



PRINCIPAL COMPONENT ANALYSIS FOR AGRO-MORPHOLOGICAL AND QUALITY CHARACTERS IN GERMPLASM OF RICE (*Oryza sativa* L.).

^aRavi Yugandhar P., ^bSuneetha Kota ^cUsha Kiran, B. and ^dSridhar, M.

^aAgricultural College, Bapatla-522101, Acharya N.G. Ranga Agricultural University, Andhra Pradesh, INDIA

^bICAR-Indian Institute of Rice Research, Rajendra Nagar-500030, Hyderabad, Telangana, INDIA

^cICAR-Indian Institute of Oil Seeds Research, Rajendra Nagar-500030, Hyderabad, Telangana, INDIA

^dAgricultural college, PJTSAU, Rajendra Nagar, Hyderabad-500030, Telangana, INDIA

Corresponding author email- raviyugandhar@gmail.com

ABSTRACT

The present investigation was carried out to determine the genetic diversity among ninety five rice germplasm lines along with six checks by using principal component analysis. Principal component analysis was utilized to examine the variation and to estimate the relative contribution of various traits for total variability. There are six axes which accounted for 71.37% cumulative variance of the total variability for twenty agro-morphological and quality traits. PC1 accounted 23.48% of the total variability contributed by the traits like amylose content, days to maturity, days to 50% flowering, total grains per panicle, filled grains per panicle, grain weight per panicle, elongation ratio and chaffy grains per panicle. PC2 accounted 12.45% of the total variation and the traits *viz.* total grains per panicle, filled grains per panicle, chaffy grains per panicle, kernel breadth and alkali spreading value contribute to the variation. Component 3 had the contribution from the characters like chaffy grains per panicle, grain yield per plant, kernel length, length/breadth ratio, total grains per panicle, alkali spreading value, water uptake and filled grains per panicle which accounted for 10.62% of the total variation. Grain quality characters like kernel length after cooking and elongation ratio had contributed 9.97% of the total variation in PC4. PC5 and 6 accounted 8.03% and 6.82% of the total variability respectively and contributed by the traits like spikelet fertility, filled grains per panicle, ear bearing tillers per plant, total grains per panicle and alkali spreading value. Thus, the results revealed vast genetic variability exists in the studied germplasm lines and can be used for various breeding programmes for improvement in yield and quality.

KEY WORDS: Principal Component Analysis, Rice, Genetic variability, Germplasm.

INTRODUCTION

Rice (*Oryza sativa* L.) is the important staple food crop for more than half of the world's population. About 90% of the world's rice is grown and consumed in Asia, whereas 50% of the population depends on rice for food (Tenorio *et al.*, 2013). In India, rice accounts for more than 43% of food grain production. It is cultivating in 44.8 million hectare under four main ecosystems *viz.* irrigated, rainfed lowland, rainfed upland and flood-prone with an average annual production of 100 million tons (Song *et al.*, 2007). By 2030, 40% raise in production of rice is required in view of the escalation rate of the world population and food security with reduction of arable land in the world (Khush, 2005). The success of any crop improvement programme is highly dependent on the efficient manipulation 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 and the association among them are essential for shaping breeding strategies, classification of parental lines, defining the heterotic groups and to predict the future hybrid performance (Acquaah, 2012). Morphological markers have played an essential role in crop improvement since the beginning of modern breeding programmes (Mignouna *et al.*, 1996).

Statistical process of categorization is generally by multivariate methods as it has wide use in summarizing and describing the innate discrepancy among the genotypes. Principal Component Analysis (PCA), Cluster analysis and discriminate analysis are the important multivariate analysis methods (Oyelola, 2004). Cluster analysis is concerned with classifying earlier unclassified materials, whereas PCA can be used to find out the resemblance between the variables and classify the genotypes (Leonard and Peter, 2009). PCA may be used to disclose the patterns and eradicate redundancy in data sets as variations regularly arise in crop species for yield and grain quality (Maji and Shaibu, 2012). The aim of PCA is to dig up the key information from the table, to signify it as a set of novel orthogonal variables called principal components and to exhibit the blueprint of similarity of the observations and of the variables as points in maps. 'Proper values' compute the weight and role of each component to total variability, while each coefficient of proper vectors indicates the extent of contribution of each original variable with which each principal component is associated. The superior the coefficients, apart from of the direction, the more efficient they will be in discerning among the accessions. The PCA condense the magnitude of a multivariate data to a small number of principal axes, generates an Eigen vector for every axes and produces

component scores for the traits (Sneath *et al.*, 1973 and Ariyo & Odulaja, 1991). PCA has been used by various workers like Maji & Shaibu (2012), Gana *et al.* (2013), Asfaq *et al.* (2014), Ravikumar *et al.* (2015) and Kumar *et al.* (2015) for characterization different rice germplasm lines.

Therefore, bearing in mind the value of PCA, 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.

MATERIALS & METHODS

Experimental material for characterization of germplasm for agro-morphological and quality characteristics consisted of hundred and one rice germplasm collections. The collection includes indigenous cultures from various districts of Assam state, Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Jammu & Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Manipur, Odisha, Rajasthan, Tamil Nadu, Uttaranchal, Uttar Pradesh and West Bengal maintained at ICAR-Indian Institute of Rice Research (ICAR-IIRR), Hyderabad (Table 1).

TABLE 1: List of genotypes used in the study

.No.	Germplasm line	Common name	Origin (District)
1	SR1	Joha Bora	Sivasagar (Assam)
2	SR2	Ranga Bora	Golaghat (Assam)
3	SR3	Sunga Bora	Golaghat (Assam)
4	SR4	Noldong Bora	Nagaon (Assam)
5	SR5	Tegori Bora	Kamrup (Assam)
6	SR6	Bongari Bora	Kamrup (Assam)
7	SR7	Kola Ampathi Bora	Sivasagar (Assam)
8	SR8	Bora-1	Tinsukia (Assam)
9	SR9	Dadhara Bora	Morigaon (Assam)
10	SR10	Chokura Bora	Kamrup (Assam)
11	SR11	Sakoi bhanu Bora	Darrang (Assam)
12	SR12	Kola Bora	Sonitpur (Assam)
13	SR13	Misiri Chokua	Darrang (Assam)
14	SR14	Boka Chokua	Dhubri (Assam)
15	SR15	Bora- Chokua	Goalpara (Assam)
16	SR16	Kagori- Chokura	Goalpara (Assam)
17	SR17	Kola Boka Chokura	Dibrugarh (Assam)
18	SR18	Haru Chokua	Dhemaji (Assam)
19	SR19	Boga Chokua	Kokrajhar (Assam)
20	SR20	Lahi Chokua	Golaghat (Assam)
21	SR21	Sam Chokua	Sivasagar (Assam)
22	SR22	Maju Chokua	Golaghat (Assam)
23	SR23	Ham Chokua	Sonitpur (Assam)
24	SR24	Hampori Chokua	Dibrugarh (Assam)
25	SR25	Agnoni Bora	Dibrugarh (Assam)
26	SR26	Bogali Bora	Dibrugarh (Assam)
27	ASG73	Seetabhog	Darjeeling (WB)
28	ASG1	Shukla Phool	Chhattisgarh
29	ASG33	Maguraphulla	Odisha
30	ASG30	RAU 3043	Bihar
31	ASG9	Barang	Raipur (Chhattisgarh)
32	ASG200	Dubraj (Raipur)	Deogarh (Odisha)
33	ASG12	Kali Muchh	Gwalior, Bhind (M.P)
34	ASG138	Munibhog	Chhattisgarh
35	ASG36	Krushnabhoga	Nimapara (Odisha)
36	ASG193	Champaran Basmati 2	Bihar
37	ASG8	Chini Kapoor	Raigarh, Bastar (Chhattisgarh)
38	ASG68	Thakurabhog	Puri (Odisha)
39	ASG71	Parbatjira	Odisha
40	ASG191	Bhanta Phool B	Sidhi (M.P)
41	ASG204	Kala Joha	Rajasthan
42	ASG47	Barijunja	Odisha
43	ASG148	Ganjo	Rajnandgaon (MP)
44	ASG35	Deulabhog	Odisha
45	ASG107	Pimpdibasa	Keonjhar (Odisha)
46	ASG44	Nalidhan	Odisha
47	ASG62	Basuabhava	Odisha
48	ASG5	Til Kasturi	Chhattisgarh
49	ASG38	Tulasiphulla	Puri (Odisha)
50	ASG58	Gatia	Odisha

51	ASG110	Chhabiswa	Odisha
52	ASG32	RAU 3044	Bihar
53	ASG87	Kankjeer	Barabanki (UP)
54	ASG20	Sitabhog	West Bengal
55	ASG90	KB-13	Uttar Pradesh
56	ASG203	Kamod	Bundi, Baran, Hanumangarh (Rajasthan)
57	ASG43	Dhoiabankoi	Odisha
58	ASG70	Kalikati	Kalahandi (Odisha)
59	ASG137	Bayasa Bhog	Chhattisgarh
60	ASG199	Bor Joha 1	Assam
61	ASG78	Randhumpaughal	West Bengal
62	ASG54	Kalanamak (Birdpur)	Basti, Sidarthnagar, Maharajgamj, Gorakhpur, Gonda (UP), Blarampur (WB)
63	ASG103	Kalajauvan	Odisha
64	ASG39	Neelabati	Odisha
65	ASG55	Basnadhan	Odisha
66	ASG162	Sonth	Shahdol (MP)
67	ASG14	Moongphali - B	Ghajipur (UP)
68	ASG24	Badshaha	Uttaranchal, Uttar Pradesh
69	ASG60	Barangamali	Odisha
70	ASG67	Muhulakuchi	Odisha
71	ASG53	Nagri Dubraj	Odisha
72	ASG182	Jhingisiali	Odisha
73	ASG6	Bans patri	Vidarbha (MR)
74	ASG85	Jiraphool	Chhattisgarh
75	ASG26	RAU 3049	Bihar
76	ASG130	Kalanamak 1	Basti, Sidarthnagar, Maharajganj, Gorakhpur, Gonda(UP), Balarampur (WB)
77	ASG86	RAU 3056	Bihar
78	ASG111	Kheerasai	Odisha
79	ASG69	Basnasapuri	Odisha
80	ASG82	Gangabarud	Bastar (CG)
81	ASG190	Bhanta Phool A	Sidhi (M.P)
82	ASG83	Atmashital	Bastar (CG)
83	ASG66	Kalakanhu	Odisha
84	ASG113	IGSR3-1-5	Chhattisgarh
85	ASG77	Karnal local-B	Haryana
86	ASG81	Sonachoor	Bhojpur, Rothas (Bihar)
87	ASG195	Champaran Basmati 4	Bihar
88	ASG63	Khosakani	Odisha
89	ASG22	Dhaniya-B2	Basti, Gorakhpur, Gonda (UP)
90	ASG52	Seetakeshari	Odisha
91	ASG93	RAU 3048	Bihar
92	ASG167	Sheetalkani	Odisha
93	ASG49	Jaiphulla	Odisha
94	ASG10	Bishnu bhog	Chhattisgarh
95	ASG104	Kalajeevan	Odisha
96	SWARNA	Released variety	-
97	IR64	Released variety	-
98	VASUMATI	Released variety	-
99	KASTURI	Released variety	-
100	JAYA	Released variety	-
101	BPT5204	Released variety	-

Six checks were included viz. Swarna, Jaya, IR 64, Vasumati, Kasturi and BPT 5204 for combination of yield and different quality characters. The experiment was laid in augmented randomized complete block design for ninety five germplasm lines along with six checks in Rabi, 2014-15. Checks were replicated in each block. All genotypes were sown in nursery beds and transplanted to field after 30 days after the germination consisting of 25 plants each with spacing of 20 X 20cm. After transplanting, 5-7 cm of standing water was maintained in the field until draining before harvest. The recommended dose of fertilizers @

100: 50: 50 kg N: P: K/ha was applied. The full dose of P₂O₅ and K₂O and half dose of nitrogen were applied as basal dose at the time of transplanting and rest of the nitrogen was top dressed in two split doses at the time of maximum tillering stage and between panicle initiation and flowering.

Ten competitive plants of each genotype were selected randomly and data collected on phenotypic characters like days to fifty percent flowering, days to maturity, panicle length, total grains per panicle, filled grains per panicle, chaffy grains per panicle, spikelet fertility, ear bearing

tillers per plant, grain weight per panicle, test weight and grain yield. Analysis for grain quality traits like kernel length, kernel breadth, length/breadth ratio, water uptake, kernel length after cooking, elongation ratio, alkali spreading value, amylose content and gel consistency was done as per IIRR standard protocols (IIRR, 2013). Analysis for principal components, Eigen values, Eigen vectors and 2D biplot between PC1 & 2 were done by using R-software-3.4.3 (available in <http://cran.r-project.org>).

RESULTS & DISCUSSION

The results of PCA explained the genetic variation among the genotypes for all agro-morphological and quality

characters under study. Data were considered in each component with Eigen values more than 1 as per the suggestions given by Brejda *et al.* (2000), which determines as a minimum 10% of the variation. Superior Eigen values are considered as best attributes in principal components. In our study, six components exhibited Eigen values of >1 and showed cumulative variation of 71.37%. It indicates that the identified characters within these components exhibited immense influence on the phenotype of the genotypes. Table 2 presents principal components, Eigen values and percentage contribution of each component to the total variation in the rice germplasm.

TABLE 2: Eigen values, per cent variance and cumulative variance values of rice germplasm

	PC1	PC2	PC3	PC4	PC5	PC6
Eigen value	4.92	2.48	2.21	1.80	1.45	1.32
Total Variance (%)	23.48	12.45	10.62	9.97	8.03	6.82
Cumulative Variance (%)	23.48	35.93	46.56	56.53	64.55	71.37
Trait	Eigen vectors					
DFP	0.537	-0.116	0.267	-0.608	-0.144	0.277
PL	-0.015	0.198	-0.386	0.017	0.078	-0.436
DM	0.608	-0.099	0.279	-0.570	-0.132	0.167
FGP	0.497	0.626	0.338	0.114	0.433	-0.115
CG	0.435	0.460	0.572	0.226	-0.343	-0.244
TGP	0.515	0.634	0.393	0.142	0.339	-0.143
SF	-0.122	0.068	-0.397	-0.185	0.789	0.296
EBT	0.216	0.016	-0.028	-0.275	0.381	0.017
GWP	0.487	-0.395	0.169	0.109	0.123	0.133
TW	-0.729	0.279	-0.083	-0.062	-0.047	0.161
GY	-0.347	-0.156	0.519	-0.038	0.162	0.161
KL	-0.621	-0.466	0.462	0.196	0.130	-0.125
KB	-0.679	0.426	0.018	-0.065	-0.204	0.291
LBR	-0.114	-0.716	0.417	0.239	0.254	-0.295
WU	-0.809	0.150	0.364	0.014	0.014	0.076
KLAC	-0.030	-0.013	-0.232	0.496	0.078	0.062
ER	0.439	0.158	-0.339	0.477	-0.204	0.248
ASV	-0.482	0.322	0.392	0.090	0.029	0.351
AC	0.769	-0.263	-0.015	0.248	-0.041	0.208
GC	-0.287	0.133	-0.139	-0.504	-0.048	-0.598

The six components viz. PC1, PC2, PC3, PC4, PC5 and PC6 showed 23.48%, 12.45%, 10.62%, 9.97%, 8.03% and 6.82% of variations among the characters respectively. Similar results were reported by Mahendran *et al.* (2015) and Ojha *et al.* (2017). Only well loaded characters in each component values within 10% of the highest factor loading were retained for further explanation. Results revealed by rotated component matrix showed that the PC1 which accounted for the maximum variability (23.48%) and highly loaded with characters such as amylose content (0.769), days to maturity (0.608), days to fifty percent flowering (0.537), total grains per panicle (0.515), filled grains per panicle (0.497), grain weight per panicle (0.487), elongation ratio (0.439) and chaffy grains per panicle (0.435) contributed in positive direction whereas water uptake (-0.809), test weight (-0.729), kernel breadth (-0.679) and kernel length (-0.621) contributed in negative direction. It clearly indicated that the variation in PC1 is mainly contributed by yield characters except test weight. PC2 accounted 12.45% of the total variation and loaded with the traits viz. total grains per panicle (0.634), filled grains per panicle (0.626), chaffy grains per panicle (0.460), kernel breadth (0.426) and alkali spreading value

(0.322). PC3 had the contribution from the characters like chaffy grains per panicle (0.572), grain yield per plant (0.519), kernel length (0.462), length/breadth ratio (0.417), total grains per panicle (0.393), alkali spreading value (0.392), water uptake (0.364) and filled grains per panicle (0.338) which accounted for 10.62% of the total variation. It clearly showed that the variation in this component is contributed by the combination of yield and quality characters. Grain quality characters like kernel length after cooking (0.496) and elongation ratio (0.477) had contributed 9.97% of the total variation in PC 4. PC5 is accounted for 8.03% of total variation and it was loaded by the traits such as spikelet fertility (0.789), filled grains per panicle (0.433), ear bearing tillers per plant (0.381) and total grains per panicle (0.339) in which the quality characters were not responsible for contribution of the variability. PC6 is contributed 6.82% of the total variation which is loaded by alkali spreading value (0.351). Scree plot showed in the Fig.1 explained the percentage of variation associated with each principal component obtained by drawing a graph between Eigen values and principal component numbers. PC1 showed 23.48% variability with the Eigen value of 4.92. The Eigen values

are gradually declined from PC1 to PC6. The Eigen values are 2.48, 2.21, 1.80, 1.45 and 1.32 for PC2, PC3, PC4, PC5 and PC6 respectively. Elbow type line is obtained after PC6 tended to straight with minute difference observed in each PC and it is clearly showed that the utmost variation was observed in PC1. Fig. 2 showed the distribution of germplasm lines accounted by different variables from component 1 and 2. The loading plot

depicted that the traits such as water uptake, kernel length, kernel breadth, test weight, amylose content, L / B ratio and grain weight per panicle showed high degree of variation compared to others. Comparable studies and results in rice by the researchers like Nachimuthu et al. (2014), Gour *et al.* (2015), Kumar *et al.* (2015) and Ojha *et al.* (2017).

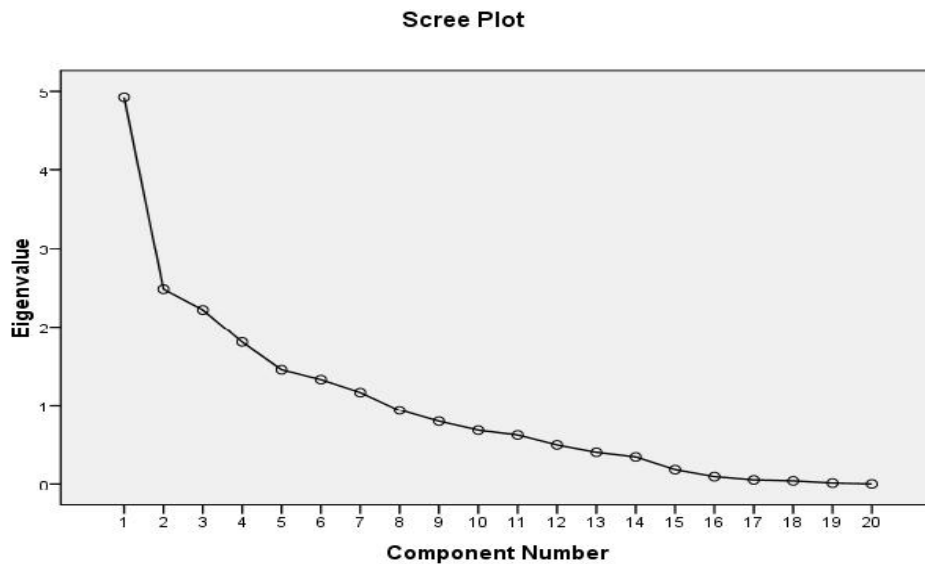


FIGURE 1: Scree plot of principal component analysis of rice germplasm between Eigen values and principal components

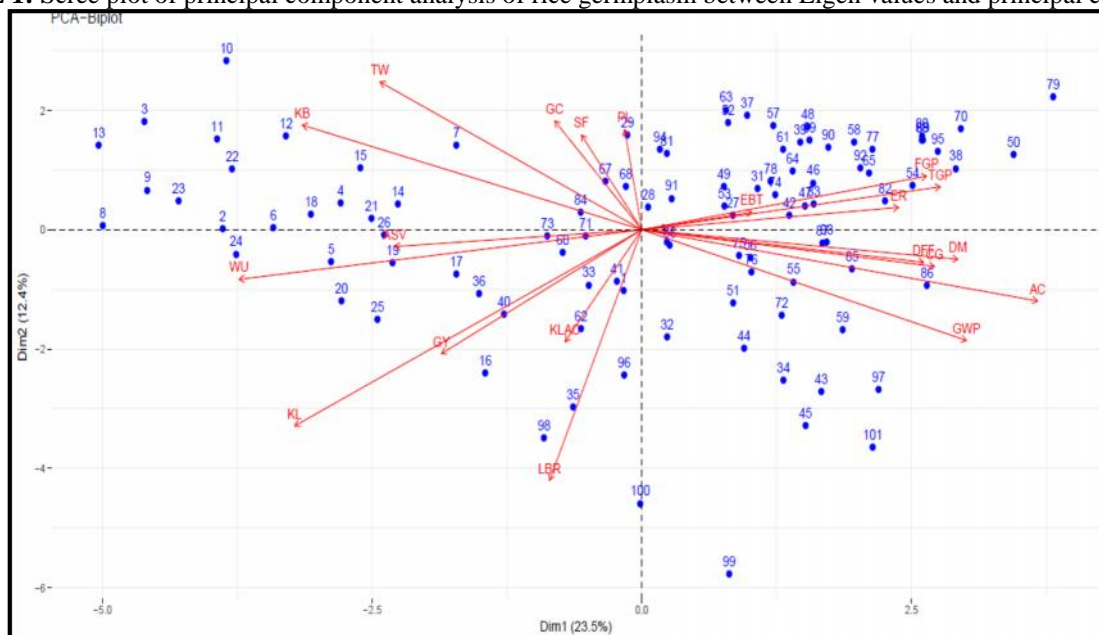


FIGURE 2: Distribution of germplasm accessions across first two components based on PCA

CONCLUSION

Principal component analysis was utilized to examine the variation and to estimate the relative contribution of various traits for total variability. Results of the study revealed that there is a large quantity of variability in the rice germplasm. PCA identified only some characters that plays prominent role in classifying the variation existing in the germplasm. The results of the PCA revealed that the 71.37% of the total variability was explained by the first six principal components. Agro-morphological traits related to yield are used as a preliminary evaluation tool due to their easiness and the improvement in the yield will

be the prime goal of any crop improvement programme. Present days yield as well as quality is very essential in rice. Therefore, we included grain quality characters along with the agro-morphological traits to study variability in the rice germplasm. Greater level of variability existing in the varieties and the characters will craft the scope for additional enhancement of the cultivars in crop improvement programmes in rice.

ACKNOWLEDGEMENT

The authors are very much thankful to The Director, ICAR- Indian Institute Rice Research (formerly DRR),

Hyderabad for providing all the facilities required in the present studies.

REFERENCES

- Acquaah, G. (2012) Principles of plant genetics and breeding. Second edition. Wiley-blackwell publications. 127-129.
- Ariyo, Q.J. and Odulaja, A. (1991) Numerical analysis of variation among accessions of Okra. *Abelmoschus esculentus*. Ann Bot. 67, 527-31.
- Ashfaq, M., Saleem Haider, M., Ali, A., Ali, M., Hanif, S. and Mubashar, U. (2014) Screening of diverse germplasms for genetic studies of drought tolerance in rice (*Oryza sativa* L.). *Caryologia*, 67(4), 296-304.
- Brejda, J.J., Moorman, T.B., Karlen, D.L., Dao, T.H. (2000) Identification of regional soil quality factors and indicators. I. Central and Southern High- Plains. Soil Sci. Soc. Am. J. 64, 2115-2124.
- Gana, A.S., Shaba, S.Z. and Tsado, E.K. (2013) Principal component analysis of morphological traits in thirty-nine accessions of rice (*Oryza sativa*) grown in a rainfed lowland ecology of Nigeria. *J. Plant Breed. Crop Sci.* 5, 120-126.
- 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. *IJCS*, 5(4), 975-983.
- Khush, G.S. (2005) *Plant Mol Biol.* 59, 1-6.
- Kumar, S., Dwivedi, S.K., Singhm, S.S., Jha, S.K., Lekshmy, S. and Elanchezhian, R. (2015) Identification of drought tolerant rice genotypes by analyzing drought tolerance indices and morpho-physiological traits. *SABRAO J. Breed. Gene.* 46(2), 217-230.
- Laboratory manual on Rice Grain Quality, September, (2013) Indian Institute of rice research, Rajendranagar, Hyderabad.
- 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, 87-93.
- Mignouna, H.D., Fatokun, C.A. and Thottappily, G. (1996) Choice of DNA marker system in DNA assisted improvement of the stable crops of sub-Saharan Africa. Proceedings of the workshop on DNA markers at IITA, Ibadan, Nigeria. 9-15.
- 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 panel by principal component analysis. *Indian J Sci Technol.* (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, *IJCS*. 5(5), 1875-1878.
- Oyelola, B.A. (2012) The Nigerian Statistical Association preconference workshop, University of Ibadan. 20-21.
- Ravikumar, B.N.V.S.R., Kumari, P.N., Rao, P.V.R., Rani, M.G., Satyanarayana, P.V., Chamundeswari, N., Vishnuvardhan, K.M., Suryanarayana, Y., Bharatha lakshmi, M. and Reddy, A.V. (2015) Principal component analysis and character association for yield components in rice (*Oryza sativa* L.) cultivars suitable for irrigated ecosystem. *Current Biotica*, 9(1). 25-35.
- Sneath Peter. H.A. and Sokal Robert, R. (1973) Numerical Taxonomy: The Principles and Practice of Numerical Classification.
- Song, X.J., Huang, W., Shi, M., Zhu, M.Z. and Lin, H.X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING- type E3 Ubiquitin ligase. *Nat. Genet.* 39, 623-630.
- Tenorio, F.A., Ye, C., Redona, E., Sierra, S. and Laza, M. (2013) Screening rice genetic resources for heat tolerance. *SABRAO J. Breed. Gene.* 45(3), 371-381.