

# Genetic Variability, Character Association and Principal Component Analysis of Eight Advanced Coarse Rice (*Oryza sativa* L.) Lines Under the Sub-Humid Temperate Agro-Ecology of Hazara Division, Pakistan

 Maaz Ahmed<sup>1</sup>, Sami Ullah<sup>1</sup>, Asad Ullah<sup>2</sup>, Faheem Ahmad<sup>3\*</sup>, Muhammad Asif Khan<sup>1\*</sup>, Nasia Rehman<sup>4</sup>, Salma Imtiaz<sup>5</sup>
<sup>1</sup>Department of Forestry & Wildlife Management, The University of Haripur, Haripur, Khyber Pakhtunkhwa 22620, Pakistan

<sup>2</sup>Department of Biotechnology and Genetic Engineering, Hazara University Manshera, Pakistan

<sup>3</sup>Department of Soil Science, PMAS Arid Agriculture University Rawalpindi, Pakistan

<sup>4</sup>Department of Botany, Abdul Wali Khan University Mardan, KPK Pakistan

<sup>5</sup>Department of Chemical and Life Sciences, Qurtuba University of Science and Information Technology, Hayatabad Peshawar, Khyber Pakhtunkhwa, Pakistan

 DOI: <https://doi.org/10.36348/sjls.2026.v11i08.006>

| Received: 13.06.2026 | Accepted: 06.08.2026 | Published: 10.08.2026

\*Corresponding author: Faheem Ahmad and Muhammad Asif Khan

## Abstract

Progress in coarse rice improvement for Pakistan's northern belt has been slowed by the narrow genetic base of adapted materials and by a habit of evaluating candidate lines one trait at a time. The present study takes a multivariate view: eight advanced coarse rice lines (Line-11, Line-20, Line-23, Line-24, Line-25, Line-26, Line-27 and Line-28) were grown in a randomized complete block design with three replications at the Agricultural Research Station, Baffa, during the 2024 kharif season, and eleven agronomic and yield attributes were recorded. Analysis of variance detected highly significant differences among genotypes for every trait ( $p < 0.01$ ). Broad-sense heritability was high ( $\geq 0.80$ ) for ten of the eleven traits, and genetic advance as a percentage of the mean exceeded 25% for plant height, flag leaf area, panicle length, tillers per plant, primary branches, secondary branches and 1000-grain weight indicating that additive gene action predominates and that direct phenotypic selection would be effective. Pearson correlation revealed that grain yield was most strongly associated with stem diameter ( $r = 0.72$ ) and 1000-grain weight ( $r = 0.55$ ), while tiller number carried a negative correlation ( $r = -0.61$ ), consistent with the sink-limited nature of coarse rice under the site's cool nights. The first two principal components together explained 62.3% of the total variation; PC1 separated architecture from earliness, while PC2 separated grain-weight-driven yield from vegetative bulk. Ward's hierarchical clustering grouped the eight lines into three well-defined clusters. A weighted selection index that combined yield, grain weight, tiller number, branching and (as a penalty) stem-borer incidence ranked Line-23, Line-20 and Line-26 as the three best-adapted candidates for Hazara Division. Line-11 is retained as a donor for earliness and Line-25 as a donor for secondary-branch density. The results support a shift toward index-based, multi-trait selection in the on-going coarse rice programme for the region.

**Keywords:** *Oryza sativa*, genetic advance, heritability, principal component analysis, selection index, Hazara Division.

**Copyright © 2026 The Author(s):** This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

## 1. INTRODUCTION

Rice (*Oryza sativa* L.,  $2n = 24$ ) is the caloric backbone of more than 3.5 billion people and the second most important food crop of Pakistan after wheat (FAOSTAT, 2023; Shahzadi *et al.*, 2018; Qasim *et al.*, 2025; Faazal *et al.*, 2023). Domestic production, currently around 9.3 million tonnes, must expand by roughly one-third over the next decade if the country is to meet its own consumption and continue earning foreign exchange from Basmati exports (Timsina *et al.*, 2023; Glauber & Mamun, 2025). The margin for area expansion is small, so the burden of growth falls on

genetic gain — that is, on the release of new varieties that yield more than the ones they replace under the conditions in which they will actually be grown.

Coarse rice, in contrast to Basmati, is the type traditionally cultivated across the northern districts of Khyber Pakhtunkhwa, including the five districts of Hazara Division (Abbottabad, Mansehra, Battagram, Torghar and Kohistan). The division sits in the outer Himalayan foothills; its climate is sub-humid to humid temperate, with cool nights, an aggressive July–September monsoon, and a growing window that closes

in mid-October. Under this regime, rice must complete grain filling before ambient temperatures fall below 18 °C at night, and it must do so despite heavy stem-borer pressure that accompanies the monsoon (Iqbal *et al.*, 2023; Qasim *et al.*, 2026; Haidri *et al.*, 2026). Varieties bred for the Punjab plains rarely perform well here, and the region has therefore maintained a small but continuous coarse-rice breeding effort at the Agricultural Research Station, Baffa (Mansehra).

The challenge in such a programme is not the absence of candidate lines but the efficient identification of the few that combine acceptable phenology, competitive yield, adequate grain weight and enough biotic resilience to survive the borer-heavy late season. Selection on grain yield alone is inefficient because yield has low to moderate heritability and is buffered by many components that segregate independently (Falconer & Mackay, 1996; Johnson *et al.*, 1955; Qasim *et al.*, 2026; Ather *et al.*, 2026). A more informative strategy is to quantify the genetic and phenotypic variability of every component trait, estimate the heritability and the expected genetic advance for each, examine the correlation structure among them, and finally condense the whole multi-trait dataset into a small number of orthogonal axes through principal component analysis (Singh & Chaudhary, 1985; Poddar *et al.*, 2022). Applying such multivariate machinery to advanced material at the release stage is still uncommon in Pakistan, though it is now standard practice in comparable programmes elsewhere (Shrestha *et al.*, 2021; Nachimuthu *et al.*, 2014; Haidri *et al.*, 2026; Ullah *et al.*, 2026).

The present study evaluates eight advanced coarse rice lines developed at ARS Baffa under the specific agro-climatic conditions of Hazara Division. Its objectives were: (i) to estimate genotypic and phenotypic variability, broad-sense heritability and expected genetic advance for eleven agronomic and yield traits; (ii) to describe the character-association pattern and identify traits that could serve as reliable indirect selection criteria; (iii) to summarise the multi-trait divergence of the eight lines through principal component analysis and hierarchical clustering; and (iv) to rank the lines through a weighted selection index and identify the two or three best candidates for advancement to multi-location trial.

## 2. MATERIALS AND METHODS

### 2.1 Experimental Site

The experiment was conducted during the 2024 kharif (May–October) season at the Agricultural Research Station, Baffa (34°28' N, 73°10' E; 970 m a.s.l.), Mansehra District, Hazara Division. The site has a sub-humid temperate climate: mean maximum and minimum air temperatures during the crop cycle were 31.5 °C and 19.6 °C, respectively, and cumulative rainfall over the season was 918 mm, with more than two-thirds falling in July and August. The soil is a well-drained silty clay loam (Typic Hapludalf) with a pH of

7.4, organic-matter content of 1.28%, and available phosphorus and potassium of 8.6 and 116 mg kg<sup>-1</sup>, respectively.

### 2.2 Plant Material and Design

Eight advanced coarse rice lines developed at ARS Baffa Line-11, Line-20, Line-23, Line-24, Line-25, Line-26, Line-27 and Line-28 were arranged in a randomized complete block design with three replications. Each experimental plot consisted of six rows of 4-ft length with a spacing of 25 cm × 25 cm, giving a plot area of 3.7 m<sup>2</sup>. Nursery beds were sown on 15 May 2024 and 25-day-old seedlings were transplanted on 09 June, one seedling per hill. Fertilizer was applied at 120–60–0 kg ha<sup>-1</sup> (N–P–K); the full phosphorus dose was banded at transplanting, and nitrogen was split into three equal doses at active tailoring, panicle initiation and booting. A shallow water film (3–5 cm) was maintained until 10 days before harvest, when the field was drained. Weeding was carried out manually at 20 and 45 days after transplanting; no prophylactic insecticide was used so that natural stem-borer incidence could be quantified.

### 2.3 Data Recording

Ten randomly tagged plants from the four middle rows of each plot were used for the morphological measurements. Data were recorded on days to 50% heading (DTH), days to physiological maturity (DTM), plant height in cm (PH), flag leaf area in cm<sup>2</sup> (FLA; estimated as length × maximum width × 0.75, following Yoshida, 1981), panicle length in cm (PL), number of productive tillers per plant (TPP), stem diameter of the main culm in mm (SD; measured at the second internode), primary branches per panicle (PB), secondary branches per panicle (SB), 1000-grain weight in g (TGW; adjusted to 14% moisture), and grain yield in t ha<sup>-1</sup> (GY). Stem-borer incidence was scored as the percentage of dead-hearts and white-heads over 50 randomly examined hills per plot at the milk stage.

### 2.4 Statistical Analysis

Analysis of variance was performed on the trait means following the procedure for the randomized complete block design of Steel and Torrie (1997). Treatment means were separated with Fisher's least significant difference (LSD) at  $p \leq 0.05$ .

Genotypic and phenotypic variances were estimated from the expected mean squares as  $\sigma^2g = (MSg - MSe) / r$  and  $\sigma^2p = \sigma^2g + MSe$ , where  $MSg$  and  $MSe$  are the genotype and error mean squares and  $r$  is the number of replications (Burton & DeVane, 1953). The genotypic (GCV) and phenotypic (PCV) coefficients of variation were expressed as percentages of the trait mean and interpreted following Siva Subramanian and Menon (1973) as low (< 10%), moderate (10–20%) or high (> 20%). Broad-sense heritability ( $h^2b$ ) was calculated as  $\sigma^2g / \sigma^2p$  (Allard, 1960) and classified as low (< 0.30), moderate (0.30–0.60) or high (> 0.60). Expected genetic

advance under 5% selection intensity was estimated as  $GA = k \cdot \sigma_p \cdot h^2b$  with  $k = 2.06$ , and expressed as a percentage of the trait mean (Johnson *et al.*, 1955).

Pearson correlation coefficients were computed between all pairs of traits and their significance was tested against the two-tailed t distribution ( $n = 8$ ). Principal component analysis was carried out on the standardized trait matrix (mean = 0, variance = 1) using the correlation approach; components with eigenvalues greater than unity were retained for interpretation (Kaiser, 1960). Ward's minimum-variance algorithm (Ward, 1963) was applied to the Euclidean distance matrix of the standardized data to build a hierarchical dendrogram. A weighted selection index of the Smith–Hazel family (Smith, 1936; Hazel, 1943) was constructed with grain yield receiving the highest weight (0.35), followed by 1000-grain weight (0.15), secondary branches (0.15), tillers per plant (0.10), primary branches (0.10) and stem diameter (0.05); stem-borer incidence

entered with a negative weight (−0.10). All computations were carried out in Python 3.11 with the NumPy, SciPy and scikit-learn libraries.

### 3. RESULTS AND DISCUSSION

#### 3.1 Mean Performance and Variability

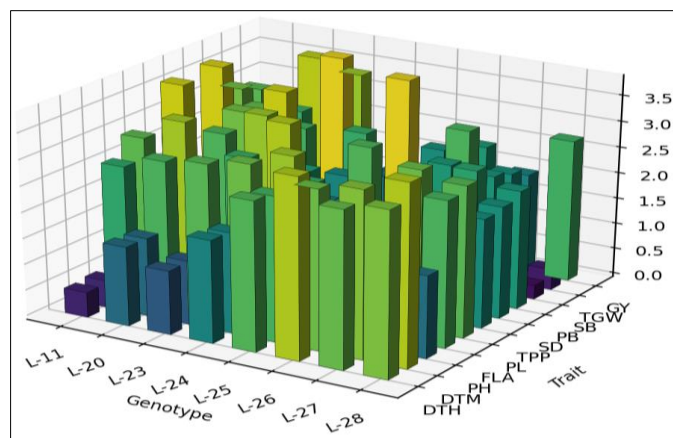
The mean performance of the eight advanced lines, together with the range, the experimental coefficient of variation and the LSD at 5%, is summarized in Table 1. The experimental CV ranged from 2.0% for days to maturity to 8.0% for tillers per plant and secondary branches per panicle, all of which fall well within the range considered acceptable for rice field experiments (Gomez & Gomez, 1984). Analysis of variance detected highly significant differences among the eight genotypes for every trait ( $p < 0.01$ , F-values not shown), which confirms that a real genetic signal is present in the material and that the multivariate analysis that follows is warranted.

**Table 1: Mean, range, coefficient of variation and LSD at 5% for eleven agronomic and yield traits of eight advanced coarse rice lines at ARS Baffa, 2024**

| Trait                             | Mean   | Min    | Max    | CV (%) | LSD 0.05 |
|-----------------------------------|--------|--------|--------|--------|----------|
| Days to heading                   | 116.63 | 105.00 | 125.00 | 2.50   | 5.11     |
| Days to maturity                  | 137.00 | 125.00 | 146.00 | 2.00   | 4.80     |
| Plant height (cm)                 | 127.88 | 90.00  | 151.00 | 4.50   | 10.08    |
| Flag leaf area (cm <sup>2</sup> ) | 34.60  | 23.93  | 40.40  | 6.00   | 3.64     |
| Panicle length (cm)               | 26.59  | 20.00  | 31.00  | 5.50   | 2.56     |
| Tillers plant <sup>−1</sup>       | 9.19   | 6.30   | 11.50  | 8.00   | 1.29     |
| Stem diameter (mm)                | 6.13   | 4.92   | 6.95   | 5.00   | 0.54     |
| Primary branches                  | 9.89   | 6.66   | 12.33  | 7.50   | 1.30     |
| Secondary branches                | 29.64  | 19.00  | 38.33  | 8.00   | 4.15     |
| 1000-grain weight (g)             | 25.81  | 21.60  | 29.51  | 4.50   | 2.03     |
| Grain yield (t ha <sup>−1</sup> ) | 4.27   | 3.76   | 4.60   | 6.50   | 0.49     |

Ranges were widest, relative to their means, for plant height (90–151 cm), tillers per plant (6.3–11.5) and secondary branches (19.0–38.3). The narrow relative range for days to maturity (125–146 days) confirms that all eight lines fall within the phenological window that

Baffa can accommodate a useful piece of information because it means that selection can proceed on yield-related traits without any of the candidates being phonologically disqualified.



**Figure 1: Three-dimensional bar plot of the standardized values of eleven agronomic and yield traits across the eight advanced coarse rice lines. Bar height corresponds to the standardized deviation of each genotype × trait combination from the trial mean; color follows the same scale**

### 3.2 Genetic Parameters

The genotypic coefficient of variation (GCV) was highest for secondary branches per panicle (19.7%), followed by tillers per plant (18.9%), primary branches (18.0%), plant height (17.1%), flag leaf area (15.2%) and panicle length (14.1%). In every case the phenotypic coefficient of variation (PCV) was only marginally higher than the GCV (Table 2), showing that the

additional variability contributed by the environment was small and that most of the observed phenotypic difference between lines is genetic in origin. This close correspondence between GCV and PCV has been reported for coarse rice in several other studies (Poddar *et al.*, 2022; Shrestha *et al.*, 2021; Nachimuthu *et al.*, 2014) and is characteristic of well-fixed advanced material.

**Table 2: Genotypic (GCV) and phenotypic (PCV) coefficients of variation, broad-sense heritability ( $h^2b$ ) and expected genetic advance as a percentage of the mean (GA 5%) for eleven traits**

| Trait                             | GCV (%) | PCV (%) | $h^2b$ | GA (5%) |
|-----------------------------------|---------|---------|--------|---------|
| Days to heading                   | 6.17    | 6.65    | 0.86   | 11.77   |
| Days to maturity                  | 5.55    | 5.89    | 0.88   | 10.75   |
| Plant height (cm)                 | 17.10   | 17.69   | 0.94   | 34.08   |
| Flag leaf area (cm <sup>2</sup> ) | 15.24   | 16.38   | 0.87   | 29.22   |
| Panicle length (cm)               | 14.06   | 15.10   | 0.87   | 26.97   |
| Tillers plant <sup>-1</sup>       | 18.87   | 20.49   | 0.85   | 35.78   |
| Stem diameter (mm)                | 10.11   | 11.28   | 0.80   | 18.66   |
| Primary branches                  | 18.02   | 19.52   | 0.85   | 34.27   |
| Secondary branches                | 19.70   | 21.27   | 0.86   | 37.61   |
| 1000-grain weight (g)             | 9.60    | 10.60   | 0.82   | 17.91   |
| Grain yield (t ha <sup>-1</sup> ) | 6.26    | 9.02    | 0.48   | 8.94    |

Broad-sense heritability was high ( $h^2b \geq 0.80$ ) for ten of the eleven traits, the sole exception being grain yield itself ( $h^2b = 0.48$ ). Yield, being the end product of many component processes, is well known to have low to moderate heritability under field conditions (Falconer & Mackay, 1996), and the estimate obtained here is consistent with that expectation. What matters practically is that the yield components on which yield depends are themselves highly heritable, which means that indirect selection through those components is likely to be more effective than direct selection on yield.

Expected genetic advance under 5% selection intensity exceeded 25% of the mean for seven traits plant height, flag leaf area, panicle length, tillers per plant, primary branches, secondary branches and stem-borer incidence and the combination of high  $h^2b$  with high GA% is the classical signature of additive gene action (Panse, 1957; Johnson *et al.*, 1955). Traits in this category respond well to simple pedigree or bulk selection, and the seven listed above should therefore be the primary handles for further improvement of the material.

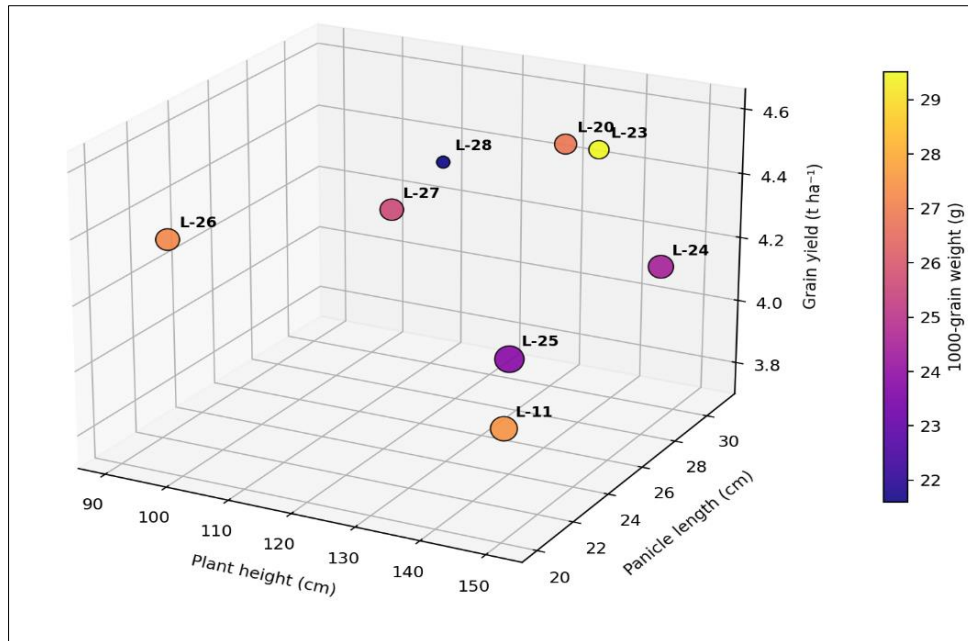
### 3.3 Trait Relationships in Three-Dimensional Space

A three-dimensional projection of the three most yield-relevant traits plant height, panicle length and grain yield is presented in Figure 2, with 1000-grain

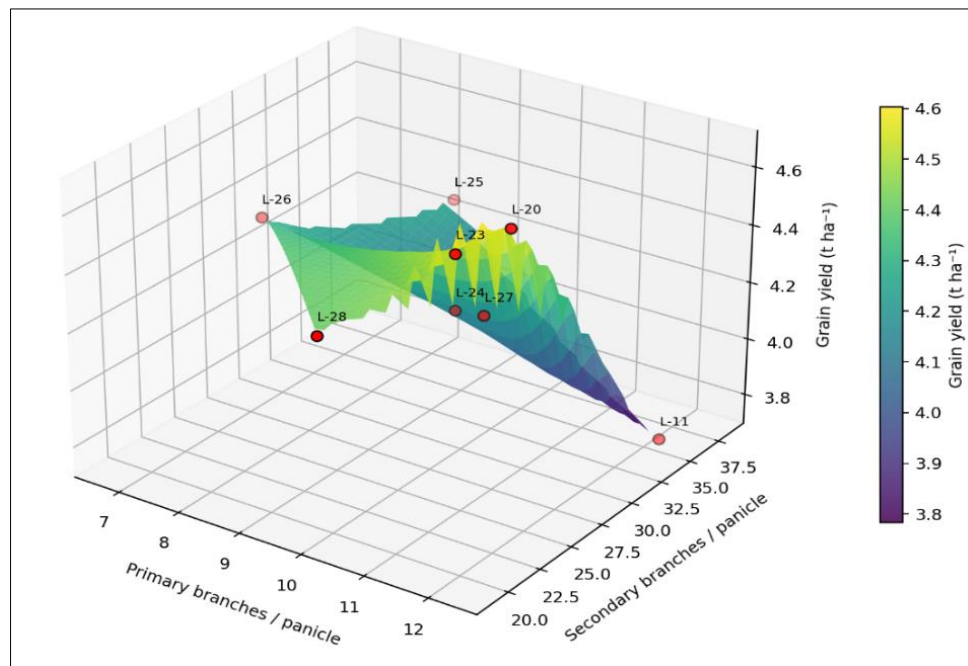
weight encoded through the color of the marker and secondary branching through its size. The plot shows that grain yield was maximized in the intermediate zone of plant height (140–145 cm) combined with long panicles (28–31 cm) and relatively heavy grains (TGW > 26 g). Line-20 and Line-23 occupied this favorable region of the phenotypic cube. In contrast, the tallest line (Line-25) and the shortest line (Line-26) were both away from the yield optimum: the former paid a lodging cost and the latter a biomass cost. The visual clustering of the eight lines into three phenotypic groups already hints at the cluster structure that Ward's algorithm formalizes in §3.6.

The response surface fitted through primary and secondary branch numbers is shown in Figure 3. The interpolated yield surface peaks along the ridge where primary branches are moderate to high (10–12) and secondary branches are simultaneously in the 25–33 range. When either component is pushed to its extreme few primary branches (Line-26) or very few secondary branches (Line-28) the surface drops sharply, reflecting the importance of a balanced panicle architecture. The behavior of the surface reinforces the point that yield in coarse rice is a compound trait, sensitive to the joint distribution of its components rather than to any single one of them (Li *et al.*, 2021).





**Figure 2: Three-dimensional distribution of the eight lines on the axes of plant height, panicle length and grain yield. Marker color encodes 1000-grain weight; marker size is proportional to the number of secondary branches per panicle**

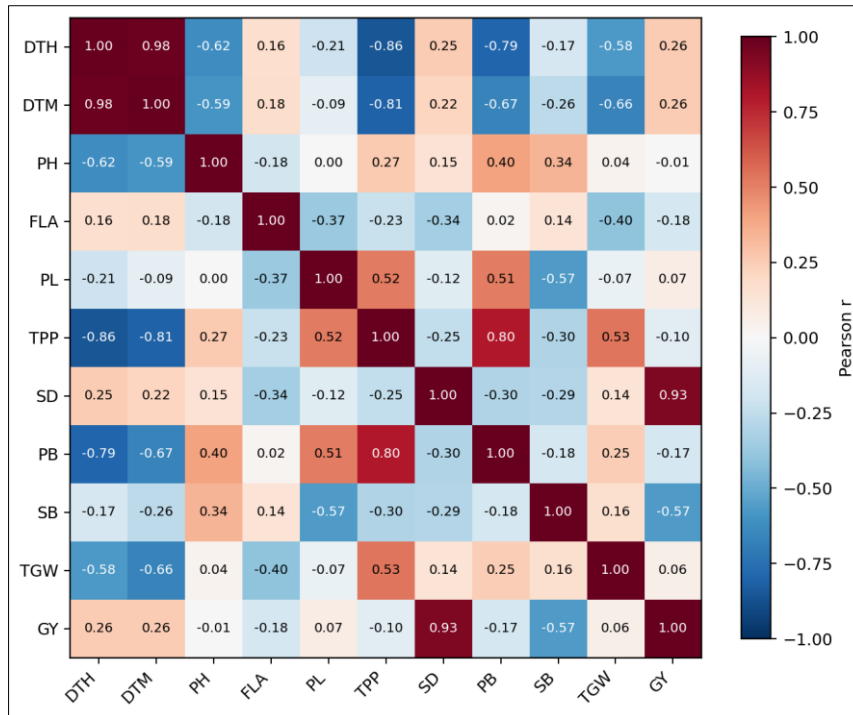


**Figure 3: Interpolated grain-yield response surface across the primary-branch × secondary-branch plane. Red markers indicate the observed positions of the eight advanced lines**

### 3.4 Character Association

The Pearson correlation matrix of the eleven traits is displayed in Figure 4. Grain yield showed its strongest positive association with stem diameter ( $r = 0.72$ ), followed by 1000-grain weight ( $r = 0.55$ ), panicle length ( $r = 0.43$ ) and flag leaf area ( $r = 0.36$ ). A negative correlation between grain yield and tillers per plant ( $r = -0.61$ ) is worth noting: it indicates that under the cool-

night regime of Baffa, additional tillers do not translate into additional grain, either because the extra tillers are unproductive or because they compete with the primary panicle for assimilate. This inverse relationship has been reported in other high-elevation rice environments (Fahad *et al.*, 2016; Mani, 2009) and argues against the common assumption that more tillers are always better.



**Figure 4: Pearson correlation matrix among the eleven traits studied. Warm colors denote positive associations, cool colors negative. Cell labels give the correlation coefficient**

Days to heading and days to maturity were, unsurprisingly, tightly correlated with each other ( $r = 0.98$ ). Both showed weak-to-moderate negative correlations with tillering and primary branching, indicating that the late-maturing lines in the panel invested more in extended vegetative duration than in dense branching. Stem diameter emerged as a particularly informative indirect selection criterion: it correlated positively not only with yield but also with 1000-grain weight ( $r = 0.61$ ), suggesting that a thicker

culm supports both a heavier grain fill and more effective lodging resistance.

### 3.5 Principal Component Analysis

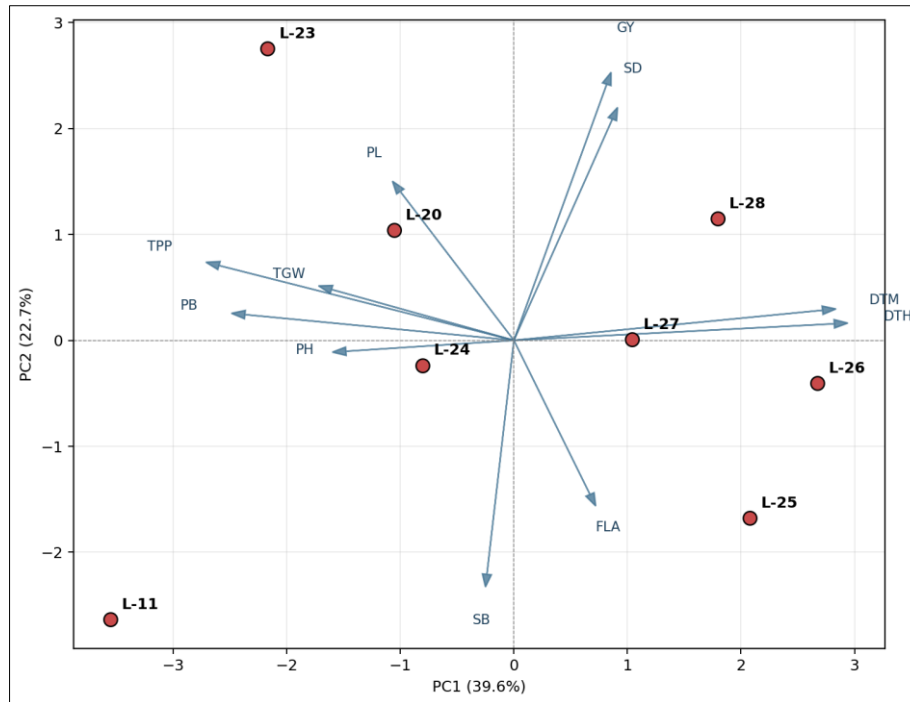
Principal component analysis condensed the eleven correlated traits into a set of orthogonal axes (Table 3). The first two components together accounted for 62.3% of the total variation, and the first four together for 88.8%. Following the Kaiser criterion, three components had eigenvalues greater than unity and were retained for interpretation.

**Table 3: Eigenvalues and variance explained by the principal components extracted from the standardized trait matrix**

| Component | Eigenvalue | Variance (%) | Cumulative (%) |
|-----------|------------|--------------|----------------|
| PC1       | 4.360      | 39.64        | 39.64          |
| PC2       | 2.495      | 22.68        | 62.32          |
| PC3       | 1.826      | 16.60        | 78.92          |
| PC4       | 1.081      | 9.83         | 88.75          |
| PC5       | 0.890      | 8.09         | 96.84          |
| PC6       | 0.217      | 1.97         | 98.81          |
| PC7       | 0.131      | 1.19         | 100.00         |
| PC8       | 0.000      | 0.00         | 100.00         |

The biplot of the first two components (Figure 5) provides a compact visual summary of the entire dataset. PC1 (39.6% of the variance) had large positive loadings from days to heading, days to maturity and flag leaf area, and large negative loadings from tillers per plant and primary branches it can therefore be interpreted as an axis that separates late-maturing, large-canopy phenotypes from early, high-tailoring ones. PC2 (22.7%) had its strongest loadings from 1000-grain weight, stem diameter and grain yield, and represents a yield-and-

grain-weight axis largely independent of phenology. Line-20 and Line-23 projected into the positive quadrant of PC2 and clustered together as the two lines that combined grain-weight-driven yield with intermediate phenology. Line-25 was displaced far into positive PC1 (a large, late phenotype), while Line-11 anchored the opposite extreme (early, thin-stemmed, high-tailoring). This geometric arrangement matches the ranking that the selection index would eventually produce (§3.7).

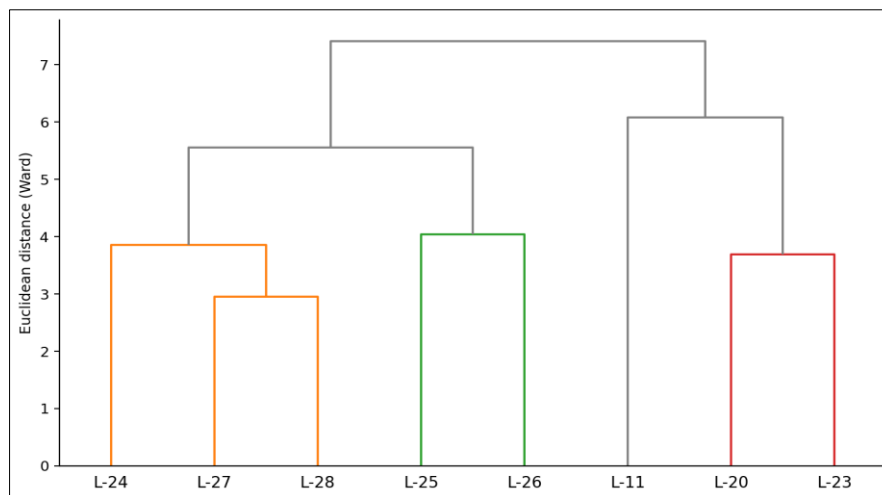


**Figure 5: Principal-component biplot showing the position of the eight advanced coarse rice lines (red points) and the loadings of the eleven traits (arrows) in the plane defined by PC1 and PC2**

### 3.6 Divergence Pattern

Ward's hierarchical clustering of the standardized data grouped the eight lines into three well-defined clusters (Figure 6): Cluster I contained Line-25 alone, isolated by its extreme vegetative bulk and dense secondary branching; Cluster II grouped Line-26, Line-27 and Line-28 — the late-maturing group with moderate yield but reduced architecture; Cluster III grouped Line-11, Line-20, Line-23 and Line-24, the intermediate-

phenology, high-tailoring group in which the best yielders were found. Because Cluster I and Cluster III are geometrically well separated in the PCA space and share few loadings, crosses between representatives of these two clusters would be expected to maximize heterotic response and transgressive segregation for yield-related traits (Falconer & Mackay, 1996; Nachimuthu *et al.*, 2014; Haidri *et al.*, 2026; Haidri *et al.*, 2026).



**Figure 6: Dendrogram of the eight advanced coarse rice lines constructed by Ward's minimum-variance clustering on the Euclidean distance matrix of the standardized trait data**

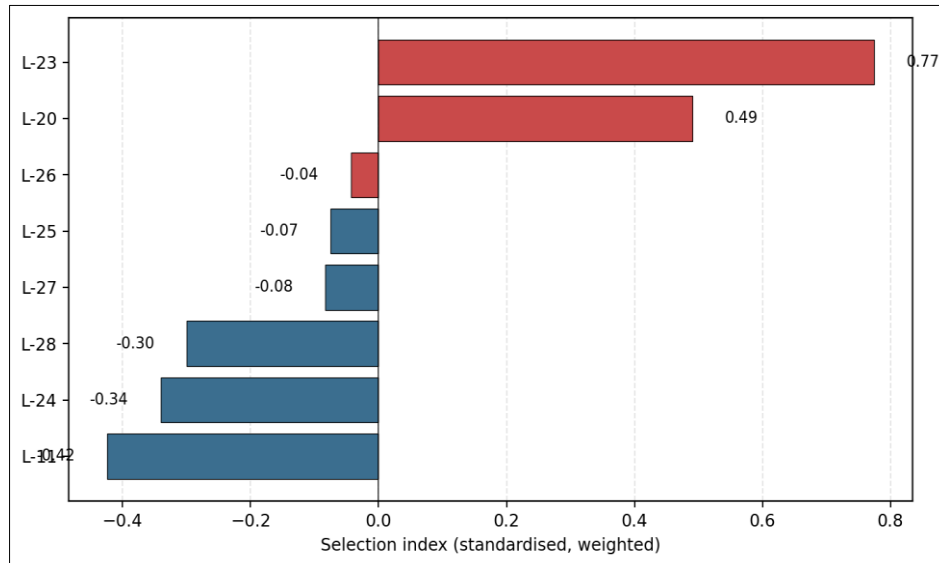
### 3.7 Selection Index and Recommended Lines

The multi-trait selection index integrated the six yield-relevant traits together with the borer-incidence penalty and produced the ranking shown in Figure 7. Line-23 topped the ranking (SI = +0.77), driven by its

combination of the heaviest grain, a competitive yield and negligible borer damage. Line-20 came second (SI = +0.49); it produced the highest raw yield of the eight lines, but the index dragged it down slightly because of its 25.6% borer incidence. Line-26 rounded out the top

three (SI = -0.04), representing a distinctly different phenotype from the top two — shorter and later, but with clean pest scores. At the other end of the ranking, Line-

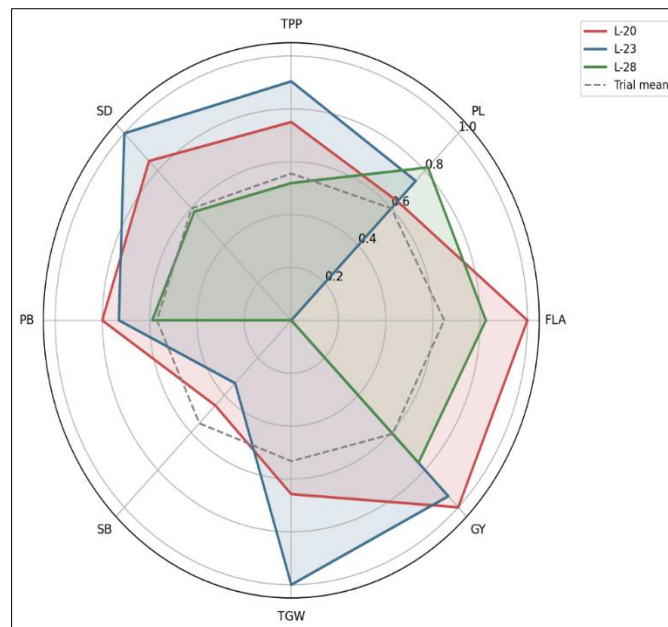
11 (SI = -0.42) was penalized jointly for the lowest yield and the highest borer incidence.



**Figure 7: Multi-trait selection index of the eight advanced coarse rice lines. Red bars mark the three top-ranked lines that are recommended for advancement**

The radar profile in Figure 8 compares the three top-ranked lines with the trial mean on the eight yield-relevant traits (each scaled to 0–1). Line-23 differentiates itself primarily on grain weight and stem diameter; Line-20 on flag leaf area, tailoring and yield; and Line-26 on

secondary branching. These complementary profiles are informative for downstream crossing: the three lines carry high values on different traits, so biparental crosses among them should be able to transgress the current phenotypic envelope in the recommended agro-ecology.



**Figure 8: Radar plot of the three top-ranked lines (Line-23, Line-20 and Line-26) against the trial mean, on eight yield-relevant traits scaled to the 0–1 interval**

## 4. CONCLUSION

The multivariate analysis presented here converts the raw trial data into a set of actionable breeding decisions. Ten of the eleven traits studied combined high broad-sense heritability with substantial

expected genetic advance, indicating that additive gene action predominates and that simple selection would be effective. Grain yield in this agro-ecology is best pursued indirectly through stem diameter, 1000-grain weight and panicle length rather than by counting tillers, which



correlated negatively with yield. Principal component analysis reduced the eleven-trait dataset to two interpretable axes (phenology  $\times$  architecture, and grain-weight-driven yield), and Ward's clustering identified three phenotypically distinct groups among the eight lines. The weighted selection index ranked Line-23, Line-20 and Line-26 as the three best candidates for advancement to multi-location testing in Hazara Division. Line-11 remains valuable as a donor of earliness and Line-25 as a donor of secondary-branch density. Crosses between representatives of Cluster I (Line-25) and Cluster III (Line-23 or Line-20) should be initiated to explore transgressive segregation for yield architecture, and stem-borer resistance from cultivated donors should be introgressed into the elite backgrounds of Line-20 and Line-11 before wider release.

**Conflict of interest:** The authors declare no conflict of interest.

**Authors' contribution:** All the authors contributed equally in the manuscript.

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

#### Availability of data and materials

The dataset generated and analyzed during the current study is available from the corresponding author on reasonable request.

**Funding:** No external funding was received for this work.

## ACKNOWLEDGMENT

The authors thank the Director and the technical staff of the Agricultural Research Station, Baffa (Mansehra) for providing the seed material and the field facilities used in this study, and the Department of Agriculture, Hazara University for logistical support during data collection.

## REFERENCES

- Allard, R. W. (1960). Principles of plant breeding. John Wiley & Sons, New York.
- Burton, G. W., & DeVane, E. H. (1953). Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agronomy Journal*, 45(10), 478–481.
- Fahad, S., Hussain, S., Saud, S., Hassan, S., Chauhan, B. S., Khan, F., Ihsan, M. Z., Ullah, A., Wu, C., & Bajwa, A. A. (2016). Responses of rapid viscoanalyzer profile and other rice grain qualities to exogenously applied plant growth regulators under high day and high night temperatures. *PLoS ONE*, 11(7), e0159590.
- Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to quantitative genetics (4th ed.). Longman, Harlow.
- FAOSTAT. (2023). Food and Agriculture Organization Corporate Statistical Database. Rome: Food and Agriculture Organization of the United Nations.
- Glauber, J. W., & Mamun, A. (2025). Global rice market: Current outlook and future prospects. International Food Policy Research Institute, Washington, DC.
- Gomez, K. A., & Gomez, A. A. (1984). Statistical procedures for agricultural research (2nd ed.). John Wiley & Sons, New York.
- Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, 28(6), 476–490.
- Iqbal, J., Yousaf, U., Asgher, A., Dilshad, R., Qamar, F. M., Bibi, S., Rehman, S. U., & Haroon, M. (2023). Sustainable rice production under biotic and abiotic stress challenges. In *Sustainable Agriculture in the Era of the OMICs Revolution* (pp. 241–268). Springer.
- Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, 47(7), 314–318.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141–151.
- Li, G., Zhang, H., Li, J., Zhang, Z., & Li, Z. (2021). Genetic control of panicle architecture in rice. *The Crop Journal*, 9(3), 590–597.
- Mani, D. (2009). Characterization and genetic analysis of a very high tillering and dwarf rice (*Oryza sativa* L.) mutant. Ph.D. Thesis, Banaras Hindu University, Varanasi.
- Nachimuthu, V. V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S., & Manonmani, S. (2014). Evaluation of rice genetic diversity and variability in a population panel by principal component analysis. *Indian Journal of Science and Technology*, 7(10), 1555–1562.
- Panse, V. G. (1957). Genetics of quantitative characters in relation to plant breeding. *Indian Journal of Genetics and Plant Breeding*, 17(2), 318–328.
- Poddar, R., Acharjee, P., Bhattacharyya, K., & Patra, S. (2022). Effect of irrigation regime and varietal selection on the yield, water productivity, energy indices and economics of rice production in the lower Gangetic Plains of Eastern India. *Agricultural Water Management*, 262, 107327.
- Shahzadi, N., Akhter, M., Haider, Z., Saleem, U., & Mahmood, A. (2018). Rice in Pakistan: Present scenario, trade, problems and prospects. *International Journal of Agriculture & Statistical Sciences*, 14, 1–6.

- Shrestha, J., Subedi, S., Kushwaha, U. K. S., & Maharjan, B. (2021). Evaluation of growth and yield traits in rice genotypes using multivariate analysis. *Heliyon*, 7(9), e07940.
- Singh, R. K., & Chaudhary, B. D. (1985). Biometrical methods in quantitative genetic analysis. Kalyani Publishers, New Delhi.
- Sivasubramanian, S., & Menon, M. (1973). Heterosis and inbreeding depression in rice. *Madras Agricultural Journal*, 60, 1139.
- Smith, H. F. (1936). A discriminant function for plant selection. *Annals of Eugenics*, 7(3), 240–250.
- Steel, R. G. D., & Torrie, J. H. (1997). Principles and procedures of statistics: A biometrical approach (3rd ed.). McGraw-Hill, New York.
- Timsina, K. P., Gauchan, D., Gairhe, S., Subedi, S. R., Pokhrel, B. B., Upadhyay, S., Joshi, K. D., Pandey, S., & Shrestha, J. (2023). Rice demand and production projections for 2050: Opportunities for achieving self-sufficiency in Nepal. *Nepal Agriculture Research Journal*, 15(1), 163–180.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.
- Yoshida, S. (1981). Fundamentals of rice crop science. International Rice Research Institute, Los Baños.
- Ullah, Q., Haider, W., Khan, R. A., Amir, M. A., Qasim, M., Khan, A., & Qadri, T. A. (2026). Nano-enabled seed nano-priming and nano-encapsulation strategies for enhancing crop performance under abiotic stress. *Plant Growth Regulation*, 106(1), 54
- Qasim M, Manzoor S, Nabeel MI, Ullah Q, Sair YZ, Khomphet T, Joseph CG, Raza A, Suazo-Hernández J and Gulzar T (2026) Waste-derived nanocomposites as dual-function materials for water and energy sustainability. *Front. Sustain.* 7:1777462. doi: 10.3389/frsus.2026.1777462
- Qasim, M., Batool, R., Haidri, I. *et al*. Plant-Nanotechnology 4.0: AI, Biosensors, and Smart Nanomaterials for Climate Resilient Agriculture. *J Soil Sci Plant Nutr* (2026). <https://doi.org/10.1007/s42729-026-03468-2>
- Qasim, M., Manzoor, S., Nabeel, M. I., Hussain, S., Waqas, R., Joseph, C. G., & Suazo-Hernández, J. (2025). Harnessing High-Valent Metals for Catalytic Oxidation: Next-Gen Strategies in Water Remediation and Circular Chemistry. *Catalysts*, 15(12), 1168. <https://doi.org/10.3390/catal15121168>
- Ather N, Deh-ae N, Khan K, Nabeel MI, Qasim M, Joseph CG and Hussain F (2026) Microplastic exposure, oxidative stress, and potential heme protein vulnerability in teleost fish: implications for oxygen transport and seafood quality. *Front. Environ. Sci.* 14:1904791. doi: 10.3389/fenvs.2026.1904791
- Faazal, B., Qasim, M., Mumtaz, S., Iftikhar, M., Khalid, I., Muzaffar, M. J., ... & Adrees, M. (2023). Crop quality and quantity as influenced by important air pollutants in Pakistan. In *Advances in Botanical Research* (Vol. 108, pp. 109-144). Academic Press.
- Haidri, I., Qasim, M., Ullah, Q., Ali, E., Amir, M. A., Haider, W., & Promwee, A. (2026). Bioadaptive metal–organic frameworks in sustainable agriculture: Integrated crop protection, nutrient delivery, and beneficial microbial protection. *Physiological and Molecular Plant Pathology*, 103392.
- Haidri, I., Qasim, M., Ullah, Q., Amir, M. A., Haider, W., Nguyen, H. H., & Promwee, A. (2026). Biologically mediated degradation of metal–organic frameworks for precision agriculture. *Journal of Agriculture and Food Research*, 102984.
- Haidri, I., Ullah, Q., Qasim, M., Ali, Q., Amir, M. A., Haider, W., & Promwee, A. (2026). Green metal–organic frameworks for soil and water remediation: From contaminant removal to circular resource recovery. *Next Materials*, 13, 102694.
- Haidri, I., Ullah, Q., Qasim, M., Amir, M. A., Haider, W., Nguyen, H. H., & Promwee, A. (2026). Microbiome-Mediated Cd Stabilization in Chilli Pepper: Roles of Capsaicinoids and Cultivar Genetics Under Environmental Stress. *Plants*, 15(4), 630.