



GENETIC DIVERSITY AND VARIATION AMONG MAJOR AGRICULTURAL CROP VARIETIES

Dr Pooja Sahu¹

¹Assistant Professional, Department: Department of Dravyaguna Vigyana, College Name: Om Ayurveda Medical College, Betul, M.P. Pincode: 460001, Email I'd: misthy041024@gmail.com, ORCID ID - <https://orcid.org/0009-0004-4582-3091>

Corresponding Author: Dr Pooja Sahu, misthy041024@gmail.com

ABSTRACT

The genetic diversity of rice germplasm is great enough to be useful for crop improvement, parental selection and conservation. In the current study phenotypic diversity and variation of rice accessions was evaluated based on 12 agronomic and morphological traits from a secondary data set of 2266 accessions. Trait variability and multivariate relationships were described by using descriptive statistics, coefficient of variation, spearman correlation analysis, principal component analysis and hierarchical cluster analysis. Significant variation was noted for all the traits evaluated, with the traits with the highest relative variability being CUNO_REPRO, ligule length and seedling height while grain length and grain width were relatively more stable. The correlation analysis showed that CULT_REPRO was highly positively correlated with leaf length, heading-related traits as well as grain width and 100-grain weight. The first two principal components accounted for 47.16% of the total phenotypic variation, with vegetative and reproductive traits having high loadings on PC1 and grain-related traits having high loadings on PC2. A total of 1,865 complete accessions were clustered into two large phenotypic groups that had different agronomic profiles. Results showed that the germplasm used in this study was highly multidimensional phenotypically diverse and the value of multivariate analysis for accessions characterization. The use of molecular markers and multi-location is recommended to enhance future genetic diversity assessment and breeding application.

Keywords: Rice accessions, Phenotypic diversity, Agronomic traits, Principal component analysis, Cluster analysis

1. Introduction

Rice (*Oryza sativa* L.) is a major staple food crop and plays significant role in food security, rural livelihoods and agricultural economies worldwide. The ability to produce rice in an environment of climate change, requires access to a diverse germplasm with useful agronomic and morphological traits. Crop diversity is the biological basis for selection, adaptation and long-term productivity, especially under environmental stresses that pose a challenge to yield stability. Traditional and locally adapted crop varieties may contain significant variability which can be leveraged to enhance resilience, boost productivity and adapt to a changing environment (Bibi & Asad, 2024). Beyond this, genetic and phenotypic variation in cultivated populations is for breeding programmes and may result in a wider variation of traits to select from for crop enhancement (Singh et al., 2023).

Characterization of germplasm is thus one of the initial tasks in the identification of useful variation among accessions and elucidation of the relationships among economically important traits. In recent years, studies on diversity have been increasingly using molecular markers, genomic information and traditional phenotypic characterization to gain a broad understanding of the variation in the crop. The use of molecular markers has significantly enhanced the characterization of germplasm, as it can distinguish closely related accessions and identify potentially valuable breeding materials (Bunjkar et al., 2024). Research with landraces has also shown that significant diversity can be found within cultivated crop populations and thus landraces can be conserved and utilized as locally adapted genetic resources (Dido et al., 2022). Genetic diversity is also considered as a basic biological resource that has an impact on adaptability and responses of populations to environmental and biological challenges (Gibson, 2022).

Phenotypic characterization is still very useful in the agricultural research, as the agronomic and morphological characteristics observed are directly related to the desired ones for crop performance and selection. Landraces and cultivars have been found to be diverse in numerous crop species, highlighting the need for maintaining diverse germplasm pools for future breeding (Guimarães et al., 2023). This diversity is even more significant in a climate variability scenario, when breeders need different parental materials to contribute traits for stress tolerance and yield stability. Studies on millet have also identified the need to increase the resilience of the crop with various germplasm under different production conditions (Krishnababu et al., 2024). In fact, research conducted on lesser-known crops have also shown that the crop diversity could be helpful for crop improvement, food security and overall food system sustainability (Muhammad et al., 2020).

Systematic identification of representative and divergent germplasm is also essential for effective conservation and utilization of crop diversity. Utilizing core collections generated from multiple accessions can aid in the maintenance of maximum diversity and in breeding and evaluation efforts (Li et al., 2023a). Plant genetic resources are, therefore, not only conservation assets but also are practical sources of traits needed for future crop improvement and sustainable agriculture (Salgotra & Chauhan, 2023). Simultaneously, it is important to balance diversity with uniformity in cultivars because too much uniformity could limit the diversity needed for future adaptation and improvement (Metakovsky et al., 2024). The use of diversity information also has been used in combination with traditional selection to speed improvement of agronomic traits through genomic-assisted breeding (Singh & Pandey, 2024). Further characterization of plant diversity is therefore crucial for comprehending the current diversity and guiding future conservation and breeding efforts (Xiong et al., 2024).

Quantitative evaluation of agronomic and morphological characteristics can offer an effective basis to distinguish accessions and identify those characteristics that add the most to the phenotypic variability in rice. Multivariate analysis, including correlation analysis, principal component analysis and cluster analysis are especially useful since they consider multiple traits simultaneously and generate patterns that can be detected only through multivariate analysis. The present study therefore aims at evaluating the phenotypic diversity and variation among the rice accessions, available with respect to the various agronomic and morphological traits, to determine major sources of variability and useful groups of accessions in agriculture.

Objectives

- To evaluate variability in major agronomic and morphological traits among rice accessions.
- To determine relationships among measured traits using correlation and multivariate analyses.
- To classify rice accessions into distinct phenotypic groups based on their overall agronomic and morphological characteristics.

2. Materials and Methods

2.1 Study Design and Data Source

A secondary data-based quantitative research design was used to assess phenotypic diversity/variation of the rice accessions using available agronomic and morphological characteristics (Shahane, 2020). A total of 2,266 individual rice accessions were included with each having a unique accession identifier. There were 12 quantitative agronomic and morphological variables for analysis. Since the study was an analysis of an existing dataset and no experimentation was done on plants, all analyses were based on the accession-level data recorded in that dataset.

2.2 Study Variables

The variables used for the analysis were all the grain, vegetative and reproductive characteristics of the rice accessions. The traits recorded were CUDI_REPRO, CULT_REPRO, CUNO_REPRO, grain length (GRLT), grain width (GRWD), 100-grain weight (GRWT100), heading-related characteristic (HDG_80HEAD), ligule length (LIGLT), leaf length (LLT), leaf width (LWD), plant posture-related characteristic (PLT_POST), and seedling or plant height (SDHT). All these variables were then combined to describe the phenotypic variability and to find the differences among the rice accessions.

2.3 Data Screening and Preparation

The data set was first checked for duplicate accession numbers and for missing data and variable formatting errors. There were no duplicate accession numbers (2,266 accession numbers were all different). There were a few missing values, and they were coded as “NA” in the original data. A relatively small percentage of grain characteristics variables had missing data, while there were a larger percentage of incomplete data for some of the vegetative and reproductive traits. When data were missing for a single observation for an individual, the entire observation was treated as missing and not included in the individual calculation for the variable. Only observations that had valid information for the variables included in the specific multivariate analysis were used in the multivariate analyses. Continuous variables were checked for extreme values and overall distributional patterns before statistical analysis.

2.4 Descriptive and Variability Analysis

Rice accessions were summarized for their phenotypic characteristics by descriptive statistics for each quantitative trait. Means, standard deviations, median, minimum and maximum values were calculated

for each variable. Coefficient of variation was also computed to compare the relative amount of variability among the different traits that were measured on different scales. Phenotypic heterogeneity of the assessed accessions was considered higher when the traits showed higher CV.

2.5 Correlation Analysis

Correlation analysis was used to determine the relationships between agronomic and morphological traits. Association between the different grain, vegetative and reproductive characters were done using pairwise correlation. Approximately continuous, normally distributed variables were correlated using Pearson correlation coefficients and Spearman rank correlation was used where distributional assumptions were not sufficiently met. Correlation coefficients were interpreted based on their values and sign and statistical significance was determined at the 5% probability level.

2.6 Principal Component Analysis

The dimensionality of the data was reduced using principal component analysis to identify the major combinations of traits that underlie the phenotypic variation among the rice accessions. Quantitative traits were standardized prior to analysis since the variables measured were on different scales and units. Principal components were extracted from the standardized data and components with the largest eigenvalues and meaningful contribution to the cumulative variance were retained and interpreted. To identify the main agronomic and morphological traits of each of the principal components, the values of trait loadings were analysed. Accessions were then plotted on the major principal components to display patterns of phenotypic differentiation in the distribution of accessions.

2.7 Hierarchical Cluster Analysis

Hierarchical cluster analysis was used to group rice accessions based on the overall phenotypic similarity. The values of traits were standardized to reduce the effect of scale differences in measurements. The dissimilarity was calculated using the euclidean distance and Ward's linkage method was used to sequentially merge accessions and/or groups minimizing the variance within each cluster. This created a hierarchical structure which was displayed in the form of a dendrogram. In the next step, the various cluster characteristics were analysed by comparing the mean values of the main agronomic and morphological traits for the different clusters.

2.8 Statistical Analysis

The cleaned secondary data was used in all statistical analyses. Appropriate measures of central tendency and dispersion were used for continuous variables. Relationships and multivariate patterns between the data were examined using correlation analysis, principal component analysis and hierarchical clustering. Inferential testing was only performed where statistical significance was set at $p < 0.05$. Descriptive tables, correlation matrix, principal component plots and hierarchical clustering visualisations were used to present results to demonstrate the integrated assessment of the phenotypic diversity of the rice accessions.

3. Results

3.1 Descriptive Characteristics and Phenotypic Variability

2,266 rice accessions were tested for 12 agronomic and morphological traits. Due to missing measurements, the number of observations was slightly different for each trait. Characteristics related to grain were available for almost all accessions, and the reproductive and vegetative traits generally had valid observations of ~1,877–1,901. A lot of phenotypic variation was observed for the characters evaluated (Table 1).

Table 1. Descriptive statistics and coefficient of variation of agronomic and morphological traits among rice accessions

Trait	n	Mean	SD	Median	Minimum	Maximum	CV (%)
CUDI_REPRO	1,899	4.87	1.00	5.00	2.00	9.10	20.51
CULT_REPRO	1,900	113.81	29.53	117.00	27.00	204.00	25.95
CUNO_REPRO	1,901	16.45	5.23	16.00	5.00	40.00	31.79
GRLT	2,261	8.62	1.02	8.60	4.70	12.70	11.81
GRWD	2,261	3.02	0.41	3.00	1.50	4.40	13.62
GRWT100	2,259	2.49	0.52	2.50	1.00	5.00	20.75
HDG_80HEAD	2,265	102.10	24.52	99.00	50.00	184.00	24.01
LIGLT	1,878	18.46	5.62	18.00	4.00	47.00	30.47

LLT	1,901	54.91	12.26	55.80	13.00	93.00	22.33
LWD	1,901	1.38	0.32	1.30	0.40	2.50	22.95
PLT_POST	1,896	25.01	3.93	25.00	13.00	37.00	15.72
SDHT	1,877	38.70	11.72	38.00	12.00	74.00	30.29

The coefficient-of-variation analysis showed that phenotypic variability was not uniform across traits. CUNO_REPRO exhibited the greatest relative variability, followed by LIGLT and SDHT, whereas GRLT and GRWD were comparatively less variable. The overall pattern of relative variability among the 12 traits is illustrated in Figure 1.

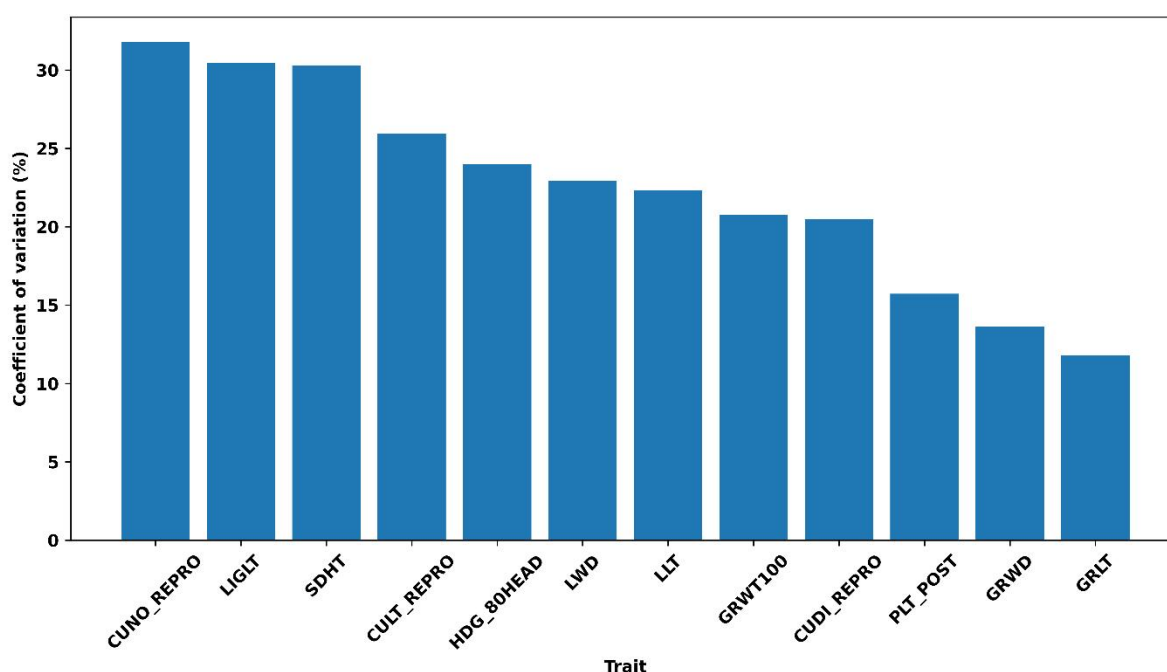


Figure 1. Coefficient of variation of major agronomic and morphological traits among rice accessions

3.2 Relationships Among Agronomic and Morphological Traits

Several moderate to strong correlations between the traits measured were observed (Table 2). The highest positive correlation was found between CULT_REPRO and LLT, and between CULT_REPRO and HDG_80HEAD and between LIGLT and LLT. The correlation of grain width with 100-grain weight was positive, showing that accessions with wider grains on average had higher 100-grain weight.

In contrast, CUNO_REPRO had some negative relationships with the other characteristics, but these relationships were relatively weak.

Table 2. Strongest correlations among agronomic and morphological traits

Trait 1	Trait 2	Spearman ρ	p-value
CULT_REPRO	LLT	0.701	<0.001
CULT_REPRO	HDG_80HEAD	0.589	<0.001
LIGLT	LLT	0.568	<0.001
CULT_REPRO	PLT_POST	0.540	<0.001
LLT	PLT_POST	0.534	<0.001
GRWD	GRWT100	0.513	<0.001
CULT_REPRO	LIGLT	0.498	<0.001
LLT	LWD	0.494	<0.001
HDG_80HEAD	LLT	0.493	<0.001
HDG_80HEAD	LIGLT	0.458	<0.001
CULT_REPRO	LWD	0.448	<0.001
CUDI_REPRO	HDG_80HEAD	0.418	<0.001

The complete pattern of positive and negative relationships among all 12 traits is presented in the correlation heatmap (Figure 2). The heatmap demonstrates a particularly coherent positive association among CULT_REPRO, HDG_80HEAD, LIGLT, LLT, LWD, and PLT_POST.

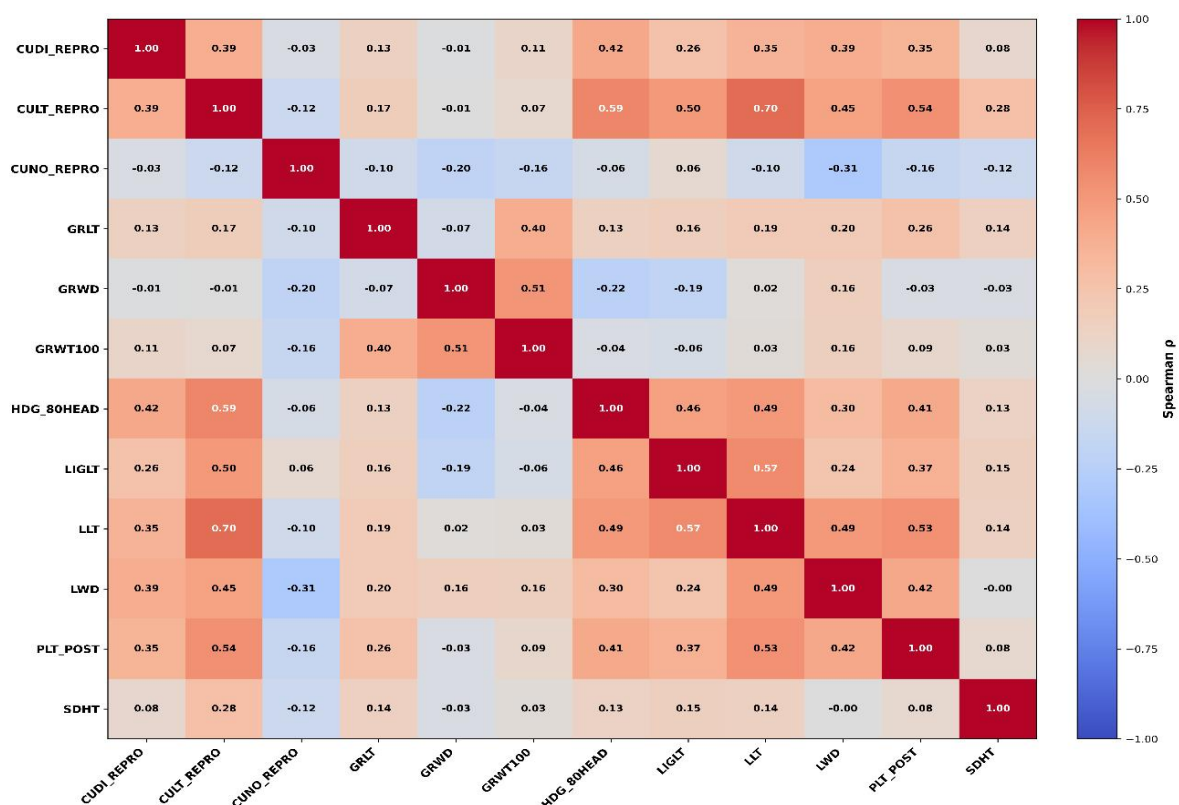


Figure 2. Spearman correlation heatmap of agronomic and morphological characteristics among rice accessions

3.3 Principal Component Analysis

A principal component analysis was performed on 1,865 accessions that have data for all 12 traits. The first six PCs together explained 79.90% of the total phenotypic variation (Table 3). The highest source of variation was found along PC1, with PC2 being the second largest source of variation. The first two components accounted for 47.16% of the overall variation.

Table 3. Eigenvalues and proportion of phenotypic variance explained by the first six principal components

Principal component	Eigenvalue	Variance explained (%)	Cumulative variance (%)
PC1	3.817	31.80	31.80
PC2	1.844	15.36	47.16

PC3	1.199	9.99	57.14
PC4	1.017	8.47	65.61
PC5	0.915	7.62	73.23
PC6	0.800	6.66	79.90

PC1 was predominantly influenced by CULT_REPRO, LLT, PLT_POST, HDG_80HEAD, LIGLT, and LWD, indicating that vegetative and reproductive characteristics represented an important dimension of phenotypic differentiation. PC2 was influenced principally by GRWT100 and GRWD, demonstrating a distinct contribution of grain-size and grain-weight characteristics to variation among accessions. The progressive decline in variance explained across components is shown in Figure 3.

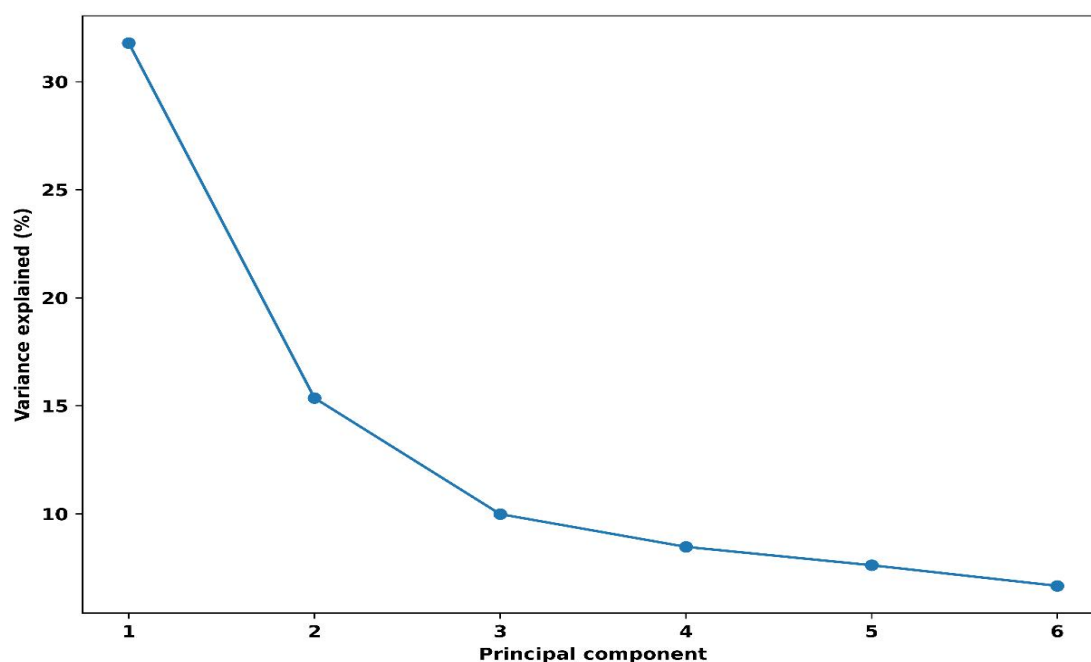


Figure 3. Scree plot showing the percentage of total phenotypic variance explained by the first six principal components

3.4 Hierarchical Classification of Rice Accessions

The 1,865 complete accessions were divided into two general phenotypic clusters by hierarchical cluster analysis of standardized measurements. The number of accessions in Cluster 1 was 955 while Cluster 2

had 910 accessions. Significant differences between the two treatments were found in several reproductive and vegetative parameters (Table 4).

Table 4. Mean agronomic and morphological characteristics of the two major phenotypic clusters

Trait	Cluster 1 (n = 955)	Cluster 2 (n = 910)
CUDI_REPRO	4.49	5.22
CULT_REPRO	95.08	133.37
CUNO_REPRO	17.18	15.57
GRLT	8.47	8.85
GRWD	3.06	3.01
GRWT100	2.46	2.50
HDG_80HEAD	85.68	115.03
LIGLT	15.61	21.42
LLT	47.51	62.68
LWD	1.25	1.51
PLT_POST	23.00	27.08
SDHT	36.69	40.77

Cluster 2 generally exhibited higher values for CULT_REPRO, HDG_80HEAD, LIGLT, LLT, LWD, PLT_POST, SDHT, and GRLT, whereas Cluster 1 showed a higher mean CUNO_REPRO and slightly greater GRWD. These differences indicate the presence of two broad phenotypic profiles among the evaluated rice accessions. The separation of accessions along the first two principal components is visualized in **Figure 4**, demonstrating that multivariate variation contributed to the differentiation of the identified groups.

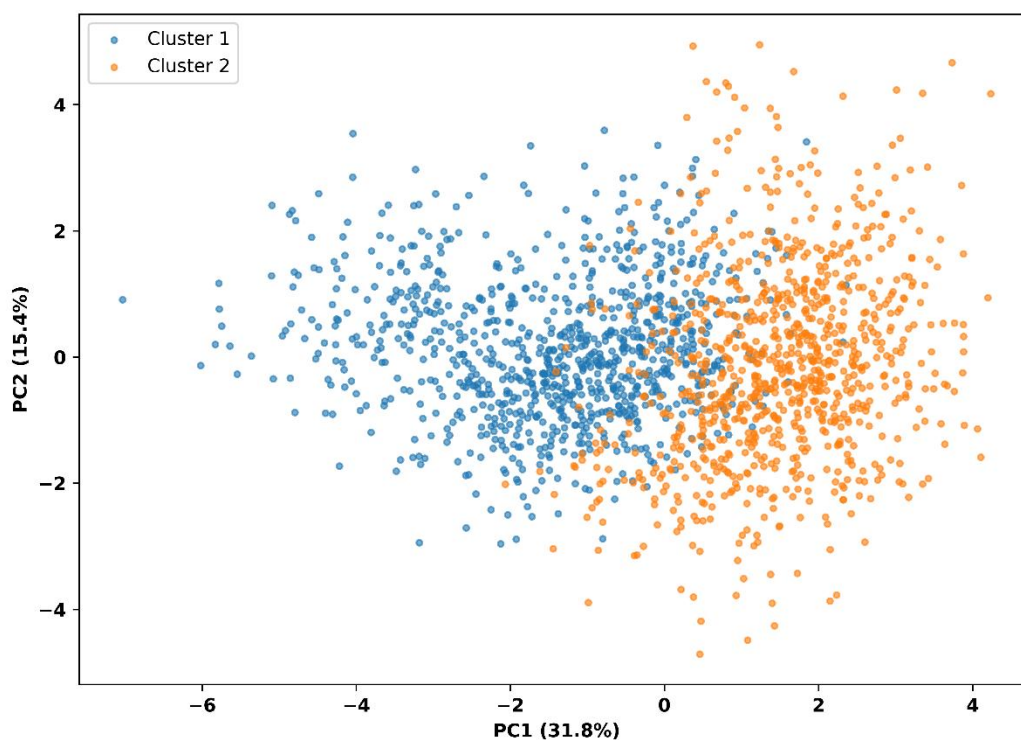


Figure 4. Distribution of rice accessions across PC1 and PC2 according to hierarchical phenotypic cluster

3.5 Overall Phenotypic Diversity Pattern

The hierarchical dendrogram also showed a significant amount of phenotypic variation within accessions and a gradual increase in the formation of larger phenotypic groups. The branching was an indication that the germplasm evaluated was not phenotypically uniform and had both relatively similar and comparatively divergent groups of accessions. The results, when combined with the variability, correlation, PCA and cluster results, indicate that there is significant multidimensional phenotypic diversity among the rice germplasm evaluated.

4. Discussion

The present study revealed high phenotypic variability in the accessions examined with significant differences in reproductive, vegetative and grain related traits. The relative variability of the grain was highest for CUNO_REPRO, LIGLT and SDHT, with lowest for grain length and grain width. This type of variation demonstrates there is significant variation within the evaluated germplasm that can be

exploited for agronomic selection. In germplasm evaluation of crops, comparable trends of high variation within quantitative traits have been observed and multivariate assessment proved beneficial in discriminating accessions with contrasting agronomic traits (Chand et al., 2023). A significant amount of on-farm diversity has also been seen in wheat populations, highlighting the need for maintaining on-farm diversity for selection and adaptation in future (Gelelcha et al., 2023).

The correlations also showed that there was synchronization of fluctuations in multiple vegetative and reproductive traits. CULT_REPRO x LLT ($\rho = 0.70$), CULT_REPRO x HDG_80HEAD ($\rho = 0.59$), LIGLT x LLT ($\rho = 0.57$) had the highest positive correlation. Grain width was also moderately positively correlated with 100-grain weight ($\rho = 0.51$), with the wider grains being associated with heavier grain weight. Similar studies have shown that the relationship assessments of different yield-related and morphological traits can help identify traits that can be used indirectly for breeding programs (Kumar et al., 2024). In addition to providing information on molecular diversity, morphological differentiation has been demonstrated to exhibit agriculturally relevant variation of cultivated materials (Wang et al., 2023).

Further evidence of multidimensional phenotypic differentiation was obtained using principal component analysis. The first principal component accounted for the 31.80% of the total variance and the second principal component accounted for the 15.36% of the total variance, the cumulative variance of the first two principal components was 47.16%. The vegetative and reproductive traits were mainly related to PC1 while grain width and 100-grain weight were mainly related to PC2. Dimensionality reduction techniques have shown their potential for identification of major axes of variation by comparison of similar utilization of high-density SNP information, which reveals substantial diversity and differentiation among wheat germplasm resources (Lei et al., 2024). Multivariate differentiation of genetically diverse wheat collections has also been reported based on SNP markers (Mazumder et al., 2024) and among soybean genotypes based on SSR-based markers (Rani et al., 2023).

The full set of accessions was divided into two major phenotypic categories in the hierarchical clusters, with accessions in Cluster 2 tend to have greater values for CULT_REPRO, heading-related traits, ligule length, leaf length, leaf width, plant posture, seedling height, and grain length. The separation of the evaluated germplasm into distinct groups indicates that there is structure in the diversity of the germplasm. Similar clustering and population differentiation have been shown for cowpea (Gumede et

al., 2022), sugar beet germplasm resources (Peng et al., 2024), cassava collections for different unique genotype (Tighankoumi et al., 2024) and cultivated and wild Avena germplasm using SSR markers (Tiruneh et al., 2024).

The results are also especially relevant for rice improvement as previous studies have shown that there is significant amount of diversity present in rice germplasm at both the molecular and varietal level. The SSR-based evaluation of Shanlan upland rice showed meaningful relationships and differentiation among accessions (Li et al., 2023b) and SSR characterization of *Oryza sativa* cultivars indicated considerable diversity and phylogenetic structuring among the cultivars of rice materials (Salem et al., 2024). Similarly, variations in the phenotype of traditional rice varieties have been observed, and morphology can be used for initial germplasm characterisation and selection of parents (Sruthi et al., 2023).

In spite of these strengths, the current analysis has the drawbacks of using secondary phenotypic data, and the lack of molecular markers, and thus is not suitable for interpretation as real genetic distances. Variations in quantitative traits may also be due to environmental effects. SNP and/or SSR data, multi-location trials and genotype by environment analysis should be combined with the present phenotypic characterization in future studies. This integration would help to provide better evidence of the underlying genetic differentiation and to better identify diverse parental materials for breeding as has been accomplished in molecular studies of rice and other crop germplasm.

5. Conclusion

In the present study, significant phenotypic diversity and variation were seen among the rice accessions using a wide range of agronomic and morphological traits. Significant variation was found among the reproductive, vegetative, and grain characteristics, among them CUNO_REPRO, ligule length and seedling height were the three traits with higher relative variability. Grain length and grain width, however, were relatively consistent among the accessions tested. Several significant relationships among traits were found from the correlation analysis, particularly between CULT_REPRO and leaf length, CULT_REPRO and heading-related traits, and grain width and 100-grain weight. Some traits may be useful for indirect selection during crop improvement, as shown in these relationships. Phenotypic variation was also found to be multidimensional with the first two principal components accounting for

a large share of the variation and differentiating vegetative, reproductive, and grain-related traits as separate sources of variation. The accessions were grouped into two broad phenotypic clusters through hierarchical clustering that showed significant differences in some agronomic traits, revealing some structure in the germplasm. The results, in general, underscore that phenotypic characterization is important for discovering and selecting diverse and potentially useful rice accessions for breeding purposes and germplasm management. Nonetheless, due to the secondary phenotypic data used in the analysis, future research should incorporate molecular markers, multi-location evaluation, and genotype-by-environment analysis to validate the genetic differentiation and enhance selection methods for sustainable rice improvement.

References

1. Bibi, R., & Asad, M. (2024). Harnessing Genetic Diversity in Traditional Crop Varieties for Climate Adaptation: Exploring Strategies for Enhancing Resilience and Productivity. *Bioscience Insights*, 1(02), 24-33
2. Bunjkar, A., Walia, P., & Sandal, S. S. (2024). Unlocking genetic diversity and germplasm characterization with molecular markers: strategies for crop improvement. *J. Adv. Biol. Biotechnol*, 27(6), 160-173.
3. Chand, S. P., Debnath, S., Rahimi, M., Purakayastha, S., & Rout, S. (2023). Genetic variability, correlation, diversity, path coefficients and principal component analysis in Indian mustard. *International Journal of Computational Biology and Drug Design*, 15(6), 445-462.
4. Dido, A. A., Krishna, M. S. R., Assefa, E., Degefu, D. T., Singh, B. J. K., & Tesfaye, K. (2022). Genetic diversity, population structure and relationship of Ethiopian barley (*Hordeum vulgare* L.) landraces as revealed by SSR markers. *Journal of Genetics*, 101(1), 9.
5. Gelelcha, F., Kumsa, F., & Kuma, T. (2023). On-farm genetic diversity of wheat (*Triticum aestivum* spp.) in Digalu Tijo District, Arsi zone, Ethiopia. *Heliyon*, 9(1).
6. Gibson, A. K. (2022). Genetic diversity and disease: The past, present, and future of an old idea. *Evolution*, 76(S1), 20-36.
7. Guimarães, J. B., Nunes, C., Pereira, G., Gomes, A., Nhantumbo, N., Cabrita, P., ... & Veloso, M. M. (2023). Genetic diversity and population structure of cowpea (*Vigna unguiculata* (L.) Walp.) landraces from Portugal and Mozambique. *Plants*, 12(4), 846.

8. Gumedde, M. T., Gerrano, A. S., Amelework, A. B., & Modi, A. T. (2022). Analysis of genetic diversity and population structure of cowpea (*Vigna unguiculata* (L.) Walp) genotypes using single nucleotide polymorphism markers. *Plants*, 11(24), 3480.
9. Krishnababu, M. E., Mandal, O., Begum, M., Saikanth, D. R. K., Nandy, R., Kaushal, K., & Mishra, P. (2024). Exploring millet genetic diversity for improved crop resilience: a review. *Int J Environ Clim Change*, 14(1), 898-905.
10. Kumar, A., Jaiswal, J. P., Chaudhary, H., Vinesh, B., Kumawat, S., Parihar, P., & Bhatt, B. (2024). Deciphering genetic diversity in advanced wheat lines (*Triticum aestivum* L.) for yield and other yield contributing traits across the locations. *Int J Plant Soil Sci*, 36(3), 427-40.
11. Lei, M. L., Liu, X., Wang, Y. Z., Cui, G. Q., Mu, Z. X., Liu, L. L., ... & Zhang, X. J. (2024). Genetic diversity analysis of winter wheat germplasm resources in Shanxi Province based on 55k SNP array.
12. Li, R., Huang, Y., Yang, X., Su, M., Xiong, H., Dai, Y., ... & Yuan, Q. (2023a). Genetic diversity and relationship of Shanlan upland rice were revealed based on 214 upland rice SSR markers. *Plants*, 12(15), 2876.
13. Li, X., Cui, L., Zhang, L., Huang, Y., Zhang, S., Chen, W., ... & Yan, C. (2023b). Genetic diversity analysis and core germplasm collection construction of radish cultivars based on structure variation markers. *International Journal of Molecular Sciences*, 24(3), 2554.
14. Mazumder, A. K., Budhlakoti, N., Kumar, M., Pradhan, A. K., Kumar, S., Babu, P., ... & Gaikwad, K. B. (2024). Exploring the genetic diversity and population structure of an ancient hexaploid wheat species *Triticum sphaerococcum* using SNP markers. *BMC Plant Biology*, 24(1), 1188.
15. Metakovsky, E., Melnik, V. A., Pascual, L., & Wrigley, C. W. (2024). How important are genetic diversity and cultivar uniformity in wheat? The case of gliadins. *Genes*, 15(7), 927.
16. Muhammad, I., Rafii, M. Y., Ramlee, S. I., Nazli, M. H., Harun, A. R., Oladosu, Y., ... & Arolu, I. W. (2020). Exploration of bambara groundnut (*Vigna subterranea* (L.) Verdc.), an underutilized crop, to aid global food security: Varietal improvement, genetic diversity and processing. *Agronomy*, 10(6), 766.
17. Peng, F., Pi, Z., Li, S., & Wu, Z. (2024). Genetic diversity and population structure analysis of excellent sugar beet (*Beta vulgaris* L.) germplasm resources. *Horticulturae*, 10(2), 120.
18. Rani, R., Raza, G., Tung, M. H., Rizwan, M., Ashfaq, H., Shimelis, H., ... & Arif, M. (2023). Genetic diversity and population structure analysis in cultivated soybean (*Glycine max* [L.] Merr.) using SSR and EST-SSR markers. *PLoS One*, 18(5), e0286099.

19. Salem, K. F., Safhi, F. A., Alwutayd, K. M., Abozahra, M. S., Almohisen, I. A., Alsharari, S. F., ... & Ibrahim, A. A. (2024). Analysis of genetic diversity, population structure and phylogenetic relationships of rice (*Oryza sativa* L.) cultivars using simple sequence repeat (SSR) markers. *Genetic Resources and Crop Evolution*, 71(5), 2213-2227.
20. Salgotra, R. K., & Chauhan, B. S. (2023). *Genetic diversity, conservation, and utilization of plant genetic resources*. *Genes*. 2023; 14: 174.
21. Shahane, S. (2020). *Rice genotype and phenotype data* [Data set]. Kaggle. <https://www.kaggle.com/datasets/saurabhshahane/rice-genotype>.
22. Singh, A., & Pandey, M. K. (2024). Advances in okra (*Abelmoschus esculentus* L.) breeding: Integrating genomics for enhanced crop improvement. *Journal of Advances in Biology & Biotechnology*, 27(5), 397-407.
23. Singh, A., Singh, A., Barb, J., Mahama, A. A., & Suza, W. (2023). Genetic variation and germplasm usage. *Crop Improvement*.
24. Sruthi, S. R., Ivin, J. J. S., Williams, G., & Anbuselvam, Y. (2023). Genetic variability and diversity analysis in traditional rice (*Oryza sativa* L.) varieties of Tamil Nadu.
25. Tighankoumi, G., Dzidzienyo, D. K., Some, K., Tongoona, P., & Asante, I. (2024). Genetic diversity and unique genotype identification in togo's cassava (*Manihot esculenta* Crantz) germplasm using SNP markers. *Cogent Food & Agriculture*, 10(1), 2429741.
26. Tiruneh, A. A., Geletu, K. T., k Yao, N., & Weldegiorgis, K. D. (2024). The genetic diversity and population structure of wild and cultivated *Avena* species in Ethiopia using a SSR markers. *Heliyon*, 10(21).
27. Wang, Z., Zhou, F., Tang, X., Yang, Y., Zhou, T., & Liu, H. (2023). Morphology and SSR markers-based genetic diversity analysis of sesame (*Sesamum indicum* L.) cultivars released in China. *Agriculture*, 13(10), 1885.
28. Xiong, Y., Zhao, Y., He, Y., Zhao, L., Hu, H., Barrett, R. L., ... & Lu, L. (2024). Current progress and future prospects for understanding genetic diversity of seed plants in China. *Biological Diversity*, 1(1), 13-21.