Research Note Electronic Journal of Plant Breeding, Vol.17(2): 188-197

Multivariate analysis for earliness, yield and fibre quality traits in cotton using correlation, PCA and cluster analysis

N. S. Navingautham[1], K. Mohan Vishnuvardhan*[2], N. Sabitha[1], G. Mohan Naidu[3] and M. Reddisekhar[1]

https://doi.org/10.37992/2026.1702.020

Abstract

The present study was undertaken to assess the genetic variability, trait associations and genetic divergence among 40 cotton genotypes using correlation, Principal Component Analysis (PCA) and cluster analysis. Analysis of variance revealed significant variability for all the traits studied, indicating ample scope for selection and improvement. Correlation analysis showed that plant height, number of bolls per plant, fibre length and elongation percentage had significant positive associations with seed cotton yield, identifying them as important yield-contributing traits. Earliness traits were positively interrelated, while fibre quality parameters such as fibre length, fibre strength, uniformity index and elongation percentage were positively correlated among themselves and with yield, suggesting the possibility of simultaneous improvement of yield and fibre quality. PCA revealed that seven principal components with eigen values greater than one explained 75.6% of the total variability, with PC-I contributing the highest proportion and largely influenced by yield and fibre quality traits. The bi-plot analysis identified several genetically divergent genotypes positioned away from the origin, indicating their potential as promising parents in hybridization programmes. Cluster analysis grouped the genotypes into five distinct clusters, with Cluster II characterized by superior yield and fibre quality traits, and Cluster III by earliness and higher lint index. The substantial inter-cluster distances suggested the potential for exploiting genetic divergence through hybridization to develop high-yielding, early-maturing and superior fibre quality cotton cultivars. Overall, the combined multivariate approaches provided valuable insights for effective parent selection and genetic improvement in cotton breeding programmes.

Keywords: *Gossypium hirsutum,* Correlation, Genetic diversity, Earliness, Seed cotton yield.

Cotton (*Gossypium* spp.), known as “white gold,” is a vital fibre crop supporting the textile industry and millions of farmers worldwide. India, the second-largest producer, contributes about 23% of global cotton output (USDA Cotton Outlook, 2026), with cultivation across diverse agro-climatic zones. The limited genetic diversity in commercial cotton cultivars remains a major bottleneck to productivity enhancement, resulting in slow yield gains (Naveed *et al.,* 2023; Ali *et al.,* 2025). Broadening the genetic base by utilizing diverse parental lines and germplasm resources is essential for developing high-yielding cultivars with superior fibre quality (Chapara *et al.,* 2025; Shrestha *et al.,* 2025). Earliness is a particularly important trait, enabling crops to escape late-season stresses, reduce costs, and support sustainable cropping systems.

Developing superior genotypes requires understanding the interactions among agronomic and fibre quality traits. Correlation analysis provides insights into the relationships among traits, while multivariate techniques such as Principal Component Analysis (PCA) and cluster analysis offer a comprehensive assessment of genetic diversity and trait interactions (Kumar *et al*., 2025; Bhanuchand and Raju, 2025; Gowda *et al*., 2026) PCA effectively reduces datasets to a few principal components, highlighting the most influential traits and grouping genotypes based on overall performance (Kumar *et al*., 2025). Cluster analysis further classifies genotypes with similar traits, aiding in the identification of genetically distinct and promising lines. Accordingly, the present study was conducted to assess diversity among *Gossypium hirsutum* genotypes with respect to earliness, yield, and fibre quality traits.

---

[1]Department of Genetics and Plant Breeding, S.V. Agricultural College, Tirupati, Andhra Pradesh -517502
[2]Cotton Breeding, Regional Agricultural Research Station, Nandyal, Andhra Pradesh -518502
[3]Department of Statistics and Computer Applications, S.V. Agricultural College, Tirupati, Andhra Pradesh -517502

Corresponding author: N. S. Navingautham, Department of Genetics and Plant Breeding, S.V. Agricultural College, Tirupati, Andhra Pradesh -517502. *E-mail: k.mohanvishnuvardhan@angrau.ac.in

How to cite this artilce: Navingautham, N. S., Mohan Vishnuvardhan, K., Sabitha, N., Mohan Naidu, G. and Reddisekhar, M. 2026. Multivariate analysis for earliness, yield and fibre quality traits in cotton using correlation, PCA and cluster analysis. *Electronic Journal of Plant Breeding.* 17(2):188-197. doi: 10.37992/2026.1702.020

Received: 15.12.2025 **Revised:** 06.04.2026 **Accepted**: 13.04.2026

In this study, 34 American cotton genotypes and six check varieties, namely Srirama, NDLH -2051-1, NDLH -2035-5, NDLH -2056-4, Aadhya-BG-II, Navaneet-BG-II, were evaluated for variability, correlation and diversity during *kharif* season of 2024-25. The genotypes were raised in Alpha Lattice Design (ALD) (Patterson and Williams, 1976) with two replications in the research farm of Regional Agricultural Research Station, Nandyal, Andhra Pradesh, India (**Table 1**). The genotypes were evaluated using an Alpha Lattice Design (ALD) with two replications with a

Table 1. List of genotypes used in the present study

S.No.	Cotton genotypes	Parentage	Source
Advanced Breeding Lines (ABLs)			
1	NDLH - 2027	NDLH-1839 × BN-1	RARS, Nandyal, AP
2	NDLH - 1981	NA-1588 × NA-1799	RARS, Nandyal, AP
3	NDLH - 2015	NA-1325 × FFLC-7	RARS, Nandyal, AP
4	NDLH - 2008	NA-1755 × CCH-5266/2	RARS, Nandyal, AP
5	NDLH - 2027-2	NDLH-1839 × BN-1	RARS, Nandyal, AP
6	NDLH - 1975	NA-1678 × CPD-448	RARS, Nandyal, AP
7	NDLH - 1982	NA-1588 × NA-1678	RARS, Nandyal, AP
8	NDLH - 2010	NDLH-1325 × FFLC-1	RARS, Nandyal, AP
9	NDLH - 2030-2	NDLH-1325 × Surabhi	RARS, Nandyal, AP
10	NDLH - 2049-4	BS-79 × HYPS-152	RARS, Nandyal, AP
11	NDLH - 2056-3	BS-79 × MCU-5	RARS, Nandyal, AP
12	NDLH - 2029-4	NDLH-1938 × BN-1	RARS, Nandyal, AP
13	NDLH - 1988	NDLH-1325 × HAG-1055	RARS, Nandyal, AP
14	NDLH - 2041-3	ADB-108 × HYPS-152	RARS, Nandyal, AP
15	NDLH - 1939	Biofemale × RAH-821	RARS, Nandyal, AP
16	NDLH - 2043-3	NDLH-1755 × NDLH-1956	RARS, Nandyal, AP
17	NDLH - 2050-2	NDLH-1931 × HYPS-152	RARS, Nandyal, AP
18	NDLH - 2077-1	NDLH-1938 × P-2151	RARS, Nandyal, AP
19	NDLH - 1937	NA-1325 × NSPHH-9	RARS, Nandyal, AP
20	NDLH - 1941	NA-1755 × MCU-12	RARS, Nandyal, AP
21	NDLH - 1945	NA–1325 × 996H-14	RARS, Nandyal, AP
22	NDLH - 1963	NA-1678 × SA-349	RARS, Nandyal, AP
23	NDLH - 1964	NA-1678 × SA-454	RARS, Nandyal, AP
24	NDLH - 1922	NA-1678 × AC-738	RARS, Nandyal, AP
25	NDLH - 1953	NA-1325 × 996H55	RARS, Nandyal, AP
26	NDLH - 2020	NA-1755 × FFLC-14	RARS, Nandyal, AP
27	NDLH - 2051-3	HAG-823 × HYPS-152	RARS, Nandyal, AP
28	NDLH - 2061-4	NDLH-1325 × NH-635	RARS, Nandyal, AP
29	NDLH - 2071-2	NDLH-1755 × GBHV-164	RARS, Nandyal, AP
30	NDLH - 2074-1	NDLH-1755 × RAH-911	RARS, Nandyal, AP
31	NDLH - 1906	ADB-320 × HYPS-152	RARS, Nandyal, AP
32	NDLH - 1918	ADB-320 × [NA-1678 × NDH-6]	RARS, Nandyal, AP
33	NDLH - 1927	NA-1325 × GBHB-146	RARS, Nandyal, AP
34	NDLH - 1929	NA-1325 × LRA-5166	RARS, Nandyal, AP
Checks			
35	Srirama ©	NDLH-1797 × NDLH 1325	RARS, Nandyal, AP
36	NDLH - 2051-1	HAG-823 × HYPS-152	RARS, Nandyal, AP
37	NDLH - 2035-5	NDLH-1905 × MCU-5	RARS, Nandyal, AP
38	NDLH - 2056-4	BS-79 × MCU-5	RARS, Nandyal, AP
39	Navaneet - BG-II	BG II cotton hybrid from Nuziveedu Seeds Pvt. Ltd	
40	Aadhya-BG-II	BG II cotton hybrid from Nuziveedu Seeds Pvt. Ltd	

spacing of 60 cm x 30 cm and row length of 5m. In each replication 4 blocks and in each block 10 genotypes were maintained. All recommended agronomic practices and inputs were properly followed throughout the experiment. For the traits *viz.,*days to first squaring formation, days to first flowering, days to 50% flowering and days to first boll opening were collected by counting the number of days on plot basis; the traits *viz.,* nodes to first fruiting branch, plant height (cm), number of monopodial branches/plant, number of sympodial branches/plant, number of bolls/plant, boll weight (g), ginning out-turn (%), lint index (g) and seed cotton yield (kg/ha) were taken on five randomly selected plants in each plot; and the fibre quality traits *viz.,* fibre length (mm), fibre strength (g/tex), micronaire (µg/inch), elongation percentage and uniformity index (%) were measured with the help of the High Volume Instrument (HVI) operated in HVI mode at ICAR-CIRCOT, Coimbatore.

Statistical analysis

The statistical procedures *viz.,* correlation (Johnson *et al.,* 1955), principal component analysis (Sneath and Sokal, 1973) and cluster analysis using Ward's minimum variance method (Ward, 1963) were carried out using R-software Version 4.4.3 (R Core Team, 2025).

Correlation studies

The basic statistical analysis indicated a considerable degree of variability among the 40 cotton genotypes for the traits studied (**Table 2**). At genotypic level, seed cotton yield showed a strong positive correlation with plant height (0.724**), bolls per plant (0.678**), fibre length (0.334*) and elongation percentage (0.341*). The results indicate that traits such as plant height and bolls per plant are major yield contributing characters, while fibre length and elongation percentage also play an important role in enhancing yield. Association among yield components, number of sympodial branches/plant showed a positive association with elongation percentage (0.632**) and uniformity index (0.490**). Plant height exhibited a significant positive correlation with number of sympodial branches/plant (0.565**) and uniformity index (0.520**) at the genotypic level. Fibre quality traits exhibited strong interrelationships. Fibre length showed a significant positive correlation with fibre strength (0.507**) and uniformity index (0.434**). Fibre strength also had a strong positive association with uniformity index (0.737**) and elongation percentage (0.987**). In addition, uniformity index was highly correlated with elongation percentage (0.948**), indicating the possibility of simultaneous improvement of fibre quality traits through selection. The similar type of association with yield and between yield and fibre traits were also studied by Kumar *et al.* (2024), Bhanuchand and Raju, (2025) and Gowda *et al.* (2026). The traits *viz.*, ginning out-turn exhibited a negative association with fibre length (-0.416**) and elongation percentage (-0.384*); Plant height showed a strong negative association with number of monopodial branches/plant (-0.543**). Similarly, days to flowering traits showed weak or negative associations with yield, indicating that earliness traits may not directly contribute to yield improvement. The negative association of these traits were studied earlier by Kumar *et al.* (2024). The correlation analysis revealed that bolls per plant, plant height, fibre length and elongation percentage are

Table 2. Trait-wise statistical overview of 40 cotton genotypes

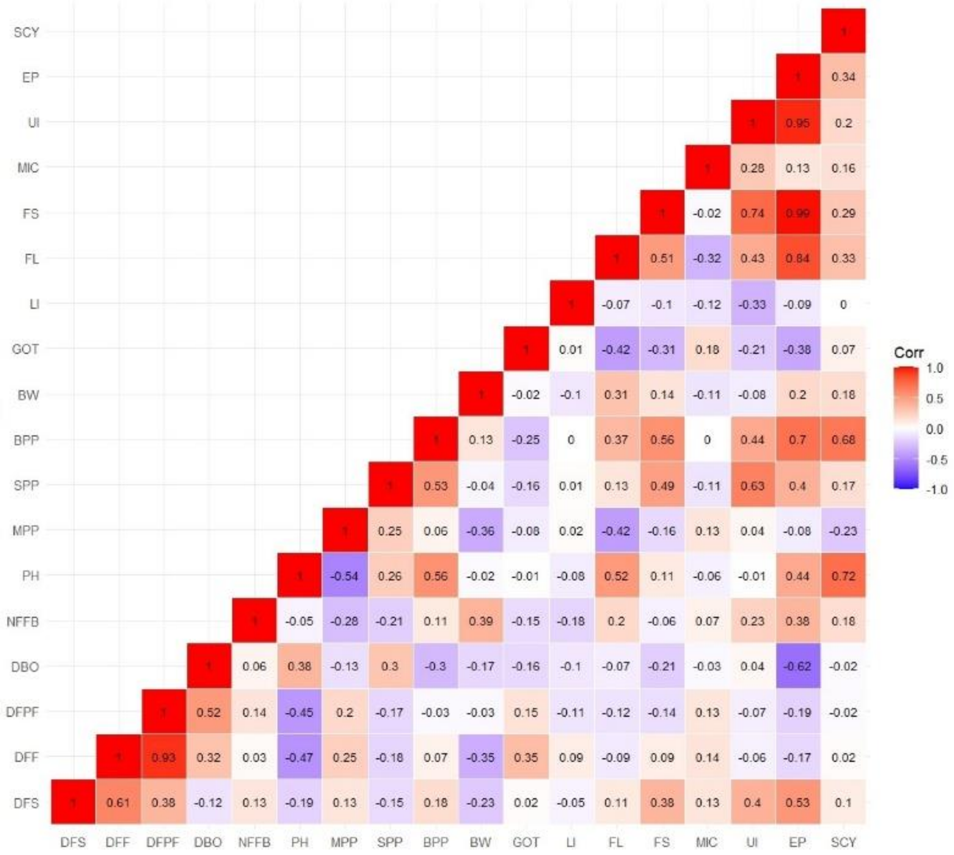
S. No.	Traits	Minimum	Maximum	Mean	S.E.Mean
1	Days to first squaring formation	30.5	36.5	33.3	0.6
2	Days to first flowering	47.0	52.0	50.3	0.6
3	Days to 50% flowering	56.0	61.5	59.1	0.6
4	Days to first boll opening	107.5	121.5	115.1	1.8
5	Nodes to first fruiting branch	5.1	6.5	5.8	0.2
6	Plant height (cm)	111.6	140.4	122.3	5.3
7	Number of monopodial branches/plant	0.6	3.1	1.7	0.2
8	Number of sympodial branches/plant	13.0	24.2	18.9	1.0
9	Number of bolls/plant	21.4	39.4	31.1	1.9
10	Boll weight (g)	3.5	4.9	4.0	0.2
11	Ginning out turn (%)	34.4	38.4	36.3	0.4
12	Lint Index (g)	4.0	6.2	4.9	0.09
13	Fibre length (mm)	25.9	30.7	28.2	0.4
14	Fibre strength (g/tex)	25.2	30.5	27.8	0.3
15	Micronaire (µg/inch)	3.5	4.3	3.9	0.04
16	Elongation percentage	5.4	5.7	5.6	0.05
17	Uniformity index (%)	79.0	85.0	81.7	1.0
18	Seed cotton yield (kg/ha)	711.7	1771.7	1224.6	104.9

Table 3. Genotypic correlation analysis of earliness, yield and fibre quality traits in upland cotton

	DFS	DFF	DFPF	DBO	NFFB	PH	MPP	SPP	BPP	BW	GOT	LI	FL	FS	MIC	UI	EP	SCY
DFS	**1.000**	0.609 **	0.385 *	-0.121	0.129	-0.185	0.132	-0.150	0.184	-0.226	0.023	-0.051	0.109	0.377 *	0.134	0.397 *	0.530 **	0.104
DFF		**1.000**	0.932 **	0.322 *	0.033	-0.469 **	0.253	-0.177	0.067	-0.348 *	0.351 *	0.092	-0.088	0.093	0.143	-0.060	-0.174	0.019
DFPF			**1.000**	0.520 **	0.145	-0.455 **	0.198	-0.169	-0.027	-0.027	0.151	-0.110	-0.117	-0.137	0.130	-0.074	-0.191	-0.021
DBO				**1.000**	0.063	0.382 *	-0.128	0.299	-0.303	-0.175	-0.159	-0.102	-0.066	-0.205	-0.032	0.044	-0.625 **	-0.020
NFFB					**1.000**	-0.049	-0.278	-0.209	0.110	0.388 *	-0.152	-0.182	0.195	-0.063	0.072	0.232	0.385 *	0.176
PH						**1.000**	-0.543 **	0.262	0.565 **	-0.021	-0.009	-0.078	0.520 **	0.113	-0.057	-0.009	0.440 **	0.724 **
MPP							**1.000**	0.249	0.057	-0.361 *	-0.078	0.015	-0.421 **	-0.155	0.134	0.036	-0.081	-0.231
SPP								**1.000**	0.530 **	-0.041	-0.162	0.011	0.135	0.490 **	-0.108	0.632 **	0.400 *	0.171
BPP									**1.000**	0.130	-0.254	0.001	0.366 *	0.557 **	0.001	0.444 **	0.699 **	0.678 **
BW										**1.000**	-0.020	-0.100	0.308	0.137	-0.105	-0.083	0.196	0.185
GOT											**1.000**	0.008	-0.416 **	-0.312	0.178	-0.210	-0.384 *	0.069
LI												**1.000**	-0.069	-0.097	-0.115	-0.334 *	-0.091	0.003
FL													**1.000**	0.507 **	-0.322 *	0.434 **	0.839 **	0.334 *
FS														**1.000**	-0.022	0.737 **	0.987 **	0.290
MIC															**1.000**	0.285	0.130	0.158
UI																**1.000**	0.948 **	0.196
EP																	**1.000**	0.341 *

* & ** Significant at 1 & 5 % probability level, respectively

DFS: Days to first squaring formation; DFF: Days to first flowering; DFPF: Days to 50% flowering; DBO: Days to first boll opening; NFFB: Nodes to first fruiting branch; PH: Plant height ; MPP: Number of monopodial branches/plant; SPP: Number of sympodial branches/plant ; BPP: Number of bolls/plant; BW: Boll Weight; GOT: Ginning out turn ; LI: Lint Index ; FL: Fibre length ; FS: Fibre Strength ; MIC: Micronaire ; UI: Uniformity Index ; EP: Elongation percentage; SCY: Seed cotton yield .



	DFS	DFF	DFPF	DBO	NFFB	PH	MPP	SPP	BPP	BW	GOT	LI	FL	FS	MIC	UI	EP	SCY
SCY																		1
EP																	1	0.34
UI																1	0.95	0.2
MIC															1	0.28	0.13	0.16
FS														1	-0.02	0.74	0.99	0.29
FL													1	0.51	-0.32	0.43	0.84	0.33
LI												1	-0.07	-0.1	-0.12	-0.33	-0.09	0
GOT											1	0.01	-0.42	-0.31	0.18	-0.21	-0.38	0.07
BW										1	-0.02	-0.1	0.31	0.14	-0.11	-0.08	0.2	0.18
BPP									1	0.13	-0.25	0	0.37	0.56	0	0.44	0.7	0.68
SPP								1	0.53	-0.04	-0.16	0.01	0.13	0.49	-0.11	0.63	0.4	0.17
MPP							1	0.25	0.06	-0.36	-0.08	0.02	-0.42	-0.16	0.13	0.04	-0.08	-0.23
PH						1	-0.54	0.26	0.56	-0.02	-0.01	-0.08	0.52	0.11	-0.06	-0.01	0.44	0.72
NFFB					1	-0.05	-0.28	-0.21	0.11	0.39	-0.15	-0.18	0.2	-0.06	0.07	0.23	0.38	0.18
DBO				1	0.06	0.38	-0.13	0.3	-0.3	-0.17	-0.16	-0.1	-0.07	-0.21	-0.03	0.04	-0.62	-0.02
DFPF			1	0.52	0.14	-0.45	0.2	-0.17	-0.03	-0.03	0.15	-0.11	-0.12	-0.14	0.13	-0.07	-0.19	-0.02
DFF		1	0.93	0.32	0.03	-0.47	0.25	-0.18	0.07	-0.35	0.35	0.09	-0.09	0.09	0.14	-0.06	-0.17	0.02
DFS	1	0.61	0.38	-0.12	0.13	-0.19	0.13	-0.15	0.18	-0.23	0.02	-0.05	0.11	0.38	0.13	0.4	0.53	0.1

Fig. 1. Genotypic correlation heat map for earliness, yield and fibre quality traits

Residual effect for genotypic correlation: (r_g) = 0.465

DFS: Days to first squaring formation; DFF: Days to first flowering; DFPF: Days to 50% flowering; DBO: Days to first boll opening; NFFB: Nodes to first fruiting branch; PH: Plant height ; MPP: Number of monopodial branches/plant; SPP: Number of sympodial branches/plant; BPP: Number of bolls/plant; BW: Boll Weight ; GOT: Ginning out turn ; LI: Lint Index ; FL: Fibre length ; FS: Fibre strength ; MIC: Micronaire ; UI: Uniformity index ; EP: Elongation percentage; SCY: Seed cotton yield .

key traits contributing to seed cotton yield. The strong interrelationships among fibre quality traits suggest that simultaneous improvement of yield and quality is feasible, although careful selection is required to manage negative associations among certain traits.

Principal component and cluster analysis

Principal Component Analysis (PCA) is a widely used multivariate statistical technique for data comparison, dimensionality reduction, and transformation, which converts correlated variables into a smaller set of uncorrelated principal components (Kumar *et al.,* 2025). This approach facilitates the identification of underlying patterns, relationships, and genetic diversity among genotypes. In this study, seven out of eighteen principal components were selected with eigen value greater than 1 (**Table 4 and Fig. 2**). Principal Component Analysis (PCA) revealed that seven principal components (eigen value >1) together explained 75.6% of the total variability among the eighteen traits studied. PC-I contributed the highest proportion (21.3%), followed by PC-II (14.2%), PC-III (10.2%), PC-IV (8.9%), PC-V (8.3%), PC-VI (7.2%) and PC-VII (5.6%). PC-I was mainly governed by elongation percentage, micronaire, number of bolls per plant, fibre length, uniformity index, seed cotton yield, number of sympodial branches per plant and plant height, indicating the predominant influence of yield and fibre quality traits on overall variability. PC-II was largely associated with earliness traits (days to first squaring, flowering and boll opening), number of monopodial branches/plant, fibre strength and uniformity index. The subsequent components (PC-III to PC-VII) reflected variation due to boll weight, nodes to first fruiting branch, days to 50% flowering, ginning out-turn, lint index, fibre strength and branch number, highlighting the combined contribution of phenological, yield and fibre traits to genetic divergence. The similar findings were also obtained by Chapara *et al.* (2022), Sagar *et al.* (2023), Pavan *et al.* (2025), Ujjainkar, (2025) and Rehman *et al.* (2025). The PC bi-plot illustrated the

Table 4. Principal component analysis of various traits in cotton genotypes

Parameter	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Eigen Value (Root)	3.826	2.552	1.831	1.601	1.499	1.295	1.002
% Var. Exp.	21.3	14.2	10.2	8.9	8.3	7.2	5.6
Cum. Var. Exp.	21.3	35.4	45.6	54.5	62.8	70.0	75.6
Characters							
Days to first squaring formation	0.2301	0.6786	0.1925	-0.2549	0.1331	0.0768	-0.2813
Days to first flowering	-0.1332	0.8308	0.2480	0.1483	0.0793	0.3233	-0.0294
Days to first boll opening	-0.2038	0.7047	0.4136	0.2677	-0.1348	0.1438	0.1944
Days to 50% flowering	-0.1528	0.1751	0.0246	0.7486	-0.5180	0.0150	-0.0456
Nodes to first fruiting branch	0.1352	-0.0161	0.6342	-0.0627	-0.1897	-0.2813	0.3293
Plant height	0.4298	-0.3642	-0.0425	0.6384	0.1827	-0.0049	-0.2469
Number of monopodial branches/plant	-0.2206	0.4870	-0.5536	-0.1745	0.0620	-0.0563	0.4328
Number of sympodial branches/plant	0.4596	0.0615	-0.5868	0.2751	-0.1965	0.0284	0.2732
Number of bolls/plant	0.7295	0.1099	-0.0659	0.1765	0.2986	0.1306	0.3694
Boll weight	0.2221	-0.3883	0.5404	-0.1388	-0.0743	-0.0612	0.3957
Ginning out turn	-0.3751	0.0466	0.1931	0.1879	0.5732	-0.0929	-0.1789
Lint Index	-0.1510	-0.1104	-0.1378	-0.0521	0.3264	0.6090	0.0817
Fibre length	0.7065	-0.1413	0.2940	-0.0559	-0.2134	0.3011	-0.1788
Micronaire	0.7973	0.2039	-0.1046	-0.1848	-0.0419	0.0961	-0.1061
Fibre strength	-0.0115	0.3131	-0.0273	0.1029	0.3946	-0.6753	-0.0266
Elongation percentage	0.8219	0.1409	0.0181	-0.2951	0.0742	-0.0465	-0.0849
Uniformity index	0.6212	0.3724	-0.1562	0.0174	-0.2621	-0.3497	-0.1764
Seed cotton yield	0.5354	-0.0683	0.1757	0.3754	0.5192	0.0103	0.1921

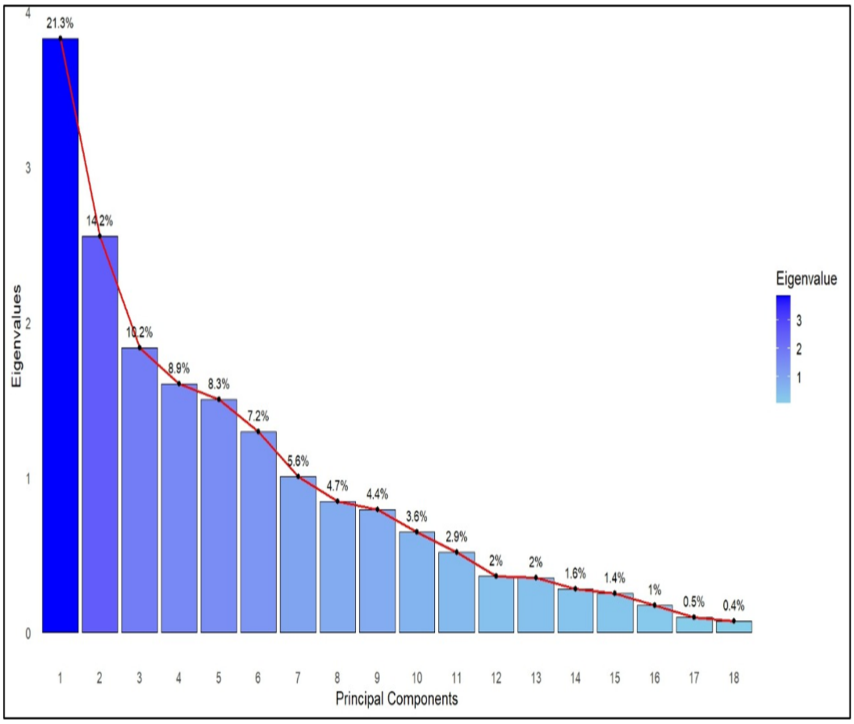


Fig. 2. Scree plot showing the eigen values and percentage of variance explained by each principal components

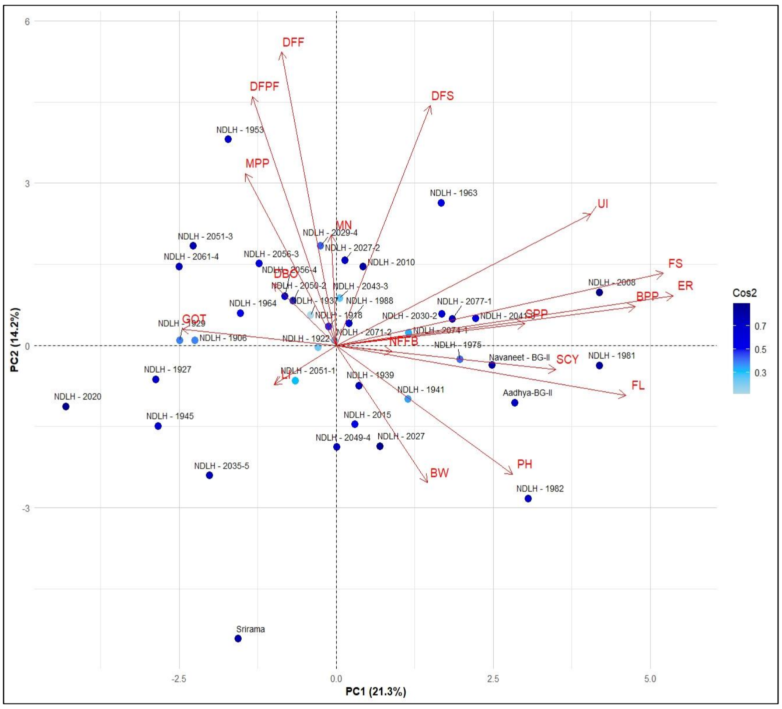


Fig. 3. Bi-plot illustrating the contribution of variables to principal components 1 and 2

distribution of genotypes and trait vectors, showing that elongation percentage, micronaire, number of bolls/plant, fibre length, uniformity index and earliness parameters contributed maximally to variability. Genotypes such as Srirama, NDLH-2020, NDLH-2061-4, NDLH-2051-3, NDLH-1953, NDLH-1963, NDLH-2008, NDLH-1981 and NDLH-1982 were positioned far from the origin, indicating greater genetic divergence and their potential use as parents in hybridization programmes for American cotton improvement. Similar findings were also recorded earlier by Ujjainkar, (2025) and Rehman *et al.* (2025). Early maturing genotypes included Srirama, NDLH-2049-4, NDLH-2027 and NDLH-1982, while Navaneet-BG-II, Aadhya-BG-II, NDLH-1982 and NDLH-2077-1 exhibited higher seed cotton yield. Overall, PCA effectively identified diverse and superior genotypes for utilization in breeding programmes aimed at enhancing yield, fibre quality and earliness.

Cluster analysis grouped the 40 cotton genotypes into five clusters. Cluster I was the largest (11 genotypes), followed by Clusters V (10), III (8), II (6), and IV (5) (**Table 5**). Clusters II and III were identified as the most promising for improvement. Cluster II exhibited superior performance for seed cotton yield, boll weight, fibre length, fibre strength, and uniformity index, while Cluster III was characterized by earliness traits (fewer days to squaring formation, flowering, and 50% flowering, and fewer nodes to first fruiting branch) along with higher sympodial branches and lint index, contributing to lint yield. These results were similar to the findings of Prakash and Suthar, (2023), Pavan *et al.* (2025), Ujjainkar, (2025) and Rehman *et al.* (2025), demonstrating key values for yield-contributing and fibre traits across clusters. Intra-cluster distances indicated the highest variability in Cluster I (5.29), suggesting greater genetic diversity within this group. The highest inter-cluster distance was observed between Clusters I and II (6.98), followed by Clusters I and III (6.40) and Clusters II and V (6.12), indicating wider genetic divergence among these groups. Crosses between genotypes from highly divergent clusters may therefore be effective for generating superior recombinants and enhancing yield and fibre traits in cotton. The similar findings were also noticed by Prakash and Suthar, (2023), Ali *et al.* (2025) and Kumar *et al.* (2025). The dendrogram (**Fig. 4**) clearly grouped the genotypes into distinct clusters and sub-clusters. Cluster II showed superiority in yield and fibre quality traits, whereas Cluster III exhibited desirable earliness and higher lint index. The complementary strengths and genetic divergence between these clusters highlight their potential as promising parents for developing high-yielding, early-maturing, and superior fibre quality cotton cultivars.

The present investigation revealed substantial genetic variability among the 40 cotton genotypes, providing significant scope for effective selection and crop improvement. The strong positive associations of plant height, number of bolls per plant, fibre length and elongation percentage with seed cotton yield indicate that these traits can serve as reliable selection criteria for yield enhancement. The positive interrelationship among fibre quality traits and their association with yield further

Table 5. Grouping of cotton genotypes based on cluster analysis

Cluster No.	Number of genotypes	Name of genotypes
I	11	NDLH-2051-3, NDLH-2061-4, NDLH-2056-4, NDLH-2035-5, NDLH-2020, NDLH-1927, NDLH-1929, NDLH-2056-3, NDLH-1906, NDLH-1964, NDLH-1953
II	6	NDLH-2077-1, Aadhya-BG-II, Navaneet-BG-II, NDLH-1981, NDLH-2008, NDLH-2030-2
III	8	Srirama, NDLH-2074-1, NDLH-1941, NDLH-2051-1, NDLH-1982, NDLH-2049-4, NDLH-2027, NDLH-2015
IV	5	NDLH-1963, NDLH-2010, NDLH-2029-4, NDLH-1988, NDLH-2043-3
V	10	NDLH-1939, NDLH-1918, NDLH-2027-2, NDLH-2041-3, NDLH-2050-2, NDLH-1945, NDLH-1922, NDLH-2071-2, NDLH-1975, NDLH-1937

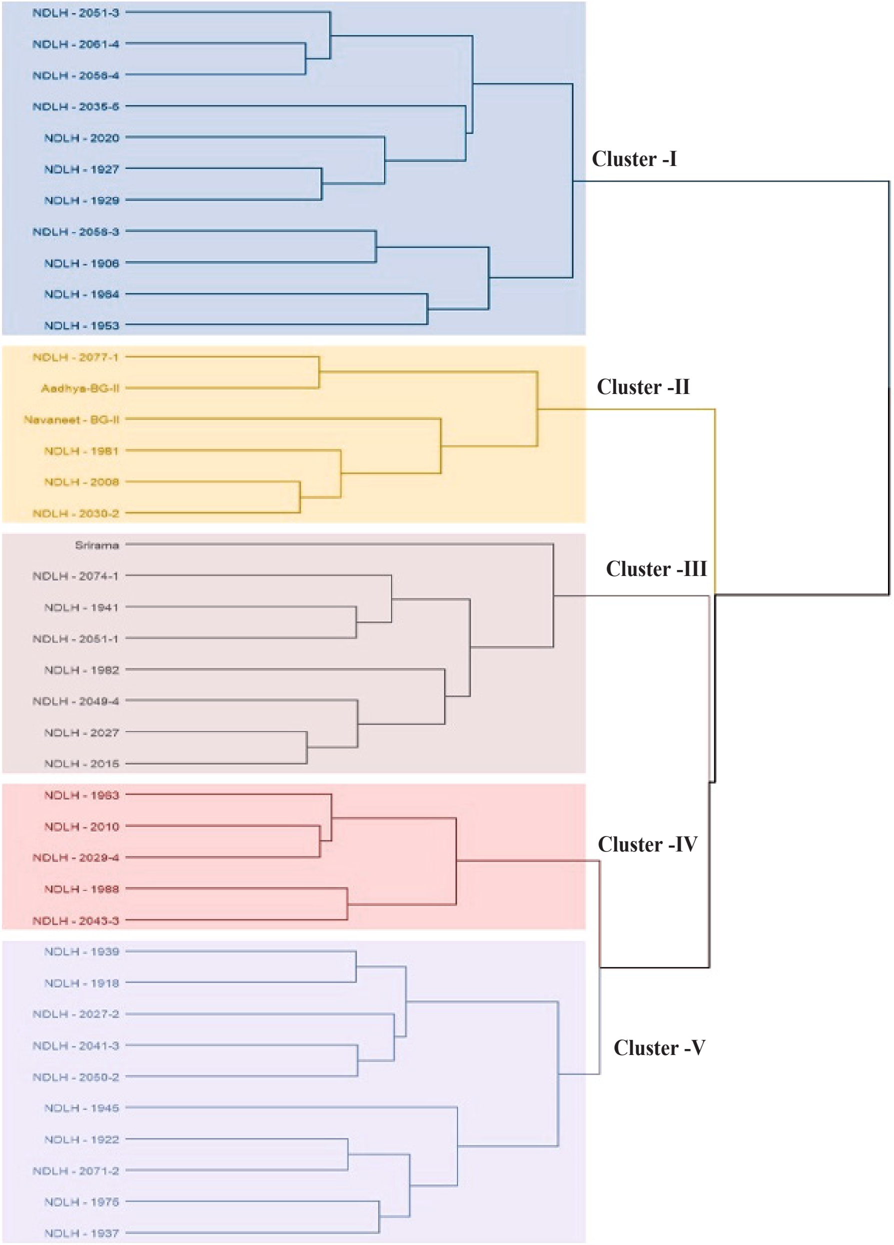


Fig. 4. Dendrogram depicting the clustering pattern of 40 cotton genotypes

Table 6. Cluster means for earliness, yield and fibre quality traits in 40 cotton genotypes

Traits	Cluster number				
	I	II	III	IV	V
Days to first squaring formation	33.00	34.00	31.94	35.10	33.30
Days to first flowering	50.82	50.33	49.13	50.80	50.45
Days to first boll opening	59.86	59.17	57.75	59.30	59.30
Days to 50% flowering	117.05	116.08	112.88	111.30	115.85
Nodes to first fruiting branch	5.82	5.90	5.61	5.72	5.76
Plant height (cm)	121.01	129.85	127.68	117.92	117.11
Number of monopodial branches/plant	1.79	1.24	1.80	1.78	1.71
Number of sympodial branches/plant	17.81	19.30	20.09	17.96	19.22
Number of bolls/plant	28.93	35.27	32.61	30.20	30.39
Boll weight (g)	3.89	4.26	4.07	3.83	4.01
Ginning out turn (%)	37.00	36.27	36.36	36.84	35.19
Lint Index (g)	5.02	4.74	5.06	4.67	4.88
Fibre length (mm)	27.03	29.52	28.11	28.18	28.62
Micronaire (µg/inch)	26.48	29.32	27.44	28.82	28.16
Fibre strength (g/tex)	4.02	3.98	3.84	4.02	3.74
Elongation percentage	5.49	5.63	5.58	5.60	5.56
Uniformity index (%)	80.73	83.17	81.50	82.60	81.60
Seed cotton yield (kg/ha)	1170.46	1527.23	1313.54	1181.67	1053.00

Table 7. Average intra and inter-cluster Eucledian^2 values among the five clusters in 40 genotypes of cotton

Clusters	I	II	III	IV	V
I	**5.29**	6.98	6.40	5.84	6.06
II		**4.87**	6.09	5.57	6.12
III			**5.2**	5.88	5.87
IV				**4.33**	5.49
V					**5.1**

suggest the possibility of simultaneous improvement of both productivity and fibre quality through appropriate breeding strategies. Multivariate analyses reinforced these findings, with PCA explaining a major proportion of total variability and identifying yield, fibre quality and earliness traits as principal contributors to genetic divergence. Cluster analysis demonstrated clear grouping of genotypes, with Cluster II excelling in yield and fibre quality traits and Cluster III characterized by earliness and higher lint index. The substantial inter-cluster distances highlight the potential of selecting parents from divergent clusters to generate superior recombinants. Overall, the study provides valuable insights for developing high-yielding, early-maturing and superior fibre quality cotton cultivars through strategic hybridization and selection.

REFERENCES

Ali, R.K., Malik, Y., Majidano, T., Abbasi, M.Z.H., Ejaz, S., Bahalkani, G.H., Aleem, N.M.M.U., Kaleri, Z.S.A., Chutto, B., Kaleri, F.A., Kaleri, K.K. and Kaleri, M.I.A. 2025. *Genetic diversity for yield and fiber traits in upland cotton (Gossypium hirsutum L.) genotypes*. *Pure Appl. Biol.,* **14**(2): 524-530. [Cross Ref]

Bhanuchand, K. and Raju, B. 2025. *Assessment of genetic divergence in cotton through principal component and cluster analysis*. *The Andhra Agric. J.,* **72(**2): 123–130.

Chapara, R., Reddy, K.S., Rani, M.S., Lakshmi, B.S., Roja, V. and Pranaya, J. 2022. Variability studies and genetic divergence in cotton (*Gossypium hirsutum* L.) germplasm using multivariate analysis. *Electron. J. Plant Breed.,* **13**(4): 1305-1311. [Cross Ref]

Chapara, R., Pavan, C., Sudha Rani, M., Diana Grace, G. A., James, M. and Priyadharshini. 2025. *Genetic variability and correlation studies for seed cotton yield and fibre quality parameters in diverse cotton genotypes*. *J. Res. Angrau,* **53**(2): 1–8. [Cross Ref]

Global Cotton Outlook 2026 (Global Cotton Outlook 2026).

Gowda, M. V. C., Ramesh, S., Kumar, A. and Shankar, A. G. 2026. *Multivariate analysis for genetic diversity and trait association in cotton (Gossypium spp.). J. Cotton Res.,* **9**(1): 1–10.

Johnson, H.W., Robinson, H.F. and Comstock, R.E. 1955. Genotypic and phenotypic correlations in soybeans and their implications in selection. *Agron. J.,* **47**: 477-483. [Cross Ref]

Kumar, D., Sangwan, O., Jattan, M., Kumar, S., Somveer, Koli, G. K. and Singh, V. 2024. *Harnessing the association between morphological, biochemical and fibre quality traits in desi cotton (Gossypium arboreum L.) genotypes. Indian J. Agric. Sci.,* **94**(2): 150–155. [Cross Ref]

Kumar, S., Reddy, K. R., Prakash, G. and Naik, V. 2025. *Genetic diversity analysis in upland cotton (Gossypium hirsutum L.) using multivariate techniques. Elect. J. Plant Breed.,* **16**(1): 45–52.

Naveed, S., Toyinbo, J., Ingole, H., Shekar, P.V., Jones, M., Campbell, B.T. and Rustgi, S. 2023. *Development of high-yielding upland cotton genotypes with reduced regrowth after defoliation using a combination of molecular and conventional approaches. Genes.,* **14**(11): 2081. [Cross Ref]

Pavan, C., Chapara, R., Sudha Rani, M., Diana Grace, G.A., James, M. and Priyadharshini. 2025. Unraveling genetic divergence using principal component analysis for yield and fibre quality traits in cotton genotypes. *Int. J. Adv. Biochem. Res.*,**9**(8): 122-127. [Cross Ref]

Prakash, V. and Suthar G. 2023. Cluster analysis in newly developed Bt cotton hybrids (G. hirsutum). *J. Progressive Agricult.,***14**(1): 23-26.

R Core Team. 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Rehman, M.M.U., Mohibullah, M., Mudasir, M., Khan, N.U., Khan, I.A., Rima Bibi, R. and Mustansir, R. 2025. Principal component analysis of yield and fibre-related traits in cotton genotypes. *Int. J. Agric.. Innov. Cutting-Edge Res.*, **3**(1): 26-35.

Sagar, U.N., Reddy B.V., Vishnuvardhan K.M., Prasanna R.A., Rao E.S.V., Sabitha N., Kalyani D.L., Reddy Y.R. and Suresh G. 2023. Genetic analysis of fibre quality and yield related traits in desi cotton (*Gossypium arboreum* L.). *Elect. J. Plant Breed.,* **14**(3): 833-840. [Cross Ref]

Shrestha, R., Acharya, S., Gyawali, S., Subedi, S. and Adhikari, B. 2025. *Genetic diversity and trait association analysis using multivariate approaches in crop plants: A review. Agronomy*, **15**(3): 512–528.

Sneath, P.H.A. and Sokal, R.R. 1973. Numerical taxonomy: The principles and practice of numerical classification. *W. H. Free-Man, San Francisco.,* 573p.

Ujjainkar, V.V. 2025. Principal Component Analysis (PCA) in cotton (*Gossypium hirsutum* L.) for estimation of diversity of yield and fibre quality. *South Asian J. Agric. Sci.,* **5**(2): 372-375. [Cross Ref]

USDA Cotton Outlook, 2026: https://www.usda.gov/oce/ag-outlook-forum/commodity-outlooks.

Ward, J.H. 1963. Hierarchical Grouping to Optimize an Objective Function. *J. Amer. Stat. Assoc.*, **58**: 236-244. [Cross Ref]