Evaluation of Rice Genetic Diversity and Variability in a Population Panel by Principal Component Analysis

Vishnu Varthini Nachimuthu¹, S. Robin¹, D. Sudhakar², M. Raveendran², S. Rajeswari¹ and S. Manonmani¹

¹Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India; popvarun@gmail.com, robin.tnau@gmail.com, rajisundar93@gmail.com, swamimano@yahoo.co.in ²Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India; sudhakard@hotmail.com, raveendrantnau@gmail.com

Abstract

A population panel of 192 rice genotypes comprising traditional landraces and exotic genotypes from 12 countries was evaluated for 12 agro - morphological traits by principal component analysis for determining the pattern of genetic diversity and relationship among individuals. Twelve quantitative characters i.e. plant height, leaf length, number of productive tillers, panicle length, number of filled grains, spikelet fertility, days to 50% flowering; days to harvest maturity, grain length, grain width, grain length width ratio, and single plant yield were measured. The largest variation was observed for number of productive tillers with Coefficient of Variation (CV) of 28.03% followed by number of filled grains per panicle, single plant yield, leaf length, grain length width ratio. Days to maturity has shown the least variation with the CV of 9.74%. Principal component analysis was utilized to examine the variation and to estimate the relative contribution of various traits for total variability. In the current study, Component 1 had the contribution from the traits such as days to 50% flowering, leaf length, plant height, panicle length, days to maturity and number of filled grains which accounted 28.46% of the total variability. Grain width and grain length width ratio has contributed 16.8% of total variability in component 2. The remaining variability of 14.4%, 11.7% and 9.3% was consolidated in component 3, component 4 and component 5 by various traits such as spikelet fertility, single plant yield, grain length and number of productive tillers. The cumulative variance of 80.56% of total variation among 12 characters was explained by the first five axes. Thus the results of principal component analysis used in the study have revealed the high level of genetic variation and the traits contributing for the variation was identified. Hence this population panel can be utilized for trait improvement in breeding programs for the traits contributing for major variation.

Keywords: Genetic Variation, Principal Component Analysis, Rice

1. Introduction

Genetic diversity represents the heritable variation within and between populations of organisms. The success of plant breeding depends on the availability of genetic variation, knowledge about desired traits, and efficient selection strategies that make it possible to exploit existing genetic resource. The pool of genetic variation within an inter-mating population is the basis for selection as well as for plant improvement. Rice is a major agriculture crop in India grown under diverse ecological conditions¹, with varied phenology and yield². History shows that technological advancement and its applications in agricultural crop plants brought great respite at times when burgeoning population desperately needed security in food fronts³.

Before exploiting a population for trait improvement, it is necessary to understand the magnitude of variability in the population which is fundamental for genetic improvement in all crop species. To develop segregating

*Author for correspondence

population, genetic distance estimates form the basis for selecting parental combinations with sufficient genetic diversity and for classifying germplasm into heterotic groups for hybrid crop breeding. Population Grouping can be based on geographical origin, agro-morphological traits, pedigree information, or molecular marker data⁴. Genetic distance estimates for population grouping can be estimated by different methods as it is crucial to understand the usable variability existing in the population panel. One of the approaches is to apply multivariate analysis.

Statistical method of classification is usually by multivariate methods as it has extensive use in summarizing and describing the inherent variation among crop genotypes. Multivariate statistical tools include Principal Component Analysis (PCA), Cluster analysis and discriminate analysis⁵. Principal Component Analysis (PCA) can be used to uncover similarities between variable and classify the cases (genotypes), while cluster analysis on the other hand is concerned with classifying previously unclassified materials⁶. Principal Component Analysis (PCA) is a multivariate technique that analyzes a data table in which observations are described by several inter-correlated quantitative dependent variables. Its goal is to extract the important information from the table, to represent it as a set of new orthogonal variables called principal components, and to display the pattern of similarity of the observations and of the variables as points in maps. 'Proper values' measure the importance and contribution of each component to total variance, whereas each coefficient of proper vectors indicates the degree of contribution of every original variable with which each principal component is associated. The higher the coefficients, regardless of the direction (positive or negative), the more effective they will be in discriminating between accessions.

The study is aimed to determine level of germplasm variation in *Oryza sativa* L. to identify and classify variation for grouping the accessions by taking into account several characteristics and relationship between them.

2. Materials and Methods

The germplasm collection consisting of 192 rice accessions was used in this study, which consist of land races and varieties collected from nine different states of India as well as from different countries. It has 82 landraces from different agro climatic zones of Tamil Nadu and two landraces from Orissa. 62 exotic genotypes from Argentina,

Bangladesh, Brazil, Bulgaria, China, Colombia, Indonesia, Philippines, Taiwan, Uruguay, Venezuela and United States and 46 varieties and improved genotypes from different states of India constitute the population panel of 192 genotypes. For easy identification and retrieval, each accession was named as RG1 to RG192. A set of 192 genotypes were grown in Paddy Breeding Station, Department of Rice, Tamil Nadu Agricultural University, India during Rabi 2013. This area is situated at latitude of 11°N and longitude of 77°E with clayey soil of pH 7.8. One hundred and ninety two genotypes were transplanted 21 days after sowing as two seedlings per hill in randomized complete block design with a spacing of 20 X 20 cm. Each plot per accession consisted of four rows each 0.8 by 3.6 m long at a distance of 40 cm between the plots. Normal cultural practices were followed as per standard recommendation. Twelve quantitative characters were measured according to methods in the descriptors for rice O. sativa7. Variables considered in the descriptive and multivariate analyses were morphological (plant height, leaf length, number of productive tillers, panicle length, number of filled grains, spikelet fertility), phenological (days to 50% flowering and days to harvest maturity from the day of seeding), and grain traits (grain length, grain width, grain length width ratio, and single plant yield). The observations recorded on 12 traits were statistically analyzed in SPSS16.0 to cluster the genotypes based on genetic similarity. The PCA analysis reduces the dimensions of a multivariate data to a few principal axes, generates an Eigen vector for each axis and produces component scores for the characters^{8,9}.

3. Results and Discussion

The First order Statistical measures *i.e.* maximum, minimum, sum, mean, Standard Deviation (SD) and Coefficient of Variation (CV) for the measured traits are presented in Table 1. The largest variation was observed for number of productive tillers with CV of 28.03% followed by number of filled grains per panicle (CV= 27), single plant yield (23.19), leaf length (23.02), grain length width ratio (22.16). Days to maturity has shown the least variation with the CV of 9.74%.

DFF- days to fifty percent flowering, LL- leaf length, PH- plant height, NPT- number of productive tillers, PLpanicle length, DM- days to maturity, NFG- number of filled grains per panicle, SF- spikelet fertility, GL- grain length, GW- grain width, GLWR- grain length width ratio, SPY – single plant yield

| Variable | Sum | Mean | SD | CV | Minimumm Value | Accessions | Maximum | Accessions |
|----------|----------|--------|-------|-------|----------------|------------|---------|------------|
| DFF | 16380.1 | 85.31 | 11.07 | 12.97 | 66.0 | RG69 | 123.0 | RG1 |
| LL (cm) | 7227.7 | 37.64 | 8.67 | 23.02 | 20.8 | RG104 | 62.4 | RG106 |
| PH (cm) | 22231.6 | 115.79 | 21.79 | 18.82 | 51.0 | RG111 | 162.3 | RG20 |
| NPT | 2547.95 | 13.27 | 3.72 | 28.03 | 5.0 | RG147 | 26.5 | RG183 |
| PL (cm) | 4237.2 | 22.07 | 2.74 | 12.42 | 11.6 | RG111 | 31.1 | RG85 |
| DM | 22556.5 | 117.48 | 11.44 | 9.74 | 94.0 | RG43, RG59 | 155.0 | RG1 |
| NFG | 13201.1 | 68.76 | 18.56 | 27.00 | 26.6 | RG189 | 112.0 | RG23 |
| SF (%) | 15907.25 | 82.85 | 6.93 | 8.37 | 54.2 | RG25 | 95.7 | RG131 |
| GL (mm) | 1592.5 | 8.29 | 0.89 | 10.68 | 5.8 | RG111 | 10.5 | RG57 |
| GW(mm) | 483.785 | 2.52 | 0.49 | 19.49 | 1.5 | RG95 | 3.7 | RG160 |
| GLWR | 661.4 | 3.44 | 0.76 | 22.16 | 1.8 | RG122 | 5.6 | RG47 |
| SPY (g) | 4832.88 | 25.17 | 5.84 | 23.19 | 12.0 | RG104 | 55.5 | RG164 |

Table 1. Characteristic means and variations of 192 accessions

The genotype RG1 (Mapillai Samba) has taken the longest days (123 days) for 50per cent flowering and RG 69 (RPHP48) (66 days) had short duration of 50per cent flowering. The overall mean for flowering for the germplasm accessions was 85 days. The skewness and kurtosis coefficients were 0.62 and 0.68 respectively with the variation of 12.97 per cent. Mean for leaf length of 192 germplasm lines were 37.64 cm with the minimum and maximum value of 20.8 cm in the genotype RG104 (Carolina Rinaldo Barsani) and 62.4 cm in the genotype RG106 (Katta samba) respectively. It also showed higher Coefficient of Variation of 23.02 per cent with positive skewness (0.543) and negative kurtosis (-0.258). The taller genotype was RG20 (Kalvalai) (162.3 cm) where as the genotype RG111 (Chigyungdo X IRGC 55466-1) (51 cm) had short stature. Mean height of 192 lines were 115.79 cm. The variation observed for this trait was 18.8 per cent with negative skewness and kurtosis (-0.332 and - 0.383).

The larger variation was observed for number of productive tillers with Coefficient of Variation of 28.03 per cent. The genotype RG183 (RPHP163) was recorded with maximum number of productive tillers per plant (26.5) and the genotype RG147 (CI11011) had registered less number of productive tillers (5). The mean was 13 productive tillers per plant. The skewness and kurtosis coefficients were 0.652 and 1.141 respectively which shows right skewed and platykurtic distribution. The trait panicle length had recorded the mean value of 22.07 cm with the coefficient of variation of 12.4 per cent. The genotype RG111 (Chigyungdo X IRGC 55466-1) recorded the lowest panicle length of 11.6cm and the genotype RG85 (RPHP104) had lengthier panicle of 31.1 cm. The skewness and kurtosis coefficients were -0.146 and 1.074 respectively. Days to maturity had observed with the least variation with the Coefficient of Variation of 9.74 per cent. The genotype RG1 (Mapillai Samba) was late in maturity (155 days) and RG43 (RPHP129) and RG59 (RPHP68) were earlier (94 days). The mean days to maturity for 192 germplasm lines were 117 days with the skewness and kurtosis coefficients of 0.312 and 0.184.

Number of filled grains per panicle has shown wide range of variation with the mean of 68.76. The genotype RG 23 (Koombalai) was recorded with the highest number of filled grains per panicle (112) and the genotype RG189 (Kaniranga) had the lowest (6.6). It also showed higher coefficient of variation of 27 per cent. The coefficient of skewness and kurtosis was -0.284 and -0.428. Spikelet fertility ranged from 54.2 per cent in the genotype RG25 (Sornakuruvai) to 95.7 per cent in the genotype RG131 (RPHP90). The mean spikelet fertility was observed as 82.85 per cent. This trait had exhibited lesser coefficient of variation (8.37 per cent). Spikelet fertility had negative skewness (-1.013) and platykurtic distribution (Kurtosis-1.664). Substantial variation was observed for grain length among 192 rice accessions. The accession with longest grain was RG57 (RPHP103) (10.5 mm) and shortest grain was measured in the genotype RG111 (Chigyungdo

X IRGC 55466-1) (5.8 mm). The mean grain length was recorded as 8.29 mm. It was registered with higher coefficient of 10.6 per cent. The coefficient of skewness and kurtosis was -0.255 and 0.354. The largest grain width was recorded in the genotype RG160 (TD25 X IRGC9146-1) (3.7 mm) which is a bold grain type and the slender grain type with lesser grain width was noticed in the genotype RG95 (Jeeraga samba) (1.5 mm). The mean grain width was 2.52mm. The skewness and kurtosis was 0.046 and -0.998. This trait had recorded Coefficient of Variation of 19.4 per cent.

Grain length Width ratio was highest for the genotype RG47 (Machakantha) (5.6) and lowest for the genotype RG122 (Kaliyan samba) (1.8). The mean ratio was 3.45. This trait had recorded enormous variation with Coefficient of Variation of 22.1. It had positive skewness (0.351) and negative kurtosis (-0.628).

The trait single plant yield had recorded the overall mean of 25.17 g per plant. The genotype RG164 (Thillainayagam) was recorded with the highest single plant yield of 55.5g and the genotype RG104 (Carolina Rinaldo Barsani) had the lowest single plant yield (12g). The coefficient of variation observed for this trait was higher (23.1%). This trait have exhibited positive skewness (0.906) and leptokurtic distribution (kurtosis = 3.07).

The genotype RG1 has taken the longest days for flowering as well as maturity. The taller genotype is RG20 where as RG111 has short stature. RG183 has more number of productive tillers but RG164 has higher single plant yield. Spikelet fertility ranges from 95.7% in RG131 to 54.2 in RG25. The accession with longest grain was RG57 (10.5) and largest grain width in RG160 (3.7) which is a bold grain type. The slim grain type with lesser grain width was RG95 (1.5) and shortest grain was RG111 (5.8). The accessions possessing extreme phenotype can be utilized in mapping population development for Quantitative trait loci identification. Coefficient of variation was high for all the traits measured. The study of distribution of quantitative traits using skewness and kurtosis provides information about nature of gene action¹⁰ and number of genes controlling the traits¹¹ respectively.

The skewed distribution of a trait in general suggests that the trait is under the control of non-additive gene action and is influenced by environmental variables. Positive skewness is associated with complementary gene interactions while negative skewness is associated with duplicate (additive x additive) gene interactions. The genes controlling the trait with skewed distribution tend to be predominantly dominant irrespective of whether they have increasing or decreasing effect on the trait¹². Frequency distribution for different traits on 192 germplasm accessions revealed different patterns of distribution as shown on Figure 1a-11.





Figure1a-11. Frequency distribution of different quantitative traits.

None of the traits had shown normal distribution. The traits such as days to fifty percent flowering, leaf length, number of productive tillers, days to maturity, grain width, grain length width ratio and single plant yield were positively skewed. Negative skewness was observed for plant height, panicle length, number of filled grains per panicle, spikelet fertility and grain length.

Kurtosis is negative or close to zero in the absence of gene interaction and is positive in the presence of gene interactions¹³. The traits with leptokurtic and platykurtic distribution are controlled by fewer and large number of genes, respectively. As none of the traits had kurtosis coefficient of 3 / mesokurtic, they were not in normal distribution. The traits such as days to fifty per cent flowering, number of productive tillers per plant, panicle length, days to maturity, spikelet fertility and grain length were Platykurtic. The trait single plant yield was leptokurtic.

Principal component analysis has shown the genetic diversity of the population panel. The cumulative variance of 80.56% by the first five axes with Eigen value of > 1.0 (Figure 2) indicates that the identified traits within the axes exhibited great influence on the phenotype of population panel (Table 2). A scatter plot was drawn between PC1 and PC2 depicted a clear pattern of grouping genotypes in the factor plane. Convex of the hull showing the outliers was occupied by the genotypes namely RG23 (Koombalai), RG24 (Tadukan), RG25 (Sorna kuruvai), RG2 (CK275), RG64 (Ottadaiyan), RG189 (Kaniranga), RG52 (ARB58), RG104 (Carolina Rinaldo Barsani), RG111 (Chigyungdo X IRGC 55466-1), RG155 (Sona mahsuri), RG142 (Sasyasree), RG51 (RPHP134) and RG71 (ARC11345 X IRGC21336-1). All the genotypes were widely scattered across different quarters (Figure3).



Figure 2. Distribution of genotypes across two components.



Figure 3. Scree plot showing Eigen value variation.

DFF- Days to Fifty percent Flowering, LL- Leaf Length, PH- Plant Height, NPT- Number of Productive Tillers, PL-Panicle Length, DM- Days to Maturity, NFG- Number of Filled Grains per panicle, SF- Spikelet Fertility, GL- Grain Length, GW- Grain Width, GLWR- Grain Length Width Ratio, SPY – Single Plant Yield

Principal Component Analysis measures the importance and contribution of each component to total variance. It can be used for measurement of independent impact of a particular trait to the total variance whereas each coefficient of proper vectors indicates the degree of contribution of every original variable with which each principal component is associated. The higher the coefficients, regardless of the sign, the more effective they will be in discriminating between accessions. There are no inferential tests to prove significance of proper values and the coefficients¹⁴. The current study was based on the Proportion of Variance Criterion by O'Rourke and Hatcher¹⁵. According to this criterion, Five principal components with cumulative variance of 80.6% was extracted which gives the clear idea of structure underlying the variables analyzed.

The different morphological traits contribute for total variation calculated for each component. For Component 1 which has the contribution of Days to 50% flowering (loadings -0.87), leaf length (0.78), plant height (0.765), panicle length (0.637), days to maturity (0.853) and number of filled grains (0.352) for 28.46% of the total variability. For component 2, grain width (0.886) and grain length width ratio (0.951) has contributed 16.8% of total variability. Similarly spikelet fertility (0.771) and single plant yield (0.542), grain length (0.81), number

| Principal Components | 1 | 2 | 3 | 4 | 5 |
|----------------------|--------|--------|--------|--------|--------|
| Eigen values | 3.415 | 2.017 | 1.724 | 1.399 | 1.112 |
| % of Variance | 28.459 | 16.810 | 14.369 | 11.658 | 9.267 |
| Cumulative % | 28.459 | 45.269 | 59.638 | 71.296 | 80.563 |
| Component Matrix | | | | | |
| DFF | 0.87 | 0.047 | -0.222 | -0.276 | -0.125 |
| LL (cm) | 0.78 | -0.297 | -0.062 | -0.063 | 0.108 |
| PH (cm) | 0.765 | -0.11 | -0.025 | 0.267 | 0.214 |
| NPT | -0.056 | 0.179 | 0.284 | -0.194 | 0.846 |
| PL (cm) | 0.637 | 0.134 | -0.084 | 0.524 | -0.029 |
| DM | 0.853 | 0.034 | -0.219 | -0.279 | -0.116 |
| NFG | 0.352 | 0.258 | 0.711 | 0.003 | -0.326 |
| SF (%) | 0.03 | 0.233 | 0.771 | -0.003 | -0.294 |
| GL (mm) | -0.147 | 0.232 | -0.256 | 0.81 | -0.031 |
| GW(mm) | -0.089 | -0.886 | 0.154 | 0.323 | -0.047 |
| GLWR | 0.024 | 0.951 | -0.228 | 0.098 | 0.034 |
| SPY (g) | 0.416 | 0.012 | 0.542 | 0.295 | 0.335 |

Table 2. Eigen value and percent of total variation and component matrix for the principal component axes

of productive tillers (0.846) has contributed for the total variation of 14.4%, 11.7% and 9.3% from component 3, component 4 and component 5 respectively.

Total variation in each principal axis is determined by number of variables. In the current study, Component 1 has the contribution from Days to 50% flowering, leaf length, plant height, panicle length, days to maturity and number of filled grains for 28.46% of the total variability. The first component has phenological and yield related variables. Similar type of performance was obtained by ^{14,16}. ¹⁷ has obtained similar pattern for phenological variables in rice. ¹⁸ reported that grain size that may be indicated by weight, volume, or length is one of the most important agronomic traits in rice. In current study also, grain width and grain length width ratio has contributed 16.8% of total variability in component 2. The remaining variability of 14.4%, 11.7% and 9.3% was consolidated in component 3, component 4 and component 5 by various traits such as spikelet fertility, single plant yield, grain length and number of productive tillers. Two Principal components (PRIN 1) and (PRIN 2) accounted for most of the variability observed among the rice germplasm collections from different locations. PRIN 1 accounted for 55% of the morphological variation in the rice germplasm collection and was loaded on plant height at vegetative

stage, leaf length, ligule length, plant height at maturity stage, days to 50% flowering, grain weight and number of grains per panicle. PRIN 2 accounted for 23 per cent of the variation and was loaded on total number of suckers at three weeks after planting, leaf width, grain weight and number of unfilled grains, while the remaining variables had weak or no discriminatory power¹⁹. ²⁰ has also reported similar results that grain characteristics along with panicle density, leaf length and plant height contributes for phenotypic diversity in a study involving Indian landraces of aromatic and non aromatic accessions. ²¹ has reported similar results that the traits such as maturity, heading, plant height, culm length, leaf length, and tillering ability were found to be the major factors contributing to the variation of parental lines of modern Philippine rice cultivars. ²² conducted a study in assessing the diversity of rice landraces in West Bengal. In that, principal component analysis revealed that six quantitative characters viz., leaf length, culm number, culm diameter, number of grains per panicle, grain length/breadth ratio and grain length significantly influenced the variation in these cultivars. The first six components in the PCA analysis with eigenvalues >1 contributed 75.9 per cent of the variability among genotypes evaluated for different agro-morphological traits. Other PCs (7-16) had eigenvalues less than 1. Thus, the prominent characters coming together in a particular principal component by contributing towards variability has the tendency to hang together offer opportunity for its utilization in crop breeding.

4. Conclusion

Principal component analysis has identified few characters that plays prominent role in classifying the variation existing in the germplasm set. The analysis identified days to 50% flowering, days to maturity plant height, number of filled grains, spikelet fertility, panicle length and grain length in different principal components are the most important for classifying the variation. Thus the results of principal component analysis used in the study have revealed the high level of genetic variation existing in the population panel and explains the traits contributing for this diversity. Hence the results will be of greater benefit to identify parents for improving various morphological traits analyzed in this study.

5. References

- Gopikannan M, Ganesh SK. Inter-Relationship and Path Analysis in Rice (*Oryza sativa* L.) under Sodicity. Indian Journal of Science and Technology. 2013; 6(9):5223–7.
- 2. Rani ABI, Maragatham N. Effect of Elevated Temperature on Rice Phenology and Yield. Indian Journal of Science and Technology. 2013; 6(8):5095–7.
- Shetty PK, Manorama K, Murugan M, Hiremath MB. Innovations that Shaped Indian Agriculture-then and Now. Indian Journal of Science and Technology. 2014; 7(8):1176–82.
- Liakat MA, McClung AM, Jia MH, Jennifar AK, McCouch SR, Georgia CE. A rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. Crop Science. 2011; 51:2021–35.
- 5. Oyelola BA. The Nigerian Statistical Association preconference workshop; 2004 Sep 20–21; University of Ibadan.
- Leonard K, Peter RJ. Finding Groups in Data: An Introduction to Cluster Analysis. Vol. 344. John Wiley & Sons; 2009.
- IRRI. IBPGR-IRRI Rice Advisory, & Resources, International Board for Plant Genetic. *Descriptors for rice, Oryza sativa L*: Int. Rice Res. Inst. 1980

- Sneath, Peter HA, & Sokal Robert R. Numerical Taxonomy: The Principles and Practice of Numerical Classification. W.H.Freeman & Co Ltd; 1973.
- Ariyo QJ, Odulaja A. Numerical analysis of variation among accessions of Okra [Abelmoschus esculentus (L.) Moench], Malvaceae. Annals of Botany. 1991; 67:527–31.
- 10. Fisher RA, Immer FR and Tedin O. The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. Genetics. 1932; 17:107 –24.
- 11. Robson DS. Application of K4 statistics to genetic variance component analysis. Biometrics 1956; 12:433–44.
- 12. Pooni HS, Jinks JL, Cornish MA. The causes and consequences of non-normality in predicting the properties of recombinant inbred lines. Heredity.1977; 38:329–38.
- 13. Kotch GP, Ortiz R, Ross WM. Genetic analysis by use of potato haploid populations. Genome. 1992; 35:103–8.
- Sanni KA, Fawole I, Ogunbayo A, Tia D, Somado, EA Futakuchi, Guei RG. Multivariate analysis of diversity of landrace rice germplasm. Crop Science. 2012; 52:494–504.
- O'Rourke N, Hatcher L. A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling. 2nd ed. Cary, NC, USA: SAS Institute; 2013.
- Kayode SA, Fawole I, Guei RG, Ojo DK, Eklou SA, Daniel TD. Geographical patterns of phenotypic diversity in Oryza sativa landraces of Côte d'Ivoire. Euphytica. 2008; 160: 389–400.
- 17. Guei RG, Sanni KA, Fawole AFJ. Genetic diversity of rice (*Oryza sativa* (L.). Agron Africaine. 2005; 5:17–28.
- Maji AT, Shaibu AA. Application of principal component analysis for rice germplasm characterization and evaluation. J Plant Breed Crop Sci. 2012; 4:87–93.
- 19. Takeda K. Inheritance of grain size and its implications for rice breeding. Rice Genetics. 1991; 2:181–9.
- 20. Rai A, Debal D, Rajasri R, Chattopadhayay B. Phenotypic characters of rice landraces reveal independent lineages of short-grain aromatic indica rice. AOB plants. 2013; 5:plt032.
- 21. Caldo RA, Sebastian LS, Hernandez JE. Morphology-based genetic diversity analysis of ancestral lines of Philippine rice cultivars. Philippine Journal of Crop Science. 1996; 21:86–92.
- 22. Chakravorty A,Ghosh PD, Sahu PK. Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. Am J Exp Agr. 2013; 3(1):110–23.