Oryza Vol. 52 No.2, 2015 (111-116)

Principal component analysis for yield and related attributes in Aerobic rice

Chandan Kumar*, Nilanjaya and Santosh Kumar

Rajendra Agricultural University, Pusa, Samastipur – 848125, Bihar, India

E-mail: chandandbg@ymail.com

ABSTRACT

The present investigation was planned to assess the principal component for yield, yield contributing traits and quality traits in thirty genotypes of rice suitable for aerobic condition. The genotypes were procured from Directorate of Rice Research, Hyderabad and were evaluated in randomized block design with three replication. Principal component analysis (PCA) indicated that three components (PC1 to PC3) accounted for about 78.809% of the total variation among traits. Out of total principal components retained PC1, PC2 and PC3 with values of 58.125 %, 13.305 % and 7.380 % respectively contributed more to the total variation. The first principal component had high positive loading for 7 characters out of 18 viz.*grain yield per plant (0.278), canopy temperature (0.278), 1000-grain weight (0.271), panicle length (0.270), critical temperature for reproductive stage (0.239), seedling vigour (0.213) and days to 50% flowering per plant (0.201) which contributed more to the diversity. Genotypes in cluster VI showed higher mean performance for most of the yield attributing characters. Therefore, selection of parents from this cluster for these traits would be effective. The result of present study could be exploited in planning and execution of future breeding programme in rice under aerobic condition.*

Key words: *PCA, K- Clustering, Aerobic rice*

Selection for seed yield and production of the cultivars with high yield potential is the main objective of breeding programs. Many researchers (Quarrie *et al*., 1999, Richards, 1996) believed that genetic improvement of seed yield must be done *via* genetic improvement of physiological traits. In determining the potential of genetically different lines and cultivars, breeders have to observe many different characters that influence yield. Accurate evaluation of these characters is made more difficult by the genotype by environment interaction (Tadesse and Bekele, 2001). Principal component analyses helps researchers to distinguish significant relationship between traits. This is a multivariate analysis method that aims to explain the correlation between a large set of variables in terms of a small number of underlying independent factors. The cluster analysis is also an appropriate method for determining family relationships but the main advantage of using PCA over cluster analysis is that each genotype can be assigned to one group only (Mohamed, 1999). The main objective of this study is to assess the potential genetic diversity among rice genotypes by using cluster analysis and cluster analysis-PCA-based methods for selection of parents in hybridization programme to obtain desirable segregants in advanced generation.

MATERIALS AND METHODS

The experimental material for the present study comprised of thirty genotypes of aerobic rice procured from Directorate of Rice Research, Hyderabad laid in randomized block design (RBD) with three replications at the Field Experimentation Centre of Department of Plant Breeding and Genetics, Rajendra Agricultural University, Pusa Samastipur Bihar during wet season 2013. Standard agronomic practices and plant protection measures were taken as per schedule. Each genotype was grown in a plot of 5x2 square meters with the spacing 20x15 cm row to row and plant to plant respectively. Observations were recorded on five randomly selected plants per replication for different quantitative traits *viz.* plant height (cm), tillers per plant, panicle length (cm), panicle per plant, spikelet fertility (%), chlorophyll content (SPAD VALUE), relative water content, canopy temperature (^{0}C), critical temperature for reproductive stage (^{0}C), root length (cm), root volume (cm^3), grains per panicle, 1000 grain weight (g) and grain yield per plant (g) while seedling vigour (1-9 scale), days to 50% flowering, days to physiological maturity and grain yield (kg/plot) were recorded on plot basis. The data were analyzed following PCA and Mahalanobis D^2 statistics as suggested by Rao (1952).

Seedling vigor was measured on 1 to 9 scale under standard evaluation system (SES) as proposed by IRRI (1996). Canopy temperature was measured using a hand-held infrared thermometer (Taleb, 2011).

Critical temperature for reproductive stage was measured by recording accumulated daily temperature till flowering and accumulated daily temperature till maturity and calculated using the equation as proposed Arnold (1960):

$$\text{CRT} = \frac{\text{Accumulated daily temp. till maturity - Accumulated daily temp. till flowering}}{\text{Number of days to maturity - number of days to flowering}}$$

Accumlulated daily temp. till maturity = Sum of daily average temperature × No. of days upto maturity

Accumulated daily temp. till flowering = Sum of daily average temperature × No. of days upto flowering

Root length and Root volume was measured by digital root scanner using Win RHIZO software (V5.0), Regent Instruments, Quebec, Canada.

Leaf relative water content (RWC) of the flag leaves was determined using the equation given by Barr and Weatherley (1962):

$$\text{RWC} = \frac{\text{F.W} - \text{D.W.}}{\text{T.W.} - \text{D.W.}} \times 100$$

Where,

F.W. = Fresh Weight of flag leaf (g)

D.W. = Dry Weight of flag leaf (g)

T.W. = Turgid Weight of flag leaf (g)

The data were analyzed using WINDOSTAT version 9.1 software for computation of correlation coefficients and path coefficient analysis.

RESULT AND DISCUSSION

Principal component analysis is a form of multivariate analysis utilized in present study, where canonical vectors and roots representing different axes of differentiation and the amount of variation accounted for each of such axes, respectively **(**Sharma *et al.* 1996). The variations were identified for eighteen characters in thirty aerobic rice genotypes. Most of the variation was accounted for the first three canonical vectors in the present study represented the primary, secondary and tertiary axes of differentiation in thirty aerobic rice genotypes. These three vectors were used to obtain, means of canonical variates and three dimensional (3D) pictorial representations of all the thirty genotypes were obtained. Genetic divergence was measured between variates in terms of spatial distance. According to Chahal and Gosal (2002), characters with largest absolute value closer to unity within the first principal component influence the clustering more than those with lower absolute value closer to zero. Therefore, in the present study, differentiation of the genotypes into different clusters was because of relatively high contribution of few characters rather than small contribution from each character.

The first three Principal components eigene value 10.462 accounted for about 78.809 % of the variation among the genotypes for all the characters (Table 1). The first Principal component explained 58.125 % of the variation and was associated mainly with grain yield per plant (0.278), canopy temperature (0.278), 1000-grain weight (0.271), panicle length (0.270), critical temperature for reproductive stage (0.239), seedling vigour (0.213) and days to 50% flowering per plant (0.201) while characters like grain yield per plant (-0.282), root volume (-0.266), relative water content (-0.265), chlorophyll content (-0.263), grains per panicle (-0.253), panicle per plant (-0.251), root length(-0.221) contributed negatively to first component The first component identified mainly phonological variables presenting negative contributions. Guei *et al.* (2006), Ashf *et al* (2012)

Table 1. Vector loadings and percent variation for different traits contributed by Principal Components

Characters	Vector 1	Vector 2	Vector 3
Seedling Vigour (log)	0.213	0.262	0.258
Days to 50% Flowering	0.201	-0.115	-0.443
Days to Physiological maturity	0.042	-0.480	-0.343
Plant Height(cm)	0.127	0.470	0.034
Tillers/ Plant	-0.259	-0.121	-0.129
Panicle Length (cm)	0.270	-0.131	0.159
Panicles/ Plant	-0.251	-0.035	0.023
Spikelet Fertility (%)	-0.099	0.369	-0.564
Chlorophyll Content	-0.263	-0.087	0.035
Relative Water Content (%)	-0.265	0.157	0.120
Canopy Temp°	0.278	-0.105	0.022
Critical Temp° for Rep.stage	0.239	-0.277	0.278
Root Length (cm)	-0.221	-0.350	0.050
Root Volume (cm^3)	-0.266	0.004	0.172
Grains/ Panicle	-0.254	0.125	0.114
1000 Grain Weight (g)	0.271	0.021	-0.299
Grain Yield/ Plant (g)	-0.282	-0.173	-0.022
Grain Yield (kg/Plot)	-0.269	0.096	-0.170
Eigene Value (Root)	**10.462**	**2.395**	**1.328**
% Var. Exp.	**58.125**	**13.305**	**7.380**
Cum. Var. Exp.	**58.125**	**71.429**	**78.809**

observed similar performance for phonological variables of rice.

The second Principal component with eigene value 2.395 was responsible for about 13.305 % of the variation. Characters which contributed positively to total variation are seedling vigor (0.262), plant height (0.470), spikelet fertility (0.369), relative water content (0.157), grains per panicle (0.125) while the characters, days to physiological maturity (-0.480), root length (-0.350), critical temperature for reproductive stage (-0.277), panicle length (-0.131), tillers per plant (-0.121), canopy temperature (-0.105), grain yield per plant (-0.173) contributed negatively to total variation.These findings agree with Caldo *et al.* (1996), Sanni *et al*. (2008).

The third principal component with eigene value 1.328 was responsible for about 4.380 % of the variation. Characters which contributed positively to total variation are critical temperature for reproductive stage (0.278), seedling vigor (0.258), root volume (0.172) panicle length (0.159), relative water content (0.120), grains per panicle (0.114) while the characters like spikelet fertility (-0.564), days to 50% flowering (-0.443), days to physiological maturity (-0.343), 1000-grain weight (-0.299), tillers per plant (-0.129) contributed negatively to total variation. These findings revealed that first three principal components were related to various morphological and physiological traits in rice mostly associated with early genotypes and also these traits can identify the diverse genotypes which could be employed in hybridization programme (Das *et al.* 2014).

Table 3 showed that thirty genotypes of rice were grouped into six clusters using K-means non-hirerichical clustering. Cluster V comprised of 8 genotypes namely, NDR 1143, RP 5310, NDR 1143, NDR 1140, CR 2658-11, RP 5333-11, CB 09515, Rewa 1208-15 followed by cluster III (7 genotypes) viz., KMP 175 E, Rewa 862-27-1, MGD 1104, GK 5022, NDR 1159, UPRI 2012-15. Cluster IV made up by 5 members of genotypes namely, NDR 1145, CR 2991-1-2-1-1-1, CB 09512, CB 10-504, Vandana. Cluster V1 Comprised of six genotypes namely IR 64, RASI, MAS 946, IR 36, PAU 3832-79-4-3-1 whereas cluster I and II having 3 and 2 genotypes namely, RP 5311, MAS 9, UPRI 2012-15 and R-RF-48, OR-2172-7-1, respectively.

Means of canonical variate (mean value of Z_1, Z_2 and Z_3) have been presented Table 2. Three dimensional (3D Plot) and 2 dimensional (2D Plot) configuration of genotypes were presented in Fig. 1 and Fig. 2 respectively. The figure depicts spatial distance between genotypes as a measure for genetic divergence between them. Genetic divergence was found maximum between 1and 19 (IR 64 and OR 2172-7-1) followed by 7 and 6 (IR 36 and KMP 175E) 2 and 26 (RASI and UPRI 2012-15) 21 and 12 (CB 10-504 and MGD 1104) and 20 and 15 (CB 09512andGK 5022) exhibited high genetic diversity, whereas genotypes 25 and 14 (RP 5333-5-2-B and NDR 1143), followed by 13 and 9 (MAS 9 and Rewa 862-27-1), 11 and 29 (CR 2991-1-2-1-1-1 and Rewa 1208-15) and 23 and 4 (NDR 1159 and RP 5311) expressed low genetic diversity.

K-cluster mean of Principal Component analysis for all the eighteen characters obtained and presented in Table 4. The individual cluster reflected superiority for various traits. Cluster I showed highest mean for plant height (109.444). Cluster II showed highest mean for seedling vigor (0.876), days to 50% flowering (98.00), days to physiological maturity (112.833), panicle length (26.500) and canopy temperature (31.217) whereas lowest mean value for

Table 2. Means of canonical variate

Genotype	Vector 1	Vector 2	Vector 3
IR 64	-14.582	7.863	0.007
RASI	-12.746	8.643	0.594
MAS 946	-12.113	8.269	0.059
RP 5311	-9.294	9.371	-0.296
NDR 1145	-9.296	8.528	0.009
KMP 175E	-6.173	7.108	-1.321
IR 36	-14.207	7.910	-0.423
KMP 153	-7.683	7.210	-0.405
Rewa 862-27-1	-6.596	8.424	0.311
RP 5310	-7.476	7.705	-0.366
CR 2991-1-2-1-1-1	-9.419	7.204	-0.083
MGD 1104	-6.999	8.810	-0.443
MAS 9	-9.184	10.322	0.818
NDR 1143	-8.172	9.039	-0.207
GK 5022	-7.499	8.984	0.051
UPRI 2012-16	-9.855	9.702	1.327
PAU 3832-79-4-3-1	-11.834	7.335	0.487
R-RF-48	-0.336	6.275	-0.540
OR 2172-7-1	-2.796	9.243	-0.495
CB 09512	-9.699	7.482	1.844
CB 10-504	-10.646	8.162	1.743
NDR 1140	-8.451	8.569	-0.263
NDR 1159	-7.760	8.751	0.225
CR 3695-1-1	-8.888	7.612	0.947
RP 5333-5-2-B	-8.867	8.594	-0.048
UPRI 2012-15	-7.270	9.583	-0.485
CR 2658-11	-7.563	8.459	0.286
CB 09516	-8.261	7.853	1.177
Rewa 1208-15	-7.947	7.943	-0.363
Vandana(check)	-10.704	8.035	0.453

tillers per plant (13.667), panicles per plant (12.167), spikelet fertility (68.00), relative water content (77.43), critical temperature for reproductive stage, root volume (6.92), number of grains per panicle (67.67), 1000-grain weight (21.52), grain yield per plant (8.25) and grain yield per plot (1.50). Cluster IV showed lowest mean value for seedling vigor (0.46), days to 50% flowering (80.33) and plant height (101.07). Cluster VI showed

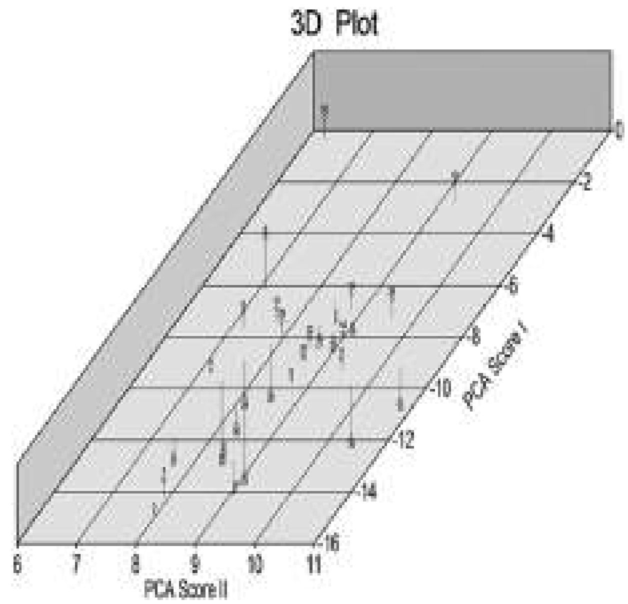

Fig. 1. Three dimensional representation of aerobic rice genotypes showing divergence based on Principal Component Analysis

highest mean value for tillers per plant(18.73), panicles per plant (16.80), spikelet fertility (81.40), chlorophyll content (43.60), relative water content (84.64), root length (12.67), root volume (9.87), grains per panicle (106.20), 1000-grain weight (25.01), grain yield per plant (3.37) and grain yield per plot (3.15) and lowest mean value for panicle length (18.73), canopy temperature (27.67) and critical temperature for reproductive stage (27.09).Clusters with desired mean value may be used in hybridization programme to achieve desired high yielding segregants (Saraswathi *et al*. 1996).

For the improvement of rice genotypes, under aerobic condition diverse parents were selected based on genetic distance, K-cluster mean for different

Table 3. Distribution of thirty genotypes into different clusters based on K-mean cluster analysis

Group K	Number of genotypes	Cluster Members
1	3	RP 5311, MAS 9, UPRI 2012-15
2	2	R- RF- 48, OR -2172-7-1
3	7	KMP 175 E, Rewa 862-27-1, MGD 1104, GK 5022, NDR 1159, UPRI 2012-15, CR2658-11
4	5	NDR 1145, CR 2991-1-2-1-1-1, CB 09512, CB 10-504, Vandana
5	8	NDR 1143, RP 5310, NDR 1143, NDR 1140, CR 2658-11, RP 5333-11, CB 09515, Rewa 1208-15
6	5	IR 64, RASI, MAS 946, IR 36, PAU 3832-79-4-3-1

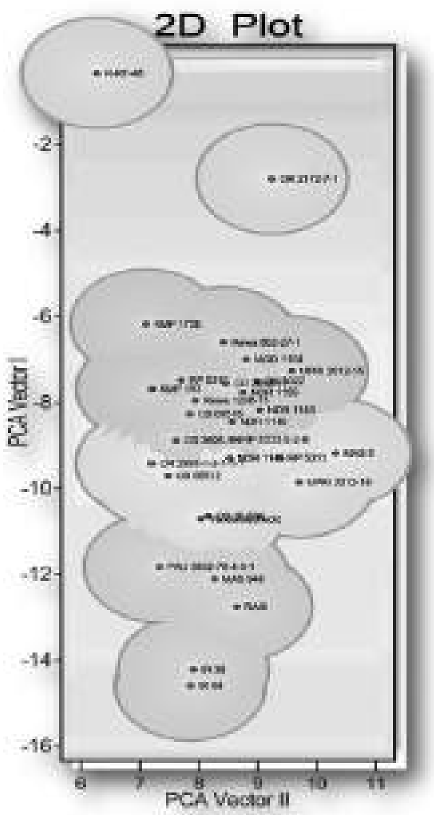

Fig. 2. Bi-plot based on first and second components for promising aerobic rice genotypes

characters as presented in Table 4. Based on cluster mean genotypes in cluster I can be selected for early maturity. Genotypes in cluster IV can be selected for vigorous seedling, early flowering and shorter plant height. Genotypes in cluster VI can be selected for more number of tillers per plant, panicles per plant, more fertile spikelet, higher chlorophyll content, relative water content, lower canopy and critical temperature for reproductive stage, longer root length, higher root volume, more number of grains per panicle and 1000-grain weight plant.

ACKNOLEDGEMENT

Authors are thankful to Directorate of Rice Research, Hyderabad, for providing genotypes of rice suitable for aerobic condition and all the members of Department of Plant Breeding and Genetics for their encouragement, support and providing necessary facilities.

Table 4. Cluster means of eighteen characters based on K-mean cluster analysis

Clusters characters	I	II	III	IV	V	VI
Seedling vigour (log)	0.551	0.875	0.562	0.458	0.550	0.301
Days to 50% Flowering	80.667	98.000	85.714	80.333	83.000	82.133
Days to Physiological maturity	99.889	112.833	105.810	102.333	104.583	102.133
Plant Height(cm)	109.444	109.000	106.667	101.067	103.167	100.800
Tillers/ Plant	15.111	13.667	15.048	17.067	15.917	18.733
Panicle Length (cm)	20.778	26.500	22.095	20.733	21.000	18.733
Panicles/ Plant	13.556	12.167	13.429	15.667	14.458	16.800
Spikelet Fertility (%)	77.222	68.000	75.000	75.067	75.042	81.400
Chlorophyll Content	39.667	34.000	39.905	41.600	40.542	43.600
Relative Water Content (%)	83.088	77.433	80.330	82.509	79.526	84.637
Canopy Temp°	28.989	31.217	29.500	28.733	28.883	27.671
Critical Temp° for Rep.stage	28.233	31.162	28.630	28.428	28.334	27.085
Root Length (cm)	9.889	7.625	9.143	10.533	9.635	12.167
Root Volume (cm³)	8.111	6.917	7.881	9.000	8.406	9.867
Grains/ Panicle	105.556	67.667	96.286	99.867	100.875	106.200
1000 Grain Weight (g)	23.556	21.517	23.774	24.307	23.848	25.005
Grain Yield/ Plant (g)	10.500	8.250	9.690	10.500	10.313	13.367
Grain Yield (kg/Plot)	2.528	1.500	2.286	2.533	2.471	3.153

REFERENCES

Abna F, Golam F and Bhassu S 2012. Estimation of genetic diversity of mungbean (*Vigna radiata L.* Wilczek) in Malaysian tropical environment. *African Journal of Microbiology* Re., 6(8):1770-1775.

Arnold CY 1960. Maximum-minimum temperatures as a basic for computing heat units. *Proc. Amer. Soc. Hort. Sci*.,76: 682-692.

Barr HD and Weatherley PE 1962. A re-examination of the relative turgidity technique for estimating water deficit in leaves.*Australian J.Biological Science*. 15: 413-428.

Das BB, Satyanarayana NH, Mukherjee S, and Sarkar KK 2014. Genetic diversity of wheat genotypes based on principal component analysis in Gangetic alluvial soil of West Bengal. *Journal of Crop and Weed*.10(2):104-107

Guei RG, Sanni KA, Abamu FJ and Fawole I 2005. Genetic diversity of rice (*Oryza sativa* L.). *Agronomie Africaine*.5: 17–28.

IRRI 1996. International network for genetic evaluation of rice: Standard Evaluation System for Rice. *International Rice Research Institute, Philippines*.

Mohamed NA 1999. Some statistical procedures for evaluation of the relative contribution for yield components in wheat. Zagazig. J. Agric. Res.26(2): 281-290.

Muhammad Ashf AQ, Khan AS, Khan SBU and Ahmad R 2012.Association of Various Morphological Traits with Yield and Genetic Divergence in Rice(*Oryza sativa*L.)*International journal of Agriculture & biology*.1560–8530

Sanni KA, Fawole I, Guei RG, Ojo DK, Somado EA, Sanchez I, Ogunbayo SA and Tia DD 2008. Geographical patterns of phenotypic diversity in *Oryza sativa* landraces of Côte d'Ivoire. *Euphytica* 160: 389–400.

Saraswathi R, Hepziba SJ and Palanisamy S 1996. Combining ability of diverse genotypes in rice. *Crop Res*. 11(2):194-199.

Sharma S. 1996. Applied multivariate techniques. *1nd ed. John Wiley and Sons*, New York. 493 pp.

Quarrie SA, Stojanovic J, Pekic S 1999. Improving drought tolerance in small-grain cereals: A case study, progress and prospects. *Plant Growth Regulation*. 29: 1-21.

Richards RA 1996. Defining selection criteria to improve yield under drought. *Plant Growth Regulation*. 20: 157-166.

Tadesse W, Bekele E 2001. Factor analysis of yield in grasspea (*Lathyrus sativus* L.). *Lathyrus Lathyrism Newsletter*. 2: 416-421.

Talebi R 2011. Evaluation of chlorophyll content and canopytemperature as indicator for drought tolerance in durum wheat (T.durum Desf.). Australian J. Basic *and Applied Science*. 5(11): 1457-1462.