



RESEARCH ARTICLE

Multivariate analyses for identification of high-yielding lines of rice based on yield and its component characters

Ujwala Chennareddy^{1*}, Mallapuram Shanthi Priya¹, Chintala Sreelakshmi², Shaik Nafeez Umar³, Pavani Pulla¹ & Chaduvula Eshwar Sai Prasad¹

¹Department of Genetics and Plant Breeding, Sri Venkateswara Agricultural College, Tirupati 517 502, Andhra Pradesh, India

²Plant Breeding, Agricultural Research Station, Nellore 524 001, Andhra Pradesh, India

³Department of Statistics and Computer Application, Sri Venkateswara Agricultural College, Tirupati 517 502, Andhra Pradesh, India

*Correspondence email - ujwalareddy0111@gmail.com

Received: 09 February 2026; Accepted: 12 June 2026; Available online: Version 1.0: 24 September 2026

Cite this article: Ujwala C, Mallapuram SP, Chintala S, Shaik NU, Pavani P, Chaduvula ESP. Multivariate analyses for identification of high-yielding lines of rice based on yield and its component characters. Plant Science Today. 2026;13(sp5):01–13. <https://doi.org/10.14719/pst.14045>

Abstract

The present study assessed 64 advanced breeding lines of rice and four checks to evaluate the selection criteria for yield improvement. Analysis of variance (ANOVA) for 15 yield-contributing attributes revealed significant genotypic differences, indicating considerable genetic variability. The genotypes NLR 3128, NLR 20002, NLR 20016, NLR 33891 and NLR 34417 performed better. Traits such as number of panicles per plant (NPP), panicle weight (PW), number of filled grains per panicle (NFGP) and number of chaffy grains per panicle (NCGP), total number of grains per panicle (TGP), straw yield (SY) and grain yield per plant (GYP) recorded high estimates of phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV), heritability and genetic advance suggesting their potential for genetic improvement. Grain yield exhibited significant positive correlations with the number of tillers per plant (NTP), panicle length (PL), test weight (TW), SY and harvest index (HI), along with their direct effects, indicating the importance of these traits in selection. A composite selection index incorporating NTP, PL, 1000-grain weight and HI showed maximum efficiency. A genetic diversity study via D² statistics grouped genotypes into 12 clusters; of these, 42 genotypes were in Cluster I, while Clusters X, XI and XII were solitary. Clusters IX and XI exhibited maximum inter-cluster divergence. The unweighted pair group method with arithmetic mean (UPGMA)-based hierarchical clustering approach classified 64 genotypes into nine clusters. Principal component analysis (PCA) revealed that the first six principal components (PCs) explain about 91.5 % of total variance. These findings highlight the presence of significant genetic diversity among genotypes for identifying key traits for targeted breeding strategies to improve rice yield potential.

Keywords: D² analysis; genetic diversity; genetic variability; principal component analysis; selection criterion; unweighted pair group method with arithmetic mean

Introduction

Rice (*Oryza sativa* L.) is the world's most diverse cereal crop that serves as a primary staple food. It is cultivated across 165 M ha worldwide, which accounts for 11 % of the global cultivable land, with an average production and productivity of 522.0 M t and 5127 kg ha⁻¹ respectively (1). Asia constitutes about 60 % of the world's population and contributes to 92 % of the world's rice production and 90 % of the world's rice consumption, followed by Africa (2).

India ranks first globally in terms of rice cultivation area and production, followed by China, and ranks second in terms of rice consumption, following China. Roughly 46.28 M ha, or one-fourth, of India's total cultivable land is used for rice. With a productivity rate of 2809 kg ha⁻¹, this results in an average yearly production of 130.29 M t (3).

In India, West Bengal is the leading producer of rice with an average annual production of 15.75 M t. Among the other Indian states, Andhra Pradesh ranks fifth with an area of 2.16 M ha under

rice cultivation, an average production of 7.49 M t (4) and an average productivity of 3896 kg ha⁻¹ (5). As the world's population is expected to reach 9.3 billion by 2050, efforts are being made to increase rice yields to necessarily meet the demand for rice.

Yield is a complex quantitative trait influenced by various contributing characters; thus, the main objective of rice breeding is to develop varieties with higher yield and improved nutritional quality. The presence of a substantial amount of variability in the rice germplasm collections and existing gene pool is essential in selecting for higher yield. Assessment of this variability is done using the proportion of phenotypic variance due to genetic factors (termed 'heritability') among the individuals in the population. The degree of projected improvement for a trait under certain selection pressures can be measured using genetic advance, which refers to the mean improvement in the genotypic value of selected plants relative to the original parental population.

Improvement of several characters simultaneously forms the basis of selection in certain breeding programmes. The understanding that improvement in one character may lead to a

corresponding increase or decrease in related traits emphasises the importance of considering all significant traits at the same time. Index selection or selection indices, is a method of artificial selection that allows for this simultaneous consideration of various advantageous characteristics (6). Since the genotypic value of an individual cannot be directly estimated, a linear function of observable phenotypic values is utilised instead.

Genetic diversity, which measures the total variation present in a population, forms the basis for various breeding strategies (7). Knowledge of genotype divergence guides heterosis breeding approaches and provides information on pedigree relationships (8). Heterosis can be maximised by using diverse parental lines in the breeding schemes which are identified using several multivariate approaches such as D^2 statistics, principal component analysis (PCA) and the UPGMA approach. These techniques not only make it easier to find the best parental combinations but also help to improve several traits, such as yield, disease resistance and many others that are essential for food security.

Materials and Methods

Experimental site

The present study was conducted during the kharif season 2023–24 to assess rice breeding lines for yield and related traits at the Agricultural Research Station, Nellore, in the Southern agro-climatic zone of Andhra Pradesh. The station is located at 14°21'N latitude and 79°59'E longitude, 20 m above mean sea level (AMSL). With a pH range of 7–8, the sandy clay loam soil at the location offers a neutral to slightly alkaline climate that is ideal for crop development.

Experimental material

The experimental material comprised 64 advanced breeding lines of rice and 4 checks obtained from the Agricultural Research Station, Nellore. The list of genotypes is presented in Table 1.

Experimental design

A nursery with raised and well-prepared beds was set up during the kharif season 2023–24 for assessing 68 advanced rice breeding lines. Thirty-day-old seedlings were transplanted in the main field using alpha lattice experimental design in two replications with a total of 68 genotypes. Blocking was imposed to control local field heterogeneity, including soil fertility gradients and micro-environmental variation, by grouping relatively homogeneous experimental units. Genotypes were randomised

within blocks, ensuring that each genotype occurred once per replication. Field view at the crop maturity stage is shown in Fig. 1. The seedlings were planted in 3.75 m long rows with a spacing of 20 × 15 cm. A baseline fertiliser dosage of 80 : 60 : 40 kg of NPK ha⁻¹ was applied one day after transplanting to ensure good crop development and all essential crop protection measures were implemented during the crop growth period.

Data collection on biometric observation

Data was collected on 15 quantitative traits in the study. Days to maturity (DTM) and days to 50 % flowering (DFF) have been recorded per plot. Plant height (PH), number of tillers per plant (NTP) and number of panicles per plant (NPP) were counted on five plants per genotype per replication and averaged, while panicle length (PL), panicle weight (PW), number of filled grains per panicle (NFGP), number of chaffy grains per panicle (NCGP), total number of grains per panicle (TGP) and spikelet fertility (SF) were measured on five panicles per genotype per replication and averaged. Straw yield (SY), harvest index (HI) and grain yield per plant (GYP) were measured on 10 plants per genotype for each replication and averaged. The TW was calculated by weighing 1000 full grains. This approach ensures accurate and representative data collection for

Table 1. List of 64 advanced breeding lines of rice along with 4 checks

Sl. No.	Genotype	Sl. No.	Genotype	Sl. No.	Genotype
1	NLR 20084	24	NLR 3111	47	NLR 3382
2	NLR 20104	25	NLR 3117	48	NLR 3339
3	NLR 20022	26	NLR 3128	49	NLR 3336
4	NLR 20002	27	NLR 3186	50	NLR 3338
5	NLR 33718	28	NLR 3197	51	NLR 3465
6	NLR 33891	29	NLR 3187	52	NLR 3461
7	NLR 33648	30	NLR 3188	53	NLR 3433
8	NLR 33166	31	NLR 3199	54	NLR 3418
9	NLR 33720	32	NLR 3380	55	NLR 3389
10	NLR 20016	33	NLR 3390	56	NLR 33891
11	NLR 33801	34	NLR 3130	57	NLR 3256
12	NLR 33810	35	NLR 3190	58	NLR 40058
13	NLR 20023	36	NLR 3291	59	NLR 3083
14	NLR 34417	37	NLR 3252	60	NLR 40021
15	NLR 20106	38	NLR 3264	61	NLR 3367
16	NLR 20123	39	NLR 20156	62	NLR 3237
17	NLR 20128	40	NLR 3328	63	NLR 3411
18	NLR 20127	41	NLR 3261	64	NLR 3302
19	NLR 20131	42	NLR 3379	Checks	
20	NLR 20136	43	NLR 3385	65	BPT 5204
21	NLR 20146	44	NLR 3374	66	MTU 1010
22	NLR 20150	45	NLR 3424	67	NLR 33892
23	NLR 20083	46	NLR 3135	68	NLR 9674

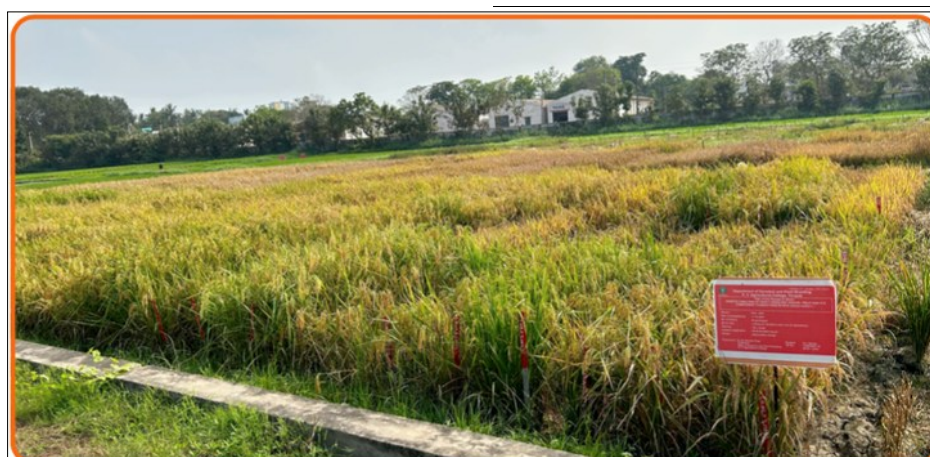


Fig. 1. Field view at the time of maturity.

meaningful genotype comparisons across all traits.

Statistical analysis

The phenotypic data was evaluated using R software for descriptive statistics and analysis of variance (ANOVA). The phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) were calculated using standard formulas (8). Heritability in the broad sense was calculated and categorised using the standard procedure (9). Genetic advances were estimated and classified using standard methods (10). Correlations among various traits were calculated and these correlation coefficients were partitioned into direct and indirect effects using path coefficient analysis (11, 12). Classical selection indices were constructed based on the discriminant function (5, 13). Genetic diversity was estimated using D² statistics genotypes were grouped into different clusters using Tocher's method in Agri analyse software (14–16). Principal component analysis was performed to determine the contribution of traits (17). Hierarchical clustering based on the distance matrix using the unweighted pair group method with arithmetic mean (UPGMA) to assess the phenotypic diversity among the advanced lines of rice was computed using DARwin software version 6 (18).

Results and Discussion

Analysis of variance

The ANOVA results for all 15 characters indicated that the mean sum of squares due to treatments is statistically significant for each trait studied, as shown in Table 2.

The results of the ANOVA revealed a statistically significant treatment mean sum of squares, suggesting that the genotypes for

the traits under study differ significantly from one another. A significant treatment mean sum of squares also suggests that there is a wide range of genetic expression in the experimental material and a significant variability in the attributes among the genotypes. Such diversity is important because it indicates that the improvement of the traits is possible through selection and breeding.

Mean performance and ranges of characters

The mean values and range values of 15 quantitative characters studied (Table 3 and 4) disclosed considerable variability among the genotypes. Wide variation was observed for DFF (90–133 days) and DTM (120–165 days), indicating the presence of both early and late-maturing types. Yield components such as NTP, NPP and PW showed substantial differences, with NLR 33720 demonstrating superior performance in tillering and panicle number.

A particularly striking variation was observed in NFGP (74.5–404.8) and TGP (88.0–453.7), highlighting differences in reproductive efficiency with a strong diversity in yield potential. Grain yield per plant varied widely (12.5–74.1 g), with NLR 33891 as the highest yielder. Overall, the observed variability suggests a good scope for selection and improvement.

Genetic variability

The estimates of PCV and GCV are presented in Table 4. In general, PCV values were slightly higher than GCV for all the traits, indicating a limited influence of environmental factors on their expression. Low PCV and GCV were observed for DFF and DTM, suggesting limited variability and relative stability of these traits. Moderate variability was recorded for PH and PL.

On the other hand, high PCV and GCV estimates were

Table 2. Analysis of variance for grain yield and its component traits

Sl. No.	Characters	Mean sum of squares			
		Replications (df: 1)	Treatments (df: 67)	Rep/block (df: 32)	Residuals (df: 35)
1	Days to 50 % flowering	39.38	182.6**	26.83	23.59
2	Days to maturity	47.70	201.26**	27.37	24.65
3	Plant height (cm)	18.80	511.30**	17.60	12.5
4	Number of tillers plant ⁻¹	0.324	20.696**	0.519	0.51
5	Number of panicles plant ⁻¹	0.417	17.756**	0.267	0.188
6	Panicle length (cm)	0.580	12.159**	0.154	0.612
7	Panicle weight (g)	0.100	5.137**	0.066	0.113
8	Number of filled grains panicle ⁻¹	246.0	9214.0**	126.0	103
9	Number of chaffy grains panicle ⁻¹	0.030	311.90**	1.06	1.28
10	Total number of grains panicle ⁻¹	111.0	11711**	95.00	87
11	Spikelet fertility (%)	17.66	51.30*	23.8	24.8
12	Test weight (g)	0.006	23.392**	0.482	0.617
13	Straw yield (g)	13.36	269.30**	4.23	8.26
14	Harvest index (%)	11.61	70.740**	13.23	10.58
15	Grain yield plant ⁻¹ (g)	3.86	300.87**	3.24	3.21

df: degree of freedom; *significant at 5 % level; **significant at 1 % level.

Table 3. Range of means of characters observed among genotypes

Sl. No.	Characters	Highest	Lowest
1	Days to 50 % flowering (days)	NLR 3256 (133)	NLR 3302 (90)
2	Days to maturity (days)	NLR 3256 (165)	NLR 3302 (120)
3	Plant height (cm)	NLR 33648 (132.4)	NLR 40021 (66.9)
4	Number of tillers plant ⁻¹	NLR 33720 (25.5)	NLR 3186 (11.1)
5	Number of panicles plant ⁻¹	NLR 33720 (21.0)	NLR 3489 (8.6)
6	Panicle length (cm)	NLR 9674 (28.6)	NLR 40021 (15.7)
7	Panicle weight (g)	NLR 3199 (8.7)	NLR 3237 (1.7)
8	Number of filled grains panicle ⁻¹	NLR 20128 (404.8)	NLR 3237 (74.5)
9	Number of chaffy grains panicle ⁻¹	NLR 3424 (60.6)	NLR 3411 (7.9)
10	Total number of grains panicle ⁻¹	NLR 33891 (453.7)	NLR 3237 (88.0)
11	Spikelet fertility (%)	NLR 3338 (94.9)	NLR 3424 (80.7)
12	Test weight (g)	NLR 3390 (31.8)	BPT 5204 (14.8)
13	Straw yield (g)	NLR 3117 (73.3)	NLR 3237 (21.5)
14	Harvest index (%)	NLR 3128 (62.1)	NLR 33891 (31.3)
15	Grain yield plant ⁻¹ (g)	NLR 33891 (74.1)	NLR 40021 (12.5)

Table 4. Variability parameters for yield and its component traits in advanced breeding lines of rice

Sl. No.	Character	Mean	Range		Coefficient of variation		Heritability (h ² b)	GA	GAM
			Max.	Min.	Genotypic	Phenotypic			
1	DFF	117.9	133.5	90.0	7.5	8.6	75.8	15.9	13.5
2	DTM	151.2	165.5	120	6.2	7.1	77.2	16.9	11.2
3	PH	104.1	132.4	66.9	15.1	15.6	94.3	31.5	30.3
4	NTP	16.2	25.5	11.1	19.7	20.2	95.2	6.4	39.5
5	NPP	13.6	21.0	8.6	21.8	22.1	97.5	6.0	44.4
6	PL	22.5	28.6	15.7	10.7	11.2	91.1	4.7	21.1
7	PW	5.5	8.7	1.7	28.8	29.3	96.5	3.2	58.2
8	NFGP	233.9	404.8	74.5	28.8	29.2	97.6	58.7	137.2
9	NCGP	26.4	60.6	7.9	47.5	47.9	99.3	25.6	97.0
10	TGP	260.3	453.7	88.0	29.3	29.5	98.5	59.9	155.8
11	SF	90.0	94.9	80.7	3.4	3.9	76.9	6.1	6.8
12	TW	20.3	31.8	14.8	16.3	16.7	95.2	6.6	32.8
13	SY	46.5	73.3	21.5	24.7	25.3	95.4	23.1	49.7
14	HI	79.4	164.0	45.8	26.9	27.4	95.8	42.9	54.1
15	GYP	12.5	74.1	12.5	33.5	33.9	97.9	24.9	68.3

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g); GA: genetic advance; GAM: genetic advance as percentage of mean.

observed for key yield contributing traits such as NTP, NPP, PW, NFGP, NCGP, TGP, SY, HI and GYP. The high variability observed for these traits may be attributed to the diverse genetic background of the genotypes and their differential response to environmental conditions, particularly for complex quantitative traits governed by multiple genes. Among these, chaffy grains per panicle exhibited the highest variability, indicating greater sensitivity to both genetic and environmental factors.

The high variability estimates observed for the NPP are in accordance with earlier reports (19, 20). The present findings on PW are consistent with the reports of earlier studies (21, 22). Similarly, variability observed for NFGP corroborates the results previously reported by other researchers (20, 23). Comparable observations for NCGP were reported in earlier studies (23, 24), whereas similar findings for TGP were also documented previously (19, 25). Furthermore, similar trends for SY, HI and GYP have also been documented in previous studies.

Genetic parameters

The estimates of heritability and genetic advance as percentage of mean (GAM) are presented in Table 4, providing insight into the nature of gene action governing the traits. All the traits under study exhibited high heritability, indicating a predominant role of genetic factors in their expression. However, the magnitude of GAM varied considerably among traits, ranging from low in spikelet fertility (6.8 %) to exceptionally high in TGP (155.8 %).

Traits such as PH, NTP, NPP, PW, NFGP, NCGP, TGP, test weight (TW), SY, HI and GYP exhibited high heritability coupled with high GAM. This combination is indicative of the predominance of additive gene action, suggesting that these traits are amenable to improvement through direct selection. The remarkably high GAM observed for total grains per panicle and filled grains per panicle underscores the presence of substantial, exploitable genetic variability.

In contrast, DFF and DTM recorded high heritability with moderate GAM, while spikelet fertility exhibited high heritability but low GAM, implying the possible involvement of non-additive gene effects and environmental influence, thereby reducing the efficiency of selection. These results suggest that the traits governed predominantly by additive gene action can be effectively improved

through simple selection strategies, offering significant potential for genetic enhancement.

Similar results for PH were reported in previous studies (26, 27), while comparable findings for the NTP were also documented earlier (28, 29). Similar results for the NPP were also reported (30, 31). The results for PL and PW were also confirmed by previous studies (22, 32). The results for the NFGP agreed with earlier studies (27, 33). Likewise, comparable results for the NCGP have also been documented previously (33, 34). The variability observed for TGP in the present investigation was comparable with earlier reports (31, 32). Earlier studies have also documented similar trends for TW (29) and SY per plant (31, 34). In addition, the present findings for HI and GYP are in close conformity with previous observations (27, 30).

Correlation analysis

The correlation coefficients among the traits are presented in Table 5, revealing the direction and magnitude of associations with grain yield. Notably, most of the traits exhibited significant positive correlations with GYP, indicating their potential importance for yield improvement. Among these, HI ($r = 0.722^{**}$), NPP ($r = 0.695^{**}$), PW ($r = 0.636^{**}$) and SY ($r = 0.638^{**}$) showed strong positive associations, suggesting their major contribution to grain yield.

Moderate but significant positive correlations were observed for NTP, NFGP, TGP, PH and DTM, indicating their indirect role in influencing yield. Test weight also showed a positive association, though of lower magnitude. Interestingly, spikelet fertility exhibited a negative and non-significant correlation ($r = -0.085$) with grain yield, suggesting its limited contribution under the present conditions.

Overall, the strong positive associations of key yield components with grain yield indicate that selection for traits such as panicle number, PW and HI would be effective in enhancing grain yield.

The present findings for DFF and DTM corroborate earlier reports (35–37). Comparable observations for PH were documented earlier (38, 39). The findings related to the NTP were consistent with earlier findings (40). Similar trends for the number of effective tillers per hill were observed in previous studies (38, 41). The observed PL and PW agreed with the earlier reports (20, 38, 42). Similarly, the results pertaining to the NFGP for TGP were consistent with previous

Table 5. Genotypic correlations coefficients among yield and its component characters in advanced breeding lines of rice

Trait	DFF	DTM	PH	NTP	NPP	PL	PW	NFGP	NCGP	TGP	SF	TW	SY	HI	GYP
DFF	1	0.969**	0.707**	0.1233	0.257*	0.667**	0.670**	0.583**	0.440**	0.585**	-0.0663	0.109	0.313**	0.236*	0.376**
DTM		1	0.711**	0.172*	0.320**	0.660**	0.694**	0.618**	0.457**	0.619**	-0.0686	0.1131	0.357**	0.304**	0.457**
PH			1	-0.184*	-0.0952	0.807**	0.759**	0.561**	0.461**	0.570**	-0.1489	0.389**	0.467**	0.1575	0.429**
NTP				1	1.097	-0.216*	0.0023	0.0724	-0.0798	0.0507	0.119	-0.183*	0.302**	0.447**	0.543**
NPP					1	-0.0763	0.1306	0.1117	-0.1152	0.0794	0.207*	-0.11	0.443**	0.549**	0.695**
PL						1	0.644**	0.355**	0.212*	0.348**	0.0132	0.440**	0.341**	0.219*	0.381**
PW							1	0.858**	0.668**	0.865**	-0.204*	0.422**	0.418**	0.450**	0.636**
NFGP								1	0.685**	0.993**	-0.1208	-0.0725	0.305**	0.448**	0.553**
NCGP									1	0.767**	-0.774**	0.0503	0.1504	0.358**	0.368**
TGP										1	-0.233*	-0.0556	0.293**	0.453**	0.547**
SF											1	-0.1512	0.025	-0.1534	-0.0855
TW												1	0.185*	0.1312	0.234*
SY													1	-0.0477	0.638**
HI														1	0.722**
GYP															1

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g).

findings (38, 41, 43). The present results for TW were in proximity with earlier reports (36, 40, 41). Likewise, the findings pertaining to SY and HI were also supported by earlier studies (38, 41).

Path coefficient analysis

The path coefficient analysis (Table 6) partitioned the correlation coefficients into direct and indirect effects, providing a clearer understanding of trait contributions to grain yield. Among the traits, TGP exerted a very high positive direct effect (3.8744), followed by NTP, PL, TW, SY and HI, all of which also exhibited significant positive correlations with grain yield. Similar findings have been reported for PL (44, 45), TGP (45) and TW, SY and HI (46). The positive direct effects of these traits indicate that selection based on these characters would result in a direct and effective improvement in grain yield.

In contrast, PW and NFGP, despite showing significant positive correlations with grain yield, exhibited very high negative direct effects, suggesting that their positive association is mainly due to indirect effects through other component traits. A similar trend for NFGP was reported earlier (45).

Interestingly, spikelet fertility showed a high positive direct effect, but a negative non-significant correlation with grain yield, indicating that its favourable direct effect is counterbalanced by

negative indirect effects. These results are consistent with the earlier findings (47). Traits such as DFF, DTM, PH and NPP exhibited negligible to low direct effects, indicating their limited direct role in yield determination.

The residual effect (0.168) indicated that most of the variability in grain yield was explained by the traits included in the analysis, with only a small proportion (16.8 %) attributed to other factors. Overall, traits with high positive direct effects and significant correlations, particularly the TGP and HI, should be considered in selection programmes for effective yield improvement.

Selection indices

The selection index technique provides information on yield components and thus aids indirect selection for yield improvement. Classical selection indices based on Fisher's discriminant function were calculated for 68 advanced breeding lines of rice and the results were presented in Table 7.

Traits with high PCV and GCV, broad sense heritability and GAM, along with a significant positive correlation and direct effects on GYP, were chosen for building the selection indices. These traits include the NTP, PL, TW, SY, HI and GYP. The economic weight assigned to each trait was determined as the reciprocal of its mean value.

Table 6. Genotypic path coefficients for yield and its contributing traits

Trait	DFF	DTM	PH	NTP	NPP	PL	PW	NFGP	NCGP	TGP	SF	TW	SY	HI	GYP
DFF	-0.0107	-0.0906	-0.062	0.0425	-0.0293	0.2274	-0.7081	-1.5809	-0.0463	2.267	-0.0199	0.0549	0.1946	0.1378	0.376**
DTM	-0.0104	-0.0936	-0.0624	0.0594	-0.0363	0.225	-0.7332	-1.676	-0.048	2.3971	-0.0207	0.0569	0.2217	0.1772	0.457**
PH	-0.0076	-0.0665	-0.0877	-0.0633	0.0108	0.2753	-0.8016	-1.5221	-0.0485	2.2066	-0.0448	0.1958	0.2904	0.0919	0.429**
NTP	-0.0013	-0.0161	0.0161	0.3448	-0.1247	-0.0736	-0.0024	-0.1964	0.0084	0.1962	0.0358	-0.092	0.1876	0.2608	0.543**
NPP	-0.0028	-0.0299	0.0084	0.3782	-0.1137	-0.026	-0.1379	-0.303	0.0121	0.3078	0.0623	-0.0554	0.2753	0.32	0.695**
PL	-0.0071	-0.0617	-0.0708	-0.0744	0.0087	0.341	-0.6805	-0.9641	-0.0223	1.3469	0.004	0.2213	0.212	0.1277	0.381**
PW	-0.0072	-0.0649	-0.0666	0.0008	-0.0148	0.2197	-1.0563	-2.3278	-0.0702	3.3503	-0.0615	0.2123	0.2595	0.2625	0.636**
NFGP	-0.0062	-0.0578	-0.0492	0.025	-0.0127	0.1211	-0.906	-2.714	-0.072	3.8467	-0.0364	-0.0365	0.1893	0.2612	0.553**
NCGP	-0.0047	-0.0427	-0.0405	-0.0275	0.0131	0.0724	-0.705	-1.8581	-0.1051	2.971	-0.2328	0.0253	0.0935	0.2089	0.368**
TGP	-0.0063	-0.0579	-0.05	0.0175	-0.009	0.1186	-0.9134	-2.6945	-0.0806	3.8744	-0.0702	-0.028	0.182	0.2643	0.547**
SF	0.0007	0.0064	0.0131	0.041	-0.0235	0.0045	0.2157	0.3278	0.0813	-0.9035	0.3009	-0.0761	0.0156	-0.0894	-0.0855
TW	-0.0012	-0.0106	-0.0341	-0.063	0.0125	0.1499	-0.4455	0.1967	-0.0053	-0.2154	-0.0455	0.5034	0.115	0.0765	0.234*
SY	-0.0033	-0.0334	-0.041	0.104	-0.0503	0.1163	-0.441	-0.8265	-0.0158	1.1345	0.0075	0.0931	0.6217	-0.0278	0.638**
HI	-0.0025	-0.0284	-0.0138	0.1542	-0.0624	0.0747	-0.4756	-1.2159	-0.0377	1.7561	-0.0462	0.066	-0.0297	0.5831	0.722**

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g).

Table 7. Selection indices for yield and its component traits and their relative efficiency in advanced breeding lines of rice

Index	Genetic gain at 5% selection intensity	Relative efficiency (%)	Ranks
GYP	24.90	100.00	-
NTP + PL	25.99	518.07	46
NTP + TW	24.12	361.89	52
NTP + SY	58.33	293.10	54
NTP + HI	56.07	663.57	33
NTP + GYP	60.79	323.51	53
PL + TW	56.08	841.13	14
PL + SY	117.43	590.04	42
PL + HI	116.26	1375.73	8
PL + GYP	117.22	623.81	39
TW + SY	81.14	407.71	50
TW + HI	80.52	952.75	13
TW + GYP	81.01	431.12	49
SY + HI	43.73	517.49	47
SY + GYP	54.96	292.47	55
HI + GYP	97.63	519.58	45
NTP + PL + TW	65.00	974.84	12
NTP + PL + SY	140.26	704.06	30
NTP + PL + HI	137.94	1632.31	4
NTP + PL + GYP	141.44	752.71	24
NTP + TW + SY	123.49	619.50	41
NTP + TW + HI	104.34	1234.26	10
NTP + TW + GYP	123.27	656.04	34
NTP + SY + HI	63.02	745.69	25
NTP + SY + GYP	72.25	384.51	51
NTP + HI + GYP	102.35	544.72	44
PL + TW + SY	123.49	620.50	40
PL + TW + HI	122.32	1447.46	6
PL + TW + GYP	123.27	656.04	34
PL + SY + HI	117.03	1384.83	7
PL + SY + GYP	119.81	637.59	37
PL + HI + GYP	139.86	739.86	26
TW + SY + HI	86.50	1023.60	11
TW + SY + GYP	91.67	487.86	48
TW + HI + GYP	119.00	633.31	38
SY + HI + GYP	121.75	647.91	36
NTP + PL + TW + SY	146.87	738.02	27
NTP + PL + TW + HI	144.56	1710.64	2
NTP + PL + TW + GYP	148.04	787.84	20
NTP + PL + SY + HI	138.55	1639.52	3
NTP + PL + SY + GYP	141.47	752.85	23
NTP + PL + HI + GYP	149.71	796.69	18
NTP + TW + HI + GYP	126.71	674.33	31
NTP + TW + SY + GYP	109.96	585.16	43
NTP + TW + SY + HI	104.92	1241.59	9
NTP + SY + HI + GYP	63.02	648.28	35
PL + TW + SY + HI	122.84	1453.53	5
PL + TW + SY + GYP	125.33	666.97	32
PL + TW + HI + GYP	142.99	760.96	22
PL + SY + HI + GYP	145.83	776.07	21
TW + SY + HI + GYP	133.86	712.36	29
NTP + PL + TW + SY + HI	145.54	1722.17	1
NTP + PL + TW + SY + GYP	141.85	788.60	19
NTP + PL + TW + HI + GYP	154.74	823.50	16
NTP + PL + SY + HI + GYP	151.01	803.63	17
NTP + TW + SY + HI + GYP	135.32	720.13	28
PL + TW + SY + HI + GYP	148.92	792.53	19
NTP + PL + TW + SY + HI + GYP	155.41	827.07	15

NTP: number of tillers plant⁻¹; PL: panicle length (cm); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g).

The relative efficiency of different selection indices is presented in Table 7. Among the combinations evaluated, the highest relative efficiency (1722.17 %) was recorded for a five-character index comprising the NTP, PL, TW, SY and HI. Closely comparable efficiencies were observed for four-character combinations, particularly the index including NTP, PL, TW and HI (1710.64 %), followed by the combination of NTP, PL, SY and HI (1639.52 %).

The marginal difference in efficiency between the five character and four character indices suggests that inclusion of additional traits beyond four may not substantially enhance selection efficiency. This is further supported by the high performance of simpler combinations such as NTP, PL and HI (1632.31 %), as well as PL with HI alone (1375.73 %). These results highlight the consistent importance of PL and HI across multiple efficient indices.

Considering both efficiency and practicality, the four character index comprising NTP, PL, TW and HI are identified as the most suitable for indirect selection, as it offers near-maximum efficiency with fewer traits. The superior genotypes identified based on this index are presented in Table 8.

Table 8. Superior genotypes of rice based on index scores for the character combination NTP + PL + TW + HI

Ranks	Genotypes	Index score	Ranks	Genotypes	Index score
1	NLR 34417	943.73	11	NLR 33801	858.78
2	NLR 9674	900.89	12	NLR 33720	854.32
3	NLR 3390	886.94	13	NLR 3117	852.02
4	NLR 20002	886.00	14	NLR 33166	847.65
5	NLR 3188	884.50	15	NLR 33892	846.83
6	NLR 3128	878.76	16	NLR 3197	846.35
7	NLR 33648	876.64	17	NLR 3261	842.28
8	NLR 3187	876.17	18	NLR 20016	832.41
9	NLR 3111	873.16	19	NLR 20022	830.49
10	NLR 33718	871.44	20	NLR 3389	824.14

NTP: number of tillers plant⁻¹; PL: panicle length (cm); TW: test weight (g); HI: harvest index (%).

The effectiveness of selection indices observed in the present study agrees with earlier reports. Similar approaches using multivariate selection indices for yield improvement have been demonstrated (48), while studies under both drought and non-stress conditions also emphasised the utility of such indices in identifying superior rice genotypes (49, 50). Furthermore, the effectiveness of combining multiple yield components for enhanced selection efficiency has been reported in indigenous rice germplasm (51).

Genetic divergence D² statistics

Clustering pattern

The clustering pattern based on D² analysis (Table 9) grouped the 68 genotypes into 12 distinct clusters, indicating substantial genetic diversity. Cluster I was the largest, comprising 42 genotypes, while clusters X, XI and XII were solitary, reflecting their unique genetic makeup. The clustering pattern suggests that genotypes within a cluster are genetically similar to those in different clusters.

Intra and inter cluster distances

The intra- and inter-cluster distances (Fig. 2, Table 10) further emphasised the extent of diversity. Among the clusters, cluster IX exhibited the highest intra-cluster distance, indicating greater heterogeneity among its genotypes, whereas clusters X, XI and XII showed zero intra-cluster distance due to their solitary nature. The maximum inter-cluster distance was observed between clusters IX and XII, followed by other widely separated cluster pairs, indicating a high degree of genetic divergence. In contrast, the minimum inter-cluster distance between clusters III and VII suggested close genetic similarity. These results indicate that selection of parents from highly divergent clusters, particularly IX and XII, would be advantageous for hybridisation to generate superior recombinants.

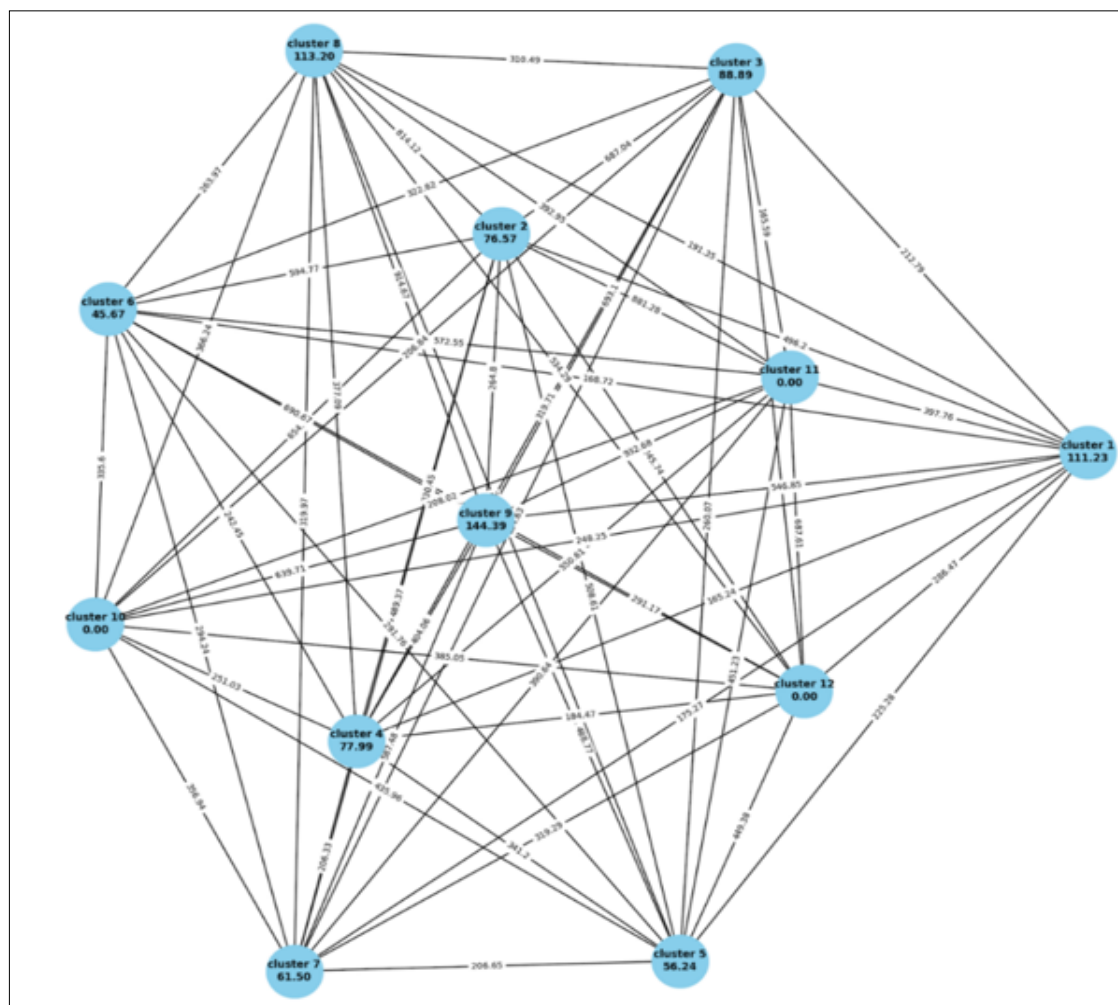
Table 9. Clustering of 68 genotypes into 12 clusters using Tochers method

Cluster	Total genotypes	Genetic material
Cluster I	42	NLR 20084, NLR 33801, NLR 20023, NLR 34417, NLR 20106, NLR 20123, NLR 20128, NLR 20127, NLR 20136, NLR 20146, NLR 3187, NLR 20022, NLR 3199, NLR 3380, NLR 3390, NLR 3190, NLR 3291, NLR 3252, NLR 3264, NLR 20156, NLR 20002, NLR 3328, NLR 3261, NLR 3379, NLR 3385, NLR 3374, NLR 3424, NLR 3135, NLR 3382, NLR 3339, NLR 33718, NLR 3338, NLR 3465, NLR 3461, NLR 3433, NLR 3418, NLR 40058, MTU 1010, NLR 9674, NLR 33648, NLR 33166, NLR 33720
Cluster II	4	NLR 3256, NLR 40058, NLR 40021, NLR 3367
Cluster III	5	NLR 33891, NLR 20131, NLR 20150, NLR 20083, NLR 3130
Cluster IV	3	NLR 33810, NLR 3187, NLR 3188
Cluster V	3	NLR 3111, NLR 3128, NLR 3186
Cluster VI	2	NLR 3302, BPT 5204
Cluster VII	2	NLR 20016, NLR 3336
Cluster VIII	2	NLR 3117, NLR 33892
Cluster IX	2	NLR 3083, NLR 3411
Cluster X	1	NLR 20104
Cluster XI	1	NLR 33891
Cluster XII	1	NLR 3237

Table 10. Average intra and inter cluster distances among rice genotypes

Cluster No.	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI	Cluster VII	Cluster VIII	Cluster IX	Cluster X	Cluster XI	Cluster XII
Cluster I	111.23	496.20	212.79	165.24	225.28	168.72	175.27	191.35	546.85	248.25	397.76	286.47
Cluster II		76.57	687.04	330.45	508.61	594.77	489.37	814.12	264.80	654.40	881.28	245.74
Cluster III			88.89	319.71	260.07	322.62	154.63	310.49	693.10	206.84	165.59	452.70
Cluster IV				77.99	341.20	242.45	206.33	377.09	404.06	251.03	550.61	184.47
Cluster V					56.24	291.76	206.65	277.33	488.77	435.96	451.23	449.38
Cluster VI						45.67	294.24	263.97	690.67	335.60	572.55	404.40
Cluster VII							61.50	319.97	587.48	356.94	390.64	319.29
Cluster VIII								113.20	914.67	366.24	392.95	534.29
Cluster IX									144.39	639.71	932.68	291.17
Cluster X										0.00	208.02	385.05
Cluster XI											0.00	687.61
Cluster XII												0.00

Diagonal values represent intra cluster distances whereas, the values above diagonal represent inter-cluster distance.

**Fig. 2.** Cluster diagram indicating intra and inter cluster distances.

Cluster means

The cluster means (Table 11) revealed considerable variation among clusters for all traits. Notably, cluster IV exhibited superior mean performance for key yield-related traits, including NFGP, TGP, HI and GYP, indicating its potential for yield improvement. Cluster VII recorded the highest NTP and NPP, while cluster II showed the highest PL. Cluster XII was characterised by late flowering and maturity along with taller plant statures. Such variation among clusters provides an opportunity to select parents with desirable and complementary traits.

Percentage contribution of various characters towards genetic divergence

The contribution of different characters towards genetic divergence (Fig. 3, Table 12) indicated that DTM contributed the highest (29.13 %), followed by GYP, PH and TW, suggesting their major role in genetic differentiation. In contrast, the NPP and the NFGP contributed negligible to divergence.

The presence of wide genetic diversity among the genotypes agrees with earlier reports, where rice germplasm was grouped into multiple clusters, such as seven clusters in Northwestern Himalayan landraces (52), 8 clusters in indigenous rice varieties of Assam (53)

and eight clusters in rice germplasm lines (54). These findings further support the effectiveness of D² analysis in assessing genetic divergence and identifying diverse parents for breeding programmes.

Principal component analysis

Principal component analysis was performed for 15 quantitative traits in 68 genotypes to assess the extent and pattern of variability (Fig. 4, Table 13). The analysis revealed six principal components (PCs) with eigenvalues greater than one, collectively accounting for 91.5 % of the total variation, indicating that most of the variability among genotypes is effectively captured by these components. Among them, PC1 contributed the highest variation (40.3 %), followed by PC2 (15.8 %) and PC3 (11.9 %), suggesting that these components play a major role in differentiating genotypes.

The trait contributions to each PC were interpreted using loading values (Fig. 5, Table 14). In PC1, most traits exhibited negative loadings, except spikelet fertility, indicating an inverse association of these traits with the principal axis of variation. In contrast, PC2 was largely influenced by the NTP and the NPP, along with notable contributions from grain yield and SY, highlighting their importance in yield determination.

Table 11. Mean values of 15 quantitative traits in 12 cluster of 68 genotypes

Trait	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
DFF	120.65	126.50	103.00	117.00	119.40	91.00	127.50	115.33	118.75	93.25	118.50	130.00
DTM	153.88	157.50	136.00	152.50	156.00	122.00	157.50	148.33	157.00	125.00	150.33	162.00
PH	108.03	108.00	83.00	121.20	113.92	80.85	95.60	81.97	89.50	74.35	107.07	125.10
NTP	15.82	13.25	16.00	14.80	19.38	14.45	24.90	20.57	18.30	14.18	14.37	13.95
NPP	13.18	13.75	13.00	12.60	16.76	11.85	20.30	17.63	14.90	10.60	12.43	12.55
PL	22.74	28.40	17.20	24.60	24.44	19.30	20.80	19.87	21.85	17.55	24.27	23.65
PW	5.69	4.25	3.00	8.75	6.81	2.45	6.10	3.85	4.93	2.89	6.82	8.68
NFGP	246.72	127.50	180.00	400.50	275.74	92.25	287.50	192.80	205.60	137.03	205.17	351.85
NCGP	27.48	9.00	20.00	53.20	27.10	18.75	29.30	20.00	24.05	9.65	29.93	54.30
TGP	274.20	136.50	200.00	453.70	302.84	111.00	316.80	212.80	229.65	146.68	235.10	406.15
SF	90.05	93.40	90.10	88.20	90.99	83.50	90.60	90.72	89.75	93.29	87.62	86.43
TW	19.89	19.10	14.80	20.10	22.01	21.03	16.25	16.30	22.50	18.80	30.13	21.05
SY	46.79	49.88	69.76	47.75	56.18	40.10	62.90	28.78	49.75	29.75	47.92	53.63
HI	42.69	36.85	34.55	60.85	52.16	39.43	52.25	51.03	43.55	35.00	45.43	41.83
GYP	34.68	29.05	36.90	74.10	60.60	25.95	68.30	30.27	38.08	16.11	39.38	37.78

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g).

Table 12. Percentage contribution of various characters towards total divergence

Traits	Singh statistic	Percent contribution	Cumulative percentage
DTM	710701.66	29.13	29.13
GYP	388993.55	15.94	45.07
PH	252585.97	10.35	55.42
TW	183283.12	7.51	62.93
TGP	152536.23	6.25	69.18
SY	140840.27	5.77	74.95
PW	100616.54	4.12	79.07
PL	96904.21	3.97	83.05
SF	95396.90	3.91	86.96
NCGP	84949.22	3.48	90.44
DFF	80837.86	3.31	93.75
HI	74563.11	3.06	96.80
NTP	66119.06	2.71	99.51
NPP	7920.83	0.32	99.84
NFGP	3929.61	0.16	100.00

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹ (g).

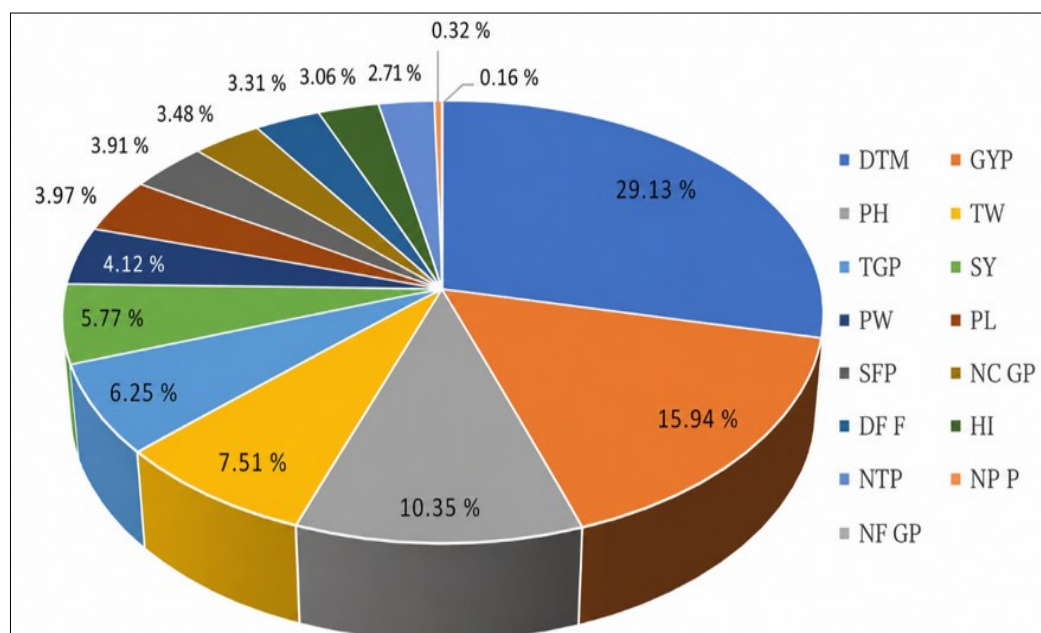


Fig. 3. Percent contribution of each trait towards total divergence.

DTM: days to maturity; GYP: grain yield plant⁻¹(g); PH: plant height (cm); (%); TW: test weight (g); TGP: total number of grains panicle⁻¹; SY: straw yield (g); PW: panicle weight (g); PL: panicle length (cm); SFP: spikelet fertility percentage; NCGP: number of chaffy grains panicle⁻¹; DFF: days to 50 % flowering; HI: harvest index (%); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; NFGP: number of filled grains panicle⁻¹.

Table 13. Eigen values and proportion of variability explained by the principal components

Principal components	Eigen vector values	Proportion of variability explained	Cumulative proportion
PC 1	6.083	0.403	0.403
PC 2	2.388	0.158	0.561
PC 3	1.804	0.119	0.68
PC 4	1.332	0.088	0.768
PC 5	1.178	0.078	0.846
PC 6	1.043	0.069	0.915
PC 7	0.446	0.03	0.945
PC 8	0.396	0.026	0.971
PC 9	0.21	0.014	0.985
PC 10	0.098	0.006	0.991
PC 11	0.056	0.004	0.995
PC 12	0.03	0.002	0.997
PC 13	0.028	0.002	0.999
PC 14	0.02	0.001	1
PC 15	0	0	1

Table 14. Eigen vectors loading values for principal components

Variables	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13	PC 14	PC 15
DFF	-0.318	0.004	0.180	-0.059	0.311	-0.368	0.249	-0.288	0.046	0.009	0.039	-0.377	-0.566	0.138	0
DTM	-0.332	0.024	0.156	-0.059	0.280	-0.335	0.2	-0.338	0.018	-0.031	-0.087	0.402	0.585	-0.068	0
PH	-0.312	-0.163	0.290	0.074	0.049	-0.016	-0.234	0.275	-0.796	-0.136	0.028	0.009	-0.017	0.009	0
NTP	-0.074	0.584	-0.122	0.109	0.128	-0.081	0.121	0.282	-0.176	0.676	-0.141	-0.024	0.026	-0.019	0
NPP	-0.094	0.576	-0.050	0.154	0.109	-0.11	0.043	0.355	0.153	-0.66	0.148	0.018	0	0.017	0
PL	-0.253	-0.082	0.403	0.099	-0.086	-0.227	-0.571	0.275	0.5	0.181	-0.116	-0.007	0.008	-0.003	0
PW	-0.369	-0.111	0.001	-0.077	-0.186	0.141	0.239	0.175	0.128	0.2	0.798	-0.005	0.101	-0.026	0
NFGP	-0.334	-0.084	-0.225	-0.355	-0.067	0.131	0.099	0.204	0.073	-0.045	-0.287	0.225	-0.152	0.175	-0.661
NCGP	-0.266	-0.234	-0.393	0.271	0.198	0.057	-0.016	0.073	0.063	-0.057	-0.152	-0.541	0.271	-0.432	-0.12
TGP	-0.341	-0.113	-0.265	-0.273	-0.028	0.126	0.086	0.194	0.076	-0.049	-0.281	0.113	-0.092	0.086	0.741
SF	0.077	0.208	0.343	-0.644	-0.27	-0.026	0.115	0.051	-0.045	-0.067	-0.116	-0.396	0.197	-0.336	0
TW	-0.099	-0.076	0.363	0.486	-0.432	0.130	0.55	0.075	0.063	-0.039	-0.314	0.001	-0.015	0.002	0
SY	-0.183	0.210	0.259	0.020	0.289	0.651	-0.109	-0.239	0.079	0.007	-0.014	0.198	-0.235	-0.421	0
HI	-0.207	0.166	-0.291	0.082	-0.547	-0.317	-0.185	-0.305	-0.121	-0.025	0.032	0.208	-0.257	-0.43	0
GYP	-0.293	0.293	-0.05	0.059	-0.254	0.289	-0.253	-0.414	-0.047	-0.028	-0.034	-0.314	0.247	0.524	0

DFF: days to 50 % flowering; DTM: days to maturity; PH: plant height (cm); NTP: number of tillers plant⁻¹; NPP: number of panicles plant⁻¹; PL: panicle length (cm); PW: panicle weight (g); NFGP: number of filled grains panicle⁻¹; NCGP: number of chaffy grains panicle⁻¹; TGP: total number of grains panicle⁻¹; SF: spikelet fertility (%); TW: test weight (g); SY: straw yield (g); HI: harvest index (%); GYP: grain yield plant⁻¹(g).

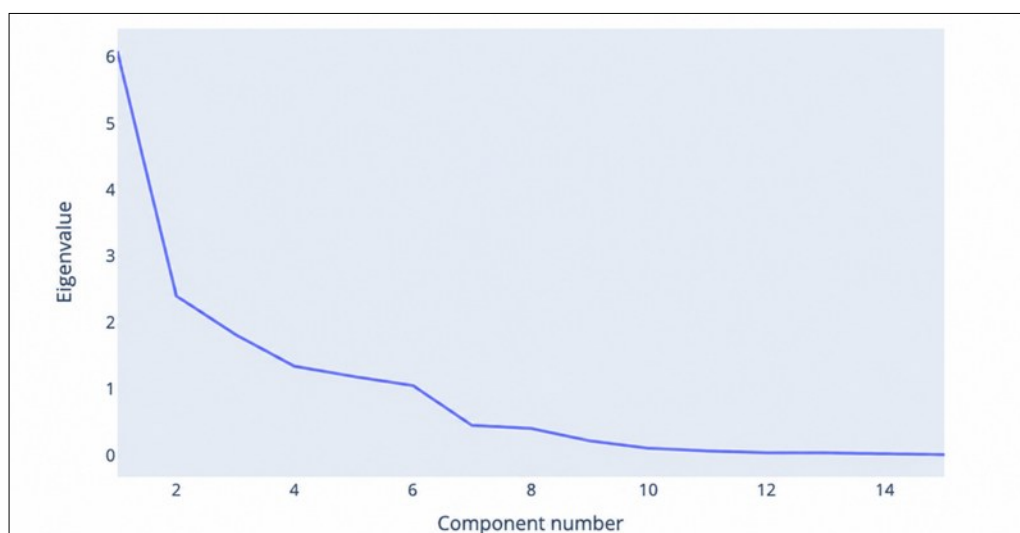


Fig. 4. Scree plot of principal components.

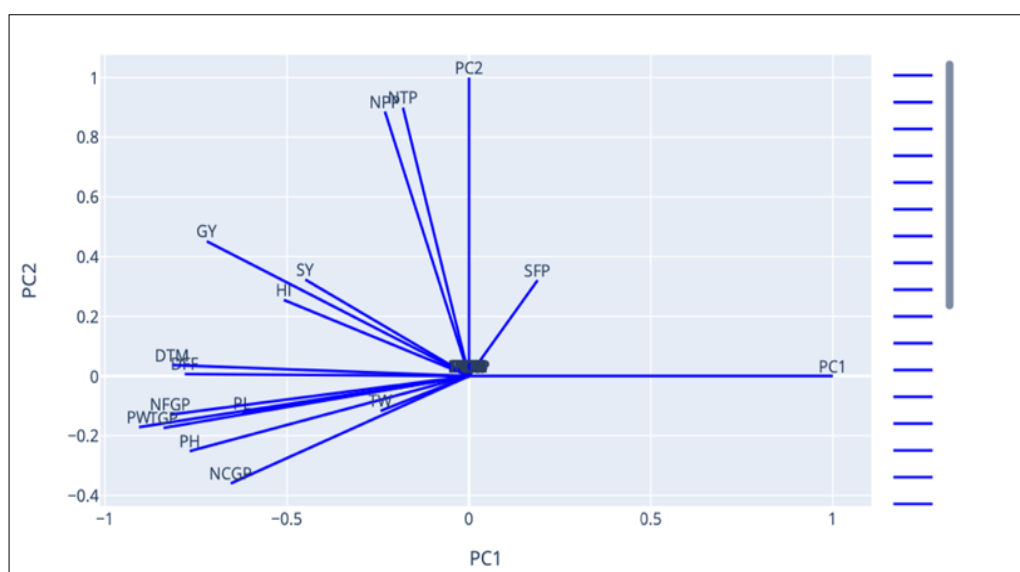


Fig. 5. Variable loading plot PC1 vs PC2.

Principal component 3 was mainly associated with PL and TW, suggesting their role in defining grain and panicle characteristics. Principal component 4 was predominantly influenced by TW and NCGP, while PC5 was characterised by DFF and DTM, indicating the contribution of phenological traits. Principal component 6 was strongly associated with SY and GYP, emphasising their significance in overall productivity.

The presence of positive loadings for key yield-contributing traits across PCs indicates their strong influence on overall genetic variability, whereas negative loadings suggest inverse relationships among traits. These results are consistent with earlier research, where similar patterns of variability and trait contributions were reported in rice germplasm (53). Comparable results were also observed in recombinant inbred line populations for submergence tolerance (55, 56) and in indigenous rice germplasm, where a limited number of PCs accounted for most of the variation (51).

Diversity pattern based on the unweighted pair group method with arithmetic mean (UPGMA) approach

The hierarchical relationship among the 68 genotypes based on 15 quantitative traits was analysed using the UPGMA method and the resulting dendrogram is presented in Fig. 6. The clustering pattern grouped the genotypes into nine distinct clusters, indicating

considerable morphological diversity among the breeding lines. Genotypes grouped within the same cluster, viz, NLR 20084, NLR 20104, NLR 20022 and NLR 20002, exhibited close genetic similarity, reflecting shared agronomic characteristics.

The dendrogram further demonstrates that genotypes connected through shorter branches (e.g., NLR 3188 and NLR 3197) were more closely related, while those linked by longer branches (e.g., NLR 33891 and NLR 20123) were highly divergent. Such divergent genotypes can be used as potential parents in hybridisation programmes to generate superior recombinants. Additionally, genotypes sharing lower node values (e.g., NLR 20136 and NLR 33810) (node 70) indicated higher genetic similarity compared to other genotypes.

The node values in the dendrogram symbolise the similarity or distance measures among genotypes, where lower values indicate closer relationships. The overall clustering pattern highlights the presence of substantial diversity, which can be effectively exploited in breeding programmes. These findings agree with earlier reports, where similar clustering patterns were observed in rice germplasm, including the grouping of Bangladesh rice genotypes (57), classification of rice landraces from hilly regions of northeastern India (58) and clustering of rice landraces based on cultivation ecosystems (59).

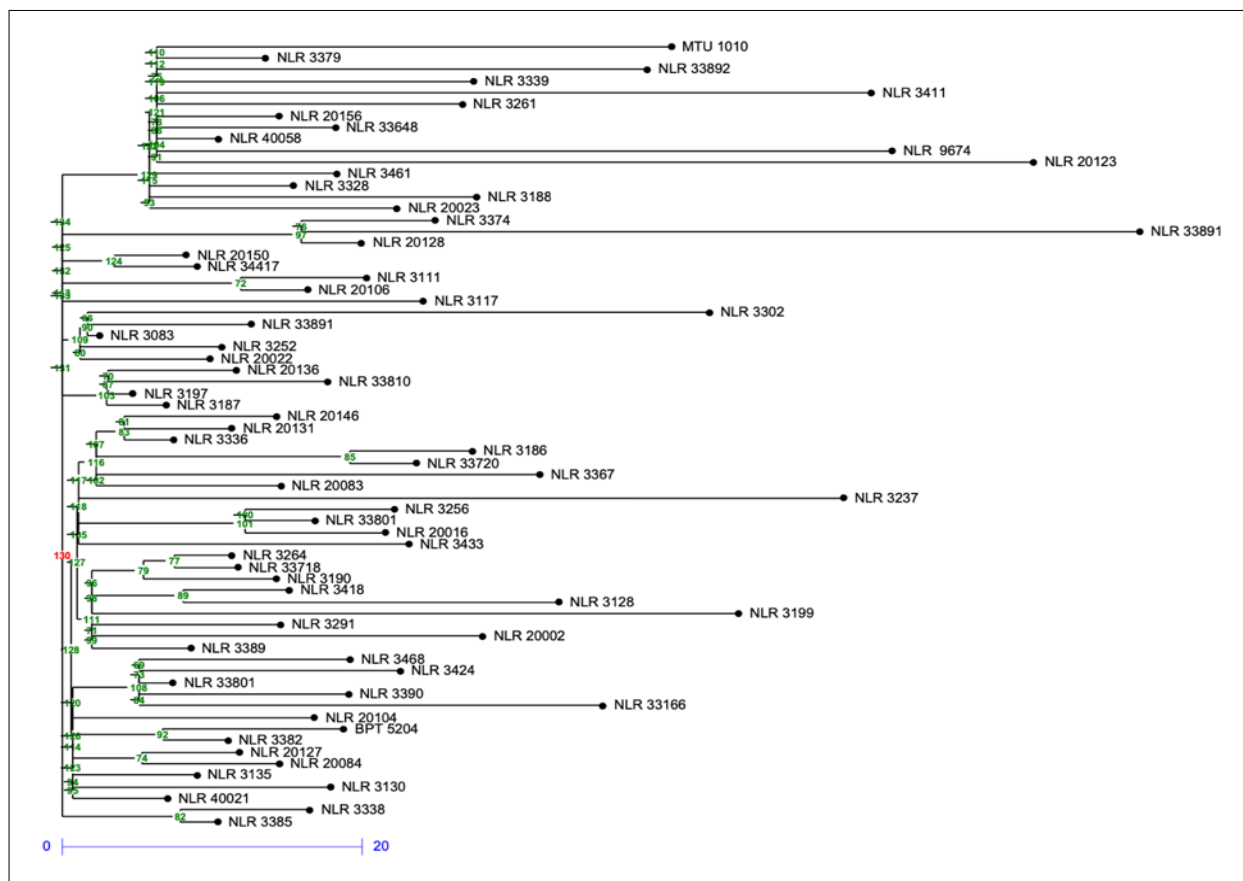


Fig. 6. Grouping of 68 advanced breeding lines of rice based on unweight pair group method with arithmetic mean (UPGMA) hierarchical clustering method.

Conclusion

The study revealed that the genotypes, viz., NLR 3128, NLR 20002, NLR 20016, NLR 33891 and NLR 34417 recorded significantly higher grain yield when compared to the checks. High heritability in broad sense coupled with high genetic advance as percentage of mean genetic advance as percentage of mean (GAM) was recorded for the traits viz., plant height (PH), number of tillers per plant (NTP), number of panicles per plant (NPP), panicle length (PL), panicle weight (PW), number of filled grains per panicle (NFGP), number of chaffy grains per panicle (NCGP), total number of grains per panicle (TGP), test weight (TW), straw yield (SY), harvest index (HI) and grain yield per plant (GYP) indicating the presence of additive gene action and that selection may be rewarding for these traits. The traits viz., NTP, PL, TW, SY and HI recorded positive significant correlation along with positive direct effects on GYP and the index including $NTP + PL + TW + HI$ recorded highest relative efficiency. The presence of significant variability among the genotypes is further supported by the formation of twelve clusters and the promising genotypes may be used as parents in future hybridisation programmes. Further, these findings can be utilised to develop high-yielding cultivars through targeted selection and crossing genetically diverse parents to exploit heterosis and generate superior recombinants.

Acknowledgements

The author extends heartfelt gratitude to the Agriculture Research Station, Nellore and the Department of Genetics and Plant Breeding Laboratory at Sri Venkateswara Agricultural College, Tirupati, for providing the essential facilities required for this research. The author

also expresses sincere appreciation to Bommireddy Reddiyami and Sai Venkat K.S. for their invaluable guidance and unwavering support throughout the course of this work.

Authors' contributions

MSP, CS and SNU conceived the study, designed the methodology and revised the manuscript. UC supervised the study, performed the experiments, conducted data collection and statistical analyses, interpreted the results and prepared the original draft. PP performed the UPGMA software analysis using DARwin software (Version 6) and revised the manuscript. CESP also revised the manuscript. Funding for this research was acquired by CS. All authors 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

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Microsoft 365 Copilot (a GPT-5-based version) and Grammarly to improve the grammar, clarity and readability of the manuscript. The tools were solely used for language improvement and did not contribute to the generation of scientific content in the manuscript. Following its use, authors have carefully reviewed, edited and validated the content and take full responsibility for integrity of the publication.

References

- Food and Agriculture Organization. FAOSTAT. Rome: Food and Agriculture Organization; 2024. <https://www.fao.org/faostat>
- Fukagawa NK, Ziska LH. Rice: Importance for global nutrition. *Journal of Nutritional Science and Vitaminology*. 2019;65(Suppl)–3. <https://doi.org/10.3177/jnsv.65.S2>
- Department of Agriculture and Farmers Welfare. Agricultural statistics at a glance. New Delhi: Ministry of Agriculture and Farmers Welfare; 2022. <http://desagri.gov.in>
- Indiastat. 2024. <https://www.indiastat.com>
- Smith HF. A discriminant function for plant selection. *Annals of Eugenics*. 1936;7(3):240–50. <https://doi.org/10.1111/j.1469-1809.1936.tb02143.x>
- Jadhav RA, Mehtre SP, Patil DK, Gite VK. Multivariate analysis using D² and principal component analysis in mung bean (*Vigna radiata* (L.) Wilczek) for study of genetic diversity. *Legume Research*. 2021;1(8). <https://doi.org/10.18805/LR-4508>
- Nanthakumar S, Kuralarasu C, Gopikrishnan A. D² analysis for assessing genetic diversity in okra (*Abelmoschus esculentus* (L.) Moench). *Electronic Journal of Plant Breeding*. 2021;12(4):1249–53. <https://doi.org/10.37992/2021.1204.171>
- Burton GW. Quantitative inheritance in grasses. *Proceedings of the 6th International Grassland Congress*. 1952;1:277–83.
- Allard RW. *Principles of plant breeding*. New York: John Wiley & Sons; 1999.
- Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*. 1955;47(7):314–8. <https://doi.org/10.2134/agronj1955.00021962004700070009x>
- Robinson HF, Comstock RE, Harvey PH. Genotypic and phenotypic correlations in corn and their implications in selection. *Agronomy Journal*. 1951;43(6):282–7. <https://doi.org/10.2134/agronj1951.00021962004300060007x>
- Dewey DR, Lu K. A correlation and path-coefficient analysis of components of crested wheatgrass seed production. *Agronomy Journal*. 1959;51(9):515–8. <https://doi.org/10.2134/agronj1959.00021962005100090002x>
- McLachlan GJ. *Discriminant analysis and statistical pattern recognition*. New York: John Wiley & Sons; 2004.
- Mahalanobis PC. On the generalized distance in statistics. *Sankhyā: The Indian Journal of Statistics, Series A*; 2018. p. 80–7. <https://doi.org/10.1007/s13171-019-00164-5>
- Rao CR. *Advanced statistical methods in biometric research*. New York: John Wiley & Sons; 1952.
- Popat R, Patel H, Popat P. *Agri Analyze*. 2024. <https://agrianalyze.com>
- Singh RK, Chaudhary BD. *Biometrical methods in quantitative genetic analysis*. New Delhi: Kalyani Publishers; 1977.
- Perrier X, Jacquemoud-Collet JP. *DARwin Software*. 2006. <http://darwin.cirad.fr>
- Lakshmi M, Shanmuganathan M, Jeyaprakash P, Ramesh T. Genetic variability and diversity analysis in selected rice (*Oryza sativa* L.) varieties. *Electronic Journal of Plant Breeding*. 2022;13(3):959–66. <https://doi.org/10.37992/2022.1303.124>
- Karim D, Sarkar U, Siddique MNA, Khaleque Miah MA, Hasnat MZ. Variability and genetic parameter analysis in aromatic rice. *Int J Sustain Crop Prod*. 2007;2(5):15–18.
- Rashmi D, Saha S, Loitongbam B, Singh S, Singh PK. Genetic variability study for yield and yield components in rice (*Oryza sativa* L.). *International Journal of Agriculture, Environment and Biotechnology*. 2017;10(2):171. <https://doi.org/10.5958/2230-732X.2017.00020.1>
- Panda D, Sahu N, Behera PK, Lenka K. Genetic variability of panicle architecture in indigenous rice landraces of Koraput region of Eastern Ghats of India for crop improvement. *Physiology and Molecular Biology of Plants*. 2020;26(10):1961–71. <https://doi.org/10.1007/s12298-020-00871-6>
- Rashid MM, Nuruzzaman M, Hassan L, Begum SN. Genetic variability analysis for various yield attributing traits in rice genotypes. *Journal of the Bangladesh Agricultural University*. 2017;15(1):15–9. <https://doi.org/10.3329/jbau.v15i1.33525>
- Babu VR, Shreya K, Dangi KS, Usharani G, Nagesh P. Genetic variability studies for qualitative and quantitative traits in popular rice (*Oryza sativa* L.) hybrids of India. *International Journal of Scientific and Research Publications*. 2012;2(6):1–5.
- Sameera SK, Rajesh AP, Jayalakshmi V, Nirmala PJ, Srinivas T. Genetic variability studies for yield and yield components in rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*. 2015;6(1):269–73.
- Chacko A, Jayalekshmy VG, Shahiba AM. Studies on PCV, GCV, heritability and genetic advance in rice genotypes for yield and yield components. *International Journal of Plant & Soil Science*. 2023;35(16):324–30. <https://doi.org/10.9734/ijpss/2023/v35i163266>
- Demeke B, Dejene T, Abebe D. Genetic variability, heritability and genetic advance of morphological, yield related and quality traits in upland rice (*Oryza sativa* L.) genotypes at Pawe, northwestern Ethiopia. *Cogent Food & Agriculture*. 2023;9(1):2157099. <https://doi.org/10.1080/23311932.2022.2157099>
- Allam CR, Jaiswal HK, Qamar A, Challa Venkateshwarlu CV, Reddy YS. Variability, heritability and genetic advance studies in some indigenous genotypes of basmati rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*. 2015;6(2):506–11.
- Umesh U, Jaiswal HK, Sravan T, Waza SA, Bhardwaj R. Estimation of genetic variability, heritability and genetic advance for yield and quality traits in some indigenous basmati rice (*Oryza sativa* L.) genotypes. *International Journal of Farm Sciences*. 2015;5(4):32–40.
- Limani PL, Gangani MK, Pandya MM. Genetic variability, heritability and genetic advance in rice (*Oryza sativa* L.). *International Journal of Pure and Applied Bioscience*. 2017;5(6):1364–71. <https://doi.org/10.18782/2320-7051.5395>
- Tripathi N, Verma OP, Tiwari R, Singh PK, Rajpoot P. Estimation of genetic variability parameters in rice (*Oryza sativa* L.) gene bank in NEP Zone. *Journal of Pharmacognosy and Phytochemistry*. 2019;8(2S):960–2.
- Russinga AM, Srividhya A, Reddy VL, Latha P. Studies on variability, heritability and genetic advance for yield and its related traits in rice (*Oryza sativa* L.) genotypes. *International Journal of Current Microbiology and Applied Sciences*. 2020;9(9):3180–6. <https://doi.org/10.20546/ijcmas.2020.909.393>
- Lipi LF, Hasan MJ, Akter A, Quddus MR, Biswas PL, Ansari A, et al. Genetic variation, heritability and genetic advance in some promising rice hybrids. *SAARC Journal of Agriculture*. 2020;18(2):39–49. <https://doi.org/10.3329/sja.v18i2.51107>
- Singh B, Gauraha D, Sao A, Nair SK. Assessment of genetic variability, heritability and genetic advance for yield and quality traits in advanced breeding lines of rice (*Oryza sativa* L.). *The Pharma Innovation*. 2021;10(8):1627–30.
- Aditya JP, Bhartiya A. Genetic variability, correlation and path analysis for quantitative characters in rainfed upland rice of Uttarakhand hills. *Journal of Rice Research*. 2013;6(2):24–34.
- Parte D, Sathish K, Sahu P, Sharma D, Bhainsa K, Das BK, et al. Correlation and path analysis studies for yield and quality traits of rice (*Oryza sativa* L.) landraces and mutants of Chhattisgarh, India. *International Journal of Environment and Climate Change*. 2022;12(12):1771–9. <https://doi.org/10.9734/ijec/2022/v12i121621>
- Lakshmi MV, Suneetha Y, Yugandhar G, Lakshmi NV. Correlation studies in rice (*Oryza sativa* L.). *International Journal of Genetic Engineering and Biotechnology*. 2014;5(2):121–6.
- Faysal AS, Ali L, Azam MG, Sarker U, Ercisli S, Golokhvast KS, et al. Genetic variability, character association and path coefficient analysis in transplant Aman rice genotypes. *Plants*. 2022;11

- (21):2952. <https://doi.org/10.3390/plants11212952>
39. Thuy NP, Trai NN, Khoa BD, Thao NH, Phong VT, Thi QV. Correlation and path analysis of association among yield, micronutrients and protein content in rice accessions grown under aerobic condition from Karnataka, India. *Plant Breeding and Biotechnology*. 2023;11(2):117–29. <https://doi.org/10.9787/PBB.2023.11.2.117>
 40. Salunkhe H, Kumar A, Janeja HS, Krishna B, Talekar N, Mehandi S, et al. Genetic variability, correlation and path-coefficient analysis for yield and yield attributing traits in aerobic rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*. 2024;15(1):226–32. <https://doi.org/10.37992/2024.1501.026>
 41. Saran D, Gauraha D, Sao A, Sandilya VK, Kumar R. Correlation and path coefficient analysis for yield and yield attributing traits in rice (*Oryza sativa* L.). *International Journal of Plant & Soil Science*. 2023;35(18):94–101. <https://doi.org/10.9734/ijpss/2023/v35i183271>
 42. Ranawake AL, Amarasinghe UG. Relationship of yield and yield related traits of some traditional rice cultivars in Sri Lanka as described by correlation analysis. *Journal of Scientific Research and Reports*. 2014;3(18):2395–403. <https://doi.org/10.9734/JSRR/2014/12050>
 43. Augustina UA, Iwunor OP, Ijeoma OR. Heritability and character correlation among some rice genotypes for yield and yield components. *Cabi digital library*. 2013;73–84.
 44. Bhargava K, Shivani D, Pushpavalli SN, Sundaram RM, Beulah P, Senguttuvel P. Genetic variability, correlation and path coefficient analysis in segregating population of rice. *Electronic Journal of Plant Breeding*. 2021;12(2):549–55. <https://doi.org/10.37992/2021.1202.077>
 45. Belete DA, Tewachew A, Bitew M, Mulualet T. Correlation and path coefficient studies for yield and its components of upland rice (*Oryza sativa* L.) in North Western Ethiopia. *Journal of Scientific Agriculture*. 2022;6:14–9. <https://doi.org/10.25081/jsa.2022.v6.7253>
 46. Dhurai SY, Reddy DM, Bhati PK. Correlation and path coefficient analysis for yield and quality traits under organic fertilizer management in rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*. 2014;5(3):581–7.
 47. Bhadru D, Reddy DL, Ramesha MS. Correlation and path analysis of yield and yield components in hybrid rice (*Oryza sativa* L.). *Agricultural Science Digest*. 2019;32(3):199–203.
 48. Sabouri H, Rabiei B, Fazlalipour M. Use of selection indices based on multivariate analysis for improving grain yield in rice. *Rice Science*. 2008;15(4):303–10. [https://doi.org/10.1016/S1672-6308\(09\)60008-1](https://doi.org/10.1016/S1672-6308(09)60008-1)
 49. Rahimi M, Ramezani M, Kordrostami M. Selecting best rice varieties under drought stress and non-stress conditions using selection indices. *Yuzuncu Yil University Journal of Agricultural Sciences*. 2017;27(4):473–80.
 50. Vengadessan V, Pavithra S. Selection for drought tolerance in rice genotypes based on principal components and selection indices. *Electronic Journal of Plant Breeding*. 2020;11(4):1032–6. <https://doi.org/10.37992/2020.1104.168>
 51. SaiVenkat KS, Sekhar MR, Vemireddy LR, Umar SN, Bommisetty R. Multivariate analyses of rice indigenous germplasm for identification of high yield potential genotypes based on yield and its attributes. *Genetic Resources and Crop Evolution*. 2024;71(8):4807–23. <https://doi.org/10.1007/s10722-024-01932-5>
 52. Thakur A, Kumar NV, Bhardwaj N. Genetic diversity studies in elite rice landraces of North-Western Himalayas. *Journal of Cereal Research*. 2023;15(3). <https://doi.org/10.25174/2582-2675/2024/143430>
 53. Thakur K, Sarma MK. Genetic diversity and principal component analysis in cultivated rice (*Oryza sativa*) varieties of Assam. *The Indian Journal of Agricultural Sciences*. 2023;93(2):145–50. <https://doi.org/10.56093/ijas.v93i2.132052>
 54. Prakash S, Reddy SS, Chaudhary S, Vimal SC, Kumar A. Multivariate analysis in rice (*Oryza sativa* L.) germplasms for yield attributing traits. *Plant Science Today*. 2024;11(1):64–75. <https://doi.org/10.14719/pst.2231>
 55. Lakshmeesha R, Mahesh HB, Basavaraj MP, HM K, Veena SA. Principal component analysis of yield and its attributing traits in recombinant inbred lines of rice under submerged condition (*Oryza sativa* L.). *Archives of Current Research International*. 2024;24(4):39–44. <https://doi.org/10.9734/acri/2024/v24i44658>
 56. Vaidya N, Nayak G, Parida M, Sharma RS, Verma R, Katara JL, et al. Principal component analysis of grain yield and its attributing traits in recombinant inbred lines of rice (*Oryza sativa* L.) developed for submergence tolerance. *International Journal of Plant & Soil Science*. 2024;36(2):143–8. <https://doi.org/10.9734/ijpss/2024/v36i24375>
 57. Ahmed MS, Khalequzzaman M, Bashar MK, Shamsuddin AK. Agromorphological, physico-chemical and molecular characterization of rice germplasm with similar names of Bangladesh. *Rice Science*. 2016;23(4):211–8. <https://doi.org/10.1016/j.rsci.2016.06.004>
 58. Roy S, Marndi BC, Mawkhlieng B, Banerjee A, Yadav RM, Misra AK, et al. Genetic diversity and structure in hill rice (*Oryza sativa* L.) landraces from the North-Eastern Himalayas of India. *BMC Genetics*. 2016;17(1):107. <https://doi.org/10.1186/s12863-016-0414-1>
 59. Touthang L, Kalita H, Makdoh B, Angami T, Singh R. Agromorphological characterization, genetic variation and heritability analysis of rice landraces (*Oryza sativa* L.) of Arunachal Pradesh, Northeast India. *Indian Journal of Plant Genetic Resources*. 2021;34(2):185–95. <https://doi.org/10.5958/0976-1926.2021.00017.6>

Additional information

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

Reprints & permissions information is available at https://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: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonpublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (<https://creativecommons.org/licenses/by/4.0/>)

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