Journal of Pharmacognosy and Phytochemistry

Available online at www.phytojournal.com

B Kusuma Kumari
Agricultural College, Bapatla, Andhra Pradesh, India

BNVSR Ravi Kumar
Regional Agricultural Research Station, Maruteru, Andhra Pradesh, India

Jyothula DPB
Agricultural College, Bapatla, Andhra Pradesh, India

N Mallikarjuna Rao
Regional Agricultural Research Station, Maruteru, Andhra Pradesh

Diversity analysis in rice breeding lines for yield and its components using principal component analysis

B Kusuma Kumari, BNVSR Ravi Kumar, Jyothula DPB and N Mallikarjuna Rao

DOI: https://doi.org/10.22271/phyto.2021.v10.i1m.13451

Abstract
One hundred and nineteen rice breeding lines with two local checks were subjected to the principle component analysis (PCA) and cluster analysis to estimate the existing genetic diversity for yield contributing characters. The first three principal components having Eigen value more than one are cumulatively contributing 68.69% to the total variability. PC1 has the contribution from the traits days to 50% flowering (0.497), days to maturity (0.484) and ear bearing tillers (0.359) which accounted 31.84% of total variability indicating these traits contributed more to the total variance. Cluster analysis revealed that the rice lines were classified into 12 divergent clusters by both PCA and Tocher's method. Among the 12 clusters, cluster 1 had highest number of breeding lines (18) and cluster 11 had least number of lines (2) in PCA cluster analysis whereas in Tocher's method highest number of lines observed in cluster 5 (34) followed by cluster 1(33). This analysis reveals the presence of wide genetic variance in rice breeding lines.

Keywords: Diversity analysis, principal component analysis, rice, rice breeding lines

Introduction
Rice is an important food crop for majority of global population particularly in Asia. Human population is increasing day by day but the crop production is lagging behind. So in order to meet the food requirement of growing population, development of high yielding varieties is essential. The success of any breeding programme depends on the selection of parents and the parents involved in the development of varieties should be divergent. Hence there is a need to measure diversity among population.

Multivariate methods are extensively used in summarizing and describing the inherent variability among different rice genotypes. Multivariate statistical tools include Principal Component Analysis (PCA), Cluster analysis and Discriminate analysis (Oyelola, 2004) [5].

Principal Component Analysis (PCA) is the major tool used for analyzing the inherent genetic variation. This analysis helps in identification of characters that help in separating selected genotypes based on similarities in one or more characters and classify the genotypes into separate groups (Ariyo, 1987, Nair *et al.*, 1998, Ravi Kumar *et al.*, 2015, and Sudeepthi *et al.*, 2019) [1, 4, 6, 8]. Genetic diversity analysis helps the breeder in selection of divergent parents for improving several traits.

Hence, the present investigation was undertaken with the objective of determining the amount of genetic diversity present in rice breeding lines to identify and classify variation by using yield components and relationship between them.

Materials and Methods
The experiment was carried out with 121 rice breeding lines (Table 1) in simple lattice design with two replications during *kharif* 2016 at RARS, Maruteru, Andhra Pradesh. The experiment was conducted with the recommended package of practices for healthy plant growth. Observations were recorded on nine quantitative traits *viz.*, days to 50% percent flowering, days to maturity, plant height (cm), number of ear bearing tillers per plant, panicle length (cm), number of grains per panicle, spikelet fertility (%), grain yield per plant (g), test weight (g). The data was subjected to PCA and cluster analysis using the methodology given by Gomez and Gomez (1984) [9]. The PCA analysis limits the dimensions of a multivariate data to a few principal axis, generates an Eigen vector for each axis and produces component scores for the traits. (Sneath *et al.*, 1973 and Ariyo *et al.*, 1991) [7, 2].

Corresponding Author:
BNVSR Ravi Kumar
Regional Agricultural Research Station, Maruteru, Andhra Pradesh, India

Results and Discussion
Analysis of variance indicated that the traits under study were highly significant. The distribution of all the lines under study is presented in Figure 1.

Principle component analysis
The Eigen values, percent contribution of variance and variable loading of 4 PCs were presented in Table 2. The first 3 PCs having >1 Eigen value contributing 68.69% cumulative variance to total variability in rice lines for yield contributing characters in the study. The relative discriminating power of the PCA as revealed by the Eigen values was high in PC1 (2.866) and lower in PC4 (0.961).

The PC1 added highly 31.848% to the total variability with significant loading of days to 50% flowering (0.497), days to maturity (0.484) and ear bearing tillers (0.359) which were positively correlated. Hence, it can be said that these characters contributed maximum to the total variance. The PC 2 contributed 19.553% to the total variance having positive correlation with total grains per panicle (0.417), spikelet fertility (0.395), ear bearing tillers (0.338), which were loaded significantly. Plant height (0.668), panicle length (0.453), grain yield (0.326), days to 50% flowering (0.303) are the traits majorly contributing to variability created by PC 3 which accounts 17.29% to total variance.

Days to maturity in PC 1, Days to 50% flowering in PC 1 and PC 3, ear bearing tillers per plant in PC 1 and PC 2, total grains per panicle and spikelet fertility in PC 2, plant height, panicle length, grain yield per plant in PC 3 have more positive weights in respective principal components indicating their contribution towards total diversity and selection for these traits will be rewarding.

Three PCs extracted from the nine traits studied by PCA were used for clustering genotypes (Table 3). The similar genotypes were classified in to the same cluster based on their various yield and yield attributing traits.

Cluster analysis
One hundred and twenty one rice lines were grouped into 12 different clusters through PCA analysis and the mean performance of different clusters is given in the Table 4. MTU 2330-180-1-1-1(17), MTU 2330-187-1-1-1 (20) and MTU 2409-6-3-2-1(70) lines in cluster 3 exhibited higher days to 50% flowering, ear bearing tillers and days to maturity. Cluster 8 is having 6 lines (MTU 2395-14-5-2-1(39), MTU 2463-4-1-1-1(79), MTU 2463-4-1-2-1(80), MTU 2463-14-1-1-1(81), MTU 2465-6-2-1-1(82), MTU 2491-39-2-1-1(101)) with higher mean for panicle length and total grains per panicle. Nine breeding lines were grouped in cluster 9 with higher mean performance for plant height and grain yield. MTU 2398-5-3-1-2(43), MTU 2489-11-2-1-1(96) were grouped into cluster 11 having higher average value for test weight. Cluster 1 is exhibiting higher spikelet fertility. Fifteen lines were grouped into cluster 5 having average performance for all the traits under study.

One hundred and twenty one lines were subjected to cluster analysis using both Tocher's and Ward's method. In both the methods all the lines were grouped into 12 clusters. In Tocher's method they formed 7 minor clusters (2,6,8,9,10,11 and 12) each having only one line and 5 major clusters includes cluster 5 with 34 lines, cluster 1 is having 33 lines, cluster 4 (21 lines), cluster 3 (18 lines) and cluster 7 (8 lines). (Fig.2).

In Ward's method, all the lines were grouped into two major clusters and again these two clusters are having sub clusters (Fig.3). One major cluster includes 8 sub clusters with 93 lines and another major cluster is having 4 sub clusters with the grouping of 28 breeding lines.

The 121 rice breeding lines were clustered into 12 groups based on the dendrograms of cluster analysis and PCA. The groupings of hierarchical cluster analysis exhibited a similar dendrogram topology and cluster membership to that produced using PCA analysis, thereby confirming the accuracy of the constructed dendrogram. In the similar way Khatun *et al.* (2015) [3], obtained clustering pattern for PCA and hierarchical cluster analysis for 43 upland rice genotypes.

Table 1: Code and pedigree of 119 breeding lines with 2 local checks

Code	Genotype Pedigree	Code	Genotype Pedigree	Code	Genotype Pedigree
1	MTU 2263-34-2-1-1	42	MTU 2398-5-3-1-1	82	MTU 2465-6-2-1-1
2	MTU 2263-39-1-1-1	43	MTU 2398-5-3-1-2	83	MTU 2465-11-3-1-1
3	MTU 2263-50-1-1-1	44	MTU 2398-5-3-2-1	84	MTU 2465-12-2-2-1
4	MTU 2263-60-2-1-1	45	MTU 2398-7-3-1-1	85	MTU 2465-16-1-1-1
5	MTU 2263-63-1-1-1	46	MTU 2399-1-2-1-1	86	MTU 2465-16-2-1-1
6	MTU 2263-63-2-1-1	47	MTU 2399-3-1-1-1	87	MTU 2468-25-2-1-1
7	MTU 2263-65-2-2-1	48	MTU 2399-7-1-1-1	88	MTU 2468-31-1-1-1
8	MTU 2263-67-1-1-1	49	MTU 2399-21-4-1-1	89	MTU 2469-6-1-1-1
9	MTU 2263-69-1-1-1	50	MTU 2399-24-1-1-1	90	MTU 2469-6-3-2-1
10	MTU 2263-89-2-1-1	51	MTU 2399-25-2-1-1	91	MTU 2469-14-1--1-1
11	MTU 2321-156-1-1-1	52	MTU 2399-37-2-1-1	92	MTU 2469-42-4-1-1
12	MTU 2321-158-3-1-1	53	MTU 2399-38-1-2-1	93	MTU 2469-55-2-2-1
13	MTU 2323-70-1-2-1	54	MTU 2401-7-1-1-1	94	MTU 2469-57-2-1-1
14	MTU 2324A-152-1-1-1	55	MTU 2401-7-1-2-1	95	MTU 2469-68-2-1-1
15	MTU 2324A-152-1-2-1	56	MTU 2403-8-4-1-1	96	MTU 2489-11-2-1-1
16	MTU 2328-199-2-2-1	57	MTU 2404-3-1-1-1	97	MTU 2489-14-2-1-1
17	MTU 2330-180-1-1-1	58	MTU 2404-8-2-1-1	98	MTU 2489-14-2-1-2
18	MTU 2330-183-2-2-1	59	MTU 2404-18-2-1-1	99	MTU 2491-5-1-1-1
19	MTU 2330-186-1-1-1	60	MTU 2404-28-2-1-1	100	MTU 2491-19-1-1-1
20	MTU 2330-187-1-1-1	61	MTU 2404-31-2-1-1	101	MTU 2491-39-2-1-1
21	MTU 2330-187-1-2-1	62	MTU 2408-1-3-1-1	102	MTU 2491-46-2-1-1
22	MTU 2331-210-1-1-1	63	MTU 2408-2-5-1-1	103	MTU 2491-52-2-1-1
23	MTU 2331-239-2-1-1	64	MTU 2408-3-2-1-1	104	MTU 2491-57-1-1-1
24	MTU 2336A-11-1-1-1	65	MTU 2408-7-2-1-1	105	MTU 2491-63-1-1-1

25	MTU 2336A-11-2-2-1	66	MTU 2408-11-2-1-1	106	MTU 2491-65-1-1-1
26	MTU 2337-227-2-1-1	67	MTU 2408-29-1-1-1	107	MTU 2491-66-1-1-1
27	MTU 2386-11-1-1-1	68	MTU 2408-44-2-1-1	108	MTU 2491-68-2-1-1
28	MTU 2386-15-1-1-1	69	MTU 2409-6-3-1-1	109	MTU 2491-69-1-1-1
29	MTU 2386-18-1-1-1	70	MTU 2409-6-3-2-1	110	MTU 2491-70-1-1-1
30	MTU 2386-18-1-2-1	71	MTU 2409-9-1-1-1	111	MTU 2491-70-2-1-1
31	MTU 2386-21-1-1-1	72	MTU 2410-13-2-1-1	112	MTU 2492-6-3-1-1
32	MTU 2386-28-2-1-1	73	MTU 2410-13-2-2-1	113	MTU 2492-15-2-2-1
33	MTU 2386-35-2-1-1	74	MTU 2413-5-1-1-1	114	MTU 2492-15-3-1-1
34	MTU 2387-2-2-1-1	75	MTU 2462-31-2-1-2	115	MTU 2492-25-2-1-1
35	MTU 2391-23-3-1-1	76	MTU 2462-34-1-1-2	116	MTU 2494-21-2-1-1
36	MTU 2391-23-2-1-1	77	MTU 2462-45-1-2-1	117	MTU 2494-19-1-2-1
37	MTU 2391-32-2-1-1	78	MTU 2462-49-1-1-1	118	MTU 2495-9-2-2-1
38	MTU 2391-32-4-1-1	79	MTU 2463-4-1-1-1	119	MTU 2495-9-3-1-1
39	MTU 2395-14-5-2-1	80	MTU 2463-4-1-2-1	120	MTU 7029
40	MTU 2395-14-5-3-1	81	MTU 2463-14-1-1-1	121	MTU 1075
41	MTU 2398-4-5-1-1				

Table 2: Eigen vectors and eigen values of the four principle components of 9 characters

Variables	1 Vector	2 Vector	3 Vector	4 Vector
Days to 50% flowering	0.497	0.250	0.303	0.058
Plant height (cm)	-0.060	-0.131	0.668	0.398
Ear bearing tiller plant^{-1}	0.359	0.338	-0.153	-0.486
Panicle length (cm)	-0.279	0.208	0.453	-0.227
No. of grains panicle^{-1}	-0.402	0.417	0.140	-0.024
Spikelet fertility (%)	-0.345	0.395	-0.037	-0.023
Days to maturity	0.484	0.271	0.297	0.016
Grain yield plant^{-1} (g)	-0.133	-0.249	0.326	-0.725
Test weight (g)	0.102	-0.545	0.136	-0.153
Eigen Value (Root)	2.866	1.760	1.557	0.961
% Var. Exp.	31.848	19.553	17.296	10.674
Cum. Var. Exp.	31.848	51.401	68.697	79.370

Table 3: Cluster analysis by PCA

Cluster No	No of Breeding lines in each Cluster	Cluster Members
1 Cluster	18	2 6 7 8 9 10 19 27 28 29 41 52 55 59 61 73 84 87
2 Cluster	3	17 20 70
3 Cluster	17	1 4 12 30 34 37 44 53 54 57 63 64 74 95 108 116 119
4 Cluster	12	31 32 38 40 62 65 71 76 77 78 99 100
5 Cluster	15	3 5 11 25 35 36 42 46 51 60 68 106 109 113 121
6 Cluster	4	24 47 67 89
7 Cluster	6	16 21 22 23 49 69
8 Cluster	6	39 79 80 81 82 101
9 Cluster	9	88 90 92 93 103 110 111 114 118
10 Cluster	14	66 83 85 86 91 94 97 98 102 104 105 112 115 117
11 Cluster	2	43 96
12 Cluster	15	13 14 15 18 26 33 45 48 50 56 58 72 75 107 120

Table 4: Cluster-wise mean values in PCA for 9 characters in rice

| Cluster no.	Days to 50% flowering	Plant height (cm)	EBT	Panicle length (cm)	No. of grains panicle^{-1}	Spikelet fertility (%)	Days to maturity	Grain yield plant^{-1} (g)	Test weight (g)
1 Cluster	117.00	120.00	11.00	24.72	238.00	91.92	146.00	21.00	18.87
2 Cluster	120.00	122.00	11.00	26.72	290.00	87.12	149.00	19.00	15.82
3 Cluster	114.00	126.00	10.00	24.99	234.00	89.16	143.00	26.00	20.12
4 Cluster	106.00	122.00	10.00	24.99	253.00	87.86	137.00	21.00	18.09
5 Cluster	114.00	126.00	11.00	24.25	199.00	89.58	144.00	28.00	21.30
6 Cluster	117.00	122.00	10.00	24.03	140.00	88.05	147.00	23.00	23.26
7 Cluster	117.00	121.00	12.00	26.11	270.00	87.97	145.00	20.00	17.08
8 Cluster	105.00	123.00	10.00	28.05	316.00	85.61	134.00	27.00.	16.98
9 Cluster	109.00	127.00	11.00	25.37	202.00	88.77	139.00	36.00	22.80
10 Cluster	105.00	125.00	10.00	25.86	230.00	90.00	135.00	31.00	20.58
11 Cluster	102.00	117.00	10.00	24.55	141.00	90.90	132.00	24.00	26.00
12 Cluster	118.00	122.00	10.00	23.98	164.00	89.67	148.00	23.00	20.29
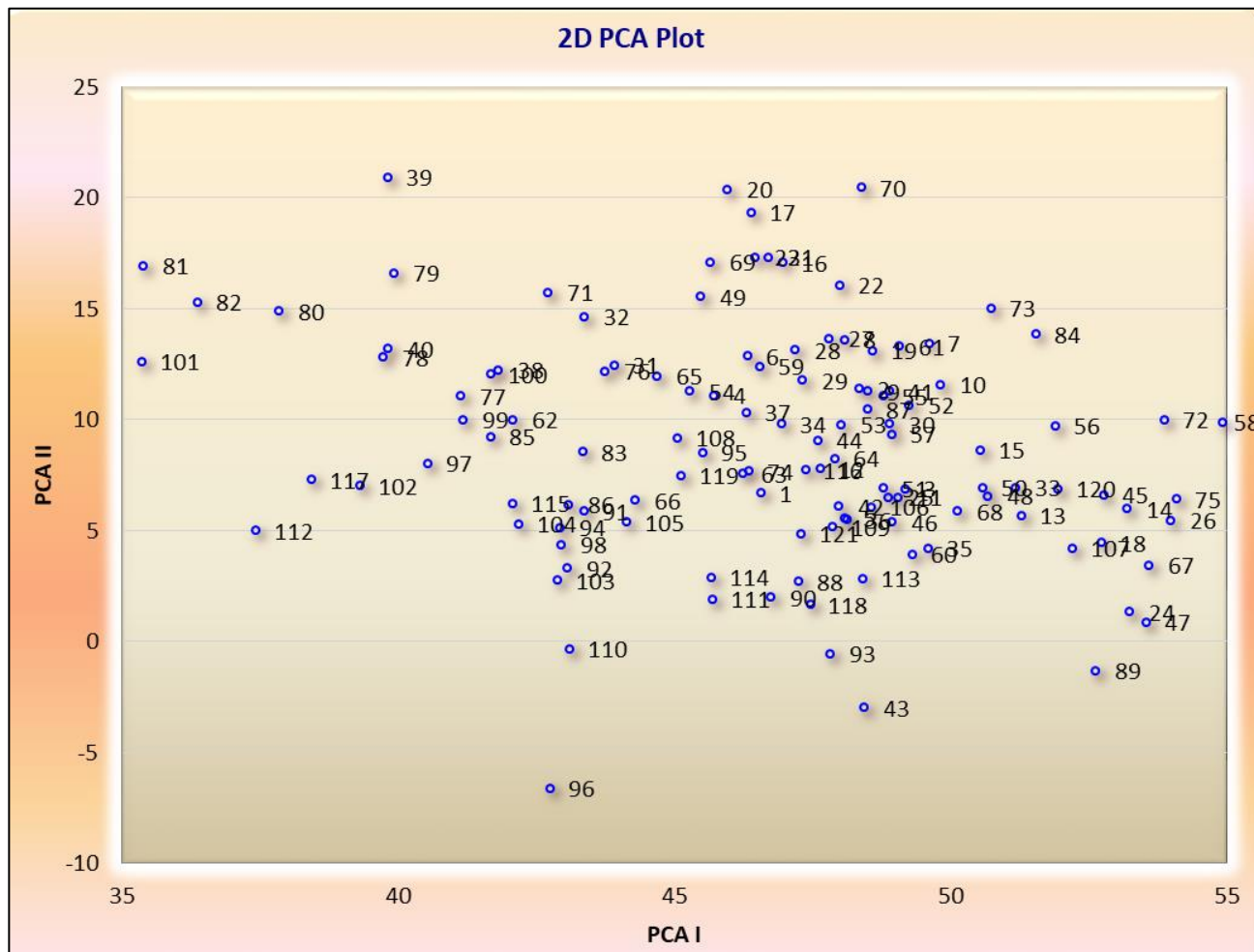
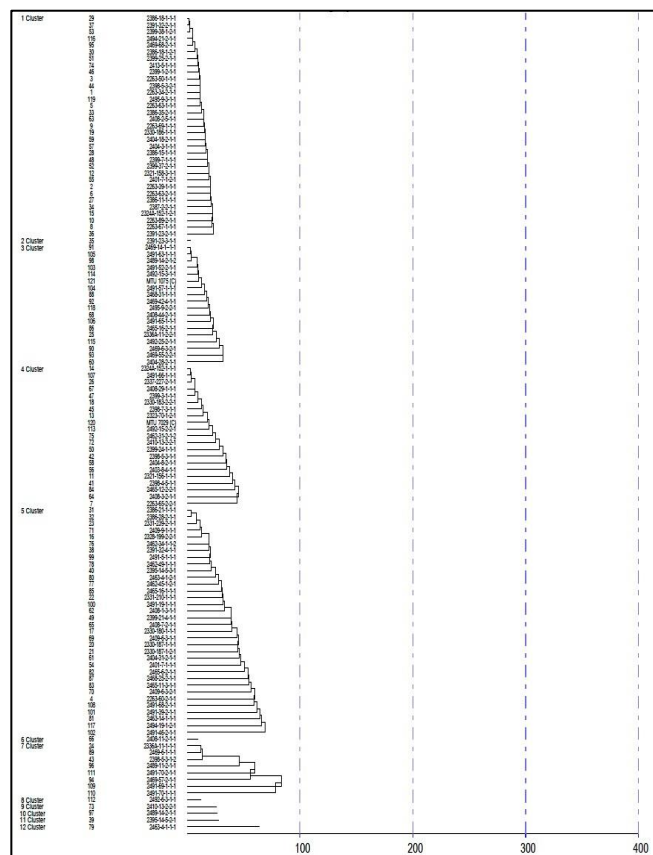**Fig 1:** Distribution of breeding lines across two components

Fig 2: Cluster analysis showing the relationship and diversity among rice genotypes based on yield and its attributing characters using Tocher's method
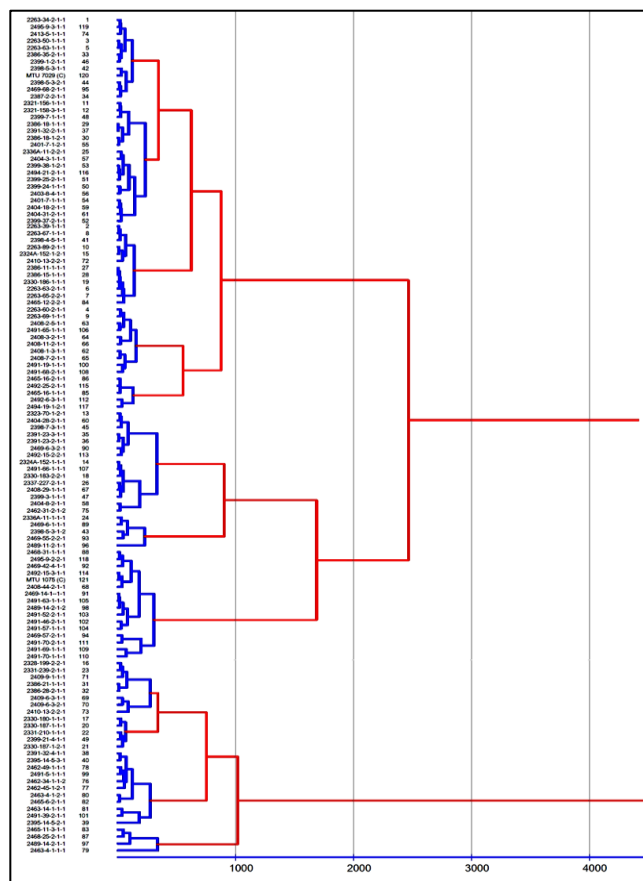

Fig 3: Classification of 121 lines of rice using Ward's minimum variance dendrogram

Conclusions

The present investigation revealed the presence of considerable level of genetic divergence among the rice breeding lines under study. Principal component analysis identified few important traits that plays predominant role in classifying the variation present in the breeding lines. The analysis identified days to 50% flowering, days to maturity, ear bearing tillers per plant, total grains and panicle length in different principal components are the most important characters for classifying the variation and selection of genotypes based on these characters would be reliable. The clustering analysis can be used for the selection of parental materials with diverse characters and further this gives a picture of similarities and dissimilarities of individual rice genotypes.

References

1. Ariyo ON. Mulitvariate analysis and choice of parents in Okra (*A. esculentus*). Annals of Applied Biology 1987;116:335-341.
2. Ariyo QJ, Odulaja A. Numerical analysis of variation among accessions of Okra (*Abelmoschus esculentus*). Annals of Botany. 1991;67:527-31.
3. Khatun MT, Hanafi M, Yusop MR, Wong MY, Salleh FM, Ferdous J. Genetic variation, heritability and diversity analysis of upland rice (*Oryza sativa* L.) genotypes based on quantitative traits. Bio Med Research International. 2015, 1-7.
4. Nair NV, Balakrishnan R, Sreenivasan TV. Variability of quantitative traits in exotic hybrid germplasm of sugarcane. Genetic Resources and Crop Evolution. 1998;45:459-464.
5. Oyelola BA. The Nigerian Statistical Association preconference workshop. University of Ibadan 2004, 20-21.
6. Ravikumar BNVSR, Kumari PN, Rao PVR, Rani MG, Satyanarayana PV, Chamundeswari N *et al*. Principal component analysis and character association for yield components in rice (*Oryza sativa* L.) cultivars suitable for irrigated ecosystem. Current Biotica 2015;9(1):25-35.
7. Sneath PHA, Sokal RR. Numerical Taxonomy: The Principles and Practice of Numerical Classification 1973.
8. Sudeethi K, Srinivas T, Ravi Kumar BNVSR, Jyothula DPB, Nafeez Umar K. Estimation of genetic diversity for anaerobic germination traits using principal component analysis in rice (*Oryza sativa* L.). The Bioscan 2019;14(4):283-287
9. Gomez KA, Gomez KK. Statistical analysis for agricultural research, 2nd Ed. John Eilly & Sons, New York 1984.