

Diversity assessment and evaluation of Nepalese rice landraces for agronomic traits in Chitwan, Nepal

Shovit Khanal

khanal.shovit@gmail.com

Agriculture and Forestry University

Madhav Prasad Pandey

Agriculture and Forestry University

Ritesh Kumar Yadav

Agriculture and Forestry University

Suraj Shrestha

Agriculture and Forestry University

Raju Kharel

Agriculture and Forestry University

Research Article

Keywords: genetic diversity, landraces, multivariate analysis, Shannon Weiner diversity index, genetic variability

Posted Date: August 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10619051/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Diversity assessment and evaluation of Nepalese rice landraces for agronomic traits in Chitwan, Nepal

Shovit Khanal^{1*}, Madhav Prasad Pandey¹, Ritesh Kumar Yadav², Suraj Shrestha¹, Raju Kharel¹

¹ Department of Genetics and Plant Breeding, Faculty of Agriculture, Agriculture and Forestry University, Chitwan 44200, Nepal

² Department of Plant Pathology, Faculty of Agriculture, Agriculture and Forestry University, Chitwan 44200, Nepal

*Corresponding author email: khanal.shovit@gmail.com

ORCID:

Shovit Khanal: <https://orcid.org/0000-0002-8608-0034>

Ritesh Kumar Yadav: <https://orcid.org/0000-0002-5333-9652>

Suraj Shrestha: <https://orcid.org/0000-0003-2537-0449>

Raju Kharel: <https://orcid.org/0000-0003-0653-7979>

Abstract

Nepal has a diversity among rice landraces whose exploration aids as a valuable source of traits for breeding programs. However, their cultivation is declining, leading to a shrinking genetic base for rice. Classifying landraces helps identify desirable traits and understand genetic diversity. This study evaluated 46 Nepalese rice landraces from different regions of Nepal along with three released varieties used as checks. A field experiment was conducted in the Plant Breeding Block of Agriculture and Forestry University, Rampur, Chitwan, using alpha-lattice design with three replicates and seven blocks with seven plots per replication, and they were evaluated to assess the diversity. Observations were taken for twenty- nine different qualitative traits and twenty-two quantitative traits. Shannon Weiner index, Simpson index, and evenness were calculated to assess the diversity in the accessions. Evaluation of variability, genetic parameters, correlation, cluster analysis and principal component analysis (PCA) based on different agronomic traits was performed in this study. All the qualitative traits under the study showed polymorphism, except for the shape of ligule, and all the quantitative traits showed significant differences. Higher estimates of phenotypic and genotypic coefficients of variability were observed for ligule length, grain width, total number of grains per panicle, thousand-grain weight, straw yield, grain yield, and harvest index. Grain yield showed a high positive association with leaf width, number of effective tillers per plant, grain length, grain width, and thousand- grain weight, which can be selection indices for the improvement of grain yield. The landraces were classified into eight different clusters. PCA revealed that seven components were significant and contributed to 77.68% of the variance. The observed genetic diversity among landraces for various qualitative and quantitative traits could be pivotal information for further study of genetic variation of rice germplasm and can be effectively used in future rice improvement programs.

Keywords: genetic diversity; landraces; multivariate analysis; Shannon Weiner diversity index; genetic variability

1. Introduction

Rice (*Oryza sativa* L.), $2n=24$, is an autogamous cereal crop of the Gramineae family [1]. Six of the 11 genomes (labeled from AA to LL) discovered in 27 species are diploid ($n=12$: AA, BB, CC, EE, FF, and GG) and 5 are tetraploid ($n=24$: BBCC, CCDD, HHJJ, HHKK, and KKLL). Two cultivated species *O. sativa*, and *O. glaberrima* are members of AA genome group, which is the primary germplasm pool for improving rice (Jena 2010; Stein et al. 2018) while 16 of 25 wild species contain two sets of chromosomes (diploid genome), and 9 with four copies of each chromosome (tetraploid genome) (Fornasiero et al. 2022). Rice is the daily meal of over two-thirds of the world's population, providing about 20% of their daily calories (Salem et al. 2021). In 2024/25, almost three billion people consumed 527.05 million tonnes of rice (Statista and The Statistics Portal 2026), with Asians accounting for more than 75 percent of the total (Ishwarya Lakshmi et al. 2019). According to the MoALD (2025), rice was cultivated on 1,420,636 ha in Nepal in 2024/2025, with a production of 5,955,476 tonnes and a productivity of 4.19 t ha⁻¹. Rice consumption per capita in Nepal is 138 kilograms per year, accounting for 74.5 percent of total cereal consumption (Joshi and Ali 2021).

Nepal ranks 16th in rice production and 64th in rice productivity in the world (FAOSTAT 2021) where it is grown in three ecological regions (Terai and inner Terai- 67 to 900 meter above mean sea level (masl); mid hills-1000 to 1500 masl; and high hills-1500 to 3050 masl:- Chhumlung of Jumla valley), under two water regimes (irrigated and rainfed), and in two topographic conditions (lowland and upland) (Malla 2008; Gauchan and Pandey 2011; Gadal et al. 2019). *O. nivara*, *O. rufipogon*, *O. granulata*, and *O. officinalis* are wild rice species found in Nepal (Joshi et al. 2008). Till now, 156 rice varieties are recommended for various domains of Nepal, of which 78 are registered, and 78 are released including two hybrid rice varieties; Hardinath Hybrid 1 and Hardinath Hybrid 3 while 30 varieties have been denotified (SQCC 2023). According to Joshi et al. (2021), National Agriculture Genetic Resources Centre (NAGRC) is preserving over 2500 rice landraces that existed in Nepal, 150 of which are aromatic. Landraces are valued genetic resources for crop improvement due to their genetic diversity and tolerance to biotic and abiotic stresses (Kovach and McCouch, 2008; Sang and Ge 2013). Landraces are primordial kinds that have evolved over millennia by natural and artificial selection but without a deliberate breeding effort.

In Nepal, just 13% of the 275 crop varieties were generated solely by landrace selection, while 16% were developed through the crossing and 71% were developed through introduction (Joshi 2017). The area covered by rice landraces in Nepal decreased from 70% to 10% in 2017 (CDD 2017) as a result genetic base of rice is shrinking. Only 12 of the 78 released rice varieties in Nepal were produced using six native landraces. Despite our vast rice diversity, foreign germplasm is often employed as parents in rice breeding operations, and Nepal is around 95% dependent on foreign rice germplasm for varietal development (Joshi 2017). Characterizing a crop species diversity is an important tool for improving it and ensuring future food security. Plant breeding activities rely heavily on germplasm agro morphological characterization. Understanding the amount of diversity and specifying desired characters for crop improvement programs necessitates landrace classification (Ahmed et al. 2016). Qualitative characteristics are employed as markers in the identification of rice landraces that are less dependent on environmental responses (Singh et al. 2018). Hence, this experiment was carried out to assess the diversity of Nepalese rice landraces based on different qualitative and quantitative agro-morphological traits and to estimate their genetic parameters.

Materials and Methods

Forty-six landraces of rice collected from the different community seed banks (CSB) of Jhapa, Bara, Nawalpur, and Dang along with three released varieties: Sabitri, Sunaulo Sugandha, and Lalka Basmati collected from National Rice Research Programme, Hardinath, Dhanusha as checks were used in this study (Table 1). The experiment was performed in the research block of the Department of Genetics and Plant Breeding, AFU, Rampur, Chitwan from June to November 2019. The experimental site was located at 27°64'N latitude and 84°35' E longitude and at an altitude of 229.2 meters above sea level. The experiment was carried out in alpha lattice design with three replications, seven blocks per replication, and seven plots on each block. Crop geometry was 20 cm between a row and 20 cm between plants within a row.

During the experimental period, the highest amount of precipitation occurred in July (519.70 mm). Likewise, the maximum temperature was recorded in June (36.40°C) while the minimum temperature was in November (17.36°C) and the average relative humidity ranged from 81.41% to 92.22% recorded at National Maize Research Programme, Chitwan (Table S1). There was separate dry nursery beds established for each genotype and the main field was applied with well-decomposed farmyard manure (FYM) at the rate of 6 tonnes per hectare. Nitrogen (N), Phosphorous (P₂O₅), and potassium (K₂O) of recommended doses (100, 50, and 50 kg ha⁻¹, respectively) were calculated and applied to the field. Nitrogen was applied in three equal split doses i.e. 1/3rd as a basal dose, 1/3rd at the active tillering stage (30 DAT), and 1/3rd at 60 DAT. Two manual weeding and regular irrigation were done to maintain a 5cm water depth up to the grain filling stage.

Table 1 The complete list of genotypes included in the study

S.N.	Accessions	District	S.N.	Accessions	District
1	Anga	Bara	26	Jaagad	Bara
2	Aanp Jhutthe	Jhapa	27	Jeerasari	Jhapa
3	Amaghauj	Bara	28	Jhorpal Basmati	Jhapa
4	Aandhi Dhan	Jhapa	29	Kalo Basmati Junge	Jhapa
5	Bagar Anadi	Jhapa	30	Kalo Nuniya 1	Jhapa
6	Baharni	Bara	31	Kalo Nuniya 10	Jhapa

7	Bakuli Anadi	Nawalpur	32	Kalo Nuniya Thulo	Jhapa
8	Balamsar	Bara	33	Karangi	Bara
9	Bangalia	Bara	34	Kariya Khera	Bara
10	Basmati Kalo Tude	Jhapa	35	Kasturi	Bara
11	Basmati Rato Tude	Jhapa	36	Lajhi	Bara
12	Bhabbhi	Bara	37	Lalka Basmati	NRRP
13	Biramphul	Jhapa	38	Lalka Jesariya	Bara
14	Budhidayan 3	Bara	39	Mansara	Bara
15	Chameli	Nawalpur	40	Nyauri	Bara
16	Champasari	Jhapa	41	Parewapwankh	Bara
17	Chhatraj	Bara	42	Ramajwine	Bara
18	Devsar	Bara	43	Rango	Dang
19	Dhankute	Jhapa	44	Rato Basmati Sano	Jhapa
20	Dhudhraj	Jhapa	45	Rato Basmati Thulo	Jhapa
21	Diruwa	Bara	46	Sabitri	NRRP
22	Gajargawl	Bara	47	Sanojeera	Jhapa
23	Gakhulachan	Bara	48	Seto Basmati	Jhapa
24	Ghyukumar	Jhapa	49	Sunaulo Sugandha	NRRP
25	Ghusara	Bara			

Note: NRRP: National Rice Research Program, Dhanusha

2.1 Observations recorded

Rice descriptor developed by Biodiversity International, IRRI, and WARDA (2007) was used to characterize the qualitative and quantitative traits of the landraces under study.

2.2 Statistical analysis

2.2.1 Diversity index

2.2.1.1 Shannon-Weiner diversity index (H')

The Shannon-Weiner diversity index is a well-known diversity index and Shannon entropy developed by Claude E. Shannon and Norbert Wiener in 1949 (Spellerberg and Fedor 2003). It is denoted by (H') and each morphological trait was calculated to explain its state of diversity quantitatively by using the following formula (Hien et al. 2007):

$$H' = - \sum_{i=1}^n pi \ln pi$$

$$pi = n/N$$

Where,

H' = Standardized Shannon- Weaner index

N = total number of observations for that trait

pi = proportion of i^{th} trait

$\ln$ = natural log of the number

n = frequency of a phenotypic class of that trait

The value of H' ranges from 0 to 5, usually ranging from 1.5 to 3 where higher values indicate higher levels of diversity of traits.

2.2.1.2 Simpson's diversity index (D)

According to Risser and Rice (1971) and Whittaker (1972), Simpson's diversity index is a measure of the dominance or relative abundance of each of the species in a group. It is denoted by " D " and was calculated as follows given by Simpson (1949):

$$\text{Simpson's index } (D) = \sum \frac{n(n-1)}{N(N-1)}$$

where n is the number of genotypes bearing a character, and N is the total number of genotypes.

The value of D ranges from 0 to 1, where 0 represents no diversity and 1, infinite diversity.

2.2.1.3 Evenness (J)

The evenness value represents equality in the distribution of qualities in the population (Magurran 2004). It is measure of closeness in numbers of each species in a community. It is denoted by "J" and was calculated as given by Magurran (2004).

$$\text{Evenness } (J) = \frac{H'}{\ln(S)}$$

Where S is the total number of variation cases.

The value of J ranges from 0 to 1. Higher values indicate higher levels of evenness. Low J indicates few species dominate the community (Pielou 1966).

2.2.2 Analysis of variance

For the quantitative characters, the analysis of variance (ANOVA) for alpha lattice design was performed using the F test, and treatment means were compared by LSD test at a 5% level (Gomez and Gomez 1984). The ANOVA model was: $y_{ikl} = \mu + g_i + r_k + b_{lk} + e_{ikl}$.

Where μ is the overall mean, g_i is the effect of the i^{th} genotype, r_k is the effect of the k^{th} replicate, b_{lk} is the effect of the l^{th} block within the k^{th} replicate and e_{ikl} the error. This model was fitted to residual maximum likelihood (ReML).

2.2.3 Genetic variability

Genetic variability parameters provide information on the degree of variation in the landraces and are therefore important foundations for the genetic development of the trait. The degree of genetic variability influences the planning and execution of any breeding program for quantitative trait improvement (Adhikari et al. 2018). The variations between individuals in a population caused by genetic composition and growing environment are known as phenotypic variability. Based on the variance component obtained from ANOVA and the significance, the genetic parameters genotypic and phenotypic variances (VG and VP), genotypic and phenotypic coefficient of variation (GCV and PCV), broad-sense heritability, expected genetic advance (GA) from the selection of the best 5% genotypes were estimated.

A higher phenotypic coefficient of variance (PCV) than the genotypic coefficient of variance (GCV) indicates the environmental influence on the expression of traits. Genotypic and phenotypic variance were obtained from the analysis of the variance table according to Bernardo (2002).

$$VG = \frac{\text{Genotypic mean sum of square} - \text{Error mean sum of square}}{r}$$

$$VP = VG + (\text{Error mean sum of square})/r$$

The phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) were estimated according to Sing and El-Bizri (1992).

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

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

Where,

σ_g^2 = genotypic variance, σ_p^2 = phenotypic variance, $\bar{X}$ = mean value of the trait

PCV and GCV were categorized as 0-10%: low, >10-20%: moderate, and above 20%: high (Sivasubramanian and Menon 1973).

2.2.4 Heritability and genetic advance

The ratio of variation resulting from genotype differences to the total phenotypic variance for a characteristic in a population is known as its heritability, and it indicates the aspect of a character that is transmitted to progenies. Higher heritability and remarkable genetic progress may be regarded as the most acceptable qualities for trait improvement and selection to create stable and high-yielding early rice genotypes under rainfed conditions (Tiwari et al. 2019).

Broad sense heritability was calculated according to Falconer (1960).

$$\text{Heritability } (h^2) = \frac{\text{Genotypic variance } (\sigma_g^2)}{\text{Phenotypic variance } (\sigma_p^2)}$$

Heritability was categorized as 0-30% as low; >30-60% as moderate; >60% as high heritability (Robinson et al. 1949).

Estimated heritability along with genetic advancement is more precise in predicting the genetic gain under selection than heritability alone. This shows that these qualities are predominantly under genetic control, and that selection for them may be accomplished through phenotypic performance. Genetic advance as % of the mean was estimated according to (Hanson et al. 1956). The genetic advance was calculated as per the following formula

$$GA = K \times \sqrt{\sigma_p^2} \times h^2$$

Where, K= Selection intensity (5%=2.06)

$$\text{Genetic advance as \% of the mean (GAM)} = \frac{GA}{\text{mean value of the trait}} \times 100\%$$

Genetic advance as % of mean is categorized as 0 to 10% low GAM, >10 to 20% moderate GAM, and above 20% high GAM (Johnson et al. 1955).

2.2.5 Character association

The degree of association and link between two variables is measured by the correlation coefficient. The measure of the correlation between multiple characters may help the plant breeder to know how the improvement of one character will bring simultaneous changes in other characters. Pearson correlation coefficients were calculated and presented through a heat map using the Metan package in R.

2.2.6 Multivariate analysis

2.2.6.1 Cluster analysis

Cluster analysis refers to a series of multivariate approaches whose major goal is to group genotypes based on the characters they possess, such that genotypes with comparable descriptions are mathematically grouped into the same cluster. Genotypes within-cluster have higher homogeneity and between clusters have higher heterogeneity. Thus, if the categorization is effective, genotypes inside a cluster will be closed, whereas genotypes from other clusters would be further apart and attributed to unequal adaptation to different agroecosystems. For cluster analysis. The landraces were grouped in different clusters using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method. Significant differences among genotypes for twenty-two characters justified further calculation of Euclidean generalized distance (D^2) based on average linkage method which measures the genetic divergence among the landraces and their grouping was done by ward. D^2 method and clustering was performed using the cluster tree method obtained through ggtree package in R for assessing the genetic diversity of genotypes collections in terms of attributes considered collectively (Sneath and Sokal 1973; Worede et al. 2014).

2.2.6.2 Principal Component Analysis (PCA)

PCA is a multivariate approach for assessing correlations between several characters which are responsible for the observed genotypic variation within a group (Sinha and Mishra, 2013). It decreases the dimensionality of multivariate data by reducing interrelationships between variables. As variation occurs often in plants for yield and yield-related variables, PCA identifies patterns and reduces redundancy in data sets (Maji and Shaibu 2012). The primary advantage of PCA is that it quantifies the significance of each dimension for linking the variability of a dataset (Shoba et al. 2019). It can be used to generate a 2- or 3-dimensional scatter plot, such that the geometrical distances between individuals in the plot represent the genetic distances between them. As observations, PCA analyzes data that consists of numerous intercorrelated quantitative dependent variables (Mahendran et al. 2015). PCA, eigenvalues, eigenvectors, and 2D biplots were obtained using FactoMineR and factoextra packages (Yan and Fregeau-Reid 2018; Banyai et al. 2020).

3. Results and Discussion

3.1 Agro-morphological characterization

Polymorphism was found in all the qualitative traits with varying degrees of contribution to the phenotypic diversity except in the shape of ligule. Descriptor states and their frequency for each 28 qualitative traits are depicted in Table

2. Agro-morphological characterization of germplasm is necessary for obtaining preliminary information during the breeding program as it describes crop species. Qualitative traits are employed as markers in the identification of rice landraces that are less dependent on environmental responses (Kinoshita et al. 2017). The uniqueness shown by a specific landrace regarding this characteristic could be further utilized for candidate variety development.

Characterization of leaf characters revealed that out of 49 landraces, 31 (63.27%) landraces exhibited green color, 9 landraces (18.37%) showed purple color, 6 landraces (12.24%) had light purple, and the remaining 3 landraces (6.12%) had purple lines. This suggests that the purple pigmentation of basal leaf sheath had been lost during evolution and the green types evolved further. However, additional studies are required for verification (Parikh et al. 2012; Gour et al. 2019). Based on the intensity of green color on the leaf blade, 30 landraces (61.22%) showed medium green color, 12 landraces (24.49 %) showed light green color, and 7 landraces (14.29%) showed dark green color. Anthocyanin coloration in leaf blade was present on only 4 landraces (8.16%) but absent in the remaining (91.8%) and its distribution was found on tips by 32 landraces (65.31%), on blotches by 9 landraces (18.37%), and on margins by 4 landraces (8.16) while none showed uniform distribution. Anthocyanin was absent in 36 landraces (73.47%) while only 9 landraces (18.37%) showed the weak intensity of anthocyanin and 4 showed medium landraces (8.16%). Anthocyanin, a group of glycosylated flavonoids, accumulates in plant tissues or organs in rice which produces diverse color ranges from red to purple phenotypes (Chaves-Silva et al. 2018). These are induced under various biotic and abiotic stresses as cell stress coping mechanisms. These pigments help reduce and mitigate photo-oxidative stress like high light and low temperatures and are also responsible for the modulation of hormonal response (Chin et al. 2016; Gould et al. 2018).

The pubescent leaf blade is the dominant type shown by 39 landraces (79.60%) whereas 5 landraces (10.20%) showed intermediate type and the remaining 5 landraces (10.20%) showed glabrous type of leaf blade. Similarly, 21 landraces (42.86%) showed drooping leaf blade attitude, 15 landraces (30.61%) showed horizontal leaf blade attitude and 13 landraces (26.53%) showed erect leaf blade attitude. Pubescence or trichomes on the leaf surface contribute to photosynthesis and are involved in transpiration, respiration & protection from pest attacks (Hu et al. 2013; Sharmitha et al. 2019).

Auricle color was absent in 13 landraces (26.53%), 21 landraces (42.86%) showed whitish auricle color, 5 landraces (10.2%) showed yellowish-green auricle color, 3 landraces (6.12%) showed purple lines, 5 landraces (10.2%) showed light purple color, and 2 landraces (4.08%) purple. Likewise, 26 landraces (53.07%) showed light green collar color, 18 landraces (36.73%) showed green-collar color, and the remaining 5 landraces (10.2%) showed purple collar color. No variation was found for the shape of ligule as all the landraces showed the two-cleft shape of ligule and color of ligule showed low variability as 32 landraces (65.3%) showed whitish ligule, 8 landraces (16.33%) showed light purple ligule, 5 landraces (10.2%) showed purple lines, 4 landraces (8.16%) showed the purple color and none showed yellowish green color of ligule. A ligule is whitish thin and tongue-like membrane that binds leaf sheath tightly against auricles to prevent rainwater or dust from entering (Lee et al. 2007).

The angle of the flag leaf is a crucial growth characteristic in which maximal photosynthesis occurs and provides a majority of carbohydrates needed for grain filling (Acevedo-Siaca et al. 2021). This may be related to increased light penetration in the crop canopy as a result of upright leaf growth (Zaman et al. 2005). Twenty-two landraces (44.9%) had erect flag leaf attitude, 10 landraces (20.41%) each had semi-erect and intermediate flag leaf attitude, 6 landraces (12.2%) had horizontal, and only one (2.04%) showed drooping flag leaf. The intermediate type was the dominant attitude of flag leaf blade during the reproductive stages showing 25 landraces (51.02%).

Sixteen landraces (32.65%) showed intermediate culm habit, 15 landraces (30.62%) showed spreading culm habit, 13 landraces (26.53%) showed open culm habit, and 5 landraces (10.2%) showed erect culm habit. Culm habits are crucial as they determine light interception and self-shading effect during the active tillering and flowering stage (Taferre and Irie 2019). Similarly, 16 landraces (32.65%) showed very weak culm lodging resistance, while 11 landraces (22.45%) showed strong, 10 landraces (20.41%) showed intermediate, 7 landraces (14.29%) showed weak and 5 landraces (10.2%) showed very strong culm lodging resistance. About 20-40% of grain carbohydrate accumulation is attributed to culm carbohydrates (Zhang et al. 2017). Generally, landraces with long culms are susceptible to lodging also due to low densities of cellulose and lignin in the culm cell wall. This makes harvesting difficult and seeds soaked in water germinate viviparous (Nomura et al. 2021). Anthocyanin coloration of nodes was absent in the majority of landraces (75.52%) while green (59.18%) was the dominant underlying node color in culm. Likewise, internode anthocyanin of culm was absent in the majority (61.22%) of landraces while 16 landraces (32.65%) showed purple lines, 2 landraces (4.08%) showed purple, and only one (2.04%) showed light gold coloration of culm internode. For underlying culm

internode anthocyanin coloration, 21 landraces showed no coloration, 16 landraces (32.65%) showed light gold color and 12 landraces (24.49%) showed green color. Similar results were found by Sharma et al. (2020) for these culm-related traits.

Awns were present in only 29 landraces while the remaining were awnless. They were present in tips only in 17 landraces (34.69%), on the upper half only in 10 landraces (20.41%), and along the whole length in 2 landraces (4.08%). Awn color was white in 11 landraces (22.45%), purple in 7 landraces (14.29%), straw in 6 landraces (12.24%), red in 4 landraces (8.16%), and brown in only one landrace (2.04%). Awns in wild rice help with seed dissemination and protection from animals and also have physiological and biological buffers to help plants adjust to the environment (Gour et al. 2019). However, they are considered unfavorable during seed storage and processing and were partially or eliminated through artificial selection during domestication (Qin et al. 2021).

Lemma and palea color showed high variability among the landraces. Eleven landraces (22.45%) showed green lemma and palea color, while purple spots, purple furrows, and brown furrows coloration were shown by 6 landraces (12.24%) each, 5 landraces (10.2%) showed purple color, 4 landraces (8.16%) showed reddish lemma and palea color. Golden furrow, yellowish-green, white, purple furrow on the green, and purple shades of lemma and palea were shown by 2 landraces (4.08%) each and the remaining one landrace with brown tawny lemma and palea color. The green color was the dominant color of the apiculus of the lemma whereas 7 landraces (14.29%) showed brown tawny color, 4 landraces (8.16%) each showed purple and purple apex color, 3 landraces (6.12%) showed red color, 2 landraces (4.08%) showed straw color and only one showed the white color of apiculus. The color of the lemma and palea determines the color of the seed (Hidayatun et al. 2020). The coloration of flowering structures: lemma, palea, and apiculus attracts insects and animals and helps in pollination and seed dispersal (Chin et al. 2016).

Panicle exertion showed 55.1% well exerted, 20.41% well exerted, 14.29% exerted, and 10.2% partly exerted panicle. Secondary branching of panicle showed 38.78% dense, 30.61% sparse, and 24.49% clustering and was absent in 6.12% of collected landraces. The attitude of the main axis of the panicle showed 53.06% drooping, 32.65% slightly drooping, 10.2% semi-upright, and 4.08 % upright attitude of the main axis among the landraces. The attitude of branches of the panicle showed 55.10 % drooping, 32.65% horizontal, and 12.24% semi-erect attitude. In addition to controlling the effectiveness of water and nutrient transmission from leaves and stems to grains, panicle exertion aids in the regulation of the panicle enclosure in rice. When panicles are enclosed, the yield is significantly decreased, particularly in hybrid rice varieties (Zhan et al. 2019). Usually, erect panicle types with more branches and spikelets are preferred during selection than the spreading types and, thus, easy harvesting and higher grain number per panicle (Khush 2000; Kumawat 2022).

Among the landraces, 36.73% showed medium slender, 32.65% medium bold, 14.29% long slender, 10.2% short slender, 4.08% short bold, and 2.04 % long bold grain shape. Similarly, observation on decorticated grain color revealed that white is the dominant (46.94%) color among the landraces with 18.37 % red, 16.33% light brown, 8.16% light red, 6.12% dark red, and 4.08% greenish-white color. These traits are important and optimal characteristics and determine the consumer preference and commercial success of variety. Different grain shapes are preferred by people from different regions (Huang et al. 2013). The pericarp of rice grains accrues different phenolic compounds which gives it a different color (Walter et al. 2013). Accumulation of proanthocyanidin produces “red rice” and anthocyanin produces “black rice” (Maeda et al. 2014). These compounds have antioxidant activity and health benefits too (Hou et al. 2010).

Table 2 Frequency distribution of different qualitative characters observed in 49 rice genotypes

Qualitative characters	Descriptor's states	Frequency	Proportion (%)
Basal leaf sheath color	1. Green	31	63.27
	2. Light purple	6	12.24
	3. Purple	9	18.37
	4. Purple lines	3	6.12
Leaf blade (Green color)	1. Dark	7	14.29
	2. Light	12	24.49
	3. Medium	30	61.22
Leaf blade (Anthocyanin coloration)	1. Absent	45	91.8
	2. Present	4	8.16
Leaf blade (Distribution of anthocyanin)	1. Absent	4	8.16
	2. Blotches only	9	18.37

Qualitative characters	Descriptor's states	Frequency	Proportion (%)
	3. Margins only	4	8.16
	4. Tips Only	32	65.31
	5. Uniform	0	0.0
Leaf sheath (Anthocyanin)	1. Absent	36	73.47
	2. Medium	4	8.16
	3. Weak	9	18.37
	4. Strong	0	0.0
Leaf blade (Pubescence)	1. Glabrous	5	10.20
	2. Intermediate	5	10.20
	3. Pubescent	39	79.60
Leaf blade attitude	1. Drooping	21	42.86
	2. Erect	13	26.53
	3. Horizontal	15	30.61
Auricle color	1. Absent	13	26.53
	2. Light Purple	5	10.20
	3. Whitish	21	42.86
	4. Yellowish green	5	10.20
	5. Purple	2	4.08
	6. Purple Lines	3	6.12
Collar color	1. Green	18	36.73
	2. Light green	26	53.07
	3. Purple	5	10.2
	4. Absent	0	0.0
Shape of ligule	1. Two-clefted	49	100.0
	2. Absent	0	0.0
	3. Truncate	0	0.0
	4. Acute	0	0.0
Color of ligule	1. Light purple	8	16.33
	2. Purple	4	8.16
	3. Purple lines	5	10.2
	4. Whitish	32	65.31
	5. Yellowish Green	0	0.0
	6. Absent	0	0.0
Flag leaf (Attitude of blade)	1. Drooping	1	2.04
	2. Erect	22	44.9
	3. Horizontal	6	12.24
	4. Intermediate	10	20.41
	5. Semi-erect	10	20.41
Culm habit (Angle)	1. Erect	5	10.2
	2. Intermediate	16	32.65
	3. Open	13	26.53
	4. Semi-erect	0	0.00
	5. Spreading	15	30.62
Awns (Colors of awns)	1. Absent	20	40.82
	2. Brown	1	2.04
	3. Purple	7	14.29
	4. Red	4	8.16
	5. Straw	6	12.24
	6. White	11	22.45
	7. Gold	0	0.0
	8. Light Green	0	0.0
	9. Black	0	0.0

Qualitative characters	Descriptor's states	Frequency	Proportion (%)
Awns (Distribution of awns)	1. Absent	20	40.82
	2. Tips only	17	34.69
	3. Whole length	2	4.08
	4. Upper Quarter Only	0	0
	5. Upper Half Only	10	20.41
	6. Upper Three Quarters	0	0.0
Lemma (Color of Apiculus)	1. Brown tawny	7	14.29
	2. Green	28	57.14
	3. Purple	4	8.16
	4. Purple apex	4	8.16
	5. Red	3	6.12
	6. Straw	2	4.08
	7. White	1	2.04
	8. Red apex	0	0
	9. Black	0	0
Lemma and Palea color	1. Brown furrows	6	12.24
	2. Brown tawny	1	2.04
	3. Gold furrows	2	4.08
	4. Green	11	22.45
	5. Purple	5	10.2
	6. Purple furrows	6	12.24
	7. Purple spots	6	12.24
	8. Reddish	4	8.16
	9. Yellowish green	2	4.08
	10. White	2	4.08
	11. Green striped white	0	0.0
	12. Blackish brown	0	0.0
	13. Black	0	0.0
	14. Purple shade	2	4.08
	15. Purple furrows on green	2	4.08
Flag leaf (Attitude of blade) (Late)	1. Descending	6	12.24
	2. Drooping	10	20.41
	3. Erect	5	10.20
	4. Horizontal	3	6.12
	5. Intermediate	25	51.02
Culm (Anthocyanin coloration of nodes)	1. Absent	37	75.52
	2. Light purple	4	8.16
	3. Purple	5	10.20
	4. Purple lines	3	6.12
Culm (underlying node color)	1. Absent	14	28.57
	2. Green	29	59.18
	3. Light gold	6	12.24
Culm (Internode anthocyanin)	1. Purple lines	16	32.65
	2. Absent	30	61.22
	3. Purple	2	4.08
	4. Light gold	1	2.04
Culm (Underlying internode coloration)	1. Green	12	24.49
	2. Light gold	16	32.65
	3. Absent	21	42.85
Culm lodging tolerance	1. Intermediate	10	20.41
	2. Strong	11	22.45
	3. Very strong	5	10.2

Qualitative characters	Descriptor's states	Frequency	Proportion (%)
	4. Very weak	16	32.65
	5. Weak	7	14.29
Panicle (Attitude of main axis)	1. Semi-upright	5	10.2
	2. Slightly drooping	16	32.65
	3. Drooping	26	53.06
	4. Upright	2	4.08
Panicle (Attitude of branches)	1. Semi-erect	6	12.24
	2. Drooping	27	55.1
	3. Horizontal	16	32.65
Panicle (Exsertion)	1. Exserted	7	14.29
	2. Moderately Well Exserted	10	20.41
	3. Well Exserted	27	55.1
	4. Partly Exserted	5	10.2
Panicle (Secondary branching)	1. Dense	19	38.78
	2. Sparse	15	30.61
	3. Absent	3	6.12
	4. Clustering	12	24.49
Shape of grain	1. Medium bold	16	32.65
	2. Medium slender	18	36.73
	3. Long slender	7	14.29
	4. Short slender	5	10.2
	5. Short bold	2	4.08
	6. Long bold	1	2.04
Decorticated grain color	1. White	23	46.94
	2. Red	9	18.37
	3. Light brown	8	16.33
	4. Light red	4	8.16
	5. Dark red	3	6.12
	6. Greenish white	2	4.08

3.2 Assessment of diversity indices

Shannon Weiner Diversity Index (H'), shown in Table 3 ranged from 0.41 to 3.28 inferring tremendous diversity in the collections for the qualitative characters among the 49 landraces used in the experiment. The highest value of H' was inferred for the color of lemma and palea (3.28) followed by culm lodging resistance (2.21), and auricle color (2.12) which indicates that the landraces were most diversified in these traits. However, relatively lower values of H' were found for leaf blade anthocyanin (0.41), followed by leaf blade pubescence (0.92) and leaf sheath anthocyanin (1.06) indicating that these traits were least diverse among the experimented landraces (Joshi 2021). Two-cleft ligule was observed in all genotypes and the value of (H') was found to be zero with no diversity at all.

Simpson's diversity index (D), presented in Table 3 ranged from 0.10 to 0.85. The lowest index ($D = 0.10$) for lemma and palea color indicates that there was the minimum abundance of a single type of lemma and palea coloration among the genotypes used while the trait with the highest index ($D = 0.85$) was shown by anthocyanin coloration of leaf blade inferring that dominance of trait absence of anthocyanin coloration was observed in forty-five genotypes among forty-nine genotypes. The common or dominant traits are given more weightage than the rare traits (Fakhar Izadi and Keshtkar 2020).

Along with this, the estimation of evenness, explained in Table 3 revealed that leaf blade attitude was most even with value ($J=0.98$) since attitude drooping, erect and horizontal were equally distributed among the landraces used in the experiment, followed by underlying internode coloration of culm ($J=0.97$), and culm lodging resistance ($J=0.95$). The most uneven traits was the anthocyanin coloration of the leaf blade ($J=0.41$) followed by the color of the ligule ($J=0.56$). The value of evenness describes the equality in the distribution of the traits in the population (Poudel et al. 2020).

Table 3 Diversity indices of qualitative characters of rice landraces

Character	Shannon-Weiner index (H')	Simpson's index (D)	Evenness (J)
Basal leaf sheath color	1.48	0.44	0.74
Leaf blade (Green color)	1.33	0.44	0.84
Leaf blade (Anthocyanin coloration)	0.41	0.85	0.41
Leaf blade (Distribution of anthocyanin)	1.44	0.46	0.62
Leaf sheath (Anthocyanin)	1.06	0.57	0.53
Leaf blade (Pubescence)	0.92	0.65	0.58
Leaf blade attitude	1.55	0.33	0.98
Auricle color	2.12	0.26	0.82
Collar color	1.34	0.41	0.67
Shape of ligule	0.00	1.00	0
Color of ligule	1.45	0.46	0.56
Flag leaf (Attitude of blade)	1.63	0.29	0.7
Culm habit (Angle)	1.88	0.27	0.81
Awns (Colors of awns)	2.19	0.24	0.69
Awns (Distribution of awns)	1.71	0.32	0.66
Lemma (Color of Apiculus)	1.87	0.35	0.59
Lemma and Palea color	3.28	0.10	0.84
Flag leaf (Attitude of blade) (Late)	1.90	0.32	0.82
Culm (Anthocyanin coloration of nodes)	1.18	0.58	0.59
Culm (underlying node color)	1.33	0.44	0.84
Culm (Internode anthocyanin)	1.26	0.47	0.63
Culm (Underlying internode coloration)	1.54	0.34	0.97
Culm lodging tolerance	2.21	0.21	0.95
Panicle (Attitude of the main axis)	1.52	0.39	0.76
Panicle (Attitude of branches)	1.36	0.41	0.86
Panicle (Exsertion)	1.66	0.36	0.83
Panicle (Secondary branching)	1.78	0.29	0.89
Shape of grain	2.09	0.26	0.81
Decorticated grain color	2.11	0.28	0.81

3.3 Analysis of variance

The mean squares for yield and yield component characters of forty-nine rice genotypes, which showed highly significant differences in all studied characters (Table S2) pointed out that there is an existence of variation in the landraces studied and there is ample scope for the selection of these characters for rice improvement (Bitew 2016).

Table 4 Variability in 22 quantitative characters of 49 rice genotypes

Characters	Range		Mean	CV (%)		h_b^2	GA	GAM
	Minimum	Maximum		PCV	GCV			
PLL	40.39	60.59	50.06 ± 0.46	9.64	8.84	0.84	8.37	16.72
PLW	0.75	1.46	1.08 ± 0.02	17.35	15.95	0.85	0.32	30.21
LL	0.97	2.65	1.78 ± 0.03	21.85	21.23	0.94	0.76	42.50
FLL	22.95	48.28	32.07 ± 0.49	16.67	15.60	0.88	9.65	30.09
FLW	0.87	2.03	1.32 ± 0.02	19.26	18.48	0.92	0.48	36.51
NTI	9.65	24.77	16.94 ± 0.33	19.09	17.03	0.80	5.30	31.31
CD	4.28	7.82	6.18 ± 0.06	11.68	11.46	0.96	1.43	23.18
DFF	84.02	131.25	117.78 ± 0.78	7.99	7.95	0.99	19.19	16.30
ET	8.63	16.46	11.49 ± 0.22	16.61	14.81	0.79	3.12	27.19
DTM	111.65	164.01	147.94 ± 1.00	8.07	8.05	0.99	24.44	16.52
PH	105.4	192.26	144.86 ± 1.76	14.60	14.48	0.98	42.87	29.60
PL	20.12	30.08	25.36 ± 0.23	7.99	6.17	0.59	2.49	9.83
GL	4.54	8.63	7.48 ± 0.08	12.42	12.31	0.98	1.88	25.11
GW	1.04	3.19	1.99 ± 0.05	28.22	27.67	0.96	1.11	55.87
GLWR	2.49	6.92	4.03 ± 0.08	25.25	24.62	0.95	1.99	49.46

GPP	93.06	224.56	141.72 ± 2.67	20.29	18.80	0.86	50.84	35.87
UGPP	8.43	58.53	21.53 ± 0.96	44.27	39.43	0.79	15.58	72.33
STR	7.32	29.81	15.38 ± 0.58	34.78	27.94	0.65	7.11	46.25
TGW	10.34	32.27	21.69 ± 0.40	21.98	21.83	0.99	9.69	44.67
SY	3613.09	11108.04	7790.72 ± 200.26	27.08	24.93	0.85	3684.22	47.29
GY	763.7	4036.06	2263.30 ± 88.68	44.73	43.38	0.94	1961.75	86.68
HI	0.08	0.45	0.22 ± 0.007	37.23	35.40	0.90	0.16	69.32

PLL= penultimate leaf length (cm), PLW= leaf width of blade (cm), LL= length of ligule (cm), FLL= flag leaf length (cm), FLW= flag leaf width (cm), NTI = number of tillers per plant, CD = culm diameter (mm), DFF= days of fifty percent flowering, ET = number of effective tillers, DTM = days to maturity, PH = plant height (cm), PL = panicle length (cm), GL = grain length (mm), GW = grain width (mm), GLWR = grain length width ratio, GPP = number of grains per panicle, UGPP= number of unfilled grains per panicle, STR % = sterility percentage, TGW= thousand grain weight (gm), SY= straw yield (kg. ha-1), GY= grain yield (kg. ha-1), HI = Harvest index, GCV= genotypic coefficient of variation, PCV= phenotypic coefficient of variation, h²_b = broad sense heritability, GA= genetic advance, GAM= genetic advance as percent of mean.

3.4 Estimate of genetic parameters

3.4.1 Phenotypic and genotypic coefficient of variation (PCV and GCV)

PCV and GCV values for quantitative characters are depicted in Table 4. Higher estimates of PCV and GCV were observed for length of ligule (21.85% and 21.23%), grain width (28.22% and 27.67%), grain length-width ratio (25.25% and 24.62%), sterility percentage (34.78% and 27.94%), thousand-grain weight (21.98% and 21.83%), straw yield (27.08% and 24.93%), grain yield (44.73% and 43.38%) and harvest index (37.23% and 35.40%) respectively which indicates the existence of wide genetic base among these landraces and possibility of genetic improvement through direct selection for these characters (Parikh et al. 2012). Moderate PCV was observed for penultimate leaf width (17.35%), flag leaf length (16.67%), flag leaf width (19.26%), numbers of tillers per plant (19.09%), culm diameter (11.68%), effective tillers per plant (16.61%), plant height (14.60%), and grain length (12.42%); however, penultimate leaf length (9.64%), days to 50% flowering (7.99%), days to maturity (8.07%), and panicle length (7.99%) display low (<10%) PCV value. Characters with moderate (>10%) GCV values were width of penultimate leaf (15.95%), flag leaf length (15.60%), flag leaf width (18.48%), culm diameter (11.46%), total tillers per plant (17.03%), effective tillers/plant (14.81%), plant height (14.48%), grain length (12.31%) and total grain per panicle (18.80). However, the length of the penultimate leaf (8.84%), days to 50% flowering (7.95%), days to maturity (8.05%), and panicle length (6.17%) display low (<10%) GCV value. PCV and GCV recorded moderate values for penultimate leaf width, flag leaf length, flag leaf width, number of tillers per plant, effective tillers per plant, culm diameter, plant height, and grain length. Similarly, lower estimates of PCV and GCV were observed in penultimate leaf length, days of flowering, days to maturity, and panicle length. The magnitude of PCV, in general, was found to be higher than GCV for all the characters studied which might be attributable to a higher degree of genotype-environment interaction but there were only slight differences indicating that these characters were less influenced by environment, genotype had a larger role in trait expression, and might contribute to high heritability (Immanuel et al. 2011).

3.4.2 Broad sense heritability (h^2_b) and genetic advance (GA)

Higher heritability (>60%) was shown by penultimate leaf length (84.22%), penultimate leaf width (84.54%), ligule length (94.44%), flag leaf length (87.65%), flag leaf width (92.01%), number of tillers per plant (79.63%), culm diameter (96.35%), days to fifty percent flowering (98.96%), numbers of productive tillers per plant (79.48%), days to maturity (99.36%), plant height (98.39%), grain length (98.13%), grain width (96.11%), grain length-width ratio (95.09%), total grains per panicle (85.82%), unfilled grains per panicle (79.31%), sterility percentage (64.55%), thousand-grain weight (98.64%), straw yield (84.79%), grain yield (94.08%), and harvest index (90.37%). Moderate heritability was observed for panicle length (59.77%). All the characters studied exhibited high heritability in a broad sense except panicle length and it indicates that the environment had less influenced the landraces in expression of phenotype (Visscher et al. 2008).

High GAM was observed in penultimate leaf width (30.21%), ligule length (42.50%), flag leaf length (30.09%), flag leaf width (36.51%), total no of tillers per plant (31.31%) culm diameter (23.18%), productive tillers per plant (27.19%), plant height (29.60%), grain length (25.11%), grain width (55.87%), grain length-width ratio (49.46%), total grain per panicle (35.87%), unfilled grains/panicle (72.33%), sterility percentage (46.25%), thousand-grain weight (44.67%), straw yield (55.13%), grain yield (86.68%) and harvest index (69.32%). Moderate GAM was observed in characters; penultimate leaf length (16.72%), days to 50% flowering (16.30%), and days to maturity

(16.52%). Low GAM was observed in trait for panicle length (9.83%). High heritability and high genetic advance as a percent of the mean found in some characters indicate the lesser influence of environment and a preponderance of additive gene action for the expression (Immanuel et al. 2011). They can be improved by mass selection and other breeding methods based on progeny testing. If a character is governed by additive gene action, high heritability along with high genetic advance as a percent of the mean (above 20%) provide good scope for further improvement (Govindaraj et al. 2011). The value with high heritability with moderate genetic advance was also recorded in some of the traits like penultimate leaf length, days to flowering, and days to maturity suggesting control of expression by both additive and non-additive gene action in the development of earlier features, heterosis breeding can be useful (Kumar et al. 2013).

3.5 Correlation of yield and attributing characters

Character association derived from a correlation coefficient forms the basis for selecting the desirable plant, aiding in the evaluation of the relative influence of various component characters on grain yield (Akter et al. 2018). The phenotypic correlation coefficient between yield and various yield-attributing traits is presented in Figure 1. The positive and highly significant correlation between grain yield and penultimate leaf width ($r=0.43^{***}$), flag leaf width ($r=0.36^{***}$), effective tillers per plant ($r=0.29^{***}$), grain length ($r=0.47^{***}$), grain width ($r=0.40^{***}$), thousand-grain weight ($r=0.54^{***}$), straw yield ($r=0.31^{***}$), and harvest index ($r=0.77^{***}$) thereby indicating that it is possible to improve grain yield by genetic improvement of all these characteristics (Padmaja et al. 2011; Reddy et al. 2013). Positive and non-significant correlation of grain yield was observed with penultimate leaf length, flag leaf length, days to maturity, culm diameter, plant height, and grain per panicle. However, grain yield showed a negative and highly significant correlation with the number of unfilled grains per panicle ($r=-0.24^{***}$) and sterility percent ($r=-0.31^{***}$). Days to maturity showed a positive and highly significant correlation with days to flowering ($r=0.89^{***}$) and culm diameter ($r=0.22^{**}$), which showed a positive and significant correlation with effective tillers per plant ($r=0.18^{*}$). While days to flowering showed a negative and highly significant correlation with flag leaf length ($r=-0.34^{***}$). A highly significant positive correlation was seen between thousand-grain weight and penultimate leaf width ($r=0.44^{***}$), flag leaf width ($r=0.42^{***}$), grain length ($r=0.55^{***}$), and grain width ($r=0.54^{**}$). While it was highly significant and negatively correlated with grain length: width ratio ($r=-0.32^{***}$), grain per panicle ($r=-0.24^{**}$), unfilled grain per panicle ($r=-0.34^{***}$), and sterility percent ($r=-0.23^{**}$). Straw yield showed positive and highly significant correlation with penultimate leaf length ($r=0.29^{***}$), penultimate leaf width ($r=0.31^{***}$), flag leaf width ($r=0.22^{**}$), plant height ($r=0.42^{***}$), culm diameter ($r=0.28^{***}$), grain length ($r=0.21^{**}$), and grain width ($r=0.42^{***}$) and negative and highly significant correlation with grain length-width ratio ($r=-0.37^{***}$). A positive and highly significant correlation was seen between plant height and culm diameter ($r=0.37^{***}$), and penultimate leaf width ($r=0.23^{**}$). On the other side, plant height showed a weak positive correlation with panicle length, grains per panicle, straw yield, and days to maturity. Overall correlation results between yield and yield attributing traits indicated that selection directed towards favoring higher panicle number per plant, effective tillers per square meter, test weight and straw weight, and medium plant height with moderate grain number and optimum days to flowering would be effective to achieve higher grain yield (Kalyan et al. 2017).

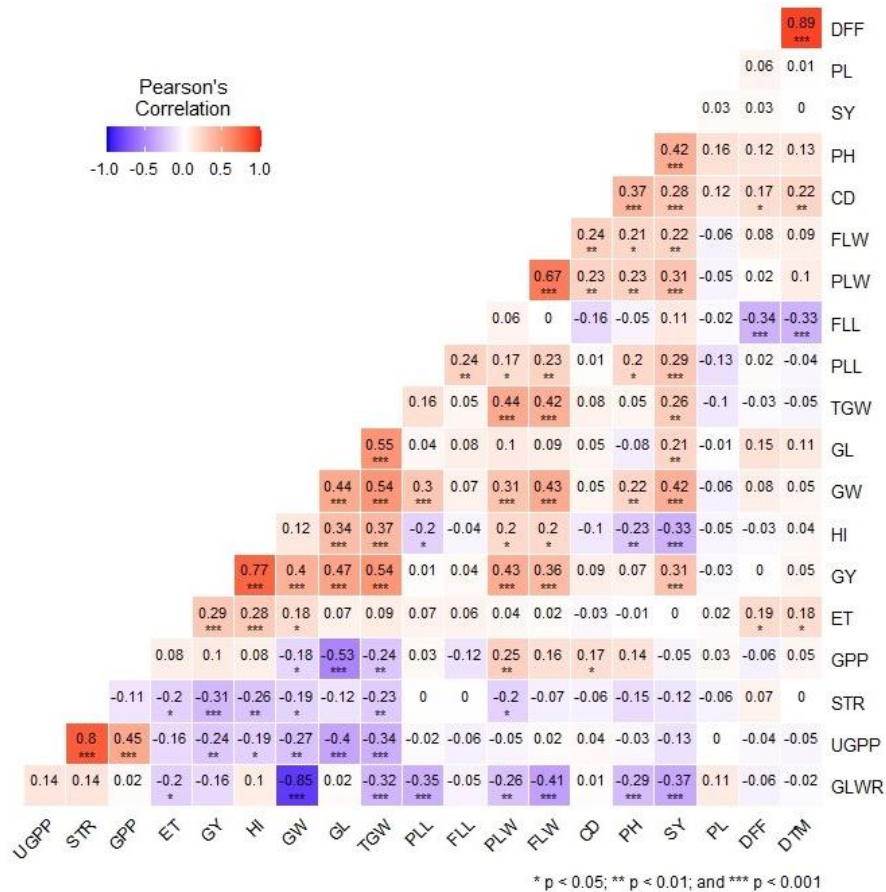


Fig. 1 Pearson correlation (r) of yield and attributing characters of rice landraces tested at Rampur, Chitwan, 2019. Negative correlations are in blue, positive correlations are in red. R-value is marked* = Significant at 5%, ** = Significant at 1% probability level and *** = Significant at 0.1% probability level, PLL= penultimate leaf length (cm), PLW= leaf width of blade (cm), FLL= flag leaf length (cm), FLW= flag leaf width (cm), CD = culm diameter (mm), DFF= days to fifty percent flowering, ET = number of effective tillers, DTM = days to maturity, PH = plant height (cm), PL= panicle length (cm), GL= grain length (cm), GW = grain width (cm), GPP= number of grains per panicle, UGPP= number of unfilled grains per panicle, SY= straw yield (kg. ha-1), TGW= thousand grain weight (gm), GY= grain yield (kg. ha-1)

3.6 Multivariate analysis

3.6.1 Cluster analysis

3.6.1.1 Clustering of genotypes

The forty-nine rice landraces were grouped into eight clusters representing the existence of considerable genetic diversity among them (Umarani 2014; Shoba et al. 2019). The percentage of representation and related traits among the twenty-two quantitative traits observed are listed in Table 5.

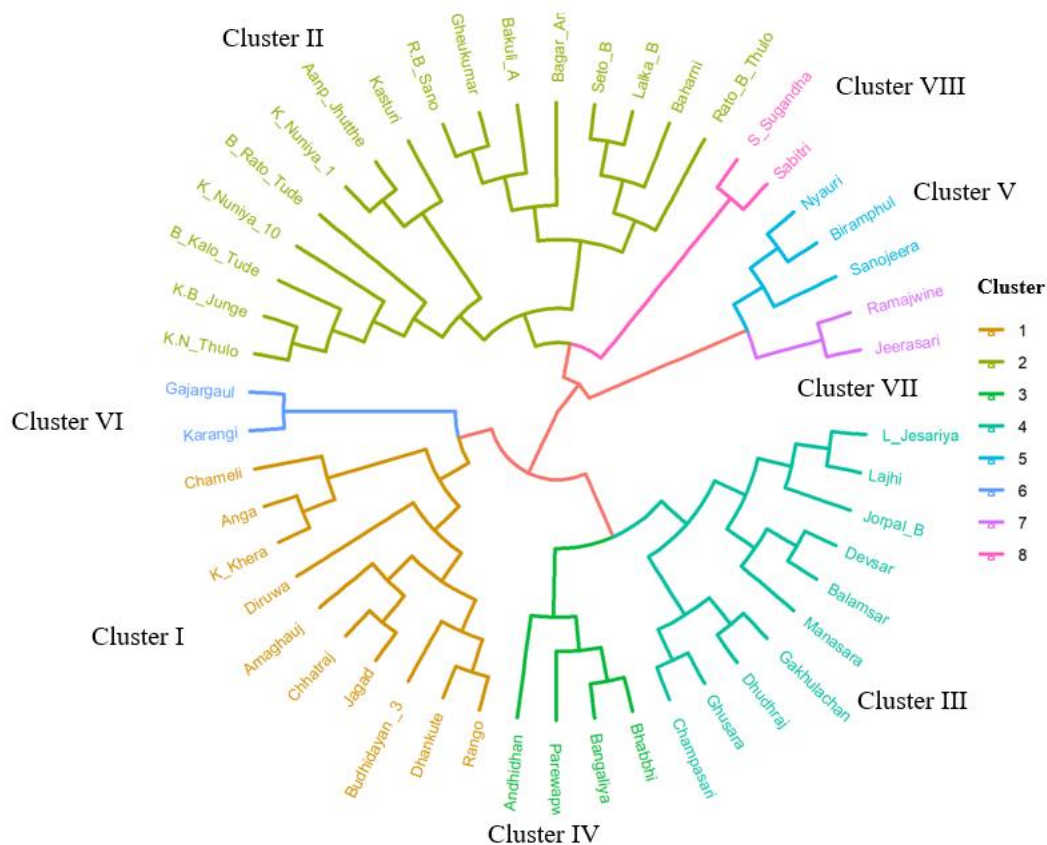


Fig. 2 Clustering of rice landraces grown at the research farm, AFU, Rampur, Chitwan, 2019

Cluster I included, ten genotypes, were grouped (Figure 2), which represent 20.42% of the total genotypes and have lower culm diameter. Cluster II was the largest cluster, which included sixteen genotypes representing 32.64% of the total genotypes with the lower number of grains per panicle. Cluster III which consisted of four landraces represented 8.16% of the total. It had been observed that penultimate leaf length, penultimate leaf width, flag leaf width, culm diameter, plant height, grain width, thousand-grain weight, and straw yield were found higher in this cluster. Likewise, in cluster IV, ten landraces were grouped, which represent 20.42% of the total genotypes. In cluster V, three landraces were grouped, which represent 6.12% of the total genotypes. Late days to flowering and maturity, the number of grains per panicle, and unfilled grains per panicle were found higher after cluster VII. Likewise, in clusters VI, VII, and VIII, two genotypes were grouped, which represent 4.08% of the total genotypes. These cluster genotypes have higher flag leaf length, the number of tillers per plant, and lower ligule length and panicle length with earlier days to flowering and maturity. In cluster VII genotypes characterized with higher ligule length, the number of grains per panicle, no of unfilled grains per panicle, and sterility percentage whereas lower penultimate leaf width, flag leaf width, no of effective tillers per plant, thousand-grain weight, grain yield, harvest index. Genotypes in cluster VIII were characterized by higher- panicle length, grain length, grain yield, harvest index, and lower- penultimate leaf length, flag leaf length, plant height, grain length-width ratio, straw yield, and sterility percentage.

Table 5 Clustering of rice genotypes used

Cluster No.	Genotypes	Distinct characters	Number of Genotypes	Percentage of Representation
I	Anga, Amaghauj, Budhidayan-3, Chameli, Chhatraj, Dhankute, Diruwa, Jagad, Kariya Khera, Rango	Lower- culm diameter	10	20.42

II	Aanp Jhutthe, Bagar Anadi, Baharni, Bakuli Anadi, Basmati Kalo Tude, Basmati Rato Tude, Ghiyukumar, Kalo Basmati Junge, Kalo Nuniya-1, Kalo Nuniya-10, Kalo NuniyaThulo, Kasturi, Lalka Basmati, Rato Basmati Sano, Rato Basmati Thulo, Seto Basmati	Lower- number of grains per panicle	16	32.64
III	Aandhi Dhan, Bangaliya, Bhabbhi, Parewapwankh	Higher- penultimate leaf length, penultimate leaf width, flag leaf width, culm diameter, plant height, grain width, thousand-grain weight, straw yield	4	8.16
IV	Balamsar, Champasari, Devsaar, Dudhraj, Gakhulachan, Ghusara, Jhorpal Basmati, Lajhi, Lalka Jesariya, Mansara	Higher- Straw yield after cluster III	10	20.42
V	Biramphul, Nyauri, Sanojeera	Higher-No of grains per panicle, unfilled grains per panicle	3	6.12
VI	Gajargaul, Karangi	Late days to flowering and maturity Higher-Flag leaf length, number of tiller per plant Lower-Ligule length, panicle length Earlier days to flowering and maturity, no of unfilled grains per panicle	2	4.08
VII	Jeerasari, Ramajwine,	Higher-Ligule length, number of grains per panicle, no of unfilled grains per panicle, sterility percentage, Lower- penultimate leaf width, flag leaf width, no of effective tillers per plant, thousand-grain weight, grain yield, harvest index	2	4.08
VIII	Sabitri, Sunaulo Sugandha	Higher- panicle length, grain length, grain yield, harvest index Lower- penultimate leaf length, flag leaf length, plant height, grain length-width ratio, straw yield, sterility %	2	4.08

Targeted traits of different geotypes from different clusters should be carefully combined through hybridization to develop a desirable variety (Baloch et al. 2016; Bose and Pradhan 2005). No cluster showed the landraces collected from a single location indicating no role of ecological distribution on genetic divergence. The possible reason for this could be a free exchange of landraces among the farmers, unidirectional selection of the landraces by the farmers, and also the sites of collection of the landraces used in this research were very near and of similar geography (Bhati et al. 2015).

3.6.1.2 Estimation of intra and inter-cluster square distances (D^2)

The estimation of intra and inter-cluster square distances (D^2) distance between clusters was determined using the Euclidean distance and is presented in Table 6. The highest inter-cluster distance was exhibited by clusters IV and VIII ($D^2 = 5165.73$) followed by clusters III and VIII ($D^2 = 5156.91$) and clusters VI and VIII ($D^2 = 4527.63$), which implied these clusters were genetically more divergent from each other than any other pairs of the cluster. The smallest inter-cluster distance was observed between Cluster III and IV ($D^2 = 1569.81$) succeeded by Cluster I and III ($D^2 = 1988.19$). The genotypes belonging to these clusters were relatively close to each other, in comparison to genotypes grouped in other clusters. The intra-cluster distance (D^2) was maximum in cluster VI ($D^2=3634.47$) followed by cluster V ($D^2=2932.76$). The cluster with the maximum number of landraces (cluster II) showed an intracluster distance ($D^2=2477.88$). The minimum intracluster distance was obtained in cluster VIII ($D^2=205.42$).

Table 6 Intracluster (**diagonal**) and intercluster distance among 8 clusters of 49 rice genotypes

	I	II	III	IV	V	VI	VII	VIII
I	2344.13							
II	3295.61	2477.88						
III	1988.19	3759.13	1440.65					
IV	2258.09	3427.54	1569.81	1687.33				
V	3650.28	2313.32	4255.13	3880.05	2932.76			
VI	2680.75	2820.93	2445.16	2184.82	3018.00	3634.47		
VII	3023.60	2073.41	3208.11	2716.18	2067.52	2023.96	1754.95	
VIII	3854.53	3300.77	5156.91	5165.73	3317.05	4527.63	3975.25	205.42

The presence of wide genetic diversity among the genotypes present within cluster VI is indicated by the highest intra-cluster distance, whereas the maximum inter-cluster distance was observed between clusters IV and VIII, indicating that there is wide genetic diversity between these clusters and crosses can be made between the genotypes of these clusters to obtain desirable transgressive segregants (Mukul et al. 2019). It is suggested to cross genotypes from clusters IV and VIII to get genotypes/varieties with high grain yield and early maturing genotypes. In hybridization programs, parental lines must be selected from the clusters that are at the highest genetic divergence because divergent parents are expected to produce wide variability and transgressive segregations with high heterotic effects (Qamar et al. 2012; Chandra et al. 2007). The landraces belonging to diverse clusters and showing high mean performance in desirable directions for different traits can be used as parents in the breeding program (Chakma et al. 2012). Genetic gains in yield or any characters can be exploited by recombining the elite germplasm followed by selection (Paudel et al. 2017). Thus, landraces with high performance from clusters IV and VIII can be used in the hybridization program as a parent line due to greater intercluster distance. Crossing genotypes belonging to the same cluster do not yield superior hybrids or segregates they might not give a higher heterotic value in F1 and a narrow range of variability in segregating the F2 population. Such analysis was meant to avoid the selection of parents from genetically homogeneous clusters to maintain a relatively broad genetic base.

The mean value of genotypes in each cluster was computed and cluster means are presented in Table 7. With the highest and the lowest cluster means values (bold) for respective traits. Such as the highest cluster mean for grain yield was observed in cluster VIII (3731.98) and lowest in cluster VII (1021.86). From the cluster analysis, it was clear that the landraces used in the study have highly diverged in the mean values of agro-morphological characters across the clusters. The promising landraces from these clusters with high mean values for different characters may be used as parents in hybridization, depending upon the objective of breeding programs to improve certain traits (Singh et al. 2021).

Table 7. The centroid of clustering of 22 variables measured in 49 rice genotypes

Cluster	I	II	III	IV	V	VI	VII	VIII	Grand Centroid
Variables									
PLL	52.74	48.94	52.96	50.19	50.62	51.11	46.37	41.00	50.06
PLW	1.13	0.94	1.34	1.12	1.14	1.15	0.86	1.14	1.07
LLL	1.60	1.64	2.04	2.16	1.63	1.29	2.31	1.56	1.78
FLL	32.53	31.16	30.62	33.15	29.79	43.26	29.84	29.13	32.07
FLW	1.46	1.19	1.64	1.33	1.40	1.28	0.98	1.19	1.32
NTI	18.93	15.97	14.29	18.18	15.48	21.81	12.98	15.16	16.94

CD	5.61	6.01	7.26	6.55	6.52	5.73	6.28	6.36	6.18
DFF	118.37	120.76	120.85	120.47	121.68	87.69	101.11	112.18	117.78
ET	13.15	10.86	10.58	12.23	11.43	9.60	9.40	10.33	11.49
DTM	147.78	151.23	153.74	149.70	155.64	111.87	126.37	148.07	147.94
PH	132.56	137.26	178.22	159.50	165.84	134.30	140.27	111.00	144.86
PL	24.62	24.48	25.52	26.84	25.73	24.15	27.05	27.35	25.36
GL	7.64	7.58	8.18	7.93	4.83	7.05	5.64	8.45	7.48
GW	2.41	1.77	2.91	2.11	1.17	2.10	1.10	1.25	1.99
GLWR	3.30	4.40	2.83	3.93	4.18	3.43	5.13	6.78	4.03
GPP	151.06	124.44	147.31	130.71	195.35	113.64	210.63	155.90	141.72
UGPP	20.70	20.71	21.09	18.58	27.27	12.59	54.87	14.99	21.53
STR	13.85	17.46	14.37	14.35	13.83	12.12	25.41	8.97	15.38
TGW	24.55	20.31	27.82	22.29	14.37	24.54	10.35	22.77	21.69
SY	8356.44	6547.36	9721.46	9541.47	5966.63	8507.40	7475.78	4628.29	7790.72
GY	3239.77	1559.07	3168.45	2395.40	1267.47	1810.41	1021.86	3731.98	2263.30
HI	0.29	0.20	0.24	0.20	0.18	0.18	0.12	0.45	0.22

PLL= penultimate leaf length (cm), PLW= penultimate leaf width (cm), LL= ligule length (cm), FLL= flag leaf length (cm), FLW= flag leaf width (cm), NTI = number of tillers per plant, CD = culm diameter (mm), DFF= days of fifty percent flowering, ET = number of effective tillers, DTM = days to maturity, PH = plant height (cm), PL = panicle length (cm), GL = grain length (cm), GW = grain width (cm), GLWR = grain length-width ratio, GPP = number of grains per panicle, UGPP= number of unfilled grains per panicle, STR % = sterility percentage, TGW= thousand grain weight (gm), SY= straw yield (kg. ha-1), GY= grain yield (kg. ha-1), HI = Harvest index

3.6.2.2 Principal component analysis (PCA)

PCA, using twenty-two quantitative traits revealed that, 77.68% variability was accounted for by the first seven principal components (PC) with ≥ 1 Eigen value, ranging from 5.032 to 1.106 (Table 8). However, the remaining component contributed 22.32% towards total diversity. Only these seven components which had an eigenvalue of more than 1 which explained 72.1% of the total cumulative variance within the axes further strengthening the current result were considered for further analysis (Tuhina-Khatun et al. 2015). The first seven components are mostly responsible for genetic variability and are predicted to offer a higher percentage of gene transfer during breeding programs (Gana et al. 2013). Scree plots shown in Figures 3a and 3b show the eigenvalue and contribution of each principal component to the total explained variance in the phenotypic diversity of rice, respectively. PC1 showed 22.87% variability with an eigenvalue of 5.032 which then declined gradually as shown in Figures 3a and 3b. The biplot of rice genotypes and the magnitude and direction of the contribution of different characters for PC1 and PC2 is shown in Figure 4. Genotypes grouped in such a layout indicate groups of genetically related individuals (Warburton and Crossa 2002).

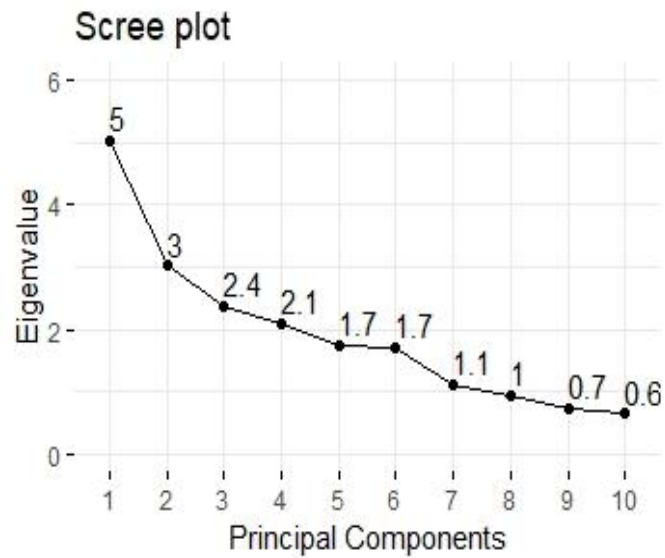


Fig. 3a Scree plot of principal component analysis of rice landraces between Eigen value and principal components

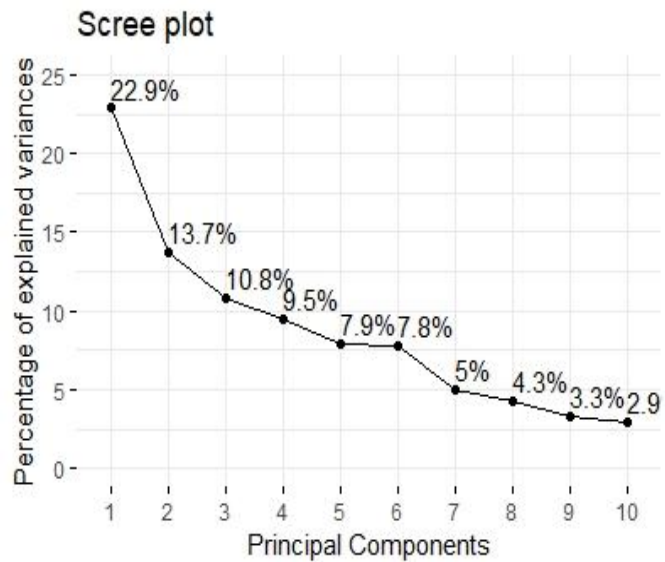


Fig. 3b Scree plot of the contribution of each principal component to the total explained variance in the phenotypic diversity of rice genotypes

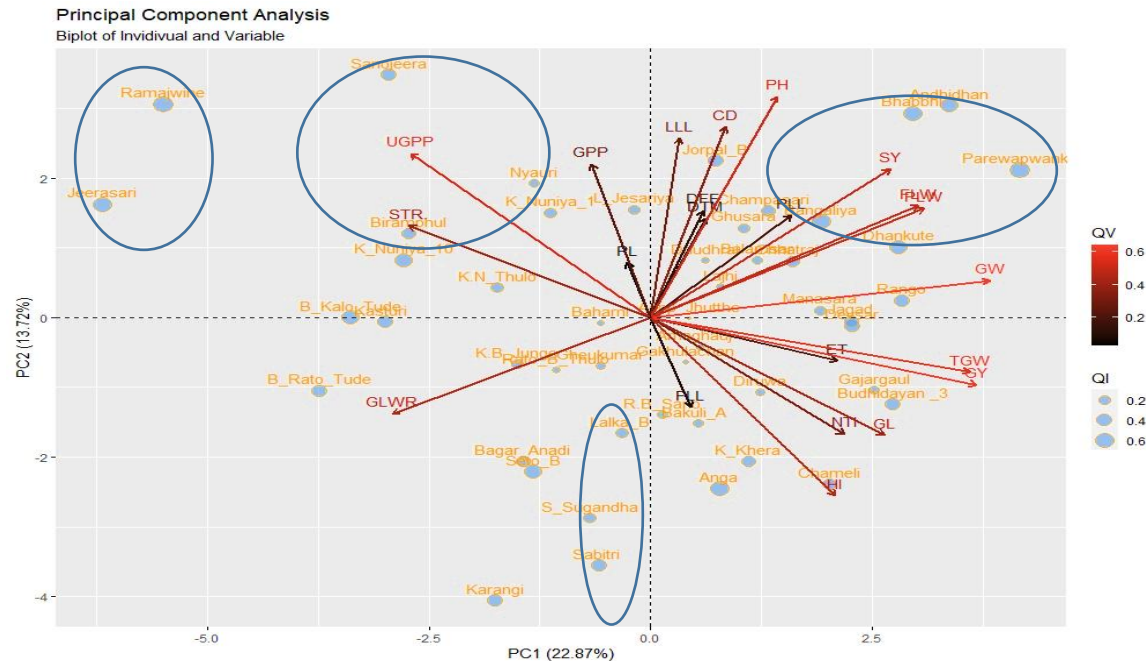


Fig. 4 Biplot of first and second principal components showing the relationship between rice genotypes
 Note: QV = Quality of representation of variables by color and size, QI Quality of representation of individuals by color and size (PLL= penultimate leaf length (cm), PLW= leaf width of blade (cm), LL= length of ligule (cm), FLL= flag leaf length (cm), FLW= flag leaf width (cm), NTI = number of tillers per plant, CD = culm diameter (mm), DFF= days of fifty percent flowering, ET = number of effective tillers, DTM = days to maturity, PH = plant height (cm), PL = panicle length (cm), GL = grain length (cm), GW = grain width (cm), GLWR = grain length width ratio, GPP = number of grains per panicle, UGPP= number of unfilled grains per panicle, STR % = sterility percentage, TGW= thousand grain weight (gm), SY= straw yield (kg. ha-1), GY= grain yield (kg. ha-1), HI = Harvest index).

The first principal component (PC1) with Eigen value 5.023 explained 22.87% of the total variation. Results from the PCA and vector loading in Table 8 showed that PC1 was contributed mainly positively by character such as number of penultimate leaf length (0.329), penultimate leaf width (0.637), flag leaf width (0.625), number of tillers per plant (0.451), number of effective tillers per plant (0.436), grain length (0.546), grain width (0.793), thousand-grain weight (0.745), straw yield (0.559) and grain yield (0.759) and negatively by grain length-width ratio (-0.602), unfilled grain per panicle (-0.559) and sterility percentage (-0.564). That means the character with positive contributions had a high correlation with PC1 which were the ones that most differentiated the clusters. If one increases, then the remaining ones tend to as well. So, the selection of any traits among these traits indirectly also selects the remaining traits. This clearly showed that the PC1 included parameters that were mainly related to grain yield, yield attributes, and grain characteristics (Gour et al. 2017). Landraces Dudhraj, Bhabbhi, Laghi, Parewapwankh, Aandhi Dhan, and Gajargaul showed good performance in these characters.

The second principal component (PC2) had Eigen value 3.019 accounted for 13.72% of the total variability in which ligule length (0.531), culm diameter (0.564), plant height (0.654), grain per panicle (0.454), unfilled grain per panicle (0.485), sterility percentage (0.271), contributed positively while the number of tillers per plant (-0.345), grain yield (-0.220) and harvest index (-0.527) contributed inversely. Landraces Sanojeera, Ramajwne, Nyauri, and Kasturi showed good performance in these characters. The third principle component (PC3), had an eigenvalue of 2.386 and contributed 10.84% to the observed variability was highly loaded with characters like days to flowering (0.645), days to maturity played (0.710) and grain length-width ratio (0.403) played a significant positive role while penultimate leaf length (-0.516) flag leaf length (-0.592), and grain width (-0.244) played a negative significant role. Landraces Karangi and Gajargaul showed good performance in these characters. PC4 with eigenvalue 2.088, accounted for 9.49% of the total variation which was mainly due to the positive association with grain per panicle (0.708), unfilled grain per panicle (0.389), and harvest index (0.467) while the negative association with flag leaf length (-0.211), days to fifty percentage flowering (-0.448), and grain length (-0.378). PC5 had an eigenvalue of 1.747 representing 7.94% of the total variation was a mainly negative association with culm diameter (-0.158). PC6 with eigenvalue 1.711,

accounted for 7.77% of the total variation which was mainly due to positive association with panicle length (0.518) while the negative association with flag leaf width (-0.256). Sterility percentage (0.525) contributed to the PC7 which accounts for 5.02% of total variation with an eigenvalue of 1.106. Principal discriminatory characters are penultimate leaf width, flag leaf width, grain length, grain width, thousand-grain weight, straw yield, and grain yield. The prominent characters coming together in different PCs and contributing towards explaining the variability tend to remain together which may be kept into consideration during the utilization of these characters in the breeding program (Islam et al. 2019).

Table 8 Contribution of first seven extracted PC in total variability and their coefficients of correlation with evaluated 22 traits measured in 49 rice genotypes

Traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7
PLL	0.329	0.303	-0.516	-0.114	0.033	-0.305	-0.055
PLW	0.637	0.323	0.041	0.475	-0.196	-0.012	-0.107
LL	0.067	0.531	0.104	-0.211	-0.075	0.482	0.308
FLL	0.095	-0.266	-0.592	0.033	0.093	0.020	0.228
FLW	0.625	0.331	-0.032	0.430	-0.173	-0.256	-0.016
NTI	0.451	-0.345	-0.091	-0.111	0.650	0.021	0.094
CD	0.176	0.564	0.265	0.077	-0.158	0.288	-0.108
DFF	0.124	0.316	0.645	-0.448	0.037	-0.434	0.004
ET	0.436	-0.128	0.213	-0.008	0.662	-0.205	0.356
DTM	0.129	0.293	0.710	-0.305	0.074	-0.429	-0.099
PH	0.295	0.654	-0.071	-0.148	0.147	0.330	-0.172
PL	-0.055	0.163	0.239	-0.190	0.251	0.518	0.306
GL	0.546	-0.348	0.175	-0.378	-0.469	0.132	0.270
GW	0.793	0.110	-0.244	-0.204	-0.100	-0.194	0.109
GLWR	-0.602	-0.285	0.403	0.133	-0.189	0.346	0.031
GPP	-0.138	0.454	0.145	0.708	0.370	0.005	-0.018
UGPP	-0.559	0.485	-0.070	0.389	-0.061	-0.207	0.475
STR %	-0.564	0.271	-0.144	-0.081	-0.325	-0.374	0.525
TGW	0.745	-0.162	-0.032	0.044	-0.400	0.014	0.033
SY	0.559	0.439	-0.277	-0.233	0.078	0.256	0.079
GY	0.759	-0.200	0.269	0.341	-0.052	0.146	0.221
HI	0.430	-0.527	0.449	0.467	-0.071	0.006	0.144
Eigen value	5.032	3.019	2.386	2.088	1.747	1.711	1.106
Variabilty (%)	22.872	13.723	10.847	9.492	7.941	7.777	5.029
Cumulative %	22.872	36.595	47.442	56.934	64.875	72.653	77.681

PLL= penultimate leaf length (cm), PLW= penultimate leaf width (cm), LL= ligule length (cm), FLL= flag leaf length (cm), FLW= flag leaf width (cm), NTI = number of tillers per plant, CD = culm diameter (mm), DFF= days of fifty percent flowering, ET = number of effective tillers, DTM = days to maturity, PH = plant height (cm), PL = panicle length (cm), GL = grain length (cm), GW = grain width (cm), GLWR = grain length-width ratio, GPP = number of grains per panicle, UGPP= number of unfilled grains per panicle, STR % = sterility percentage, TGW= thousand grain weight (gm), SY= straw yield (kg. ha-1), GY= grain yield (kg. ha-1), HI = Harvest index

4. Conclusion

In our study, polymorphism was found in all qualitative characters except for the shape of the ligule. The morphological characters like lemma and palea coloration, culm lodging tolerance, awn color, auricle color decorticated grain color, and shape of grain are more diverse while pubescence and distribution of anthocyanin in leaf blade, and anthocyanin of the leaf sheath, were found to be least diverse. Analysis of variance revealed the presence of considerable genetic variation in the landraces which can also be utilized directly by selection or as donor parents in a future breeding program. A narrow difference between the phenotypic coefficient of variation and with genotypic coefficients of variation indicates the greater contribution of genotype to trait expression leading to high heritability. There is considerable genetic variability among landraces regarding ligule length, grain width, number of grains per panicle, thousand-grain weight, straw yield, and grain yield so there is a possibility of genetic improvement through direct selection for these characters. Most of the characters show high heritability coupled with high genetic advance as a percent of the mean, we can say that there is a preponderance of additive gene action for these characters which could be improved by selection. Leaf width, number of effective tillers per plant, grain length, grain width, and

thousand-grain weight can be selection indices for the improvement of grain yield. Distribution of landraces into agronomically similar clusters and the pattern of the traits' contribution to the total variation would quicken the usefulness of estimates to rice breeders. Marked heterogeneity for qualitative and quantitative characters among the studied rice landraces as revealed in the present study could be utilized in rice improvement programs for several characters. Parental material selection should be done between clusters rather than within clusters for future breeding.

Funding

No funding was received for this study.

Data availability

The necessary data are available in supplementary files and other data that support findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

References

- Acevedo-Siaca LG, Coe R, Quick WP, Long SP (2021) Variation between rice accessions in photosynthetic induction in flag leaves and underlying mechanisms. *Journal of Experimental Botany* 72(4):1282-1294. <https://doi.org/10.1093/jxb/eraa520>
- Adhikari BN, Pokhrel BB, & Shrestha J (2018) Evaluation and development of finger millet (*Eleusine coracana* L.) genotypes for cultivation in high hills of Nepal. *Farming and Management* 3(1):37-46. <http://dx.doi.org/10.31830/2456-8724.2018.0001.7>
- Ahmed MSU, Bashir MK, Wazuddin M, Shamsuddin AKM (2016) Agro-morphological qualitative characterization of Jesso-Balam rice (*Oryza sativa* L.) accessions in Bangladesh. *International Journal of Agronomy and Agricultural Research* 8(1):50-58.
- Akter N, Khalequzzaman M, Islam MZ, Mamun MAA, Chowdhary MAZ (2018) Genetic variability and character association of quantitative characters in Jhum rice genotypes. *SAARC Journal of Agriculture* 16(1):193-203. <http://dx.doi.org/10.3329/sja.v16i1.37434>
- Baloch A, Baloch M, Baloch R, Kandhro M, Odhano I, Bughio H Rajput A (2016) Evaluation of genetic diversity in rice (*Oryza sativa* L.) genotypes. *Sindh University Research Journal-SURJ (Science Series)* 48(1).
- Banyai J, Kiss T, Gizaw SA, Mayer M, Spitkó T, Tóth V, Kuti C, Mészáros K, Láng L, Karsai I, Vida G (2020) Identification of superior spring durum wheat genotypes under irrigated and rain-fed conditions. *Cereal Research Communications* 48:355–364. <https://doi.org/10.1007/s42976-020-00034-z>
- Bernardo R (2002) Breeding for quantitative characters in plants. Woodbury, MN, Stemma press 1:369.
- Bhati K, Singh SK, Dhurai SK, Sharma A (2015) Genetic divergence for quantitative characters in rice germplasm. *Electronic Journal of Plant Breeding* 6(2):528–534.
- Biodiversity International, IRRI, & WARDA, Descriptors for wild and cultivated rice (*Oryza* spp.) Biodiversity International, Rome, Italy, 2007. https://www.biodiversityinternational.org/fileadmin/_migrated/uploads/tx_news/Rice_232_01.pdf.
- Bitew JM (2016) Estimation of genetic parameters, heritability, and genetic advance for yield-related characters in upland rice (*Oryza sativa* L. and *Oryza glaberrima* Steud) genotypes in northwestern Ethiopia. *World Scientific News* 47(2):340-350.

- Bose KL, Pradhan SK (2005) Genetic divergence in deep-water rice genotypes. *Journal of Central European Agriculture* 6(4):635–640. <https://doi.org/10.5513/jcea.v6i4.347>
- CDD (2017) Rice varietal mapping in Nepal: Implication for development and adoption. Crop Development Directorate, Department of Agriculture, Hariharbhawan, Lalitpur.
- Chakma SP, Huq H, Mahmud F, Husna A (2012) Genetic diversity analysis in rice (*Oryza sativa* L.). *Bangladesh Journal of Plant Breeding and Genetics* 25(1):31–39. <https://doi.org/10.3329/BJPBG.V25I1.17010>
- Chandra R, Pradhan SK, Singh S, Bose LK, Singh ON (2007) Multivariate analysis in upland rice genotypes. *World Journal of Agricultural Sciences* 3(3):95-300.
- Chaves-Silva S, Dos Santos AL, Chalfun-Júnior A, Zhao J, Peres LE, Benedito VA (2018) Understanding the genetic regulation of anthocyanin biosynthesis in plants—tools for breeding purple varieties of fruits and vegetables. *Phytochemistry* 153:11-27. <https://doi.org/10.1016/j.phytochem.2018.05.013>
- Chin HS, Wu YP, Hour AL, Hong CY, Lin YR (2016) Genetic and evolutionary analysis of purple leaf sheath in rice. *Rice* 9(1):1-14. <https://doi.org/10.1186/s12284-016-0080-y>
- Fakhar Izadi N, Keshtkar HR (2020) Investigating the effects of distribution patterns on ecological indices of plant species in a simulated environment. *Desert* 25(2):201-211. <https://doi.org/10.22059/jdesert.2020.79257>
- Falconer DS (1960) Introduction to quantitative genetics. The Ronald Press Co., New York.
- FAOSTAT. (2021) Food and Agriculture Organization of the United Nations. Information sheet on rice. <https://www.fao.org/faostat/en/#data/QCL>
- Fornasiero A, Wing RA, Ronald P (2022) Rice domestication. *Current Biology* 32(1):20-24. <https://doi.org/10.1016/j.cub.2021.11.025>
- Gadal, N, Shrestha J, Poudel MN, Pokharel B (2019) A review on production status and growing environments of rice in Nepal and in the world. *Archives of Agriculture and Environmental Science* 4(1):83-87 <https://doi.org/10.26832/24566.632.2019.0401013>
- Gana AS, Shaba SZ, Tsado EK (2013) Principal component analysis of morphological traits in thirty-nine accessions of rice (*Oryza sativa* L.) grown in a rainfed lowland ecology of Nigeria. *Journal of Plant Breeding and Crop Science* 4(6):120-126. <https://doi.org/10.5897/JPBCS2013.0381>
- Gauchan D, Pandey S (2011) Is Investment in Rice Research in Nepal Adequate and Balanced Across Production Environments? Some Empirical Evidence. *Journal of International Agriculture* 50(4):305-324. <http://dx.doi.org/10.22004/ag.econ.155536>
- Gomez KA, Gomez AA (1984) Statistical procedures for agricultural research. John Wiley & Sons
- Gould KS, Jay-Allemand C, Logan BA, Baissac Y, Bidet LP (2018) When are foliar anthocyanins useful to plants? Re-evaluation of the photoprotection hypothesis using *Arabidopsis thaliana* mutants that differ in anthocyanin accumulation. *Environmental and Experimental Botany* 154:11-22. <https://doi.org/10.1016/j.envexpbot.2018.02.006>
- Gour L, Koutu, GK, Singh Y, Singh SK (2019). Characterization and grouping of indigenous rice lines based on the DUS test for different qualitative differentials. *Pharma Innovative Journal* 8(4):1162-1168.
- Gour L, Maurya SB, Koutu GK, Singh SK, Shukla SS (2017) Characterization of rice (*Oryza sativa* L.) genotypes using principal component analysis including scree plot & rotated component matrix. *International Journal of Chemical Studies* 5(4):975-983.
- Govindaraj M, Selvi B, Rajarathinam S, Sumathi P (2011) Genetic variability and heritability of grain yield components and grain mineral concentration in India's pearl millet [*Pennisetum glaucum* (L) R. Br.]

- accessions. African Journal of Food, Agriculture, Nutrition and Development 11(3). <https://doi.org/10.4314/ajfand.v11i3.66627>
- Hanson C, Robinson H, Comstock R (1956) Biometrical studies of yield in segregating populations of Korean Lespedeza. Agronomy Journal 48 (6): 268-272.
- Hidayatun N, Sabran M, Ramadhani A, & Widaningsih, NA (2020) Development of physical color reference for Indonesia paddy collection. In IOP Conference Series: Earth and Environmental Science 482(1):12-46. IOP Publishing. <https://doi.org/10.1088/1755-1315/482/1/012046>
- Hien NL, Sarhadi WA, Oikawa Y, Hirata Y (2007) Genetic diversity of morphological responses and the relationships among Asia aromatic rice (*Oryza sativa* L.) cultivars. Tropics 16(4):343-355. <https://doi.org/10.3759/tropics.16.343>
- Hou Z, Qin P, Ren G (2010) Effect of anthocyanin-rich extract from black rice (*Oryza sativa* L. *japonica*) on chronically alcohol-induced liver damage in rats. Journal of agricultural and food chemistry 58(5): 3191-3196. <https://doi.org/10.1021/jf904407x>
- Hu B, Wan Y, Li X, Zhang F, Yan W, Xie J (2013) Phenotypic characterization and genetic analysis of rice with pubescent leaves and glabrous hulls. Crop Science 53. <https://doi.org/10.2135/cropsci2012.09.0522>.
- Huang R, Jiang L, Zheng J, Wang T, Wang H, Huang Y, Hong Z (2013) Genetic bases of rice grain shape: so many genes, so little known. Trends in plant science 18(4):218-226. <https://doi.org/10.1016/j.tplants.2012.11.001>
- Immanuel SC, Pothiraj N, Thiagarajan K, Bharathi M, Rabindran R (2011) Genetic parameters of variability, correlation and path coefficient studies for grain yield and other yield attributes among rice blast disease resistant genotypes of rice (*Oryza sativa* L.). African Journal of Biotechnology 10(17):3322-3334. <https://doi.org/10.5897/AJB10.2575>
- Ishwarya Lakshmi, VG, Sreedhar M, Vanisri S, Anantha MS, Subba Rao LV, Gireesh C (2019) Multivariate analysis and selection criteria for identification of African rice (*Oryza glaberrima*) for genetic improvement of *indica* rice cultivars. Plant Genetic Resources: Characterization and Utilization 1-7. <https://doi.org/10.1017/S1479262119000327>
- Islam M, Khalequzzaman M, Bashar M, Ivy N, Haque M, Mian M, Tomita M (2019) Agro-morphological characterization of Bangladeshi aromatic rice (*Oryza sativa* L.) germplasm based on qualitative traits. Bangladesh Rice Journal 22(2):41-54. <https://doi.org/10.3329/brj.v22i2.44041>
- Jena KK (2010) The species of the genus *Oryza* and transfer of useful genes from wild species into cultivated rice, *O. sativa*. Breed. Sci. 60:518-523. <https://doi.org/10.1270/jsbbs.60.518>
- Johnson HW, Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans 1. Agronomy Journal, 47(7), 314-318.
- Joshi BK (2017) Conservation and utilization of agro-biodiversity advanced from 1937 to 2017 in Nepal. Krishi Sanchar Smarika (F Devkota, ed). Agricultural Information and Communication Center (AICC), MoAD 181-208.
- Joshi BK (2021) Agrobiodiversity indicators and measurement using R for description, monitoring, comparison, relatedness, conservation, and utilization. Agrobiodiversity & Agroecology 1(1): 47-64. <https://doi.org/10.33002/aa010103>
- Joshi BK, Ghimire KH, Bista PR, Yadaw RB, Shrestha RK, Kharel GK, Paneru P, KC RB, Bhandari D (2021) Intellectual property right on basmati rice: current scenario and evidence of origin, diversity, cultivation and use values of basmati rice in Nepal. Nepal Journal of Biotechnology 9(1):93-108. <https://doi.org/10.3126/njb.v9i1.38671>

- Joshi BK, KC H, Uadhyaya MP, Gupta SR, Lu BR, Mathur PN, Sthapit B R (2008) Ex-situ and in-situ management of wild and weedy rice in Nepal using a geographical information system. *Plant Genetic Resources Newsletter* 155:69-74.
- Joshi KD, Ali J (2021) Role of IRRI in developing hybrid rice technology development and promotion in Nepal. In *Proceedings of Stakeholders Workshop on Hybrid Variety Development in Nepal*. 52.
- Kalyan B, Radha Krishna KV, Subba Rao LV (2017) Correlation coefficient analysis for yield and its components in rice (*Oryza sativa* L.) genotypes. *International Journal Current Microbiology App. Sci.* 6(7):2425-2430. <https://doi.org/10.20546/ijcmas.2017.604.287>
- Khush GS (2000) New plant types of rice for increasing the genetic yield potential. *Rice breeding and genetics* 99-108.
- Kinoshita N, Kato M, Koyasaki K, Kawashima T, Nishimura T, Hirayama Y, Takamure I, Sato T, Kato K (2017) Identification of quantitative trait loci for rice grain quality and yield-related traits in two closely related *Oryza sativa* L. *japonica* cultivars grown near the northernmost limit for rice paddy cultivation. *Breeding Science* 67(3):191-206. <https://doi.org/10.1270/jsbbs.16155>
- Kovach M, McCouch S (2008) Leveraging natural diversity: back through the bottleneck. *Curr. Opin. Pl. Biol.* 11: 93-200. <https://doi.org/10.1016/j.pbi.2007.12.006>
- Kumar A, Rangare NR, Vidyakar V (2013) Study of genetic variability of Indian and exotic rice germplasm in Allahabad agroclimate. *The Bioscan* 8(4):1445-1451.
- Lee J, Park JJ, Kim SL, Yim J, An G (2007) Mutations in the rice ligule-less gene result in a complete loss of the auricle, ligule, and laminar joint. *Plant molecular biology*. 65(4):487-499. <https://doi.org/10.1007/s11103-007-9196-1>
- Maeda H, Yamaguchi T, Omoteno M, Takarada T, Fujita K, Murata K, Iyama Y, Kojima Y, Morikawa M, Ozaki H, Mukaino N, Kidani Y, Ebitani T (2014) Genetic dissection of black grain rice by the development of a near-isogenic line. *Breeding Science* 64(2):134–141. <https://doi.org/10.1270/jsbbs.64.134>
- Magurran AE (2004) *Measuring Biological Diversity*. Hoboken, NJ, United States. Wiley.
- Mahendran R, Veerabhadhiran P, Robin S, Raveendran M (2015) Principal component analysis of rice germplasm accessions under high-temperature stress. *International Journal of Agricultural Science and Research* 5(3):355-360.
- Maji AT, Shaibu AA (2012) Application of principal component analysis for rice germplasm characterization and evaluation. *Journal of Plant Breeding and Crop Science* 4(6): 87-93. <https://doi.org/10.5897/jpbcs11.093>
- Malla G (2008) Climate change and its impact on Nepalese agriculture. *Journal of Agriculture Environment* 9: 62-71.
- MoALD (2025) *Production Forecast of Rice*, Press Release, Government of Nepal, Ministry of Agriculture & Livestock Development, Singhadurbar, Kathmandu, Nepal.
- Mukul, S, Singh, A, Singh PK, Singh SP (2019) Estimation of genetic divergence analysis for yield and bacterial leaf blight (BLB) disease tolerance in rice (*Oryza sativa* L.). *Journal of Pharmacognosy and Phytochemistry*, 8(2):1124-1128.
- Nomura T, Seki Y, Matsuoka M, Yano K, Chigira K, Adachi S, ... Ookawa T (2021) Potential of rice landraces with strong culms as genetic resources for improving lodging resistance against super typhoons. *Scientific Reports* 11(1): 1-13. <https://doi.org/10.1038/s41598-021-95268-0>

- Padmaja D, Radhika K, Rao LV, Padma V (2011) Correlation and path analysis in rice germplasm. *Oryza-An International Journal on Rice* 48(1): 69-72.
- Parikh M, Motiramani NK, Rastogi NK, Sharma B (2012) Agro-morphological characterization and assessment of variability in aromatic rice germplasm. *Bangladesh Journal of Agricultural Research* 37(1):1-8. <https://doi.org/10.3329/bjar.v37i1.11168>
- Pielou EC (1966) The measurement of diversity in different types of biological collections. *Journal of theoretical biology* 13:31-144.
- Poudel S, Poudel A, Poudel S, Neupane S, Bhandari N (2020) Diversity assessment of rice (*Oryza sativa* L.) landraces adopted to Terai, Nepal. *GSJ* 8(2).
- Qamar ZU, Akhtar J, Ashraf M, Akram M, Hameed A (2012) A multivariate analysis of rice genetic resources. *Pakistan Journal of Botany* 44(4):1335–1340.
- Qin B, Lu T, Xu Y, Shen W, Liu F, Xie X, ... Li R (2021) Identification and characterization of An-4, a potential quantitative trait locus for awn development in rice. *BMC Plant Biology* 21(1):1-13. <https://doi.org/10.1186/s12870-021-03055-w>
- Reddy, GE, Suresh BG, Sravan T, Reddy PA (2013) Interrelationship and cause-effect analysis of rice genotypes in the north east plain zone. *The Bioscan* 8(4): 1141-1144.
- Risser P, Rice EL (1971) Diversity in tree species in Oklahoma upland forest. *Ecology* 52: 876-880.
- Robinson H, Comstock R, Harvey P (1949) Estimates of heritability and the degree of dominance in corn. *Agronomy Journal* 41(8):353-359.
- Salem KF, Saleh MM, Aldahak L, Elabd AB (2021) Assessment phenotypic diversity of rice (*Oryza sativa* L.) genotypes by multivariate analysis. *Journal of Aridland Agriculture* 7:52–59. <https://doi.org/10.25081/jaa.2021.v7.6516>
- Sang T, Ge S (2013) Understanding rice domestication and implications for cultivar improvement. *Curr. Opin. Plant Biol.* 16:139-146. <https://doi.org/10.1016/j.pbi.2013.03.003>
- Sharma S, Pokhrel A, Dhakal A, Poudel A (2020) Agro-morphological characterization of rice (*Oryza sativa* L.) landraces of Lamjung and Tanahun district, Nepal. *Annals of Plant Sciences* 9(2):3731-3741.
- Sharmitha T, Justin CGL, Roseleen SSJ, Ramesh T (2019) Assessment of biophysical and biochemical attributes conferring resistance in rice accessions/varieties to yellow stem borer, *Scirpophaga incertulas* Walker (Lepidoptera: Crambidae). *The Pharma Innovation Journal* 8(6) 421-426.
- Shoba D, Vijayan R, Robin S, Manivannan N, Iyanar K, Arunachalam P, Nadarajan N, Pillai MA, Geetha S (2019) Assessment of genetic diversity in aromatic rice (*Oryza sativa* L.) germplasm using PCA and cluster analysis. *Electronic Journal of Plant Breeding* 10(3):1095-1104. <https://doi.org/10.5958/0975-928X.2019.00140.6>
- Simpson EH (1949) Measurement of diversity. *Nature* 163:688.
- Sing M, El-Bizri KS (1992) Phenotypic Correlation: Its Estimation and Testing Significance. *Biometrical Journal* 34(2):165–171.
- Singh B, Singh N, Mishra S, Tripathi K, Singh BP, Rai V, ... Singh NK (2018) Morphological and molecular data reveal three distinct populations of Indian wild rice *Oryza rufipogon* Griff. species complex. *Frontiers in Plant Science* 9:272865. <https://doi.org/10.3389/fpls.2018.00123>

- Sinha AK, Mishra PK (2013) Agro-morphological characterization and morphology-based genetic diversity analysis of landraces of rice variety (*Oryza sativa* L.) of Bankura district of West Bengal. *International Journal of Current Research* 5(10):2764-2769.
- Sivasubramanian S, Menon M (1973) Heterosis and inbreeding depression in rice. *Madras Agric. J* 60(7):1139-1140.
- Sneath PH, Soka RR (1973). Numerical taxonomy. The principles and practice of numerical classification.
- Spellerberg IF, Fedor PJ (2003) A tribute to Claude Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity, and the 'Shannon-Wiener' index. *Global Ecology and Biogeography* 12:177-179.
- SQCC (2023) Notified and Denotified Varieties of Different Crops, Government of Nepal, Ministry of Agricultural and Livestock Development. Seed Quality Control Center, Lalitpur, Nepal.
- Statista, The Statistic Portal. Total rice consumption worldwide from 2008/2009 to 2025/2026 (in 1,000 metric tons)* <https://www.statista.com/statistics/255977/total-global-rice-consumption/> Accessed 6 August 2026
- Stein JC, Yu Y, Copetti D, Zwickl DJ, Zhang L, Zhang C, ... Wing RA (2018) Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover, and innovation across the genus *Oryza*. *Nature Genetics* 50(2):285-296. <https://doi.org/10.1038/s41588-018-0040-0>
- Tafere C, Irie K (2019) Photosynthesis efficiency of culm habit near-isogenic lines rice (*Oryza sativa* L.) to different planting densities. *Int. J. Res. Agric. Sci* 6: 219-229.
- Tiwari DN, Tripathi SR, Tripathi MP, Khatri N, Bastola BR (2019) Genetic variability and correlation coefficients of major characters in early maturing rice under rainfed lowland environments of Nepal. *Advances in Agriculture*: 1-9. <https://doi.org/10.1155/2019/5975901>
- Tuhina-Khatun M, Hanafi MM, Rafii YM, Wong MY, Salleh FM, Ferdous J (2015) Genetic variation, heritability, and diversity analysis of upland rice (*Oryza sativa* L.) genotypes based on quantitative characters. *BioMed Research International*:1-7. <https://doi.org/10.1155/2015/290861>
- Umarani E (2014) Evaluation of landraces of rice (*Oryza sativa* L.) for genetic diversity. Acharya, N. G. Ranga Agricultural University. <http://krishikosh.egranth.ac.in/handle/1/75824>
- Visscher PM, Hill WG, Wray NR (2008) Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics* 9(4):255-266. <https://doi.org/10.1038/nrg2322>
- Walter M, Marchesan E, Massoni PFS, da Silva LP, Sartori GMS, Ferreira RB (2013) Antioxidant properties of rice grains with light brown, red and black pericarp colors and the effect of processing. *Food Research International* 50(2):698-703. <https://doi.org/10.1016/j.foodres.2011.09.002>
- Wang MY, Haberer G, Marri P, Fan C, Goicoeche LJ, Zuccolo A, ... Wing AR (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nature Genet* 46:982-988. <https://doi.org/10.1038/ng.3044>
- Warburton M, Crossa J (2002). Data analysis in the CIMMYT Applied Biotechnology Center for Fingerprinting and Genetic Diversity Studies. CIMMYT, Mexico. 255p.
- Whitaker RH (1972) Evolution and measurement of species diversity. *Taxon* 21:213-251.
- Worede F, Sreewongchai T, Phumichai C, & Sripichitt P (2014) Multivariate analysis of genetic diversity among some rice genotypes using morpho-agronomic characters. *J. Plant Sci* 9(1):14-24. <http://scialert.net/qredirect.php?doi=jps.2014.14.24&linkid=pdf>

- Yan W, Fregeau-Reid J (2018) Genotype by yield×trait (GYT) biplot: a novel approach for genotype selection based on multiple characters. *Scientific Reports* 8(1): 8242. <https://doi.org/10.1038/s41598-018-26688-8>.
- Zaman MR, Paul DNR, Kabir MS, Mahbub MAA, Bhuiya MAA (2005) Assessment of character contribution to the divergence for some rice varieties. *Asian Journal of Plant Sciences* 4:388-391. <https://doi.org/10.3923/ajps.2005.388.391>
- Zhan C, Hu J, Pang Q, Yang B, Cheng Y, Xu E, ... Cheng J (2019) Genome-wide association analysis of panicle exertion and uppermost internode in rice (*Oryza sativa* L.). *Rice* 12(1):1-13. <https://doi.org/10.1186/s12284-019-0330-x>
- Zhang J, Li G, Huang Q, Liu Z, Ding C, Tang S., ... Zhang W (2017) Effects of culm carbohydrate partitioning on basal stem strength in a high-yielding rice population. *The Crop Journal* 5(6):478-487. <https://doi.org/10.1016/j.cj.2017.08.008>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.docx](#)