster analysis of yield and its contributing traits in sesame (*Sesamum indicum* L.) genotypes

B Venkata Prasad[1], Swapnil[1], Sanghamitra Rout[1*], Saurav Singla[2], Digvijay Singh[3], Siddhartha Shankar Sharma[4], P Sushmita[5], P Jahnavi[1], M Chandhana Ishwarya[1], N Mary Lydia[1] & G Indumathi[1]

[1]Department of Genetics and Plant Breeding, Centurion University of Technology and Management, Paralakhemundi 761 211, Odisha, India

[2]Sri Karan Narendra Agriculture University, Jobner 302 018, Rajasthan, India

[3]Department of Seed Science and Technology, Acharya Narendra Dev University of Agriculture and Technology, Kumarganj, Ayodhya 224 229, Uttar Pradesh, India

[4]Department of Agriculture and Allied Sciences, C.V. Raman Global University, Bhubaneswar 752 054, Odisha, India

[5]Vignan Institute of Agriculture and Technology, Vignan's Foundation for Science, Technology and Research, Guntur 522 213, Andhra Pradesh, India

*Correspondence email - sanghamitra.rout49@gmail.com

Received: 31 August 2025; Accepted: 09 October 2025; Available online: Version 1.0: 05 November 2025; Version 2.0: 18 November 2025

Cite this article: Venkata Prasad B, Swapnil, Sanghamitra R, Saurav S, Digvijay S, Siddhartha SS, Sushmita P , Jahnavi P, Chandhana IM, Mary LN, Indumathi G. Principal component and cluster analysis of yield and its contributing traits in sesame (*Sesamum indicum* L.) genotypes. Plant Science Today. 2025; 12(4): 1-12. https://doi.org/10.14719/pst.11556

Abstract

Sesamum indicum L. (sesame) is an important oilseed crop valued for its high-quality oil and adaptability to diverse agro-climatic conditions. The present study assessed genetic variability, heritability, genetic advance, correlation, path analysis, genetic divergence and principal component analysis for 14 yield contributing traits across 28 sesame genotypes. The genotypes were evaluated over two years using a randomised complete block design. Analysis of variance revealed highly significant differences ($p<0.01$) among genotypes for all traits, indicating the presence of substantial genetic variability. High genotypic and phenotypic coefficients of variation were observed for seed yield per plant and number of capsules per plant. Traits such as seed yield per plant, oil content, harvest index, number of capsules per plant and days to 75 % maturity exhibited high heritability coupled with high genetic advance, suggesting predominant additive gene action. Genotypic correlation analysis indicated significant positive associations of seed yield per plant with stem height of the first capsule, plant height, number of seeds per capsule, 1000 seed weight, capsule length and harvest index. Path analysis identified plant height, number of capsules per plant, number of seeds per capsule, oil content and 1000-seed weight as key traits with the highest direct effects on seed yield. Mahalanobis D^2 analysis grouped genotypes into seven distinct clusters, with the highest inter-cluster distance observed between Clusters IV and V, indicating substantial genetic divergence between these groups. Principal component analysis (PCA) revealed six components accounting for 77.76 % of the total variation, with PC1 driven primarily by seed yield per plant and number of capsules per plant. Diverse genotypes such as IC-203945, NIC-1005 and PRACHI were identified as promising parents for hybridisation programs aimed at generating transgressive segregants.

Keywords: genetic variability; heritability; Mahalanobis D^2; path analysis; PCA; seed yield; sesame

Introduction

Sesamum indicum L. (2n = 26), commonly known as sesame, is an ancient and economically important oilseed species cultivated widely across tropical and subtropical regions. Often referred to as the "queen of oilseeds," sesame is prized for its high oil content, oxidative stability and health benefits due to naturally occurring antioxidants such as sesamin and sesamol (1). Its oil, rich in unsaturated fatty acids such as oleic and linoleic acids, contributes significantly to dietary and industrial value (2). In addition to oil, sesame seeds provide considerable protein, energy and essential micronutrients, making them nutritionally and economically valuable (3). Sesame is cultivated worldwide on about 11.7 million hectares, with a total production of 6.02 million tonnes and an average productivity of 512 kg ha^{-1} (4). The crop is predominantly grown in Asia and Africa, which together account for nearly 97 % of global production, while smaller contributions are made by North America and Europe (5). Among sesame producing countries, India ranks first in total sesame seed output, followed by Myanmar and China. Myanmar alone contributes nearly 9.5 % of global exports. However, productivity remains low due to factors such as seed shattering, disease susceptibility and limited access to improved cultivars. Despite its potential, sesame lags behind other oilseeds in yield gains, necessitating focused crop improvement efforts (6).

Improving yield in sesame requires understanding the extent of genetic variability, which reflects the differences among genotypes for quantitative and qualitative traits. It forms the fundamental basis of any crop improvement program and provides a platform for both natural and artificial selection. Greater variability provides a higher probability of identifying promising recombinants in segregating generations. Heritability plays a key role in

determining the degree to which genetic factors influence trait expression. When considered along with genetic advance, it helps assess the expected gain from selection. High heritability with high genetic advance indicates additive gene action, implying that such traits can be improved effectively through selection. High heritability and genetic advance for key traits like seed yield, oil content, number of capsules per plant and 1000-seed weight in sesame have been reported (7). Correlation analysis quantifies the degree and direction of association between yield and its contributing traits. Positive and significant correlations between seed yield and other morphological traits such as plant height, number of capsules and seed weight indicate the potential for indirect selection (8). However, correlation alone does not unravel the exact nature of relationships, particularly when traits are interdependent.

To overcome this limitation, path coefficient analysis is used to partition correlation coefficients into direct and indirect effects, revealing the actual contributors to yield. This technique aids breeders in identifying traits that influence yield independently versus those that exert influence through other components (9). Furthermore, assessing genetic divergence among genotypes is crucial for selecting parents in hybridisation programs. Mahalanobis D^2 statistics help quantify inter-genotypic distances based on multiple traits, allowing efficient identification of diverse and promising combinations. Selecting genetically distant parents increases the likelihood of heterotic effects and transgressive segregants in subsequent generations (10). For instance, the utility of D^2 analysis in grouping sesame genotypes and identifying superior parental combinations (11). PCA is another multivariate technique that helps reduce dimensionality and extract the most informative traits contributing to variation. By identifying major components governing overall diversity, PCA supports the selection of key traits for breeding and helps prioritise breeding targets. PCA has been found effective in sesame for identifying and clustering traits that significantly contribute to genetic variability (12).

With evolving breeding priorities such as improving oil quality, seed yield and climate resilience, there is an increasing need to re-evaluate existing germplasm. Therefore, integrating genetic variability estimation with correlation, path analysis and multivariate techniques can provide a comprehensive understanding of trait architecture and enhance breeding efficiency. Given the complex nature of yield governed by multiple genes and environmental interactions, simultaneous selection for multiple traits using integrated approaches becomes vital. This study was therefore undertaken to estimate genetic variability, heritability and genetic advance among 28 sesame genotypes, determination of trait associations and their direct and indirect effects on seed yield through correlation and path analysis and identification of genetically divergent parents using multivariate analysis for use in future breeding programs.

Materials and Methods

The present study was conducted during the *Rabi* seasons of 2023-24 and 2024-25 at the Post Graduate Research Farm, Ranadevi, M.S. Swaminathan School of Agriculture, Centurion University of Technology and Management, Paralakhemundi, Odisha. The experimental site is situated in the southeastern region of Odisha at an elevation of 145 m above mean sea level, with geographic coordinates of 18°46′41.8586″ N latitude and 84°53′1.1436″ E longitude. The experimental site experiences a sub-humid subtropical climate. During the *Rabi* seasons of 2023 and 2024, the average temperature ranged from 28 °C to 33 °C, with relative humidity between 65 % and 80 %. Rainfall during this period remained low, generally below 50 mm per month, indicating that sesame cultivation in the region primarily depends on supplemental irrigation for successful crop establishment and growth. The experimental soil was sandy loam, slightly alkaline in reaction and low organic carbon and available nitrogen - characteristics typical of the region's agro-ecology. Before sowing, the field was ploughed thoroughly and levelled to ensure a uniform seedbed and soil fertility status was maintained through appropriate nutrient management.

Twenty-eight genetically diverse sesame (*Sesamum indicum* L.) genotypes, procured from the Agricultural Research Station, Mandor, Rajasthan, India, were evaluated. The list of sesame genotypes is presented in Table 1. The experiment was laid out in a Randomised Block Design (RBD) with three replications to ensure statistical reliability. Each genotype was sown in three rows per plot, maintaining a spacing of 30 cm between rows and 15 cm between plants. Sowing was conducted during the early morning hours to avoid heat stress and ensure uniform germination.

Table 1. List of sesame genotypes

S. No.	Genotypes	S. No.	Genotypes
1.	RT-372	15	IC203945
2.	RT-399	16	IC-203970
3.	RT-351	17	IC-203938
4.	RT-346	18	NIC-7816
5.	RT-54	19	NIC-7968
6.	ES-384-1-84	20	NIC-1005
7.	PRACHI	21	EC-376966
8.	GT-10	22	TKG-22
9.	NIC-16386	23	K-3-K-92
10.	NIRMALA	24	K-22-RT-103
11.	NIC-16368	25	K-22-RT-102
12.	IC1415A	26	IC-203980
13.	IS-214	27	RT-103
14.	RT-391	28	IS1162

Source of germplasm is Agriculture Research Station, Mandor, Rajasthan, India

Standard agronomic practices were followed uniformly across all plots throughout the cropping period. A recommended basal dose of fertilisers, i.e. 40 kg N, 20 kg P_2O_5 and 20 kg K_2O per hectare, was applied, with remaining nitrogen top-dressed in two equal splits at 30 and 50 days after sowing. Irrigation was provided as and when required, particularly during the critical growth stages of flowering and capsule development, using the furrow irrigation method to ensure adequate soil moisture. Intercultural operations, including thinning, weeding and plant protection measures, were carried out as per standard protocols to maintain healthy crop growth and minimise biotic stress.

Observations were recorded on 14 quantitative traits from five randomly selected representative plants per plot in each replication and their average values were used for analysis. Days to flower initiation (DFI) was noted as the number of days from sowing to the appearance of the first flower, observed daily across all plots. Days to 50 % flowering (D50F) was recorded as the number of days from sowing until half of the plants in a plot exhibited at least one open flower. Days to 75 % maturity (D75M) was measured as the

number of days from sowing until 75 % of the plants reached physiological maturity, identified by the yellowing and drying of capsules. Stem height to first capsule (SHFC) was measured in centimetres from the soil surface to the node bearing the first capsule on the main stem using a meter scale. Plant height (PH) was recorded from the base to the terminal bud of the main stem at full maturity using a measuring tape. The number of primary branches per plant (PB) was counted manually as the branches arising directly from the main stem. Number of capsules per plant (NCP) was assessed by counting all mature capsules on the main stem and its branches. The number of locules per capsule (NLC) was determined by counting the number of locules in five randomly selected mature capsules, one per sampled plant and averaging the data. Similarly, the number of seeds per capsule (NSC) was recorded by counting seeds in five mature capsules per sampled plant and calculating the average. 1000 seed weight (TSW) was recorded in grams using a digital precision balance after manually counting and weighing 1000 well-dried, cleaned seeds. Capsule length (CL) and capsule width (CW) were measured in mm using a digital vernier calliper from five mature capsules per plot. Harvest index (HI) was calculated as the ratio of seed yield per plant to total above-ground biomass and expressed as a percentage (13). Seed yield per plant (SYPP) was obtained by harvesting, drying, threshing, cleaning and weighing the seeds from the five sampled plants. Finally, oil content (OC) was estimated using the Soxhlet extraction method with cold petroleum ether (boiling point range 60-80 °C) (14). The extracted sesame oil was clear, slightly aromatic and golden-yellow in appearance. These comprehensive observations helped evaluate the genotypic performance in terms of both yield and physiological efficiency.

For statistical analysis, the average values from each replication were used. To assess the significance of variation among genotypes, the experimental data were analysed using a RBD, treating genotypes as treatments (15). Analysis of variance (ANOVA) was performed using the 'agricolae' package in R software (16). The F-test was used to determine the significance of treatment effects; significance was assumed when the calculated F-value exceeded the tabulated F-value at the 5 % probability level.

The genotypic coefficient of variation (GCV) and phenotypic coefficient of variation (PCV) were calculated following (17). GCV and PCV are classified as high (> 25 %), moderate (10–25 %) and low (< 10 %) (18). Heritability in the broad sense (h^2B) was estimated as the ratio of genotypic variance to phenotypic variance and categorised following (19) as low (< 30 %), moderate (30-60 %) and high (> 60 %). Genetic advance (GA) was computed as per to estimate the expected gain under selection pressure (20). All genetic parameters were computed using the 'variability' package in R software (21).

To understand the nature of association among traits, genotypic and phenotypic correlation coefficients were calculated through Pearson's correlation through the 'metan' package in R (version 4.3.2) (22). The relationships among traits were visualised using the 'corrplot' package (23). PCA was conducted to partition the correlation coefficients into direct and indirect effects of various component traits on seed yield (24, 25). Path coefficients and residual effects were computed using the 'variability' package in R. Genetic divergence among the 28 sesame genotypes was evaluated using Mahalanobis D^2 statistics and genotypes were grouped based on inter- and intra-cluster distances to identify divergent parents (26). The hierarchical clustering was performed by using the package dendextend in R software (27). PCA was employed to reduce data dimensionality and to identify the traits contributing most to total variation among genotypes. PCA was carried out using the 'FactoMineR' package in R software (28). Components with eigenvalues greater than one were considered significant. Biplots and dendrograms illustrating genotype distribution were constructed using the 'factoextra' package (29).

Results and Discussion

ANOVA was performed to evaluate the differences among 28 sesame (*Sesamum indicum* L.) genotypes for 14 yield and yield contributing traits, as presented in Table 2. The results demonstrated that the genotypic differences were highly significant ($p < 0.01$) for all observed traits. Significant genotypic variation was observed for all 14 traits, including phenological traits DFI, D50F, D75M, plant morphology [SHFC, PH, PB] and yield components [NCP, NLC , NSC, SW, CL, HI, OIL and SYPP]. In most traits, the replication effect was non-significant, which indicates the experiment was precise. For some traits such as HI, PB and SYPP, the effect of year and the interaction of year × genotype (Y × G) were also significant. This means that the performance of genotypes for these traits changed across the two years. On the other hand, traits like OC, SW and PH were stable with no significant year × genotype interaction. The observed significant variation among genotypes indicates substantial genetic variability, offering strong potential for selecting superior lines in breeding programs. The considerable genetic variability for yield and related traits in sesame highlights the potential for genetic improvement (30).

The genetic parameters for 14 agronomic and yield-related traits in sesame are presented in Table 3. Analysis of PCV and GCV revealed substantial genetic variability across traits. High PCV and GCV (> 20 %) were recorded for SYPP (35.09 % and 29.88 %), indicating high genetic variability and suggesting strong potential for improvement through direct selection. The NCP also showed high PCV (21.19 %) and moderate GCV (16.87 %), reflecting considerable genetic potential, though partly influenced by environmental factors. Traits like D75M (17.42 % PCV, 16.88 % GCV), PH (12.91 %, 11.09 %), PB (14.46 %, 10.61 %), NSC (11.53 %, 10.31 %), HI (16.20 %, 13.67 %) and OC (19.40 %, 18.42 %) exhibited moderate variability (10-20 %), indicating reasonable genetic control and scope for selection. In contrast, traits such as DFI (10.28 % PCV, 7.40 % GCV), SHFC (10.59 %, 8.31 %) and NLC (10.29 %, 7.57 %) displayed moderate PCV but lower GCV, suggesting greater environmental influence. Meanwhile, D50F (6.24 %, 4.17 %), CL (2.58 %, 2.08 %) and SW (11.71 % PCV, 9.63 % GCV) showed low variability, implying limited scope for direct improvement through selection. Importantly, the narrow differences between PCV and GCV for many traits, especially seed yield and oil content, highlight that their variability is largely genetic, favouring early-generation selection. Similar results of high PCV and GCV for SYPP and NCP were reported (31, 32). Moderate PCV and GCV were recorded for PH, OC and D50F (33).

Heritability estimates further support these observations. High heritability (> 60 %) was observed for D75M (93.90 %), OC (90.10 %), NSC (79.90 %), SYPP (72.50 %), HI (71.30 %), PH (73.70 %), SW (67.70 %), CL (64.90 %), NCP (63.40 %) and SHFC (61.60 %). These high values suggest these traits are predominantly controlled by additive gene action and are suitable for selection. Moderate heritability (30-60 %) was found for DFI (51.80 %),

Table 2. ANOVA for 14 traits in sesame (*Sesamum indicum* L.) genotypes over two years

Year	Sources of variation (SV)	d.f.	Mean Sum of Squares						
			DFI	D50F	D75M	SHFC	PH	PB	NCP
First Year	Replication	2	0.235	0.080	15.048	0.149	0.298	0.001	0.139
	Genotypes (G)	27	20.506	15.273	932.897	9.173	71.002	0.075	75.902
	Error	54	8.5040	1.930	1.801	1.8120	8.730	0.026	4.063
Second Year	Replication	2	0.5200	0.668	7.370	0.181	0.593	0.000	0.557
	Genotypes (G)	27	45.554	21.525	909.675	11.510	92.625	0.446	155.978
	Error	54	6.664	12.104	47.012	1.818	8.750	0.012	8.268
Pooled	Replication	2	0.377	0.374	11.209	0.165	0.446	0.001	0.348
	Years (Y)	1	0.006	0.011	27.410	0.005	0.000	0.148	3106.492
	Genotypes (G)	27	57.693	34.951	1841.729	18.998	155.141	0.367	178.273
	Y × G	27	8.367	1.848	0.844	1.684	8.485	0.155	53.607
	Pooled Error	108	7.584	7.017	24.407	1.815	8.740	0.019	6.166

** Significant at 1 % level of probability

DFI= Days to flower initiation, D50F= Days to 50 % flowering, D75M=Days to75 % maturity, SHFC= Stem height of first capsule (cm), PH= Plant height (cm), PB= No. of primary branches per plant, NCP= No. of capsules per plant

Year	Sources of variation (SV)	d.f.	Mean Sum of Squares						
			NLC	NSC	SW	CL	HI	OIL	SYPP
First Year	Replication	2	0.003	54.316	0.531	0.000	10.227	0.245	0.019
	Genotypes (G)	27	0.298	84.483	0.380	0.008	2.339	261.748	6.923
	Error	54	0.086	0.308	0.121	0.000	0.155	14.564	0.209
Second Year	Replication	2	0.004	0.939	0.001	0.000	0.039	0.036	0.078
	Genotypes (G)	27	0.298	89.269	0.696	0.008	4.919	264.358	13.471
	Error	54	0.098	17.036	0.007	0.003	0.526	8.287	0.399
Pooled	Replication within Year	4	0.003	27.627	0.266	0.000	5.133	0.140	0.048
	Years (Y)	1	0.001	1.099	0.076	0.000	410.281	0.052	127.264
	Genotypes (G)	27	0.595	173.582	0.976	0.016	6.555	524.537	16.652
	Y × G	27	0.000	0.170	0.100	0.000	0.703	1.570	3.742
	Pooled Error	108	0.092	8.672	0.065	0.002	0.340	11.426	0.304

** Significant at 1 % level of probability. NLC= No. of locules per capsule, NSC= No. of seeds per capsule, SW= 1000 seed weight (g), CL= Capsule length (cm), HI= Harvest Index (%), OIL= Oil content (%) and SYPP= Seed yield per plant (g)

Table 3. Genetic parameters for 14 traits in sesame (*Sesamum indicum* L.) genotypes

Characters	Range		GCV	PCV	h^2 (Broad sense)	Genetic advance as % of mean
	Lowest	Highest				
Days to flower initiation	33.728	45.760	7.399	10.279	51.800	10.973
Days to 50 % flowering	48.643	56.643	4.168	6.238	44.700	5.739
Days to75 % maturity	80.667	135.502	16.882	17.420	93.900	33.700
Stem height of first capsule (cm)	15.157	23.517	8.312	10.592	61.600	13.438
Plant height (cm)	33.933	58.525	11.087	12.910	73.700	19.614
No. of primary branches	1.867	2.800	10.607	14.461	53.800	16.027
No. of capsules per plant	22.600	40.633	16.872	21.191	63.400	27.671
No. of locules per capsule	3.002	4.002	7.571	10.287	54.200	11.479
No. of seeds per capsule	43.332	62.065	10.310	11.532	79.900	18.988
1000 seed weight (g)	2.782	4.728	9.630	11.705	67.700	16.322
Capsule length (mm)	2.300	2.400	2.075	2.575	64.900	3.445
Harvest Index (%)	5.100	9.278	13.674	16.198	71.300	23.778
Oil content (%)	33.053	63.907	18.415	19.402	90.100	36.003
Seed yield per plant (g)	2.400	8.030	29.880	35.090	72.500	52.390

GCV= Genotypic coefficient of variation and PCV= Phenotypic coefficient of variation

D50F (44.70 %) and PB (53.80 %), indicating moderate genetic control. Genetic advance as a percentage of mean revealed that traits like SYPP (52.39 %), OC (36.00 %), D75M (33.70 %), NCP (27.67 %) and HI (23.77 %) had high genetic advance (> 20 %), confirming their suitability for improvement through direct phenotypic selection due to additive gene effects. Moderate genetic advance (10-20 %) was recorded for DFI (10.97 %), SHFC (13.44 %), PH (19.61 %), PB (16.03 %), NSC (18.98 %) and SW (16.32 %). Low genetic advance (< 10 %) was noted for D50F (5.74 %) and CL (3.44 %), suggesting limited response to selection. Overall, traits like SYPP, OC, NCP, HI and D75M, which combine high heritability with high genetic advance, are mainly governed by additive gene action and hold the greatest promise for effective selection and genetic improvement in sesame. Similarly high heritability coupled with high genetic advance for SYPP, NCP and HI was observed (34).

In the present study, genotypic correlation coefficients among 14 quantitative traits in sesame were estimated and the results are presented in Table 4 and illustrated in Fig. 1. The correlation analysis revealed that SYPP exhibited highly significant and positive associations with PH (0.531), NSC (0.588), SW (0.592), CL (0.260), HI (0.433) and SHFC (0.179). These results indicate that genotypes with taller stature, larger seed size, better capsule development and efficient assimilate partitioning tend to produce higher yields. These traits can be prioritised during selection as they are likely to contribute directly or indirectly to enhanced productivity. A positive but non-significant correlation was also observed between SYPP and the PB (0.115), NCP (0.565) and days to flower initiation (0.032), suggesting these traits may influence yield under specific environmental or genetic backgrounds, although not universally across all genotypes or seasons. In previous studies, PH and TSW were regarded as the key components of seed yield (35, 36).

Table 4. Association (correlation) between the yield and its attributing traits at genotypic level in sesame (*Sesamum indicum* L.) genotypes

	D50F	D75M	SHFC	PH	PB	NCP	NLC	NSC	SW	CL	HI	OIL	SYPP
DFI	0.290**	-0.223**	-0.712**	-0.179*	-0.299**	0.048	-0.326**	-0.043	-0.004	-0.079	0.583**	-0.108	0.032
D50F		0.086**	-0.338**	-0.504**	-0.248**	-0.386**	-0.153*	-0.297**	-0.111	0.167*	-0.109	-0.130	-0.438**
D75M			0.148	-0.018	0.082**	-0.292**	0.356**	0.146**	0.121	-0.137	0.181*	0.059	-0.106
SHFC				0.435**	-0.035	0.048	-0.171	0.036	0.584**	0.095	0.027	0.067	0.179*
PH					0.235**	0.528**	0.181*	0.159*	0.278**	-0.093	0.572**	-0.395**	0.531**
PB						0.088**	0.104	0.047**	0.164	0.095	0.419**	-0.383**	0.115
NCP							-0.046	0.224**	0.221**	0.194*	0.347**	-0.159*	0.565
NLC								-0.406**	0.010	0.137	0.453**	-0.343**	-0.149
NSC									0.139	0.166*	0.134	0.109	0.588**
SW										0.337**	0.424**	-0.070	0.592**
CL											-0.185	0.111	0.260**
HI												-0.234**	0.433**
OIL													-0.107

**, *significant at 1 % and 5 % level of probability

DFI= Days to flower initiation, D50F= Days to 50 % flowering, D75M=Days to75 % maturity, SHFC= Stem height of first capsule (cm), PH= Plant height (cm), PB= No. of primary branches per plant, NCP= No. of capsules per plant, NLC= No. of locules per capsule, NSC= No. of seeds per capsule, SW= 1000 seed weight (g), CL= Capsule length (cm), HI= Harvest Index (%), OIL= Oil content (%) and SYPP= Seed yield per plant (g)

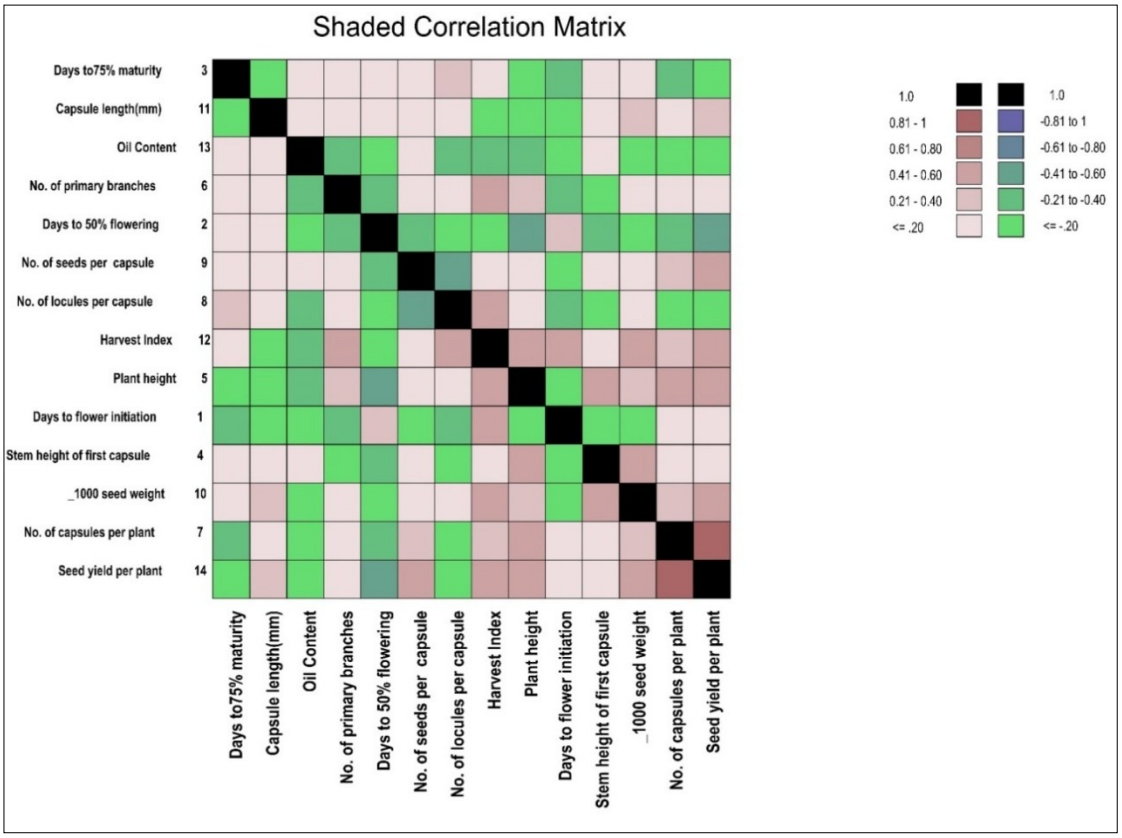

Fig. 1. Genotypic correlation coefficient among the 14 yield attributing traits sesame (*Sesamum indicum* L.) genotypes.

The trait PH was significantly correlated with the NCP, SW, HI and SHFC, indicating a strong contribution of overall plant architecture to reproductive efficiency. However, PH was negatively correlated with OC (-0.395), suggesting a trade-off between biomass and OC. NSC was significantly and positively correlated with CL (0.166) and showed a non-significant positive correlation with SW (0.139), HI (0.134) and OC (0.109). SW exhibited highly significant positive associations with CL (0.337) and HI (0.424), confirming its central role in determining seed size and overall assimilate allocation. CL was also positively but non-significantly correlated with OC (0.111) and HI exhibited a negative, significant correlation with OC (-0.234). SHFC exhibited a highly significant positive correlation with PH (0.435) and SW (0.584) and a non-significant positive correlation with NCP (0.048), NSC (0.036), CL (0.095), HI (0.027) and OC (0.067). These associations highlight their potential contribution to overall plant productivity through better seed fill and capsule structure.

Correlation analysis provides useful insights into trait associations but does not distinguish whether these relationships are direct or mediated through other characters. To resolve this, PCA was carried out, using seed yield per plant as the dependent variable and 13 other traits as independent variables. The genotypic path coefficient results are presented in Table 5, which effectively partitions the correlation into direct and indirect effects, helping identify traits with the most substantial influence on yield. Among all traits, PH exerted the highest positive direct effect (1.2766) on SYPP, confirming its critical and independent role in yield determination. This was followed by NCP (0.5688), NSC (0.6986), OC (0.7789) and SW (1.941). These traits also showed strong positive correlations with yield (0.531, 0.565, 0.588, -0.107 and 0.592, respectively), indicating

Table 5. Direct (diagonal) and indirect (off-diagonal) effects of different attributing traits on seed yield in sesame (*Sesamum indicum* L.) genotypes

	DFI	D50F	D75M	SHFC	PH	PB	NCP	NLC	NSC	SW	CL	HI	OIL	Genotypic correlated with SYPP
DFI	**-0.7352**	-0.2132	0.164	0.6405	0.1316	0.2199	-0.0353	0.2396	0.0316	0.0031	0.0579	-0.4289	0.0793	0.032
D50F	0.2183	**0.7527**	0.065	-0.2547	-0.3792	-0.1867	-0.2904	-0.1152	-0.2232	-0.0832	0.1253	-0.0823	-0.0975	-0.438**
D75M	0.0877	-0.0339	**-0.393**	-0.0583	0.0069	-0.0324	0.1146	-0.14	-0.0572	-0.0475	0.0538	-0.071	-0.0232	-0.106
SHFC	1.5145	0.5883	-0.2578	**-1.7384**	-0.7555	0.0603	-0.0841	0.2973	-0.0633	-1.0153	-0.1648	-0.0469	-0.1161	0.179*
PH	-0.2286	-0.6431	-0.0224	0.5548	**1.2766**	0.2996	0.674	0.2314	0.2026	0.3548	-0.1181	0.7297	-0.5043	0.531**
PB	-0.0378	-0.0313	0.0104	-0.0044	0.0297	**0.1264**	0.0112	0.0131	0.0059	0.0208	0.012	0.053	-0.0484	0.115
NCP	0.0273	-0.2195	-0.1659	0.0275	0.3003	0.0503	**0.5688**	-0.0259	0.1273	0.1258	0.1105	0.1971	-0.0904	0.857
NLC	-0.1573	-0.0739	0.172	-0.0826	0.0875	0.0502	-0.022	**0.4827**	-0.196	0.0047	0.066	0.2188	-0.1657	-0.149
NSC	-0.0301	-0.2071	0.1016	0.0254	0.1109	0.0327	0.1563	-0.2837	**0.6986**	0.0974	0.1161	0.0937	0.0764	0.588**
SW	-0.0082	-0.2146	0.2348	1.1337	0.5395	0.3189	0.4292	0.019	0.2705	**1.941**	0.6531	0.822	-0.1362	0.592**
CL	0.0734	-0.155	0.1275	-0.0883	0.0862	-0.0884	-0.1809	-0.1274	-0.1547	-0.3134	**-0.9313**	0.1723	-0.1035	0.260**
HI	-0.608	0.1139	-0.1882	-0.0281	-0.5958	-0.4371	-0.3611	-0.4725	-0.1398	-0.4414	0.1928	**-1.0423**	0.244	0.433**
OIL	-0.084	-0.1009	0.046	0.052	-0.3077	-0.2985	-0.1238	-0.2674	0.0852	-0.0547	0.0865	-0.1824	**0.7789**	-0.107

Residual effect = 0.2682

**, *significant at 1 % and 5 % level of probability

DFI= Days to flower initiation, D50F= Days to 50 % flowering, D75M=Days to75 % maturity, SHFC= Stem height of first capsule (cm), PH= Plant height (cm), PB= No. of primary branches per plant, NCP= No. of capsules per plant, NLC= No. of locules per capsule, NSC= No. of seeds per capsule, SW= 1000 seed weight (g), CL= Capsule length (cm), HI= Harvest Index (%), OIL= Oil Content (%) and SYPP= Seed yield per plant (g)

their consistent contribution to productivity. Their strong direct effects suggest that selection based on these characters will result in immediate genetic gains without relying heavily on indirect influences. These results are in agreement with (37-39).

Conversely, negative direct effects were observed for DFI (-0.735), D75M (-0.393), SHFC (-1.738), HI (-1.042) and CL (-0.931). Although some of these traits were moderately correlated with yield, their suppressive direct effects suggest that they may not be suitable for direct selection. For instance, delayed flowering or extended maturity may divert assimilates away from reproductive development, lowering productivity. Similarly, higher stem height or longer capsules might increase biomass without proportional seed gains. Research indicates that negative direct effects of maturity duration and capsule traits, suggesting such traits may act as physiological burdens on yield potential (37). Despite their negative direct contributions, several traits demonstrated favourable indirect effects. For instance, SHFC, though negatively affecting yield directly, contributed positively via DFI (1.5145) and D50F (0.5883). Similarly, DFI, while suppressive on its own, had positive indirect effects through SHFC (0.6405) and NLC (0.2396). These traits may not be suitable for direct selection but could be included in multi-trait selection indices due to their indirect support of yield-enhancing components. SW, which had a substantial positive direct effect, also contributed indirectly through PH (0.5395) and NCP (0.4292), further reinforcing its central role in yield physiology by influencing seed development and sink strength. Similarly, HI, despite its negative direct effect, had a supportive role through indirect effects via PH (0.7297) and SW (0.822), suggesting that optimising these components can improve source-sink dynamics. Traits such as CL and NLC, although unfavourable in direct effects, demonstrated positive indirect contributions *via* HI, PH and other reproductive characters, indicating their potential supportive role in certain breeding contexts.

The results of the path analysis suggest that PH, NCP, NSC, OC and SW are reliable indicators for direct selection in breeding programs targeting yield enhancement. Additionally, traits with significant indirect effects, such as PH, HI and flowering duration, may be used in index-based selection strategies to exploit favourable interactions among yield components. The residual effect of the PCA was 0.2682, indicating that the set of traits studied accounted for approximately 73.18 % of the total variation in seed yield per plant. The relatively low residual effect suggests that the majority of the variability in yield was adequately explained by the traits included in the analysis, highlighting the effectiveness of the chosen characters in determining yield potential. Comparable outcomes of low residual effect have been documented (40).

Genetic divergence analysis using Mahalanobis D^2 statistics revealed considerable variability among the 28 sesame (*Sesamum indicum* L.) genotypes studied. The genotypes were grouped into seven distinct clusters was presented in Table 6) and the dendrogram was shown in Fig. 2, indicating a wide genetic base within the material. Clusters I (RT-372, ES-384-1-84, IC1415A, IS-214, IC-203938 and EC-376966) and VI (NIC-16368, IC203945, NIC-1005, K-3-K-92, K-22-RT-103 and IS1162) were the largest, each comprising six genotypes, followed by Clusters II (RT-399, RT-351, RT-54 and IC-203970), III (RT-346, RT-391, TKG-22 and IC-203980) and VII (NIC-7816, NIC-7968, K-22-RT-102 and RT-103) with 4 genotypes each. Cluster V (GT-10, NIC-16368 and Nirmala) contained 3 genotypes, while Cluster IV was monogenotypic, represented solely by 'PRACHI'. The unique presence of PRACHI in a separate cluster underlines its genetic distinctness, making it a valuable candidate for breeding programs aimed at incorporating novel or underutilised traits. Such clustering patterns have been similarly reported in sesame (41). They highlighted the importance of including genetically unique genotypes to broaden the genetic base and achieve meaningful recombination in hybrid progeny.

The intra-cluster (diagonal) and inter-cluster (off-diagonal) D^2 distances among the 28 genotypes was shown in Table 7. It further confirmed the presence of substantial genetic diversity. Cluster III exhibited the highest intra-cluster distance (13.969), suggesting it encompasses genetically diverse entries. In contrast, Cluster IV, being monogenotypic, showed zero intra-cluster distance. The maximum inter-cluster distance was observed between Cluster IV and Cluster V (41.106), followed by Cluster III vs. IV (38.589) and

Table 6. Distribution of 28 sesame (*Sesamum indicum* L.) genotypes in seven clusters

Cluster number	No. of genotypes	Genotypes
I	6	RT-372, ES-384-1-84, IC1415A, IS-214, IC-203938 and EC-376966
II	4	RT-399, RT-351, RT-54 and IC-203970
III	4	RT-346, RT-391, TKG-22 and IC-203980
IV	1	PRACHI
V	3	GT-10, NIC-16368 and NIRMALA
VI	6	NIC-16368, IC203945, NIC-1005, K-3-K-92, K-22-RT-103 and IS1162
VII	4	NIC-7816, NIC-7968, K-22-RT-102 and RT-103

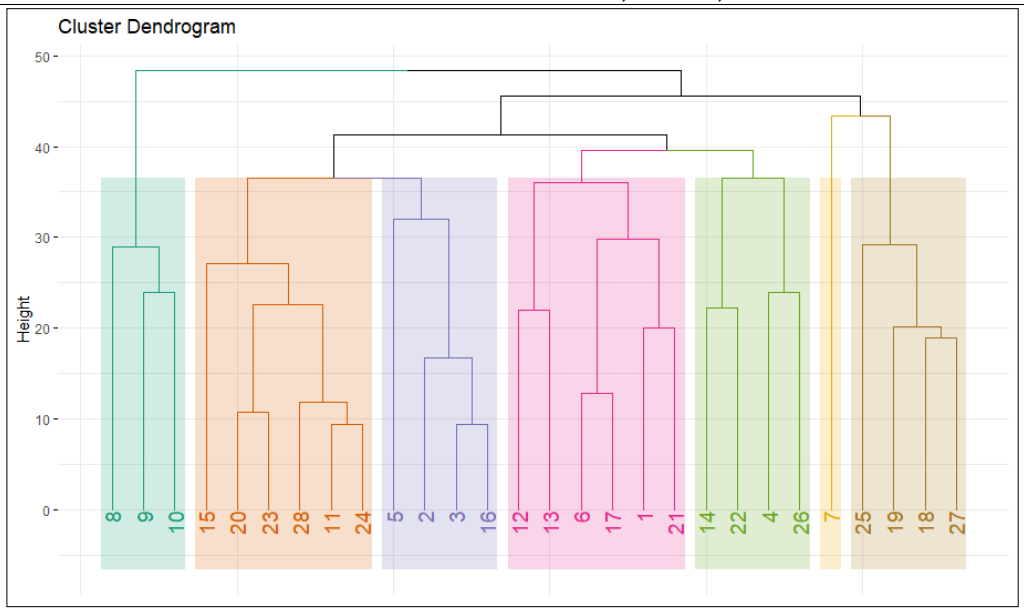

Fig. 2. Dendrogram showing the clustering of 28 genotypes of sesame.

Table 7. Average intra and inter D^2 values of sesame genotypes

Cluster	I	II	III	IV	V	VI	
I	**12.386**	27.141	29.248	38.169	32.078	24.940	29.201
II		**9.914**	28.753	36.928	31.004	23.578	27.856
III			**13.969**	38.589	33.857	27.007	30.885
IV				**0**	41.106	33.525	36.871
V					**13.229**	28.843	32.736
VI						**8.456**	25.769
VII							**11.467**

Cluster I vs. IV (38.169). These large distances imply high genetic divergence**,** indicating that crosses between these clusters could lead to maximum heterosis and recombination potential**.** Conversely, the minimum inter-cluster distance between Cluster I and II (27.141) suggests relatively narrow variability and their combinations may offer limited improvement. Such relationships are consistent with previous studies, which emphasised that selecting genetically distant parents based on D^2 analysis results in heterotic hybrids and transgressive segregants**,** enhancing breeding outcomes.

Specific genotype combinations with high D^2 values were identified and presented in Table 8. The table include IC-203945 (Cluster VI) × GT-10 (Cluster V) with D^2= 40.28, NIC-1005 (Cluster VI) × GT-10 (Cluster V) with D^2 = 40.28, K-3-K-92 (Cluster VI) × NIC-16368 (Cluster V) with D^2=34.61, IC-203970 (Cluster III) × PRACHI (Cluster IV) with D^2 = 42.75 and RT-54 (Cluster III) × PRACHI (Cluster IV) with D^2= 38.53. These pairs represent the most genetically divergent genotypes and should be prioritised in hybridisation programs for improving yield, oil content and adaptability.

Cluster mean analysis for the 14 agronomic traits was presented in Table 9, which revealed significant inter-cluster variation, offering useful insights for trait-specific improvement. Cluster IV, despite being monogenotypic, was found superior in DFI, D50F, PH, PB, NCP, CL, OC and HI, making it an ideal parent for earliness, plant vigour and seed quality traits**.** Cluster III recorded the highest SYPP (3.165 g), establishing it as a strong candidate for yield enhancement. Cluster V showed the highest NSC (12.463) and good OC (20.814 %), while Cluster VII was promising for OC (22.650 %) and SW (0.500 g)**.** These findings suggest that crosses between clusters with complementary trait means can lead to genetic gain through recombination. These observations are in line with the results, which advocate combining high-performing clusters for strategic breeding.

The contribution of each trait to total divergence was presented in Table 10, which revealed that CL (22.550 %), D75M (13.821 %) and D50F (12.319 %) were the major contributors to genetic divergence. These were followed by PB (9.412 %) and NCP (7.672 %), indicating that reproductive duration and plant architecture traits are the key differentiators among the genotypes. Traits like SW (0.046 %), HI (0.594 %) and SYPP (0.870 %) contributed least to divergence, suggesting relative uniformity among the genotypes for these attributes. This suggests that while yield is the ultimate goal, its component traits and growth duration play a more decisive role in defining genotypic differences.

PCA was conducted to reduce the dimensionality of the multivariate dataset and to identify the key traits contributing to genetic variation among the 28 sesame (*Sesamum indicum* L.) genotypes. The analysis identified six principal components (PCs) with eigenvalues greater than one, shown in Table 11, which

Table 8. Description of the genetically divergent clusters and distance (D^2 value) between the genotypes selected

Cluster combination	Inter-cluster distance	Genotype selected from the cluster	Distance between the genotypesselected (D^2 value)
Cluster VI and cluster V	41.106	IC-203945 from cluster VI and GT-10 from cluster V	40.28
Cluster VI and cluster V	41.106	NIC-1005 from cluster VI and GT-10 from cluster V	40.28
Cluster VI and cluster V	41.106	K-3-K-92 from cluster VI and NIC-16368 from cluster V	34.61
Cluster III and cluster IV	38.589	IC-203970 from cluster III and Prachi from cluster IV	42.75
Cluster III and cluster IV	38.589	RT-54 from cluster III and Prachi from cluster IV	38.53

Table 9. Cluster mean for the14 traits in sesame (*Sesamum indicum* L.) genotypes

Traits	I	II	III	IV	V	VI	VII
Days to flower initiation	3.090	2.461	6.607	42.988	9.250	3.187	5.501
Days to 50 % flowering	1.783	2.555	4.768	54.957	6.496	2.562	2.896
Days to75 % maturity	22.432	30.071	22.235	88.165	21.636	9.389	28.642
Stem height of first capsule (cm)	1.677	2.357	1.204	19.372	5.760	2.173	1.876
Plant height (cm)	4.660	4.394	11.281	43.187	6.577	3.994	3.761
No. of primary branches	0.116	0.221	0.282	2.002	0.048	0.141	0.209
No. of capsules per plant	5.944	5.928	10.526	24.065	7.094	5.218	10.359
No. of locules per capsule	4.000	4.000	4.000	4.002	3.000	4.000	4.000
No. of seeds per capsule	6.045	11.219	8.331	51.700	12.463	4.560	8.372
1000 seed weight (g)	0.260	0.290	0.307	2.782	1.767	0.282	0.500
Capsule length (mm)	0.054	0.080	0.000	2.300	0.143	0.054	0.092
Harvest Index (%)	1.594	0.594	1.444	6.448	0.591	0.901	1.098
Oil Content (%)	6.112	13.861	19.660	53.680	20.814	7.723	22.650
Seed yield per plant (g)	1.806	2.520	3.165	1.927	3.617	1.764	1.286

Table 10. Contribution of different traits towards genetics divergence in sesame (*Sesamum indicum* L.) genotypes

S. No.	Source	Contribution %
1	Days to flower initiation	4.040
2	Days to 50 % flowering	12.319
3	Days to75 % maturity	13.821
4	Stem height of first capsule (cm)	7.766
5	Plant height (cm)	7.272
6	No. of primary branches	9.412
7	No. of capsules per plant	7.672
8	No. of locules per capsule	2.606
9	No. of seeds per capsule	6.116
10	1000 seed weight (g)	0.046
11	Capsule length (cm)	22.550
12	Harvest Index (%)	0.594
13	Oil Content (%)	4.914
14	Seed yield per plant (g)	0.870

collectively explained 77.757 % of the total variability present in the population. This indicates a significant amount of the total information was retained using a reduced number of components. Among them, PC1 had the highest eigenvalue (2.982) and explained 21.299 % of the total variance, followed by PC2 (2.162; 15.441 %), PC3 (1.774; 12.671 %), PC4 (1.558; 11.126 %), PC5 (1.296; 9.256 %) and PC6 (1.115; 7.963 %). These findings indicate that most of the genetic diversity could be captured using the first few components, consistent with earlier reports who found that a limited number of PCs effectively summarise complex trait variation in sesame (44, 45). The scree plot shown in (Fig. 3) visually supported these results, showing a steep drop after PC1 and a gradual decline afterwards, confirming that PC1 and PC2 were the most informative components for variation. Similar patterns were reported in sesame, validating the interpretability of early PCs (46) .

The factor loadings on the first five PCs were presented in Table 12, which revealed that PC1 (21.299 %) was primarily influenced by SYPP (0.844), followed by NCP (0.626), PH (0.458), SW (0.260) and NSC (0.222). These are direct yield-related traits, indicating that PC1 serves as a selection axis for productivity traits. Similar conclusions were drawn by (47, 48), who found these traits as critical yield drivers in sesame. PC2 (15.441 %) was dominated by NLC (0.531), HI (0.388) and PB (0.369). These traits represent plant architecture and efficiency in biomass partitioning, essential for source-sink balance in sesame. This component thus complements PC1 in understanding physiological efficiency. PC3 (12.671 %) captured phenological traits, with high loadings for DFI (0.576), D50F (0.448) and D75M (0.350). PC4 (11.126 %) was associated with SHFC (0.345), CL (0.292), SW (0.261) and NSC (0.064), indicating its relevance for reproductive morphology and seed development. PC5 (9.256 %) was defined by NSC (0.339) and D75M (0.232). This component explained residual variation in seed setting and crop duration, contributing to finer differentiation among genotypes.

The biplot in Fig. 4 showed PC1 versus PC2 revealed relationships among traits and genotypes. Traits such as SYPP, NCP and PH were grouped in the direction of PC1, indicating their strong association with productivity. On the other hand, HI, NLC and PB were aligned more closely with PC2, suggesting these traits influence structural and physiological efficiency. The biplot helps in visual clustering of genotypes and identification of those combining favourable values for yield and associated traits (49). The contribution of different traits to significant PCs is further visualised through the corrplot in (Fig. 5). In this analysis, PC1 (21.299 %) was associated with SYPP, NCP, PH, SW and NSC, clearly indicating its dominance in yield expression. PC2 (15.441 %) was primarily influenced by NLC, HI and PB, all of which are indicative of efficient reproductive development and source sink balance. PC3 (12.671 %)

Table 11. Proportion of variance, cumulative proportion and eigen values in sesame (*Sesamum indicum* L.) genotypes

	Eigen value	Proportion of variance (%)	Cumulative proportion of variance (%)
PC1	2.982	21.299	21.299
PC2	2.162	15.441	36.740
PC3	1.774	12.671	49.412
PC4	1.558	11.126	60.538
PC5	1.296	9.256	69.794
PC6	1.115	7.963	77.757

Table 12. Five principal components along with their factor's loadings

	PC 1	PC 2	PC 3	PC 4	PC 5
Days to flower initiation	0.062	0.132	0.576	0.019	0.048
Days to 50 % flowering	0.022	0.056	0.448	0.167	0.055
Days to75 % maturity	0.048	0.071	0.350	0.043	0.232
Stem height of first capsule (cm)	0.191	0.010	0.066	0.345	0.113
Plant height (cm)	0.458	0.221	0.019	0.013	0.119
No. of primary branches	0.011	0.369	0.001	0.006	0.126
No. of capsules per plant	0.626	0.016	0.081	0.049	0.020
No. of locules per capsule	0.000	0.531	0.013	0.056	0.013
No. of seeds per capsule	0.222	0.138	0.079	0.064	0.339
1000 seed weight (g)	0.260	0.004	0.047	0.261	0.001
Capsule length (mm)	0.135	0.053	0.042	0.292	0.033
Harvest index (%)	0.090	0.388	0.039	0.116	0.083
Oil content (%)	0.011	0.095	0.012	0.119	0.094
Seed yield per plant (g)	0.844	0.077	0.000	0.007	0.021

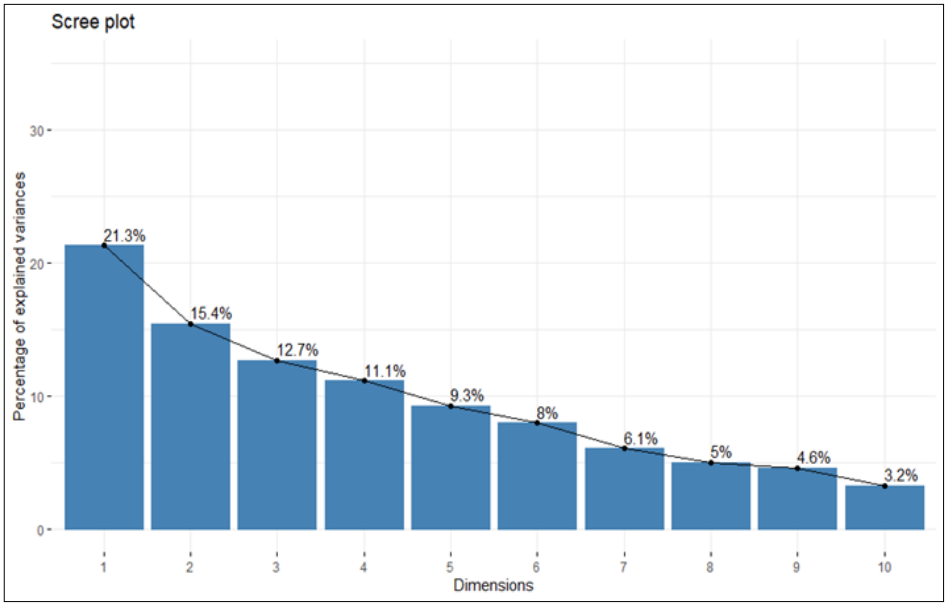

Fig. 3. Screen plot.

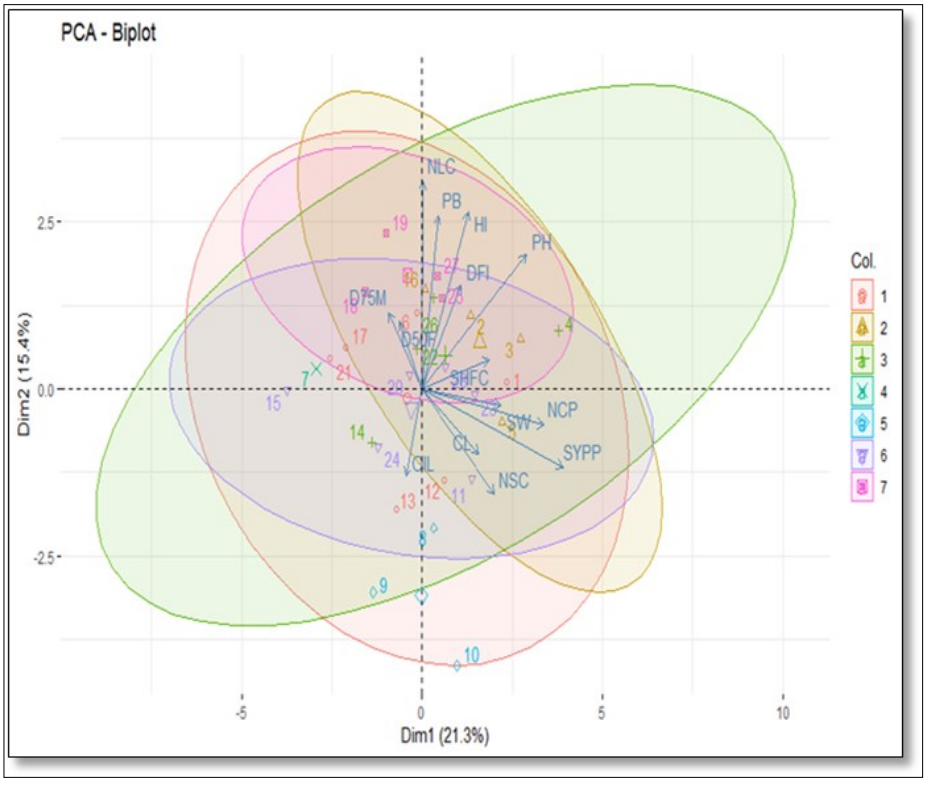

Fig. 4. Biplot among 14 traits and 28 in sesame genotypes

captured variation in DFI, D50F and D75M, highlighting its relevance for phenological classification and breeding for duration-specific genotypes. PC4 (11.126 %) captured variability due to SHFC, CL, SW and D50F, while PC5 (9.256 %) represented NSC, D75M and SHFC traits. These results suggest that any improvement in the traits heavily loaded on PC1 and PC2 will have a major impact on sesame improvement. This finding is consistent with the earlier studies that emphasised seed yield and structural traits are the principal drivers of variation in sesame and should be emphasised in breeding strategies.

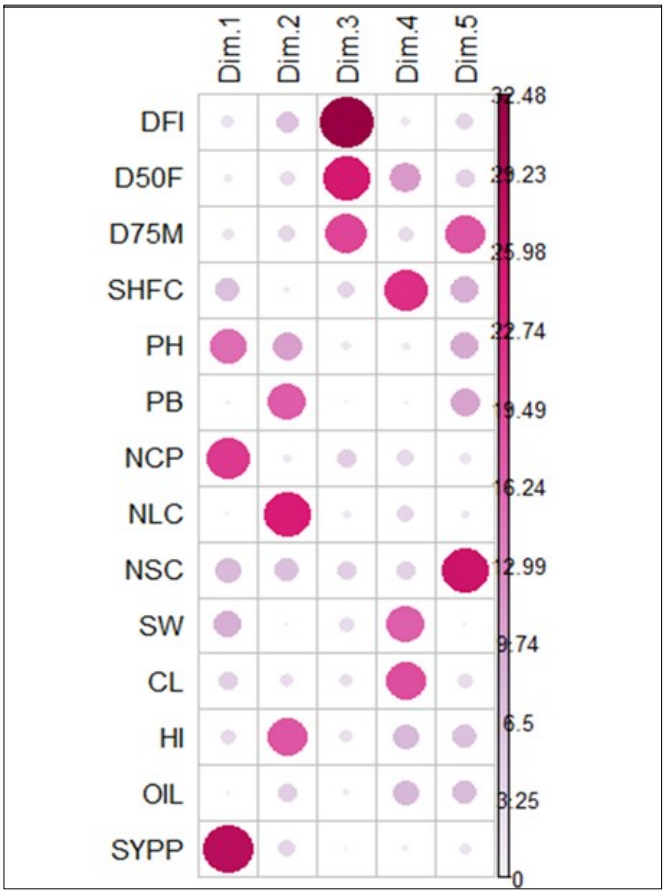

Fig. 5. Contribution of different yield attributing traits to significant.

Conclusion

The multivariate analysis of 28 sesame (*Sesamum indicum* L.) genotypes revealed significant genetic variability across 14 yield contributing and associated traits. High PCV, GCV, heritability and genetic advance for seed yield, OC, HI and DM indicated strong additive gene action, making these traits ideal for selection. Correlation and path analysis identified PH, CL and HI as key contributors to seed yield. Genetic divergence grouped genotypes into seven clusters. High inter-cluster distances suggested that crosses between diverse genotypes like IC-203945, NIC-1005, K-3-K-92, RT-54 and PRACHI could enhance genetic recombination. PC1 had the highest eigenvalue and explained 21.299 % of the total variance. Overall, these findings provide a strong foundation for the strategic selection of parents and key traits in breeding programs aimed at developing high-yielding and superior sesame cultivars.

Acknowledgements

The authors sincerely acknowledge the Department of Genetics and Plant Breeding and Seed Science and Technology, M. S. Swaminathan School of Agriculture, Centurion University of Technology and Management, Parlakhemundi, for their support in conducting this research and for providing the necessary infrastructure and agricultural facilities. The authors also extend their gratitude to the Agriculture Research Station, Mandor, Rajasthan, India, for facilitating access to plant genetic resources. Special thanks are due to the professors of the Department of Genetics and Plant Breeding for their continuous guidance, supervision and valuable contributions that enhanced the quality of this research.

Authors' contributions

SR and S conceptualised and designed the research. BVP conducted the field experiments. SS and DS did a formal statistical analysis. SSS and PS drafted the manuscript. PJ, MCI, NML and GI edited the manuscript. The findings were discussed and the manuscript was written by all of the authors. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflicts of interest to declare.

Ethical issues: None

References

1. Yadav R, Kalia S, Rangan P, Pradheep K, Rao GP, Kaur V, et al. Current research trends and prospects for yield and quality improvement in sesame, an important oilseed crop. Front Plant Sci. 2022;13:863521. https://doi.org/10.3389/fpls.2022.863521
2. Pathak N, Rai AK, Kumari R, Bhat KV. Value addition in sesame: A perspective on bioactive components for enhancing utility and profitability. Pharmacogn Rev. 2014;8(16):147–55. https://doi.org/10.4103/0973-7847.134249
3. Coniglio S, Shumskaya M, Vassiliou E. Unsaturated fatty acids and their immunomodulatory Properties. Biology (Basel). 2023;12 (2):279. https://doi.org/10.3390/biology12020279
4. Wei P, Zhao F, Wang Z, Wang Q, Chai X, Hou G, et al. Sesame (*Sesamum indicum* L.): a comprehensive review of nutritional value, phytochemical composition, health benefits, development of food and industrial applications. Nutrients. 2022;14(19):4079. https://doi.org/10.3390/nu14194079
5. Tripathi SM, Kalia A, Mishra BP, Mishra S, Mishra D, Shukla G. Trends in area production and productivity of sesame in bundelkhand region of Uttar Pradesh. Indian J Ext Educ. 2023;59(4):40–3.
6. Rauf S, Basharat T, Gebeyehu A, Elsafy M, Rahmatov M, Ortiz R, et al. Sesame, an underutilized oil seed crop: Breeding achievements and future challenges. Plants. 2024;13(18):2662. https://doi.org/10.3390/plants13182662
7. Hukumchand, Parameshwarappa SG. Genetically potential sesame (*Sesamum indicum* L.) breeding lines. J Farm Sci. 2020;33(03):302–5. https://doi.org/10.61475/jfm.v33i03.224
8. Rout S, Sur B, Sadhu S, Ghimiray TS, Mondal HA, Hijam L, et al. Trait's association, cause and effect analyses in Indian mustard [*Brassica juncea* (L.) Czern & Coss]. Electron J Plant Breed. 2019;10 (4):1482–94.
9. Sur B, Rout S, Singla S, Mandal R, Nath S, Maying B, et al. Evaluation of mustard genotypes [*Brassica juncea* (L.) Czern and Coss] for quantitative traits and character association of seed yield and yield components at sub Himalayan region of West Bengal (India). Plant Sci Today. 2023;10(1):166–78. https://doi.org/10.14719/pst.1948

10. Mawcha KT, Gebru MM, Assefa MK, Mesfin M, Gebre GG. Morphological characterization and genetic diversity of sesame (*Sesamum indicum* L.) varieties cultivated in Ethiopia. Open Agric J. 2020;14:117–29.
11. Mohanty TA, Singh UK, Singh SK, Singh D, Kushwaha N. Assessment of genetic diversity in Sesame (*Sesamum indicum* L.) based on agro-morphological traits. Curr J Appl Sci Technol. 2020;39(25):100–7. https://doi.org/10.9734/cjast/2020/v39i2530890
12. Umamaheswari D, Suganthi S, Thirumeni S, Kumar PS, Bhuvaneswari R. Principal component analysis, a multivariate technique to assess yield-contributory traits in the sesame germplasm. Ecol Environ Conserv. 2024;30(04):1939–43. https://doi.org/10.53550/EEC.2024.v30i04.077
13. Rout S, Roy SK, Mandal R, Singla S, Rahimi M, Sur B, et al. Genetic analysis and heterosis breeding of seed yield and yield attributing traits in Indian mustard (*Brassica juncea* (L.) Czern & Coss.). Sci Rep. 2025;15:2911. https://doi.org/10.1038/s41598-025-86621-8
14. Soxhlet F. The weight analytic determination of milk fat. Polytechnisches J. 1879;232:461–5.
15. Panse VG, Sukhatme PV. Statistical methods for agricultural workers. New Delhi: Indian Council of Agricultural Research; 1954.
16. Felipe M, Yaseen M. Agricolae: Statistical Procedures for Agricultural Research [Internet]. R package version 1.40.; 2020 [cited 2025 Oct 13]. https://myaseen208.github.io/agricolae/https://cran.r-project.org/package=agricolae
17. Burton GW. Quantitative in-heritance in grass. In: Proceedings of the Sixth International Grassland Congress 1952, Pennsylvania; 1952.
18. Sivasubramanian V, Madhavamenon P. Path analysis for yield and yield component of rice. Madras Agric J. 1973;60:1217–21.
19. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in Soybeans 1. Agron J. 1955;47. https://doi.org/10.2134/agronj1955.00021962004700070009x
20. Allard RW. Principles of Plant Breeding. New York, London: J. Wiley & Sons Ltd; 1961.
21. Popat R, Patel R, Parmar D. Genetic variability analysis for Plant Breeding research [Internet]. R package [internet]; 2020 [cited 2025 Oct 13]. https://cran.r-project.org/web/packages/variability/variability.pdf
22. Olivoto T, Lúcio AD. metan: An R package for multi-environment trial analysis. Methods Ecol Evol. 2020;11:783–9. https://doi.org/10.1111/2041-210X.13384
23. Wei T, Simko V, Levy M, Xie Y, Jin YJ, Zemla J. Visualization of a Correlation Matrix. R package "corrplot. Statistician. 2017;56.
24. Wright S. Correlation and Causation. J Agric Res. 1921;20:557–85.
25. Dewey DR, Lu KH. A Correlation and path-coefficient analysis of components of Crested Wheatgrass Seed Production 1. Agron J. 1959;51. https://doi.org/10.2134/agronj1959.00021962005100090002x
26. Mahalanobis PC. Reprint of: Mahalanobis, P.C. (1936) On the Generalised Distance in Statistics. Sankhya A. 1936;80:1–7. https://doi.org/10.1007/s13171-019-00164-5
27. Galili T. Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics. 2015;31(22):3718–20. https://doi.org/10.1093/bioinformatics/btv428
28. Lê S, Josse J, Husson F. FactoMineR: An R Package for multivariate analysis. J Stat Softw. 2008;25. https://doi.org/10.18637/jss.v025.i01
29. Kassambara A, Mundt F. Package 'factoextra' [Internet]. CRAN- R Package.; 2020 [cited 2025 Oct 13]. Available from: https://cran.r-project.org/web/packages/factoextra/factoextra.pdf
30. Vamshi P, Ramya KT, Prasad MS, Ramana J V, Kumhar SR. Genetic variability studies for yield and yield attributes in breeding lines of sesame (*Sesamum indicum* L.). J Oilseeds Res. 2021;38(3):240–3.
31. Abhijatha A, Arya K, Madhukar K, Gogineni S. Evaluation of Sesame (*Sesamum indicum* L.) genotypes to the shaded uplands of southern region. Int J Curr Microbiol App Sci. 2017;6(7):332–9. https://doi.org/10.20546/ijcmas.2017.607.039
32. Padmaja D, Babu TK, Rani TS, Reddy R U. Genetic variation in sesame genotypes (*Sesamum indicum* L.) grown in Telangana. J Oilseeds Res. 2020;37(Special issue). https://doi.org/10.56739/jor.v37iSpecialissue.139128
33. Mohanty TA, Singh UK, Singh SK, Kushwaha N, Singh D. Study of genetic variability, heritability and genetic advance in Sesame (*Sesamum indicum* L.) Genotypes. Int J Curr Microbiol Appl Sci. 2020;9(2):347–56. https://doi.org/10.20546/ijcmas.2020.902.044
34. Ranjithkumar G, Bisen R, Kumar V. Genetic variability studies for yield and its attributing traits in dark brown Sesame (*Sesamum indicum* L.). Int J Plant Soil Sci. 2022;34(8):43–9. https://doi.org/10.9734/ijpss/2022/v34i830897
35. Wang C, Niu J, Miao H, Li C, Duan Y, Ju M, et al. Genetic variation, correlation and association mapping of seed yield and its component traits in sesame. Genet Resour Crop Evol. 2024;71:603–19. https://doi.org/10.1007/s10722-023-01644-2
36. Gedifew S, Demelash H, Abate A, Abebe TD. Association of quantitative traits and genetic diversity in Ethiopian sesame (*Sesamum indicum* L.) genotypes. Heliyon. 2024;10:e26676. https://doi.org/10.1016/j.heliyon.2024.e26676
37. Abate M, Mekbib F. Study on genetic divergence in low-altitude sesame (*Sesamum indicum* L.) germplasm of Ethiopia based on agro morphological traits. J Agric Biol Environ Stat. 2015;2(3):78–90.
38. Sasipriya S, Paimala K, Eswari KB, Balram M. Correlation and path analysis for seed yield and its components in sesame (*Sesamum indicum* L.). Electron J Plant Breed. 2018;9(4):1594. https://doi.org/10.5958/0975-928X.2018.00199.0
39. Hukumchand, Parameshwarappa SG. Character Association and Path Analysis for Quantitative Traits in Advanced Breeding Lines of Sesame (*Sesamum indicum* L.). Int J Curr Microbiol Appl Sci. 2019;8(12):2980–90. https://doi.org/10.20546/ijcmas.2019.812.346
40. Srikanth K, Ghodke MK. Correlation and Path Analysis in Sesame (*Sesamum indicum* L.) Genotypes. Int J Plant Soil Sci. 2022;34(18):228–35. https://doi.org/10.9734/ijpss/2022/v34i1831075.
41. Swapna B, Thippani S, Padmaja D, Nayak SS. Genetic diversity for yield and yield contributing characters in Sesame (*Sesamum indicum* L.). Int J Environ Clim Change. 2023;13(8):1717–25.
42. Yadav S, Thiruvengadam BP V, Sasikala R, Rajendran L. Analysis of genetic diversity in sesame (*Sesamum indicum* L.) germplasm for yield and its attributing traits. Electron J Plant Breed. 2022;13(3):927–31. https://doi.org/10.37992/2022.1303.121
43. Nikfekr R, Kazemitabar SK, Ranjbar GA, Hashemi SH, Mehraban Joubani P. Study of genetic diversity and evaluation of some sesame genotypes under salinity stress. Environ Stress Crop Sci. 2023;16(3):575–86.
44. Baraki F, Berhe M. Evaluating Performance of Sesame (*Sesamum indicum* L.) Genotypes in Different Growing Seasons in Northern Ethiopia. Int J Agron. 2019;1–7. https://doi.org/10.1155/2019/7804621
45. Shim KB, Shin SH, Shon JY, Kang SG, Yang WH, Heu SG. Classification of a collection of sesame germplasm using multivariate analysis. J Crop Sci Biotechnol. 2016;19(2):151–5. https://doi.org/10.1007/s12892-015-0099-4
46. Baraki F, Gebregergis Z, Belay Y, Berhe M, Teame G, Hassen M, et al. Multivariate analysis for yield and yield-related traits of sesame (*Sesamum indicum* L.) genotypes. Heliyon. 2020;6(10):e05295. https://doi.org/10.1016/j.heliyon.2020.e05295
47. Biabani AR, Pakniyat H. Evaluation of seed yield-related characters in Sesame (*Sesamum indicum* L.) using factor and path analysis. Pak J Biol Sci. 2008;11(8):1157–60. https://doi.org/10.3923/pjbs.2008.1157.1160
48. Asghari A, Dargahi Y, Rasoulzadeh A, Ahmadian M. Study of sesame

(*Sesame indicum* L.) cultivars based on morphological characteristics under water deficit stress condition using factor analysis. [journal name unknown]. 2014;14(11):593–607.

49. Kumar A, Swapnil, Perween S, Singh RS, Singh DN. Prospects of molecular markers in precision Plant Breeding. In: Jha AK, Kumar U, editors. Recent advances in Chemical Sciences and Biotechnology. New Delhi: New Delhi Publishers; 2020. p. 132–42.

50. Swapnil, Jaishreepriyanka R, Singh D, Mandal SS. Principal component analysis in maize (*Zea mays* L.) under normal sown condition of Bihar. Pharma Innov J. 2021;10:641–4.

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.