PLANT SCIENCE TODAY
Vol x(x): xx–xx
https://doi.org/10.14719/pst.12078

eISSN 2348-1900

RESEARCH ARTICLE

Genetic variability, heritability and multivariate analysis of 89 rice (*Oryza sativa* L.) genotypes under sodic soil

Poonam Sharma[1], Subhash Chandra Vimal[2*], Uma Bharti[1], Anamish Tyagi[3], Sonali Srivastava[3], Priya Raj[3], Avanji Saini[4], Aman Srivastava[3] & Ajeet Kumar Gupta[5]

[1]Department of Genetics and Plant Breeding, Chaudhary Sarwan Kumar Himachal Pradesh Krishi Vishvavidyalaya, Palampur 176 062, Himachal Pradesh, India

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

[3]Department of Genetics and Plant Breeding, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya 224 229, Uttar Pradesh, India

[4]Department of Genetics and Plant Breeding, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur 208 002, Uttar Pradesh, India

[5]Faculty of Agriculture Sciences, Mandsaur University, Rewas Dewda Road, Mandsaur 458 001, Madhya Pradesh, India

*Correspondence email - scvimalndgpb@gmail.com

Received: 02 October 2025; Accepted: 13 April 2026; Available online: Version 1.0: 31 July 2026

Cite this article: Poonam S, Subhash CV, Uma B, Anamish T, Sonali S, Priya R, Avanji S, Aman S, Ajeet KG. Genetic variability, heritability and multivariate analysis of 89 rice (*Oryza sativa* L.) genotypes under sodic soil. Plant Science Today (Early Access). https://doi.org/10.14719/pst.12078

Abstract

Sodic soils, characterised by excessive sodium ion accumulation, significantly hamper rice productivity in major cultivation zones. This study aimed to assess genetic variability, divergence and inheritance parameters among 89 rice (*Oryza sativa* L.) genotypes under sodic soil conditions to identify promising lines donor lines for breeding programs. The experiment was conducted during kharif 2021 at a single site (ANDUAT, Ayodhya, Uttar Pradesh) using an augmented block design (ABD) with 5 blocks and 13 quantitative traits were recorded. Principal component analysis (PCA) and Mahalanobis D^2 cluster analysis used to determine genetic divergence and trait contributions to variability. Significant differences were observed among genotypes for all traits, indicating substantial genetic variability. Eight clusters were formed among which cluster II comprised the highest number of genotypes was 16, while cluster III exhibited the highest inter-cluster distance (3.15), revealing its potential for identifying genetically diverse parental lines. The first 5 principal components showed 62.06 % of the total phenotypic variance, with PC1 (19.36 %) and PC2 (13.18 %) contributing most to total variation. Key traits contributing to this variation included grain yield per plant (GYP), biological yield (BY), plant height (PH), flag leaf area (FLA) and spikelet fertility (SF). The PC1 was primarily associated with the GYP (-0.5721) and biological yield per plant (-0.5187), whereas PC2 was influenced mainly by spikelet fertility (0.4130) and days to maturity (0.4241). Broad-sense heritability was ranged from 3.20 % to 96.63 % across different traits. Biological yield per plant exhibited the highest heritability (96.63 %), followed by grain yield and flag leaf area, which also showed high heritability coupled with high genetic advance as percent of mean, indicating predominant additive gene action and high selection efficiency. Promising genotypes such as WHM-1024 (25.96 g) and WHM-1016 (25.63 g) demonstrated high grain yield, whereas TCN-2080 (84.00), TCN-2066 (90.00) RAU-3074 (90.00), and NDR-2065 (119.00) showed early flowering and maturity. The integration of multivariate and genetic parameter analyses provides a robust basis for selection of genetically diverse parents and key yield-contributing traits for sodicity-tolerant rice breeding programs.

Keywords: genetic variability; heritability; multivariate analysis; sodic soil tolerance; sustainable agriculture

Introduction

Of the approximately 300000 plant species that humans could possibly consume, only roughly 100 are commonly grown (1). Nearly, 60 % of global population caloric intake is derived from three major crops are wheat, rice and maize where the rice is being the staple diet for nearly half of the world's population (2, 3). Despite its significance, rice productivity is increasingly constrained by abiotic stresses, among which soil sodicity is one of the most severe constrain. Rice does best in well-drained clay or loamy soils with a pH of 5.5 to 6.5 and grows best in tropical and subtropical climates, preferring temperatures between 20 °C and 30 °C. However, by separating important minerals and interacting with clay particles, sodic soils, which are defined by an excess of sodium ions can weaken soil structure and eventually hinder plant growth and nutrient uptake (4). High sodium ion concentrations in the soil cause this situation, which degrades the soil's structure, limits the capacity to absorb water and makes it more difficult for vital nutrients to be available. Sodicity is a major concern, particularly in lowland and rainfed rice environments and is found in many important rice-producing regions, including Egypt, Bangladesh, Australia and India. According to the FAO, salinity and sodicity will affect about 1.4 billion hectares, or 10.7 % of the world's land surface, by 2025. About 2.1 % of India's land area, or 6.72 million hectares, are classified as sodic soils, which include 3.77 million hectares of sodic soils and 2.95

million hectares of saline soils. Under sodic conditions, rice yields may decline by 20–50 % depending on stress severity (5). To increase the productivity of rice under sodic conditions by creating high-yielding, stress-tolerant cultivars continues to be the key breeding goal due to the restricted horizontal spread of cultivable land (4). Restoring soil health or developing rice varieties that combine high yield potential with tolerance to salt and sodicity are the two most important ways to alleviate the negative effects of sodic soils on rice agriculture. Of these, genetic improvement provides a viable and profitable approach for degraded and marginal lands. Finding viable lines for crop improvement in this situation requires assessing genetic variability, heritability and genetic diversity among genotypes (6). Effective selection depends on knowing the degree of genetic variation within a population since complex traits like yield and stress tolerance are controlled by several genes. Genetic distances between genotypes are measured using methods like Mahalanobis D^2 analysis, which makes it easier to classify genotypes into discrete clusters. By acting as a guide when selecting genetically distinct parents, these clusters raise the possibility of producing superior recombinants in later generations. Furthermore, by emphasising the traits that account for the majority of overall variation, principal component analysis (PCA) is a reliable multivariate technique for simplifying big datasets. This facilitates the classification of genotypes according to their performance and adaptability in addition to identifying trait associations (7). Although several studies have been conducted for rice diversity under normal conditions, but there is limited systematic characterisation of diverse rice germplasm under sodic soil environments in this region. Identification of genetically divergent and high-performing donor lines suitable for sodicity-tolerant breeding is still insufficiently documented (8). Breeders can also estimate the effectiveness of selection and identify traits that are more likely to be passed down through the generations by calculating heritability. Together, these genetic and multivariate techniques offer a thorough framework for creating accurate and effective breeding programs. To find desirable donor genotypes and important selection qualities for breeding programs, the current study was conducted to evaluate genetic variability, divergence and inheritance parameters among diverse rice genotypes under sodic soil conditions.

Materials and Methods

Experimental site

The experiment was conducted during the kharif season of 2021 at the Genetics and Plant Breeding Research Farm, Acharya Narendra Deva University of Agriculture and Technology (ANDUAT), Kumarganj, Ayodhya, Uttar Pradesh, India. The research site is geographically located at 26.47° N latitude and 82.12° E longitude, with an elevation of 113 m above mean sea level. The location experiences a semi-arid, subtropical climate characterised by hot summers and cold winters. The soil at the experimental site was alkaline in nature, with a pH of 8.9, an electrical conductivity (EC) of 2.21 dS m^{-1} and an exchangeable sodium percentage (ESP) of 45 %.

Experimental materials

A total of 89 rice genotypes, comprising 85 diverse germplasm lines and 4 standard check varieties, were used in this study. The 85 germplasm lines were obtained from the institutional breeding stock of the Department of Genetics and Plant Breeding, ANDUAT, Ayodhya and included advanced breeding lines and locally adapted selections developed for stress environments. The 4 check varieties (Pusa Sugandha-5, Sarjoo-52, Narendra usar-3 and NDR-2065) were locally cultivated varieties recommended for sodic or normal conditions in Uttar Pradesh. These genotypes were selected to assess genetic diversity and breeding performance under sodic-soil conditions.

Experimental design and layout

The experiment was conducted in an augmented block design (ABD), consisting of five blocks. Each block contained 17 unreplicated test genotypes and 4 replicated check varieties, resulting in 21 plots per block. Check varieties were randomised within each block and used to adjust block effects. Each block had three rows of 3 m in length, with a spacing of 20 cm between rows and 10 cm between plants, 10 plants per genotype were planted and standard agronomic procedures recommended for rice cultivation for best possible crop growth and development. After 25 days, seedlings cultivated in a nursery were transplanted. The prescribed fertiliser dose was given at 120:60:40 kg ha^{-1} (N : P_2O_5 : K_2O). The half of nitrogen was top-dressed in two equal splits at the active tillering and panicle start stages, while the other half of nitrogen, complete phosphorus and potassium were applied as basal at transplanting. Throughout vegetative growth, irrigation was maintained to guarantee shallow submergence (3–5 cm). One pre-emergence spray of butachlor (1.5 kg a.i. ha^{-1}) and hand weeding at 13 days after transplanting (DAT) were part of standard weed control. When necessary, plant preservation measures were implemented in accordance with university instructions.

Data collection

Thirteen quantitative traits were observed morphologically. Data were recorded on 5 randomly selected genotypes from the central row of each plot to avoid border effects. Days to maturity (DM), which was recorded on a plot basis, were the single exception. The following characteristics were measured: Days to 50 % flowering (DFF), DM, plant height (PH) in centimetres, flag leaf area (FLA) in centimetres, number of productive tillers per plant (PTP) in centimetres, spikelet fertility (SF) in percentage, number of panicles bearing per plant (PBP), number of panicle length (PL) in centimetres, number of grains per panicle (GP), biological yield per plant (BYP) in grams, harvest index (HI) in grams, 1000 seed weight (TSW) in grams and grain yield per plant (GYP) in grams were among the thirteen quantitative traits that were measured. Measurement procedures were according to IRRI Standard Evaluation System with specific protocols detailed above to ensure reproducibility (9).

Statistical analysis

The data collected was statistically analysed to determine the genotypes' genetic diversity and variability. The ABD's analysis of variance (ANOVA) was carried out using SAS software version 9.2's General Linear Model (GLM), as described by previous researchers (10). Analysis of variance for the ABD was performed using the following linear model:

$$Y_{ij} = \mu + B_i + G_j + e_{ij}$$

Where: μ= overall mean

B_i = fixed effect of ith block

G_j= genotype effect (checks treated as fixed; test entries treated as random for variance estimation)

e_{ij} = random error assumed ~ $N(0, \sigma^2 e)$

Adjusted means for unreplicated genotypes were estimated using block correction based on replicated check performance within each block, following Federer's ABD adjustment procedure.

Estimation of genetic parameters

Phenotypic variance (σ^2p), genotypic variance (σ^2g) and environmental variance (σ^2e) were estimated as:

$\sigma^2g = (MSg-MSe)/r$

$\sigma^2p = \sigma^2g + \sigma^2e$

PCV and GCV were calculated as:

$$GCV = \frac{\sqrt{\sigma_g^2}}{\bar{X}} \times 100$$

$$PCV = \frac{\sqrt{\sigma_p^2}}{\bar{X}} \times 100$$

Broad-sense heritability:

$$h^2 = \frac{\sigma_g^2}{\sigma_p^2}$$

Genetic advance was calculated as

$$GA = k \times \sigma_p \times h^2$$

Where k = 2.06 (selection intensity at 5 %).

Genetic advance as percent of mean:

$$GA\% = \frac{GA}{\bar{X}} \times 100$$

Heritability in the broad sense was computed using the method of previous researchers (11) and genetic advance was estimated following the approach of previous studies (12). Phenotypic and genotypic variances, their respective coefficients, heritability and genetic advance were calculated using formulas proposed previously (13). Non-hierarchical Euclidean cluster analysis was carried out to assess the genetic diversity among the genotypes (14, 15). A dendrogram illustrating genetic relationships was generated using windostat software version 9.3 (16). The PCA was performed on standardised data (mean-centred and scaled). Eigenvalues > 1 were retained following Kaiser's criterion. Negative values indicate inverse association with the principal axis due to centring and do not imply negative biological relationships whereas, the positive values indicates that trait increases in the direction of that PC axis.

Results and Discussion

Analysis of variance is a statistical technique that is frequently used in plant breeding to evaluate the significance of genotype-to-genotype variations for different traits. It assists in separating the overall variance seen in an experiment into components caused by environmental effects, genetic variations and experimental error. A significant genotype effect in ANOVA indicates the presence of genetic variability, which is essential for effective selection and crop improvement. The Table 1 presents an analysis of variance (ANOVA) for 13 agronomic traits and yield related traits of rice examined under sodic soil conditions using an augmented design in a field experiment. Significant differences were observed among the check varieties for all traits, as indicated by the double asterisks (**), showing the considerable genetic variations. High variation was seen in traits such as PH, biological yield and grain production, highlighting their significance in selection for higher yield. Significant block effects observed for most of the characters indicate the influence of environmental variation across the replication. However, the use of replicated checks in each block allowed adjustment of genotype means, thereby minimising environmental variations. The relatively small error mean squares confirm the accuracy and dependability of the experimental findings.

Mean performance of genotypes for various traits in rice

The Fig. 1, summarises the mean performance, range, standard error and critical difference (CD) for thirteen quantitative traits for several rice genotypes. These findings provide first understanding of the degree of genetic variation among the genotypes under study. The mean performance of 89 genotypes of rice (*O. sativa*) across 13 agronomic and yield-related traits is displayed in the bar graph. These traits are crucial indicators of plant productivity, growth and development. It is evident from the compiled results in Fig. 1 that there is a significant amount of genetic variation among the genotypes.

Phenological traits

The Fig. 1 illustrates that the average number of DFF was 97.07 and the mean for number of DM was 127.08. This supports the previous

Table 1. Analysis of variance of augmented design for 13 characters in rice under sodic condition

Source of variation	DF	Days to 50 % flowering	Days to maturity	Plant height (cm)	Flag leaf area(cm) 2	No. of productive tillers/plant	Panicle length (cm)	No. of fertile spikelets/ panicles
Block	4	8.01	8.20	537.87**	53.01**	1.89**	9.66**	56.00**
Checks	3	168.78**	238.83**	1498.38**	80.36**	20.71**	38.04**	442.38**
Error	12	12.20	9.01	18.96	0.61	0.10	0.57	7.49

Source of variation	DF	Spikelet fertility(%)	Biological yield/plant (g)	Harvest index (%)	1000 grain weight (g)	L:B ratio	Grain yield /plant (g)
Block	4	114.32**	25.19**	31.28**	4.70**	0.43**	12.53**
Checks	3	20.24**	1540.41**	125.59**	29.97**	0.56**	156.56**
Error	12	6.88	0.99	1.75	0.39	0.01	0.29

DF: degrees of freedom; * and ** indicate significance at $p \leq 0.05$ and $p \leq 0.01$, respectively; NS: non-significant difference; L:B ratio: length: breadth ratio.

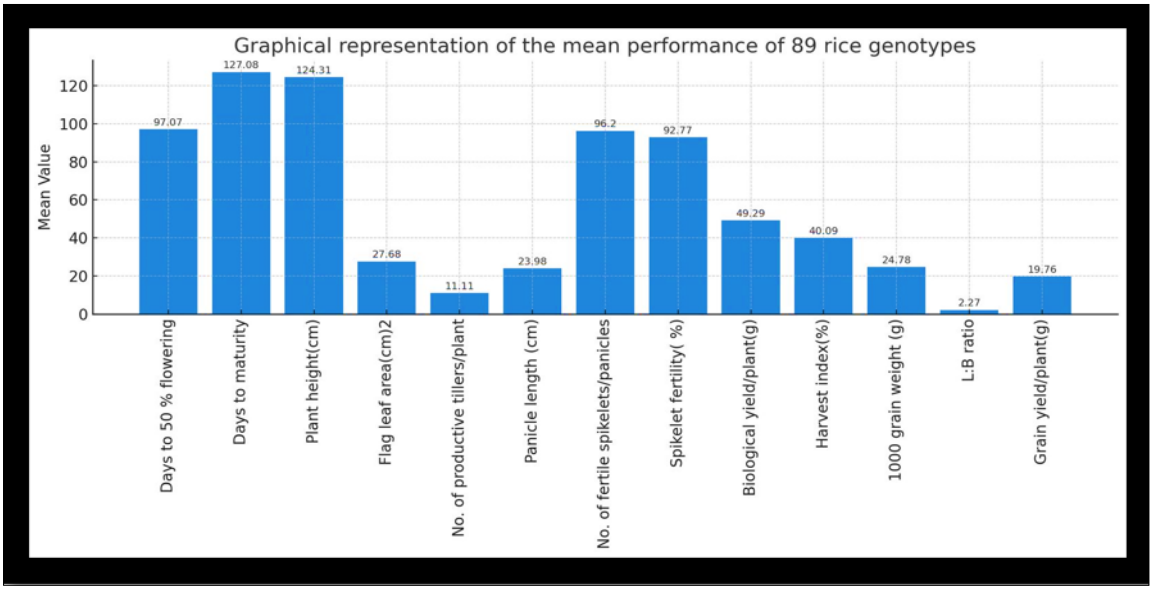


Fig. 1. Graphical representation of the mean performance of eight nine rice (*Oryza sativa*) genotypes.

finding that flowering and maturity times differed significantly between genotypes, ranging from 84 to 103 days and 119 to 135 days, respectively. This variance is best illustrated by genotypes like TCN-2080 (earliest flowering) and NDR-2065 (earliest maturity). Early maturing genotypes are ideal for time-bound cropping systems because they assist shorten the crop cycle without sacrificing yield potential.

Plant architecture traits

Across genotypes, the average PH was 124.31 cm, with extremes ranging from 88.58 cm (Narendra Usar-3) to 167.22 cm (TCN-2108). The mean of FLA was 27.68 cm^2, which again showed significant variance. The largest FLA was 48.39 cm^2 for TCN-2092 and the smallest was 18.09 cm^2 for TCN-2072. Large flag leaf genotypes are preferred for high-yielding hybrids because they improve photosynthetic supply and light interception.

Yield contributing traits

According to the graph, there were on average 11.11 productive tillers per plant. With the greatest number of tillers (more than 15), WHM-1001 and WHM-1003 were the most notable, confirming the conclusion that this trait varies widely. The average panicle length was 23.98 cm, with WHM-1000 having the shortest length (12.07 cm) and TCN-2058 having the longest (37.88 cm). Higher sink capacity genotypes, such as WHM-1001 (17.72 tillers), ought to be utilised in breeding programs that prioritise yield.

Reproductive traits

The bar graph shows a mean SF of 92.77 % and an average of 96.2 fertile spikelets per panicle. In programs aimed at increasing production, TCN-2118 and WHM-1028 with more than 120 fertile spikelets are greatly desired. When breeding climate-resilient rice, high -performing genotypes such as WHM-1016 can be given priority since they guarantee reproductive success even in stressful situations.

Biomass and harvest index

The average BY per plant was 49.29 g, but genotypes like TCN-2071 and WHM-1001 had maximums of 70.12 g and 65.98 g, respectively. With an average HI of 40.09 %, there was a wide variety; WHM-1006 had the highest index at 57.55 %, while TCN-2104 had the lowest at 30.04 %. These genotypes are appropriate to produce both grain and straw because they combine substantial growth with yield potential and have high biomass and harvest index.

Grain traits

The average 1000-grain weight (D^2) was 24.78 g, while the length-to-breadth (L/B) ratio averaged 2.27, indicating moderate grain slenderness. With the highest L/B ratio (3.64), TCN-2085 might be an ideal option for long-grain preferences. The highest-yielding genotypes were WHM-1024 (25.96 g) and WHM-1016 (25.63 g), with an average grain production per plant of 19.76 g. Consistent grain weight genotypes guarantee homogeneity and increased farmer acceptance. High yielders like as TCN-2085, WHM-1016 and WHM-1024 (25.96 g) are ideal for hybrid or direct release development.

Overall, this graphical summary highlights that the genotypes demonstrate diverse performance, particularly in key yield-contributing traits like SF, panicle traits and biomass partitioning. These results can guide further selection and breeding programs aimed at enhancing rice yield potential under specific agro-ecological conditions.

These findings were also reported in previous studies (17). It showed a wide range of traits across rice genotypes, with CO (R) 50 having the most grains per panicle and CO 53 having the highest yield. ASD 17 had the earliest flowering but the lowest yield, whereas IR 64 produced the most tillers. These differences show important genetic variety for agricultural improvement in the future.

Estimation of genetic differences using coefficients of phenotypic and genotypic variation

Crop improvement requires the evaluation of genetic variability, which is often done with the use of genotypic and phenotypic coefficients of variation. The phenotypic coefficient of variation (PCV) accounts for both genetic and environmental variables whereas, the genotypic coefficient of variation (GCV) calculates the amount of trait variation attributable solely to genetic factors. The trait is mostly controlled by hereditary factors with little influence from the environment, as indicated by small variations between PCV and GCV levels. In plant breeding programs, these factors are useful tools for identifying traits with high heritability and selection potential.

The Important genetic indices for several traits in rice (*O. sativa*) genotypes are summarised in Fig. 2: heritability (h^2bs), GA % and phenotypic and genotypic coefficients of variation (PCV and GCV). The PCV values were typically only slightly higher than GCV, indicating that environmental influences had little effect on trait expression. In this study, the GCV varied from 0.65 % to 11.91 %. Yield

and yield-contributing characters showed moderate to high GCV and PCV. Grain yield per plant (11.10 % and 11.43 %), BY per plant (10.84 % and 11.02 %), FLA (10.79 % and 11.15 %) and the L:B ratio (11.91 % and 12.87 %, respectively) had the highest GCV and PCV values. These findings suggest that these qualities have significant innate genetic variability. On the other hand, DFF (0.65 %) and DM (1.11 %) showed extremely low GCV values, indicating small genetic variability for these phenological traits. For most parameters (such as GYP, BYP, L:B ratio and productive tillers per plant), the comparatively small differences between GCV and PCV values show little influence of environment and confirm that the phenotypic expression closely reflects genotypic potential. However, larger differences for traits such as SF suggest greater environmental sensitivity.

Similar results have been reported by previous researchers, who investigated 19 rice hybrids and discovered that the PH, DM and panicle length had minor genotypic and phenotypic coefficients of variation (18). Days to heading, DM, PH and panicle length all have genotypic and phenotypic coefficients of variation that vary slightly, indicating that there is sufficient genetic diversity in the variables for favourable selection (19). The number of panicles per square meter, the number of tillers per plant, the weight of the panicles, the number of grains per panicle, the TGW and the grain yield all showed high to moderate differences in genotypic and phenotypic coefficients of variation, indicating that the environment has a major impact on the traits.

Heritability and genetic advance

Heritability varied from 3.20 to 96.63 % among the thirteen traits evaluated; the greatest values were found for BY (96.63 %), GYP (94.32 %), FLA (93.57 %), TP (88.81 %), L:B ratio (85.68 %) and panicle length (82.20 %). High heritability indicates that the observed variation is largely genetic and less influenced by environmental factors. Whereas, moderate heritability was observed for SF (42.26 %) and number of fertile spikelets per panicle (67.73 %), while very low heritability was observed for DFF (3.20 %) and relatively low for DM (18.16 %), suggesting greater environmental influence on these parameters.

However, great genetic gain is not always guaranteed by high heritability alone. A more accurate estimate of the effectiveness of selection can be obtained by integrating genetic advances and heritability estimates, as highlighted by previous researchers (20). The GA as percent of mean provides proper insight for the nature of gene action.

High GA coupled with high heritability was observed L:B ratio (22.71 %), GYP (22.21 %), BY (21.94 %), FLA (21.50 %) and PTP (15.74 %). This combination strongly indicating that most of the variation is genetically controlled and there were largely due to additive gene action and implying that direct phenotypic selection in early segregating generations (F2 and F3) would be effective. Therefore, direct phenotypic selection would be highly effective in improving them. In contrast, traits such as DFF and DM exhibited low heritability and very low genetic advance, indicating limited scope for improvement through simple selection and suggesting possible non-additive gene effects or strong environmental influence. Similar results were found earlier reports that showed low broad sense heritability of 0.16 for the number of panicles per plant and 0.20 to 0.33 for the number of spikelets per panicle, while revealed high narrow sense heritability for grain weight, moderate for the number of spikelets per panicle and low for the number of panicles per plant (21, 22).

These results show that selection was able to improve these traits, which were mostly governed by genetic factors. Because of their increased transmissibility and variability, traits with high heritability and GA are considered suitable for selection and breeding objectives. These findings are in line with the earlier research (6).

Cluster formation conducted through Mahalanobis squared distance metric

Mahalanobis D^2 distance cluster analysis, which takes into account both trait variance and covariance, was an excellent way to quantify genetic divergence. Lower D^2 values were used to group genotypes with similar agro-morphological features, whereas higher distances were associated with greater genetic diversity. In breeding programs, this approach works well for choosing parents from a variety of backgrounds. Similarities in agro-morphological traits led

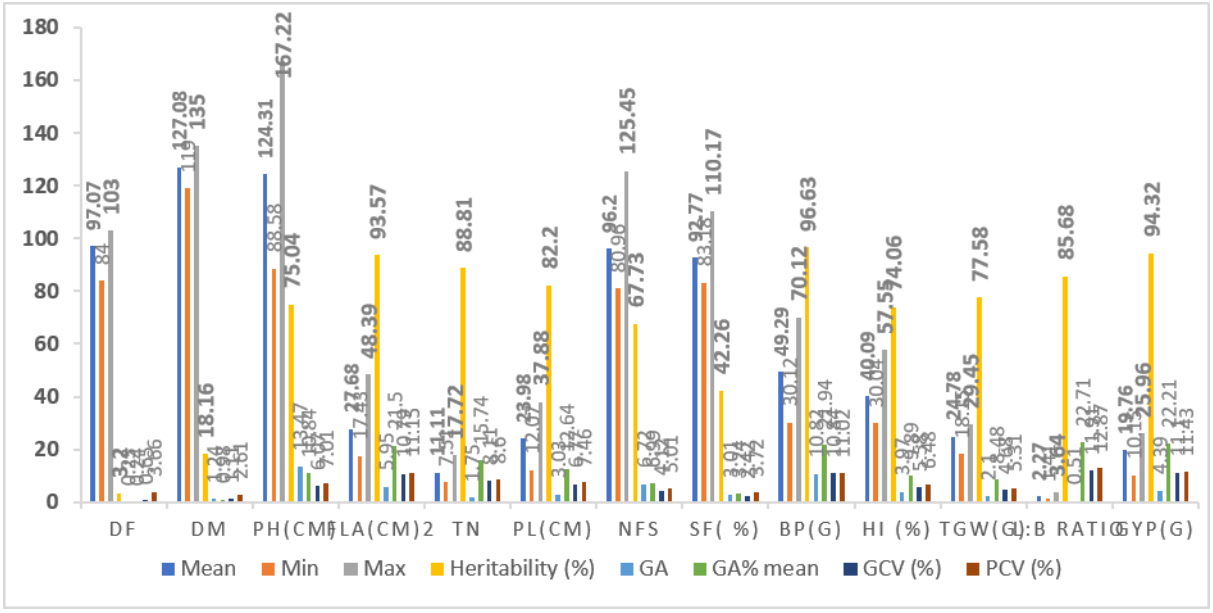


Fig. 2. Graphical representation of heritability and genetic advance as percent of mean in rice (*Oryza sativa*) genotypes.

PCV: phenotypic coefficient of variation; GCV: genotypic coefficient of variation; ECV: environmental coefficient of variation; GA: genetic advance (%); h^2: heritability (%).

Table 2. Number of genotypes in each clusters

Clusters	No of genotypes	Genotypes
I	12	1,6,14,21,31,32,49,67,68,71,84,85
II	16	2,7,8,9,12,13,16,17,25,34,48,73,76,77,83,86
III	6	3,23,46,87,88,89
IV	13	22,30,35,36,37,39,40,41,44,64,66,75,79
V	8	4,5,24,26,45,56,57,60
VI	12	33,47,50,61,62,63,65,72,74,80,81,82
VII	9	10,20,27,28,38,51,52,53,69
VIII	13	11,15,18,19,29,42,43,54,55,58,59,70,78

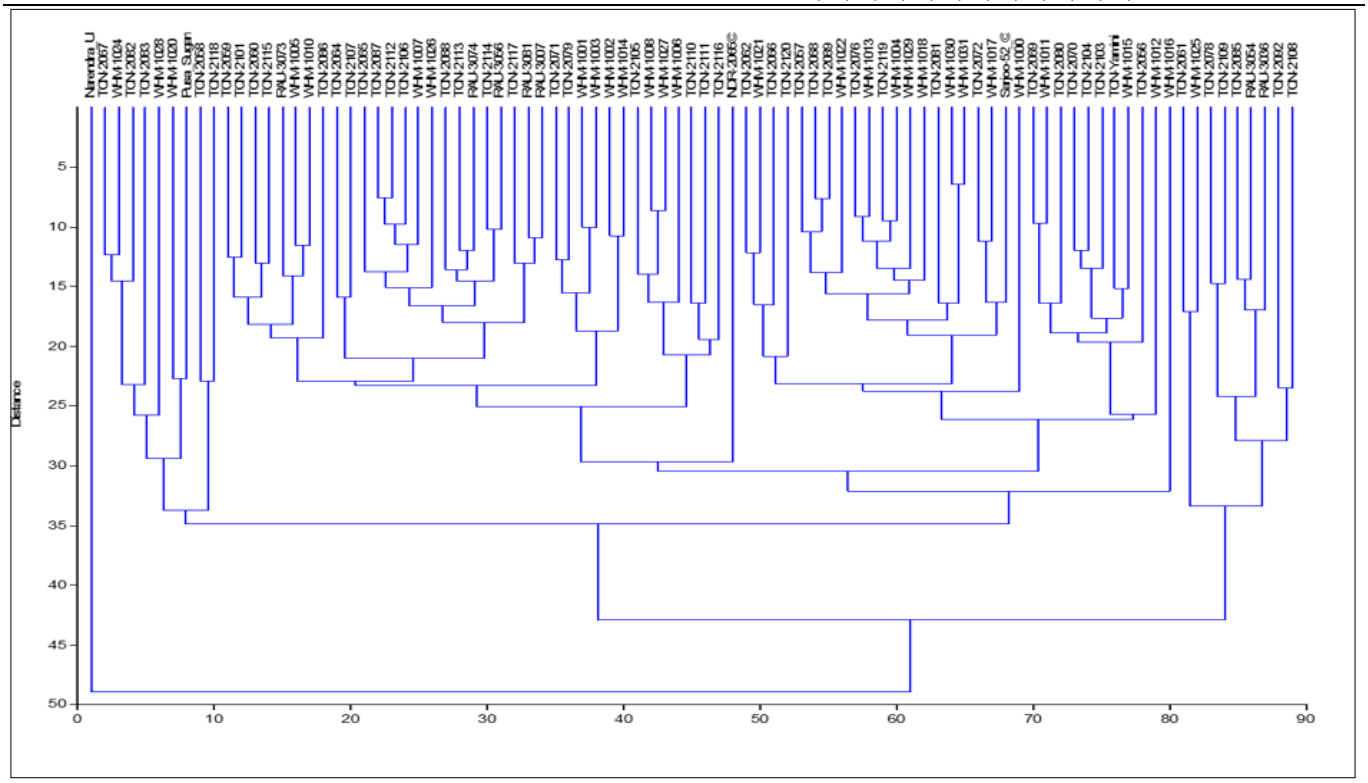


Fig. 3. Clustering patterns of 89 rice genotypes under sodic soil based on Mahalnobis D^2 analysis.

to the identification of eight different clusters (Fig. 3). As mentioned in Table 2, Cluster II comprised the highest number of genotypes (16) suggesting that there was limited divergence among entries, followed by Clusters IV and VIII with 13 genotypes each and Clusters I and VI with 12 genotypes. Nine genotypes were included in Cluster VII, which was one of the larger clusters. By comparison, Cluster III had the smallest group with 6 genotypes, while Cluster V had 8 genotypes. These findings align with those presented in earlier reports (23).

Intra-Inter cluster distances among eight clusters

Intra- and inter-cluster distances among eight distinct clusters formed from 89 rice genotypes. The Fig. 4, helps in understanding the genetic diversity and relationships among the clusters. The identification of clusters with high inter-cluster distances indicates the presence of substantial genetic divergence among the grouped genotypes. Such genetically distant clusters are very important in various breeding programs because hybridising these parents from the clusters is more likely to produce higher heterosis and wider variability in segregating generations which enhances the chances of getting superior recombinants and transgressive segregants. Therefore, genotypes showed the highest inter-cluster distances should be selected as parents for hybridisation to broaden the genetic base and accelerate genetic improvement.

Understanding the genetic diversity and relationships between the clusters is made clear by Fig. 4. The maximum intra cluster distance was recorded for Cluster III indicating maximum genetic divergence followed by Cluster VIII. Cluster V, II, VI showed the equal intra class cluster distance than followed by Cluster VII, Cluster I. Maximum inter cluster distance was found between Cluster III-VIII (4.90) followed by Cluster III-V (4.84), Cluster I-VIII (4.83), Cluster III-VII (4.62), Cluster I-V (4.42) Cluster III-IV (4.34) and cluster I-VI (4.27) whereas, minimum inter-cluster distance was recorded for cluster IV -VIII (2.68) followed by cluster II-V (2.76), cluster IV-VI (2.83). This indicates that the cultivars in these clusters are highly genetically diverse, making them suitable for use in hybridization programs to produce desirable recombinants for the development of high-yielding varieties. To increase heterosis and variability, it was proposed to select parental lines from clusters with significant inter-cluster distances (24, 25).

Clusters means for 13 characters in rice

All measured traits showed a large range of inter-cluster variation, indicating significant genetic diversity among the rice genotypes. As mentioned in Table 3, Cluster IV showed the latest flowering at 98.62 days, whereas Cluster VII showed the earliest flowering at 90.22 days. Cluster II matured the earliest (125.05 days), while Cluster III had the greatest mean value (130.54 days), followed by Cluster VI (129.25 days). Significant variations in PH were noted, with Cluster III having the shortest plants (107.74 cm) and Cluster VIII having the tallest plants (138.20 cm). Cluster V had the highest FLA (33.70 cm^2), whereas Cluster II had the smallest (23.17 cm^2). The PTP ranged from

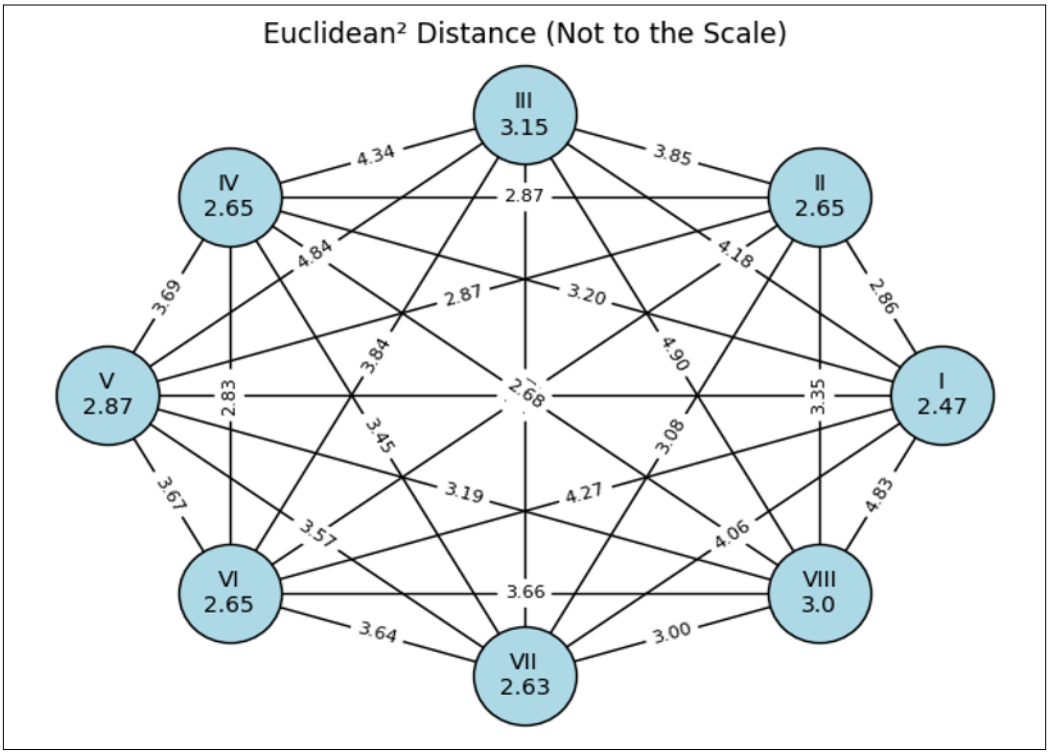


Fig. 4. Graphical representation of intra- and inter-cluster distances among eight groups formed from 89 rice (*Oryza sativa*) genotypes.

Table 3. Cluster means for 13 characters in rice germplasm by D^2 analysis

Characters		Days to 50 % flowering	Days to maturity	Plant height (cm)	Flag leaf area (cm^2)	No. of productive tillers/ plant	Panicle length (cm)	No. of fertile spikelets/ panicles	Spikelet fertility (%)	Biological yield/ plant (g)	Harvest index (%)	1000-grain weight (g)	L : B ratio	Grain yield/ plant (g)
I	**Mean**	97.75	127.92	119.55	25.56	10.35	22.27	91.18	90.64	34.80	34.30	24.14	2.14	11.89
	± SE	0.97	3.18	12.42	4.96	1.40	2.24	5.81	5.12	5.89	2.91	2.09	0.33	1.96
II	**Mean**	98.01	125.05	114.10	23.17	11.21	23.63	93.06	89.53	47.61	40.67	25.58	2.40	19.34
	± SE	1.57	3.27	9.09	4.77	1.52	2.81	8.50	3.20	5.13	4.28	1.99	0.41	2.68
III	**Mean**	97.38	130.54	107.74	27.83	11.26	30.04	112.87	91.99	39.64	39.33	26.26	2.62	15.39
	± SE	2.18	1.55	13.99	6.41	1.46	6.53	8.49	6.05	7.17	2.23	2.40	0.51	2.17
IV	**Mean**	98.62	128.08	133.10	30.90	9.35	23.38	96.76	94.36	51.67	38.35	23.98	1.72	19.71
	± SE	1.12	2.84	18.13	5.18	0.97	2.51	5.66	3.77	7.51	2.95	3.06	0.25	2.36
V	**Mean**	96.50	125.25	129.63	33.70	11.93	21.72	93.93	91.19	54.03	41.03	25.35	3.30	22.08
	± SE	2.33	3.49	12.15	5.95	2.19	4.02	4.19	6.03	6.51	2.40	2.10	0.26	2.18
VI	**Mean**	98.33	129.25	122.43	24.69	10.96	24.17	99.95	97.91	47.20	47.33	25.68	2.18	22.19
	± SE	1.23	2.86	11.57	2.68	1.22	1.98	7.63	5.01	6.20	5.10	2.58	0.43	2.58
VII	**Mean**	90.22	126.22	122.35	26.22	11.17	25.08	93.30	92.82	51.25	40.53	23.10	2.19	20.78
	± SE	2.59	3.19	10.38	4.59	1.98	2.09	3.96	4.07	6.88	2.84	2.75	0.29	3.22
VIII	**Mean**	98.23	126.85	138.20	32.52	13.13	24.15	98.16	94.30	62.15	39.45	24.84	2.06	24.43
	± SE	2.31	3.72	14.32	7.61	2.05	1.45	8.50	4.98	4.68	3.04	2.57	0.42	1.51

9.35 in Cluster IV to 13.13 in Cluster VIII. Panicle length varied across clusters, with Cluster III showing the longest (30.04 cm) and Cluster V the shortest (21.72 cm). Cluster III also had the highest number of fertile spikelets per panicle (112.87), while Cluster I had the fewest (91.18). Spikelet fertility was greatest in Cluster VI (97.91 %) and lowest in Cluster II (89.53 %). The highest BY was recorded in Cluster VIII (62.15 g) and the lowest in Cluster I (34.80 g). Similarly, Cluster VI had the highest HI (47.33 %), while Cluster I had the lowest (34.30 %). The TGW ranged from 23.10 g in Cluster VII to 26.26 g in Cluster III. The L : B ratio peaked in Cluster V (3.30) and was lowest in Cluster IV (1.72). For GYP, Cluster VIII stood out with the highest value (24.43 g) and Cluster I recorded the lowest (11.89 g).

Notably, our findings support earlier findings that genetic diversity and geographic origin don't correspond, as documented previously (26, 27).

Contribution of different characters genetic divergence

The traits that showed the greatest variability across clusters among the evaluated traits were PH, BY, spikelet fertility per panicle (SFP), HI, GYP and FLA. This suggests that these traits play a significant role in contributing to genetic divergence and useful for hybridisation programs. Plant height, for instance, varied significantly between Cluster III (107.74 cm) and Cluster VIII (138.20 cm). The BY, which peaked in Cluster VIII at 62.15 g, also displayed a broad range. The HI ranged from 47.33 to 34.30 % and SFP varied from 112.87 to 91.18. Flag leaf area ranged from 23.17 cm^2 to 33.70 cm^2, indicating wide phenotypic diversity and GYP varied from 11.89 g in Cluster I to 24.43 g in Cluster VIII. Days to 50 % flowering (DFF), which ranged from 90.22 to 98.62 days, likewise shown significant variation. There were clear variations in flowering behaviour and reproductive traits between genotypes, as evidenced by the range of panicle length (PL) from 21.72 cm to 30.04 cm and SF values from 89.53 % to 97.91 %. On the other hand, traits like TGW, L:B ratio, PTP and DM varied only

little between clusters. For example, TSW varied from 23.10 g to 26.26 g and DM ranged narrowly from 125.05 to 130.54 days, indicating a restricted involvement in genetic differentiation. These less variable characteristics had a negligible impact on the study's overall genotype divergence.

Most promising rice genotypes for different traits

After evaluating 85 rice genotypes, several potential lines for important attributes were found. The result are presented in Fig. 5, exhibited that the largest grain yields were observed by WHM-1024, WHM-1016 and TCN-2071, which made them appropriate donor parents for hybridisation. While TCN-2058 and TCN-2083 displayed superior panicle lengths, TCN-2085, TCN-2116 and TCN-2117 had the greatest FLA. There was high SF in TCN-2085, WHM-1016 and WHM-1018. Genotypes like as TCN-2080, RAU-3074 and TCN-2066 were among the earliest to flower. These genotypes present significant promise for future efforts aimed at improving rice.

Principal component analysis

Principal components that explain a significant amount of the overall variation in the dataset are considered crucial if their eigenvalues are greater than 1 according to rotated values. For plant breeders looking to increase genetic improvement, choosing such components is crucial because they represent traits that significantly add to overall diversity. The variation seen across 13 quantitative measures was broken down in this study using PCA, which allowed for the identification of genotypes with desired traits. While keeping the important details from the original dataset, this multivariate method aids in dimensionality reduction. The Kaiser criterion, which recommends keeping components with eigenvalues greater than one, was satisfied by the study, which showed that the top five main components, with eigenvalues of 2.5, 1.7, 1.4, 1.2 and 1.2 respectively, captured most of the variability. When taken as a whole, these elements show the most informative features and have significant effects on breeding efficiency. They therefore play a crucial role in directing selection choices, enabling scientists to concentrate on genotypes that present the best potential for genetic advancement in rice improvement programs.

The PCA identified significant elements of variation in the 13 traits evaluated in the genotypes of rice. As shown in Table 4 and Fig. 6, with an eigenvalue of 2.5, PC1 was the biggest contributor to the explanation of total variation (19.36 %), followed by PC2 (13.18 %), PC3 (10.87 %), PC4 (9.61 %) and PC5 (9.04 %). Associations between each component and agronomic trait were found. PC1's influence on variability was minimal, as evidenced by its mild associations with DM and the L:B ratio. The PC2 is an essential component in determining productive genotypes because of its strong correlation with yield-contributing traits such phenological traits, SF and the number of fertile spikelets per panicle. Grain quality and efficiency have been emphasised by PC3 through its correlation with harvest index, TGW and L:B ratio. PC4 captured panicle-related variation, particularly panicle length and spikelet number, while PC5 reflected changes in reproductive efficiency via HI and SF. Overall, PCA effectively dissected trait contributions to genetic diversity, highlighting the utility of multivariate analysis for guiding selection and improvement in rice breeding programs. Similar finding was observed in previous studies (28).

The screen plot (Fig. 7) illustrates the spatial distribution of rice genotypes based on the first 2 principal components-PC1 (19.36 %) and PC2 (13.18 %)- which together account for 32.54 % of the total variance. A particular genotype is represented by each point in the plot and variations in the trait composition of the points are indicated by the distance between them. A small number of genotypes, including TCN-2092, WHM-1003 and NDR-2065©, appear distant from the centre, although the majority are grouped close to the origin, suggesting shared or similar trait profiles. This implied that certain genotypes differ from the others in that they display unique combinations of traits. Thus, the PCA score plot is helpful in identifying distinct candidates that would be useful for breeding or more genetic research, as well as in highlighting genotypic diversity.

The Fig. 8 represents the correlation circle plot illustrates the relationships between agronomic traits and the first two principal components (PC1: 19.36 %, PC2: 13.18 %). Traits like SF, number of fertile spikelets per panicle, DM and DFF are closely aligned and strongly projected in the positive PC2 direction, indicating strong positive correlation and significant contribution to that axis. Meanwhile, GYP and BYP point negatively along PC1, suggesting these traits contribute to variability in a different direction. Traits closer to the origin contribute less to variation captured by PC1 and



Fig. 5. Contribution of different 13 characters in genetic divergence.

Table 4. Eigenvalues and corresponding eigenvectors for the first five principal components derived from the analysis of 13 agronomic traits in rice (*Oryza sativa*)

Traits	PC1	PC2	PC3	PC4	PC5
Days to 50 % flowering	-0.0048	0.3415	0.046	-0.2645	-0.5232
Days to maturity	0.0668	0.4241	0.0372	-0.1589	0.1478
Plant height	-0.3253	0.0632	-0.4413	-0.0047	-0.193
Flag leaf area	-0.3031	0.1904	-0.2701	0.0662	-0.2633
No. of productive tillers plant	-0.2114	-0.1813	0.2416	-0.1107	-0.3095
Panicle length	-0.0667	0.1176	0.2097	0.6695	-0.0611
No. of fertile spikelets panicles	-0.1247	0.4399	0.2546	0.4578	0.0687
Spikelet fertility	-0.2814	0.413	-0.043	-0.069	0.3568
Biological yield per plant	-0.5187	-0.2326	-0.0461	0.121	-0.0864
Harvest index	-0.238	0.0648	0.3805	-0.3817	0.3986
1000-grain weight	-0.0145	0.2251	0.4629	-0.198	-0.4248
L : B ratio	0.0265	-0.3368	0.4228	0.1328	-0.0673
Grain yield per plant	-0.5721	-0.1734	0.1412	-0.0916	0.1221
Eigen value	2.5	1.7	1.4	1.2	1.2
The proportion of variation %	19.357	13.184	10.873	9.613	9.042
Cumulative variation %	19.357	32.54	43.413	53.026	62.068

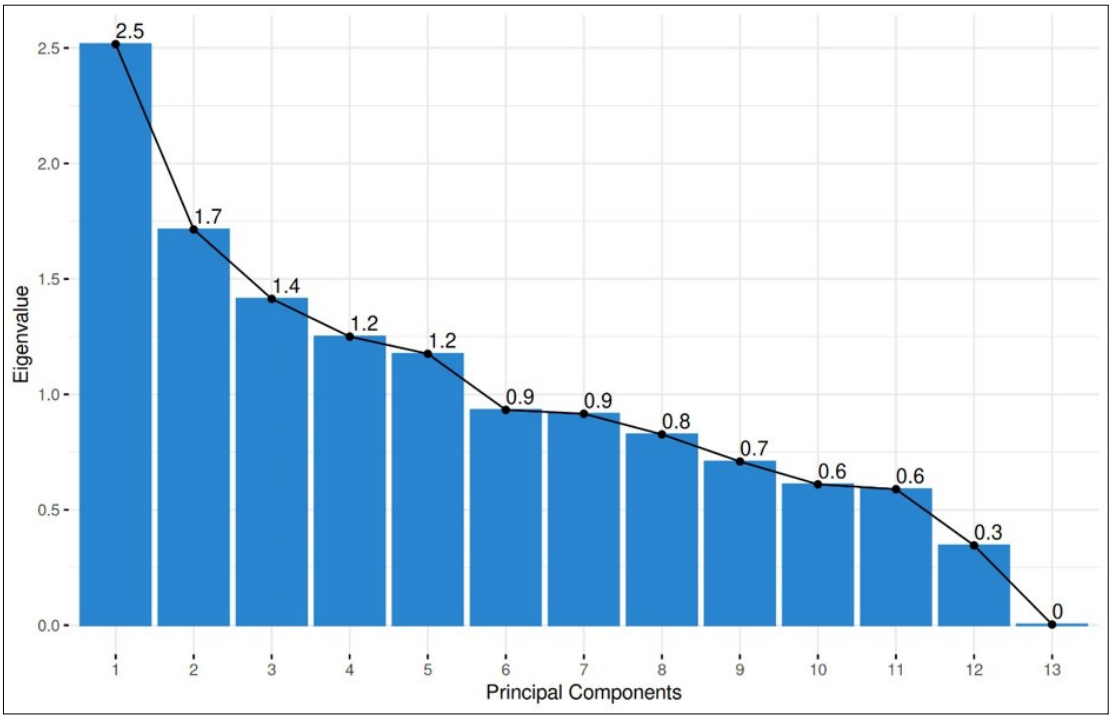


Fig. 6. Scree plot diagram of Eigen values constructed on thirteen traits recorded in 89 rice genotypes.

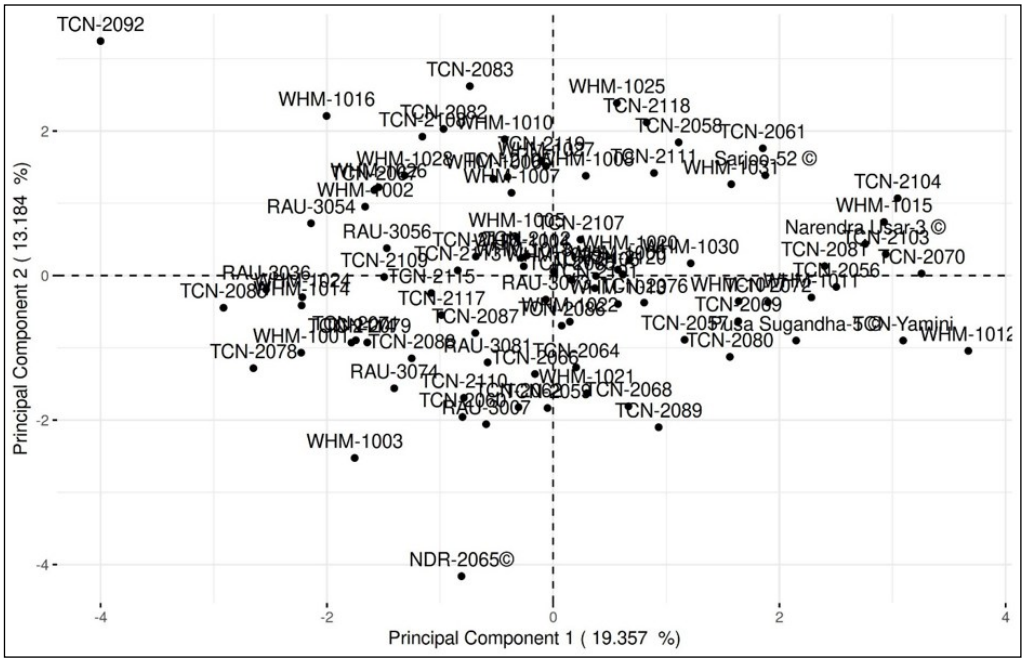


Fig.7. Biplot showing the distribution of 89 rice genotypes and associated traits along the principal component axes PC1 and PC2.

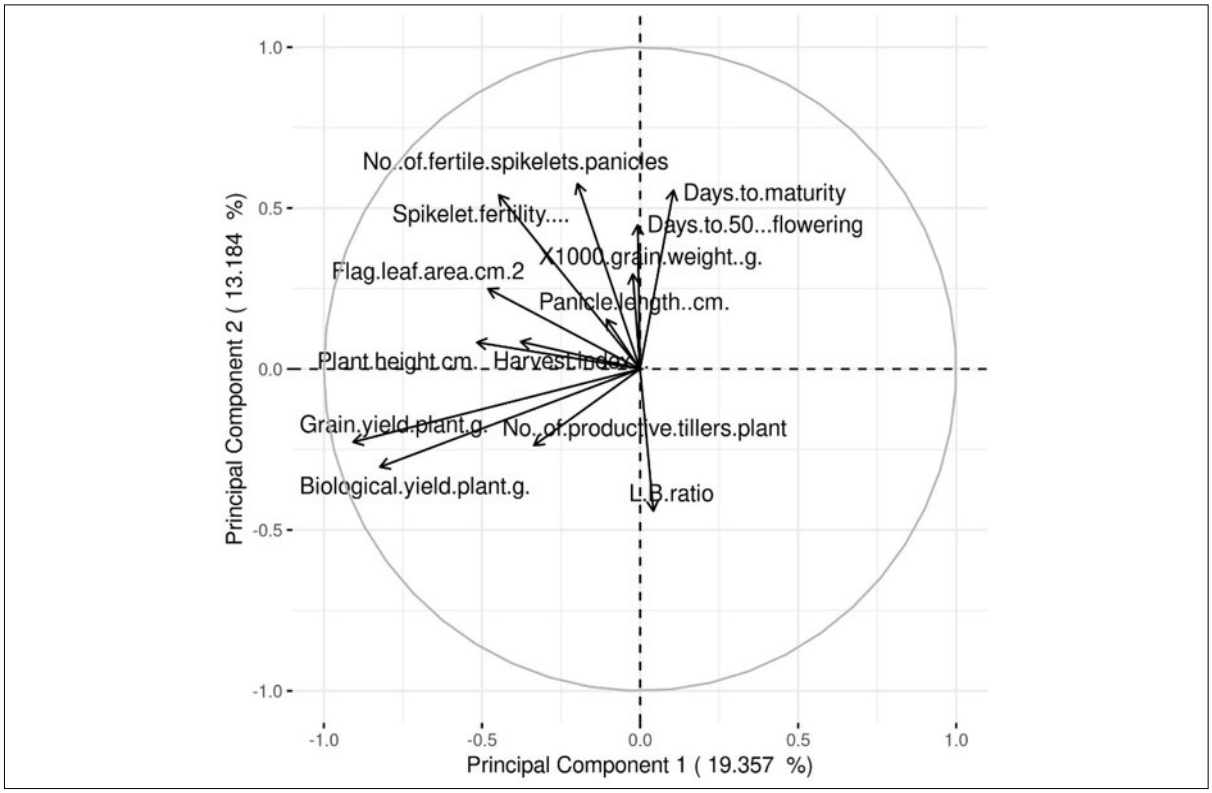


Fig. 8. Principal component biplot of 89 rice genotypes and 13 agronomic traits.

PC2. This plot reveals that phenological and fertility-related traits predominantly influence PC2, while yield-related traits define variation along PC1. This insight helps prioritise traits for genotype selection in breeding programs focused on yield and maturity.

The present findings correspond with those of earlier reports (7, 5, 29). Several researchers, have utilised PCA to characterise distinct rice germplasm lines (30, 31). The PCA aided in the identification of the traits that significantly influence the phenotypic of various landraces of rice, which is crucial for the breeding program's selection procedure.

Conclusion

The 89 rice genotypes assessed in this study under sodic soil conditions demonstrate a significant degree of genetic diversity. Key agronomic and yield-related traits showed significant variation, suggesting a high potential for development and selection. WHM-1024, WHM-1016 and TCN-2071 were the best for increasing yield because of their high biomass and grain yield. For short-duration cropping, TCN-2080 and NDR-2065 shown early blooming and maturity In traits including flag leaf area, biological yield, grain yield and L:B ratio, high heritability and significant genetic advance imply that they are mostly controlled by additive gene effects and are suitable to successful selection. Principal component and cluster analysis identified many genetic groups and significant factors influencing phenotypic variability, offering helpful information for hybridisation program parent selection. Genotypes such as WHM-1024, WHM-1016 and TCN-2071 showed potential candidates because of their high yield performance and adaptability. These genotypes may serve as potential donors for developing sodicity-tolerant, as well as high-yielding varieties. However, performance confirmation in replicated multi-location trials is necessary before varietal recommendation. Further solidifying the identification of elite lines was the incorporation of multivariate technologies such as PCA and Mahalanobis D^2 analysis. Ultimately, the results provide a strong basis for creating high-yielding, sodicity-tolerant rice cultivars that are appropriate for stressful conditions, promoting food security and sustainable rice cultivation.

Acknowledgements

The authors sincerely thank their colleagues for their valuable support in bringing this study to fruition.

Authors' contributions

PS coordinated the experiment, managed analytical work and organised data collection and led data organisation and manuscript drafting, with support from AS^1. SCV, PR contributed to the experimental planning, design, conceptual refinement and manuscript revision. AS^2 and AT oversaw the analytical process to ensure methodological rigor. UB and PS performed the statistical analyses using RStudio, which informed essential manuscript revisions and improved its clarity and accuracy. All authors have reviewed and approved the final version of the manuscript. [AS^1 stands for Avanji Saini and AS^2 stands for Aman Srivastava].

Compliance with ethical standards

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

Ethical issues: None

References

1. Ali Mekouar M. Food and agriculture organization of the United Nations (FAO). Yearb Int Environ Law. 2017;28:50620. https://doi.org/10.1093/yiel/yvy073
2. Giri D, Dhital M, Chaudhary B, Pandey R, Bastakoti B, Shrestha S. Effect of different nitrogen levels on yield and yield attributes of

different rice varieties in DDSR condition at Kanchanpur, Nepal. Arch Agric Environ Sci. 2022;7(3):310–7. https://doi.org/10.26832/24566632.2022.070302

3. Ashraf A. Identify the challenges of rice production and coping strategies in Narowal. Pak J Soc Res. 2023;5(4):10–5. https://doi.org/10.52567/pjsr.v5i04.1327
4. Maurya BK, Singh PK, Verma OP, Mandal DK. Genetic variability and divergence analysis in rice (*Oryza sativa* L.) under sodic soil. Int J Curr Microbiol Appl Sci. 2017;6(10):2865–9. https://doi.org/10.20546/ijcmas.2017.610.336
5. Kumar P, Sharma PK. Soil salinity and food security in India. Front Sustain Food Syst. 2020;4:533781. https://doi.org/10.3389/fsufs.2020.533781
6. Prakash S, Reddy SS, Chaudhary S, Vimal S, Kumar A. Multivariate analysis in rice (*Oryza sativa* L.) germplasms for yield attributing traits. Plant Sci Today. 2024;11(1):64–75. https://doi.org/10.14719/pst.2231
7. Burman M, Nair SK, Sarawgi AK. Principal component analysis for yield and its attributing traits in aromatic landraces of rice (*Oryza sativa* L.). Int J Bioresour Stress Manag. 2021;12(4):303–8. https://doi.org/10.23910/1.2021.2348a
8. Kumari N, Kumar R, Kumar A. Genetic variability and association of traits in mutant lines of rice (*Oryza sativa* L.) for submergence tolerance. Curr J Appl Sci Technol. 2019;33(2):1–7. https://doi.org/10.9734/cjast/2019/v33i230062
9. Shiragudikar SK, Bharamagoudar G, Manohara KK, Malati SY. Insight analysis of deep learning and a conventional standardized evaluation system for assessing rice crop's susceptibility to salt stress during the seedling stage. SN Comput Sci. 2023;4(3):262. https://doi.org/10.1007/s42979-022-01656-2
10. Federer WT, Raghavarao D. On augmented designs. Biometrics. 1975;31:29–35. https://doi.org/10.2307/2529707
11. Allard RW. Relationship between genetic diversity and consistency of performance in different environments. Crop Sci. 1961;1:127–33. https://doi.org/10.2135/cropsci1961.0011183X000100020012x
12. Robinson HF, Comstock RE, Harvey PH. Estimates of heritability and the degree of dominance in corn. Agron J. 1949;41:353–59. https://doi.org/10.2134/agronj1949.00021962004100080005x
13. Burton GW, Devane DE. Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. Agron J. 1953;45(10):478–81. https://doi.org/10.2134/agronj1953.00021962004500100005x
14. Beale EML, Lones PR, Cooley WW. Introduction to statistical procedures. Rev Int Stat Inst. 1969;37(3):324. https://doi.org/10.2307/1402126
15. Sparks DN. Algorithm AS 58: Euclidean cluster analysis. J R Stat Soc. 1973;22(1):126–30. https://doi.org/10.2307/2346321
16. Singh SK, Singh P, Korada M, Khaire AR, Singh DK, Habde SV, et al. Character association and path-coefficient analysis for yield and yield-related traits in 112 genotypes of rice (*Oryza sativa* L.). Curr J Appl Sci Technol. 2020;39(48):545–56. https://doi.org/10.9734/cjast/2020/v39i4831275
17. Lakshmi M, Shanmuganathan M, Jeyaprakash P, Ramesh T. Genetic variability and diversity analysis in selected rice (*Oryza sativa* L.) varieties. Electron J Plant Breed. 2022;13(3):959–66. https://doi.org/10.37992//2022.1303.124
18. Iftekharuddaula KM, Hassan MS, Islam MJ, Badshah MA, Islam MR, Akhter K. Genetic evaluation and selection criteria of hybrid rice in irrigated ecosystem of Bangladesh. Pak J Biol Sci. 2001;4(7):790–2. https://doi.org/10.3923/pjbs.2001.790.792
19. Shrivastav SP, Singh V, Lal K, Chaudhary AK, Hitaishi SK. Assessing genetic variability, heritability and genetic advance in rice (*Oryza sativa* L.) under sodic soil. Int J Curr Microbiol Appl Sci. 2020;9(2):3108–15. https://doi.org/10.20546/ijcmas.2020.902.359
20. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. Agron J. 1955;47:314–18. https://doi.org/10.2134/agronj1955.00021962004700070009x
21. Kato T. Selection responses for the characters related to yield sink capacity of rice. Crop Sci. 1997;37(5):1472–5. https://doi.org/10.2135/cropsci1997.0011183X003700050009x
22. Surek SH, Korkut KZ. Diallel analysis of some quantitative characters in F1 and F2 generations in rice (*Oryza sativa* L.). Egypt J Agric Res. 1998;76(2):651–63. https://doi.org/10.21608/ejar.1998.352470
23. Prasad KRK, Suneetha Y, Srinivas T. Genetic diversity studies in rice (*Oryza sativa* L.). Electron J Plant Breed. 2018;9(4):1335–41.https://doi.org/10.5958/0975-928X.2018.00166.7
24. Mishra A, Shahani M, Bijarania S. Genetic divergence studies of submergence tolerant rice genotypes. Int J Environ Clim Chang. 2023;13(3):195–99. https://doi.org/10.9734/ijecc/2023/v13i31697
25. Palaniyappan S, Arunachalam P, Banumathy S, Mini ML, Muthuramu S. Genetic divergence and clustering studies in advanced breeding lines of rice (*Oryza sativa* L.). Electron J Plant Breed. 2020;11(2):499–504. https://doi.org/10.37992/2020.1102.084
26. Shanmugasundaram P, Souframanien J, Sadasivam S. Genetic divergence among rice varieties released from paddy breeding station, Coimbatore, India. Oryza. 2000;37(3):225–7.
27. Liakat Ali M, McClung AM, Jia MH, Kimball JA, McCouch SR, Eizenga GC. A rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. Crop Sci. 2011;51(5):2021–35. https://doi.org/10.2135/cropsci2010.11.0641
28. Lalhruaitluangi B, Sharma MB, Ananda A, Bordoloi D. Yield and quality variants studies in Nagaland lines of Oryza sativa through principal component analysis and genetic divergence assessments. Indian J Plant Genet Resour. 2025;38(1):131–6.
29. Debsharma SK, Syed MA, Ali MH, Maniruzzaman S, Roy PR, Brestic M, et al. Harnessing genetic variability and diversity of rice (*Oryza sativa* L.) genotypes based on quantitative and qualitative traits for desirable crossing materials. Genes. 2022;14(1):10.https://doi.org/10.3390/genes14010010
30. Gana AS, Shaba SZ, Tsado EK. Principal component analysis of morphological traits in thirty-nine accessions of rice (*Oryza sativa* L.) grown in a rainfed lowland ecology of Nigeria. J Plant Breed Crop Sci. 2013;5(10):120–6. https://doi.org/10.5897/JPBCS12.065
31. Kashyap A, Yadav VK. Principal component analysis and character association for yield components in rice (*Oryza sativa* L.) genotypes of salt tolerance under alkaline condition. Int J Curr Microbiol Appl Sci. 2020;9:481–95. https://doi.org/10.20546/ijcmas.2020.910.059

Additional information

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

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

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

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

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.