**Utilizing Multivariate Analysis to Enhance Selection Methods for Seed Quality Characteristics in Rice (*Oryza sativa* L.)**

**ABSTRACT**

The experiment was carried out at the genetics and plant breeding laboratory, lovely professional university, Jalandhar (Punjab). Forty-five genotypes of rice (*Oryza sativa* L.) were investigated. Analysis of variance showed significant variation among the genotypes for all seven quantitative traits under evaluation, indicated presence of variation in the populations. The genotypes showed high variability for most traits, reflecting high potential for selection to enhance yield. Genotypes Ram Lakshman, IET-22020, SHIVANTH, DDR-119 had better mean seed vigou index, while Ashoka 200, Ruchi Dhan, PR 131 and HUR-36 showed the maximum germination speed, revealed as early types of maturing and hence the promising lines of choice for any breeding program. The investigation recorded high GCV and PCV seedling dry weight, which was trailed by speed of germination, seedling length and root length, implying large genetic variation for these traits and thus effective direct selection. The traits had high heritability and genetic advance, showing that they are regulated by additive gene effects, making them ideal for selection. Correlation analysis indicated seedling vigour index had positive correlations with standard germination, shoot length, root length and seedling length. The same traits also had maximum direct effects on seed yield at both genotypic and phenotypic levels. Cluster analysis showed that the genotypes are classified in 8 clusters having the maximum intra cluster distance in cluster 8 followed by cluster V, III and IV. Principal component analysis resulted in five principal components, which accounted for 80.11% of the total variation.

***Keywords:*** *GCV, PCV, correlation coefficient, Cluster analysis, Principal component analysis.*

1. **INTRODUCTION**

Rice (*Oryza sativa* L.) having chromosome number 2n = 24, an important cereal crop of the Poaceae family and Oryzoidea subfamily, is commonly known as the "Global Grain" because it is used as a staple food in over 100 nations. Global rice production and consumption of approximately 90% take place in Asia alone, where it is a major food for close to half the world's population (Yugandhar *et al*., 2018). Rice supplies around 20% of the world's dietary energy supply, made up mostly of starch, with about 78–79% amylose and amylopectin content. Besides carbohydrates, rice is a rich source of protein and vital vitamins and also contains especially high levels of B vitamins like thiamine and niacin, and also essential minerals including iron (Fe), phosphorus (P), and magnesium (Mg). With the population of the world increasing fast, there is an immediate need to increase rice production to cater to future food needs. The world's population, as estimated by 2050, will be 9.1 billion, necessitating a 70% increase in food production (Singh *et al*., 2021). Rice is grown in an area of 165.03 million ha in the world with a production of 776.46 million metric tons and productivity of 4.7 tons/ha (FAOSTAT 2023).

Genetic variability among traits is the secret to choosing desirable genetic material for producing high-yielding types, which is the primary objective in plant breeding (Tiwari *et al*., 2019). The genetic improvement of a genotype with success needs knowledge of some genetic parameters, such as Phenotypic Coefficient of Variation, Genotypic Coefficient of Variation, broad-sense heritability, and genetic advance of the rice (Kishore *et* *al*., 2015). Rice is highly genetically variable due to genetic variation among individuals within a population, which is essential for the selection of genetically variable parents for the generation of new genetic stocks (Nath and Kole, 2021).

Correlation coefficient is a statistical quantity employed to establish the magnitude and direction of the relationship between two or more variables. When the relation is noticeable directly, it is termed as phenotypic correlation, encompassing genetic and environmental factors. Genotypic correlation results from genes' pleiotropic effects, genetic linkage, or both. Depending on whether the involved traits are favourable or unfavourable, this correlation may help or hamper genetic improvement (Babu *et* *al*., 2012).

Path analysis represents a more wide-ranging approach by identifying cause-and-effect relationships. It breaks down correlations into direct and indirect effects, demonstrating how independent traits influence a dependent trait. This technique is especially valuable in breeding studies, where understanding the specific contributions of individual traits to yield or other key outcomes is essential for effective selection and crop improvement (Mahendran *et al*., 2024).

Mahalanobis cluster analysis is a useful statistical method in plant breeding for evaluating genetic diversity based on several traits at a time. Agro-morphological traits are usually favoured by breeders in this analysis since they are easy, economical, and straightforward to measure. Mahalanobis D² statistic assists in the identification of genetically diverse genotypes, which are very important in efficient parent selection and hybridization. Such genotypes can be obtained from various areas or bred to improve crop performance and variability. (Ranjith *et* *al*., 2018; Kumar et al., 2022). Principle component analysis is one of the foremost multivariate analysis and processing of data methods which identifies the minimum number of components that account for the maximum total variability in a data set. it is also used to arrange genotypes in order of their PC scores. These PCs are usually calculated from the correlation matrix or covariance matrix (Christina *et al*., 2021).

**2. MATERIALS AND METHODS**

**2.1 Experimental Site and Materials**

The experiment was conducted at the genetics and plant breeding laboratory of School of Agriculture, Lovely Professional University, Jalandhar, Punjab (31° 15' 47" N latitude and 75° 41' 20" E longitude). Forty-five genotypes of Rice germplasms from Kapurthala district were used as experimental material with PR-131, Ruchi Dhan and Ashoka 200 checks. Genotypes were collected from the department of genetics and plant breeding, Lovely professional university. The details for materials used is present in **Table 1.**

**Table 1:** Experimental material details for rice

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype code** | **Genotype** | **Genotype code** | **Genotype** | **Genotype code** | **Genotype** |
| **G1** | PR-131(check) | **G16** | AAU DR-1 | **G31** | R-2370-115-1-93-1 |
| **G2** | Ruchi Dhan (check) | **G17** | DDR-100 | **G32** | Ram Lakshman |
| **G3** | Ashoka 200(check) | **G18** | IET-20860 | **G33** | Haldi Ghati |
| **G4** | DDR-119 | **G19** | DDR-130 | **G34** | Tama Koni |
| **G5** | DDR-101 | **G20** | IET-22021 | **G35** | Goyadi |
| **G6** | Safari-2 | **G21** | Moha Dhan | **G36** | Kani Dhan |
| **G7** | Bami Sundari | **G22** | URG-1 | **G37** | Fara Dhan |
| **G8** | Shivanth | **G23** | Katak Dhan | **G38** | Sonagathi |
| **G9** | Sanchorma | **G24** | Matko Dhan | **G39** | kalamati |
| **G10** | Safari | **G25** | Khajur | **G40** | IR 167 1662 |
| **G11** | Bhushu | **G26** | Mala Gauri | **G41** | IR 82635-B-B-75-2 |
| **G12** | IET-22020 | **G27** | Panki guda | **G42** | CO-51 |
| **G13** | PR-110 | **G28** | R-2449-1037-1-463-1 | **G43** | SWARNA SUB-1 |
| **G14** | Kanchan | **G29** | R-2307-43-3-1-38-1 | **G44** | SAMBA SUB-1 |
| **G15** | Asamiya Dhan | **G30** | R-2370-112-2-90-1 | **G45** | HUR-36 |

**2.2 Observations recorded**

The important seven seed quality traits *viz*., Speed of germination (SOG, %), Standard germination (SG, %), Shoot length (SL, cm), Root length (RL, cm), Seedling length (SDL, cm), Seedling dry weight (SDW, g) and Seedling vigour index (SVI) under laboratory condition were observed and recorded from each lot of each replication.

**2.3 Statistical analysis**

The statistical analysis was carried out using latest version of R software. Analysis of variance, variability, correlation coefficient, path analysis and PCA parameters were calculated in accordance to Robinson *et al*., (1955), Dewey and Lu (1959), Pearson (1901).

**3. RESULTS AND DISCUSSION**

**3.1 Analysis of variance (ANOVA) for seed quality traits in rice**

For improving seed quality traits, plant breeders depend on natural genetic variability existing in the populations or opt for the generation of new variability through different plant breeding methods. The mean sum of squares due to genotypes were found to be significant for all the traits under study indicating the presence of significant differences in the traits across the genotypes, while mean squares due replications were non-significant for all the traits as represented in (**Table 2)**. The similar finding was obtained by (Hoque *et al*., 2021).

**Table 2** Analysis of variance for all seed quality traits in rice

|  |  |  |  |
| --- | --- | --- | --- |
| **Characters** | **Source of Variation** | | |
| **Replication** | **Genotypes** | **Error** |
| **Degree of freedom** | **2** | **44** | **87** |
| **Speed of Germination (%)** | 3.98 | 505.87\*\* | 0.65 |
| **Standard Germination (%)** | 0.26 | 17.37\*\* | 1.03 |
| **Shoot Length (Cm)** | 0.07 | 4.60\*\* | 0.30 |
| **Root Length (CM)** | 0.16 | 3.83\*\* | 0.25 |
| **Seedling Length (CM)** | 0.07 | 9.45\*\* | 0.38 |
| **Seedling Dry Weight (g)** | 0.02 | 0.05\*\* | 0.06 |
| **Seedling Vigour Index** | 3741.00 | 83186\*\* | 4241.00 |

**Note: \*\* significant at 1% probability level, respectively, Where, \*\*=highly significant.**

**3.2 Mean performance**

Comparative analysis of genotypic means with the overall population mean serves as an effective strategy for identifying superior and inferior performers based on laboratory data. The general mean for SOG was found to be 102.00 which ranged from 123.95 (Ashoka 200) to 79.96 (DDR 119). However, genotypes Ashoka 200, Ruchi Dhan, PR 131 and HUR-36 showed the maximum germination speed making them fir for the selection in rice breeding. Average of SG was found to be 92.92% and the genotypes PR 110 (97.60%) followed by Haldi Ghati (96.74%), SWARNA SUB-1 (96.68%), Shivanth (96.49%) which was found to be best suited for the selection in rice breeding programme.

For SL the summed mean was found highest in genotypes IET-22020 (14.42 cm) followed by RAM LAKSHMAN (14.14 cm), DDR-119 (13.72cm), SHIVANTH (13.69 cm) respectively.Root length had a mean ranging between 6.97 (ASHOKA 200) to 11.40 (RAM LAKSHMAN) with average value 9.10 cm. The trait SDW had general mean of 0.12g among which the genotypes KHAJUR (0.17 g) followed by TAMA KONI (0.16 g), R-2370-112-2-90-1 (0.15 g), MALA GAURI (0.14 g) which were identified as the best genotypes, represented in (**Fig 1**). The high amount of genetic variation for most of these traits, including seed yield per plant, has been reported earlier by (Hoque *et al*., 2021), (Singh *et al*., 2019).

The mean comparisons for seed vigour index among genotypes revealed that certain genotypes significantly outperformed the overall mean, highlighting their potential as promising candidates for yield improvement in breeding programs *viz.;* Ram Lakshman, IET-22020, SHIVANTH, DDR-119 (**Table 3**).

The CV had a mean value of 4.11 with a range of 0.06 to 10.95. According to **Gomez** and **Gomez** (1984), the CV is still a good indicator of how accurately the behaviors are related to each other and of the consistency of the research. The average CD value was 15.6 with a range of 0.02 to 105.07. The top four performing genotypes also exhibited desirable expressions for key seed quality traits relevant to selection. Notably, the genotype RAM LAKSHMAN emerged as a superior performer, characterized by high speed of germination, seeding length and seed vigour index (**Table 3**).

**Fig 1** representation of mean performance of all seed quality parameters of rice.

**Table 3** Top 5 performers for seed vigour index along with the supporting traits.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Genotypes** | **SOG (%)** | **SG (%)** | **SL (cm)** | **RL (cm)** | **SDL (cm)** | **SDW (g)** | **SVI** |
| Ram Lakshman | 107.96 | 90.86 | 14.14 | 11.4 | 25.75 | 0.1 | 2339.98 |
| IET-22020 | 87.99 | 93.13 | 14.42 | 9.39 | 23.82 | 0.11 | 2257.95 |
| Shivanth | 84.02 | 96.49 | 13.69 | 9.9 | 23.52 | 0.09 | 2240.14 |
| DDR-119 | 79.96 | 94.45 | 13.72 | 8.89 | 22.61 | 0.09 | 2131.79 |
| **Overall Mean** | 102 | 92.92 | 11.35 | 9.1 | 20.35 | 0.12 | 1902.06 |
| **Maximum** | 123.95 | 97.6 | 14.42 | 11.4 | 25.75 | 0.17 | 2339.98 |
| **Minimum** | 79.96 | 88.01 | 8.46 | 6.97 | 16.37 | 0.06 | 1506.3 |
| **SE(d) ±** | 0.03 | 0.59 | 0.31 | 0.29 | 0.35 | 0.01 | 37.39 |
| **CD at 5%** | 0.09 | 1.65 | 0.88 | 0.81 | 0.99 | 0.02 | 105.07 |
| **CV %** | 0.06 | 1.09 | 4.78 | 5.51 | 3.01 | 10.95 | 3.4 |

**3.3 Determination of Genetic variability**

**3.3.1 Genotypic and phenotypic coefficient of variations**

The PCV values were relatively higher than the GCV values, with only small differences, indicating that the environmental influence on the expression of the traits was minimal (Sudeepthi *et al*., 2020**),** (Hoque *et al*., 2021). The highest value ofGCV and PCV was exhibited by SDW (20.46, 22.59) that is (> 20%) and traits including SOG (12.87), SL (10.71, 11.73), RL (12.01, 13.22) had moderate values of GCV and PCV. Standard germination (2.5, 2.73), SDL (8.58, 9.09), SVI (8.59, 9.24) possessed low (<10%) values of PCV as well as GCV. The results of the presented study in respect of genotypic and phenotypic coefficient of variation are broadly in the findings of earlier workers (Sadhana *et* *al*., 2022), (Hoque *et* *al*., 2022), (Islam *et* *al*., 2020), (Nayak *et* *al*., 2016) (**Table 4**).

**3.3.2** **Heritability and genetic advance**

The highest estimate of heritability in broad sense was recorded for SOG (99.34%), SG (84.67%), SL (83.00%), RL (82.60%), SDL (89.00%) and SVI (86.40%), and moderate estimate of heritability (>60% to <80%) was observed for remaining trait *viz.,* SDW (71.30%). This finding was in accordance with (Akshitha *et* *al*., 2020), (Reddy *et al*., 2022) for high heritability in shoot length, root length, seedling vigour index I, seedling length **(Table 4)**.

The trait SVI had high heritability estimates along with the high genetic advance, is more helpful in predicting gain under selection than heritability estimates alone, similar with (Kumar *et al*., 2024). The estimates of genetic advance can help to understand the type of gene action of various polygenic characters. Further highest value of genetic advance in per cent of mean noticed in SDW, which was similar with the findings of (Singh *et al*., 2019). Moderate genetic advance was found for speed of germination, root length, seedling length and seedling vigour. This finding was in accordance with (Kumar *et* *al*., 2024).

**Table 4** Genetic variability parameters for all for seed quality parameters rice.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Traits** | **GCV** | **PCV** | **ECV** | **h2 (broad sense)** | **Genetic advancement 5%** | **Gen. Adv as % of Mean 5%** |
| **Speed of Germination (%)** | 12.87 | 12.87 | 0.05 | 99.34 | 27.04 | 26.51 |
| **Standard Germination (%)** | 2.5 | 2.73 | 1.09 | 84.67 | 4.4 | 4.73 |
| **Shoot Length (cm)** | 10.71 | 11.73 | 4.78 | 83 | 2.28 | 20.15 |
| **Root Length (cm)** | 12.01 | 13.22 | 5.51 | 82.6 | 2.04 | 22.5 |
| **Seedling Length (cm)** | 8.58 | 9.09 | 3,00 | 89 | 3.39 | 16.68 |
| **Seedling Dry Weight (g)** | 20.46 | 22.59 | 10.94 | 71.3 | 0.03 | 33.22 |
| **Seedling Vigour Index** | 8.59 | 9.24 | 3.4 | 86.4 | 33.29 | 16.47 |

**3.4 Correlation coefficient analysis**

According to genotypic and phenotypic correlation SG, SL, RL and SDL was significant and positively correlated with SVI. This finding was in accordance with (Jan *et* *al*., 2019), (Sofiya *et al*., 2020).Positive correlated outcome revealed that increase in one trait will lead to increase in the correlated trait(Kampe *et al*., 2018)**.** Moreover, SL was also highly significant and positively correlated with SDL and RL was positivelysignificant with SDL and SDW. Similarly, SDL was highly significant and positively correlated with RL and SL, SDW was with RL. Similar kinds of associations were reported by **(**Sofiya *et al*., 2020) (**Table 5**).

**Table 5** Correlation coefficient analysis at genotypic and phenotypic levels among all seed quality parameters of rice

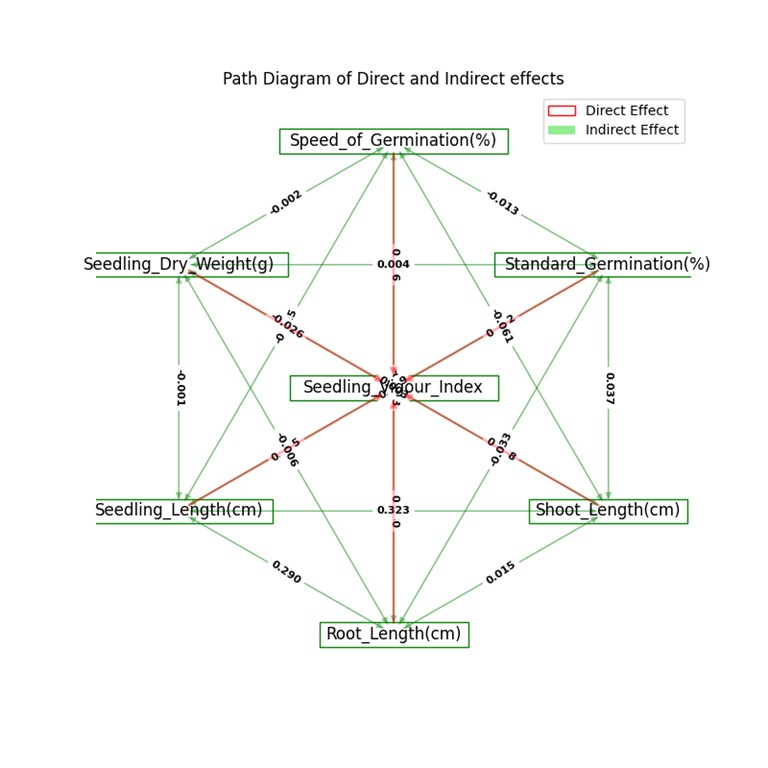
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Characters** | **Level** | **SOG** | **SG** | **SL** | **RL** | **SDL** | **SDW** | **SVI** |
| **SOG** | **rg** | 1 \*\* | -0.063 | -0.247 | -0.1307 | -0.2721 | 0.0935 | -0.2244 |
| **rp** | 1 \*\* | -0.0574 | -0.2255 \*\* | -0.1189 | -0.2569 \*\* | 0.0812 | -0.2083 \* |
| **SG** | **rg** |  | 1 \*\* | 0.121 | -0.1045 | 0.0526 | -0.1914 | 0.3083 \* |
| **rp** |  | 1 \*\* | 0.135 | -0.1108 | 0.0297 | -0.1598 | 0.2522 \*\* |
| **SL** | **rg** |  |  | 1 \*\* | 0.0488 | 0.7768 \*\* | -0.1364 | 0.7569 \*\* |
| **rp** |  |  | 1 \*\* | 0.0513 | 0.6642 \*\* | -0.1109 | 0.637 \*\* |
| **RL** | **rg** |  |  |  | 1 \*\* | 0.6754 \*\* | 0.273 | 0.6328 \*\* |
| **rp** |  |  |  | 1 \*\* | 0.5984 \*\* | 0.2471 \*\* | 0.5492 \*\* |
| **SDL** | **rg** |  |  |  |  | 1 \*\* | 0.0459 | 0.9565 \*\* |
| **rp** |  |  |  |  | 1 \*\* | 0.0364 | 0.8313 \*\* |
| **SDW** | **rg** |  |  |  |  |  | 1 \*\* | 0.0049 |
| **rp** |  |  |  |  |  | 1\*\* | -0.0028 |
| **SVI** | **rg** |  |  |  |  |  |  | 1 \*\* |
| **rp** |  |  |  |  |  |  | 1 \*\* |

**3.5 Direct and Indirect effects**

The path coefficient analysis results revealed maximum direct effect on SVI was shown by SDL followed by SG and SL and RL had negative direct effect on SVI. Similar results were also reported by (Jan *et al.*, 2019), (Katiyar *et al*., 2019), **(**Sofiya *et al*., 2020). The highest positive indirect effect on SVI via were exhibited by SL followed by RL and SVI Showed negative indirect effect on SVI via SL, followed by SOG and SG. This result was in accordance with **(**Sofiya *et al*., 2020), (Kumar *et al*., 2024), (Salman *et al*., 2022). The direct and indirect effect is represented in (**Fig 2**) (**Table 6**).

**Table 6** Phenotypic and genotypic direct and indirect effects of component traits on seed vigour index

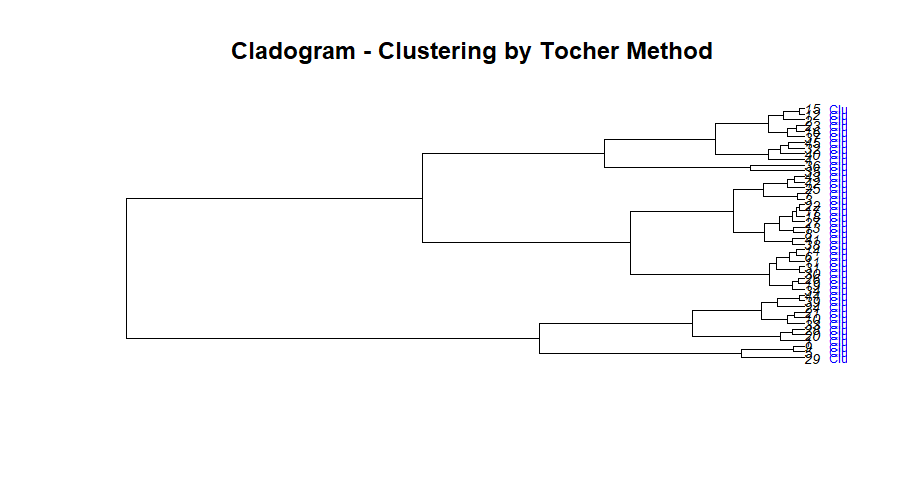
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Characters** | **Level** | **SOG** | **SG** | **SL** | **RL** | **SDL** | **SDW** | **SVI** |
| **SOG** | **rg** | **0.055** | -0.015 | 0.143 | 0.063 | -0.471 | 0.002 | -0.224 |
| **rp** | **0.026** | -0.013 | -0.062 | -0.033 | -0.123 | -0.002 | -0.208 |
| **SG** | **rg** | -0.003 | **0.244** | -0.070 | 0.050 | 0.091 | -0.004 | 0.308 |
| **rp** | -0.001 | **0.23** | 0.037 | -0.031 | 0.014 | 0.003 | 0.252 |
| **SL** | **rg** | -0.013 | 0.029 | **-0.578** | -0.024 | 1.345 | -0.003 | 0.757 |
| **rp** | -0.006 | 0.031 | **0.276** | 0.014 | 0.319 | 0.002 | 0.637 |
| **RL** | **rg** | -0.007 | -0.025 | -0.028 | **-0.481** | 1.170 | 0.005 | 0.633 |
| **rp** | -0.003 | -0.026 | 0.014 | **0.282** | 0.287 | -0.005 | 0.549 |
| **SDL** | **rg** | -0.015 | 0.013 | -0.449 | -0.325 | **1.732** | 0.001 | 0.956 |
| **rp** | -0.007 | 0.007 | 0.183 | 0.169 | **0.48** | -0.001 | 0.831 |
| **SDW** | **rg** | 0.005 | -0.046 | 0.079 | -0.131 | 0.079 | **0.019** | 0.005 |
| **rp** | 0.002 | -0.034 | -0.031 | 0.071 | 0.02 | **-0.03** | -0.003 |



**Fig 2** Graphical representation of direct and indirect for all characters in rice

**3.6 Cluster analysis**

In this study, the genotypes were classified into eight distinct, non-overlapping clusters, indicating substantial genetic variability (**Fig 3**). The highest number of germplasms was presented in Cluster III i.e. (14) genotypes, are found to be more diverse in the same cluster. followed by cluster VI (7), Cluster II (6), cluster VII (6), cluster I, IV, V, VIII (1) entries, respectively. The cluster I, IV, V, VI contains 1 entries which was very least number of entries. Similar findings were also recorded by (Farheen *et al*., 2022), (Dinesh *et al*., 2023), (Mishra *et al*., 2023; Sadhukhan et al., 2024).



**Fig 3** Graphical representation of cluster analysis by tocher method.

**3.6.1** **Intra-Inter cluster distances among eight clusters**

The estimates of intra cluster distance for 8 clusters is shown in (**Table 7)**. The maximum intra- cluster distance observed in case of cluster VIII (47.47) followed by cluster V (46.99), cluster III (42.39), cluster IV (33.22) and while the minimum intra- cluster distance showed by cluster I (21.11).

The maximum inter cluster was found between cluster V and VIII (726.48), indicating that hybridization of genotypes of these clusters would yield required segregants with the accumulation of favourable genes in segregating generations (Singh *et al*., 2021b), followed by cluster IV and V (575.12) and cluster I and VIII (569.69) exhibited very high inter cluster distance from reaming 8 cluster. The cluster VII and VIII (471.68), cluster III and V (429.26) II and VI (88.08), cluster I and IV (418.51) also exhibited high inter cluster distances. To realize high heterotic effect, (Mukesh *et al*., 2018) recommended that parents should be selected from clusters having wider inter-cluster distance.

**Table 7** Estimates of average inter and intra cluster distances for eight clusters in Rice.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cluster** | **I** | **II** | **III** | **IV** | **V** | **VI** | **VII** | **VIII** |
| **I** | 21.11 | 348.26 | 272.82 | 418.51 | 158.53 | 174.65 | 101.60 | 569.69 |
| **II** | 348.26 | 28.41 | 81.34 | 77.53 | 504.88 | 175.67 | 251.52 | 223.35 |
| **III** | 272.82 | 81.34 | 42.39 | 148.96 | 429.26 | 101.64 | 176.08 | 298.43 |
| **IV** | 418.51 | 77.53 | 148.96 | 33.22 | 575.12 | 245.71 | 320.82 | 153.86 |
| **V** | 158.53 | 504.88 | 429.26 | 575.12 | 46.99 | 330.92 | 256.37 | 726.48 |
| **VI** | 174.65 | 175.67 | 101.64 | 245.71 | 330.92 | 27.32 | 79.45 | 396.36 |
| **VII** | 101.60 | 251.52 | 176.08 | 320.82 | 256.37 | 79.45 | 32.95 | 471.68 |
| **VIII** | 569.69 | 223.35 | 298.43 | 153.86 | 726.48 | 396.36 | 471.68 | 47.47 |

**3.6.2 Estimation of mean values of seven clusters**

Cluster analysis indicated significant difference across traits is represented in (**Table 8**). Cluster VI had the highest mean for germination speed (114.50), while cluster VIII had the lowest (93.32). For germination standard, cluster II had the highest (95.02), and cluster I the lowest (91.65). Cluster VIII had the highest mean shoot length (14.08) and seedling length (24.37), with the lowest shoot length being in cluster VI (10.17). Root length was highest in cluster II (10.32) and lowest in cluster VI (7.33). Dry weight in seedlings was highest in cluster VIII (2279.36) and lowest in cluster VI (1553.74), reflecting considerable variation among clusters.

**Table 8** Mean value of Eight clusters by torcher’s method for all genotypes of rice.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Cluster** | **Speed of Germination (%)** | **Standard Germination (%)** | **Shoot Length (cm)** | **Root Length (cm)** | **Seedling Length (cm)** | **Seedling Dry Weight (g)** | **Seedling Vigour Index** |
| **I** | 101.55 | 91.65 | 11.16 | 8.45 | 19.66 | 0.106 | 1820.92 |
| **II** | 95.34 | 95.02 | 11.86 | 10.32 | 22.19 | 0.125 | 2122.51 |
| **III** | 100.36 | 92.79 | 11.54 | 9.71 | 20.85 | 0.14 | 1949.3 |
| **IV** | 103.89 | 92.11 | 10.29 | 8.65 | 18.64 | 0.115 | 1734.25 |
| **V** | 107.5 | 94.43 | 11.85 | 9.48 | 21.32 | 0.11 | 2024.98 |
| **VI** | 114.5 | 92.12 | 10.17 | 7.33 | 17.4 | 0.113 | 1553.74 |
| **VII** | 100.35 | 93.5 | 11.11 | 8.89 | 20.11 | 0.12 | 1878.84 |
| **VIII** | 93.32 | 93.5 | 14.08 | 10.23 | 24.37 | 0.098 | 2279.36 |

**3.6.3 Percent contribution of traits**

Percent contribution in genetic divergence in rice is presented in (**Fig 4**)**.** In all combinations, each character was ranked based on their contribution towards divergence between two entries, the result was in accordance with (Pallavi *et al*., 2023).The highest percent contribution in genetic divergence of rice was reported by speed of germination (15.42%) followed by shoot length (cm) (14.13%), seedling length (13.08%), Seedling Vigour Index (12.87%), Root length (cm) (11.88%), seedling dry weight (10.1%) and standard germination (7.1%).

**Fig 4** Percent contribution of traits towards the divergence.

**3.7 Principal component analysis (PCA)**

The first five principal components PC1, PC2, PC3, PC4, PC5 explain 80.116% of the total variance, making them the most important for distinguishing genotypes. The extent to which each principal component contributed to the total variation was directly related to its eigenvalue, with the contribution decreasing gradually across components. The PC1 accounted for 43.10 % of the total variation, while the PC2, PC3, PC4 and PC5 explained 20.50 %, 13.30 %, 0.118% and 0.098% respectively. These findings are consistent with those reported by (Talekar *et al*., 2022,) (Sheela *et al*., 2020), (Chandraker *et al*., 2024) and (Kumari *et al*., 2023) in their respective studies (**Table 8**).

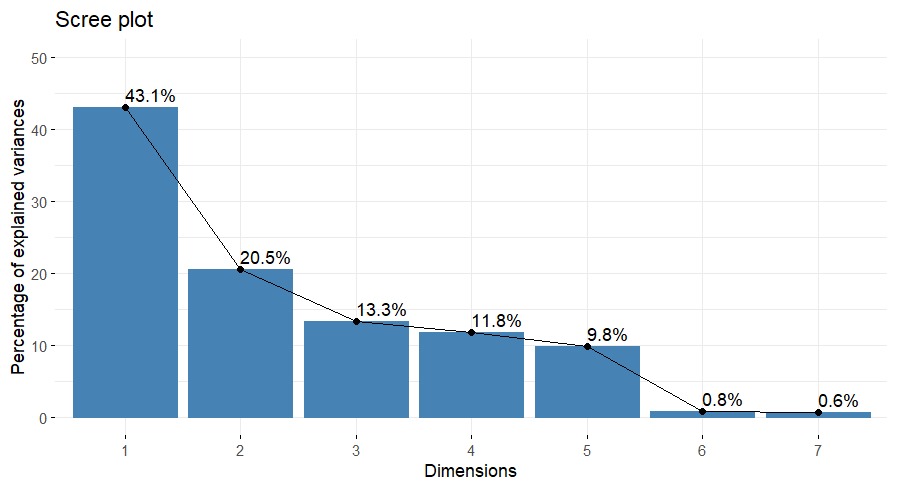
Factor loadings revealed that PC1 showed high and significantly positive loading for SOG (0.211) while PC2 enabled high and significantly positive loading for SG (0.498), SL (0.284) and SVI (0.033) and PC3 having high and significant positive loading for SL (0.072) and SDL (0.011) indicating that these three components can be designated as quality parameters attributing factors and thus, can be used to differentiate rest of the genotypes on the basis of yield and its attributing traits (**Table 9**).

Scree plot explained the percentage of variation associated with each principal component and is obtained by drawing a graph between principal component numbers (X-axis) and percentage of variation explained (Y-axis) (**Fig 5**). The PC1 showed 43.10 per cent variability with eigen value 2 which then declined gradually. From scree plot, it is clear that the maximum variation was observed in PC1.

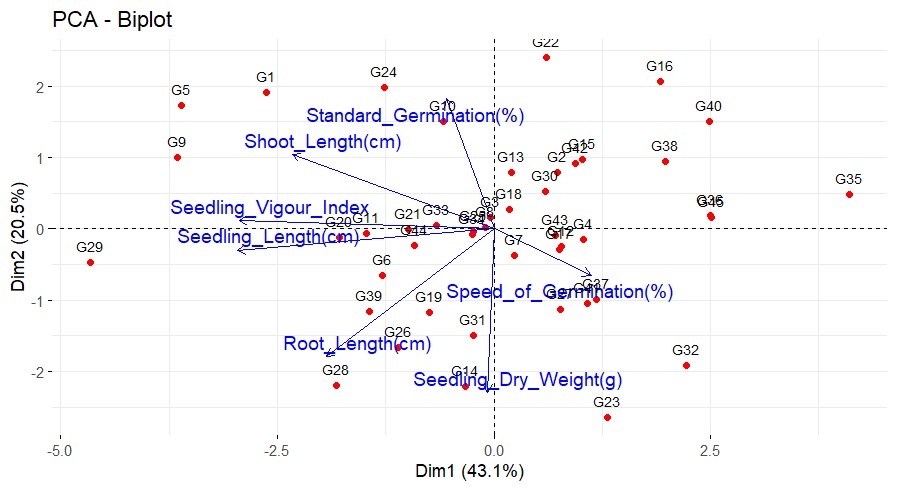
The first two principal component biplot (**Fig 6**) shows the interaction between genotypes and traits. The extent of each vector for a trait indicates its contribution to genetic divergence in total longer vectors contribute more. The longest vectors were those of traits seedling vigour index and seedling length, showing their high contribution to overall diversity, followed by harvest index and seed yield per plant. The PCA biplot identifies genotype clusters with distinct trait advantages. Genotypes G23, G32, and G17 show superior seedling dry weight and germination speed, while G1, G5, and G9 exhibit strong seedling vigor and shoot length. G14, G28, and G26 are linked with better root length. These patterns help in selecting high-performing genotypes and genetically diverse parents, supporting effective breeding strategies for improved seedling traits in rice.

**Table 9** Standard deviation, proportion of variance, cumulative proportion for principal components.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **PC’s** | **PC1** | **PC2** | **PC3** | **PC4** | **PC5** |
| **Eigen value** | 3.016 | 1.436 | 0.933 | 0.826 | 0.689 |
| **Standard deviation** | 1.737 | 1.198 | 0.966 | 0.909 | 0.830 |
| **Proportion of Variance** | 43.100 | 23.500 | 13.300 | 0.118 | 0.098 |
| **Cumulative Proportion** | 0.431 | 0.636 | 0.769 | 0.887 | 0.986 |
| **Speed of Germination (%)** | 0.211 | -0.178 | -0.753 | 0.583 | 0.125 |
| **Standard Germination (%)** | -0.105 | 0.498 | -0.594 | -0.578 | -0.07 |
| **Shoot Length (cm)** | -0.439 | 0.284 | 0.072 | 0.419 | -0.446 |
| **Root Length (cm)** | -0.365 | -0.490 | -0.069 | -0.235 | 0.526 |
| **Seedling Length (cm)** | -0.558 | -0.081 | 0.011 | 0.148 | 0.039 |
| **Seedling Dry Weight** | -0.014 | -0.626 | -0.183 | -0.269 | -0.708 |
| **Seedling Vigour Index** | -0.554 | 0.033 | -0.193 | 0.011 | 0.046 |



**Fig 5** Scree plot explained the percentage of variation associated with each principal component and is obtained by drawing a graph between principal component numbers (X-axis) and percentage of variation explained (Y-axis).



**Fig 6** PCA biplot graph of different traits for first two principal components

**4. CONCLUSION**

Following the results of this study, enormous genetic variation among the 45 rice genotypes for all the seven seed quality traits was found to show enormous possibilities of selection and genetic improvement. Genotypes Ram Lakshman, IET-22020, SHIVANTH, DDR-119 had better mean seed vigou index, while Ashoka 200, Ruchi Dhan, PR 131 and HUR-36 showed the maximum germination speed, revealed as early types of maturing and hence the promising lines of choice for any breeding program. High heritability was found for all the traits and the trait SVI had high heritability estimates along with the high genetic advance, is more helpful in predicting gain under. Correlations of SG, SL, RL and SDL was significant and positively correlated with SVI and path coefficient analysis results revealed maximum direct effect on SVI was shown by SDL, SG and SL, confirm their significance in improving yield. Eight distinct, non-overlapping clusters were obtained with cluster III having highest number of diverse germplasms. The maximum inter cluster was found between cluster V and VIII indicating that hybridization of genotypes of these clusters would yield required segregants with the accumulation of favourable genes in segregating generations. PCA revealed that five principal components accounted for 80.116% of the total variability.

**DISCLAIMER (ARTIFICIAL INTELLIGENCE)**

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

**REFERENCES**

Akshitha, B., Senguttuvel, P., Latha, V. H., Yamini, K. N., Rani, K. J. and Beulah, P. (2020). Variability and correlation analysis for seedling vigour traits in rice (*Oryza sativa* L.) genotypes. *Int. J. Curr. Microbiol. App. Sci*., 9(7): 2877–2887.

Chandraker, P., Sharma, B., Parikh, M., & Saxena, R. R. (2024). Assessment of Genetic Diversity in Aromatic Short Grain Rice (*Oryza sativa* L.) Genotypes using PCA and Cluster Analysis. *International Journal of Plant & Soil Science*, 36(5), 82–94. <https://doi.org/10.9734/ijpss/2024/v36i54504>

Christina, G. R., Thirumurugan, T., Jeyaprakash, P., & Rajanbabu, V. (2021). Principal component analysis of yield and yield related traits in rice (*Oryza sativa* L.) landraces. *Electronic Journal of Plant Breeding*, 12(3), 907-911.

Dewey, D.R. and Lu K.H. (1959). A correlation and path coefficient analysis of components of crested wheat grassseed production. *Agron. J*., 51, 511-518.

Dinesh, K., Devi, M. S., Sreelakshmi, C. and Paramasiva, I. (2023). Exploring the genetic diversity for yield and quality traits in indigenous landraces of rice (*Oryza sativa* L.). *Electron. J. Plant Breed*., 14(2): 502–510.

FAO. (2023). FAOSTAT database. Food and Agriculture Organization of the United Nations. <http://www.fao.org/faostat>

Farheen, M., Murthy, K. G. K., Mohan, Y. C., & Kumar, J. H. (2022). Genetic Diversity Analysis for Yield and Morpho-physiological Traits in Elite Rice (*Oryza sativa* L.) Genotypes under Semi-dry System Through D2 Statistics. *International Journal of Environment and Climate Change,* 827–835. <https://doi.org/10.9734/ijecc/2022/v12i1131045>

Gomez K.A., and A.A. Gomez. Statistical procedures for agricultural research. 2nd ed. John Wiley and Sons. 1984; New York, NY

Hoque, N., Islam, M. Z., Zarin, F., Mahmud, N., Rahman, M. and Biswas, B. (2021). Analysis of genetic variability and character relationship in rice (*Oryza sativa* L.) seed and seedling traits. *J. Biosci. Agric. Res*., 28(2): 2389–2398.

Hoque, N., Islam, M. Z., Zarin, F., Mahmud, N., Rahman, M., & Biswas, B. (2021). Analysis of genetic variability and character relationship in rice (*Oryza sativa* L.) seed and seedling traits. *Journal of Bioscience and Agriculture Research*, 28(02), 2389-2398.

Islam, M., Akter, N., Prince, M. D., Khan, M. and Khalequzzaman, M. (2020). Genetic variability in genotypes of Khagrachhari district in Bangladesh. *The Agriculturists*, 18(1): 1–9.

Jan, N. and Kashyap, C. S. (2019). Correlation and path analysis in rice (*Oryza sativa* L.) for seed and seed vigour traits. *J. Pharmacogn. Phytochem*., 8(1): 222–226.

K.R.V. Sathya Sheela, S. Robin, and S. Manonmani. (2020). Principal component analysis for grain quality characters in rice germplasm. *Electronic Journal of Plant Breeding*, 11 01, 127-131.

Kampe, A. K., Tassew, A. A. and Gezmu, A. T. (2018). Estimation of phenotypic and genotypic correlation and path coefficients in rainfed upland rice (*Oryza sativa* L.) genotypes at Guraferda, Southwest Ethiopia. *J. Rice Res*., 6: 195.

Katiyar, D., Srivastava, K. K., Prakash, S., Kumar, M., & Gupta, M. (2019). Study correlation coefficients and path analysis for yield and its component characters in rice (*Oryza sativa* L.). *Journal of Pharmacognosy and Phytochemistry*, 8(1), 1783-1787.

Kishore, N. S., Srinivas, T., Nagabhushanam, U., Pallavi, M., & Sameera, S. K. (2015). Genetic variability, correlation and path analysis for yield and yield components in promising rice (*Oryza sativa* L.) genotypes. *SAARC Journal of Agriculture*, 13(1), 99-108.

Kumar, K. Y., Singh, V., Singh, S. K., & George, S. G. (2022). Assessment of Rice (*Oryza sativa* L.) Hybrids on Growth and Yield under Agro-climatic Conditions of Prayagraj, U. P. International Journal of Plant & Soil Science, 34(18), 45–49. <https://doi.org/10.9734/ijpss/2022/v34i1831051>

Kumar, S.; Vimal, S.C.; Singh, J.; Gupta, R.; Yaduvanshi, N. and Singh, Y.P. (2024). Path coefficient analysis of rice (*Oryza sativa* L.) genotype under laboratory conditions. *Int. J. Plant & Soil Sci*., 36 (1): 155–162. <https://doi.org/10.9734/ijpss/2024/v36i14344>

Kumari, S., Chakrabarty, S. K., & Paul, D. (2023). Assessment of genetic diversity for seed vigor traits in rice (*Oryza sativa* L.) through principal component and cluster analysis. *Indian Journal of Genetics and Plant Breeding*, 83(03),343–354. <https://doi.org/10.31742/isgpb.83.3.6>

Mahendran, R., Vanitha J., Jegadeeswaran M., Mathivathana M.K. and Ashraf A.M. (2024). Association analysis in rice germplasm lines for high temperature tolerance. *Agricult. Sci. Digest - a Res. J*., <https://doi.org/10.18805/ag.d-5898>

Mishra, A., Shahani, M., & Bijarania, S. (2023). Genetic divergence Studies of submergence tolerant rice genotypes. *International Journal of Environment and Climate Change*, 13(3), 195–199. <https://doi.org/10.9734/ijecc/2023/v13i31697>

Mukesh, M., Vidyabhushan, J., Anand, K., Mankesh, K., & Shweta, K. (2018). Genetic divergence in mid early rice genotypes. *Journal of Pharmacognosy and Phytochemistry*, 7, 274–276.

Nath, S. and Kole, P. C. (2021). Genetic variability and yield analysis in rice. *Electron. J. Plant Breed*., 12(1), 253-258.

Nayak, R., Singh, V. K., Singh, A. K. and Singh, P. K. (2016). Genetic variability, character association and path analysis of rice genotypes. *Ann. Plant Soil Res*., 18(2): 161–164.

Pallavi, M., Shanthi, P., Reddy, V. L. N., Kumar, A. N., & Prasad, B. M. (2023). Divergence Studies for Early Seedling Vigour Related Traits, Yield and Yield Components in Rice (*Oryza sativa* L.). *Int. J. Plant Soil Sci*, 35(22), 525-537.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11), 559572.

Ranjith, P.; Sahu, S.; Dash, S. K.; Bastia, D. N. and Pradhan, B. D. (2018). Genetic diversity studies in rice (*Oryza sativa* L.). *Journal of Pharmacognosy and Phytochemistry*, 7(2), 2529-2531.

Reddy, Y. M., Vasanthi, R. P., Reddy, B. R. K., Murthy, B. R. and Vemireddy, L. R. (2022). Genetic parameters for quantitative and early seedling vigour related traits in rice (*Oryza sativa* L.) under dry direct seeded rice conditions. Andhra Pradesh *J. Agril. Sci*., 8(3): 170–173.

Robinson, H.F., Comstock R.E. and Johnson H.W. (1955). Estimates of heritability and the degree of dominance in corn. *Agron. J*., 47(7), 314–318.

Sadhana, P., Raju, C. D., Bhargava, K., Rao, L. S. and Kuna, A. (2022). Studies on variability, correlation and path analysis for seedling vigour traits in rice (*Oryza sativa* L.). *Int. J. Environ. Clim. Change*, 12: 121–130.

Sadhukhan, D., Mukherjee, T., Sarkar, A., Devi, N. D., Bisarya, D., Kumar, V., & Jincy, M. (2024). A Comprehensive Analysis of Drought Stress Responses in Rice (*Oryza sativa* L.): Insights into Developmental Stage Variations from Germination to Grain Filling. International Journal of Environment and Climate Change, 14(7), 141–158. <https://doi.org/10.9734/ijecc/2024/v14i74260>

Salman, M.; Vimal, S.C.; Tarkeshwar, G.M.; Jangid, D. and Meena, M.B. (2022). Correlation and path coefficient analysis in seed and seedling characters for yield in popular rice varieties (*Oryza sativa* L.). *Pharma Innovation*, 11 (7): 534–541.

Singh, B., Gauraha, D., Sao, A., & Nair, S. K. (2021). Assessment of genetic variability, heritability and genetic advance for yield and quality traits in advanced breeding lines of rice (*Oryza sativa* L.). *The Pharma Innovation*, 10(8), 1627-1630.

Singh, P., Singh, S., Korada, M., Khaire, A., Singh, D., Sonali, V., Habde, P., Kumar, M., & Naik, R. (2021b). Exploring variability and genetic diversity among rice genotypes in Eastern Uttar Pradesh. *Electronic Journal of Plant Breeding*, 12(4). <https://doi.org/10.37992/2021.1204.187>

Singh, R. K., Yadav, H. C., Kumar, M., Lal, K., & Amir, M. (2019). Genetic variability, heritability and genetic advance analysis for seed yield and its physiological quality parameters in rice (*Oryza sativa* L.). *Journal of. Pharmacognosy and Phytochemistry*, 8, 511-513.

Singh, T.V.J.; Mohan, Y.C.; Damodar Raju, C.H.; Jagadeshwar, R. and Balram, M. (2019). Character association and path coefficient studies in rice (*Oryza sativa* L.). *Int*. *J. Pharmacogn*., 8 (3): 383–387.

Sofiya, M., Eswaran, R. and Silambarasan, V. (2020). Correlation and path coefficient analysis in rice (*Oryza sativa* L.) genotypes under normal and cold condition. *Indian J.* *Agric. Res*., 54(2): 237–241.

Sudeepthi, K., Srinivas, T. V. S. R., Kumar, B. R., Jyothula, D. P. B. and Umar, S. N. (2020). Assessment of genetic variability, character association and path analysis for yield and yield component traits in rice (*Oryza sativa* L.). *Electron. J. Plant Breed*., 11(1): 144–148.

Talekar, S.C.; Praveena, M.V. and Satish, R.G. (2022). Genetic diversity using principal component analysis and hierarchical cluster analysis in rice. *Int. J. Plant Sci*., 17 (2): 191–196. <https://doi.org/10.15740/has/ijps/17.2/191-196>

Tiwari, D. N.; Tripathi, S. R.; Tripathi, M. P.; Khatri, N. and Bastola, B. R. (2019). Genetic variability and correlation coefficients of major traits in early maturing rice under rainfed lowland environments of Nepal. *Advan. Agric*., 23(2), 1-9.

Yugandhar RP, Kota S, Kiran UB, Shridhar M. Principal component analysis for agro-morphological and quality characters in germplasm of rice (*Oryza sativa* L.). *Int. J Adv. Biol. Res* 2018;8(2):268-273.