RESEARCH ARTICLE

ASSESSMENT OF GENETIC VARIATION AND DIVERSITY IN RICE GERMPLASM BASED ON PRINCIPAL COMPONENT ANALYSIS

Richa Sao, Ritu R. Saxena, and Parmeshwar K. Sahu*

Department of Genetic and Plant Breeding,
Indira Gandhi Krishi Vishwavidyalaya, Raipur-492012 (C.G.)
Corresponding authors' email: parmeshwarsahu1210@gmail.com

Received-04.12.2019, Revised-25.12.2019

Abstract: The present investigation was carried out to determine the relationship and genetic diversity among 47 rice genotypes based on the 27 agro morphological and quality traits using principal component analysis. Analysis of variance revealed significant and ample variation for all the studied traits in the germplasm lines. In this study, first eight principal components have more than one eigen values and more than 4% variations which explained the total 80.01% cumulative variance among 27 traits. PC1 had the contribution from traits *viz.*, grain weight/plant, harvest index, panicle weight, and head rice recovery, which accounted for 29.16% of the total variability. Plant height and grain dimension parameters were contributed 12.11% to the total variability in PC2. The remaining 38.74% variability was consolidated by PC3, PC4, PC5, PC6, PC7 and PC8. Scatter diagram plotted against PC1 and PC2 revealed that genotypes Basmati I, IGKVR-1244, MTU1010, Dhaur, Badalphool, Parra, Dhaura, Jaya, Beo-I, and Dhamna panda were very divergent for the traits under study. Thus, the results of principal component analysis revealed the wide genetic variation in rice genotypes. Identified accessions may be used as donors to improve the yield and quality traits in varietal development program.

Keywords: Genetic diversity, Germplasm, Principal component analysis, Rice, Variance

INTRODUCTION

Rice *(Oryza sativa* L.*)* is the most important cereal crop of the world meeting the dietary requirement of the people living in tropics and sub-tropics. The rising demand, saturation of cultivable field and low gross domestic production of rice are likely to cause a supply shortage of a crop in the near future. About 90% of the world's rice is grown and consumed in Asia, whereas 50% of the population depends on rice for food (Yugandhar *et al.,* 2018). Therefore being the most important staple food of the population, improving its productivity with good grain quality by exploiting the genetic diversity has become a crucial importance. Civilization could not exist without agriculture, and agriculture could not sustain the civilized world without sustainable crop varieties having good yielding capacity, better quality and nutritional traits, resistance to biotic and abiotic stresses. Sustainable crop varieties with desired traits can be only developed by exploiting the existing germplasm/ landraces by using plant breeding methods. The success of any crop improvement programme is highly dependent on the efficient utilization of the genetic variability present in the germplasm and the selection of genotypes with all possible desirable yield and quality contributing traits. Information on the genetic diversity and distance among the germplasm lines are essential for shaping breeding strategies.

Systematic study and characterization of rice germplasm is not only important for utilizing the desired trait based donors, but also essential in the present era for protecting the unique rice. Importance of germplasm can never be denied in agriculture system, because improvement in existing variety depends upon desirable genes which are possibly present in landraces and wild varieties only. These landraces could be fully exploited only after their morpho-genetical characterization and grouping based on their different traits (Sinha and Mishra, 2013). Several multivariate statistical techniques *viz.*, principal components analysis, cluster analysis and discriminant function analysis are available which can differentiate the landraces into different groups based on their morpho-genetical traits/ biochemical traits/ molecular data (Abdi and Williums, 2010).

Multivariate statistical techniques analyze multiple measurements on each individual under study, are widely used in analysis of genetic diversity. These techniques simultaneously analyse multiple measurements on each individual under investigation are widely used in analysis of genetic diversity irrespective of whether it is morphological, biochemical or molecular marker-based data and subsequently, classification of germplasm into different clusters (Maji and Shaibu, 2012). 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 (Nachimuthu *et al.,* 2014). 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 (Leonard and Peter, 2009). The higher the coefficients, regardless of the direction (positive or negative), the more effective they will be in discriminating between accessions (Abimbola *et al.,* 2016). Therefore, bearing in mind the value of PCA,

*Corresponding Author

the present research is conducted on rice germplasm accessions with an intention to identify the agro-morphological and quality traits liable for variations among the genotypes. The information, thus obtained, would be helpful to identify the diverse rice genotypes and to develop an effective rice-breeding program.

MATERIALS AND METHODS

Experimental material comprises total 47 genotypes including land races and released varieties of rice, which were procured from Rice Germplasm Section of Department of Genetics and Plant Breeding, Indira Gandhi Krishi Vishwavidyalaya, Raipur (C.G.) (Table 1). The experiment was conducted at research farm of Department of Genetics and Plant Breeding, IGKV, Raipur (C.G.), during the *Kharif-2*016 by following the Randomized Block Design (RBD) with two replications. Standard agronomic practices were adopted for all the treatments. Observations were recorded on five random plants of each genotype. Total 27 characters viz., Leaf:Length of blade, Leaf: Width of blade, Plant height, Stem: Length(excluding panicle), Panicle length, Panicle weight (g), Number of panicles per plant, 100 seed weight (g), Number of filled grains/panicle, Number of unfilled grains/panicle, Total spikelets/ panicle, Spikelet Fertlity %, Grain weight/plant(g), Shoot dry weight/plant (g), Harvest Index (%), Paddy length, Paddy width, Decorticated grain length (mm), Decorticated grain width(mm), Milled grain length, Milled grain width, Hulling %, Milling %, HRR%, Gel consistency, Alkali spreading value, Content of amylose % were measured according to DUS guidelines of PPV & FR Authority, GOI, 2007. Statistical analysis for analysis of variance and principal components, Eigen values, Eigen vectors and scatter diagram between PC1 & PC2 were done by using software OPSTAT, PAST v3.14 and XLSTAT v2018.

Table 1. List of rice genotypes used in the study

1	Dhaur	13	Satka	25	Dhamna panda	37	Karma mahsuri
2	JS-5	14	Shyamjir	26	No :21 (A)	38	Badshabhog selection 1
3	Badalphool	15	Banko –II	27	Badibarik	39	Dubraj selection 1
4	Amakoyali	16	Changadi	28	Basmati	40	Tarunbhog selection 1
5	Rambhoj	17	Dhaura	29	Kondiajan	41	Mahamaya
6	Churhaladhan	18	Kadamphool	30	Basmati (I)	42	Safri 17
7	Lalbatra	19	Kakdo	31	Beo (I)	43	IGKVR 2
8	Lallo-14	20	Kanji	32	Bodi	44	Indira aerobic dhan 1
9	Machhri kata	21	Machhariankhi	33	Bhataajan	45	TN 1
10	Parra dhan	22	Nallawadlu	34	Swarna sub 1	46	Jaya
11	Pinna basengi	23	Parra	35	IGKVR 1	47	MTU-1010
12	Sakra	24	Basmati	36	IGKVR 1244	48	

RESULTS AND DISCUSSION

The analysis of variance of 27 yield and quality traits in 47 rice germplasm accessions is presented in Table 2. The statistical procedure which separates or splits the total variation into different components is known as analysis of variation. It is useful in estimating the different components of variance. Such analysis divides the total variation into two main *viz*, variation between varieties and variation within varieties *i.e.* environmental variation into genotypic and environmental components. The results of the analysis of variance indicated that the mean sum of squares due to genotypes/ treatments was found to be highly significant for all the traits. This clearly indicates that variability does exist in all the genotypes for all the traits. The significant and relatively large percentage of the total variation attributable to G x E interaction suggests that genotypes responded differently to environment of rice. These results are in agreement with the findings of Sarawgi (2014); Ojha *et al*., (2017); Gaur *et al*., (2017).

Table 2. Analysis of variance of 27 yield and quality traits of 47 rice germplasm accessions

SV	DF	Mean Sum of Squares									
		1	2	3	4	5	6	7	8	9	10
Rep	1	23.58	0.01	739.4	282.6	107.7	4.53	1.53	0.00	29.88	53.63
Treat	46	102.24**	0.064**	721.72**	717.31**	12.45**	83.42**	16.52**	0.55**	5907.92**	588.09**
Error	46	2.21	0.01	34.28	24.56	4.79	1.70	3.25	0.00	21.41	8.85

SV	DF	Mean Sum of Squares									
		11	12	13	14	15	16	17	18	19	20
Rep	1	163.57	43.53	0.14	1.41	5.82	74.00	0.02	0.03	0.01	0.23

Treat	46	9922.71**	109.68**	61.99**	137.41**	152.87**	1.93**	0.19**	2.78**	102.24**	0.064**
Error	46	34.03	9.11	1.12	5.05	4.89	65.27	0.01	0.01	2.21	0.01

SV	**DF**	**Mean Sum of Squares**						
		21	**22**	**23**	**24**	**25**	**26**	**27**
Rep	1	0.01	13.41	13.18	38.15	7.19	0.01	4.04
Treat	46	721.72**	717.31**	12.45**	83.42**	16.52**	0.55**	19.66**
Error	46	0.03	24.56	4.79	1.70	3.25	0.25	21.41

****** Significant at 1% level of significance; ***** Significant at 5% level of significance

1= Leaf:Length of blade ; 2 = Leaf: Width of blade ; 3= Plant height ;4= Stem: Length(excluding panicle ; 5= Panicle length ; 6= Panicle weight (g) ; 7= Number of panicles per plant ; 8= 100 seed weight (g) ; 9= Number of filled grains/panicle ;10= Number of unfilled grains/panicle ;11= Total spikelets/ panicle; 12= Spikelet Fertlity % ; 13= Grain weight/plant (g) ; 14= Shoot dry weight/plant (g) ; 15=Harvest Index (%) ; 16= Paddy length ; 17= Paddy width ; 18= Decorticated grain length (mm) ; 19=Decorticated grain width(mm) ; 20= Milled grain length ; 21 ; Milled grain width ; 22= Hulling% ; 23= Milling% ; 24=HRR% ; 25= Gel consistency ; 26=Alkali spreading value ; 27= Content of amylose %

The results of PCA explained the genetic variation among the genotypes for all agro-morphological and grain quality characters under study. Principal components with eigen values more than 1 and variation percent more than 4% were considered as main PC (Brejda *et al.*, 2000; Abimbola *et al.*, 2016). The outcome of the PCA described the genetic diversity among rice genotypes for the studied traits. 'Eigen values' measure the importance and contribution of each component to total variance, whereas each coefficient of eigen 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 the genotypes. There are no standard tests to prove significance of Eigen values and the coefficients (Jolliffe, 2002).

In this study, eight components exhibited Eigen values of >1 and showed cumulative variation of 80.01% (Table 3). The first PC (29.61%) and second PC (12.11%) accounted for total 41.27% variation. Principal components, Eigen values, percentage contribution of each component to the total variation and main contributing traits for each principal component is presented in Table 2. Panicle weight, grain weight/plant, harvest index and head rice recovery were the important traits contributing to the PC1. As a result, the first component differentiated those genotypes that had high panicle weight, grain weight/plant, and harvest index and head rice recovery. The second principal component was contributed mostly by milled grain width, decorticated grain width, milled grain length and plant height which explained that PC2 discriminated those genotypes having higher grain dimensions and plant height. Gel consistency, amylose content and number of panicles per plant were the main traits contributing to PC3 whereas spikelet fertility %, leaf: length of blade and alkali spreading value were the main contributing traits for PC4. PC5 was discriminated by paddy width and number of panicles per plant whereas PC6 was discriminated by leaf: Length of blade and panicle length. Similarly, paddy length and leaf: width ratio was the main contributing traits for PC7 whereas spikelet fertlity % and gel consistency were the main contributing traits for PC8. Traits that have positive as well as negative impact on the PCs can be said to be the main source of variation and contributed mostly to differentiate the rice genotypes. Characters with high variability are expected to provide high level of gene transfer during breeding programmes (Nachimuthu *et al.*, 2014). Therefore, these traits can be used in the selection of diverse genotypes from particular principal component. The results of present study were in accordance with the findings of Ojha *et al.*, (2017), Gaur *et al.*, (2017) and Nachimuthu *et al.*, (2014).

The distribution of rice genotypes based on the first and second PC exhibited the phenotypic variation among the population and explains how these widely dispersed along both the axes**.** The genotype wise scatter diagram of the 47 rice genotypes (scores) across the first two PC axes (Figure 1) revealed that genotypes Basmati I, IGKVR-1244, MTU1010, Dhaur, Badalphool, Parra, Dhaura, Jaya, Nallawadlu, Beo-I, and Dhamna panda were very divergent for the traits under study. The dispersion of the different genotypes across the plot showed variations among the genotypes, although most accessions from the same PC were closely distributed and overlapped with each other.

Table 3. Eigen value, variability percentage and eigen vectors of 47 rice germplasm accessions for yield and quality characters

Traits	**Principal Components**							
	PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**
Eigen value	7.87	3.27	2.84	2.03	1.61	1.50	1.34	1.13
Variability (%)	29.16	12.11	10.53	7.53	5.96	5.56	4.97	4.19
Cumulative variability %	29.16	41.27	51.80	59.33	65.29	70.85	75.81	80.01
Eigen vectors								
leaf: width of blade	-0.02	0.03	**0.29**	0.22	-0.05	-0.18	**0.38**	-0.26
Plant height (cm)	-0.14	**0.27**	**0.21**	0.12	-0.35	-0.36	-0.06	-0.04
Stem: Length(excluding panicle)	-0.15	0.26	0.20	0.10	-0.33	-0.40	-0.11	0.01
Panicle length (cm)	0.04	0.09	0.11	0.20	-0.16	**0.31**	**0.38**	-0.41
Panicle weight (g)	**0.33**	-0.06	-0.04	0.13	-0.08	0.02	-0.06	-0.11
Number of panicles per plant	0.10	0.06	-0.30	**0.24**	**0.29**	-0.24	0.03	-0.09
100 seed weight (g)	-0.03	0.25	-0.23	0.21	0.19	-0.12	-0.08	**0.27**
Number of filled grains/panicle	**0.31**	0.04	0.19	0.07	0.12	-0.05	-0.12	0.07
Number of unfilled grains/panicle	0.27	-0.03	**0.24**	-0.02	**0.22**	-0.05	-0.16	-0.07
Total spikelets/ panicle	**0.31**	0.02	**0.21**	0.05	0.15	-0.05	-0.13	0.04
Spikelet Fertlity %	0.14	0.17	-0.05	**0.33**	-0.12	-0.09	0.07	**0.42**
Grain weight/plant (g)	**0.33**	-0.02	-0.02	0.11	-0.04	-0.01	-0.08	-0.08
Shoot dry weight/plant (g)	0.27	0.00	-0.03	0.19	-0.17	0.00	-0.16	-0.19
Harvest Index (%)	**0.30**	-0.02	0.00	0.02	0.06	0.00	-0.02	0.02
Paddy length (mm)	-0.02	0.12	-0.14	0.00	**0.22**	-0.12	**0.62**	**0.20**
Paddy width (mm)	-0.13	0.20	-0.05	0.07	**0.41**	-0.15	-0.08	-0.19
Decorticated grain length(mm)	0.17	**0.31**	-0.17	-0.24	-0.23	**0.19**	0.06	0.17
Decorticated grain width(mm)	0.09	**0.43**	-0.01	-0.26	0.12	0.00	-0.03	-0.28
Milled grain length (mm)	0.18	**0.30**	-0.17	-0.23	-0.25	**0.18**	0.06	0.18
Milled grain width (mm)	0.09	**0.43**	-0.02	-0.27	0.12	0.00	-0.03	-0.27
Hulling%	0.24	-0.15	-0.05	-0.11	-0.06	-0.11	0.24	0.01
Milling%	0.15	-0.26	-0.11	-0.24	-0.17	-0.31	0.16	-0.14
Head Rice Recovery %	0.27	-0.15	-0.04	-0.16	-0.14	-0.28	0.07	0.11
Gel consistency (mm)	0.04	0.08	**0.48**	-0.14	0.16	-0.02	0.03	**0.29**
Alkali spreading value	0.17	0.12	0.07	**0.28**	0.09	0.10	**0.28**	0.06
Content of amylose %	0.01	0.02	-0.45	0.23	-0.12	-0.11	-0.14	-0.17

#bold figures in eigen vectors are representing the main contributing traits for respective PC

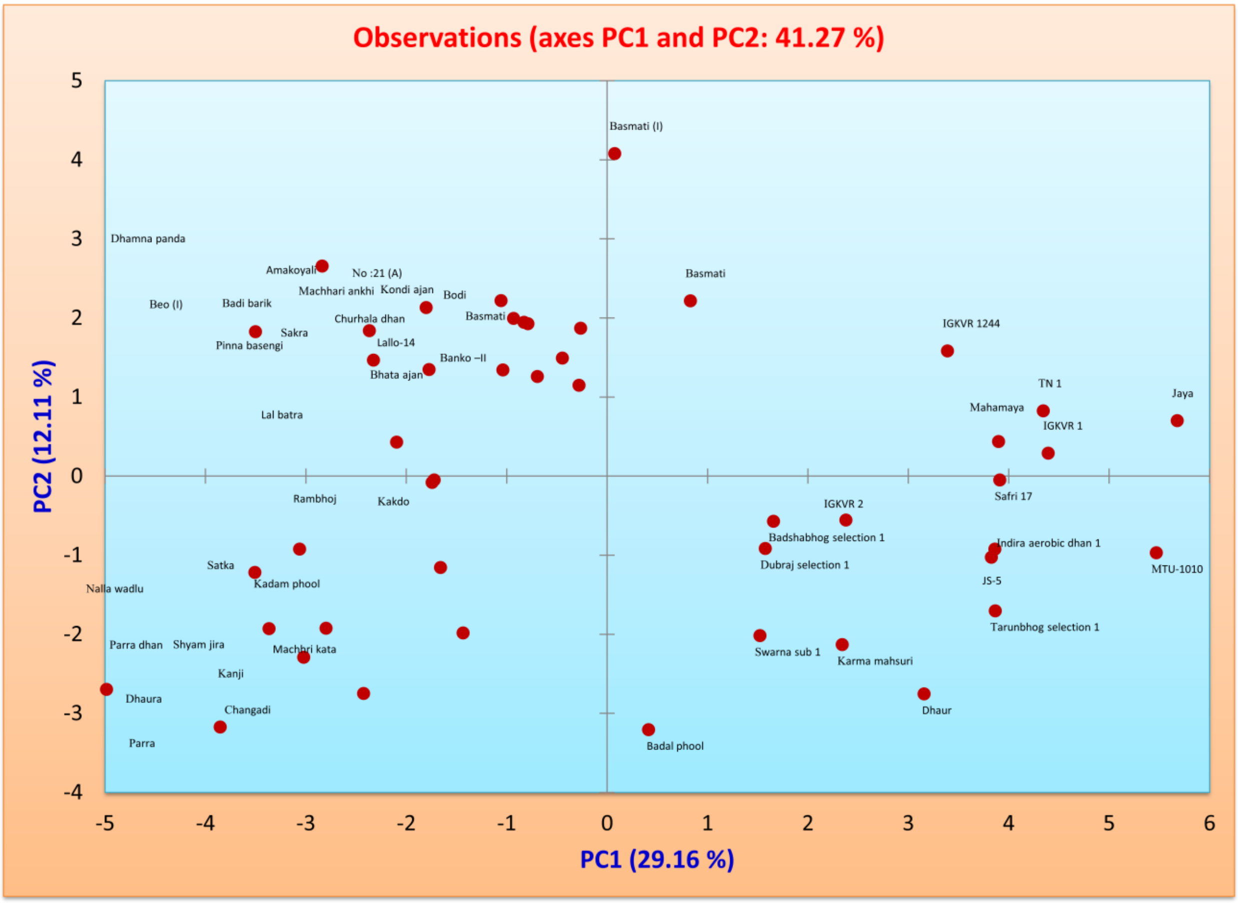

Fig. 1. PCA scatter diagram showing the dispersion of rice genotypes across PC1 and PC2

CONCLUSION

Analysis of variance revealed the existence of significant variability for all the characters included for study. It is cleared that the principal component analysis highlights the characters with maximum variability. So, intensive selection procedures can be designed to bring about rapid improvement of yield and quality traits. PC also helps in ranking of genotypes on the basis of PC scores in corresponding component. Genotypes Dhaur, Badalphool, Parra, Dhaura, Jaya, MTU-1010, IGKVR-1, and Safri-17 were the diverse accession for both quality and yield attributing traits. Identified accessions may be used as donors to improve the yield and quality traits in varietal development program.

REFERENCES

Abdi, H. and Williums, L.J. (2010). Principal component analysis. WIREs Computational Statistics. John Wiley & Sons, Inc. 2: 433-458.

Abimbola, L.H.M., Din, K.A., Nookiah, R., Hassan, M.S., Ngah, C.W. and Sukaimi, J. (2016). Genetic variability for yield, yield components and fatty acid traits in oil palm (*Elaeisguineensis*jacq.) germplasm using multivariate tools. Int. J. Agril. Fores. Plantat., 2: 219-226.

Brejda, J.J., Moorman, T.B., Karlen, D.L. and Dao, T.H. (2000). Identification of regional soil quality factors and indicators in Central and Southern High- Plains. Soil Sci. Soc. Am. J. 64: 2115-2124.

Gour, L., Maurya, S.B., Koutu, G.K., Singh, S.K., Shukla, S.S. and Mishra, D.K. (2017). Characterization of rice (*Oryza sativa* L.) genotypes using principal component analysis including scree plot & rotated component matrix. Int. J. Chem. Std.. 5(4): 975-983.

Jolliffe, I.T. (2002). Principal component analysis. 2nd edition, Springer-Verlag, New York, USA.

Leonard, K. and Peter, R.J. (2009). Finding Groups in Data: An Introduction to Cluster Analysis. 344.

Maji, A.T. and Shaibu, A.A. (2012). Application of principal component analysis for rice germplasm characterization and evaluation. J. Plant Breed. Crop Sci., 4(6): 87-93.

Nachimuthu, V.V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S. and Manonmani, S. (2014). Evaluation of rice genetic diversity and variability in a population pannel by principal component analysis. Indian J. Sci. Technol., 7(10): 1555-1562.

Ojha, G.C., Sarawgi, A.K., Sharma, B. and Parikh, M. (2017). Principal component analysis of morpho-physiological traits in rice germplasm accessions (*Oryza sativa* L.) under rainfed condition, Int. J. Chem. Std. 5(5): 1875-1878.

Sarawgi, A.K., Parikh, M., Sharma, B. and Sharma, D. (2014). Phenotypic divergence for agro-morphological traits among dwarf and medium duration rice germplasms and inter-relationships between their quantitative traits. The Bioscan, 9 (4): 1677-1681.

Sinha, A.K. and Mishra, P.K. (2013). Morphology based multivariate analysis of phenotypic diversity of landraces of rice (*Oryza sativa* L.) of Bankura district of West Bengal. J. Crop Weed, 9(2):115-121.

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