Multivariate Studies On Diverse Rice (*Oryza Sativa* L.) Genotypes For Agro-Morphological Characters Under *Terai* Region Of West Bengal

Umamaheswar N[1], Kundu A[2], Roy S K[1*], Mandal R[3], Sen S[4], Hijam L[1], Chakraborty M[1], Das B[5], Barman R[6], Vishnupriya S[1], Thapa B[7], Maying B[8] and Rout S[9]

[1]Department of Genetics and Plant Breeding, Uttar Banga Krishi Viswavidyalaya, Cooch Behar, West Bengal– 736165, India

[2]AICRN on Potential Crops, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal– 736165, India

[3]Regional Research Station, Terai Zone, Uttar Banga Krishi Viswavidyalaya, Cooch Behar, West Bengal– 736165, India

[4]AINP on Jute and Allied Fibres, Uttar Banga Krishi Vishwavidyalaya, Cooch Behar, West Bengal– 736165, India

[5]Department of Genetics and Plant Breeding, Uttar Banga Krishi Vishwavidyalaya, College of Agriculture (Extended Campus), Majhian, Dakshin Dinajpur, West Bengal– 733133, India

[6]Regional Research Station (OAZ), Uttar Banga Krishi Vishwavidyalaya, Majhian, Dakshin Dinajpur, West Bengal– 733133, India

[7]Regional Research Station (Hill Zone), Uttar Banga Krishi Vishwavidyalaya, Kalimpong, West Bengal– 734301, India

[8]College of Agriculture, Central Agricultural University, Pasighat Arunachal Pradesh– 791102, India

[9]Department of Genetics and Plant Breeding, Centurion University of Technology and Management, Paralakhemundi, Odisha– 761211, India

ABSTRACT

A study was conducted to analyze trait variations among rice genotypes in the *Terai* region of West Bengal and to select high-performing genotypes based on specific characteristics. The study was conducted during the Kharif (*Aman*) seasons of 2019 and 2020, focussing on 42 rice genotypes at Uttar Banga Krishi Viswavidyalaya, Cooch Behar, West Bengal. The Mahalanobis D^2 analysis revealed four distinct clusters, with significant variation observed for grain length, plant height, and grain yield per plant. Additionally, the maximum Mahalanobis D^2 distance was observed for Dudeswar, Baramshall, and Khara. The Principal Component Analysis identified spikelet fertility, grain yield per plant, filled grains per plant, and test weight as principal discriminatory characteristics, with Dudeswar exhibiting the highest index score of 2.49. It was followed by Geetanjali with a score of 1.92, according to the Smith selection index. The significant characters identified through the D^2 analysis and PCA, such as grain length, plant height, spikelet fertility and others played a crucial role in revealing the diversity among the genotypes. The maximum D^2 distances for specific genotypes, coupled with high index scores, suggested a strong association with discriminatory characteristics identified through the Smith selection index, emphasizing their importance in genotypic classification and selection.

INTRODUCTION

Rice serves as a staple food for about more than three billion people (Zeigler and Adam, 2008). In order to maintain self-sufficiency, it is essential to develop new varieties or hybrids with high yield potential and resilience in challenging conditions (Papademetriou *et al*., 2000). There is a pressing requirement for novel rice varieties with greater genetic diversity, high yield, resilience to biotic and abiotic stresses, and superior grain quality to meet the needs of future consumers.

Genetic diversity, which refers to heritable variation within and between populations,

*Corresponding author's e-mail: suvendukumarroy@gmail.com

plays a crucial role in determining breeding strategies as it is a pre-requisite for the initiation of any crop-breeding programme. It is crucial to comprehend the available variability within the population being studied (Nachimuthu *et al.,* 2014). Due to their adaptability to diverse environments, traditional rice varieties can effectively cope with changing climates and serve as repositories of genes resistant to pests and diseases. Thus, employing traditional rice varieties would be the optimal and sustainable approach to create climate-smart rice varieties with resistance to significant biotic and abiotic stresses. Assessing genetic diversity is a critical element in developing effective crop improvement breeding programs. This evaluation also aids in establishing genetic relationships and estimating genetic variability during germplasm collection, by doing so, it helps to ensure that parental combinations in segregating populations have increased genetic variability, leading to the creation of new recombination for the selection and incorporation of desirable genes into superior germplasm (Thompson *et al.,* 1998; Islam *et al.,* 2012).

Multivariate analysis is one methodology for measuring genetic distance estimates for a population since it is important to recognise the useful variability present in the population (Nachimuthu *et al.,* 2014). It is commonly used to summarize and characterize the intrinsic diversity among genotypes.

Mahalanobis D^2 analysis of quantitative characters is a powerful tool for measuring genetic divergence among the material selected even from the same geographic region as reported by Mahalanobis (1936) followed by Rao (1952). A high level of genetic diversity helps the plant breeder in selecting genotypes having a desired specific character or a combination of characters. D^2 statistic remains as the most effective method for quantifying the degree of genetic diversity among genotypes. Analyzing D^2 values, breeders can identify groups with similar genetic characteristics and also assess the genetic diversity within and between groups or clusters, which is crucial for accurately selecting parental lines, leading to more effective exploration of heterosis, as emphasized by Murty and Arunachalam (1966).

As variation occurs often in plants for yield and yield-related characteristics (Maji *et al.,* 2012), Principal Component Analysis (PCA) identifies patterns and reduces redundancy in datasets. According to Anderson (1972) and Morrison (1978), PCA is a powerful and well-known multivariate statistical technique used for dimension reduction. It determines the smallest number of components that can explain the greatest amount of variability out of the total variability. Principal components (PCs) are commonly derived from either a covariance matrix or a correlation matrix. When variables are measured in different units, scale effects can change the composition of derived components, emphasizing the importance of standardizing the variables in these situations. The primary advantage of PCA lies in its ability to quantify the importance of each dimension in capturing the variability of a dataset, as highlighted by Shoba *et al.,* 2019.

In many breeding schemes, genotype selection is carried out entirely based on grain yield, neglecting other yield-determining characters in commercial breeding programmes. The application of selection indices, such as those initially suggested by Smith (1936) and Hazel (1943) known as the classical index, in a single study is an effective method for simultaneously and effectively including multiple characters. Smith (1936) contended that the genetic value could not be accurately assessed through individual characters alone, but rather it could be more effectively estimated through a linear combination of observable phenotypic values. Therefore, the application of a selection index would optimize the genetic improvement for intricate characteristics such as grain yield.

The purpose of the present study is to determine genetic diversity among rice genotypes in *terai* region of West Bengal based on D^2 analysis, PCA and Smith selection Index. Best combinations of these yield attributes and genotypes can be used as selection criteria for creating high yielding rice genotypes based on agro-morphological characters for future crop improvement programmes.

MATERIALS AND METHODS

The field experiment was conducted at the experimental farm of Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal, during Kharif, 2019 and 2020. The location receives high annual rainfall (3200 mm). Moreover, there is a wide distribution of rainfall coupled with high temperatures. It is located at geographical coordinates at 26°34′19″ N latitude, 88°08′51″ E longitude at an elevation of 113 meters above mean sea level (MSL). During the

crop growth period, the total amount of rainfall from June to November, 2019 was 2100 mm and during the second season i.e. June to November, 2020 was 3000 mm. The mean annual temperature is 30 °C and it has sandy loam soil type. A total of forty-two diverse rice genotypes were used for the study (Table 1.). The genotypes collected from two states, namely West Bengal and Andhra Pradesh, with wider adaptability in areas of their recommendation were used. The field trial was laid out in a Randomized Complete Block Design (RCBD) with two replications with a spacing of 20 cm × 15 cm row to row and plant to plant, respectively. The recommended package of practices was followed during the crop season to raise a good crop in the main field. Twenty-eight days seedlings were transplanted in the main field. The observations were recorded on randomly selected five competitive plants in the inner middle rows of each plot in all the two replications for nine morphological characters namely plant height (cm)- [PH], panicles plant^{-1}- [PPP], filled grains spikelet^{-1}- [FGPP], spikelet fertility (%)-[SF], grain length (mm)- [GL], grain breadth (mm)- [GB], grain length: breadth ratio (mm)- [LBR], test weight (g)- [TW] and grain yield plant^{-1} (g)- [GYP].

The chi-square test indicated that the rice genotypes were divergent and therefore the Mahalanobis D^2 (Mahalanobis 1928, 1936) analysis was carried out. GENSTAT software was used for D^2 analysis.

In order to classify the patterns of variation, principal component analysis (PCA) was performed. Those PCs with Eigen values greater than one were selected as proposed by Jeffers (1967). Correlations between the original characters and the respective Principal Components (PCs) were calculated. The mean data of the characters were used to perform principal component analysis (PCA) using software FactoMiner package (Lê *et al*., 2008) on a matrix of nine morphological characters followed by visualization by FactoExtra package in RStudio. GraphPad Prism 7 software (GraphPad Software Version 9.0, La Jolla California USA) was used for the visualization plot of proportion of variance.

Table 1. Details of the rice genotypes used in the experiment

Sl. No.	Code	Name of the Genotype	Sl. No.	Code	Name of the Genotype
1	G1	Balam	22)	G22	Khalia Eulo
2	G2	Baramshall	23)	G23	Kalonunia
3	G3	Baskathi	24)	G24	Khara
4	G4	Basmati	25)	G25	Lal Badsahbhog
5	G5	Kharadhan	26)	G26	Patnai
6	G6	Chamarmani	27)	G27	Sagar Sugandhi
7	G7	Chamatkar	28)	G28	Tulsi Mukul
8	G8	Dehradun Gandheswari	29)	G29	Nonabokra
9	G9	Dudeswar	30)	G30	BPT 2295
10	G10	Gopalbhog	31)	G31	BPT 5204
11	G11	Indulshall	32)	G32	CR 910
12	G12	Jhara	33)	G33	Geetanjali
13	G13	JP 90	34)	G34	NL 44
14	G14	JP 120	35)	G35	NL 46
15	G15	Zugal	36)	G36	NLR 0106
16	G16	Kakri	37)	G37	NLR 3242
17	G17	Kalavati	38)	G38	MTU 1061
18	G18	Kalo Aush	39)	G39	NLR 20084
19	G19	Kamal	40)	G40	NLR 40058
20	G20	Kanakchur	41)	G41	NLR 145
21	G21	Kerala Sundari	42)	G42	BPT 2411

The weights obtained from the eigen values were utilized in the construction of selection indices, following the approach outlined by Smith (1936). The weights assigned to each trait was determined using the PCA loading values, with a scale ranging from 1 to 10. Notably, grain yield received the highest weightage of 10 within this framework. Selection index for the recorded data was computed using the software PB Tools v. 1.4 (PB Tools, 2014).

RESULTS AND DISCUSSION

Mahalanobis D^2 statistic measured genetic divergence by clustering the genotypes into four clusters (Table 2.) based on nine morphological characters. The significant difference indicates the appropriateness of the use of D^2 statistics for clustering the genotypes into different groups. In this regard, Shanmugam et al. (2023), Singh et al. (2020), Pavani et al. (2018), Singh et al. (2017), Karuppaiyan et al. (2013), and Shanmugasundaram et al. (2000) identified four clusters for various numbers of genotypes in their study involving rice genotypes.

The genotypes were grouped into four clusters and they contained a variable number of genotypes. Cluster III contained the maximum number of 28 rice genotypes (Balam, Baramshall, Baskathi, Basmati, Kharadhan, Chamarmani, Chamatkar, Dehradun Gandheswari, Dudeswar, Gopalbhog, Indulshall, Jhara, JP 90, JP 120, Zugal, Kakri, Kalavati, Kalo Aush, Kamal, Kanakchur, Kerala Sundari, Khalia Eulo, Kalonunia, Khara, Lal Badsahbhog, Patnai, Sagar Sugandh and Tulsi Mukul) followed by 10 genotypes in Cluster IV (Nonabokra, BPT 2295, BPT 5204, Geetanjali, NLR 0106, NLR 3242, MTU 1061, NLR 20084, NLR 40058 and BPT 2411), two genotypes in Cluster I (NL 44 and NL 46) as well as Cluster II (CR 910 and NLR 145).

The clustering pattern of genotypes showed that the genotypes of different origins, collected from Uttar Banga Krishi Viswavidyalaya (Majhian Campus and Pundibari Campus) and Acharya N.G. Ranga Agricultural University (Bapatla Campus and Tirupati Campus) were clubbed in one cluster, whereas the genotypes belonging to same origin were grouped in different clusters indicating that the geographical distribution need not always be considered to be the sole criterion for genetic diversity. The genotypes included in the same cluster were considered genetically similar with respect to the aggregate effect of the characters examined. From the pattern of clustering, it could be inferred that sufficient divergence was present to enable the formation of individual clusters.

In the present investigation, the inter-cluster and intra-cluster distance was estimated among the nine

Table 2. Grouping of 42 rice genotypes into different clusters on the basis of D^2 analysis for nine morphological characters (Combined over 2 years)

Cluster No.	Total no. of germplasm accessions	Source	Name of germplasm accessions
I	2	A1	(G34) and (G35)
II	2	A2	(G32) and (G41)
III	28	B	(G1), (G2), (G3), (G4), (G5), (G6), (G7), (G8), (G9), (G10), (G11), (G12), (G13), (G14), (G15), (G16), (G17), (G18), (G19), (G20), (G21), (G22), (G23), (G24), (G25), (G26), (G27) and (G28)
IV	10	A1+A3+B	(G29), (G30), (G31), (G33), (G36), (G37), (G38), (G39), (G40) and (G42)

A1- Acharya N.G. Ranga Agricultural University (Tirupati Campus), A2- Acharya N.G. Ranga Agricultural University (Agricultural Research Station, Nellore), A3- Acharya N.G. Ranga Agricultural University (Bapatla), B- Uttar Banga Krishi Viswavidyalaya (Majhian Campus and Pundibari Campus).

Table 3. Average intra (diagonal) and inter-cluster (off-diagonal) D² values of 42 rice genotypes (Combined over 2 years)

Cluster	I	II	III	IV
I	100.02	2403.40	10262.47	2798.98
II		145.19	6592.36	2243.97
III			1978.97	7248.52
IV				3222.97

characters (Table 3). Members of cluster IV were more dissimilar in morphological features and performance than those of other clusters, as indicated by the highest value of the intra-cluster distance (3222.974). The maximum intra cluster distance was recorded in cluster VI (3222.974) followed by cluster III (1978.974), cluster II (145.187), and cluster I (100.023), indicating greater genetic divergence between the genotypes in these clusters. The maximum intra cluster distance in cluster VI was because of wide genetic diversity among its genotypes.

The maximum inter cluster distance was observed between cluster III and I (10262.474) followed by cluster IV and III (7248.518), cluster III and II (6592.362), cluster IV and IV (3222.974), cluster IV and I (2798.984), cluster II and I (2403.396), cluster IV and II (2243.974) and cluster III and III (1978.974). So, genotypes can be selected as parents between cluster III and I because of maximum inter cluster genetic distance.

The larger inter-cluster distances indicated more diversity among the rice genotypes grouped in different clusters with respect to the characters considered for hybridization programme in rice. The estimates of average intra and inter cluster distance value of four clusters revealed that the genotypes belonging to the same cluster (intra cluster) have less genetic divergence as compared to genetic diversity between the genotypes of different clusters (inter cluster). When crossing is done between genotypes belonging to the same cluster, no transgressive segregants are expected from such combinations because same cluster genotypes display the lowest degree of divergence from one another. Therefore, a hybridization programme should always be formulated in such a way that parents belonging to different clusters with maximum genetic distance can be utilized to obtain desirable transgressive segregants.

The cluster means for various characters are presented in Table 4 which showed that each cluster had its own uniqueness that separated it from the other clusters. Cluster I was characterized by highest means for filled grains spikelet⁻¹ (1.965) and lowest for grain length: breadth ratio (0.460). Cluster II, consisting of only two genotypes, was characterized by the highest value for plant height (2.023) and the lowest for grain breadth (0.412). Cluster III had the highest value for plant height (2.178) and lowest for grain breadth (0.543). Cluster IV had the lowest value for grain breadth (0.468) and highest for filled grains spikelet⁻¹ (2.030). Cluster mean analysis indicated the extent of diversity among different clusters, which can be of practical value in rice breeding.

Table 4. Cluster means for nine characters of rice genotypes (Combined over 2 years)

| Cluster	PH	PPP	FGPP	SF	GL	GB	LBR	TW	GYP	Total
I	1.81	0.91	1.97	1.93	0.81	0.60	0.46	1.29	1.21	10.97
II	2.02	0.92	2.00	1.92	0.74	0.41	0.58	1.16	1.27	11.03
III	2.18	1.07	2.04	1.91	0.87	0.54	0.56	1.19	1.24	11.60
IV	1.95	1.05	2.03	1.93	0.81	0.47	0.59	1.22	1.32	11.37
Population Mean	2.10	1.05	2.03	1.92	0.85	0.52	0.56	1.20	1.26	11.49
PC (%)	27.18	2.67	2.09	0.00	31.48	2.32	0.12	8.48	25.67	100.01

PH - Plant height (cm), PPP - Panicles plant⁻¹, FGPP - Filled grains spikelet⁻¹, SF - Spikelet fertility (%), GL - Grain length (mm), GB - Grain breadth (mm), LBR - Grain length: breadth ratio (mm), TW - Test weight (g) and GYP - Grain yield plant⁻¹ (g), PC – percent contribution.

The cluster means ranged from 1.81 (Cluster I) to 2.18 (Cluster III) for plant height (cm); 0.91 (Cluster I) to 1.07 (Cluster III) for panicles plant^{-1}; 1.97 (Cluster I) to 2.04 (Cluster III) for filled grains spikelet^{-1}; 1.91 (Cluster III) to 1.93 (Cluster I and Cluster IV) for spikelet fertility; 0.74 (Cluster II) to 0.81 (Cluster I) for grain length; 0.41 (Cluster II) to 0.60 (Cluster I) for grain breadth; 0.46 (Cluster I) to 0.59 (Cluster IV) for grain length: breadth ratio; 1.16 (Cluster II) to 1.29 (Cluster I) for test weight; 1.21 (Cluster I) to 1.32 (Cluster IV) for grain yield plant^{-1} in the present study.

This implies that the grain yield plant^{-1} showed a consistent range across all the clusters like cluster IV (1.322), cluster II (1.270), cluster III (1.241), and cluster I (1.213). This uniformity in the range of grain yield across the clusters suggest the presence of divergent genotypes within each cluster, indicating a broad spectrum of genetic diversity within the population. Thus, the clustering pattern could be utilized in selection of parents for crossing and deciding the best cross combinations, which may generate the highest possible variability for various characters.

The contribution to divergence (Figure 1..) has been the maximum by grain length (31.48 %), followed by plant height (27.18 %) and grain yield plant^{-1} (25.67). In contrast, remaining characters contributed very little towards genetic divergence i.e., test weight (8.48 %), panicles plant^{-1} (2.67), grain breadth (2.32), filled grain panicle^{-1} (2.09) and grain length: breadth ratio (0.12). The interesting point in percentage contribution is that spikelet fertility (0.00) showed no contribution towards genetic divergence.

Figure 1. The percent contribution of nine characters for combined over the years
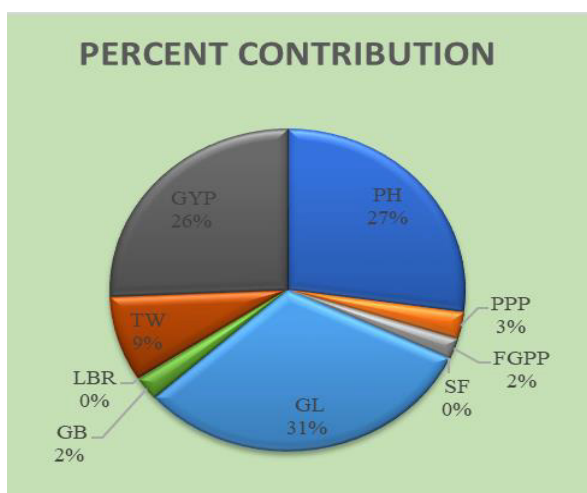
Using D^2 statistics, all the rice genotypes could be distinguished from one another considering all the characters collectively. The results suggested that some genotypes performed better than others used in this investigation over the years. A suitable crossing programme may be justifiable to exploit genetic divergence in characters such as grain length, plant height, and grain yield plant^{-1}. This can be based on the percentage contribution of these characters and the identification of rice genotypes, namely Dudeswar, Baramshall, and Khara, through intra-cluster distance, inter-cluster distance, and the D^2 distance between individual genotypes.

PCA was conducted using yield and yield attributes on diverse rice genotypes. It has been proposed that the genetic variation among the rice accessions, as indicated by their agro-morphological characters, should be considered to advance the improvement programme (Shahidullah *et al.*, 2009). Out of the nine, only four principal components possessed more value than 1.0 eigen value and showed about 75% of total variability among the characters studied. Summary of the contribution of the principal components to variability are given in Table 5.

To choose the variable parents, the main components with multiple eigen values demonstrated greater variation among the rice genotypes. So, these four principal components were given more importance. The PC1 shared high proportion of total variation of 25.86 % and the rest of the three principal components viz., PC2, PC3 and PC4 contributed 18.48%, 17.16% and 15.16% of the total variance respectively. Similar findings were also reported by Sar and Kole (2023), Mushtaq *el al.* (2023), Shanmugam *et al.* (2023), Pushpa *et al.* (2022), Akhtar *et al.* (2022), Dhakal *et al.* (2020), Tuhina-Khatun (2015) and Shoba *et al.* (2010).

Scree plot (Figure 2.) was plotted for illustrating the percentage of variance to each principal component, with PC1 exhibiting the highest variation (25.86%) than the other principal components. Thus, selecting characteristics based on PC1 would be effective.

Characters like grain length: breadth ratio, grain length, and grain breadth had relatively longer vector suggesting that the characters had relatively larger effects on grain yield (Figure 3.). On the contrary,

Table 5. Summary of the contribution of the principal components to variability

Components	Combined over the years			
	Standard deviation	Eigen Value	Proportion of Variance	Cumulative Proportion
PC1	1.533	2.350	26.11	26.11
PC2	1.291	1.668	18.53	44.64
PC3	1.241	1.539	17.11	61.74
PC4	1.155	1.335	14.83	76.57

Figure 2. Scree plot for illustrating the percentage of variance to each principal component

Proportion of variance
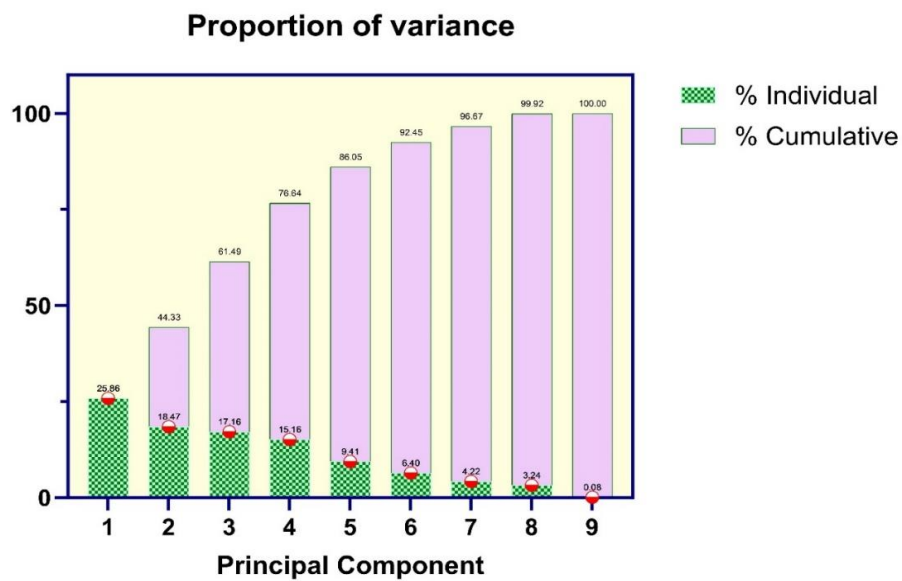
Figure 3. Variables of PCA
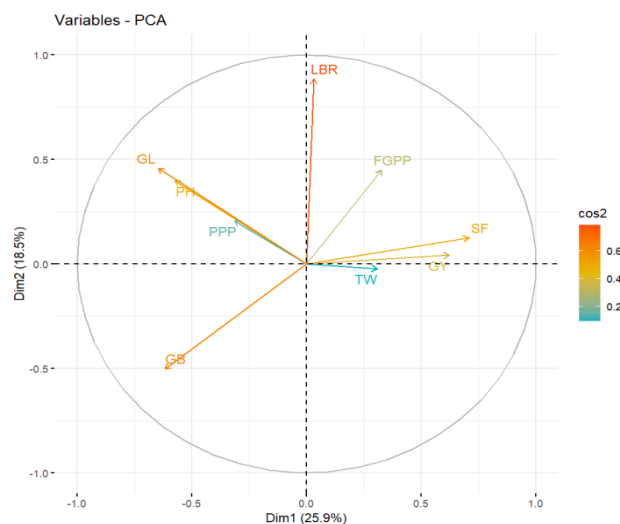
test weight and panicles plant⁻¹ had shorter vector length, indicating little association with grain yield. The cosine of the angles between the vectors of the two characters measures their similarity relative to their effect on grain yield. Characteristics such as filled grains spikelet⁻¹, spikelet fertility and grain length: breadth ratio vectors were positively strong correlation with grain yield plant⁻¹ because the degree

between two lines is acute angle (<90°) on PC1 but grain breadth shows extreme negative correlation with grain yield plant[-1] on PC1 as indicated by obtuse angle (>90°). Characters namely grain length, plant height and panicles plant[-1] displayed very close correlation among themselves along with negligible positive correlation while test weight had minimal negative correlation with grain yield plant[-1] because they form an acute angle (<90°) on PC2.

Percent contribution of variables on principal components is given in Table 6. The findings demonstrated that in PC1, spikelet fertility (0.460) and grain yield plant[-1] (0.407) had the highest positive values in PC1. While in PC2, grain length: breadth ratio had the highest positive score (0.606), followed by filled grains spikelet[-1] (0.386), grain length and plant height (0.377). Surprisingly, only grain length: breadth ratio had the positive value in PC3. In PC4, filled grains spikelet[-1] (0.560) had highest positive value followed by panicles plant[-1] (0.429). From these findings, specific selection strategies can be formulated to enhance characters such as grain yield. The agro-morphological diversity and variability among rice genotypes are pivotal for crop improvement (Seetharam *et al.*, 2009).

Biplot is the merger of PCA score plot and loading plot. The covariate effect of biplot based on correlation among the characters is presented in Figure 5 and the same explained 44% of the total variation and thus can be considered as a good approximation, as far as the effect of characters on yield as well as their similarities were concerned. Because the first two principal components (PC1 and PC2) contain the majority of

Table 6. Four principal components along with their factor loadings for combined over the years

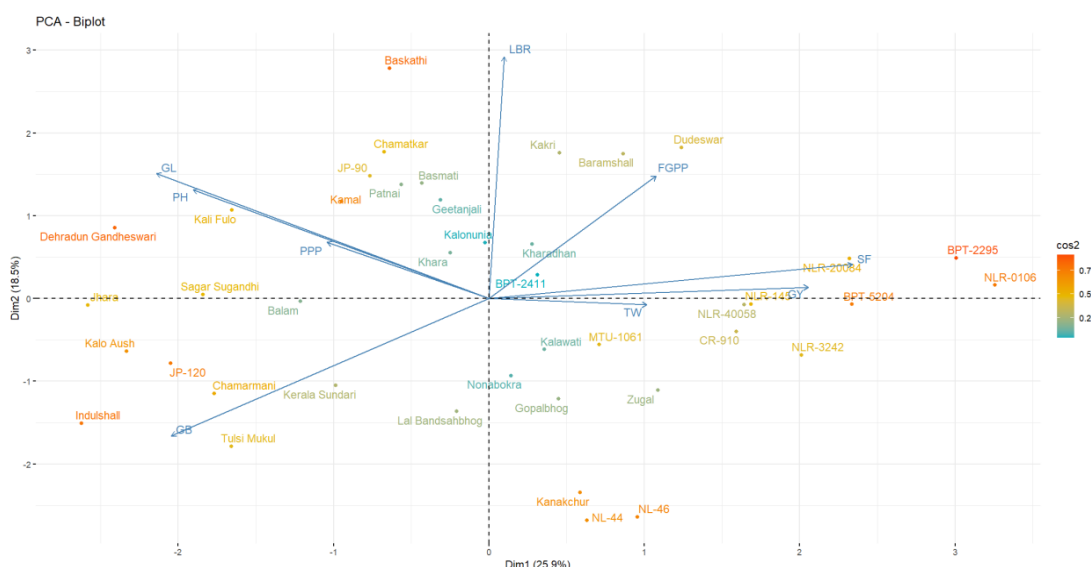| Characters	Combined over the years			
	PC1	PC2	PC3	PC4
Plant height (cm)	-0.371	0.377	-0.151	0.247
Panicles plant[-1]	-0.180	0.287	-0.059	0.429
Filled grains panicle[-1]	0.232	0.386	-0.050	0.560
Spikelet fertility (%)	0.460	0.084	-0.038	0.243
Grain length (mm)	-0.424	0.377	-0.185	-0.280
Grain breadth (mm)	-0.418	-0.267	-0.469	0.155
Grain L/B ratio	0.040	0.606	0.321	-0.405
Test weight (g)	0.208	0.137	-0.615	-0.325
Grain yield (g plant[-1])	0.407	0.144	-0.483	-0.106

Figure 4. Biplot of PCA

Table 7. Values of smith selection index for all the rice genotypes (Combined over the years)

Genotype	PH	PPP	FGPP	SF	GL	GB	LBR	TW	GYP	Smith index
1. Balam	2.22	1.26	2.01	1.93	0.90	0.60	0.53	1.25	1.29	0.63
2. Baramshall	2.12	0.91	2.28	1.94	0.90	0.51	0.62	1.16	1.19	1.36
3. Baskathi	2.22	1.10	2.09	1.94	0.95	0.49	0.69	1.16	1.19	1.29
4. Basmati	2.22	1.07	1.92	1.91	0.94	0.51	0.65	1.35	1.37	1.09
5. Kharadhan	2.16	1.17	2.16	1.93	0.86	0.53	0.56	1.13	1.37	1.38
6. Chamarmani	2.20	0.92	1.86	1.92	0.91	0.61	0.52	1.13	1.18	-1.69
7. Chamatkar	2.18	0.95	2.06	1.89	0.89	0.46	0.67	1.16	1.20	0.07
8. Dehradun Gandheswari	2.20	1.11	2.02	1.91	0.97	0.59	0.60	1.11	1.15	-0.34
9. Dudeswar	2.18	1.16	2.28	1.93	0.85	0.49	0.59	1.17	1.39	2.49
10. Gopalbhog	2.18	0.89	2.06	1.91	0.84	0.59	0.48	1.27	1.42	0.35
11. Indulshall	2.11	1.13	1.83	1.86	0.89	0.60	0.52	1.2	1.18	-2.32
12. Jhara	2.19	1.31	2.07	1.90	0.92	0.61	0.52	1.17	1.17	-0.30
13. JP 90	2.19	1.09	2.06	1.90	0.84	0.45	0.65	1.13	1.14	-0.32
14. JP 120	2.18	1.14	1.93	1.88	0.9	0.59	0.53	1.26	1.22	-0.99
15. Zugal	2.20	0.94	2.10	1.94	0.82	0.59	0.47	1.34	1.38	0.85
16. Kakri	2.26	1.08	2.29	1.94	0.85	0.51	0.58	1.20	1.22	1.59
17. Kalavati	2.19	0.89	1.92	1.93	0.83	0.53	0.54	1.26	1.26	-0.51
18. Kalo Aush	2.15	1.17	1.91	1.87	0.86	0.54	0.54	1.11	1.14	-2.05
19. Kamal	2.15	1.14	2.01	1.90	0.90	0.49	0.63	1.17	1.21	0.12
20. Kanakchur	2.10	1.03	1.98	1.93	0.76	0.60	0.42	1.19	1.39	-0.56
21. Kerala Sundari	2.17	1.04	2.17	1.92	0.84	0.61	0.47	1.14	1.15	-0.77
22. Khalia Eulo	2.18	0.91	2.00	1.89	0.94	0.52	0.64	1.18	1.16	-0.66
23. Kalonunia	2.20	1.29	2.08	1.94	0.79	0.47	0.55	1.17	1.17	0.26
24. Khara	2.20	1.10	2.07	1.90	0.85	0.50	0.58	1.28	1.30	0.62
25. Lal Badsahbhog	2.12	1.04	2.04	1.88	0.73	0.51	0.48	1.09	1.30	-1.32
26. Patnai	2.23	0.93	1.99	1.90	0.95	0.52	0.65	1.36	1.34	1.01
27. Sagar Sugandhi	2.13	0.95	1.85	1.90	0.95	0.56	0.61	1.16	1.17	-1.36
28. Tulsi Mukul	2.16	1.12	2.07	1.90	0.78	0.60	0.43	0.97	1.07	-2.41
29. Nonabokra	1.87	0.86	1.90	1.92	0.85	0.50	0.58	1.01	1.16	-1.93
30. BPT 2295	1.98	0.97	2.19	1.95	0.79	0.45	0.58	1.23	1.40	1.89
31. BPT 5204	1.93	1.01	2.02	1.95	0.79	0.45	0.59	1.18	1.34	0.75
32. CR 910	2.04	0.91	1.97	1.92	0.73	0.41	0.58	1.13	1.28	-0.63
33. Geetanjali	1.93	1.13	2.04	1.91	1.02	0.58	0.64	1.26	1.42	1.92
34. NL 44	1.82	0.86	1.95	1.92	0.82	0.59	0.46	1.29	1.22	-1.51
35. NL 46	1.80	0.96	1.98	1.93	0.8	0.59	0.45	1.29	1.21	-1.19
36. NLR 0106	1.91	1.03	1.98	1.94	0.72	0.38	0.61	1.20	1.42	1.03
37. NLR 3242	2.02	1.10	2.09	1.92	0.79	0.51	0.52	1.37	1.42	1.30
38. MTU 1061	2.04	1.07	1.99	1.92	0.84	0.52	0.55	1.28	1.31	0.25
39. NLR 20084	1.96	1.15	2.11	1.91	0.78	0.43	0.60	1.32	1.43	1.74
40. NLR 40058	1.85	1.06	1.99	1.94	0.78	0.43	0.60	1.17	1.16	-0.36
41. NLR 145	2.01	0.94	2.02	1.92	0.75	0.42	0.59	1.19	1.25	-0.18
42. BPT 2411	2.03	1.12	2.02	1.91	0.79	0.44	0.60	1.14	1.17	-0.57

PH - Plant height (cm), PPP - Panicles plant^{-1}, FGPP - Filled grains spikelet^{-1}, SF - Spikelet fertility (%), GL - Grain length (mm), GB - Grain breadth (mm), LBR - Grain length: breadth ratio (mm), TW - Test weight (g) and GYP - Grain yield plant^{-1} (g).

Table 8. Ranking of the 42 rice genotypes based on the Smith index for each year and combined over the years basis

Genotype	2019		2020		Combined over the years	
	S.I.	Rank	S.I.	Rank	S.I.	Rank
1. Balam	0.76	12	0.68	5	0.63	15
2. Baramshall	1.34	7	0.64	6	1.36	7
3. Baskathi	0.80	11	0.54	7	1.29	9
4. Basmati	0.36	16	0.70	4	1.09	10
5. Kharadhan	1.07	9	0.30	13	1.38	6
6. Chamarmani	-1.39	36	0.07	20	-1.69	38
7. Chamatkar	-0.70	30	-0.03	24	0.07	21
8. Dehradun Gandheswari	-0.77	31	0.47	9	-0.34	25
9. Dudeswar	2.40	1	0.37	10	2.49	1
10. Gopalbhog	1.04	10	0.32	12	0.35	17
11. Indulshall	-2.13	40	-0.48	33	-2.32	41
12. Jhara	0.19	21	0.34	11	-0.30	23
13. JP 90	-0.62	29	-0.43	32	-0.32	24
14. JP 120	-1.00	34	0.20	16	-0.99	33
15. Zugal	1.65	6	0.72	3	0.85	13
16. Kakri	1.97	2	0.54	8	1.59	5
17. Kalavati	-0.19	25	0.03	22	-0.51	27
18. Kalo Aush	-2.45	42	-0.55	34	-2.05	40
19. Kamal	-0.19	26	-0.09	26	0.12	20
20. Kanakchur	-0.15	24	-0.05	25	-0.56	28
21. Kerala Sundari	0.20	20	0.13	17	-0.77	32
22. Khalia Eulo	-1.05	35	0.10	18	-0.66	31
23. Kalonunia	0.37	15	-0.16	28	0.26	18
24. Khara	0.35	17	0.27	14	0.62	16
25. Lal Badsahbhog	-1.95	39	-0.69	40	-1.32	35
26. Patnai	0.51	13	0.82	2	1.01	12
27. Sagar Sugandhi	-1.74	37	0.03	23	-1.36	36
28. Tulsi Mukul	-1.78	38	-0.65	37	-2.41	42
29. Nonabokra	-2.21	41	-0.97	42	-1.93	39
30. BPT 2295	1.81	3	0.10	19	1.89	3
31. BPT 5204	0.10	22	-0.16	29	0.75	14
32. CR 910	-0.90	33	-0.94	41	-0.63	30
33. Geetanjali	1.30	8	0.88	1	1.92	2
34. NL 44	0.02	23	-0.40	31	-1.51	37
35. NL 46	0.32	18	-0.30	30	-1.19	34
36. NLR 0106	0.32	19	-0.63	35	1.03	11
37. NLR 3242	1.77	4	0.25	15	1.30	8
38. MTU 1061	0.46	14	0.06	21	0.25	19
39. NLR 20084	1.70	5	-0.09	27	1.74	4
40. NLR 40058	-0.51	28	-0.63	36	-0.36	26
41. NLR 145	-0.23	27	-0.66	39	-0.18	22
42. BPT 2411	-0.84	32	-0.65	38	-0.57	29

S.I - Smith index

Table 9. Ranking of the best 20% rice genotypes on the basis of Smith index

Year	Genotypes	Mean values for the characters									Smith index	Rank
		PH	PPP	FGPP	SF	GL	GB	LBR	TW	GY		
2019	G9	2.18	1.16	2.28	1.93	0.85	0.49	0.60	1.21	1.46	2.4	1
	G16	2.26	1.10	2.30	1.94	0.85	0.51	0.58	1.23	1.24	1.97	2
	G30	1.99	0.95	2.20	1.95	0.79	0.46	0.58	1.26	1.43	1.81	3
	G37	2.02	1.13	2.13	1.92	0.79	0.52	0.52	1.37	1.45	1.77	4
	G39	1.96	1.19	2.17	1.90	0.78	0.45	0.59	1.35	1.47	1.70	5
	G15	2.20	0.95	2.11	1.94	0.82	0.60	0.47	1.38	1.39	1.65	6
	G2	2.12	0.94	2.29	1.94	0.9	0.52	0.60	1.17	1.20	1.34	7
	G33	1.94	1.14	2.05	1.91	1.02	0.58	0.64	1.28	1.45	1.30	8
	G5	2.17	1.20	2.18	1.92	0.86	0.54	0.55	1.16	1.38	1.07	9
	MSI	2.09	1.08	2.19	1.93	0.85	0.52	0.57	1.27	1.39		
	MAI	2.10	1.06	2.05	1.91	0.85	0.53	0.56	1.22	1.29		
	SDi	-0.01	0.02	0.14	0.02	0.01	-0.10	0.01	0.05	0.10		
	EGG	0.02	0.04	0.11	0.02	0.01	-0.10	0.01	0.07	0.10		
2020	G33	1.93	1.12	2.03	1.92	1.02	0.57	0.65	1.24	1.38	0.88	1
	G26	2.23	0.91	1.98	1.91	0.95	0.51	0.65	1.34	1.31	0.82	2
	G15	2.20	0.93	2.09	1.94	0.81	0.58	0.47	1.30	1.36	0.72	3
	G4	2.21	1.06	1.91	1.91	0.94	0.51	0.65	1.34	1.34	0.70	4
	G1	2.21	1.25	1.99	1.93	0.90	0.59	0.53	1.23	1.28	0.68	5
	G2	2.11	0.88	2.28	1.95	0.90	0.50	0.63	1.15	1.18	0.64	6
	G16	2.26	1.06	2.28	1.94	0.85	0.50	0.58	1.17	1.20	0.54	7
	G3	2.22	1.08	2.06	1.94	0.95	0.48	0.68	1.14	1.17	0.54	8
	G8	2.19	1.09	2.01	1.92	0.97	0.58	0.61	1.10	1.14	0.47	9
	MSI	2.17	1.04	2.07	1.93	0.92	0.54	0.61	1.22	1.26		
	MAI	2.10	1.03	2.01	1.92	0.85	0.51	0.57	1.17	1.23		
	SDi	0.08	0.01	0.06	0.01	0.07	0.02	0.04	0.05	0.03		
	EGG	0.11	0.04	0.07	0	0.06	0.03	0.02	0.07	0.05		
Combined over the years	G9	2.18	1.16	2.28	1.93	0.85	0.49	0.59	1.17	1.39	2.49	1
	G33	1.93	1.13	2.04	1.91	1.02	0.58	0.64	1.26	1.42	1.92	2
	G30	1.98	0.97	2.19	1.95	0.79	0.45	0.58	1.23	1.4	1.89	3
	G39	1.96	1.15	2.11	1.91	0.78	0.43	0.6	1.32	1.43	1.74	4
	G16	2.26	1.08	2.29	1.94	0.85	0.51	0.58	1.2	1.22	1.59	5
	G5	2.16	1.17	2.16	1.93	0.86	0.53	0.56	1.13	1.37	1.38	6
	G2	2.12	0.91	2.28	1.94	0.9	0.51	0.62	1.16	1.19	1.36	7
	G37	2.02	1.1	2.09	1.92	0.79	0.51	0.52	1.37	1.42	1.3	8
	G3	2.22	1.1	2.09	1.94	0.95	0.49	0.69	1.16	1.19	1.29	9
	MSI	2.09	1.08	2.17	1.93	0.86	0.5	0.6	1.22	1.34		
	MAI	2.10	1.05	2.03	1.92	0.85	0.52	0.56	1.2	1.26		
	SDi	-0.10	0.04	0.14	0.02	0.02	-0.10	0.03	0.02	0.08		
	EGG	0.03	0.05	0.10	0.01	0.01	-0.10	0.03	0.06	0.10		

PH - Plant height (cm), PPP - Panicles plant[-1], FGPP - Filled grains spikelet[-1], SF - Spikelet fertility (%), GL - Grain length (mm), GB - Grain breadth (mm), LBR - Grain length: breadth ratio (mm), TW - Test weight (g) and GYP - Grain yield plant[-1] (g), MSI- Mean of Selected Individuals, MAI- Mean of all individuals, SDi: Selection differential, EGG: Expected genetic gain.

the variance, a biplot was constructed in the present study to explore the relationship among the 45 rice genotypes based on their observed yield and yield components. The top right corner of the biplot between PC1 and PC2 revealed a group of genotypes including Dudeswar, NLR 20084, BPT 2295, Kakri, Baramshall, and NLR 0106. These genotypes displayed positive values for both principal components and characters such as filled grains spikelet^{-1}, grain yield plant^{-1}, spikelet fertility, and grain length: breadth ratio (mm), indicating their occupation of the same quadrant and influence on grain yield. Similar findings are observed by Dhakal *et al.* (2020) for filled grains spikelet^{-1} and grain yield plant^{-1}, and Shanmugam *et al.* (2023) for grain length: breadth ratio and spikelet fertility.

In numerous breeding programmes, genotypes are typically selected based solely on grain yield. However, a plant's economic value is contingent upon its various characteristics. Therefore, it is crucial for plant breeders to concurrently consider the selection of multiple characteristics to optimize the economic value of a plant. The use of a selection index (Smith, 1936) aids in computing these characteristics to facilitate the development of an optimal genotype. The calculated index scores for all forty-two genotypes grown under *terai* region varied from -2.41 (Tulsi Mukul) to 2.49 (Dehradun Gandheswari) in combined over year basis. The top nine genotypes were chosen based on their high selection index scores, with Dudeswar, an optimal genotype, achieving the highest rank under the study. Similar results are reported by Pavithra *et al.* (2020) under drought environment and Venmuhil *et al*. (2020) for various characteristics in rice. Sabouri *et al.* (2008) and Habib *et al*. (2007) reported greater genetic improvements through selection based on multiple characters compared to selection based on a single trait in rice.

The values of smith selection index for all the genotypes are given for combined over the years is given in Table 7, list of the genotypes with rank based on smith index value for each year and combined over the years is given in Table 8 and the best nine genotypes based on smith index for each year and combined over the years is given in Table 9.

CONCLUSION

Based on the findings of the present study, it is evident that the D^2 analysis has successfully categorized the rice genotypes into four distinct clusters, each exhibiting significant variation in grain length, plant height, and grain yield per plant. This suggests that these genotypes, representing different clusters, hold potential as donors for enhancing various agronomic characters through hybridization programs. By incorporating these beneficial characters into modern high-yielding varieties, it becomes possible to develop new varieties with improved adaptability and resilience. Furthermore, the principal component analysis highlights spikelet fertility, grain yield per plant, filled grains per plant, and test weight as the principal discriminatory characteristics. This emphasizes their importance in influencing the overall performance and suitability of the rice genotypes. Additionally, the findings from the smith selection index for multiple characters has identified that genotypes, Dudeswar and Geetanjali are found to be the best performing genotypes, as evidenced by their highest index scores. Furthermore, utilizing indexing through PCA-selected characters could enhance the dependability of the selection process. These conclusions provide valuable insights and implications for future research and agricultural applications, particularly in the selection and breeding of rice varieties. The identified genotypes and their associated characters can serve as valuable resources for the development of improved rice varieties, ultimately contributing to the advancement of the agricultural sector.

ACKNOWLEDGMENT

Ethics Statement:

There was no human participants and or/animal included in this research.

Consent for publication:

All the authors agree to publish the content.

Competing interests:

The authors declare that there is no conflict of interest for publishing this content.

Authors contribution

N. Umamaheswar: Conducted the field experiment along with collection and analysis of data and drafting manuscript; A. Kundu: Co-supervisor, guided in planning and implementation of research work and data analysis; S.K. Roy: Supervisor, guided in planning and implementation of research work; S. Sen: Helped in implementation of research work; L. Hijam: Guided in planning and implementation of research work; M. Chakraborty: Guided in layout of the field experiment and writing the manuscript; B. Das: Helped in collection of rice genotypes from different parts of West Bengal; R. Barman: Helped in collection of rice genotypes from different parts of West Bengal and Andhra Pradesh; Vishnupriya S: Helped in data collection; B. Thapa: Helped in data compilation; S. Rout: Helped in writing the manuscript B. Maying: Helped in writing the manuscript.

REFERENCES

Akhtar, R., Iqbal, A. and T. Dasgupta. 2022. Genetic diversity analysis of aromatic rice (*Oryza sativa* L.) germplasm based on agro-morphological characterization. *Oryza,* **59**(2): 141-149. https://doi.org/10.35709/ory.2022.59.2

Dhakal, A., Pokhrel, A., Sharma S. and A. Poudel. 2020. Multivariate analysis of phenotypic diversity of rice (*Oryza sativa* L.) landraces from Lamjung and Tanahun Districts, Nepal. *Int. J. Agron.,* 1-8. https://doi.org/10.1155/2020/886 7961

Habib, S. H., Iftekharuddaula, K. M., Bashar, M. K., Akter, K. and M. K. Hossain. 2007. Genetic variation, correlation and selection indices in advanced breeding lines of rice (*Oryza sativa* L.). *Bangladesh J. Plant Breed. Genet.,* **20**(1): 25-32. https://doi.org/10.3329/bjpbg.v20i1.17024

Hazel, L.N. 1943. The Genetic Basis for Constructing Selection Indexes. *Genetics,* **28**(6): 476-90. https://doi.org/10.1093/genetics/28.6.476

Jeffer, J. N. 1967. Two case studies in the application of principal component analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* **16**(3): 225-36. https://doi.org/10.2307/2985919

Kang, M. S. 2015. Efficient SAS Programs for Computing Path Coefficients and Index Weights for Selection Indices. *Journal Crop Improvement.,* **29**: 6-22. https://doi.org/10. 1080/15427528.2014.959628

Lakshmi, V.I., Sreedhar, M., Vanisri, S., Anantha, M. S.,

Rao L. L. S. and C. Gireesh. 2019. Multivariate analysis and selection criteria for identification of African rice (*Oryza glaberrima* L.) for genetic improvement of indica rice cultivars. *Plant Genetic Resources,* **17**(6): 499-505. https://doi.org/10.1017/S147926211900 0327

Mahalanobis, P. C. 1936. On the generalized distance in statistics, in *Proceedings of the National Institute of Science of India*, Calcutta, India, **2**: 49-55. http://hdl. handle.net/10263/6765.

Maji, A. T. and A. A. Shaibu. 2012. Application of principal component analysis for rice germplasm characterization and evaluation. *J. Plant Breed. Crop Sci.,* **4**(6): 87-93. http://doi.org/10.5897/JPBCS11.093

Morrison, D.E. 1978. Multivariate Statistical Methods (2nd ed. 4th Print, McGraw Hill Kogakusta Ltd. https://lib.ugent.be/catalog/rug01:000004122

Mushtaq, D. and Kumar, B. 2022. Multivariate analysis among advanced breeding lines of rice (Oryza sativa L.) Under sub-tropical ecology of Jammu region of Jammu and Kashmir. Electronic Journal of Plant Breeding, **13**(4): 1387-1394. https://doi.org/10.37992/2022.1304.179

Nachimuthu, V. V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S. and S. Manonmani. 2014. Evaluation of rice genetic diversity and variability in a population panel by principal component analysis. *Indian J. Sci. Technol.,* **7(**10): 1555-1562. https://dx.doi.org/10.17485/ijst/2014/v7i10.14

Papademetriou, M. K., Dent, F. J. and E. M. Herath. 2000. Rice production in the Asia-Pacific region: issues and perspectives. Bridging the rice yield gap in the Asia-Pacific region. Food and agriculture organization of the United Nations. Regional office for Asia and the pacific, Bangkok, Thailand, **220**: 4-25. https://www.fao.org/3/x6905e/x6905e.pdf

Pavani, M., Sundaram, R. M., Ramesham M. S., Kishor P. B. K. and K. B. Kemparaju. 2018. Prediction of heterosis in rice based on divergence of morphological and molecular markers. *J. Genet.*, **97**: 1263-79. https://doi.org/10.1007/s12 041-018-1023-8

Pushpa, R., Sassikumar, D., Suresh, R., and Iyanar, K. (2022). Evaluation of nutritional and grain quality diversity in rice (Oryza sativa L.) germplasm

based on multivariate analysis. Electronic Journal of Plant Breeding, **13**(4), 1187-1197. https://doi.org/10.37992/2022.1304.171

R Studio and Inc. shiny. 2013. Web Application Framework for R. R package version 0.5.0. https://doi.org/10.32614/CRAN.package.shiny

Sar, P. and P. C. Kole. 2023. Principal component and cluster analyses for assessing agro-morphological diversity in rice. *Oryza*, **60**(1): 117-24. https://doi.org/10.35709/ ory.2023.60.1.2

Seetharam, K., Thirumeni, S. and Paramasvum, K. 2009. Estimation of genetic diversity in rice (*Oryza sativa* L.) genotypes using SSR markers and morphological characters. *African J. Biotechnol.*, **8**: 2050-2059. https://doi.org/10.5897/AJB2009.000-9275

Shahidullah, S. M., Hanafi, M. M., Ashrafuzzaman, M., Ismail, M. R. and M. A. Salam. 2009. Phenological characters and genetic divergence in aromatic rices. *African J. Biotechnol.,* **8**: 3199-3207. https://doi.org/10.5897/AJB09.686

Shanmugam, A., Suresh, R., Ramanathan, A., Anandhi, P. and D. Sassikumar. 2023. Unravelling genetic diversity of South Indian rice landraces based on yield and its components. Electronic Journal of Plant Breeding, **14**(1): 160-9. https://doi.org/10.37992/2023.1401.007

Shoba, D., Vijayan, R., Robin, S., Manivannan, N., Iyanar, K., Arunachalam, P., Nadarajan, N., Pillai, M. A. and S. Geetha. 2019. Assessment of genetic diversity in aromatic rice (*Oryza sativa* L.) germplasm using PCA and cluster analysis. *Electron. J. Plant Breed.,* **10**(3): 1095-104. https://doi.org/10.5958/0975-928X.2019.00140.6

Singh, M., Singh, S. K., Vennela, P. R., Singh, D. K. and D. Kumar. 2017. Genetic divergence studies for drought tolerance in rice (*Oryza sativa* L.) using morphological traits and molecular markers. *Oryza*, **54**(4): 385-91. https://doi.org/10.5958/2249-5266.2017.00053.4

Singh, S. K., Pandey, V., Mounika, K., Singh, D. K., Khaire, A. R., Habde, S., & Majhi, P. K. (2020). Study of genetic divergence in rice (Oryza sativa L.) genotypes with high grain zinc using Mahalanobis' D2 analysis. Electronic Journal of Plant Breeding, **11**(02), 367-372. https://doi.org/10.37992/2020.1102.065

Smith, H. F. 1936. A Discriminant Function for Plant Selection. *Annals Eugen.,* **7**, 240-50. https://doi.org/10.1111/j.1469-1809.1936.tb02143.x

Sparks, D.N. 1973. Euclidean cluster analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* **22**(1): 126-30. https://doi.org/10.2307/2346321

Tuhina-Khatun, M., Hanafi, M. M., Rafii Yusop, M., Wong, M. Y., Salleh, F. M. and J. Ferdous. 2015. Genetic variation, heritability, and diversity analysis of upland rice (*Oryza sativa* L.) genotypes based on quantitative traits. *Biomed Res. Int.,* 1-7. http://dx.doi.org/10.1155/2015/290861

Venmuhil, R., Sassikumar, D., Vanniarajan, C. and R. Indirani. 2020. Selection indices for improving the selection efficiency of rice genotypes using grain quality traits. *Electron. J. Plant Breed.,* **11**(2): 543-549. https://doi.org/10.37992/2020.1102.091

Zeigler, R.S., and B. Adam. 2008. The relevance of rice. *Rice*, **1**: 3-10. https://doi.org /10.1007/ s12284-008-9001-z