Original Research Article

Principal component analysis of yield and its attributing traits in recombinant inbred lines of Rice developed for submergence tolerance

ABSTRACT

Submergence is an important abiotic stress affecting yield and productivity of rice. The present investigation was studied in a RIL (recombinant inbred lines) population consisting of 120 lines where principal component analysis was carried out on different agromorphological traits viz. days to fifty percent flowering, plant height, panicle length, total tillers per plant, number of effective tillers per plant, number of filled grains and Spikelet fertility, test weight, grain yield per plant, biomass per plant and harvest index to assess the relative contribution of various traits towards total variability. It was found that first five components with Eigen values greater than 1 accounted for 74.33% of total explained variability. The highest positive eigenvalue was observed for plant height and panicle length in PC1, revealing their maximum contribution to variation in the lines.

***Keywords**: submergence, principal component analysis, biplot, scree plot, eigen values.*

1. INTRODUCTION

Rice (*Oryza sativa* L) is the chief cereal which is cultivated throughout the world. It contributes to 70% of calorie intake in the developing countries (Dawe,2000). The rising global population leads to increased demand for the food crop which necessitates improved productivity of rice. However, rice is affected by several abiotic and biotic stresses that reduces the yield and productivity of rice. One such important abiotic stress is submergence, which has been recognized as the third most significant stress causing total yield loss in rice (Sarkar *et al.,* 2006). The climate change has aggravated flooding in lowland areas especially in South-Asian countries which led to an estimated economic loss of one billion US dollars in rice production (Mackill *et al.,* 2006). There are many traditional varieties and landraces such as FR13A, AC 42087, AC 42088, AC 20431B, Thavalu, Kurkaruppan and Goda Heenati which exhibit high rate of survival percentage when exposed to submergence stress (Vergara *et al.,*1975; Sarkar *et al*., 2011). However, these genotypes showed reduced yield as well as poor quality as compared to popular varieties in farmers' fields. These cultivars can be utilized as donors in various breeding programs to develop improved varieties that have increased yield as well as tolerance to submergence. It is necessary to study the yield and yield related attributes such as plant height, panicle length etc of the population developed from such genotypes for evaluation of variability and breeding potential of the population. A large number of traits recorded may not be adequate for differentiating the lines. Hence, principal component analysis (PCA) is employed to report different patterns and also reduce redundancy in data sets (Upadhyay *et al.*, 2022). Principal component analysis is a multivariate statistical method that determines and reduces duplicate genotypes with similar attributes. It further allows characterization of different genotypes where each genotype may be assigned to only one group (Singh *et al.*, 2016; Khodadadi *et al.,*2011).The large number of variables were classified into different components and their contribution to the total variance was also estimated by PCA (Holme *et al.*, 2019). The present investigation was carried out in a RIL population to study the principal component analysis for different yield related attributes.

2. MATERIALS AND METHODS

The experimental material consisted of a RIL population consisting of 120 lines derived from Swarna-Sub 1 X AC 20431B along with four landraces AC 38575, AC 42088, AC 42087 and FR13A.The experiment was conducted in Randomized complete block design (RCBD) with two replications at ICAR-National Rice Research Institute (NRRI), Cuttack, Odisha during *kharif* season of 2022. The lines were planted in two blocks and two parents were used as checks. Recommended package of practices were followed. Three randomly selected plants per replication were chosen for taking data for agro-morphological characters such as days to 50% flowering, plant height, panicle length, total number of tillers, number of effective tillers per plant, number of filled grains and Spikelet fertility, test weight, grain yield per plant, biomass per plant and harvest index. The Principal Component analysis was performed on mean of each trait taken to determine the quantitative traits that contribute most of the observed variations among the genotypes. The analysis was carried out using Factoextra package of R software version 4.3.0.

3. RESULTS AND DISCUSSION

Principal component analysis (PCA) was conducted on eleven morphological traits of 120 recombinant inbred lines to find independent effects of each trait under study. Eigen values, percentage of variance and cumulative percentage of all the principal components are compiled in Table 1. The principal components having eigen values greater than 1 were considered for significant and were included for interpretation of results. In the present investigation, first five components contributed 74.33% of explained variability. Similar findings were shown by Singh *et al.*, 2020; where five principal components contributed to 73.84% of the total variability.

Table1: Eigen values and contribution of variability for the principal component axes

Principal components	Eigenvalue	Percentage of variance	Cumulative Percentage
PC1	2.185	19.87	19.87
PC2	1.998	18.16	38.03
PC3	1.478	13.43	51.47
PC4	1.375	12.5	63.98
PC5	1.138	10.35	74.33
PC6	0.93	8.46	82.79
PC7	0.74	6.76	89.55
PC8	0.6	5.45	95.01
PC9	0.448	4.07	99.09
PC10	0.086	0.78	99.87
PC11	0.013	0.12	100

A scree plot was plotted using the percentage of explained variances and principal components and depicted in Figure 1. The graph clearly shows the presence of maximum variance in PC1, PC2, PC3, PC4 and PC5. These results were in conformity with the results shown by Kashyap *et al.,* 2020;whichdepicted that five principal components PC1, PC2, PC3,PC4 and PC5 showed maximum percentage variance in the scree plot.

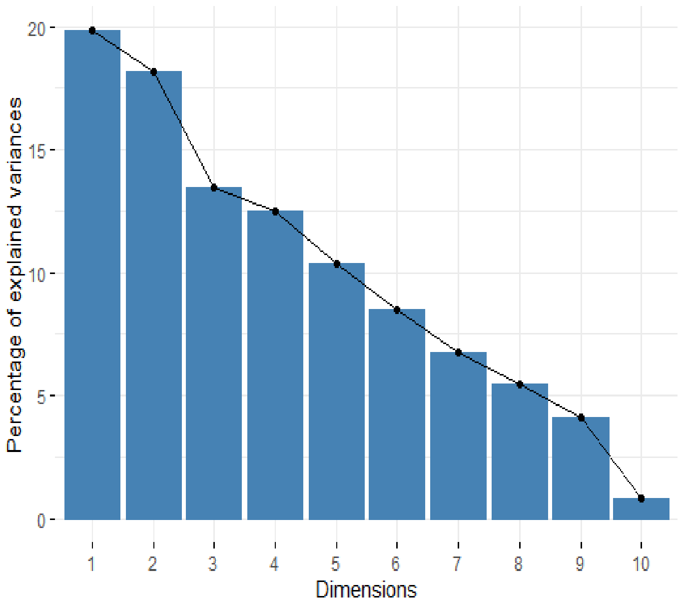

Fig 1. Scree plot diagram of principal components of Recombinant Inbred Lines

The contribution of eleven quantitative traits to the principal components is shown in Table 2. The first principal component (PC1) contributed to 19.87% of total variability which is the maximum among all principal components. Quantitative traits such as plant height (0.2332), panicle length (0.1835), spikelet fertility (0.0428), test weight (0.0313), biomass per plant (0.1067) and days to 50% flowering (0.0107) showed positive loading while rest traits showed negative loadings.Plant height and panicle length were relatively higher contributors to PC1. This result was found to be in agreement with Singh *et al.*, 2020 and Mahendran *et al.*, 2015 who reported the positive contribution of plant height, panicle length and days to 50% flowering to PC1. Reetisana*et al.*,2022 also reported positive influence of spikelet fertility and test weight on the first principal component (PC1).

The second principal component (PC2) showed 18.16% contribution to the total variability. PC2 was positively affected by traits such as harvest index (0.3439) and panicle length (0.1838) while rest all traits showed negative influence.

The third principal component (PC3) contributed to 13.43% of total variability. PC2 was positively affected by tiller number per plant (0.121), number of effective tillers per plant (0.1138), days to 50% flowering (0.0901) and biomass per plant (0.0844). Other traits such as plant height, panicle length, filled grains per panicle, spikelet fertility, test weight and harvest index showed negative loadings. The major contribution of tiller number and effective tiller number per plant to principal component (PC3) was also reported by Madhubabu *et al.*,2020.

The fourth principal component (PC4) showed 12.5% contribution to the total variability. Traits such as biomass per plant (0.4217), days to 50% flowering (0.438), plant height (0.2998), grain yield per plant (0.2957), panicle length (0.1809), number of effective tillers per plant (0.1125), tiller number per plant (0.0927) and test weight (0.0639) showed positive loadings into the component. Only harvest index showed negative influence on the component.

Table 2. Factor loadings (Eigen vectors) for different agro-morphological traits for the principal components

Traits	PC1	PC2	PC3	PC4	PC5
DFF	0.0107	-0.2162	0.0901	-0.0804	0.4385
PH	0.2332	-0.1077	-0.3314	0.463	0.2998

PL	0.1835	0.1838	-0.3544	0.427	0.1809
TN	-0.6132	-0.133	0.121	0.261	0.0927
ETN	-0.6094	-0.1245	0.1138	0.275	0.1125
FG	-0.0239	-0.4472	-0.2676	-0.1251	-0.4203
SF	0.0428	-0.4823	-0.099	0.1202	-0.4398
TW	0.0313	-0.3441	-0.3045	0.2564	0.0639
GYP	-0.2322	-0.0489	-0.5211	-0.48	0.2957
BYP	0.1067	-0.4504	0.0844	-0.3191	0.4217
HI	-0.309	0.3439	-0.5252	-0.1570	-0.1556

DFF-Days to 50% flowering, PH-Plant height, PL-Panicle length, TN-Tiller number, ETN-Effective tiller number, FG-filled grains per panicle, SF-Spikelet Fertility, TW-Test weight, GYP- grain yield per plant, BYP-Biomass yield per plant, HI-Harvest Index

A biplot was plotted between PC1 and PC2 using variability of all agromorphological traits taken to study the interaction between both the components. The biplot is presented in figure no.2. The first quadrant comprised of only one trait panicle length that showed positive correlation with PC1 and PC2 while second quadrant involved plant height, test weight, days to 50% flowering, spikelet fertility and biomass per plant which were positively correlated with PC1 and negatively correlated with PC2. Third quadrant comprised of tiller number per plant, number of effective tillers per plant and filled grains per panicle which were negatively correlated with both PC1 and PC2. Fourth quadrant contained harvest index which was positively correlated with PC2 but negatively correlated with PC1.

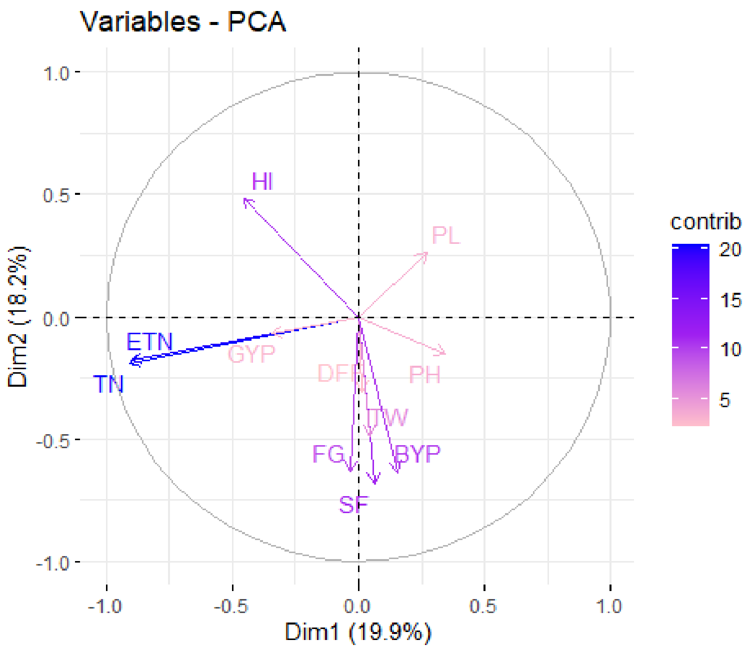

Fig 2: Biplot of 11 variables across the first two principal components

CONCLUSION

Principal component analysis (PCA) is widely applied in various breeding programs to study contribution of different quantitative traits to total variability. PCA analysis in present investigation reduced a data set of eleven variables into 5 principal components describing 74.33% of cumulative variability. It also showed that plant height, panicle length and test weight can be used for selection of genotypes. The RILs can be further evaluated in different locations for recording yield and yield related attributes in different environments. The selected lines can be further utilized as donors in different breeding programs.

REFERENCES

1. Dawe D. The contribution of rice research to poverty alleviation. In J.E. Sheehy et al. (ed.) Redesigning rice photosynthesis to increase yield. International Rice Research Institute, Elsevier Science, Makati City (Philippines), Amsterdam.2000;3-12.
2. Holme IB, Gregersen PL, Brinch-Pedersen H. Induced genetic variation in crop plants by random or targeted mutagenesis: convergence and differences. Frontiers in Plant Science.2019;10:1468.
3. Kashyap A, Yadav VK. Principal component analysis and character association for yield components in rice (*Oryza sativa* L.) genotypes of salt tolerance under alkaline condition. International Journal of Current Microbiology and Applied Sciences.2020;9(10):481-495.
4. Khodadadi M, Fotokian MH, Miransari M. Genetic diversity of wheat (*Triticum aestivum* L.) genotypes based on cluster and principal component analyses for breeding strategies. Australian Journal of Crop Science. 2011;5(1):17–24.
5. Mackill DJ, Collard BCY, Neeraja CN, Rodriguez RM, Heuer S, Ismail AM. QTLs in rice breeding: examples for abiotic stresses. In: Brar DS, DJ Mackill, B Hardy, editors. In Rice Genetics.2006; 55-167.
6. Madhubabu P, Surendra R, Suman K, Chiranjeevi M, Fiyaz RA, Rao DS, Chaitanya U, Rao LVS, Babu VR, Neeraja CN. Assessment of genetic variability for micronutrient content and agro-morphological traits in rice (Oryza sativa L.). Indian Journal of Genetics. 2020;80(2): 130-139.
7. Mahendran R, Veerabadhiran P, Robin S, Raveendran M. Principal component analysis of rice germplasm accessions under high temperature stress. International Journal of Agricultural Science and Research (IJASR).2015; 5(3): 355-359.
8. Reetisana N, Sastry EVD, Julia T, Pyngrope A, Devi YS. Principal component analysis in screening of submergence tolerance in rice (Oryza sativa L.) genotypes of Manipur. The Pharma Innovation Journal.2022; 11(6):283-286.
9. Sarkar RK and Bhattacharjee B. Rice Genotypes with SUB1 QTL Differ in Submergence Tolerance, Elongation Ability during Submergence and Re-generation Growth at Re-emergence. Rice. 2011; 5:7.
10. Sarkar RK, Reddy JN, Sharma SG, Ismail AM. Physiological basis of submergence tolerance in rice and implications for crop improvement. Current Science. 2006; 91:899–6.
11. Singh S, Prakash A, Chakraborty NR, Wheeler C, Agarwal PK, Ghosh A. Trait selection by path and principal component analysis in J*atropha curcas* for enhanced oil yield. Industrial Crops and Products.2016;86:173-179.
12. Singh SK, Habde SV, Khaire A, Mounika K, Majhi PK, Singh DK. Principal component analysis in F2 mapping population for high grain zinc content in rice (Oryza sativa L). Plant Archives.2020;20(2):6935-6941.
13. Upadhyay S, Rathi S, Choudhary M, Snehi S, Singh V, Singh PK, Singh RK. Principal Component Analysis of Yield and its Attributing Traits in Advanced Inbred Lines of Rice under Sodicity condition (*Oryza sativa* L.). Biological Forum – An International Journal.2022; 14(2): 1273-1276.
14. Vergara BS, MazaredoA.Screening for resistance to submergence under greenhouse conditions. In Proceedings International Seminar on Deepwater Rice. Dhaka, Bangladesh: Bangladesh Rice Research Institute.1975; 67–70.