



Comprehensive genetic diversity analysis of rice (*Oryza sativa* L.) genotypes for yield and cooking quality traits using multivariate approaches

M Pavan Satya Sai¹, Vishal Kumar Gupta^{1*}, Surekha², Ashish Kumar Banjare³, Prakriti Meshram³, Rita Chopkar³, Geetanjali Karchal³ & Anjan Roy¹

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

²Department of Plant Molecular Biology and Biotechnology, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Krishak Nagar, Raipur 492 012, Chhattisgarh, India

³Department of Genetics and Plant Breeding, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Krishak Nagar, Raipur 492 012, Chhattisgarh, India

*Correspondence: Vishal Kumar Gupta, vishalgupta030794@gmail.com

Abstract

Rice (*Oryza sativa* L.) holds global significance owing to its substantial genetic diversity and widespread consumer acceptance. Nevertheless, the crop encounters persistent threats from both biotic and abiotic stresses, necessitating the systematic utilisation of its available genetic diversity through selection-based breeding strategies. The present investigation was undertaken at the Postgraduate Research Farm of Centurion University of Technology and Management, Odisha, during the kharif season of 2025. Forty rice genotypes were evaluated in a randomised complete block design with 3 replications. Mahalanobis D² analysis grouped all forty genotypes into six clusters using Tocher's method. The maximum inter-cluster distance between clusters III and IV identifying crosses between these groups as promising namely RNR 15435 × Pusa Basmati 1850 and Kalma × MTU 1081 for recovering potentially superior transgressive segregants. Seed protein content was found to be the highest contributor to total genetic divergence. Principal component analysis yielded eight significant components, explaining 79.16 % of cumulative variability. The PC1 and PC2 together accounted for 38.50 %, driven primarily by grain weight, breadth and kernel length traits. Biplot analysis identified Mawaii, Bhuraj, RGL 1880 and Kalorat as superior for grain shape and cooking quality and Judi dhan, RNR 15435 and Manipur local as promising for yield components indicating their potential for exploiting maximum genetic distance and developing improved rice varieties suited to Odisha.

Keywords: biplot; cluster; principal component analysis; rice; Tocher method

Received: 25 June 2026. **Accepted:** 12 September 2026. **Published:** Version 1.0: 29 September 2026; Version 2.0: 01 October 2026.

Cite this article: Pavan SSM, Vishal KG, Surekha, Ashish KB, Prakriti M, Rita C, Geetanjali K, Anjan R. Comprehensive genetic diversity analysis of rice (*Oryza sativa* L.) genotypes for yield and cooking quality traits using multivariate approaches. Plant Science Today. 2026; 13(4): 1-11. <https://doi.org/10.14719/pst.16348>

1. Introduction

Rice (*Oryza sativa* L.), a monocotyledonous crop belonging to the genus *Oryza* ($2n = 24$), is taxonomically placed under the family Poaceae and subfamily Oryzoideae. Within the genus *Oryza*, there are only two species have been successfully domesticated for agricultural purposes. These include *Oryza sativa* L., found mostly in Asia and *Oryza glaberrima* Steud., which is endemic to Africa. Out of these, *O. sativa* has gained the distinction of being the most widely cultivated crop in the world (1). Cultivated rice can be broadly classified into two main subspecies, namely *indica* and *japonica*. Of these, the *indica* subspecies, which thrives under tropical and subtropical climatic conditions, is distinguished by the characteristic feature of long and slender shaped grains (1). In India, the area under rice cultivation occupies about

47.82 million hectares, resulting in an estimated annual production of almost 137.83 million tonnes with an average yield of 2882 kg ha⁻¹ (2). Rice is one of the principal crops grown in states like Uttar Pradesh, Madhya Pradesh, Punjab and West Bengal, which together contribute to about 90 % of total Indian rice production. Considerable effort has been directed towards increasing rice production while, at the same time ensuring grain quality through proper breeding strategies (3).

Knowledge regarding genetic variation among different types of rice plants is essential before planning breeding strategies. Genetic diversity, which refers to the variation in genetic constitution among individuals within a species, serves as a critical foundation for the simultaneous improvement of multiple desirable traits, including grain yield, resistance to biotic stresses such

as diseases and pests, tolerance to abiotic stresses like drought and enhancement of nutritional quality (4). Genetic divergence analysis of important traits such as grain yield per plant, elongated kernel ratio, amylose content and protein content allows plant breeders to identify excellent parental pairs that will produce superior-performing progenies. Rice quality-related characters such as kernel length, kernel width, length after cooking, width after cooking and elongated kernel ratio are instrumental in determining market value and consumer preferences. These are directly associated with the cooking characteristics, texture and colour of cooked rice grains, factors that become very important in locations where rice is consumed as the main food source (5). Effective genetic divergence analysis of these quality-linked traits, together with genetic parameter analysis, will help plant breeders to choose genotypes with good grain quality properties. In addition to grain quality, some other quantitative traits such as stem thickness, biological yield, harvest index and productive tiller number, grain size and test weight become equally important in assessing the agronomic performance of rice plants.

In this regard, the use of multivariate techniques such as principal component analysis (PCA), genotype by trait biplot and cluster analysis has emerged as an essential approach for the selection of genetically divergent genotypes and to facilitating breeding programmes involving yield and quality improvement in the future (6). Although various previous studies have used morphological characters for genetic divergence analysis in rice, greater emphasis has been placed on yield traits and little attention has been paid to cooking quality or biochemistry of rice grains. There is a limited number of studies that have simultaneously evaluated the diversity of rice genotypes concerning yield, grain quality and nutritive value, especially for the purpose of selecting parental varieties for developing

high-yielding varieties of rice with improved cooking quality. Furthermore, the use of these three statistical techniques for evaluating trait relationships amongst rice genotypes is poorly studied. In light of these research gaps, the present investigation was designed to assess the magnitude of genetic divergence among a set of rice genotypes for yield-attributing, cooking-quality and biochemical traits. The findings are expected to facilitate the identification of promising parental combinations for the concurrent improvement of productivity and grain quality in rice breeding programmes.

2. Materials and Methods

Experimental site

The present study was carried out at the Postgraduate Research Farm of M S Swaminathan School of Agriculture, Centurion University of Technology and Management (CUTM), Paralakhemundi, Odisha, during the kharif season of 2025–26. The station is situated 91 m above mean sea level at 18.799469° N latitude and 84.142037° E longitude. The experiment was conducted in an area that falls under Odisha's Northeastern Ghat zone. Nonetheless, the stability and adaptability of the reported traits and promising genotypes can be evaluated by assessing the genotypes in other locations and seasons.

Experimental materials

Forty rice genotypes were obtained from Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, Bidhan Chandra Krishi Viswavidyalaya, Nadia, West Bengal and Chinsura Rice Research Centre. Table 1 presents a comprehensive list of the genotypes. The study was carried out during the kharif season of 2025–26. The soil at the experimental site was characterised by sandy loam texture, optimal drainage and a pH of 6.3.

Table 1. List of rice genotypes

Sl. no.	Genotypes	Source	Sl. no.	Genotypes	Source
1	Bamspatni	IGKV, Raipur	21	Ankur sona	IGKV, Raipur
2	Thupi khajurchudi	IGKV, Raipur	22	NDLR 25	Chinsura RCS
3	KNM 118	IGKV, Raipur	23	NLR 3449	IGKV, Raipur
4	Bhuraj	BCKV, Nadia	24	NLR 40054	IGKV, Raipur
5	Judi dhan	IGKV, Raipur	25	Jai sri ram	IGKV, Raipur
6	Nal dhan	BCKV, Nadia	26	RGL 1880	IGKV, Raipur
7	CR dhan 308	BCKV, Nadia	27	RNR 15043	IGKV, Raipur
8	Tulaipanji	BCKV, Nadia	28	Mohini	IGKV, Raipur
9	Nassi jata	IGKV, Raipur	29	MTU 1081	BCKV, Nadia
10	Kalorat	BCKV, Nadia	30	RGL 2624	Chinsura RCS
11	Maryel 1011	IGKV, Raipur	31	MTU 1121	BCKV, Nadia

12	Supreme sona	IGKV, Raipur	32	BPT 284	BCKV, Nadia
13	CR 313	IGKV, Raipur	33	MTU 1010	BCKV, Nadia
14	Sourya super	IGKV, Raipur	34	Puti khali	Chinsura RCS
15	Poultry sona	IGKV, Raipur	35	Mawaii	Chinsura RCS
16	MTU 1232	BCKV, Nadia	36	RNR 15435	IGKV, Raipur
17	Manipur local	Chinsura RCS	37	Kalma	Chinsura RCS
18	MTU 1075	BCKV, Nadia	38	Sitasal	Chinsura RCS
19	MTU 1210	BCKV, Nadia	39	RNR 15048 (C)	IGKV, Raipur
20	MTU 1293	BCKV, Nadia	40	RGL 1880	IGKV, Raipur

Note: C: check variety; IGKV: Indira Gandhi Krishi Viswavidyalaya; BCKV: Bidhan Chandra Krishi Viswavidyalaya; RCS: Rice Research Centre.

Experimental layout

In order to assess 40 genotypes for 25 characteristics related to yield and cooking quality, an experiment was conducted in a randomised block design (RBD) with 3 replications. The study focused on important traits that demonstrated enhanced agronomic performance in terms of quality and production. Similar findings were also reported in earlier rice breeding researches (7–9). Each plot was 2 × 1.5 m in size, with 15 cm between plants and 20 cm between rows. Each experimental plot consisted of 90 plants and each replication of the genotypes under investigation was assigned at random. Ten plants were randomly selected from the central rows of each plot to avoid boundary effects when assessing yield and other characteristics.

Observations recorded

In order to fully utilise the genetic diversity found in the set of germplasm under examination, the current study included both yield and cooking-quality parameters in order to evaluate the genotypes of rice from both perspectives. The characters recorded were days to 50 % flowering, plant height (cm), number of tillers plant⁻¹, panicle length (cm), stem thickness (mm), number of filled grains panicle⁻¹, 1000-grain weight (g), flag leaf length (cm), biological yield (g), harvest index (%), days to 100 % maturity, grain yield plant⁻¹ (g), grain length (mm), grain breadth (mm), grain length-to-breadth (L/B) ratio, kernel length (mm), kernel breadth (mm), kernel length -to-breadth ratio, kernel length after cooking (mm), kernel breadth after cooking (mm), kernel L/B ratio after cooking, gelatinization temperature, amylose content (%), seed protein (%), kernel elongation ratio.

Estimation of cooking quality and biochemical parameters

Pre- and post-cooking kernel dimensions: International Rice Research Institute-Standard Evaluation System for Rice (IRRI SES)

Kernel length and breadth of 10 randomly selected dehusked grains per genotype were recorded using a digital slide calliper and their ratio (L/B) was computed.

Grain size and shape classification followed the International Rice Research Institute-Standard Evaluation System for Rice (IRRI SES) regulations (10). For post-cooking measurements, 5 g of milled rice was soaked in 15 mL of distilled water for 10 min and cooked in a 50 mL centrifuge tube for 20 min. Length and breadth of 10 randomly chosen cooked kernels were subsequently measured and the cooked L/B ratio was calculated accordingly.

Kernel elongation ratio: Elongation ratio was computed as the quotient of cooked kernel length to raw kernel length (11).

Gelatinisation temperature: Gelatinisation temperature was indirectly assessed through alkali spreading value and classified according to IRRI-SES standards (12).

Grain protein content: Grain protein content was estimated colourimetrically with bovine serum albumin as the standard (13).

Amylose content: Amylose content was estimated using the iodine-binding method and genotypes were classified accordingly (14, 15).

Statistical analysis

Mean data recorded across 3 replications were subjected to analysis of variance under a randomised complete block design to assess variation among genotypes. Genetic divergence among the 40 rice genotypes based on 25 traits was estimated using Mahalanobis D² statistics (16). The genotypes were grouped into clusters using Tocher's method and intra- and inter-cluster distances were calculated. The relative contribution of individual traits to total genetic divergence was also determined. Prior to multivariate analyses, trait data were standardised to minimise the influence of differences in measurement scales. Principal component analysis was performed on the standardised data and principal components with eigenvalues greater than one were retained for interpretation (17). All statistical analyses were performed using R statistical software.



3. Results and Discussion

Genetic divergence study for rice yield and cooking quality attributing traits

Using the mean values of 25 characters recorded across all 3 replications, Mahalanobis D² statistics was used, in accordance with Tocher's technique, to evaluate genetic divergence among 40 rice genotypes (18). The D² values for all possible genotype combinations were arranged in ascending order, as previously mentioned (19). Based on the contribution to overall genetic divergence, the relative contribution of each character was calculated.

Cluster analysis of rice germplasm

Genetic diversity among breeding materials is a prerequisite for achieving meaningful genetic advancement and developing superior varieties. In the present investigation, genetic divergence for yield and

cooking quality-attributing traits was assessed across 40 rice genotypes during the kharif season of 2025–26. Based on Tocher's method, the 40 genotypes were grouped into six distinct clusters as presented in Table 2. Cluster I contained the largest number of genotypes (17), whereas Cluster VI was monogenotypic and represented by Bamspatni. The occurrence of genotypes from different sources within the same clusters suggests that geographical origin was not necessarily associated with phenotypic divergence, as illustrated in Fig. 1. The lack of correspondence between geographical origin and genetic divergence suggests that phenotypic diversity is governed more by genetic constitution than by collection site, a pattern consistently observed in diverse rice germplasm evaluated through multivariate clustering (20). An earlier study of genetic divergence among 20 rice genotypes resulted in seven distinct clusters where maximum inter cluster distance was observed between clusters V and VI indicates distinctiveness between them (21).

Table 2. Clustering of rice genotypes for yield and cooking quality attributing character in rice genotypes

Cluster No.	Number of genotypes	Name of genotypes
I	17	CR dhan 308, Maryel 1011, MTU 1232, CR 313, MTU 1075, Nal dhan, Thupi khajurchudi, KNM 118, Manipur local, Sourya super, Nassi jata, Judi dhan, Puti khali, BPT 284, Supreme sona, Ankur sona, RGL 2624
II	9	Thulaippanji, MTU 1210, Poultry sona, NDLR 25, Jai sri ram, NLR 3449, RNR 15048, MTU 1010, NLR 40054
III	4	RNR 15435, Kalma, Mawaii, Kalorat
IV	6	Pusa basmati 1850, MTU 1081, Mohini, RNR 15043, RGL 1880, MTU 1121
V	3	Bhuraj, MTU 1293, Sitalal
VI	1	Bamspatni

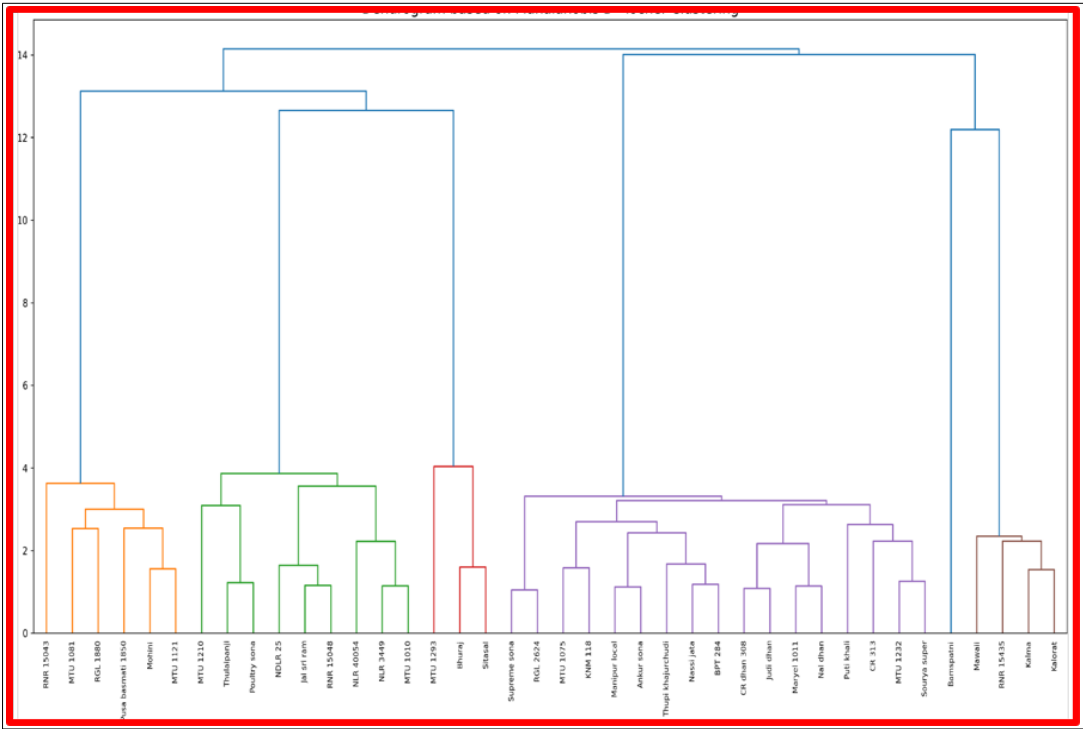


Fig. 1. Hierarchical clustering pattern of rice genotypes.



Intra and inter cluster distances for yield cooking quality attributes

The intra and inter cluster distances are presented in Table 3. Cluster V recorded the highest intra-cluster distance (1081.18), followed by cluster IV (962.96), cluster I (614.40), cluster III (596.98) and cluster II (563.80). Cluster VI recorded an intra-cluster distance of 0.00, as expected for a monogenotypic cluster. The maximum inter-cluster distance was observed between clusters III and IV (4983.52), followed by clusters IV and V (3926.32), clusters IV and VI (3889.51) and clusters II and IV (2700.66). The minimum inter-cluster distance was noted between clusters I and II (1089.20), indicating closer genetic affinity between genotypes in these two clusters. Comparable patterns of inter-cluster distances identifying the most genetically divergent clusters as optimal for cross hybridisation were reported previously where, 30 rice landraces were grouped into nine clusters with maximum inter-cluster distance observed between cluster I and VII suggesting future potential for hybridisation programmes (22). Greater inter-cluster distance reflects wider genetic divergence and crosses between highly separated clusters are expected to generate broader recombination and increased opportunities for transgressive segregants in subsequent generations (23). Hierarchical clustering grouped the studied genotypes into seven distinct clusters with a 63 % dissimilarity index. Clusters II, IV and VI were found to be most prominent among all clusters in terms of their potential for hybridisation to generate transgressive segregants (20).

Table 3. Intra and inter clustering distance for yield and cooking quality attributing characters in rice genotypes

Clusters	I	II	III	IV	V	VI
I	614.40	1089.20	2091.10	2704.37	1144.69	1143.31
II		563.80	1322.54	2700.66	1599.10	858.58
III			596.98	4983.52	1884.57	2174.76
IV				962.96	3926.32	3889.51
V					1081.18	1890.55
VI						0.00

Cluster mean analysis study

For yield and cooking-related traits, cluster III exhibited the highest plant height (135.29 cm), grain length (8.78 mm) and flag leaf length (115.18 cm), while cluster IV recorded the highest number of tillers hill⁻¹ (11.39) and grain L/B ratio (4.61) (Table 4). Cluster VI showed the highest number of days to maturity (143.33 days)

and flag leaf length (95.36 cm). Cluster V recorded the highest grain length (8.80 mm) and 1000 grain weight (32.54 g), while cluster I showed moderate performance across most yield traits. For cooking quality characters, cluster IV recorded the highest seed protein content (5.05 %) suggesting its superiority in biochemical quality traits. Cluster III had the highest mean kernel length after cooking (7.21 mm) and kernel width after cooking (2.25 mm), thereby representing better post-cooked kernels. On the other hand, cluster III recorded the highest amylose content (14.06 %) and highest kernel elongation ratio (3.11), thus demonstrating excellence in terms of cooking quality. While cluster VI was monogenotypic, it had a high mean number of days to 100 % maturity (143.33) and a high gelatinisation temperature (12.77), in addition to the moderate kernel elongation ratio (2.00). The results indicate significant levels of genetic variability among the studied genotypes. The monogenotypic nature of cluster VI (Bamspatni) suggests uniqueness in terms of its genetics, especially with respect to its superior gelatinisation temperature and days to maturity. Cluster III exhibited desirable values for several grain and cooking quality attributes, including kernel length after cooking and amylose content, whereas cluster IV showed higher values for tiller number, grain L : B ratio and seed protein content. The substantial divergence between these clusters, together with their contrasting trait profiles, suggests that parental combinations between clusters III and IV may be useful for combining complementary yield and grain-quality attributes. Seed protein content contributed the highest proportion (10.26 %) to total genetic divergence, followed by number of tillers hill⁻¹ (5.38 %) and amylose content (5.26 %). These traits were therefore major contributors to phenotypic differentiation among the evaluated genotypes and may be particularly useful for distinguishing diverse parental materials within the present germplasm set. Similar observations of monogenotypic clusters representing highly divergent accessions have been reported in a previous study of indigenous aromatic rice divergence which identified the maximum inter-cluster distance between solitary clusters (24). The character of seed protein (10.26 %) contributed the highest towards total divergence followed by numbers of tillers per hill (5.38 %) and amylose content (5.26 %) as presented in Fig. 2. Similarly it was reported that genotypes placed in divergent clusters based on yield and growth traits offered greater breeding potential for genetic improvement in rice (25). Overall, strategic hybridisation among genotypes from clusters III, IV, V and VI was recommended to exploit genetic diversity and develop superior rice lines with enhanced yield.

Table 4. Mean values of eight clusters and contribution (%) of yield and cooking quality attributing characters of rice genotypes

Clusters	DFF	PH	NTH	PL	ST	NFGP	GL	GB	GL/B	TGW	FLL	BY	HI
Cluster I	102.76	120.64	10.63	23.21	4.61	94.92	7.82	1.83	4.33	27.14	28.78	91.41	21.03
Cluster II	104.15	110.65	11.10	23.30	4.36	100.47	7.34	1.81	4.04	26.36	29.37	87.65	21.55
Cluster III	104.58	135.29	10.80	24.85	5.65	94.92	8.78	1.96	4.44	35.62	33.39	115.18	16.00



Cluster IV	97.61	102.34	11.39	23.56	4.54	94.80	8.12	1.76	4.61	25.80	31.10	94.95	18.63
Cluster V	101.22	114.02	10.20	24.57	5.13	79.24	8.80	1.96	4.50	32.54	30.43	86.63	19.46
Cluster VI	106.33	123.45	12.73	24.29	4.17	84.73	7.32	1.58	4.64	15.84	36.43	95.36	19.00
Contribution (%)	4.23	4.23	5.38	5.13	4.49	4.23	2.44	3.59	3.85	3.21	5.13	3.72	5.00

Note: DFF: days to 50 % flowering; PH: plant height(cm); NTH: number of tiller hill⁻¹; PL: panicle length (cm); ST: stem thickness (mm); NFGP: number of filled grain panicle⁻¹; GL: grain length (mm); GB: grain breadth (mm); GL/B: grain length-to-breadth ratio; TGW: 1000 grain weight (g); FLL: flag leaf length (cm); BY: biological yield (g); HI: harvest index (%).

Clusters	DM	KL	KB	KLBR	KLAC	KBAC	KLBAC	KER	GT	Amylose	Protein	Yield
Cluster I	138.63	6.10	1.61	3.79	6.46	1.83	3.50	1.05	3.59	7.82	2.21	18.55
Cluster II	138.22	5.69	1.60	3.57	6.08	1.93	3.11	1.05	3.11	13.43	2.60	18.68
Cluster III	140.75	6.57	1.80	3.59	7.21	2.25	3.22	1.09	3.50	14.06	1.97	17.63
Cluster IV	135.83	6.31	1.62	3.89	6.44	1.91	3.37	1.01	3.67	8.81	5.05	17.51
Cluster V	136.33	6.10	1.77	3.48	6.49	2.37	2.88	1.03	3.67	6.37	1.90	15.84
Cluster VI	143.33	4.67	1.40	3.33	4.98	1.64	2.85	1.06	2.00	12.77	1.84	17.90
Contribution (%)	4.87	2.31	1.92	2.18	1.54	3.85	2.31	3.46	3.46	5.26	10.26	3.97

Note: DM: days to 100 % maturity; KL: kernel length (mm); KB: kernel breadth (mm); KLBR: kernel length-to-breadth ratio; KLAC: kernel length after cooking (mm); KBAC: kernel breadth after cooking(mm); KLBAC: kernel length-to-breadth ratio after cooking; KER: kernel elongation ratio; GT: gelatinisation temperature; AMYL: amylose content (%); PRO: seed protein (%); Yield: seed yield plant⁻¹ (g).

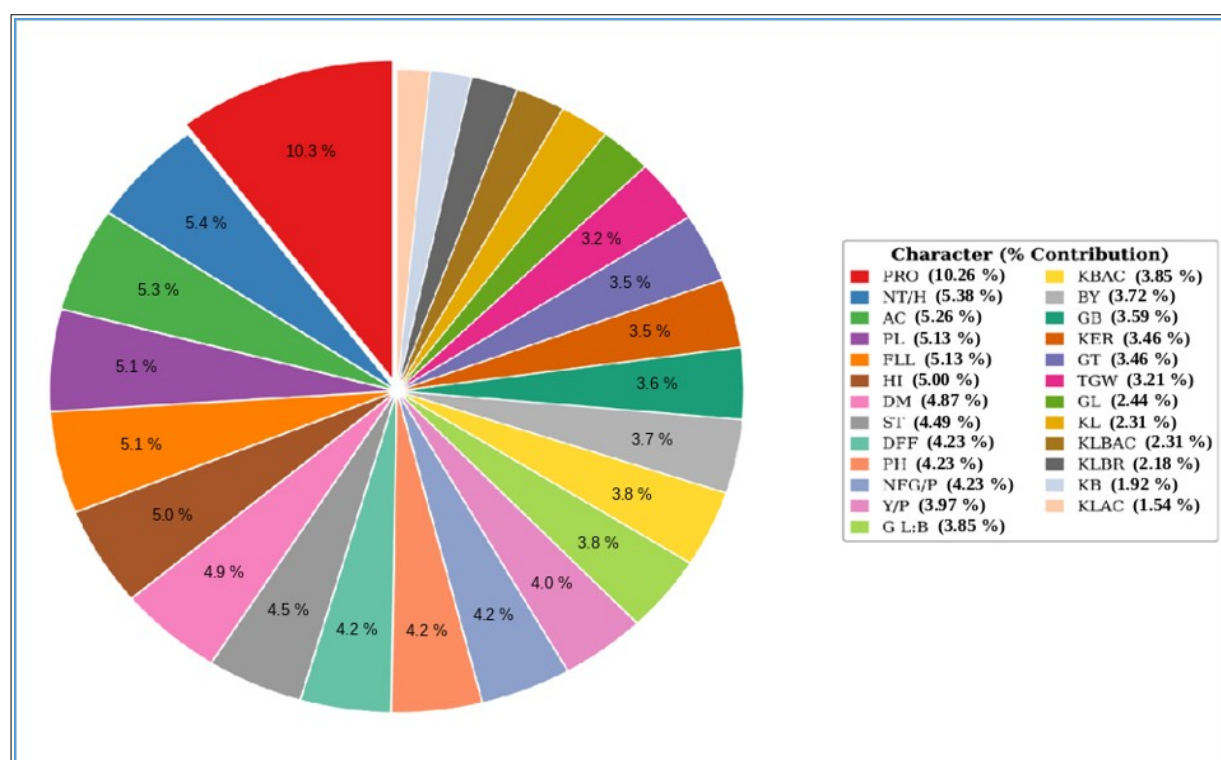


Fig. 2. Contribution (%) of different characters towards total divergence in rice genotypes.

Note: DFF: days to 50 % flowering; PH: plant height(cm); NT/H: number of tiller plant⁻¹; PL: panicle length (cm); ST: stem thickness (mm); NFG/P: number of filled grain panicle⁻¹; TW: 1000 grain weight (g); FLL: flag leaf length (cm); BY: biological yield (g); HI: harvest index (%); DM: days to 100 % maturity; Y/P: grain yield plant⁻¹ (g); GL: grain length (mm); GB: grain breadth (mm); G L : B: grain L : B ratio; KL: kernel length (mm); KB: kernel breadth; KLBR: kernel length-to-breadth ratio; KLAC: kernel length after cooking (mm); KBAC: kernel breadth after cooking(mm); KLBAC: kernel length-to-breadth ratio after cooking; GT: gelatinisation temperature; AC: amylose content (%); PRO: seed protein (%); KER: kernel elongation ratio (mm).

Principal component analysis

Principal component analysis (PCA) was carried out on 25 morphological and quality characters measured across 40 rice genotypes evaluated during the kharif season of 2025–26. The analysis yielded eight principal components, PC1 to PC8 with eigenvalues exceeding unity, which together accounted for 79.16 % of the cumulative variability present in the dataset (Table 5). The first principal component (PC1) recorded the highest eigenvalue of 4.95, contributing 19.78 % of the total variability, followed by PC2 with an eigenvalue of 4.68, explaining an additional 18.72 %, thereby giving a combined contribution of 38.50 %. Principal components 3 and 4 contributed 10.02 and 8.07 %, raising the cumulative proportion to 48.53 and 56.59 %, respectively. Principal components 5–8 constituted 6.68, 6.21, 5.06 and 4.62 % of the total variation, respectively, with an accumulative total of 79.16 %. The gradual fall in the eigenvalue in each successive component was illustrated graphically using the scree plot which showed that there was a pronounced inflection after PC2, suggesting that the major amount of biologically significant variation could be found within the first two components (Fig. 3). Earlier reports have suggested that up to 40–55 % of variation in rice germplasm was attributed to the first two to three components, wherein the traits related to rice grain quality and morphology

occupied most of the initial components (26). The dominance of morphological traits in PC1 and PC2 in the present investigation is consistent with the established importance of such characteristics in distinguishing varieties, especially in relation to grain quality improvement in basmati and non-basmati varieties. The presence of heavy loadings of 1000-grain weight and kernel breadth on PC1 indicated a clear distinction between genotypes having large-sized grains and those of small-sized grains. This trend is often seen in different rice panels wherein morphology and weight are important discrimination factors where characters like days to 50 % flowering, days to maturity and ear bearing tillers showed the maximum amount of variability towards total variance in rice (27). Out of the eight principal components, three PCs had an eigenvalue > 1 and accounted for a total variability of 78.71 % (28). The predominance of the first two principal components indicates that a relatively small number of latent variables effectively capture the majority of phenotypic variability among rice genotypes. Similar findings were reported, where it was observed that PC1 and PC2 together explained 78.19 and 76.13 % of the total variation across two growing seasons and successfully discriminated high-yielding rice cultivars based on yield-contributing morphological traits, confirming the robustness of PCA for genotype classification and parental selection in rice breeding (29).

Table 5. Eigen vectors and associated variances for principal component in rice genotypes

Eigen vectors	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Days to 50 % flowering	-0.066	-0.321	0.105	-0.212	-0.226	-0.043	-0.107	-0.334
Plant height (cm)	-0.215	-0.160	0.319	-0.181	-0.044	0.022	0.112	0.189
Number of tiller hill ⁻¹	-0.048	-0.048	-0.040	0.262	-0.320	-0.257	-0.313	-0.394
Panicle length (cm)	-0.169	-0.123	0.098	0.091	0.197	-0.088	-0.556	0.118
Stem thickness (mm)	-0.211	-0.075	0.018	0.051	-0.271	-0.026	-0.280	0.315
Number of filled grains panicle ⁻¹	0.016	0.063	0.244	0.491	-0.081	0.003	0.265	-0.072
Grain length (mm)	-0.144	0.373	-0.083	0.057	-0.139	-0.010	-0.025	0.031
Grain breadth (mm)	-0.384	-0.007	-0.041	0.001	0.241	0.142	0.072	-0.060
Grain L : B ratio	0.107	0.348	-0.031	0.013	-0.276	-0.093	-0.167	0.118
1000 grain weight (g)	-0.394	-0.040	0.006	-0.050	0.153	0.037	-0.085	-0.209
Flag leaf length (cm)	-0.085	-0.089	0.032	0.210	-0.028	-0.236	-0.260	0.461
Biological yield (g)	-0.179	0.101	0.405	0.055	0.115	-0.355	0.106	0.049
Harvest Index (%)	0.124	-0.195	-0.110	0.280	0.056	0.519	-0.183	-0.137
Days to 100 % maturity	-0.097	-0.240	0.137	-0.369	-0.223	-0.101	-0.070	-0.212
Kernel length (mm)	-0.194	0.372	0.026	-0.055	-0.039	0.155	-0.141	-0.143
Kernel Breadth (mm)	-0.397	0.041	-0.190	-0.014	0.016	0.049	0.017	-0.111
Kernel length and breadth (L : B) ratio	0.207	0.318	0.215	-0.058	-0.063	0.110	-0.171	-0.025
Kernel length after cooking (mm)	-0.261	0.297	0.082	0.027	-0.214	0.222	-0.011	-0.094
Kernel breadth after cooking (mm)	-0.316	0.131	-0.300	0.057	-0.005	0.038	0.088	0.071

Kernel L: B after cooking	0.144	0.172	0.397	-0.168	-0.136	0.184	-0.137	-0.119
Kernel elongation ratio (mm)	-0.180	-0.127	0.213	0.167	-0.380	0.130	0.315	0.069
Gelatinization temperature	-0.075	0.150	0.270	0.000	0.315	-0.234	0.138	-0.123
Amylose content (%)	-0.025	-0.144	-0.144	0.290	-0.303	-0.201	0.203	-0.082
Seed protein (%)	0.095	0.143	-0.162	0.113	0.155	-0.398	-0.054	-0.364
Grain yield plant ⁻¹ (g)	-0.014	-0.119	0.328	0.414	0.230	0.194	-0.150	-0.136
Eigenvalues	4.95	4.68	2.51	2.02	1.67	1.55	1.27	1.16
Variability (%)	19.78	18.72	10.02	8.07	6.68	6.21	5.06	4.62
Cumulative proportion (%)	19.78	38.50	48.53	56.59	63.27	69.47	74.54	79.16

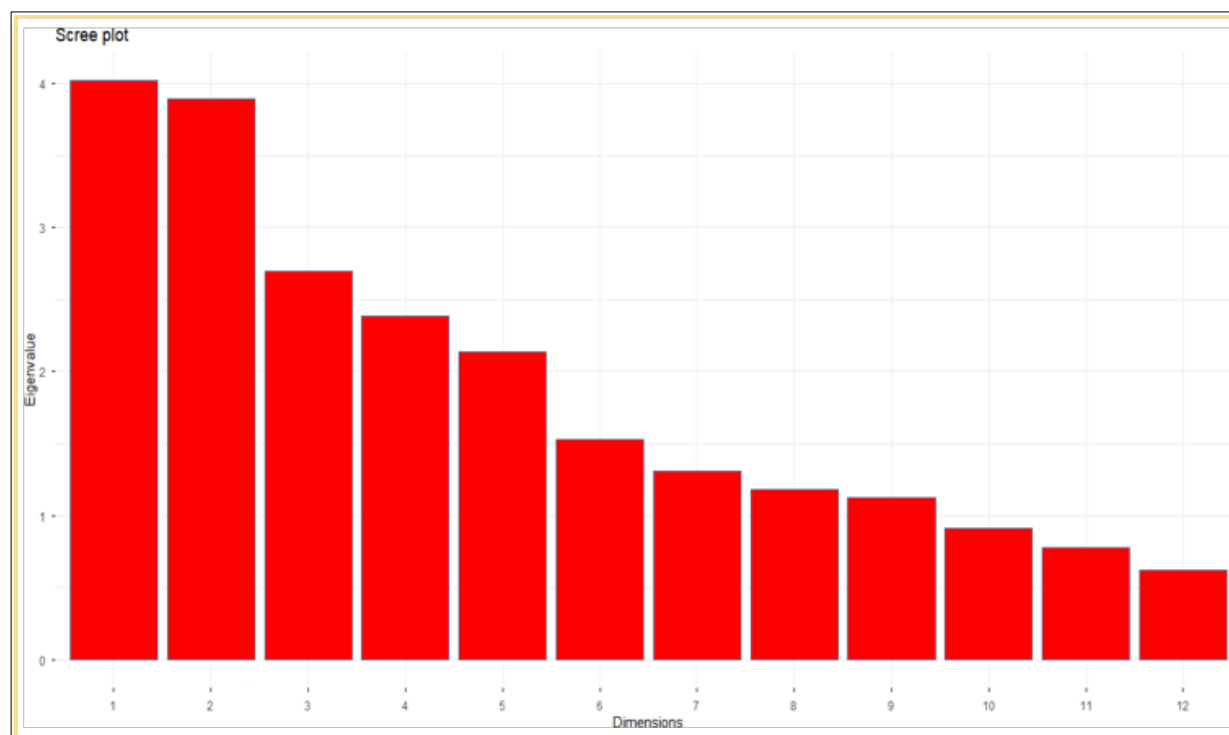


Fig. 3. Scree plot for principal component of analysis in rice genotypes.

Biplot based principal component

The PCA biplot, based on the first two principal components accounting for 19.8 and 18.7 % of the total variance, respectively illustrated the interrelationships among genotypes and traits (Fig. 4). Grain length, kernel length, cooked kernel length and cooked kernel breadth vectors showed a direction towards the upper right, implying that genotypes located in that region (Mawaii, Bhuraj, RGL 1880 and Kalorat) had better traits in terms of grain dimensions and cooking quality. On the other hand, the grain breadth and 1000-grain weight vectors showed a direction towards the positive dim1 direction, indicating that genotypes 36, 38 and 4 had higher grain weights. The PCA biplot differentiated genotypes according to their association with yield and grain-quality traits. Mawaii, Bhuraj, RGL 1880 and Kalorat were positioned closer to vectors representing grain and cooking quality attributes, suggesting their potential utility as donors for quality improvement. In

contrast, Judi dhan, RNR 15435 and Manipur local were more closely associated with biological yield, grain yield and filled grains per panicle, indicating their potential usefulness in breeding. The scree plot further validated the retention of eight principal components by showing a gradual flattening of the eigenvalue curve after PC2, supporting the interpretive focus on the first two dimensions. Overall, these findings are in accordance with reports that emphasised the utility of PCA biplots in identifying trait-specific donor genotypes for targeted improvement of rice (30). Similarly, the application of PCA across a diverse rice panel reported that the first two components accounted for the majority of phenotypic variation, with grain dimension traits and biochemical quality attributes loading prominently on PC1 and PC2, respectively, a pattern closely consistent with the present findings (31).

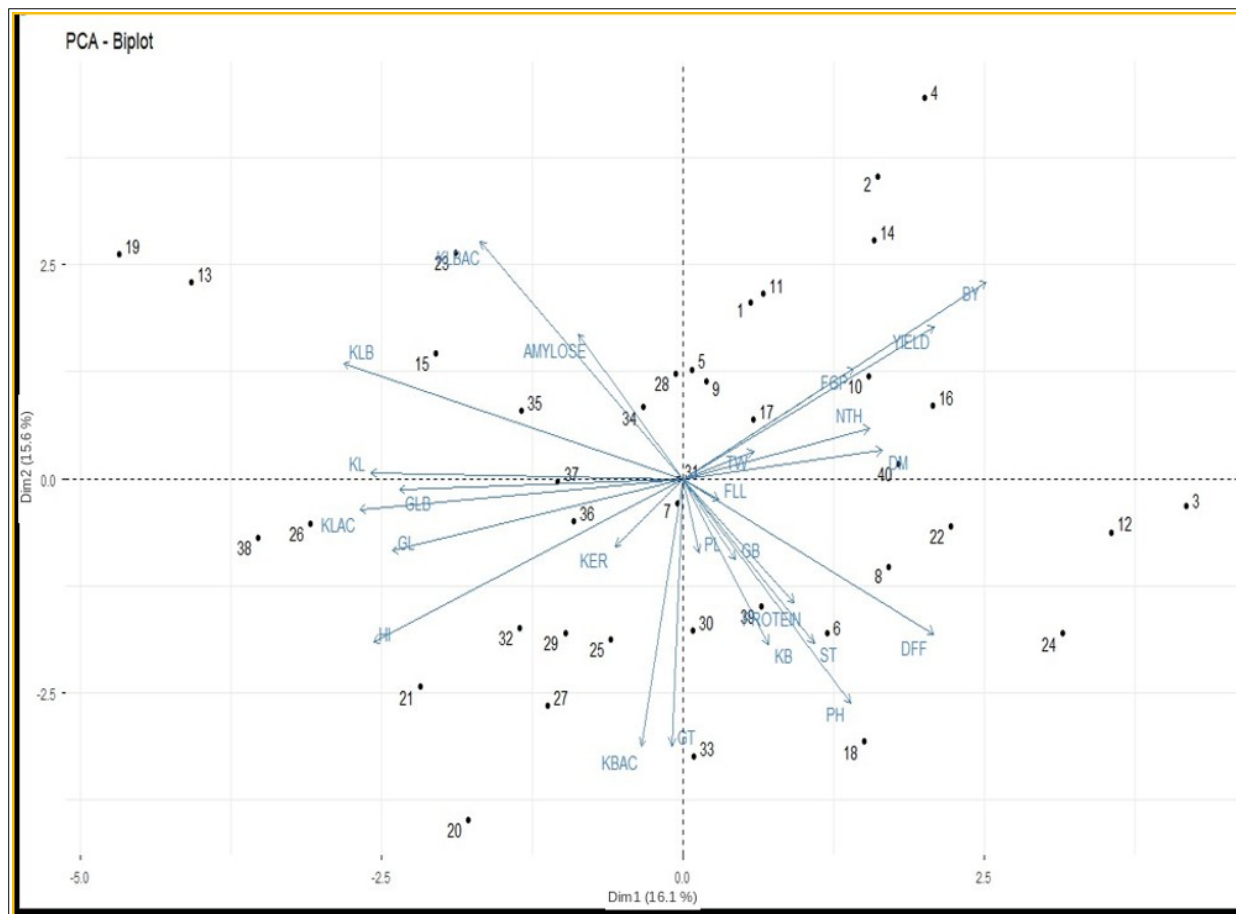


Fig. 4. Biplot for different characters in rice genotypes.

Note: PH: plant height (cm); DFF: days to 50 % flowering; TL: number of tiller plant⁻¹; PL: panicle length (cm); ST: stem thickness (mm); FGP: number of filled grain panicle⁻¹; GL: grain length (mm); GB: grain breadth (mm); L/B: grain length-to-breadth ratio; TW: 1000 grain weight (g); FLL: flag leaf length (cm); DM: days to 100 % maturity; BY: biological yield (g); HI: harvest index (%); yield: grain yield plant⁻¹ (g); KL: kernel length (mm); KB: kernel breadth (mm); KLAC: kernel length after cooking (mm); KBAC: kernel breadth after cooking (mm); PRO: seed protein (%); AMYL: amylose content (%); KER: kernel elongation ratio (mm).

4. Conclusion

Mahalanobis D² analysis grouped 40 rice genotypes into 6 clusters, with the maximum inter-cluster distance recorded between clusters III and IV (4983.52), identifying crosses among genotypes such as RNR 15435 × Pusa Basmati 1850 and Kalma × MTU 1081 as potentially most promising for transgressive segregation. Seed protein content (10.26 %) emerged as the principal contributor to overall genetic divergence. Principal component analysis extracted eight significant components collectively accounting for 79.16 % of the total variability, with grain weight, kernel breadth and grain breadth serving as the primary discriminating traits in PC1 while grain length and kernel length accounted for the primary dominating traits for PC2. Moreover, biplot visualisation clearly distinguished between those genotypes such as Mawaii, Bhuraj and RGL 1880 that were best suited for enhancing grain quality and those like Judi dhan, RNR 15435 and Manipur local that could be helpful for improving yield. Overall, these observations provide a sound foundation for selecting a wide variety of parental lines to enhance the desired traits in rice.

Acknowledgements

The authors sincerely acknowledge the Department of Genetics and Plant Breeding, M S Swaminathan School of Agriculture, Centurion University of Technology and Management, Odisha for providing financial and technical assistance, laboratory facilities and continuous support throughout the course of this experiment.

Authors' contributions

VKG conceptualised and designed the research, supervised the study, evaluated the data and finalised the manuscript. MPSS conducted the field experiments and contributed to drafting the initial manuscript. S and AKB performed data evaluation and provided critical comments during manuscript revision. RC and PM assisted in the data evaluation and interpretation of the results. AR revised and edited the manuscript for technical accuracy. GK and AR contributed to improving the structure, language and logical flow of the manuscript. All authors have read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest

Authors do not have any conflict of interests to declare.

Ethical issues

None

References

- Khush GS. Origin, dispersal, cultivation and variation of rice. *Plant Mol Biol.* 1997;35(1–2):25–34. <https://doi.org/10.1023/A:1005810616885>
- Directorate of Economics and Statistics. Agricultural Statistics at a Glance 2022–23. New Delhi: Department of Agriculture and Farmers Welfare, Government of India; 2023.
- Verma V, Vishal B, Kohli A, Kumar PP. Systems-based rice improvement approaches for sustainable food and nutritional security. *Plant Cell Rep.* 2021;40(11):2021–36. <https://doi.org/10.1007/s00299-021-02790-6>
- Nachimuthu VV, Muthurajan R, Duraialaguraja S, Sivakami R, Pandian BA, Ponniah G, et al. Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in *Oryza sativa*. *Rice.* 2015;8(1):30. <https://doi.org/10.1186/s12284-015-0062-5>
- Juliano BO. Rice Chemistry and Technology. 2nd ed. St Paul, Minnesota: American Association of Cereal Chemists; 1985.
- Yan W, Kang MS. GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists and Agronomists. Boca Raton, Florida: CRC Press; 2003. <https://doi.org/10.1201/9781420040371>
- Liu M, Fan F, He S. Creation of elite rice with high yield, superior quality and high resistance to brown planthopper based on molecular design. *Rice.* 2022;15:17. <https://doi.org/10.1186/s12284-022-00563-7>
- Durbha SR, Siromani N, Jaldhani V. Dynamics of starch formation and gene expression during grain filling and its possible influence on grain quality. *Sci Rep.* 2024;14:6743. <https://doi.org/10.1038/s41598-024-57010-4>
- Champagne ET, Bett-Garber KL, McClung AM, Bergman C. Sensory characteristics of diverse rice cultivars as influenced by genetic and environmental factors. *Cereal Chem.* 2004;81(2):237–43. <https://doi.org/10.1094/CCHEM.2004.81.2.237>
- International Rice Research Institute. Standard Evaluation System for Rice (SES). 4th ed. Manila, Philippines: IRRI; 1996.
- Hussain A, Bhatti MB, Majid A. Cooking and eating qualities of rice varieties. *Pak J Agric Res.* 1987;8(4):521–5.
- Little RR, Hilder GB, Dawson EH. Differential effect of dilute alkali on 25 varieties of milled white rice. *Cereal Chem.* 1958;35(2):111–26.
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. *J Biol Chem.* 1951;193(1):265–75. [https://doi.org/10.1016/S0021-9258\(19\)52451-6](https://doi.org/10.1016/S0021-9258(19)52451-6)
- Juliano BO. A simplified assay for milled rice amylose. *Cereal Sci Today.* 1971;16(11):334–40.
- Kumar I, Khush GS. Genetics of amylose content in rice (*Oryza sativa* L.). *J Genet.* 1986;65(1–2):1–11. <https://doi.org/10.1007/BF02923530>
- Mahalanobis PC. A statistical study at Chinese head measurement. *J Asiatic Soc Bengal.* 1928;25:301–7.
- Banfield CF. Multivariate analysis in Genstat. *J Stat Comput Simul.* 1978;6(3–4):211–22. <https://doi.org/10.1080/00949657808810190>
- Rao CR. Advanced Statistical Methods in Biometrical Research. New York: John Wiley & Sons; 1952.
- Singh RK, Chaudhary BD. Biometrical Methods in Quantitative Genetic Analysis. New Delhi: Kalyani Publishers; 1977.
- Parikh M, Kumar D, Gauraha D, Sharma B, Saxena RR, Sahu PK. Unveiling the genetic variability and diversity in rice germplasm based on phenological traits and grain quality parameters. *J Adv Biol Biotechnol.* 2025;28(6):1142–57. <https://doi.org/10.9734/jabb/2025/v28i62473>
- Roy A, Rout S, Hijam L, Sadhu S, Pavithra S, Ghosh A, et al. Multivariate genetic analyses unveil the complexity of grain yield and attributing traits diversity in *Oryza sativa* L. landraces from North-Eastern India. *Plant Sci Today.* 2024;11(2):29–37. <https://doi.org/10.14719/pst.2500>
- Beninu BS, Ekka JP, Sherring A, Sahi VP, Lavanya GR, Masih S. Study on genetic diversity and character associations in traditional and indigenous rice (*Oryza sativa* L.) genotypes under rainfed conditions. *Int J Plant Soil Sci.* 2025;37(5):22–7. <https://doi.org/10.9734/ijpss/2025/v37i55425>
- Nidhi DC, Deo BI, Kumar A, Lohani P, Singh S, Tripathi T, et al. Genetic divergence and quantitative trait analysis of yield-related attributes in rice (*Oryza sativa* L.) through Mahalanobis D² analysis. *J Adv Biol Biotechnol.* 2025;28(12):1054–63. <https://doi.org/10.9734/jabb/2025/v28i123449>
- Mondal S, Purohit A, Hazra A, Das S, Chakrabarti M, Khan MR, et al. Intraspecific variability of rice root-knot nematodes across diverse agroecosystems for sustainable management. *Sci Rep.* 2024;14(1):30032. <https://doi.org/10.1038/s41598-024-73980-x>
- Shrestha J, Subedi S, Kushwaha UKS, Maharjan B. Evaluation of growth and yield traits in rice genotypes using multivariate analysis. *Heliyon.* 2021;7(10):e08078. <https://doi.org/10.1016/j.heliyon.2021.e07940>
- Rani MH, Faruquee M, Khanom MS, Begum SN. Genetic variability and multivariate studies on grain physical properties of rice (*Oryza sativa* L.) landraces. *SABRAO J Breed Genet.* 2022;54(1):1–10. <https://doi.org/10.54910/sabao2022.54.1.1>
- 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):905–9. <https://doi.org/10.22271/phyto.2021.v10.i1m.13451>
- Roy A, Hijam L, Rout S. Diversity analysis for cooking quality traits in North-East Indian rice landraces. *ORYZA.* 2023;60(3):388–96. <https://doi.org/10.35709/ory.2023.60.3.2>
- Al Galib MA, Farzana S, Chakrobarty T, Islam MZ, Shirazy BJ, Rahman MA, et al. Multivariate analysis for identifying high-yielding rice cultivars based on seed yield and morphological traits. *Discover Plants.* 2025;2:256. <https://doi.org/10.1007/s44372-025-00345-7>
- Behera PP, Singh SK, Sivasankarreddy K, Majhi PK, Reddy BJ, Singh DK. Yield-attributing traits of high-zinc rice (*Oryza*



sativa L.) genotypes with special reference to principal component analysis. Environ Conserv J. 2022;23(3):458–70. <https://doi.org/10.36953/ECJ.10302233>

31. Praveen N, Srikanth B, Lal AB, Kumar A. Principal component analysis for assessment of genetic variability in rice genotypes for yield and grain quality traits. Plant Genet Resour. 2023;21(2):88–97.

Additional information

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

Reprints & permissions information: Available at https://horizonpublishing.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: See

https://horizonpublishing.com/journals/index.php/PST/indexing_abstracting.

Copyright: © The Author(s). This is an open-access article distributed under the Creative Commons Attribution License CC BY 4.0.

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

