

RESEARCH ARTICLE

Yield and Quality Variants Studies in Nagaland lines of *Oryza sativa* through Principal Component Analysis and Genetic Divergence Assessments

B Lalhruaitluangi¹, Malini Borthakur Sharma¹, Ashwini Ananda^{2*}, Dibosh Bordoloi³

Abstract

An experiment was conducted in the field to assess the existing variability in rice germplasm based on D² statistics for 29 quantitative traits. The whole germplasm was classified into 5 clusters, in which the highest inter-cluster distance was recorded between cluster III and cluster II. Cluster I showed the highest mean value for germination percentage, plant height, panicle weight, spikelet fertility, root length, total nitrogen, crude protein, grain N%, NutE, 100-grain weight, and grain yield per plant. It has been found that cluster II showed the highest mean value for days to 50% flowering, days to maturity, panicles per plant, stem dry weight, number of ear-bearing tillers, harvest index, chlorophyll a, and grain yield per plant and cluster III shown highest mean value for root dry weight, amylose content, PNUE, biological yield, and NHI. Cluster IV recorded the highest mean value for panicle length, chlorophyll B, and total chlorophyll, whereas cluster V recorded the highest mean value for flag leaf length, flag leaf breadth and flag leaf area. Principal component analysis revealed a total variability of 34.4% among genotypes contributed by PC₁ (18.2%) and PC₂ (16.2%). Chlorophyll a and total chlorophyll (7.41%) contributed maximum towards the observed diversity. Thupfu Lha and RCM from cluster I and Taposen Youli from cluster III have desirable characteristics, viz., grain N%, PNUE, NutE, BY, and NHI and may be selected for intercrossing. Hence, simultaneous selection of these traits is advised for genetic improvement in the rice breeding program.

Keywords: D² statistics, Euclidean distance, Genetic divergence, Principal component analysis, Contribution%.

¹Department of Genetics & Plant Breeding, School of Agricultural Sciences, Nagaland University, Nagaland, India.

²Department of Horticulture, School of Agricultural Sciences, Nagaland University, Nagaland, India.

³Subject Matter Specialist (Genetics & Plant Breeding), Karimganj, AAU-Jorhat, India.

***Author for correspondence:**

ashwiniagrian1992@gmail.com

Received: 22/04/2024 **Revised:** 13/08/2024

Accepted: 28/11/2024

How to cite this article: Lalhruaitluangi B, MB Sharma, A Ananda and D Bordoloi (2025) Yield and Quality Variants Studies in Nagaland lines of *Oryza sativa* through Principal Component Analysis and Genetic Divergence Assessments. *Indian J. Plant Genet. Resour.* 38(1),131-136.**DOI:** 10.61949/0976-1926.2025.v38i01.14

Introduction

In Asia, rice serves as a plate filler crop since it is a staple food for millions of families. According to Bollinedi *et al.* (2020), the per capita consumption of rice in the rice-dependent country ranges from 62 to 192 kg annually, accounting for roughly 20% of per capita energy and 13% of dietary protein. Many breeders are drawn to improving the quantity and quality of rice grains due to market demand. India is estimated to produce 122.27 million tonnes of rice this year, down from 118.87 million tonnes in 2019–2020 (economictimes.indiatimes.com, 2022). Breeders primarily aim to enhance yield and integrate innovative breeding strategies amidst the ever-changing climate, which necessitates the discovery of new resistance genes for both biotic and abiotic stresses from untapped genetic diversity. Genetic diversity within genotypes can emerge from geographical isolation or reproductive barriers. Varied germplasm collections serve as a vital foundation for creating improved crop varieties. It is crucial to categorize or classify genotypes according to an appropriate scale of divergence to achieve superior offspring. Devi *et al.* (2019) noted that the greatest distance between clusters signifies extensive diversity,

whereas the smallest distance points to a tight linkage among the groups. Evaluating the proportionate influence of individual traits on overall diversity aids in understanding polygenic variation. Principal component analysis (PCA), as described by Nachimuthu *et al.* (2014), is a multivariate, nonparametric method used to scrutinize a dataset composed of multiple interrelated quantitative variables that characterize observations, with the goal of distilling significant insights. For quantification of genetic divergence, various biometrical protocols have been developed that can be used in isolating genetically different parents to carry out an effective hybridization plan (Uddin *et al.*, 2014). Therefore, the assessment of genetic diversity is crucial for pinpointing genetic resources to be utilized in hybridization programs (Amin *et al.*, 2014). The extensive genetic variation within genetic material presents an excellent opportunity for obtaining diverse traits in subsequent generations, thereby enhancing crop improvement. Selecting parents based on genetic divergence holds significant importance for optimizing hybridization programs (Kumar and Gurumurthi, 2000). Mahalanobis D^2 statistic, which is based on multivariate analysis of quantitative traits, is a powerful tool for measuring the divergence among set populations. Therefore, an attempt was made to study the multivariate analysis of genetic divergence in rice.

Materials and Methods

The experiment was conducted during *Kharif* 2021 and *Kharif* 2022 at the experimental farm, Department of GPB, SAS, Nagaland University. A set of 28 germplasm of rice from SARS, Mokokchung, Nagaland and ICAR for NEH region, Jharnapani, Nagaland was used in the present study and were sown in a randomized block design (RBD) with three replications. About 29 agronomic attributes, namely, germination% (GER), days to 50% flowering (D50), days to maturity (DM), plant height (PH), flag leaf length (FLL), flag leaf breadth (FLB), Flag leaf area (FLA), no. of ear bearing tillers (NEBT), panicles/plant (PPP), panicle length (PL), panicle weight (PW), spikelet fertility (SF), root length (RL), root dry weight (RDW), stem dry weight (SDW), harvest index (HI), total nitrogen (TN), crude protein (CP), chlorophyll a (CHL A), chlorophyll b (CHL B), total chlorophyll (TCHL), amylose content (AC), grain N% (GN%), PNUE, NutE, biological yield (BY), NHI, 100 grain weight (HSW), grain yield/plant (GYPP). The data were analyzed in conformity with the multivariate analysis of genetic divergence using D^2 statistics by Mahalanobis (1936). The average inter- and intracluster D^2 mean values were estimated according to the procedure given by Singh and Chaudhary (1977), and the grouping of genotypes was done following Tocher's method, while the methodology followed for principal component analysis was done using the statistical package R with the help of Agricola package (De Mendiburu, 2014).

Results & Discussion

The 28 germplasm were grouped into 5 clusters based on Tocher's method (Table 1 and Figure 1). Among all the clusters, cluster V is the largest having maximum number of genotypes consisting of 12 germplasm namely, Thangmo Red, Ongpangsuk, Sulijak, Thangma White, Longkhum Tsuk (SASRS-2), Chahashye, Apuapa (SARS-61), Kezie (SASRS-94), Sungmangtsuk (SARS-1), Korea Tsuk, Chali and Chishoghi while Cluster II consist of 9 germplasm viz. Moyatsuk, Kedayishye, Moya Chali, Ngoni, Pfukhi Lha, Tungo, Tsungmiki, Manen Red (SARS-5) and Rosho Lha. There were two genotypes (Thupfu Lha and RCM-9) and four genotypes (Yarba (SARS-3), Shyekenyii, Tsushvuri, and Amusu) in Clusters I and IV, respectively. The sole genotype in the remaining cluster III is Taposen Youli, as shown in Table 1. The formation of distinct, solitary clusters may have occurred because their ancestor's geographic barriers prevented gene flow, or intensive natural and human selection for diverse and adaptable gene complexes must be the cause of this genetic diversity. According to the study of genetic divergence among 28 germplasm, genetic drift and natural selection under different environmental conditions may cause greater diversity than geographical distance. It is clear that genotypes from different ecogeographic areas were grouped into the same cluster, while genotypes from the same geographic origin were included in different clusters, implying that geographic diversity does not necessarily represent genetic diversity (Dwivedi and Mitra, 1995). The inter-cluster D^2 – values varied from 6.48 to 10.00, where maximum inter-cluster distance (10.00) was noticed between cluster III and cluster II, showing the selection of the divergent parents from these clusters will yield good segregation for the traits of interest (Chippy *et al.*, 2021). However, the minimum inter-cluster distance (6.48) was noticed between cluster I and cluster V (Table 2). The criterion used for the selection of genotypes as parents for crop improvement programs using D^2 analysis is the inter-cluster distance. Those genotypes included in clusters with maximum inter-cluster distance are obviously genetically more divergent. Hence, it would be logical to incorporate genotypes from these clusters in the breeding program or selection for new genotypes.

A perusal of the cluster means for different traits indicated that the magnitude of differences among genotypes was significant, as shown in Table 3. The lowest cluster mean for days to flowering was observed in cluster V (108.14), which indicated the genotypes have early flowering. Cluster III has having highest amylose content (10.58), PNUE (13.31), biological yield (32.14), and nitrogen harvest index (13.81). The highest cluster mean for panicle length (28.71), chlorophyll b (33.92), and total chlorophyll (30.29) were recorded in cluster IV. However, maximum days to maturity (165.58) and days to 50% flowering (142.05) were observed in

Table 1: Grouping of genotypes into different clusters

Cluster	No. of genotypes	Name of genotypes
Cluster I	2	Thupfu Lha and RCM-9
Cluster II	9	Moyatsuk, Kedayishye, Moya Chali, Ngoni, Pfukhi Lha, Tungo, Tsungmiki, Manen Red (SARS-5) and Rosho Lha
Cluster III	1	Taposen Youli
Cluster IV	4	Yarba (SARS-3), Shyekenyii, Tsushvuri and Amusu
Cluster V	12	Thangmo Red, Ongpangsuk, Sulijak, Thangma White, Longkhum Tsuk (SASRS-2), Chahashye, Apuapa (SARS-61), Kezie (SASRS-94), Sungmangtsuk (SARS-1), Korea Tsuk, Chali and Chishoghi

Table 2: Average intra (underlined) and inter-cluster distances among 28 genotypes

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Cluster I	<u>6.46</u>	9.12	7.96	7.53	6.48
Cluster II		<u>6.45</u>	10.00	8.45	9.03
Cluster III			<u>7.30</u>	8.47	7.37
Cluster IV				<u>6.15</u>	7.08
Cluster V					<u>5.24</u>

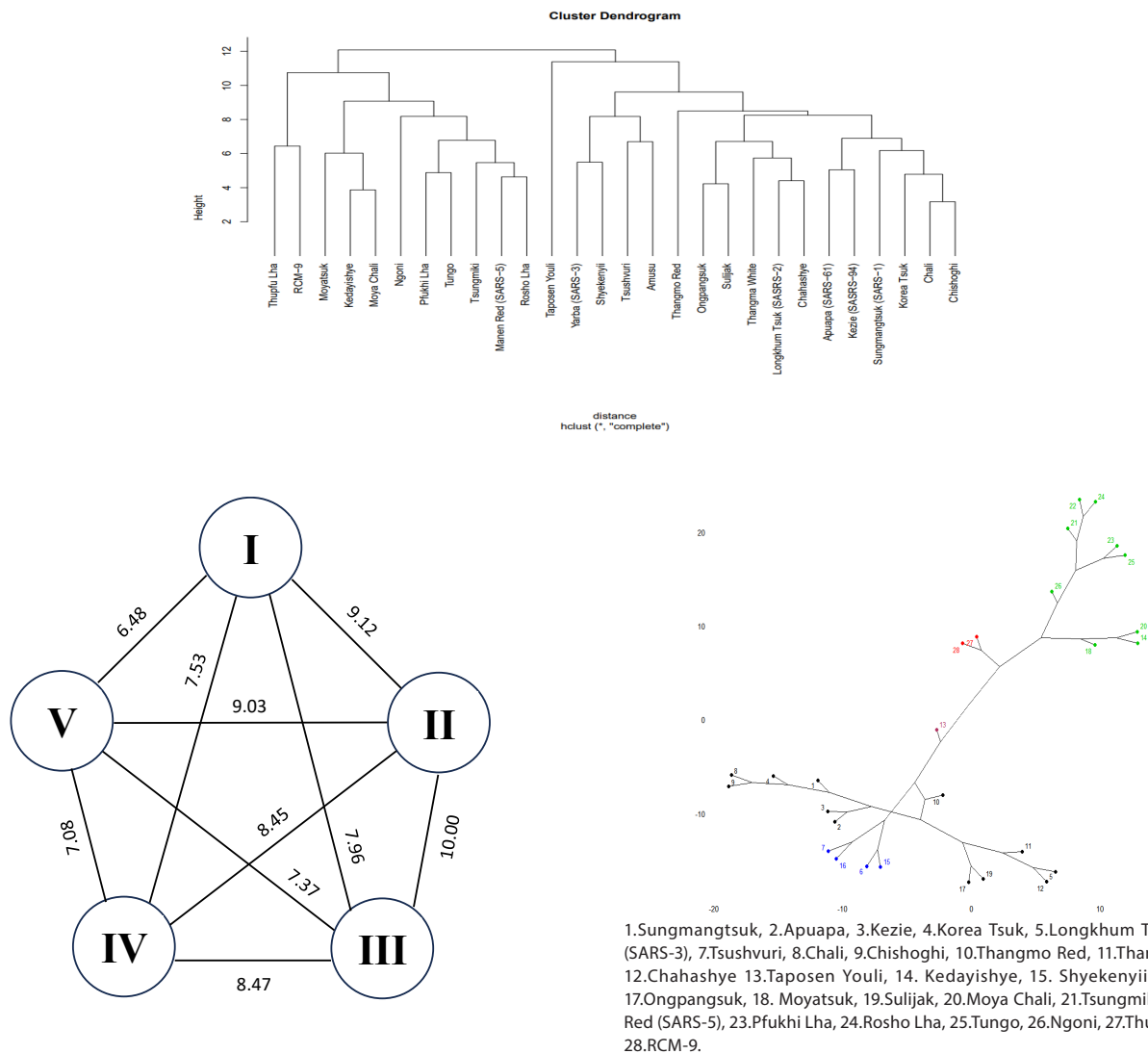


Figure 2: Inter-cluster distance among five groups of rice germplasm distances of 29 quantitative traits

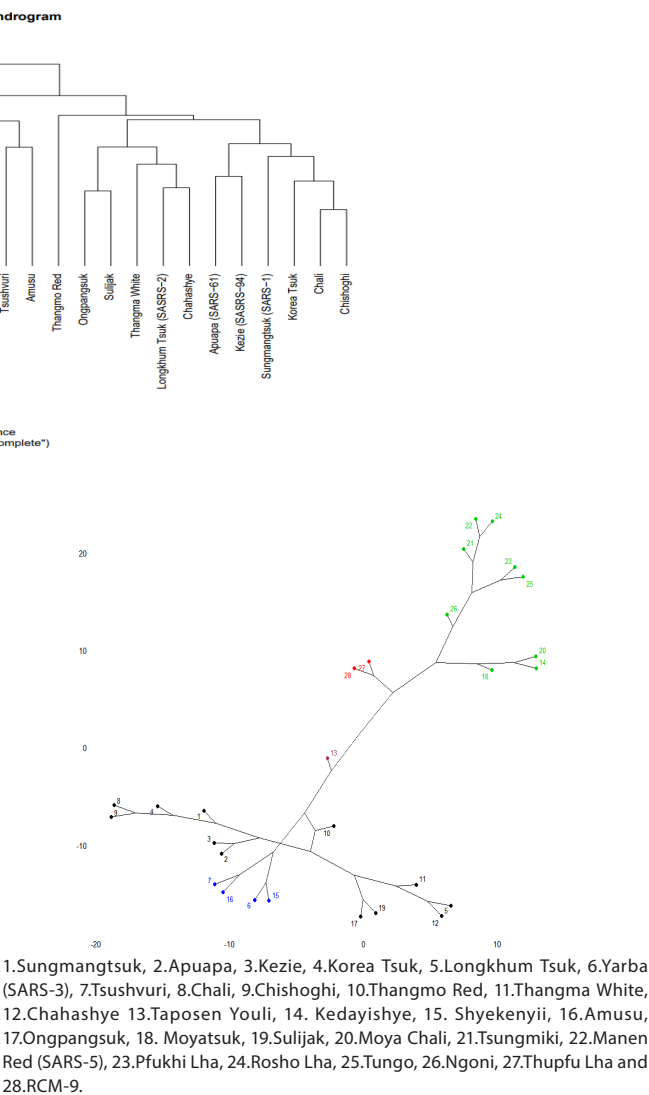


Figure 3: Cluster tree showing the clustering pattern of 28 genotypes of rice germplasm

cluster II, which indicated the genotypes have late flowering as well as late harvesting. Moreover, cluster I have recorded maximum cluster mean value for germination percentage (83.49), plant height (135.43), panicle weight (4.58), spikelet fertility (82.37), root length (27.91), total nitrogen (2.31), crude protein (6.40), grain N% (1.85), NutE (8.21), 100-grain weight (5.02) and grain yield per plant (5.23) and minimum for days to maturity (135.19). Hence, better recombinants can be obtained by selecting genotypes with their phenotypic dissimilarity. Thus, it is quite clear that the genotypes in cluster I are exceptionally divergent from the remaining genotypes of other clusters. As a result, crossing genotypes from this cluster to remaining genotypes in other clusters may be recommended for the best utilization of heterosis for different economic traits and isolation of transgressive segregants in the rice improvement program.

The principal component analysis was used to divide the combined variance into principal components in order to select the best germplasm based on the average value of different traits. The principal factor PC is a powerful tool that provides flexibility in obtaining suitable parental genotypes to design useful breeding programs (Nazir *et al.*, 2013). The 1st principle component (PC₁) contributed maximum towards total variability (18.2%). GER, D50, DM, FLL, FLA, PW, HI, TN, CP, CHL A, CHL B, TCHL, GN%, HSW, GYPP were positively loaded while PH, NEBT, PPP, PL, SF, RL, RDW, SDW, AC, PNUE, NutE, BY were negatively loaded (Figure 4). The second principal component (PC₂) contributed 16.2% towards total variability. The traits including GER, flag leaf length (FLL), flag leaf breadth (FLB), FLA, PW, TN, CP, HSW, GYPP, PPP, PL, RL, RDW, SDW and PNUE were positively loaded D50, DM, HI, CHL A, CHL B, TCHL, GN%, PH, NEBT, SF, AC, NutE and BY negatively loaded (Figure 4). The PC bi-plot shows the variables distributed and the gap within characters concerning PC₁ and PC₂ exhibited the contribution of these traits in creating variation among the germplasm (Figure 4).

Among the 29 quantitative traits, chlorophyll a and total chlorophyll (7.41%) contributed maximum towards the observed diversity (Table 4), followed by flag leaf area, chlorophyll b, flag leaf length, PNUE, flag leaf breadth, crude protein, NHI, days to maturity, amylose content, root dry weight, number of ear bearing tillers, harvest index and grain N %.

Therefore, these traits should be emphasised during hybridization and selection in the population that is segregating. Plant height, panicles per plant, days to 50% flowering, panicle weight, stem dry weight, NutE, panicle length, spikelet fertility, root length, total nitrogen, grain yield per plant, and 100-grain weight are the traits that contribute to less divergence. According to Reddy *et al.*, 2020, the least genetic diversity was contributed by grain yield per plant, plant height, panicles per plant, and panicle length. Singh and Singh (2003) also noted that among

Table 3: Mean performance of different clusters among 29 quantitative traits

	GER	D50	DM	PH	FLL	FLB	FLA	NEBT	PPP	PL	PW	SF	RL	RDW	SDW	HI	TN	CP	CHL A	CHL B	TCHL	AC	GN%	PNUE	NutE	BY	NHI	HSW	GYPP
I	83.49	114.21	135.19	135.43	37.03	1.96	55.03	14.52	9.91	27.56	4.58	82.37	27.91	6.75	14.58	29.78	2.31	6.40	12.32	21.85	19.39	7.99	1.85	12.30	8.21	24.50	12.38	5.02	5.23
II	78.33	142.05	165.58	91.09	31.06	1.67	45.81	15.56	10.53	22.63	4.13	77.01	27.28	6.11	15.52	38.96	1.91	6.24	24.03	31.61	29.98	7.26	1.84	12.23	7.94	23.92	12.54	4.19	5.23
III	54.54	108.31	135.95	128.36	28.26	2.01	42.86	14.69	9.99	26.76	3.30	78.44	27.75	7.20	15.23	25.71	2.02	6.07	11.90	20.93	18.57	10.58	1.62	13.31	8.20	32.14	13.81	3.57	4.23
IV	67.22	109.47	136.28	122.89	31.64	2.03	48.64	14.13	9.52	28.71	3.92	70.82	27.33	6.76	13.86	37.30	2.24	6.27	23.59	33.92	30.29	8.07	1.75	12.17	7.91	24.23	12.04	4.23	4.58
V	74.37	108.14	135.34	116.94	38.96	2.17	63.35	14.09	9.87	26.84	3.81	75.25	26.79	6.78	13.59	26.50	1.96	6.24	12.44	21.06	18.73	7.62	1.75	12.15	7.63	22.87	12.21	3.87	4.39

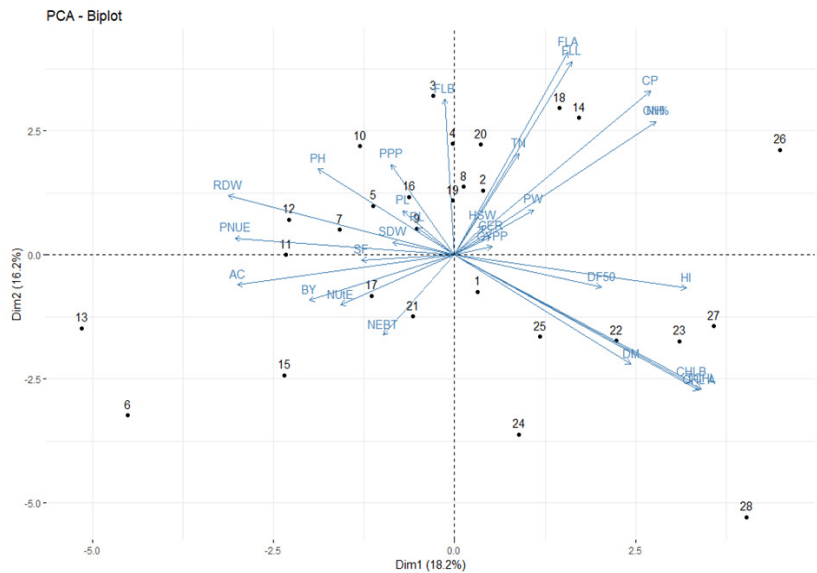


Figure 4: Biplot of 28 genotypes of rice using PCA analysis

Table 4: Contribution of individual quantitative trait towards divergence

S. No.	Characters	%Contribution
1	Chlorophyll ^a	7.41
2	Total chlorophyll	7.41
3	Flag leaf area (cm ²)	7.15
4	Chlorophyll ^b	7.08
5	Flag leaf length (cm)	6.28
6	PNUE	5.55
7	Flag leaf breadth (cm)	5.45
8	Crude protein	5.08
9	NHI	5.06
10	Days to maturity	4.88
11	Amylose content	4.85
12	Root dry weight (g)	4.81
13	No. of ear bearing tillers (EBT)	4.78
14	Harvest Index	4.46
15	Grain N (%)	3.41
16	Biological Yield	2.81
17	Plant height (cm)	2.66
18	Panicles/plant	2.38
19	Days to 50% flowering	2.23
20	Panicle weight (g)	1.21
21	Stem dry weight (g)	1.12
22	NUtE	0.89
23	Panicle length (cm)	0.62
24	Spikelet fertility	0.58

25	Root length (cm)	0.58
26	Total nitrogen	0.58
27	Grain yield/ plant (g)	0.32
28	100 grain weight (g)	0.28
29	Germination %	0.08

all characters, the minimum contribution of days to 50% flowering and grain yield. During selection, the qualities that cause divergence should be taken into account.

Thus, it can be concluded that when planning for a hybridization program or the development of heterotic hybrids and better transgressive segregants are done then one should select genotypes from cluster III (Taposen Youli) to develop high-yielding N use efficient genotypes because of high nitrogen harvest index. Similarly, Thupfu Lha and RCM-from cluster I and Taposen Youli from cluster III have desirable characteristics viz. grain N%, PNUE, NUtE, BY, and NHI and may be selected for intercrossing.

Conclusion

Ans: Based on the study, the evaluation of rice germplasm using D² statistics and principal component analysis provided crucial insights into the variability and trait distribution among genotypes. Cluster I exhibited superior performance in key traits such as grain yield per plant, crude protein, and nitrogen utilization traits, making its genotypes promising for breeding programs. Cluster II demonstrated the highest mean values for maturity and harvest index, while cluster III showed advantages in root traits and amylose content, critical for stress adaptation and grain quality. Principal component analysis revealed that PC1 and PC2, collectively

explained 34.4% of the total variability, with chlorophyll-related traits contributing most to diversity. Genotypes such as Thupfu Lha, RCM-9, and Taposen Youli from clusters I and III were identified as ideal candidates for intercrossing due to their desirable characteristics, including grain N%, PNUE, and NUtE. These findings suggest that simultaneous selection of these traits could facilitate the development of high-yielding, nitrogen-efficient rice genotypes. Moreover, the observed diversity in clusters underscores the potential for targeted hybridization strategies to enhance genetic improvement in rice breeding programs.

References

- Amin MF, M Hasan, NCD Barma, MG Rasul and MM Rahman (2014) Genetic diversity analysis in spring wheat (*Triticum aestivum* L.). *Bangladesh J. Agril. Res.* 39(2): 189-96.
- Bollinedi H, KK Vinod, K Bisht, A Chauhan, S Gopala Krishnan, K Prolay, M Bhowmick, D Nagarajan, S Rao, RK Ellur and AK Singh (2020) Characterising the diversity of grain nutritional and physico-chemical quality in Indian rice landraces by multivariate genetic analyses. *Indian J. Genet. Plant Breed.* 80(1): 26-38. doi: 10.31742/ISGPB.80.1.4.
- Chippy AK, MA Pillai and D Shoba (2021) Genetic diversity analysis in black gram [*Vigna mungo* (L.) Hepper]. *Electronic J. Plant Breeding.* 12(1): 37-45.
- De Mendiburu and F Agricolae (2014) Statistical procedures for agricultural research. R package version. 1(1).
- Devi M, DPB Jyothulal, B Krishnaveni and VS Rao (2019) Genetic Divergence Studies in Rice (*Oryza sativa* L.) Hybrids for Yield, Yield Component Traits and Quality Parameters. *Int. J. Curr. Microb. Appl. Sci.* 8(6): 1577-83.
- Dwivedi AK and SK Mitra (1995) Genotypic correlation and path coefficient analysis in litchi (*Litchi chinensis* Sonn.). *Indian Agri.* 39: 57-61.
- <http://economictimes.indiatimes.com>, 2022
- Kumar A and K Gurumurthi (2000) Genetic divergence studies on clonal performance of *Casuarina equisetifolia*. *Silvae Genetica.* 49(2): 57-59.
- Mahalanobis P. C. 1936. *Proc. Nat. Inst. of Science*, 2: 49-55.
- Nazir R, E Momeni, DJ Armaghani, and MM Amin (2013) Correlation between unconfined compressive strength and indirect tensile strength of limestone rock samples. *Electron J Geotech Eng.* 18(1): 1737-46.
- Nachimuthu VV, S Robin, D Sudhakar, M Raveendran, S Rajeswari 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): 1555-62.
- Prasad KR, KV Radha-Krishna, MHV Bhav and LV Subba-Rao (2017) Genetic variability, heritability and genetic advance in boro rice (*Oryza sativa* L.) germplasm. *Int. J. Curr. Microb. Appl. Sci.* 6(4): 1261-66.
- Reddy NR, SB G and GR Lavanya (2020) Genetic diversity studies in elite rice (*Oryza sativa* L.) germplasm for yield and quality traits. *J. of Pharmacognosy and Phytochemistry.* 9(3): 316-20.
- Singh P K and RD Choudhary (1977) Biometrical methods. In: Quantitative Genetic Analysis, Kalyani Publishers, New Delhi, 178-85 p.
- Singh RK and US Singh (2003) A treatise on the scented rice of India. New Delhi: Kalyani Publishers.
- Uddin MS, MM Rahman, MM Hossain and MAK Mian (2014) Genetic diversity in eggplant genotypes for heat tolerance. *SAARC J. of Agri.* 12(2): 25-39.