



## Evaluation of Morphological Diversity in Rice Lines through Multivariate Analysis

Abu Sayeed Md Hasibuzzaman<sup>1</sup>, Mirza Mofazzal Islam<sup>2</sup>, Shamsun Nahar Begum<sup>1</sup>, Masud Parves<sup>1</sup> and Mohammad Hasanuzzaman Rani<sup>1\*</sup>

<sup>1</sup>Plant Breeding Division, Bangladesh Institute of Nuclear Agriculture, Bangladesh, Email: [romi.hasib@gmail.com](mailto:romi.hasib@gmail.com)

<sup>2</sup>Former Director General, Bangladesh Institute of Nuclear Agriculture, Bangladesh, Email: [mirza\\_islam@yahoo.com](mailto:mirza_islam@yahoo.com)

\*Correspondence: [hasan.hasibpbd@gmail.com](mailto:hasan.hasibpbd@gmail.com); Tel: +8801975217767

Received: 18/09/2024

Accepted: 26/12/2024

Available online: 31/12/2024



Copyright: ©2024 by the author(s).

This work is licensed under a Creative Commons Attribution 4.0 License.

<https://creativecommons.org/licenses/by/4.0/>

**Abstract:** Enhancing rice productivity is a continuous challenge due to the increasing demand. To improve rice grain yield and quality attributes, knowledge of genetic diversity is crucial. Therefore, this study used multivariate analysis to reveal the genetic diversity of twenty-five R lines. This study observed the most significant divergence in productive tiller number (PT), having a CV of 19.23%, which was followed by total grains panicle<sup>-1</sup> (19.15%) and filled grains panicle<sup>-1</sup> (16.94%). Principal Component Analysis (PCA) revealed that 98.21% of the total variability was contributed by the first two components. The genotype-by-trait biplot revealed that total grain panicle<sup>-1</sup> (TG/P), filled grain panicle<sup>-1</sup> (FG/P), and fertility rate (FR) are the three traits that contribute more to genotype variation. Strong, significantly positive associations were found between FG/P and TG/P, plant height (PH) and TG/P, FR, and grain length (GL), FR and grain length breadth ratio (L/B), GL and L/B. Strong, significantly negative associations were found between TG/P and FR, TG/P and GL, TG/P and L/B, PH and FR, and grain breadth (GB) and L/B. Cluster analysis and genetic dissimilarity study suggest that genotypes A14, A17, and A20 are more diverse and could be potential sources of genetic diversity in breeding programs.

**Keywords:** Genetic diversity; PCA; Correlation; Genetic dissimilarity.

### INTRODUCTION

Rice is the most versatile crop, cultivated in a range of environments. It is the primary nourishment for the majority of the world's population and primary dietary source of natural calories worldwide (Long and Ort, 2010). Rice encompasses nearly one-fifth of the land area designated for grains (Chakravarthi and Naravaneni, 2006). The country's growing population necessitates enhanced national food security, which requires improving rice varieties in both quality and quantity. Therefore, enhancing rice productivity and yield is crucial.

Genetic variety implies the inheritable diversity within and among populations of a particular organism. The efficacy of plant breeding relies on the accessibility of genetic variety, understanding of preferred features, and effective selection methodologies that facilitate the utilization of available genetic resources. The reservoir of genetic diversity within an interbreeding population serves

as the foundation for the selection and enhancement of plants (Nachimuthu, 2014). Understanding the nature and extent of genetic variation influencing the inheritance of quantitative traits in rice is essential for a plant breeder. The rice plant demonstrates significant morphological diversity in different traits (Sohrabi *et al.*, 2012). The most stable characteristics of any genotype are grain length, grain breadth, and grain length-to-breadth ratio, which exhibit substantial heritability (Bao, 2014). As improving grain yield as well as quality remains fundamental, the breeding of rice aims to increase genetic diversity of base population (Vanaja and Babu, 2004).

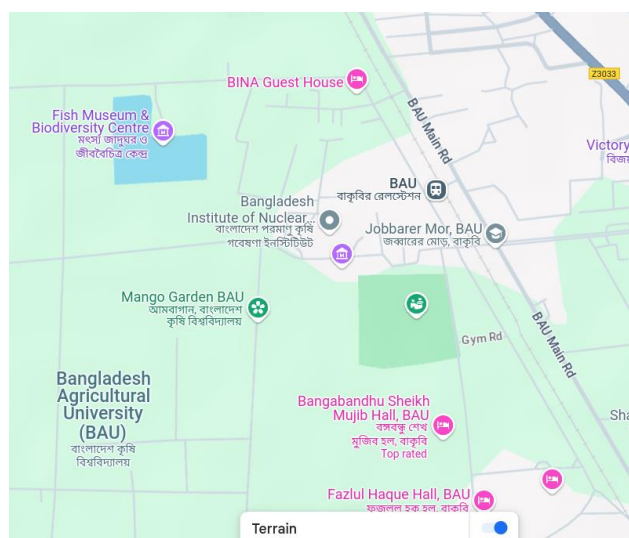
In any crop improvement initiative, to augment production, breeders must sustain a repository of genetically diverse, intriguing donors (Joshi *et al.*, 2013). A thorough insight into heritable variation is vital in effectively managing and applying genetic tools in rice breeding. Combining preferred genotypes, followed by selection, principally relies on choosing parents with genetic

divergence for specific traits. Additionally, multivariate analysis extensively summarizes and describes the genetic divergence among crop germplasm. Multivariate statistical tools, including Cluster analysis, PCA, etc., can identify divergence among the genotypes. Thus, this study focused on evaluating the diversity of the 25 Restorer (R) lines for morphological features and classifying them so that these can be used in further breeding programs.

**MATERIALS AND METHODS**

**Study location**

The experiment was conducted at the field laboratory of the Plant Breeding Division of the Bangladesh Institute of Nuclear Agriculture (BINA) during the Boro season of 2023 (Figure 1).



**Figure 1.** Map of the study area

**Plant materials and data collection**

Twenty-five restorer (R) lines (Table 1) obtained from the International Rice Research Institute (IRRI) were utilized to evaluate genetic diversity. Trait selection and measurement methodologies were derived from the IRRI standard evaluation system for rice. Quantitative traits encompassed days to 80% flowering (DF), days to 80% maturity (DM), filled grains panicle<sup>-1</sup> (FG/P), total grains panicle<sup>-1</sup> (TG/P), spikelet fertility rate as a percentage (FR), phenotypic acceptability (PACP), plant height in centimeters (PH), productive tillers per plant (PT), grain length in centimeters (GL), grain breadth in centimeters (GB), ration of grain length-breadth (L/B), and grain type. The spacing between plants and rows was maintained at 20 cm. All suggested cultural methods, irrigation schedules, fertilizer applications, and pesticide treatments were appropriately implemented.

**Data analysis**

Data analysis was conducted via the software Past 4.03 (Hammer et al., 2001).

**Table 1.** List of restorer (R) lines used in the study

Serial Number	IRRI ID	Genotype Code
1.	IR112897-8-1-1	A1
2.	IR112898-42-1-1	A2
3.	IR112899-3-1-1	A3
4.	IR112930-2-1-1	A4
5.	IR101999-31-1-2	A5
6.	IR112925-5-2-2-B	A6
7.	IR101921-BK-BK-BK-17-1-1	A7
8.	IR101922-BK-BK-BK-14-2-1	A8
9.	IR101924-BK-BK-BK-5-1-1	A9
10.	IR101930-BK-BK-BK-6-2-1	A10
11.	IR101865-2-1-1-1-1-1	A11
12.	IR98153-18-1-1-1-1-1-1	A12
13.	IR98155-68-1-1-1-1-2-1	A13
14.	IR98178-7-2-1-1-2-1-1	A14
15.	IR98184-27-1-1-1-1-1-1	A15
16.	IR98201-16-1-1-1-1-1-1	A16
17.	IR98206-51-2-1-1-1-1-1	A17
18.	IR98212-57-2-1-1-1-1-1	A18
19.	IR98220-21-2-1-1-1-1-1	A19
20.	IR98244-14-2-1-1-1-1-1	A20
21.	IR96479-66-4-1-1-1-1-1-1	A21
22.	IR96532-29-1-1-1-1-2-1-1	A22
23.	IR96542-152-2-1-1-1-1-1-1	A23
24.	IR96553-12-1-3-1-1-1-1-1	A24
25.	IR96581-4-1-4-1-1-2-1-1	A25

**RESULTS AND DISCUSSION**

**Descriptive statistics**

Table 2 presents the first-order statistical measures, i.e., maximum, minimum, sum, mean, standard error (SE), standard deviation (SD), and coefficient of variation (CV) for the measured traits. The most significant variation was observed for PT, with a CV of 19.23%, followed by total

grains per panicle (19.15%) and filled grains per panicle (16.94%). The least variation was observed in DM with the CV of 1.42%. There was no variation in phenotypic

acceptability score (PACP), so data were not presented in the table.

**Table 2.** Trait variations of 25 R lines

Traits	Min.	Genotype	Max.	Genotype	Mean	SE	SD	CV
DF	114	A23	124	A12, A19, A20, A21, A24	120	0.60	2.99	2.49
DM	147	A17, A20, A22	156	A6, A10	152	0.43	2.15	1.42
FG/P	102	A20	194	A2	146	4.95	24.75	16.94
TG/P	130	A17	252	A14	188	7.19	35.96	19.15
FR (%)	52	A14	91	A16	78.64	1.86	9.28	11.80
PH (cm)	82	A20	101	A15	90	1.08	5.38	6.00
PT	5	A1, A22	11	A14	7.8	0.30	1.50	19.23
GL (mm)	6	A14	7.73	A17	6.78	0.08	0.41	6.02
GB (mm)	2.07	A17	2.43	A1, A14	2.23	0.02	0.09	4.08
L/B	2.47	A14	3.74	A17	3.05	0.05	0.25	8.14

The number of days required for flowering ranged from 114 to 124, with an average of 120 days (Table 1). The R line A23 took the least days to complete its flowering. The number of days required for maturity ranged from 147 to 156, with an average of 152 days. The R lines A17, A20, and A22 took the least days for maturity. Filled grain per panicle varied from 102 (A20) to 194 (A2), with a mean of 146. Total grains per panicle showed more significant variability, ranging from 130 (A17) to 252 (A14). However, the fertility rate was the lowest in A14 (52%). The plant height ranged from 82 to 101 cm, averaging 90 cm. The number of productive tillers per plant ranged from 5 to 11. The grain length varied from 6.0 to 7.73 mm, averaging 6.78 mm. The grain breadth ranged from 2.07 to 2.47 mm. The highest grain length and breadth ratio was shown by A17 (3.74).

Among the studied traits, the fertility rate showed two outliers, 52% (A14) and 59% (A18), indicating a deviation from the median (Figure 2). Other traits did not show any outliers.

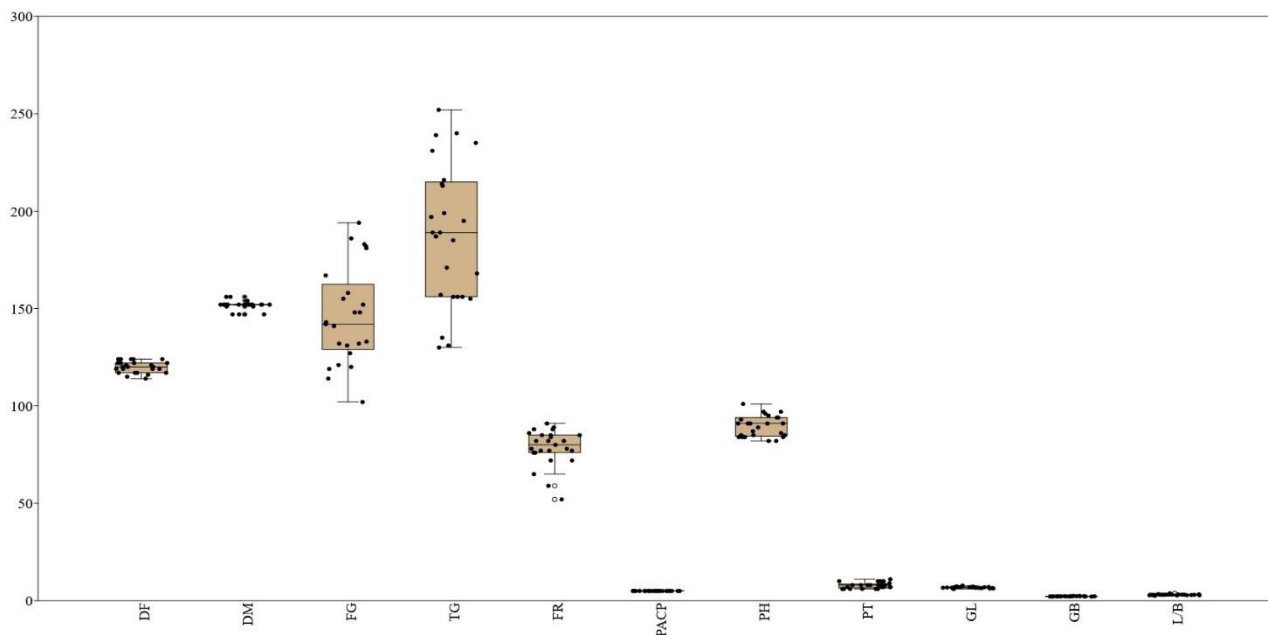
**Principal Component Analysis**

The principal component analysis revealed 11 principal components in the study (Figure 3). The first seven components had eigenvalues greater than 1 (Table 3). The first two components contributed 98.21% of the total

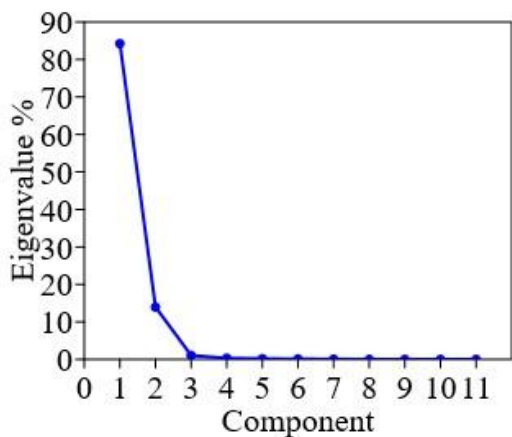
variability, and these two components were used to construct the biplot. These results are in correspondence with the findings of (Nachimuthu, 2014) where he found the the first five axes had an eigenvalue of > 1.0 with cumulative variance of 80.56%.

**Table 3.** Eigenvalues (%) of the eleven principal components

PC	Eigenvalue	% Variance
1	1716.60	84.26
2	284.11	13.95
3	21.67	1.06
4	6.73	0.33
5	3.78	0.19
6	2.71	0.13
7	1.49	0.07
8	0.14	0.01
9	0.02	0.00
10	0.00	0.00
11	0.00	0.00



**Figure 2.** Box plot and jitter showing variability in the studied traits.



**Figure 3.** Eigenvalue (%) of the principal components.

The genotype-by-trait biplot revealed that TG, FG, and FR are the three traits that contribute more to genotype variation (Figure 4). This information is valuable for breeders who can focus on these traits when selecting genotypes for further improvement.

Moreover, the biplot also identified the best genotypes for every trait. In the case of FG, the R lines A1, A2, A4, A5, A12, A15, A22, and A24 performed better. These genotypes had higher FG than the average, indicating they are

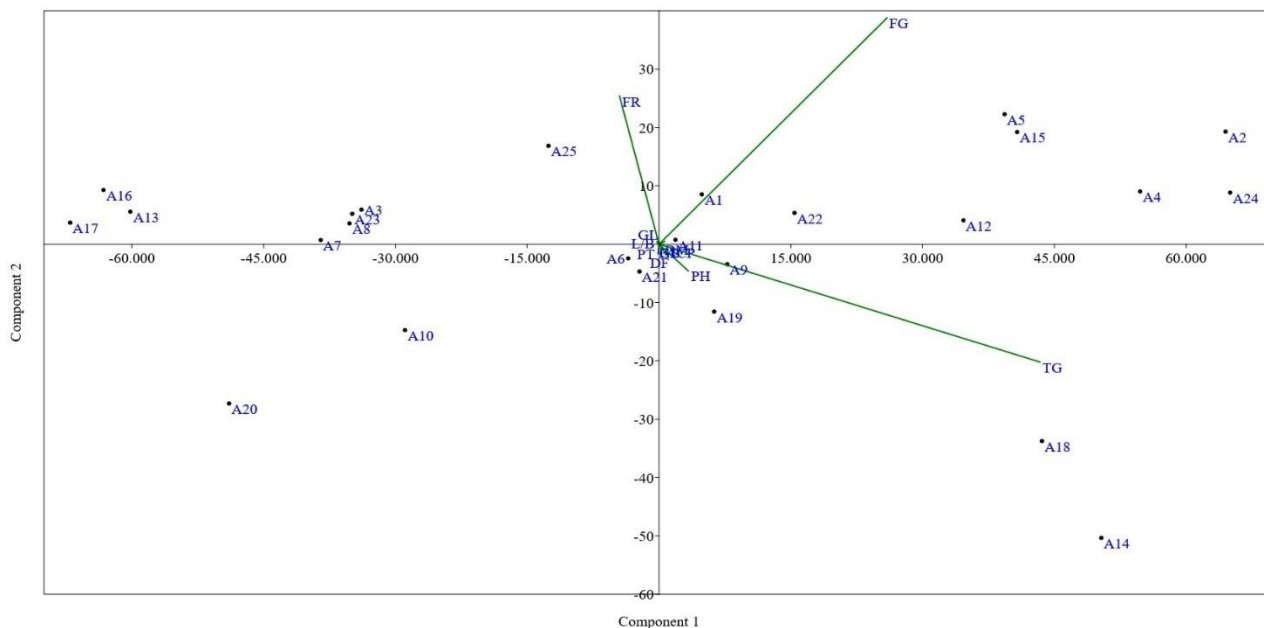
promising candidates for breeding programs to increase grain yield.

Considering the TG A9, A14, and A18 performed better. However, for FR A3, A7, A8, A13, A16, A17, A23, and A25 showed better results (Figure 3). These genotypes had higher FR than the average, indicating that they are promising candidates for breeding programs to increase resistance to sterility.

Overall, the PCA biplot provided valuable insights into the relationships between genotypes and traits and identified the best genotypes for each trait. This information can guide breeders in selecting genotypes for further improvement and developing new varieties with improved grain yield and resistance to sterility. A similar type of study for rice was performed by (Saha et al., 2022; Sanni et al., 2008), and they identified genotypes suitable for different traits.

**Correlation**

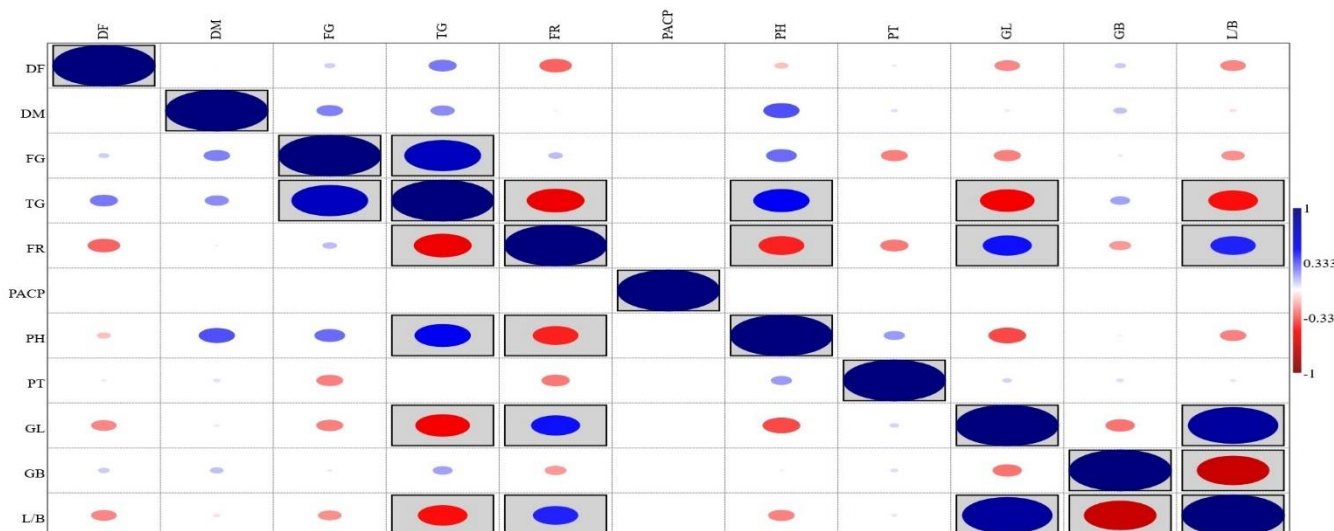
The correlation coefficients between the traits examined in this study are presented in Figure 5. Strong, significantly positive associations were found between FG/P and TG/P, PH and TG/P, FR and GL, FR and L/B, GL and L/B. Strong, significantly negative associations were found between TG/P and FR, TG/P and GL, TG/P and L/B, PH and FR, GB and L/B.



**Figure 4.** Genotype by trait biplot showing the variability of the traits and genotypes.

The correlation coefficients between the traits examined in this study provide valuable insights into their relationships. The strong positive associations observed between FG/P and TG/P, PH and TG/P, FR and GL, FR and L/B, and GL and L/B suggest that these traits are closely related and tend to vary together. For instance, the positive

correlation between FG/P and TG/P indicates that as the number of FG/P increases, so does the TG/P. This relationship is likely because genetic makeup, environmental conditions, and management practices influence these two traits.



**Figure 5.** Correlation among the studied traits. The boxed circles indicate the significant correlation at 5% level of significance.

Similarly, the positive correlation between PH and TG/P suggests that taller plants produce more grains per panicle. This relationship could be attributed to taller plants having more nodes and branches in panicles, allowing for the development of more panicles and, ultimately, more grains.

On the other hand, the strong negative associations found between TG/P and FR, TG/P and GL, TG/P and L/B, PH and FR, and GB and L/B indicate that these traits tend to vary inversely.

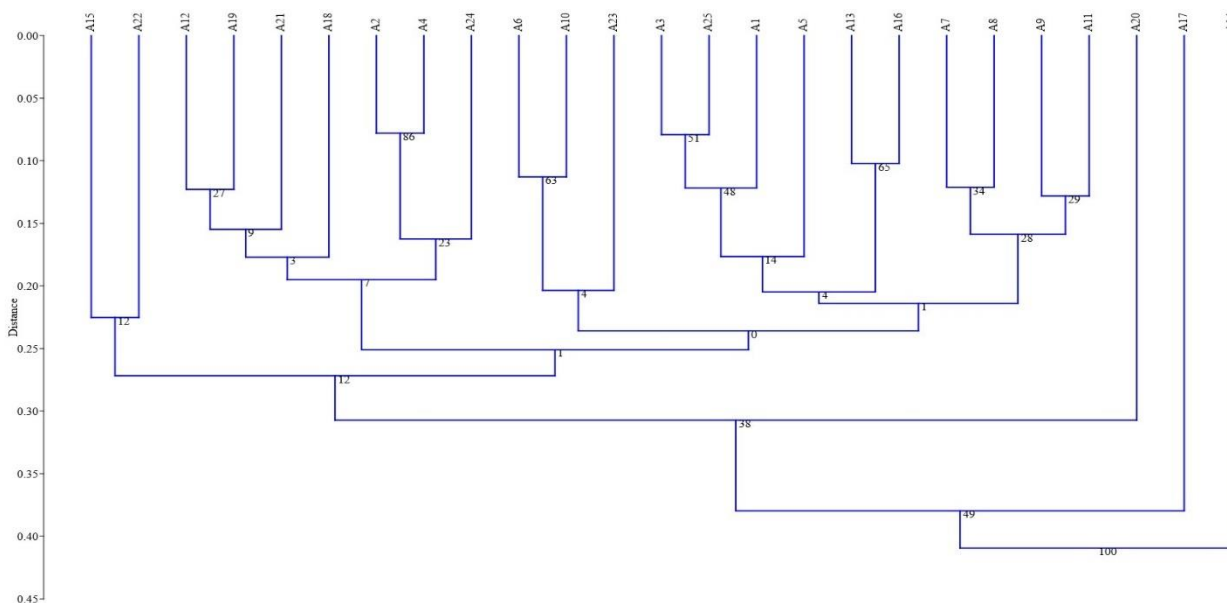
For example, the negative correlation between TG/P and FR suggests that the spikelet fertility rate tends to decrease as the number of total grains per panicle increases. This relationship could be because when plants produce more grains, there may be competition for resources within the panicle, leading to reduced fertility.

Understanding these correlations among traits can benefit plant breeders and geneticists working on rice improvement programs. It allows them to identify traits

likely to be inherited together and make informed decisions when selecting parents for hybridization and breeding purposes.

**Cluster analysis and genetic dissimilarity index**

The UPGMA (unweighted pair group method with arithmetic mean) cluster analysis revealed five clusters in the studied genotypes (Figure 6). The cluster I comprises of A15 and A22; Cluster II comprises of A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A16, A18, A19, A21,



**Figure 6.** UPGMA cluster of the studied R lines showing their clusters

A23, A24 and A25; Cluster III, Cluster IV, and Cluster V each had only one genotype: A20, A17, and A14, respectively. Gower’s genetic dissimilarity index supports the cluster analysis findings. The R lines A14 and A17 showed higher dissimilarity with other genotypes (Figure 7). A20 has also demonstrated considerable distance with different genotypes.

and Cluster V each had only one genotype (A20, A17, and A14, respectively), indicating that these genotypes are genetically distinct from the other genotypes.

The cluster analysis and dissimilarity index results provide valuable insights into the genetic relationships among the studied genotypes. These analyses can help breeders identify genetically diverse parents for crossing and developing new varieties with improved traits.

Gower's genetic dissimilarity index supports the cluster analysis findings. Genotypes A14 and A17 showed higher dissimilarity with other genotypes (Figure 7), indicating that they are genetically distinct and diverse. A20 has also demonstrated considerable distance from different genotypes, suggesting that it is also genetically unique.

The UPGMA cluster analysis revealed five distinct clusters among the genotypes (Figure 6)—cluster I comprises A15 and A22, indicating that these two genotypes are closely related. Cluster II comprises 19 genotypes (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A16, A18, A19, A21, A23, A24 and A25), suggesting that these genotypes share a common genetic ancestry. Cluster III, IV,

These results suggest that genotypes A14, A17, and A20 could be potential sources of genetic diversity in breeding programs. Breeders could utilize these genotypes to introduce new and valuable traits into existing varieties or to develop new varieties with enhanced performance.

Additionally, the cluster analysis results can be used to identify genetically similar parents for crossing, which can help to reduce the risk of inbreeding and maintain genetic diversity in breeding populations.

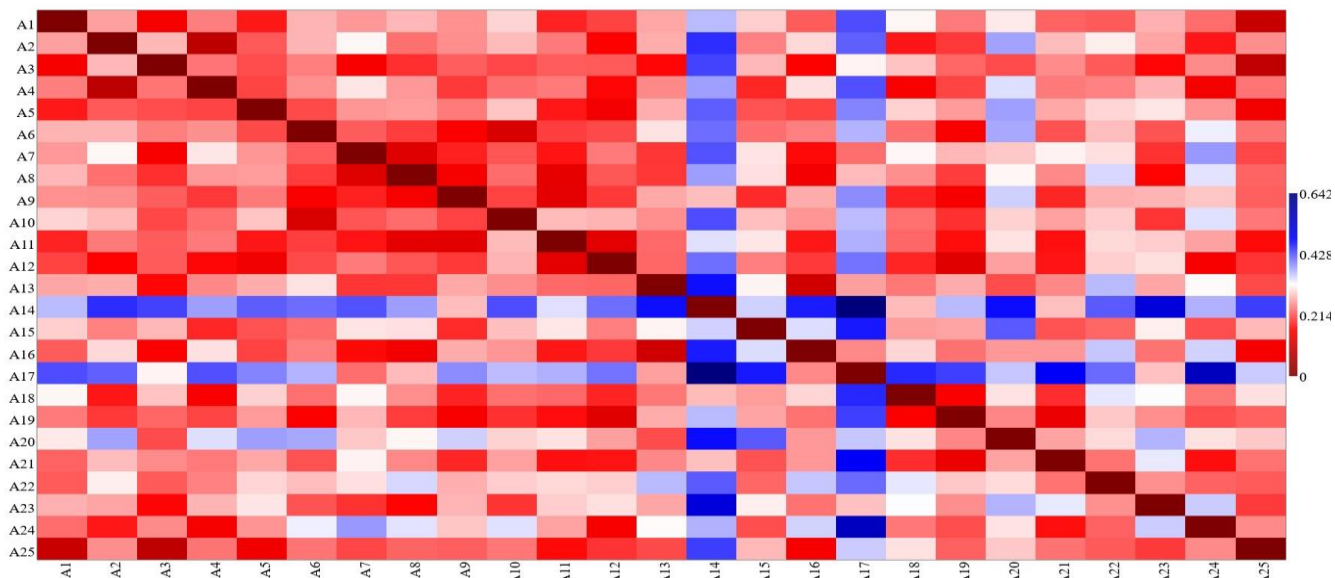


Figure 7. Genetic dissimilarity among studied R lines

Overall, the cluster analysis and dissimilarity index results provide valuable information about the genetic relationships among the studied genotypes and can guide breeders in making informed decisions about parental selection and breeding strategies.

**CONCLUSION**

This study employed multivariate analysis to uncover genetic divergence in the 25 R lines. We found that PT varied the greatest, along with FG/P and TG/P. Moreover, TG/P, FG/P, and FR are the three variables that contribute most to genotype variation, according to the genotype-by-trait biplot. Based on the findings it was evident that the line genotypes A14, A17, and A20 are more diverse and could be potential sources of genetic diversity in breeding programs.

**Conflict of Interest**

The authors declared no conflicts of interest.

**REFERENCES**

Bao, J. (2014). Genes and QTLs for Rice Grain Quality Improvement. In *Rice—Germplasm, Genetics and Improvement*. IntechOpen. <https://doi.org/10.5772/56621>  
 Chakravarthi, B. K., and Naravaneni, R. (2006). SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa*. L). *African Journal of Biotechnology*, 5(9), Article 9. <https://www.ajol.info/index.php/ajb/article/view/42772>  
 Hammer, O., Harper, D. A. T., and Ryan, P. D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 4(1), 4–9.

Joshi, M., Verma, S. K., Singh, J. P., and Barh, A. (2013). Genetic diversity assessment in lentil (*Lens culinaris medikus*) genotypes through ISSR marker. *The bioscan*, 8(supplement 4), article supplement 4.  
 Long, s. P., and ort, d. R. (2010). More than taking the heat: crops and global change. *Current opinion in plant biology*, 13(3), 241–248. <https://doi.org/10.1016/j.pbi.2010.04.008>  
 Nachimuthu, v. V. (2014). Evaluation of rice genetic diversity and variability in a population panel by principal component analysis. *Indian journal of science and technology*, 7(10), 1555–1562. <https://doi.org/10.17485/ijst/2014/v7i10.14>  
 Saha, S. R., Ferdousi, A., Hassan, L., Haque, M. A., Begum, S. N., yasmin, F., and Akram, W. (2022). Rice landraces from haor areas of bangladesh possess greater genetic diversity as revealed by morpho-molecular approaches along with grain quality traits. *Cogent food and agriculture*, 8(1), 2075130. <https://doi.org/10.1080/23311932.2022.2075130>  
 Sanni, K. A., Fawole, I., Guei, R. G., Ojo, D. K., Somado, E. A., Tia, D. D., Ogunbayo, S. A., and Sanchez, I. (2008). Geographical patterns of phenotypic diversity in *Oryza sativa* landraces of côte d’ivoire. *Euphytica*, 160(3), 389–400. <https://doi.org/10.1007/s10681-007-9549-5>  
 Sohrabi, M., Rafii, M. Y., Hanafi, M. M., Siti Nor Akmar, a., and Latif, M. A. (2012). Genetic diversity of upland rice germplasm in malaysia based on quantitative traits. *The scientific world journal*, 2012, 416291. <https://doi.org/10.1100/2012/416291>  
 Vanaja, T., and Babu, L. C. (2004). Heterosis for yield and yield components in rice (*Oryza sativa* L.). *Journal of tropical agriculture*, 42(1–2), 43–44.

