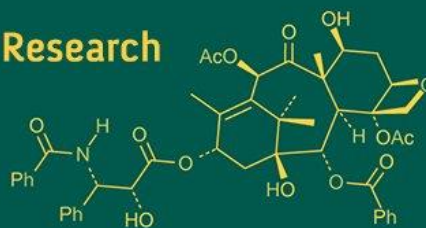
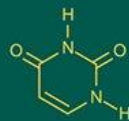


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
NAAS Rating (2026): 5.29
IJABR 2026; SP-10(7): 458-464
www.biochemjournal.com
Received: 07-05-2026
Accepted: 11-06-2026

Kunal P Chetule
PG Scholar, Genetics and
Plant breeding, College of
Agriculture, Nagpur,
Maharashtra, India

Dr. MP Meshram
Senior Rice Breeder,
Agriculture Research Station,
Sakoli, Bhandara,
Maharashtra, India

Dr. SR Kamdi
Mustard Breeder of AICRP on
Linseed and Mustard, College
of Agriculture, Nagpur,
Maharashtra, India

Dr. PV Shende
Professor, Agricultural Botany
Section, College of Agriculture,
Nagpur, Maharashtra, India

Dr. SA Patil
Professor, Agricultural Botany
Section, College of Agriculture,
Nagpur, Maharashtra, India

Arjun S Verma
Genetics and Plant breeding,
College of Agriculture, Nagpur,
Maharashtra, India

Bhushan R Pawar
Genetics and Plant breeding,
College of Agriculture, Nagpur,
Maharashtra, India

Corresponding Author:
Kunal P Chetule
PG Scholar, Genetics and
Plant breeding, College of
Agriculture, Nagpur,
Maharashtra, India

Multivariate analysis of rice genotypes under direct-seeded condition using principal component, principal factor and hierarchical cluster analyses

Kunal P Chetule, MP Meshram, SR Kamdi, PV Shende, SA Patil, Arjun S Verma and Bhushan R Pawar

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i7Sf.9019>

Abstract

Rice is one of the most important staple food crops, feeding more than half of the world's population and contributing substantially to global food security. However, conventional transplanted rice cultivation is increasingly constrained by labour shortages, escalating production costs and dwindling water resources. Consequently, direct-seeded rice (DSR) has emerged as a resource-conserving and economically viable alternative. The present study aimed to characterize the genetic diversity among 56 rice genotypes under direct-seeded conditions using multivariate statistical approaches. Principal component analysis revealed that four principal components with eigen values greater than unity accounted for 79.41% of the total variability, with PC1, PC2, PC3 and PC4 contributing 29.17%, 19.19%, 18.01% and 13.04% of the variation, respectively. Principal factor analysis identified four major factors explaining 65.60% of the total variation. Hierarchical cluster analysis grouped the rice entries into eight distinct clusters, indicating substantial genetic diversity among the experimental material. Cluster III recorded the highest grain yield per plant, whereas Cluster VII exhibited the highest number of panicles per plant. The genotypes SKL-19-23-40-05-26-08, SKL-19-15-134-138-95-60, SKL-19-18-45-60-44-78, SKL-19-15-40-32-12-15 and SKL-19-12-35-48-48-50 were identified as promising donors for direct-seeded rice improvement. These genotypes may serve as potential parents for developing high-yielding and widely adapted rice cultivars for direct-seeded ecosystems.

Keywords: Direct-seeded rice, genetic diversity, principal component analysis, principal factor analysis, hierarchical cluster analysis

Introduction

Rice (*Oryza sativa* L.) is a major cereal crop and the principal source of calories for a large proportion of the world's population. The crop plays a pivotal role in ensuring food and nutritional security and contributes significantly to the agricultural economy of many Asian countries, including India (FAOSTAT, 2024). Despite remarkable gains in productivity, sustaining rice production under changing climatic and socio-economic conditions remains a major challenge.

The conventional puddled transplanted method of rice cultivation has been widely adopted because of its stable performance and effective weed suppression. Nevertheless, this system is highly dependent on irrigation water and intensive labour inputs. Rapid urbanization, increasing labour costs and declining groundwater resources have necessitated the development of alternative crop establishment methods that are both productive and resource efficient. Direct-seeded rice (DSR) has emerged as a promising alternative because it minimizes labour requirements, reduces water consumption and lowers the cost of cultivation while ensuring timely crop establishment (Kumar and Ladha, 2011; Rao *et al.*, 2007). Consequently, DSR is increasingly being considered an important strategy for sustainable rice production [15, 30, 44].

The successful adoption of DSR largely depends on the availability of rice cultivars possessing desirable traits such as early vigour, adaptability, high yield potential and efficient resource utilization. The existence of genetic variability among rice genotypes provides opportunities for identifying superior lines and developing improved cultivars adapted to direct-seeded

ecosystems. Therefore, a comprehensive assessment of genetic diversity is essential for the effective utilization of germplasm resources in breeding programmes (Anandan *et al.*, 2011)^[5].

Multivariate statistical techniques offer powerful approaches for analysing complex datasets and elucidating the genetic relationships among traits and genotypes. Principal component analysis (PCA) summarizes the total variability into a smaller number of independent components and identifies the traits responsible for genetic divergence (Jolliffe, 2002). Principal factor analysis (PFA) helps in identifying the underlying factors governing trait associations and facilitates the interpretation of multivariate datasets (Johnson and Wichern, 2007). Hierarchical cluster analysis (HCA) classifies genotypes into groups based on their degree of similarity and assists in the identification of genetically diverse parents for hybridization programmes (Mohammadi and Prasanna, 2003)^[24, 23, 36].

Although multivariate analyses have been extensively employed in rice improvement programmes, information on the characterization of rice genotypes under direct-seeded conditions remains inadequate. Therefore, the present study was undertaken to characterize rice genotypes under direct-seeded conditions using principal component analysis, principal factor analysis and hierarchical cluster analysis and to identify promising genotypes and potential parental lines

for the development of high-yielding rice cultivars adapted to direct-seeded ecosystems.

Materials and Methods

The present investigation was conducted during Kharif 2025 at the Research Farm of the All India Coordinated Research Project (AICRP) on Linseed and Mustard, College of Agriculture, Nagpur, Maharashtra, India under direct-seeded rice (DSR) conditions. The experimental material comprised 51 rice (*Oryza sativa* L.) genotypes derived from twelve crosses developed at the Agricultural Research Station, Sakoli, Bhandara, Maharashtra, along with five check varieties, viz., MTU-1010, Vandana, Rajeshwari, PDKV-Sadhana and Samleshwari. The experiment was laid out in a Randomized Block Design (RBD) with two replications. The genotypes were grown in plots measuring 3.0 m × 1.5 m with a spacing of 20 cm × 15 cm following the recommended package of practices for rice cultivation. Observations were recorded on nine quantitative traits, namely days to 50 per cent flowering, days to maturity, plant height (cm), panicle length (cm), number of panicles plant⁻¹, number of spikelets panicle⁻¹, spikelet fertility (%), 1000-grain weight (g) and grain yield plant⁻¹ (g). Data on days to 50 per cent flowering and days to maturity were recorded on a plot basis, whereas observations on the remaining traits were recorded from ten randomly selected plants of each genotype in each replication and averaged for statistical analysis.

Table 1: List of rice genotypes and checks evaluated under direct-seeded conditions

Sr. No.	Genotypes	Sr. No.	Genotypes	Sr. No.	Genotypes
1	SKL-19-16-22-12-18-06	20	SKL-19-14-45-74-134-70	39	SKL-19-21-108-105-80-54
2	SKL-19-16-49-46-88-50	21	SKL-19-06-35-101-84-108	40	SKL-19-21-108-105-80-60
3	SKL-19-16-22-18-34-18	22	SKL-19-04-64-17-28-24	41	SKL-19-21-108-127-88-68
4	SKL-19-16-120-100-102-97	23	SKL-19-04-103-225-108-90	42	SKL-19-21-108-186-108-75
5	SKL-19-19-118-51-30-43	24	SKL-19-02-68-112-69-50	43	SKL-19-21-119-206-127-80
6	SKL-19-19-118-51-30-50	25	SKL-19-02-77-221-98-78	44	SKL-19-21-119-206-127-88
7	SKL-19-19-124-154-30-58	26	SKL-19-02-80-238-72-99	45	SKL-19-23-40-05-26-08
8	SKL-19-18-30-15-18-12	27	SKL-19-12-35-48-48-50	46	SKL-19-23-40-05-26-15
9	SKL-19-18-30-42-34-48	28	SKL-19-12-35-48-60-64	47	SKL-19-23-142-40-34-30
10	SKL-19-18-45-60-44-60	29	SKL-19-12-40-112-126-110	48	SKL-19-23-142-40-34-37
11	SKL-19-18-45-60-44-78	30	SKL-19-11-65-01-05-14	49	SKL-19-23-100-82-50-80
12	SKL-19-15-40-32-12-15	31	SKL-19-11-65-01-18-33	50	SKL-19-23-100-82-50-84
13	SKL-19-15-40-32-12-28	32	SKL-19-11-65-42-35-49	51	SKL-19-23-100-98-74-90
14	SKL-19-15-40-32-34-35	33	SKL-19-11-180-56-50-68	52	Vandana
15	SKL-19-15-107-100-58-48	34	SKL-19-11-180-111-67-84	53	Sambleshwari
16	SKL-19-15-134-138-95-60	35	SKL-19-11-194-116-89-99	54	MTU-1010
17	SKL-19-14-30-15-18-12	36	SKL-19-21-21-08-35-28	55	PDKV Sadhana
18	SKL-19-14-30-42-90-48	37	SKL-19-21-21-08-35-38	56	Rajeshwari
19	SKL-19-14-45-60-120-60	38	SKL-19-21-32-42-45-48		

This format is suitable for inclusion in a research manuscript (MS Word, thesis, or journal article).

Analysis of variance was performed according to the procedure described by Panse and Sukhatme (1985). Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) were carried out using R software version 4.4.1 (R Core Team, Vienna, Austria) with the packages FactoMineR, factoextra, cluster and dendextend, whereas Principal Factor Analysis (PFA) with varimax rotation was performed using XLSTAT 2024.1 (Addinsoft, Paris, France). Principal components with eigenvalues greater than one were retained according to the Kaiser criterion. Genetic divergence among the genotypes was estimated using Euclidean distances and clustering was performed by Ward's minimum variance method. The

multivariate analyses were used to identify the major sources of variation and to classify the rice genotypes into distinct groups for the identification of diverse and superior genotypes for rice improvement.

Result and Discussion

Principal component analysis (PCA) is an effective multivariate statistical tool that reduces the dimensionality of complex datasets and identifies the major sources of variability among genotypes by transforming correlated variables into a smaller number of uncorrelated principal components (Jolliffe and Cadima, 2016)^[25].

Table 2: Eigenvalues, Percent Variability and Cumulative Variability of Principal Components

Principal Component	Eigen Value	Percent Variability (%)	Cumulative Variability (%)
PC1	2.625	29.167	29.167
PC2	1.727	19.189	48.356
PC3	1.620	18.005	66.361
PC4	1.174	13.044	79.405
PC5	0.717	7.964	87.369

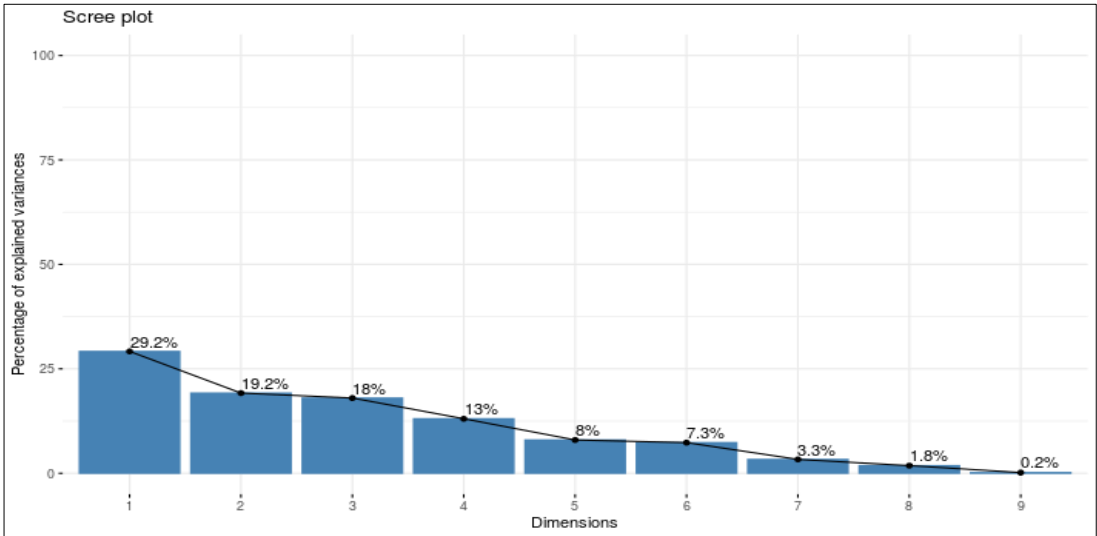


Fig 1: Screen plot showing Principal Components and percentage of variation

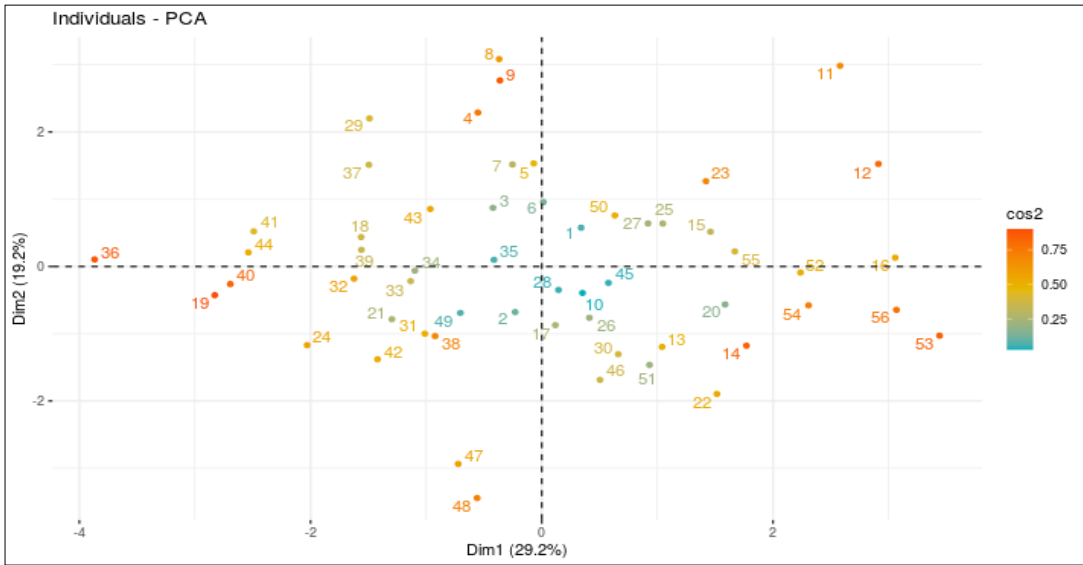


Fig 2: Scatter plot of 56 rice genotypes based on first two principal components

In crop improvement programmes, PCA has been extensively used to characterize genetic diversity and to identify divergent parents for hybridization (Mohammadi and Prasanna, 2003; Rudresh *et al.*, 2021) [36, 45]. The eigenvalues, percentage variability and cumulative variability explained by different principal components are presented in Table 2. The total variability among the 56 rice entries was partitioned into nine principal components, of which the first four components exhibited eigenvalues greater than one and were therefore retained according to the Kaiser criterion. These four principal components collectively explained 79.41% of the total variation, indicating that a substantial proportion of the variability among the genotypes could be explained by a limited number of components. Among the extracted components, PC1 accounted for the

highest proportion of variability (29.17%), followed by PC2 (19.19%), PC3 (18.01%) and PC4 (13.04%). The cumulative contribution of these four components reached 79.41%, suggesting the existence of considerable phenotypic diversity among the rice genotypes under direct-seeded conditions. Similar findings have been reported by Rudresh *et al.* (2021), Kumari *et al.* (2021) and Choudhary *et al.* (2022), who observed that the first few principal components explained most of the variability in rice germplasm [45, 31, 10]. The scree plot (Fig. 1) exhibited a distinct elbow after the fourth principal component, indicating that the first four components captured most of the variation, whereas the remaining components contributed only a small proportion of the total variability. Therefore, retention of the first four

principal components was considered adequate for explaining the genetic structure of the experimental material.

The scatter plot based on the first two principal components (Fig. 2) revealed a wide distribution of genotypes in different quadrants, indicating substantial genetic diversity among the rice entries. Genotypes located close to each other were relatively similar, whereas those positioned further apart on the first two principal component axes were likely to be more divergent. Such genetically divergent genotypes can be exploited as potential parents in hybridization programmes for the development of superior recombinants and transgressive segregants. Similar observations were reported by Hartmann *et al.* (2018), Walle *et al.* (2019) and Rudresh *et al.* (2021) [21, 45].

Principal factor analysis (PFA)

Principal factor analysis (PFA) was performed to identify the underlying latent factors governing the relationships among the studied traits and to determine the major contributors to phenotypic variability. Factor analysis has been widely used in plant breeding to simplify complex datasets and to identify groups of correlated traits that can be simultaneously improved through selection (Johnson and Wichern, 2007). The factor loadings and variance explained by different factors are presented in Table 3. Four factors with eigenvalues greater than one were extracted and together accounted for 65.60% of the total variability among the rice genotypes under direct-seeded conditions. Factor 1 possessed an eigenvalue of 2.104 and explained 23.38% of the total variation. This factor exhibited very high positive loadings for days to 50% flowering (0.988) and days to maturity (0.986), indicating that crop duration traits constituted a major source of phenotypic variability among the rice genotypes. Factor 2 accounted for 15.10% of the total

variation with an eigenvalue of 1.359 and was predominantly represented by grain yield per plant (0.870) and number of spikelets per panicle (0.676), suggesting the importance of these yield-related traits in explaining productivity differences under direct-seeded conditions. Factor 3 explained 14.85% of the total variability with an eigenvalue of 1.337 and was characterized by high positive loadings for plant height (0.819) and panicle length (0.805), indicating the importance of plant architecture and panicle morphology in genotype differentiation. Factor 4 exhibited an eigenvalue of 1.104 and contributed 12.27% of the total variability and was primarily associated with test weight (0.950), demonstrating the independent contribution of grain weight to the overall phenotypic variability. The cumulative contribution of these four factors reached 65.60%, indicating that a major portion of the variability was governed by a limited number of latent factors. Furthermore, the non-significant chi-square value ($\chi^2 = 5.656$; $df = 6$; $P = 0.4628$) confirmed the adequacy of the four-factor model and indicated that the extracted factors sufficiently explained the correlation structure among the studied traits.

The present investigation revealed that crop duration traits (days to 50% flowering and days to maturity), yield attributes (grain yield per plant and number of spikelets per panicle), plant architectural traits (plant height and panicle length) and test weight were the principal sources of phenotypic variability among the rice genotypes under direct-seeded conditions. Therefore, these traits may be given greater emphasis in selection and hybridization programmes aimed at developing high-yielding and widely adapted rice cultivars for direct-seeded ecosystems. Similar findings were reported by Behera *et al.* (2022), Gedifew *et al.* (2022) and Srinivas *et al.* (2025) [8, 18].

Table 3: Factor loadings of characters with respect to different principal factors (varimax rotation)

Trait	Factor 1	Factor 2	Factor 3	Factor 4
Days to 50% Flowering	0.988	-0.001	-0.023	0.010
Days to Maturity	0.986	-0.013	-0.021	-0.023
Spikelet Fertility (%)	-0.318	-0.031	-0.115	-0.066
Grain Yield per Plant (g)	-0.044	0.870	0.011	0.342
Number of Seeds per Panicle	0.067	0.676	-0.039	-0.055
Number of Panicles per Plant	-0.042	0.335	0.040	-0.271
Plant Height (cm)	0.128	-0.016	0.819	-0.039
Panicle Length (cm)	-0.173	0.008	0.805	0.038
Test Seed Weight (g)	-0.005	0.178	-0.001	0.950

Parameter	Factor 1	Factor 2	Factor 3	Factor 4
SS Loadings (Eigenvalue)	2.104	1.359	1.337	1.104
Proportion of Variance	0.234	0.151	0.149	0.123
Variance Explained (%)	23.38	15.10	14.85	12.27
Cumulative Variance (%)	23.38	38.48	53.33	65.60

Hypothesis Tested	Chi-square (χ^2)	df	P-value
Ho: Four factors are sufficient	5.656	6	0.4628

Hierarchical cluster analysis

Hierarchical cluster analysis based on Euclidean distances and Ward's minimum variance method grouped the 56 rice entries into eight distinct clusters, indicating the existence of substantial genetic diversity among the experimental material under direct-seeded conditions. The hierarchical relationships among the genotypes and the pattern of genetic divergence are illustrated in the dendrogram (Fig. 3), while the distribution of genotypes among different clusters is

presented in Table 4. Cluster I was the largest, comprising 13 genotypes, followed by Cluster II with 11 genotypes and Cluster III with 10 genotypes, whereas Clusters IV, VII and VIII each contained four genotypes and Clusters V and VI comprised five genotypes each (Table 4). The occurrence of genotypes in different clusters irrespective of their pedigree suggests the presence of considerable genetic divergence among the entries. Similar observations have been reported by Arunachalam (1981), Mohammadi and Prasanna (2003),

Amegan *et al.* (2020) and Choudhary *et al.* (2022)^[7, 36, 4, 10], who emphasized that clustering based on multivariate traits is

useful for identifying genetically diverse parents for crop improvement programmes.

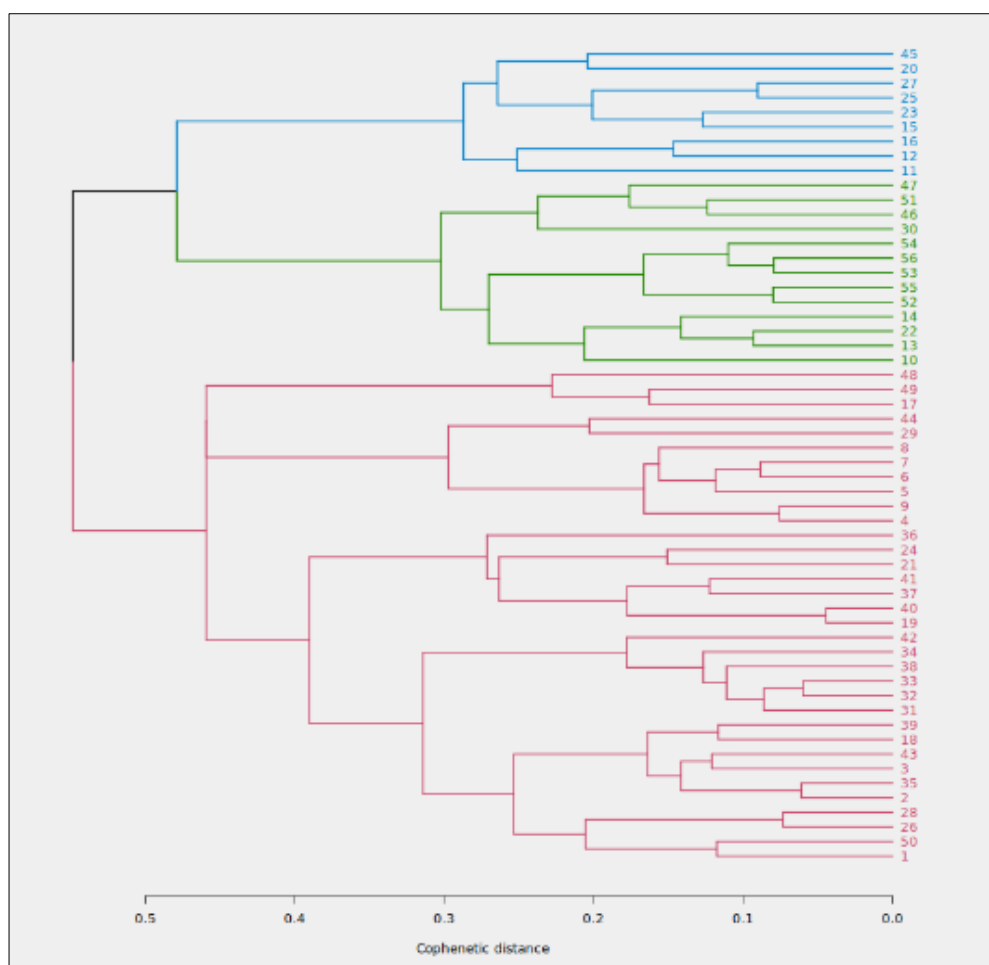


Fig 3: Hierarchical cluster dendrogram showing genetic divergence among rice genotypes based on multivariate data

Table 4: Distribution of rice genotypes into different clusters based on hierarchical cluster analysis (HCA)

Cluster	No. of Rice genotypes	Cluster Members
I	13	24, 36, 21, 37, 40, 19, 42, 34, 38, 33, 32, 31, 41.
II	11	39, 18, 43, 3, 35, 2, 28, 26, 50, 1, 9
III	10	45, 20, 27, 25, 23, 15, 16, 12, 11, 4
IV	4	8, 7, 6, 5
V	5	54, 56, 53, 55, 52
VI	5	48, 49, 17, 44, 29
VII	4	47, 51, 46, 30
VIII	4	14, 22, 13, 10

Table 5: Cluster-wise Mean Performance of Rice Genotypes for Nine Quantitative Traits

Cluster	Days to 50% flowering	Days to Maturity	Plant height (cm)	No. of panicles plant ⁻¹	Panicle length (cm)	No. of spikelet Panicle ⁻¹	Spikelet fertility (%)	1000 Grain weight	Grain yield plant ⁻¹
I	102.84	133.69	117.58	7.86	21.31	156.11	83.950	18.50	23.14
II	98.05	130.10	115.10	7.11	22.08	167.78	84.58	21.39	27.93
III	84.27	116.27	114.37	7.90	21.85	174.21	84.41	23.85	37.98
IV	92.00	124.08	101.29	6.85	19.49	164.52	84.60	21.87	26.53
V	71.80	103.70	110.38	7.44	21.91	153.05	86.95	25.12	28.32
VI	92.20	122.60	109.51	70.9	21.25	130.22	83.01	19.42	21.22
VII	79.00	112.00	114.76	8.10	21.96	141.56	86.84	16.8	19.69
VIII	79.25	112.00	118.68	6.62	22.07	137.30	84.25	23.75	28.03
S.D.	10.598	10.118	5.584	0.538	0.870	15.617	1.376	2.904	5.668
Variance	112.327	102.37	31.183	0.290	0.757	243.896	1.894	8.435	32.128

Table 6: Average Intra- and Inter-Cluster Distances among Rice Genotypes

Cluster	I	II	III	IV	V	VI	VII	VIII
I	(0.180)	0.205	0.318	0.276	0.350	0.271	0.281	0.283
II		(0.156)	0.235	0.214	0.270	0.282	0.287	0.236
III			(0.204)	0.274	0.245	0.348	0.310	0.241
IV				(0.126)	0.280	0.270	0.330	0.276
V					(0.114)	0.330	0.230	0.191
VI						(0.279)	0.284	0.265
VII							(0.175)	0.241
VIII								(0.144)

The cluster mean performance of the eight clusters for nine quantitative traits is presented in Table 5. Considerable variation was observed among the clusters for all the characters studied. Cluster V recorded the earliest flowering (71.80 days) and maturity (103.70 days), indicating its suitability as a source of earliness. Cluster VIII exhibited the highest plant height (118.68 cm), while Cluster VII recorded the maximum number of panicles per plant (8.10). The highest number of spikelets per panicle (174.21), highest grain yield per plant (37.98 g) and comparatively high-test weight (23.85 g) were observed in Cluster III, indicating the superiority of genotypes belonging to this cluster for yield improvement under direct-seeded conditions. Cluster V exhibited the highest spikelet fertility (86.95%) and maximum 1000-grain weight (25.12 g), suggesting its potential use in improving grain quality and fertility traits. The superior performance of these clusters further corroborated the results of principal factor analysis, wherein crop duration traits, yield attributes, plant architectural traits and test weight were identified as the major sources of phenotypic variability among the rice genotypes.

The average intra- and inter-cluster distances among the rice genotypes are presented in Table 6 and further confirmed the existence of substantial genetic divergence among the experimental material. The highest intra-cluster distance was observed in Cluster VI (0.279), indicating greater heterogeneity among genotypes within this cluster, whereas the lowest intra-cluster distance was recorded in Cluster V (0.114), suggesting a high degree of similarity among its constituent genotypes (Table 6). Among the inter-cluster distances, the maximum divergence was observed between Clusters I and V (0.350), followed by Clusters III and VI (0.348) and Clusters IV and VII (0.330), whereas the minimum inter-cluster distance was recorded between Clusters V and VIII (0.191), indicating a close genetic relationship between these clusters (Table 6). The high inter-cluster distances observed among the divergent cluster combinations suggest that hybridization involving genotypes from these clusters may generate greater heterosis and wider variability in segregating generations, as also reported by Arunachalam (1981), Mohammadi and Prasanna (2003) and Rudresh *et al.* (2021) [7, 36, 45].

Considering both inter-cluster distances (Table 6) and cluster mean performance (Table 5), genotypes SKL-19-23-40-05-26-08, SKL-19-12-35-48-48-50 SKL-19-15-134-138-95-60, SKL-19-18-45-60-44-78 and SKL-19-15-40-32-12-15 exhibited desirable performance for important yield-contributing traits and may serve as valuable parental lines in breeding programmes. Therefore, crosses involving genetically divergent and agronomically superior genotypes from clusters exhibiting yielding and widely adapted rice cultivars under direct-seeded ecosystems.

Acknowledgement: I sincerely express my heartfelt gratitude to everyone who supported and encouraged me throughout the successful completion of this research work. I am especially thankful to my Research Guide for providing valuable guidance, constant encouragement and constructive suggestions during the course of this study. I also extend my sincere thanks to the All India Coordinated Research Project (AICRP) on Linseed and Mustard for providing the necessary facilities and research support.

References

1. Abdel-salam KMH, Shaalan AM, Abou-El-Soud GM, El-Dail MAE, Marei DAAE, El-Moneim AAAE, *et al.* Comprehensive quality profiling and multivariate analysis of rice (*Oryza sativa* L.) cultivars: integrating physical, cooking, nutritional, and micronutrient characteristics for enhanced varietal selection. BMC Plant Biol. 2025;25:492.
2. Ali Z, Naeem M, Rafiq I, Nusrat AT, Khaerani PI, Bahrin AH, *et al.* Multivariate analysis in detecting the selection criteria for rice breeding under a direct seeding system. Sarhad J Agric. 2025;41(5):71-83.
3. Al-Khayri JM, El-Malky MM. Genetic parameters and principal component analysis for quantitative traits in rice (*Oryza sativa* L.). SABRAO J Breed Genet. 2023;55(5):1547-1560.
4. Amegan E, Efisue A, Akoroda M, Shittu A, Tonegnikes F. Genetic diversity of Korean rice (*Oryza sativa* L.) germplasm for yield and yield-related traits for adoption in rice farming systems in Nigeria. Genomics. 2020;8(1):19-28.
5. Anandan A, Eswaran R, Sabesan T, Prakash M. Diversity in rice genotypes under salt affected soil based on multivariate analysis. Pertanika J Trop Agric Sci. 2011;34:33-40.
6. Anandan A, Rajiv G, Eswaran R, Prakash M. Genotypic variation and relationships between quality traits and trace elements in traditional and improved rice (*Oryza sativa* L.) genotypes. J Food Sci. 2011;76(4):H122-H130.
7. Arunachalam V. Genetic distance in plant breeding. Indian J Genet Plant Breed. 1981;41(2):226-236.
8. Behera PP, Das TR, Mishra CHP, Bastia DN, Sahu RK. Genetic diversity and multivariate analysis in high-zinc rice (*Oryza sativa* L.) genotypes across different environments. Electron J Plant Breed. 2022;13(4):1070-1080.
9. Bhardwaj DN. Statistical techniques in agricultural research. New Delhi: New India Publishing Agency; 2012.
10. Choudhary M, Kumar B, Singh P, Sharma M. Diversity analysis for yield and its contributing traits in rice germplasm (*Oryza sativa* L.) using principal component analysis approach. Int J Environ Clim Change. 2022;12(9):143-150.

11. Choudhary M, Sharma SR, Sharma A, Singh AK. Principal component analysis and genetic divergence studies in rice (*Oryza sativa* L.) genotypes under direct-seeded conditions. *Electron J Plant Breed*. 2022;13(2):456-464.
12. Chuchert S, Nualsri C, Junsawang N, Soonsuwon W. Genetic diversity, genetic variability, and path analysis for yield and its components in indigenous upland rice (*Oryza sativa* L. var. *glutinosa*). *Songklanakarin J Sci Technol*. 2018;40(3):609-616.
13. Dash M. Statistical procedures in agricultural research. New Delhi: Kalyani Publishers; 2016.
14. Everitt BS, Landau S, Leese M, Stahl D. Cluster analysis. 5th ed. Chichester: Wiley; 2011.
15. FAOSTAT. FAOSTAT Statistical Database. Rome: Food and Agriculture Organization of the United Nations; 2024.
16. Farooq M, Siddique KHM, Rehman H, Aziz T, Lee DJ, Wahid A. Rice direct seeding: experience, challenges and opportunities. *Soil Tillage Res*. 2011;111:87-98.
17. Fisher RA, Yates F. Statistical tables for biological, agricultural and medical research. 6th ed. Edinburgh: Oliver and Boyd; 1967.
18. Gedifew G, Abebe T, Bitew M. Multivariate analysis of lowland rice (*Oryza sativa* L.) genotypes under direct-seeded conditions. *Heliyon*. 2022;8:e11875.
19. Gill MS, Kumar A, Kumar P. Growth and yield of direct-seeded rice under different tillage and crop establishment methods. *Indian J Agron*. 2006;51(4):286-289.
20. Harman HH. Modern factor analysis. 3rd ed. Chicago: University of Chicago Press; 1976.
21. Hartmann K, Krois J, Waske B. E-Learning Project SOGA: Statistics and Geospatial Data Analysis. Berlin: Department of Earth Sciences, Freie Universität Berlin; 2018.
22. Hotelling H. Analysis of a complex of statistical variables into principal components. *J Educ Psychol*. 1933;24(6):417-441.
23. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th ed. Upper Saddle River (NJ): Pearson Education; 2007.
24. Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer-Verlag; 2002.
25. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans R Soc A Math Phys Eng Sci*. 2016;374(2065):20150202.
26. Kaiser HF. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*. 1958;23(3):187-200.
27. Kesh H, Kharb R, Ram K, Munjal R, Kaushik P, Kumar D. Adaptability and AMMI biplot analysis for yield and agronomical traits in scented rice genotypes under diverse production environments. *Indian J Tradit Knowl*. 2021;20(2):550-562.
28. Khatun M, Islam MR, Khan MAR, Hossain MK, Islam AKMA. Principal component analysis and estimated breeding values for selecting suitable parental genotypes in rice (*Oryza sativa* L.). *SABRAO J Breed Genet*. 2023;55(6):1886-1896.
29. Kumar MS, Kumar B, Sheera A. Characterization and principal component analysis of rice germplasm (*Oryza sativa* L.) under sub-tropical ecology of Jammu region of India. *Plant Arch*. 2024;24(2):1782-1791.
30. Kumar V, Ladha JK. Direct seeded rice: recent development and future research needs. *Adv Agron*. 2011;111:297-413.
31. Kumari BK, Kumar BR, Jyothula D, Rao NM. Diversity analysis in rice breeding lines for yield and its components using principal component analysis. *J Pharmacogn Phytochem*. 2021;10(1):905-909.
32. Kumari S, Chakrabarty SK, Paul D. Assessment of genetic diversity for seed vigor traits in rice (*Oryza sativa* L.) through principal component and cluster analysis. *Indian J Genet Plant Breed*. 2023;83(3):343-354.
33. Lal GM. Genetic divergence and cluster analysis in crop plants. New Delhi: New India Publishing Agency; 2016.
34. Lewis GJ, Lisle AT. Towards better canola yield: a principal components analysis approach. In: *Proceedings of the 9th Australian Agronomy Conference*; 1998; Wagga Wagga, New South Wales, Australia.
35. Ministry of Agriculture and Farmers Welfare. Agricultural Statistics at a Glance 2024. New Delhi: Government of India; 2024.
36. Mohammadi SA, Prasanna BM. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Sci*. 2003;43(4):1235-1248.
37. Oyelola BA. The Nigerian Statistical Association preconference workshop. Ibadan: University of Ibadan; 2004. p. 20-21.
38. Panse VG, Sukhatme PV. Statistical methods for agricultural workers. 2nd ed. New Delhi: Indian Council of Agricultural Research; 1961.
39. Parsad R. NARS statistical package (online statistical analysis). New Delhi: Indian Agricultural Statistics Research Institute; 2009.
40. PepsiCo International. Direct seeded rice cultivation technology and farmer adoption report. New Delhi: PepsiCo International; 2011.
41. Prakash S, Reddy SS, Chaudhary S, Vimal SC, Kumar A. Multivariate analysis in rice (*Oryza sativa* L.) germplasms for yield attributing traits. *Plant Sci Today*. 2023.
42. Prasanth HP, Jasmin AS, Selvaraj R, Kumar DP, Devasena N, Sharmili K, *et al*. Principal component and cluster analysis on eating and cooking quality parameters in rice (*Oryza sativa* L.) germplasm. *Biol Forum*. 2023;15(5):382-388.
43. Ramesh Naik M, Chetariya CP, Akkenapally JS. Characterization of the genetic variability in rice (*Oryza sativa* L.) and its assessment using principal component analysis (PCA). *Int J Environ Clim Change*. 2022;12(10):1257-1269.
44. Rao AN, Johnson DE, Siva Prasad B, Ladha JK, Mortimer AM. Weed management in direct-seeded rice. *Adv Agron*. 2007;93:153-255.
45. Rudresh NS, Jayamani P, Vijayakumar E, Manonmani S, Gangashetti M, Jeyakumar P, *et al*. Genetic analysis of rice germplasm suitable for direct and transplanted establishments. *Electron J Plant Breed*. 2021;12(4):1148-1156.
46. Salem KFM, Saleh MM, Aldahak L, Elabd AB. Assessment of phenotypic diversity of rice (*Oryza sativa* L.) genotypes by multivariate analysis. *J Arid-land Agric*. 2021;7:52-59.
47. Sneath PHA, Sokal RR. Numerical taxonomy: the principles and practice of numerical classification. San Francisco: W. H. Freeman and Company; 1973.