emergent
Life Sciences Research

Research Article

Application of principal component analysis for rice f5 families characterization and evaluation

K. L. Y. Tejaswini, Srinivas Manukonda, B. N. V. S. R. Ravi Kumar, P. V. Ramana Rao, Lal Ahamed M, S. Krishnam Raju

Abstract

The type and level of genetic variance in one hundred and fourteen F5 families of rice obtained from six different crosses was estimated along with their seven parents using Mahalanobis D^2 – statistics by considering 10 characters. Mahalanobis D^2 analysis revealed considerable amount of diversity in the material. The genotypes were grouped into twelve clusters. Cluster IX constituted maximum number of genotypes (26). The genotypes falling in cluster VII had the maximum divergence, which was closely followed by cluster VIII and cluster XI. The maximum inter cluster D2 values was observed between cluster X and XI (931.276) followed by cluster VIII and XI (814.784) suggesting that the genotypes constituted in these clusters may be used as parents for future hybridization programs. Principal component analysis revealed that families MTU 2462-1-5-2, MTU 2462-15-1-1, MTU 2468-30-2-2, MTU 2468-29-3-1, MTU 2462-1-9-2, MTU 2469-23-2-1, MTU 2469-6-2-1, MTU 2469-36-1-1, MTU 2469-32-2-1 were more divergent and hence, can be used in breeding programs. Hence, the results of cluster analysis were supported by principal component analysis.

Keywords D^2 statistic, genetic divergence, principal component analysis

Received: 10 April 2018
Accepted: 18 June 2018
Online: 27 June 2018

Authors:
K. L. Y. Tejaswini✉, L. Ahamed M.,
Department of Genetics and Plant Breeding,
Agriculture College, Bapatla, Guntur District,
Andhra Pradesh,
India

S. Manukonda, B.N.V.S.R. R. Kumar, P. V. R. Rao, S. K. Raju
Andhra Pradesh Rice Research Institute & RARS, Maruteru, West Godavari,
Andhra Pradesh,
India

✉tejaswini.kraleti@gmail.com

Emer Life Sci Res (2018) 4(1): 72-84

E-ISSN: 2395-6658
P-ISSN: 2395-664X

DOI: http://doi.org/10.31783/ELSR.2018.417284

Introduction

Rice (*Oryza sativa*, 2n = 2x = 24, family: Graminaea) is the major food crop grown across all the continents in the world, except Antarctica. As a global food, it has a large influence on human nutrition and food security. It is the staple food for half of the world's population. In Asia alone, more than 2 billion people obtain 60-70% of caloric intake from rice and its derived products and almost a billion households in Asia, Africa and America depends on rice based cropping systems for their livelihood and main source of employment. Therefore, Rice is on the frontline in the fight against world hunger and poverty. Rice is also a symbol of both cultural identity and global unity. For all of these reasons, it is apt to say "Rice is Life" [1].

Globally, it is planted on about 158 million hectares of land with an annual production of 478 million tons. India ranks first in area (43.85 million hectares) and second in the rice production (104.78 million tonnes) with a productivity of 2185 kg ha-1. Andhra Pradesh occupies an area of 3.80 million hectares with a production of 11.56 million tones and productivity of 3244 kg ha-1 (www.indiaagristat.com, 2014-15).

Demand for rice is increasing day by day and keeping in view of the future demand of rice as a food for human, there is a dire need to develop new varieties which can break the yield ceiling. Variation present in the

population an important prerequisite for improvement of any crop species. The choice of parents is of greater importance in breeding programs. Assessment of a large number of germplasm for genetic diversity is of immense importance in selection of diverse genotypes. Realizing the importance of germplasm in the development of desirable varieties, breeders are now looking for more diverse forms from various sources to further augument the yield potential of the genotypes. Mahalanobias D^2 statistic [2], PCA are few powerful tools for quantifying genetic divergence among germplasm collections with respect to characters. Therefore, this study was conducted to show variation among the 114 F5 families of rice and identify the traits that can contribute to the variability in this population and for their possible exploitation in breeding programs.

Methodology

The experimental material consisted of one hundred and fourteen F5 families belonging to six crosses of rice along with their seven parents (Table 1 and 2) obtained from Andhra Pradesh Rice Research Institute and Regional Agricultural Research Station, Maruteru. Those were sown in randomized block design with two replications during kharif, 2015. Thirty days old seedlings were planted with a spacing of 20 cm and 15 cm between rows and plants, respectively. Observations were recorded on five randomly selected plants for eight characters viz., plant height, number of panicles per plant, panicle length, grain yield per plant, test weight, kernel length, kernel breadth and L/B ratio; while for other two characters viz., days to 50 per cent flowering and days to maturity, observations were recorded on plot basis. The mean values over two replications were used for statistical analysis to measure the genetic divergence using Mahalonobias D2 followed by principal component analysis as suggested by Rao [3].

Test with Wilk's criterion

Wilk's 'V' (statistic) criterion was used to test the significant differences between the groups based on the pooled effects of all the characters. The significance of 'V' (statistic) value was tested by percent (%) at 1200 degrees of freedom. The 'V' statistic value was highly significant indicating that the genotypes differed significantly when all the characters were considered simultaneously. The significance of 121 lines (114 F5 families along with their seven parents) in the analysis of variance of dispersion clearly indicated the significant pooled effect of all the characters studied among the different genotypes. Hence, further analysis was made to estimate D^2 analysis.

Mahalanobis' D^2 values

In order to assess the genetic diversity among 121 lines (114 F5 families along with their seven parents), D^2 statistic was used following the procedure given by Rao (1952). Since the entire 10 yield component characters were correlated, they were transformed into uncorrelated linear combination through pivotal condensation method. The statistical distance (D^2) between the pairs of genotypes were obtained as the sum of squares of the difference between the pairs of corresponding uncorrelated values of any two genotypes considered at one time.

Table 1. Supplementary material

Cross	Number of Families Studied
MTU 7029/ MTU 1081	45
MTU 7029/ JGL 11118	5
MTU 7029/ NLR 34449	18
MTU 7029/ MTU 1121	3
MTU 7029/ PAU 3116-25-5-1	15
MTU 7029/ PAU 3140-126-1	28
Total - Six crosses	**114**

Table 2. List of F5 families studied

S.No.	Code	Entry	Cross combination		S.No.	Code	Entry	Cross combination
1	TSM-1	MTU 2462-1-1-1	MTU 7029/ MTU 1081		62	TSM-99	MTU 2465-16-2-1	MTU 7029/ NLR 34449
2	TSM-2	MTU 2462-1-1-2	MTU 7029/ MTU 1081		63	TSM-101	MTU 2465-16-2-3	MTU 7029/ NLR 34449
3	TSM-3	MTU 2462-1-2-1	MTU 7029/ MTU 1081		64	TSM-102	MTU 2465-22-1-1	MTU 7029/ NLR 34449
4	TSM-4	MTU 2462-1-5-1	MTU 7029/ MTU 1081		65	TSM-103	MTU 2465-22-2-1	MTU 7029/ NLR 34449
5	TSM-5	MTU 2462-1-5-2	MTU 7029/ MTU 1081		66	TSM-104	MTU 2465-22-2-2	MTU 7029/ NLR 34449
6	TSM-6	MTU 2462-1-8-1	MTU 7029/ MTU 1081		67	TSM-108	MTU 2465-24-3-2	MTU 7029/ NLR 34449
7	TSM-7	MTU 2462-1-8-2	MTU 7029/ MTU 1081		68	TSM-110	MTU 2465-27-2-1	MTU 7029/ NLR 34449
8	TSM-8	MTU 2462-1-9-1	MTU 7029/ MTU 1081		69	TSM-114	MTU 2466-4-2-1	MTU 7029/ MTU 1121
9	TSM-9	MTU 2462-1-9-2	MTU 7029/ MTU 1081		70	TSM-115	MTU 2466-4-2-2	MTU 7029/ MTU 1121
10	TSM-10	MTU 2462-2-1-1	MTU 7029/ MTU 1081		71	TSM-116	MTU 2466-4-3-1	MTU 7029/ MTU 1121
11	TSM-12	MTU 2462-4-1-1	MTU 7029/ MTU 1081		72	TSM-118	MTU 2468-1-1-1	MTU 7029/ PAU 3116-25-5-1
12	TSM-16	MTU 2462-6-2-1	MTU 7029/ MTU 1081		73	TSM-120	MTU 2468-2-1-1	MTU 7029/ PAU 3116-25-5-1
13	TSM-17	MTU 2462-8-1-1	MTU 7029/ MTU 1081		74	TSM-128	MTU 2468-8-2-1	MTU 7029/ PAU 3116-25-5-1
14	TSM-18	MTU 2462-8-2-1	MTU 7029/ MTU 1081		75	TSM-132	MTU 2468-18-1-1	MTU 7029/ PAU 3116-25-5-1
15	TSM-19	MTU 2462-8-2-2	MTU 7029/ MTU 1081		76	TSM-133	MTU 2468-18-1-2	MTU 7029/ PAU 3116-25-5-1
16	TSM-20	MTU 2462-12-1-1	MTU 7029/ MTU 1081		77	TSM-134	MTU 2468-20-1-1	MTU 7029/ PAU 3116-25-5-1
17	TSM-22	MTU 2462-12-2-1	MTU 7029/ MTU 1081		78	TSM-138	MTU 2468-21-4-1	MTU 7029/ PAU 3116-25-5-1
18	TSM-29	MTU 2462-15-1-1	MTU 7029/ MTU 1081		79	TSM-141	MTU 2468-25-2-1	MTU 7029/ PAU 3116-25-5-1
19	TSM-32	MTU 2462-15-3-1	MTU 7029/ MTU 1081		80	TSM-146	MTU 2468-27-2-1	MTU 7029/ PAU 3116-25-5-1
20	TSM-33	MTU 2462-15-3-2	MTU 7029/ MTU 1081		81	TSM-147	MTU 2468-28-1-1	MTU 7029/ PAU 3116-25-5-1
21	TSM-35	MTU 2462-17-2-1	MTU 7029/ MTU 1081		82	TSM-148	MTU 2468-29-2-1	MTU 7029/ PAU 3116-25-5-1
22	TSM-36	MTU 2462-17-3-1	MTU 7029/ MTU 1081		83	TSM-149	MTU 2468-29-3-1	MTU 7029/ PAU 3116-25-5-1
23	TSM-37	MTU 2462-18-1-1	MTU 7029/ MTU 1081		84	TSM-150	MTU 2468-29-4-1	MTU 7029/ PAU 3116-25-5-1
24	TSM-39	MTU 2462-21-1-1	MTU 7029/ MTU 1081		85	TSM-152	MTU 2468-30-2-2	MTU 7029/ PAU 3116-25-5-1
25	TSM-42	MTU 2462-22-2-1	MTU 7029/ MTU 1081		86	TSM-153	MTU 2468-31-1-1	MTU 7029/ PAU 3116-25-5-1
26	TSM-43	MTU 2462-22-3-1	MTU 7029/ MTU 1081		87	TSM-164	MTU 2469-6-1-2	MTU 7029/ PAU 3140-126-1
27	TSM-48	MTU 2462-23-3-1	MTU 7029/ MTU 1081		88	TSM-165	MTU 2469-6-2-1	MTU 7029/ PAU 3140-126-1
28	TSM-50	MTU 2462-26-3-1	MTU 7029/ MTU 1081		89	TSM-166	MTU 2469-6-3-1	MTU 7029/ PAU 3140-126-1
29	TSM-51	MTU 2462-26-3-2	MTU 7029/ MTU 1081		90	TSM-167	MTU 2469-6-3-2	MTU 7029/ PAU 3140-126-1
30	TSM-54	MTU 2462-27-3-1	MTU 7029/ MTU 1081		91	TSM-168	MTU 2469-6-5-1	MTU 7029/ PAU 3140-126-1
31	TSM-56	MTU 2462-29-2-1	MTU 7029/ MTU 1081		92	TSM-169	MTU 2469-7-1-1	MTU 7029/ PAU 3140-126-1
32	TSM-57	MTU 2462-30-1-1	MTU 7029/ MTU 1081		93	TSM-171	MTU 2469-8-1-1	MTU 7029/ PAU 3140-126-1
33	TSM-58	MTU 2462-31-2-1	MTU 7029/ MTU 1081		94	TSM-174	MTU 2469-10-2-1	MTU 7029/ PAU 3140-126-1
34	TSM-61	MTU 2462-34-1-1	MTU 7029/ MTU 1081		95	TSM-175	MTU 2469-11-1-1	MTU 7029/ PAU 3140-126-1
35	TSM-62	MTU 2462-34-4-1	MTU 7029/ MTU 1081		96	TSM-178	MTU 2469-14-1-1	MTU 7029/ PAU 3140-126-1
36	TSM-63	MTU 2462-35-2-1	MTU 7029/ MTU 1081		97	TSM-183	MTU 2469-23-2-1	MTU 7029/ PAU 3140-126-1
37	TSM-64	MTU 2462-35-2-2	MTU 7029/ MTU 1081		98	TSM-184	MTU 2469-23-2-2	MTU 7029/ PAU 3140-126-1
38	TSM-65	MTU 2462-37-1-1	MTU 7029/ MTU 1081		99	TSM-190	MTU 2469-32-1-1	MTU 7029/ PAU 3140-126-1
39	TSM-66	MTU 2462-37-2-1	MTU 7029/ MTU 1081		100	TSM-191	MTU 2469-32-2-1	MTU 7029/ PAU 3140-126-1
40	TSM-67	MTU 2462-37-2-2	MTU 7029/ MTU 1081		101	TSM-200	MTU 2469-36-1-1	MTU 7029/ PAU 3140-126-1
41	TSM-68	MTU 2462-41-2-1	MTU 7029/ MTU 1081		102	TSM-204	MTU 2469-38-4-1	MTU 7029/ PAU 3140-126-1
42	TSM-69	MTU 2462-42-2-1	MTU 7029/ MTU 1081		103	TSM-211	MTU 2469-41-2-2	MTU 7029/ PAU 3140-126-1
43	TSM-72	MTU 2462-45-1-2	MTU 7029/ MTU 1081		104	TSM-213	MTU 2469-42-1-1	MTU 7029/ PAU 3140-126-1
44	TSM-73	MTU 2462-45-2-1	MTU 7029/ MTU 1081		105	TSM-215	MTU 2469-42-3-1	MTU 7029/ PAU 3140-126-1
45	TSM-74	MTU 2462-49-2-1	MTU 7029/ MTU 1081		106	TSM-216	MTU 2469-42-4-1	MTU 7029/ PAU 3140-126-1
46	TSM-76	MTU 2463-4-1-2	MTU 7029/ JGL 11118		107	TSM-219	MTU 2469-55-1-1	MTU 7029/ PAU 3140-126-1
47	TSM-77	MTU 2463-4-2-1	MTU 7029/ JGL 11118		108	TSM-220	MTU 2469-55-2-1	MTU 7029/ PAU 3140-126-1
48	TSM-79	MTU 2463-8-2-1	MTU 7029/ JGL 11118		109	TSM-221	MTU 2469-55-2-2	MTU 7029/ PAU 3140-126-1
49	TSM-82	MTU 2463-14-1-1	MTU 7029/ JGL 11118		110	TSM-223	MTU 2469-57-1-2	MTU 7029/ PAU 3140-126-1
50	TSM-83	MTU 2463-15-1-1	MTU 7029/ JGL 11118		111	TSM-228	MTU 2469-68-1-1	MTU 7029/ PAU 3140-126-1
51	TSM-85	MTU 2465-4-1-1	MTU 7029/ NLR 34449		112	TSM-229	MTU 2469-68-1-2	MTU 7029/ PAU 3140-126-1
52	TSM-86	MTU 2465-4-2-1	MTU 7029/ NLR 34449		113	TSM-230	MTU 2469-68-2-1	MTU 7029/ PAU 3140-126-1
53	TSM-87	MTU 2465-6-2-1	MTU 7029/ NLR 34449		114	TSM-235	MTU 2469-74-2-1	MTU 7029/ PAU 3140-126-1
54	TSM-88	MTU 2465-8-1-1	MTU 7029/ NLR 34449		**Parents**			
55	TSM-90	MTU 2465-10-1-1	MTU 7029/ NLR 34449		115		MTU 7029	
56	TSM-92	MTU 2465-11-3-1	MTU 7029/ NLR 34449		116		MTU 1081	
57	TSM-93	MTU 2465-11-3-2	MTU 7029/ NLR 34449		117		MTU 1121	
58	TSM-94	MTU 2465-12-2-1	MTU 7029/ NLR 34449		118		JGL 11118	
59	TSM-95	MTU 2465-12-2-2	MTU 7029/ NLR 34449		119		NLR 34449	
60	TSM-96	MTU 2465-13-3-1	MTU 7029/ NLR 34449		120		PAU 3116-25-5-1	
61	TSM-97	MTU 2465-13-3-2	MTU 7029/ NLR 34449		121		PAU 3140-126-1	

Results and Discussion

Relative contribution of individual characters towards divergence

The percent contribution towards genetic divergence by all the 10 characters is presented in Table 3 and Fig 1.The trait test weight showed maximum contribution (23.79) towards genetic divergence followed by days to maturity (22.40), plant height (19.68), grain yield per plant (16.46), kernel length (7.71), days to 50% flowering (4.82), number of panicles per plant (2.20), panicle length (1.96), kernel breadth (0.92) and L/B ratio (0.06). Senapati and Sarkar [4] reported that test weight contributed mostly to genetic divergence.

Table 3. Contribution of different characters towards genetic divergence among 121 lines (114 F5 families and seven parents) of rice (Oryzasativa L.)

S.No.	Source	Contribution %
1	Days to 50% Flowering	4.82
2	Days to Maturity	22.40
3	Plant Height	19.68
4	Number of panicles per plant (EBT)	2.20
5	Panicle Length	1.96
6	Grain Yield	16.46
7	Test Weight	23.79
8	Kernel Length	7.71
9	Kernel Breadth	0.92
10	L/B Ratio	0.06

Grouping of genotypes into various clusters

The 121 lines were assembled into 12 clusters employing the Tocher's method with the standard that the intra-cluster average D2 values should be less than the inter cluster D2 values. The distribution of 121 lines into 12 clusters is presented in Table 4. The distribution of 121 lines into 12 clusters was at random with maximum number of 26 genotypes in cluster IX. Cluster VIII and cluster XI were found to be the second largest with 13 families each followed by cluster VII with 12 genotypes.

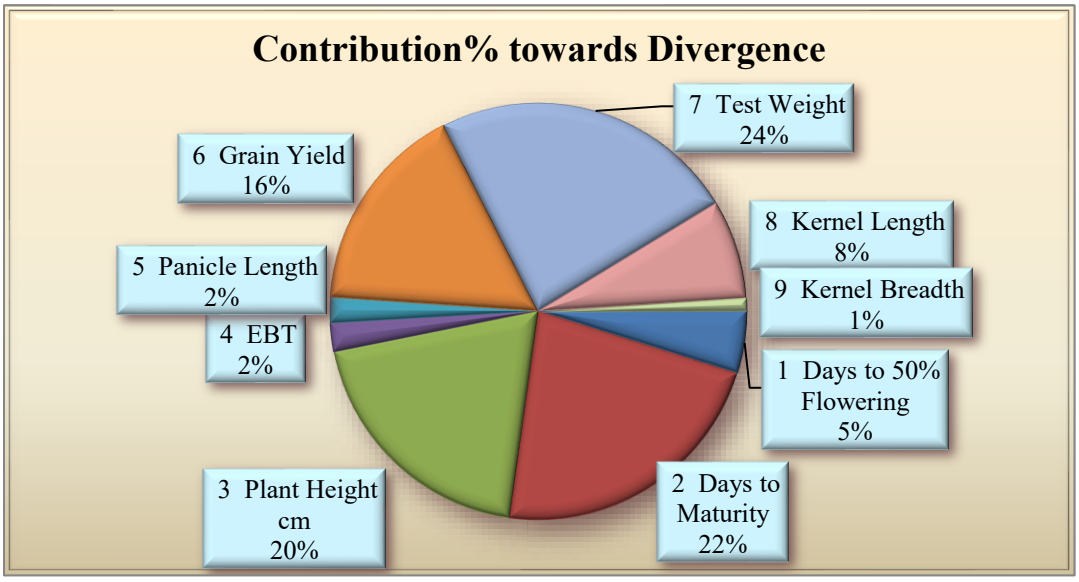


Figure 1. Percent contribution of different characters towards genetic divergence

Table 4. Clustering pattern of 121 lines (114 F5 families and seven parents) of rice (*Oryzasativa* L.) by Tocher's method

Cluster number	Number of genotypes	Genotypes
Cluster I	4	TSM-2, TSM-171, PAU-3116-25-5-1, PAU-3140-126-1
Cluster II	9	TSM-10, TSM-12, TSM-17, TSM-33, TSM-73, TSM-93, TSM-116, MTU 7029, MTU 1121
Cluster III	3	TSM-5, TSM-9, TSM-29
Cluster IV	11	TSM-4, TSM-54, TSM-58, TSM-67, TSM-141, TSM-147, TSM-149, TSM-184, TSM-204, TSM-219, TSM-230
Cluster V	11	TSM-1, TSM-8, TSM-32, TSM-50, TSM-51, TSM-57, TSM-77, TSM-82, TSM-83, TSM-97, TSM-221
Cluster VI	10	TSM-7, TSM-42, TSM-43, TSM-72, TSM-79, TSM-85, TSM-86, TSM-87, TSM-90, TSM-99
Cluster VII	12	TSM-3, TSM-16, TSM-69, TSM-74, TSM-88, TSM-92, TSM-94, TSM-95, TSM-101, TSM-115, TSM-120, TSM-228
Cluster VIII	13	TSM-63, TSM-68, TSM-108, TSM-138, TSM-148, TSM-150, TSM-164, TSM-166, TSM-168, TSM-174, TSM-175, TSM-183, TSM-190
Cluster IX	26	TSM-6, TSM-19, TSM-35, TSM-36, TSM-39, TSM-48, TSM-61, TSM-62, TSM-64, TSM-65, TSM-66, TSM-76, TSM-96, TSM-102, TSM-103, TSM-104, TSM-128, TSM-132, TSM-146, TSM-167, TSM-178, TSM-211, TSM-216, TSM-220, TSM-229, TSM-235
Cluster X	6	TSM-152, TSM-153, TSM-165, TSM-191, TSM-200, TSM-213
Cluster XI	13	TSM-18, TSM-20, TSM-22, TSM-37, TSM-56, TSM-110, TSM-114, TSM-118, TSM-133, TSM-134, TSM-169, TSM-215, TSM-223
Cluster XII	3	MTU 1081, JGL 11118, NLR 34449

Average intra and inter cluster D^2 values

The average intra and inter cluster D^2 values were estimated as per the procedure given by Singh and Chaudhary [5] and were presented in Table 5. Intra- and inter -cluster distance of 121 lines of rice are showed in Figure. 2. Dendrogram representing the relationship among 121 lines in twelve clusters based on Mahalanobis's D^2 values are presented in Figure. 3

The maximum intra cluster D^2 value was 72.356 for cluster VII followed by 64.619 for cluster VIII, 61.669 for cluster XI, 57.027 for cluster IV, 48.215 for cluster V, 43.698 for cluster II and 30.608 for cluster I; while it was zero for clusters III, VI, IX, X and XII. The high intra cluster distance in cluster VII indicated the presence of wide genetic diversity among the genotypes viz., MTU 2462-1-2-1, MTU 2462-6-2-1, MTU 2462-42-2-1, MTU 2462-49-2-1, MTU 2465-8-1-1, MTU 2465-11-3-1, MTU 2465-12-2-1, MTU 2465-12-2-2, MTU 2465-16-2-3, MTU 2466-4-2-2, MTU 2468-2-1-1 and MTU 2469-68-1-1. Wide range of cluster mean values indicates the presence of variation among the studied genotypes.

Cluster I comprised of 4 families and was closest to cluster III (41.01) and it was farthest from cluster XI (536.701). Cluster II comprised of 9 families and was closest to cluster X (65.107) and it was farthest from cluster XI (679.476). Cluster III comprised of 3families and was closest to cluster I (41.018) and it was farthest from cluster XI (452.117). Cluster IV comprised of 11families and was closest to cluster III (84.127) and it was farthest from cluster X (318.359). Cluster V comprised of 11 families and was closest to cluster I (52.586) and it was farthest from cluster XI (540.565). Cluster VI comprised of 10 families and was closest to cluster III (50.08) and it was farthest from cluster XI (451.248). Cluster VII comprised of 12 families and was closest to cluster I (75.122) and it was farthest from cluster XII (654.458). Cluster VIII comprised of 13 families and was closest to cluster I (136.941) and it was farthest from cluster XI (814.784). Cluster IX comprised of 26 families and was closest to cluster IX (56.339) and it was farthest from cluster XI (803.024). Cluster X comprised of 6 families and was closest to cluster IX (56.339) and it was farthest from cluster XI (931.276). Cluster XI comprised of 13 families and was closest to cluster IV

Table 5. Average intra-and inter –cluster D^2 values among twelve clusters in 121 lines of rice (*Oryzasativa* L.) genotypes

Inter and Intra Cluster Distances : Tocher Method												
	Cluster I	**Cluster II**	**Cluster III**	**Cluster IV**	**Cluster V**	**Cluster VI**	**Cluster VII**	**Cluster VIII**	**Cluster IX**	**Cluster X**	**Cluster XI**	**Cluster XII**
Cluster I	**30.608**	137.36	41.018	103.582	52.586	68.121	75.122	136.941	71.94	178.791	536.701	117.145
Cluster II		**43.698**	171.401	200.015	144.139	303.303	193.106	265.297	73.57	65.107	679.476	345.638
Cluster III			**0.000**	84.127	54.125	50.08	102.476	150.798	115.175	256.925	452.117	120.639
Cluster IV				**57.027**	116.407	114.294	168.857	248.865	215.247	318.359	276.044	115.183
Cluster V					**48.215**	99.364	115.242	143.398	93.505	192.116	540.565	164.796
Cluster VI						**0.0000**	102.999	168.357	197.546	372.949	451.248	63.771
Cluster VII							**72.356**	180.408	113.984	227.211	654.458	125.238
Cluster VIII								**64.619**	172.828	274.49	814.784	287.62
Cluster IX									**0.000**	**56.339**	803.024	276.517
Cluster X										**0.0000**	931.276	451.742
Cluster XI											**61.669**	369.336
Cluster XII												**0.0000**

Diagonal bold values indicate intra cluster distances

(276.044) and it was farthest from cluster X (931.276). Cluster XII comprised of 3 families and was closest to cluster VI (63.771) and it was farthest from cluster X (451.742).

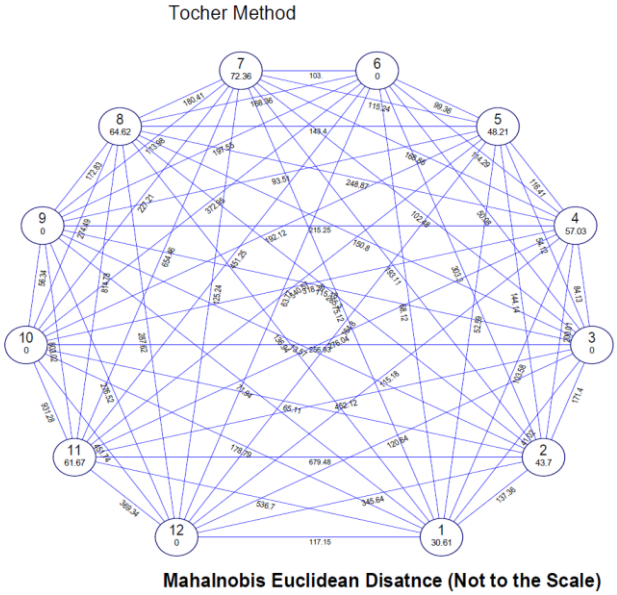


Figure 2. Intra-and Inter-cluster distances of 121 lines of rice (*Oryza sativa* L.) in twelve clusters based on Euclidean^2 distances

The intra and inter cluster D^2 values revealed that inter cluster D^2 values were greater than intra cluster D^2 values. The maximum inter cluster D^2 values was observed between cluster X and XI (931.276) followed by cluster VII and XI (814.784). Based on these studies crosses may be made between genotypes of clusters X and cluster XI followed by genotypes of clusters VII and XI to obtain new desirable recombinants in rice.

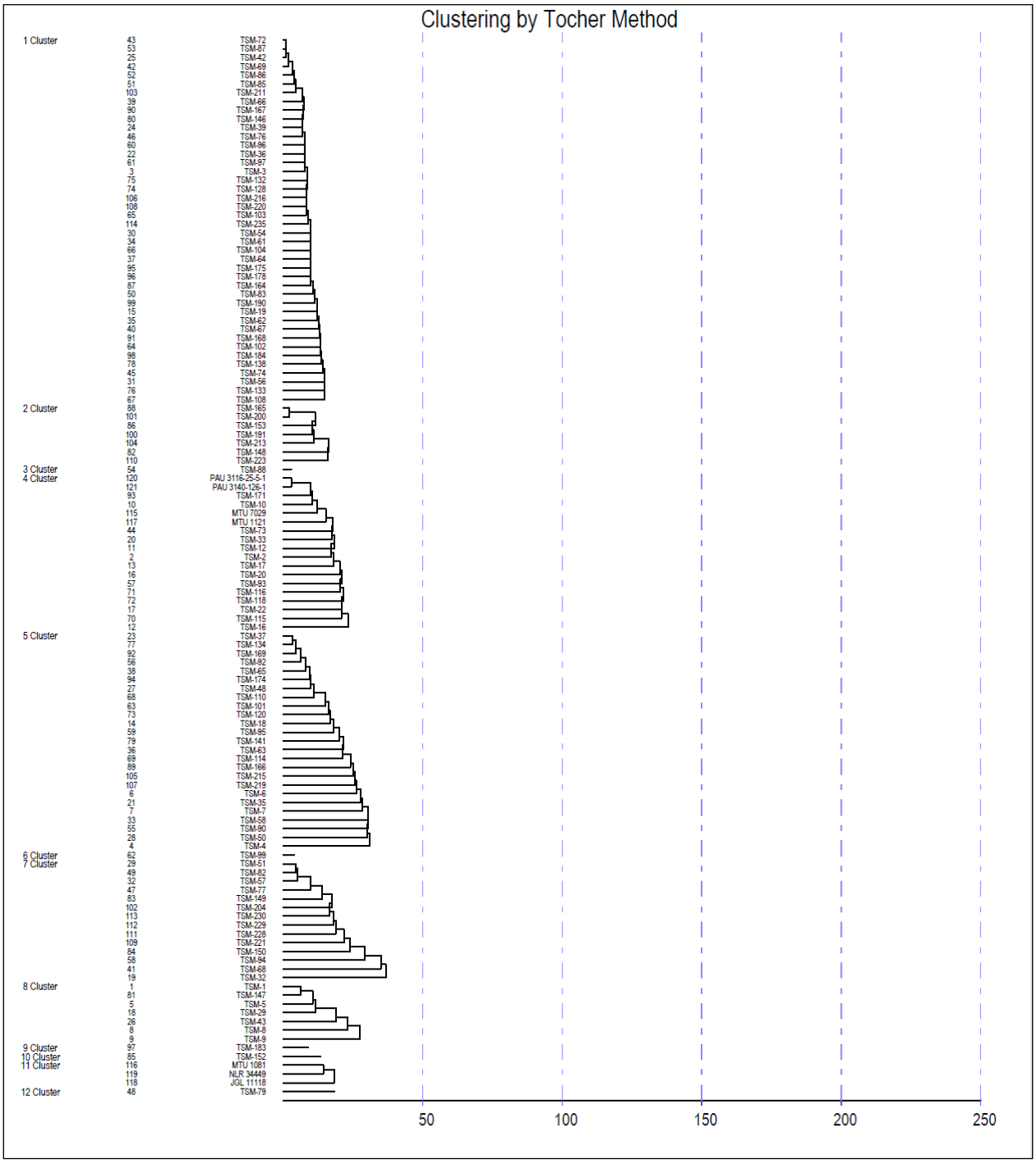


Figure 3. Dendrogram showing relationship among 121 lines of rice (*Oryza sativa* L.) genotypes in twelve clusters based on Mahalanobis's D^2 values

Cluster mean values of characters

The cluster mean values for 10 characters are presented in Table 7. The data indicated a wide range of mean values between the characters. Days to 50% flowering had a range of 126 days for cluster IX to 92 days for cluster XI; days to maturity had a range of 150 days for cluster IX & X to 118 days for cluster XI; Plant height (cm) varied from 124.29 cm for cluster VII to 99.07 cm for cluster XI; Number of panicles per plant had a range of 16 for cluster X to 9 for cluster III; Panicle length had a range of 27.05 cm for cluster III to 22.35 cm for cluster X; Grain yield per plant had a range of 33.12 g for cluster VIII to 15.40 g for cluster XII; Test weight (g) recorded as high as 36.70 g in cluster X to as low as 11.50 g in cluster VI; Kernel length had a range of 6.9 mm for cluster XII to 5.7 mm for cluster III; Kernel breadth varied from 2.5 mm for cluster II, III and X to 2.1 mm for cluster VI, IX and XI; L/B ratio had a range of 3.2 for cluster XII to 2.3 for cluster III. Cluster mean values showed wide range indicating the presence of variation among the studied genotypes. It was determined that none of the clusters possessed at least one genotype with all the anticipated traits, which ruled out the likelihood of selecting directly one genotype for instant usage. Therefore, hybridization between the selected genotypes from divergent clusters is essential to judiciously combine all the targeted traits.

Table 7. Mean values of twelve clusters by Tocher's method for 121 lines of rice (*Oryza sativa* L.)

Cluster Means : Tocher Method										
	Days to 50% Flowering	**Days to Maturity**	**Plant Height**	**Number of panicles per plant**	**Panicle Length**	**Grain Yield per plant**	**Test Weight**	**Kernel Length**	**Kernel Breadth**	**L/B Ratio**
Cluster I	118	144	110.45	11	23.82	18.02	20.41	6.2	2.3	2.7
Cluster II	118	144	108.86	11	22.59	17.64	35.24	6.4	2.5	2.6
Cluster III	117	144	105.80	9	27.05	18.80	18.25	5.7	2.5	2.3
Cluster IV	110	134	108.38	10	23.73	20.81	20.57	6.1	2.3	2.7
Cluster V	117	144	106.36	10	23.33	19.71	20.91	6.1	2.4	2.6
Cluster VI	113	142	115.20	12	25.25	19.40	11.50	5.7	2.1	2.8
Cluster VII	118	144	124.29	12	25.51	17.72	21.29	6.4	2.3	2.8
Cluster VIII	120	147	112.76	11	25.51	33.12	19.19	6.5	2.3	2.9
Cluster IX	126	150	110.00	12	23.80	16.40	28.90	6.2	2.1	3.0
Cluster X	125	150	111.30	16	22.35	18.60	36.70	6.3	2.5	2.6
Cluster XI	92	118	99.07	9	25.37	18.03	15.80	5.8	2.1	2.9
Cluster XII	110	135	119.30	12	26.75	15.40	12.80	6.9	2.2	3.2

The cluster III is having highest mean value for panicle length and kernel breadth, Cluster VII for plant height, cluster VIII for grain yield per plant, cluster IX for days to 50% flowering and days to maturity, cluster X for ear bearing tillers per plant and test weight and cluster XII for kernel length and L/B ratio. The genotypes from these clusters having high mean values may be directly used for adaptation or may be used as parents in future hybridization programs as for a successful breeding program selection of genetically diverse parents is an important prerequisite to obtain better and desirable recombinants.

Application of principal component analysis in genetic divergence

The eigen values (variance), percent variability, cumulative percent variability and factors loading of different variables are given in Table 8. In the present study, the first five principal components

Table 8. Eigen values, proportion of total variance represented by first five principal components, cumulative per cent variance and component loading of different characters in rice (*Oryza sativa* L.)

	PC1	PC 2	PC 3	PC 4	PC 5
Eigene Value (Root)	3.579	1.677	1.218	1.171	0.746
% Var. Exp.	35.788	16.767	12.176	11.707	7.458
Cum. Var. Exp.	35.788	52.555	64.731	76.438	83.896
Days to 50% Flowering	0.482	0.049	0.183	0.054	0.085
Days to Maturity	0.483	0.047	0.170	0.073	0.106
Plant Height	0.302	-0.166	-0.496	0.293	-0.356
No.of panicles per plant	0.058	-0.009	0.097	-0.812	-0.495
Panicle Length	-0.295	-0.452	-0.098	0.135	-0.396
Grain Yield per plant	0.090	-0.199	0.647	0.375	-0.484
Test Weight	0.054	0.658	0.140	0.022	-0.303
Kernel Length	0.367	-0.046	0.006	-0.016	0.022
Kernel Breadth	-0.218	0.536	-0.222	0.286	-0.288
L/B Ratio	-0.399	0.031	0.430	0.073	0.209

contributed 83.89 percent towards the total variability. The principal component with the eigen values less than one were considered as non-significant and hence were ignored as they are unlikely to have any practical significance. It was therefore inferred that the essential features of the data set had been represented in the first four principal components, explaining 76.43 percent of total variability. These results were in accordance with the reports of Yadav et al. [6] where four principal components were formed to describe the maximum variance of the data set.

The first principal component (PC 1) contributed maximum towards the variability (35.788). Characters viz., days to 50% flowering (0.482), days to maturity (0.483), plant height (0.302), number of panicles per plant (0.058), grain yield per plant (0.090), test weight (0.054) and kernel length (0.367) were positively loaded while panicle length (-0.295), kernel breadth (-0.218) and L/B ratio (-0.399) were negatively loaded.

The second principal component (PC 2) described 16.767 percent of total variance. Characters viz., days to 50% flowering (0.049), days to maturity (0.047), test weight (0.658), kernel breadth (0.536) and L/B ratio (0.030) contributed positively to the variation while plant height (-0.166), number of panicles per plant (-0.009), panicle length (-0.452), grain yield per plant (-0.199) and kernel length (-0.046) contributed negatively to the variation.

The third principal component (PC 3) was characterized by 12.176 per cent contribution towards the total variability. Characters viz., days to 50% flowering (0.183), days to maturity (0.170), number of panicles per plant (0.097), grain yield per plant (0.647), test weight (0.140), kernel length (0.006) and L/B ratio (0.430) were positively loaded while plant height (-0.496), panicle length (-0.098) and kernel breadth (-0.222) were negatively loaded.

The fourth principal component (PC 4) was characterized by 11.707 per cent contribution towards the total variability. Characters viz., days to 50% flowering (0.054), days to maturity (0.073), plant height (0.293), panicle length (0.135), grain yield per plant (0.375), test weight (0.022), kernel breadth (0.286) and L/B ratio (0.073) contributed positively to the total variation while number of panicles per plant (-0.812) and kernel length (-0.016) contributed negatively to the total variation.

The fifth principal component (PC 5) was characterized by 7.458 per cent contribution towards the total variability. Characters viz., days to 50% flowering (0.085), days to maturity (0.106), kernel length (0.022) and L/B ratio (0.209) were positively loaded while plant height (-0.356), number of panicles per plant (-0.495), panicle length (-0.396), grain yield per plant (-0.484), test weight (-0.303) and kernel breadth

(-0.288) were negatively loaded. The analysis thus identified the maximum contributing variables for diversity. Characters viz., days to 50% flowering, days to maturity, grain yield per plant, test weight, kernel length and L/B ratio.

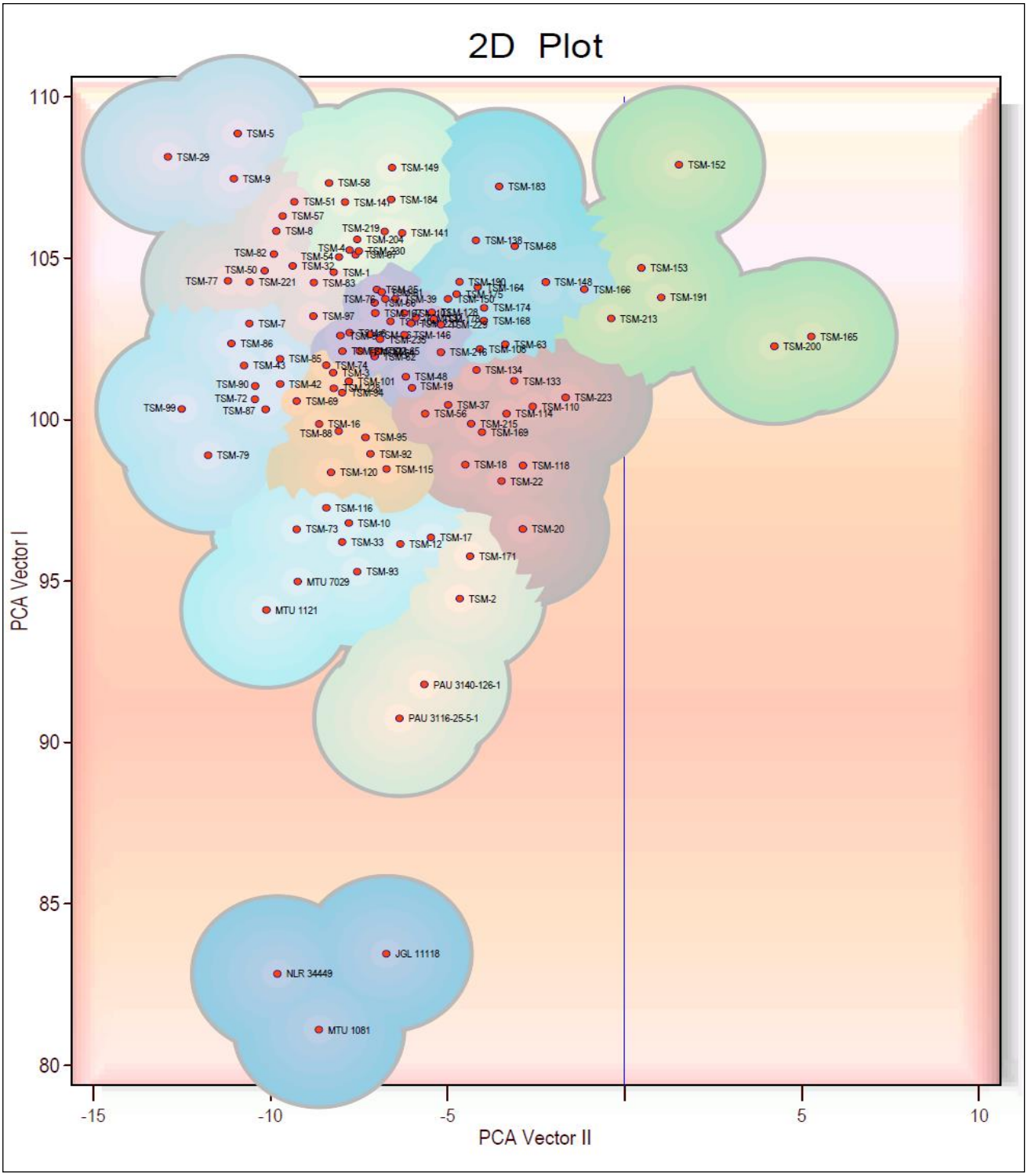


Figure 4. 2D plot showing scattering in 64 genotypes of rice (*Oryza sativa* L.)

The PCA scores for 121 lines in the first three principal components were computed and were considered as three axes as X, Y and Z and squared distance of each genotype from these three axes were calculated and presented in Table 9. These three PCA scores for 121 lines were plotted in graph to get two dimensional and three dimensional scatter diagrams (Fig 4 and Fig 5).

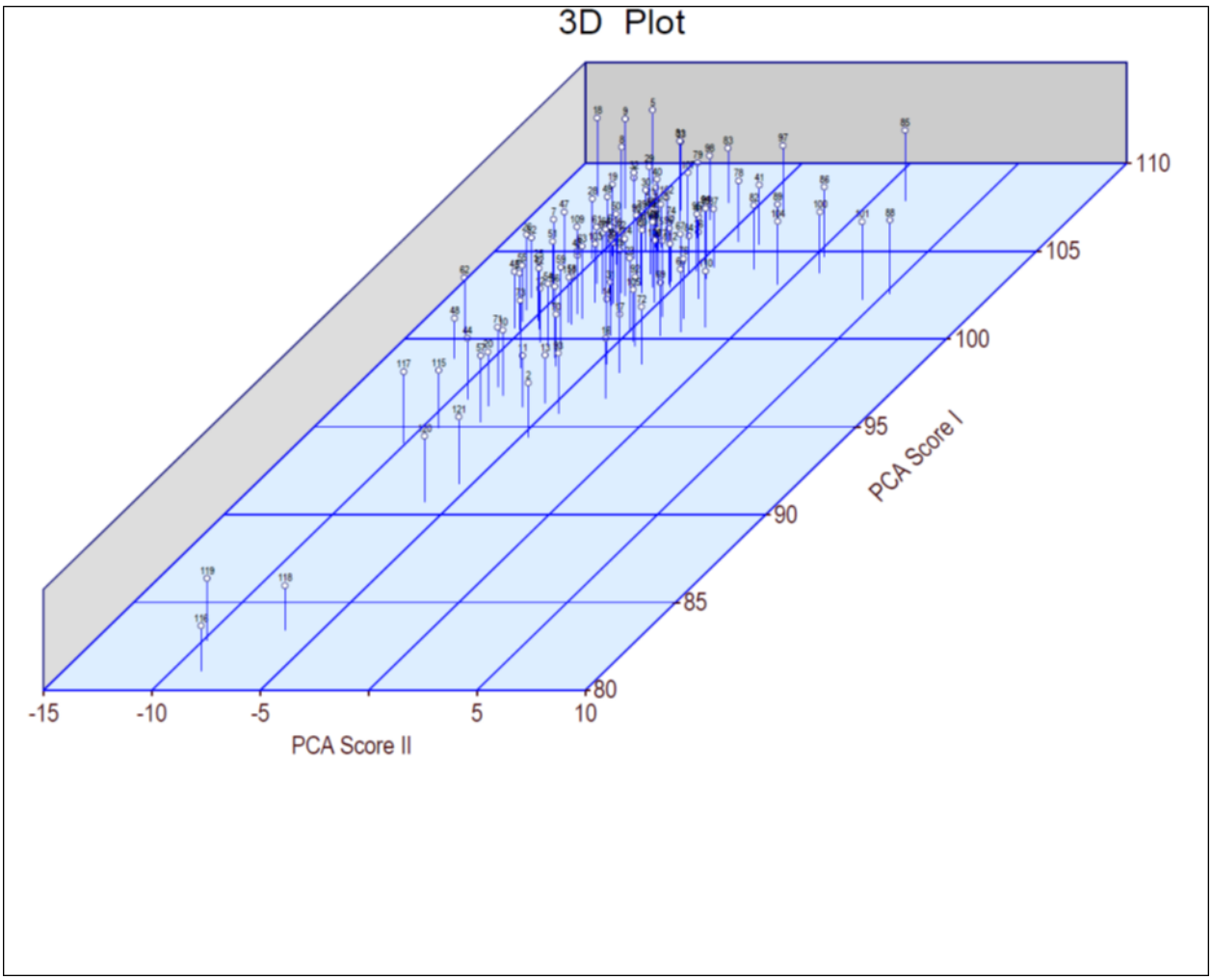


Figure 5. 3D plot showing scattering in 64 genotypes of rice (*Oryza sativa* L.) based on PCA score

The genotypes identified on extreme positive side on both the axis were considered to be the better genotypes i.e. families MTU 2462-1-5-2 (108.855), MTU 2462-15-1-1 (108.138), MTU 2468-30-2-2 (107.887), MTU 2468-29-3-1 (107.801), MTU 2462-1-9-2 (107.460), MTU 2462-31-2-1 (107.318), MTU 2469-23-2-1 (107.220) along PCA I axis and families MTU 2469-6-2-1 (5.226), MTU 2469-36-1-1 (4.229), MTU 2469-32-2-1 (1.028), MTU 2468-31-1-1 (0.470) along PCA II axis.

In 3D plot, families MTU 2462-1-5-2, MTU 2462-15-1-1, MTU 2468-30-2-2, MTU 2468-29-3-1, MTU 2462-1-9-2, MTU 2469-23-2-1, MTU 2469-6-2-1, MTU 2469-36-1-1, MTU 2469-32-2-1 were found to be away from the center. These genotypes might be used in hybridization programmes.

Table 9. PCA scores of 121 lines

Genotypes	PCA X-Axis	PCA Y-Axis	PCA Z-Axis		Genotypes	PCA X-Axis	PCA Y-Axis	PCA Z-Axis
TSM-1	104.559	-8.214	24.191		TSM-99	100.325	-12.506	16.580
TSM-2	94.446	-4.654	15.961		TSM-101	101.181	-7.785	21.442
TSM-3	101.452	-8.228	17.315		TSM-102	102.123	-7.484	21.610
TSM-4	105.250	-7.772	17.603		TSM-103	103.293	-6.227	17.637
TSM-5	108.855	-10.923	21.863		TSM-104	103.041	-6.612	18.620
TSM-6	102.686	-7.768	19.404		TSM-108	102.180	-4.087	19.860
TSM-7	102.969	-10.596	20.089		TSM-110	100.401	-2.594	18.675
TSM-8	105.833	-9.832	26.625		TSM-114	100.178	-3.335	15.856
TSM-9	107.460	-11.032	26.519		TSM-115	98.464	-6.724	15.410
TSM-10	96.791	-7.786	19.426		TSM-116	97.264	-8.419	17.838
TSM-12	96.140	-6.335	15.237		TSM-118	98.572	-2.870	17.144
TSM-16	99.856	-8.629	15.803		TSM-120	98.360	-8.289	20.107
TSM-17	96.346	-5.466	14.384		TSM-128	103.311	-5.444	18.673
TSM-18	98.600	-4.496	19.246		TSM-132	103.151	-5.893	16.211
TSM-19	101.126	-6.017	20.476		TSM-133	101.195	-3.116	17.670
TSM-20	96.606	-2.875	17.966		TSM-134	101.306	-4.201	22.355
TSM-22	98.090	-3.479	17.359		TSM-138	105.548	-4.198	18.071
TSM-29	108.138	-12.891	23.233		TSM-141	105.774	-6.281	22.264
TSM-32	104.756	-9.371	21.132		TSM-146	102.619	-6.220	18.753
TSM-33	96.200	-7.977	16.053		TSM-147	106.732	-7.892	23.804
TSM-35	104.024	-7.001	20.345		TSM-148	104.021	-2.227	18.818
TSM-36	102.630	-7.196	17.976		TSM-149	107.801	-6.567	16.001
TSM-37	100.451	-4.985	21.753		TSM-150	103.735	-4.988	11.177
TSM-39	103.732	-6.475	18.836		TSM-152	107.887	1.532	20.907
TSM-42	101.097	-9.726	17.375		TSM-153	104.692	0.470	20.783
TSM-43	101.671	-10.747	22.308		TSM-164	104.098	-4.148	17.296
TSM-48	101.320	-6.177	19.473		TSM-165	102.570	5.266	21.975
TSM-50	104.605	-10.162	17.701		TSM-166	104.029	-1.142	18.995
TSM-51	106.737	-9.325	16.070		TSM-167	103.295	-7.040	19.628
TSM-54	105.035	-8.065	18.032		TSM-168	103.063	-3.976	20.308
TSM-56	100.177	-5.633	16.059		TSM-169	99.610	-4.031	20.533
TSM-57	106.300	-9.656	16.575		TSM-171	95.760	-4.365	17.975
TSM-58	107.318	-8.345	20.450		TSM-174	103.456	-3.972	21.299
TSM-61	103.945	-6.861	17.923		TSM-175	103.883	-4.746	16.945
TSM-62	101.948	-7.062	19.072		TSM-178	103.109	-5.384	16.690
TSM-63	102.327	-3.371	19.725		TSM-183	107.220	-3.545	19.849
TSM-64	102.068	-7.106	17.866		TSM-184	106.812	-6.593	18.962
TSM-65	102.115	-6.950	21.546		TSM-190	104.263	-4.662	16.598
TSM-66	103.615	-7.061	18.965		TSM-191	103.774	1.028	18.132
TSM-67	105.103	-7.594	20.908		TSM-200	102.267	4.229	23.073
TSM-68	105.371	-3.103	17.781		TSM-204	105.579	-7.550	12.975
TSM-69	100.566	-9.261	18.134		TSM-211	102.114	-7.966	17.289
TSM-72	100.626	-10.439	16.835		TSM-213	103.132	-0.383	18.665
TSM-73	96.594	-9.260	18.063		TSM-215	99.870	-4.334	15.445
TSM-74	101.680	-8.427	17.319		TSM-216	102.080	-5.191	18.569
TSM-76	103.735	-6.758	18.178		TSM-219	105.823	-6.772	19.097
TSM-77	104.291	-11.202	15.372		TSM-220	102.981	-6.025	19.245
TSM-79	98.887	-11.763	12.001		TSM-221	104.263	-10.588	11.133
TSM-82	105.120	-9.901	15.573		TSM-223	100.677	-1.665	16.662
TSM-83	104.240	-8.775	15.010		TSM-228	100.966	-8.211	13.346
TSM-85	101.876	-9.732	19.293		TSM-229	102.935	-5.186	13.025
TSM-86	102.349	-11.102	17.898		TSM-230	105.214	-7.508	12.847
TSM-87	100.536	-10.136	16.850		TSM-235	102.484	-6.906	16.799
TSM-88	99.636	-8.076	18.389		MTU 7029	94.978	-9.229	16.907
TSM-90	101.035	-10.431	16.430		MTU 1081	81.099	-8.632	13.400
TSM-92	98.937	-7.177	21.205		MTU 1121	94.099	-10.116	21.130
TSM-93	95.285	-7.553	19.697		JGL 11118	83.449	-6.730	13.098
TSM-94	100.829	-7.975	14.367		NLR 34449	82.827	-9.804	18.553
TSM-95	99.439	-7.316	24.285		PAU 3116-25-5-1	90.739	-6.358	19.518
TSM-96	102.600	-8.026	17.896		PAU 3140-126-1	91.794	-5.655	19.779
TSM-97	103.198	-8.790	16.585					

Comparative study of D2 analysis and principal component analysis

Mahalanobis' D^2 statistic and Jackson's principal component analysis are the tools for analyzing multivariate data. The advantage of PCA over D^2 analysis is that it reduces the dimensionality of the data set by creating significant principal components which contributed towards maximum variability of the genotypes. In PCA, standardization of data made attributes to contribute equally towards the divergence studies irrespective of the units taken. The principal component analysis sorted only five significant

principal components out of the 10 attributes. The contribution of the main characters for variance was easily identified by the characters loaded on the PC1 with high loading values. PCA facilitates the in-depth analysis for genetic diversity. In D^2 analysis, test weight (23.79) days to maturity (22.40), plant height (19.68), grain yield per plant (16.46) contributed maximum towards divergence. In PCA the characters viz., days to 50% flowering, days to maturity, grain yield per plant, test weight, kernel length and L/B ratio loaded significantly and contributed more towards variability. It can be concluded that characters contributed more towards divergence in D^2 analysis were loaded in PC1 and PC 3. Families in cluster X (MTU 2468-30-2-2, MTU 2468-31-1-1, MTU 2469-6-2-1, MTU 2469-32-2-1, MTU 2469-36-1-1, MTU 2469-42-1-1) and cluster XI (MTU 2462-8-2-1, MTU 2462-12-1-1, MTU 2462-17-3-1, MTU 2462-18-1-1, MTU 2462-29-2-1, MTU 2465-27-2-1, MTU 2466-4-2-1, MTU 2468-1-1-1, MTU 2468-18-1-2, MTU 2468-20-1-1, MTU 2469-7-1-1, MTU 2469-42-3-1, MTU 2469-57-1-2) showed maximum inter-cluster distance in Mahalanobis'D2 analysis, while PCA analysis revealed that families MTU 2462-1-5-2, MTU 2462-15-1-1, MTU 2468-30-2-2, MTU 2468-29-3-1, MTU 2462-1-9-2, MTU 2469-23-2-1, MTU 2469-6-2-1, MTU 2469-36-1-1, MTU 2469-32-2-1 were more divergent.

Conclusion

The conclusion drawn from the cluster analysis is that in the studied population high variability was observed between the genotypes in different clusters for different traits. Recombination breeding among the genotypes belonging to cluster VII having maximum intra cluster distance can improve the yield potential. The maximum inter cluster D2 values was observed between cluster X and XI (931.276) followed by cluster VII and X (814.784). Based on these studies, crosses can be made between the genotypes of cluster X and cluster XI followed by the genotypes of clusters VI and XI. Crosses involving genotypes from these clusters would give wider and desirable recombinants. Principal component analysis revealed that the genotypes identified on the extreme positive sides on both the axis were considered to be the better genotypes in 2D plot i.e. families MTU 2462-1-5-2 (108.855), MTU 2462-15-1-1 (108.138), MTU 2468-30-2-2 (107.887), MTU 2468-29-3-1 (107.801), MTU 2462-1-9-2 (107.460), MTU 2462-31-2-1 (107.318), MTU 2469-23-2-1 (107.220) along PCA I axis and families MTU 2469-6-2-1 (5.226), MTU 2469-36-1-1 (4.229), MTU 2469-32-2-1 (1.028), MTU 2468-31-1-1 (0.470) along PCA II axis.In 3D plot, families MTU 2462-1-5-2, MTU 2462-15-1-1, MTU 2468-30-2-2, MTU 2468-29-3-1, MTU 2462-1-9-2, MTU 2469-23-2-1, MTU 2469-6-2-1, MTU 2469-36-1-1, MTU 2469-32-2-1 were found to be away from the center. These families might be used in hybridization programs. Hence, the results of cluster analysis were supported by principal component analysis.

References

[1] J. Biouf **(2003)**. Rice is life. International Rice Commission News Letter.**52:** 3.
[2] P. C. Mahalanobis **(1928)**. A statistical study at Chinese head measurement. Journal of Asiatic Society of Bengal. **25:** 301-307.
[3] C. R. Rao **(1952)**. Advanced statistical methods in biometric research..
[4] B. K Senapati and G. Sarkar **(2005)**. Genetic divergence in tall indica rice under rainfed saline soil of Sundarban.Oryza.**42:** 70-72.
[5] R. K. Singh and B. D. Chaudhary **(1979)**. Biometrical methods in quantitative genetic analysis. Biometrical methods in quantitative genetic analysis.
[6] S. Yadav, A. Singh, M. R. Singh, N. Goel, K. K. Vinod and T. Mohapatra et al. **(2013)**. Assessment of genetic diversity in Indian rice germplasm (*Oryza sativa* L.): use of random versus trait-linked microsatellite markers. J. Genet., **92:** 545-557.