Original Research Article

Multivariate Analysis to study Genetic Diversity for Yield and its Attributing Traits in Rice (*Oryza sativa* L.)

.

ABSTRACT

The present investigation was carried out with 72 germplasm lines and three checks of rice (*Oryza sativa* L.) were grown at Crop Research Station, Masodha conducted in the GPB farm, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya (U.P.) during *Kharif* June2022- February 2023. Data on 10 characters, including grain yield per plant focused on diversity and PCA analysis. This investigation involved the analysis of 72 rice germplasm lines alongside three checks, showcasing extensive variation in agronomic and morphological traits. The study utilized Non-hierarchical Euclidean cluster analysis to assess genetic diversity. The pseudo F-test determined the optimal grouping of 75 genotypes into six distinct clusters. Cluster distribution revealed varying genotype compositions, with Cluster V comprising the highest entries (20), followed by Clusters I, VI, and II. Intra- and inter-cluster distances illustrated significant variability among clusters, emphasizing genetic diversity. Examining agronomic traits across these clusters revealed noteworthy variations in days to 50% flowering, days to maturity, plant height, and productive tillers. Panicle length, flag leaf area, biological yield, harvest index, 1000-grain weight, and grain yield per plant also exhibited cluster-specific variations. These findings provide valuable insights for rice breeding programs, facilitating targeted enhancements of specific agronomic traits within the rice population, thus contributing to the development of more resilient and productive rice varieties.

Keywords: Diversity, Cluster, F test, PCA, Euclidean

1. INTRODUCTION

"Rice (*Oryza sativa* L.) is a short-day, C3 plant that is cultivated primarily in Asian agroecosystems. It belongs to the family Poaceae (Graminae) and is an annual, semi-aquatic, and self-pollinating crop. There are 25 recognized species in the genus Oryza, of which 23 are wild and two are domesticated. *Oryza sativa* and *Oryza glaberrima* are the cultivars"[1]**.**"There are two main domesticated species of rice: Asian rice (*Oryza sativa* L.) and African rice (*Oryza glaberrima* Steud.), both possessing the genome AA (2n = 24)" [2]. "The data for the current 2022-23 *kharif* season reveals that out of the total 3.59 million hectares under kharif crops, paddy was cultivated on 3.13 million hectares" [3]. "In India, it accounts for 20-25% of agricultural production and ensures food security for over half of the population. Rice production in India constitutes 55% of the total cereal production, with 116.48 million tons of rice being produced in the year 2018-19 from approximately 44.16 million hectares of planted rice land" [4].

"To develop varieties and hybrids based on specific needs, it is important to exploit the unique characters and existing variability present in the germplasm through crop improvement programs" [5]. "An effective breeding program requires a comprehensive understanding of the nature and extent of genetic variability and the association among different traits within a species" [23]. "Quantifying the level of divergence in experimental materials is highly valuable for identifying divergent genotypes that can be utilized in hybridization to generate new variations" [6]. The Mahalanobis D^2 [11] statistic has been widely recognized as a powerful tool for plant breeders to select suitable parental genotypes with a broader range of variability for various traits [7].“The utilization of multivariate analysis tools like principal component analysis (PCA) has been found to be effective in assessing phenotypic diversity, identifying genetically distant Cluster s of genotypes, and selecting key traits that contribute to overall variation in genotypes. PCA enables the natural grouping of genotypes and provides a reliable indicator of differences among them. Consequently, the primary objective of any plant breeding program is to develop improved genotypes that surpass the existing ones in terms of economic yield. This calls for genetic enhancement through optimal utilization of allelic resources to create an ideal genotype" [8].Standard statistical procedures were applied to the recorded data. Genetic divergence analysis was conducted using the D^2 statistics introduced by Mahalanobis and described by Rao [9] along with Principal Component Analysis (PCA).

"Genetic diversity arises either due to geographical separation or due to genetic barriers to cross ability. Genetic diversity plays an important role in plant breeding because hybrids between lines of different origin generally display a greater heterosis than those between closely related strains. Genetic divergence serves as an important parameter for successful selection of parents for hybridization programme" [10]. The D^2 statistic proposed by Mahalanobis in 1936 [11] measures the degree of diversification and determines the relative proportion of each component character to the total divergence. It measures divergence at two levels, namely, intra-cluster and inter-cluster levels, and thus helps in the selection of genetically divergent parents for exploitation in hybridization programmes [12]. It gives a result based on magnitude of divergence and is independent of sample size. The D^2 technique has been used for assessing variability in various crops.

2. MATERIAL AND METHODS

2.1 Experimentalsite

The current field experiment took place at the Crop Research Station (CRS) in Masodha, while certain laboratory experiments were carried out in the Biotechnology Laboratory at the Department of Plant Biotechnology, Narendra Deva University of Agriculture and Technology, Ayodhya (U.P.), India. Geographically, Narendra Nagar is positioned at 26.47°N latitude and 82.12°E longitude, with an elevation of 113 meters above sea level, situated in the Gangetic plains of eastern Uttar Pradesh.

2.2 Experimentaldetails

The present investigation utilized 72 rice germplasm lines, in addition to three checks *viz.* Narendra Usar Dhan-3, Sarjoo-52 and NDR-97 tabulated in (Table 1). These germplasm lines demonstrated a diverse range of variation in both agronomic and morphological traits. The cultivation of these lines took place during the *Kharif* season of 2022, employing an Augmented Block Design.

2.3 Statisticalanalysis

One of the best methods for calculating the genetic distance between genotypes based on allelic frequencies at a sample of loci is the Mahalanobis [11] D^2 statistic. For formation of clusters, Rao [9] followed the general criteria of grouping as suggested by Tocher in the present study. Furthermore, genetic divergence among genotypes was explored through Non-hierarchical Euclidean cluster analysis, following the methods of Beale [13] and Sparks [14]. Observations recorded for 10 traits *viz.* Days to 50% flowering, Days to maturity, Plant height (cm), No. of Productive Tillers/plant, Panicle length (cm), Flag leaf area (cm^2), Biological yield /plant (g), Harvest Index (%), 1000-grain weight (g), Grain yield/plant (g).

Table1: Listof72 Germplasms of rice used in investigation

Listof72 Germplasms of rice				
Bee	Geetu-II	Bilaspur	Kalakand	101/5115
Singal	Sarye	Yuvraj	K-14	102/5110

Rajmunia	Lindhinia-I	Vijets	CH-1039	103/5119
Benibhog	Dilaora Local	Palhari	KMP-41	104/1503
Kolamden	Chingard-I	Padheri	KSR-4111	106/5111
Mahechandour	Tarochani	Mangal	T-65	107/5106
Ujrous	Indrajon	Harikesh	KR-103-96	108/5113
Basahwa	Sahdlia	PR-6	GMP-12	109/5114
Mangole	Darahi	NC-495	CR-209	110/5107
Harikrishna	Turahwa	H-118	TN-1	111/5109
Bakani	Loungchoor	NC-491	KSR	112/5108
Farmahwa	Local-chl-Fzd	NC-492	CR-118	113/5127
Orissa-2	Katari bhog	T-1242(I)	W-418	Sarjoo-52©
Nandi-C	Bans	N-10-B	T-116	NDR-97©
Kerhani Local-B1	Sirdgarh	Nanital (saryu)	W-496	Narendra Usar Dhan-3 ©

3. RESULTS AND DISCUSSION

3.1 Genetic Divergence Analysis

The Non-hierarchical Euclidean cluster analysis was employed to assess the genetic diversity among 72 rice germplasm collections, including three checks, based on 10 quantitative characters. The pseudo F-test indicated that grouping the 75 genotypes into six clusters was the most effective. Thus, the 75 genotypes were classified into six non-overlapping clusters. The distribution of rice lines in these clusters is detailed in Table 2, with cluster V containing the highest number of genotypes (20 entries), followed by clusters I, VI, and II with 18, 17, and 11 entries, respectively. Clusters III and IV comprised 3 and 6 entries of germplasm, respectively.

Intra- and inter-cluster distances, presented in (Table 3), revealed that cluster V had the minimum intra-cluster distance (1.857), while the highest intra-cluster distance was in cluster III (3.244), followed by cluster II (2.443). The highest inter-cluster distance was observed between cluster VI (7.895), followed by Cluster IV (7.881). The minimum inter-cluster distances were found between cluster V and I (1.973), VI (2.101), followed by cluster II (2.137), and cluster V (2.403), and cluster VI (2.552). The intra-cluster group means for the 10 characters (Table 4) indicated significant variances between the clusters. The genotypes present in different Clusters showed high degree of diversity than the genotypes present in the same Cluster Genotypes from same geographic location fell into different Clusters indicating that Clustering of genotypes did not follow their geographic or location distribution. These findings are in conformity with the reports of [15-18].

3.2 Estimates of cluster means

1. ***Days to 50% flowering***
   For days to 50% flowering the cluster mean alters from 96.44 days (cluster I) to 105.17 days (cluster IV). Maximum mean for days to 50% flowering was observed for cluster IV (105.17days) followed by cluster II (99.00 days) and cluster III (98.70 days) while for cluster-I (96.44 days) followed by cluster V (97.25 days) minimum mean was recorded.

2. ***Days to maturity***
   For days to maturity cluster mean ranged from 120.28 (cluster I)to 130.00 (cluster IV), maximum mean observed inside cluster IV(130.00 days) succeeding by cluster VI(126.82) and cluster III (126.04) while minimum mean was founding cluster I (120.28) followed by cluster II (122.82) and cluster V (123.60).

3. ***Plant height (cm)***

For plant height, the cluster mean ranged from cluster III (10.52cm) to cluster II (149.60cm). Maximum mean for plant height was recorded for cluster II (149.60cm) followed by cluster V (137.40) and cluster VI (134.20 cm) while, minimum mean was recorded for cluster III (10.52cm) followed by cluster IV (109.78cm).

4. ***Number of productive tillers per plant***

   For the panicle bearing tillers/plant, cluster mean ranged from cluster I (9.79) to cluster III (27.94). Maximum mean for the panicle bearing tillers/plant was recorded for cluster III (27.97) followed by cluster II (10.95) and cluster IV (10.80) while minimum mean was recorded for cluster I (9.79) followed by cluster V (10.20).

5. ***Panicle Length (cm)***

   For the panicle length, cluster mean ranged from 25.56cm (cluster II) to 29.62 cm (cluster VI). Maximum mean for the panicle length was recorded for cluster VI (29.62 cm) followed by cluster IV (29.09 cm) and cluster I (27.64), while minimum mean was recorded for cluster II (25.56 cm) followed by cluster III (25.86).

6. ***Flag leaf area (cm^2)***

   For flag leaf area, the cluster mean ranged from cluster IV (31.13cm^2) to cluster III (129.91cm^2). Maximum mean for flag leaf area was recorded for cluster III (129.91 cm^2) followed by cluster II (44.61cm^2) and cluster I (37.89 cm^2), while minimum mean was recorded for cluster IV (31.13 cm^2) followed by cluster VI (32.91cm^2).

7. ***Biological yield/plant (g)***

   For the biological yield/plant, cluster mean ranged from 32.86g (cluster II) to 52.72g (cluster IV). Maximum mean for the biological yield/plant was recorded for cluster IV (52.72 g) followed by cluster V (43.92g) while minimum mean was recorded for cluster II (32.86g) followed by cluster VI (34.96 g).

8. ***Harvest index (%)***

   For the harvest index, cluster mean ranged from 26.51% (cluster V) to 64.75% (cluster II). Maximum mean for the harvest index was recorded for cluster II (64.75%) followed by cluster I (47.63%) while minimum mean was recorded for cluster V (26.51 %) followed by cluster IV (33.55%).

9. ***1000-grainweight (g)***

   For 1000-grain weight, cluster mean ranged from 22.49g (cluster IV) to 23.65g (cluster V).Maximum mean for the 1000-grain weight was recorded for cluster V (23.65g) followed by cluster I (23.27g), while minimum mean was recorded for cluster IV (22.49g) followed by cluster II (22.55 g).

10. ***Grain Yield/Plant (g)***

    For the grain yield /plant, cluster mean ranged from 15.87g (cluster III) to 21.14g (cluster II). Maximum mean for the grain yield /plant was recorded for cluster II (21.14g) followed by cluster IV (17.32g) while minimum mean was recorded for cluster III (15.87g) followed by cluster I (18.81g).

The study investigated the variation in agronomic traits across six clusters in a rice population. Notably, days to 50% flowering ranged from 96.44 days (cluster I) to 105.17 days (cluster IV), with the highest mean observed in cluster IV. Days to maturity showed a range of 120.28 (cluster I) to 130.00 days (cluster IV). Plant height exhibited significant variability, with cluster II having the maximum mean (149.60 cm) and cluster III the minimum (10.52 cm). The number of productive tillers per plant varied from 9.79 (cluster I) to 27.94 (cluster III), with the highest mean in cluster III. Panicle length, flag leaf area, biological yield per plant, harvest index, 1000-grain weight, and grain yield per plant also demonstrated cluster-specific variations, indicating the diverse genetic composition of the rice population. These findings provide valuable insights for rice breeding programs aiming to enhance specific agronomic traits. The inter-cluster distances were higher than the intra-cluster distances which indicate the existence of substantial diversity among the genotypes. Similar results of inter and intra cluster distances in rice were reported earlier by [20,22]. Therefore, judicious combination of selected genotypes from the above divergent clusters may be carried out to obtain desirable segregants with respect to nutritional properties coupled with high yield potential. The results are in broad agreement with the reports of [19] and [21].

Table 2: Clustering pattern of 75 (72+3) rice genotypes based on non-hierarchical Euclidean cluster analysis

Cluster	No of genotypes	Genotypes
Cluster I	**18**	Bee, Singal, Rajmunia, Benibhog, Ujrous, Orissa-2, Chingard-1, Tarochani, Yuvraj, Mangal, PR-6, H-118, Nainital (Saryu), K-14, CH-1039, KMP-41, KSR-411, T-65,
Cluster II	**11**	Kolamden, Mahechandour, Basahwa, Mangole, Harikrishna, Bakani, Farmahwa, Nandi-C, Karhani Local-B1, Sarya, Dilaora Local
Cluster III	**3**	Sarjoo-52 (check), NDR-97 (check), Narendra Usar Dhan-3 (check)
Cluster IV	**6**	Darhahi, Turahwa, Loungchoor, Local-chl-Fzd, Padheri, NC-491
Cluster V	**20**	Geetu-II, KR-103-96, GMP-12, CR-209, TN-1, KSR, CR-118, W-418, T-116, T-116, W-496, 101/5115, 102/5110, 103/5119, 104/1503, 106/5111, 107/5106, 108/5113, 109/5114, 111/5109, 112/5108
Cluster VI	**17**	Lindhinia-1, Indrajon, Sahdalia, Katri Bhog, Bans, Sirdgarh, Bilaspur, Vijets, Palhari, Harikesh, NC-495, NC-492, T-124(I), N-10-B, Kalakand, 110/5107, 113/5127

Table 3: Estimates of average intra and inter-cluster distances for the six clusters in rice genotypes

Clusters	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI
Cluster I	1.973					
Cluster II	2.137	2.443				
Cluster III	7.634	7.797	3.244			
Cluster IV	3.402	4.457	7.881	2.204		
Cluster V	2.403	3.909	7.829	3.160	1.857	
Cluster VI	2.552	3.357	7.895	2.910	2.277	2.101
Bold figures indicate the intra-cluster distance						

Table 4: Estimates of cluster means for 10 characters of Rice genotypes by D^2 analysis

Characters		Days to 50% flowering	Days to maturity	Plant height (cm)	No. of Productive Tillers/plant	Panicle length (cm)	Flag leaf area (cm2)	Biological yield /plant (g)	Harvest Index (%)	1000-grain weight (g)	Grain yield/plant (g)
I	**Mean**	96.44	120.28	127.11	9.79	27.64	37.89	39.91	47.63	23.27	18.81
	± SE	3.85	5.50	16.34	1.00	2.19	8.14	5.33	5.21	1.58	1.73
II	**Mean**	99.00	122.82	149.60	10.95	25.56	44.61	32.86	64.75	22.55	21.14
	± SE	2.53	6.26	31.94	1.20	1.85	13.81	4.91	11.21	2.23	4.32
III	**Mean**	98.70	126.04	10.52	27.94	25.86	129.91	37.84	41.95	23.07	15.87
	± SE	4.44	21.52	0.58	2.23	2.18	3.55	9.62	4.59	2.14	4.53
IV	**Mean**	105.17	130.00	109.78	10.80	29.09	31.13	52.72	33.55	22.49	17.32
	± SE	3.60	5.90	15.39	1.28	2.56	4.86	5.73	8.05	2.55	2.66
V	**Mean**	97.25	123.60	137.40	10.20	26.65	36.86	43.92	26.51	23.65	11.52
	± SE	1.89	4.96	18.47	1.45	2.92	5.09	4.09	4.50	1.67	1.47
VI	**Mean**	102.41	126.82	134.24	10.26	29.62	32.91	34.96	36.94	23.13	12.80
	± SE	3.02	5.73	15.43	2.06	2.59	8.36	3.96	6.77	2.04	2.23

4. CONCLUSION

The Non-hierarchical Euclidean cluster analysis effectively classified the genotypes into six distinct clusters, emphasizing the diverse genetic composition within the rice germplasm. The significant intra- and inter-cluster variations observed underscore the importance of understanding genetic diversity for targeted breeding efforts.A closer examination of agronomic traits across clusters revealed substantial differences in key characteristics such as days to 50% flowering, days to maturity, plant height, and productive tillers. Cluster-specific variations in panicle length, flag leaf area, biological yield, harvest index, 1000-grain weight, and grain yield per plant further highlight the potential for tailored improvements in specific traits.These findings provide a foundation for rice breeding programs aiming to enhance desired agronomic traits, contributing to the development of more resilient and productive rice varieties. The identified clusters offer a roadmap for targeted selection and hybridization, emphasizing the need for precision in breeding strategies to meet evolving agricultural challenges and ensure food security.

REFERENCES

1. Ndjiondjop MN, Wambugu P, Sangare JR, Dro T, Kpeki B, Gnikoua K. Oryza glaberrima Steud. The wild Oryza genomes. 2018:105-26.
2. Sanchez PL, Wing RA, Brar DS. The wild relative of rice: genomes and genomics. InGenetics and genomics of rice 2013. (pp. 9-25). New York, NY: Springer New York.
3. Anonymous (2023). Mongabay news, Retrieved on 4 April, 2023 from. https://india.mongabay.com/2023/01/punjabs-paddy-dilemma
4. AGRISTAT (2019). Agriculture statistics at a glance: Retrieved on 13 March, 2023 from https://eands.dacnet.nic.in
5. Swarup S, Cargill EJ, Crosby K, Flagel L, Kniskern J, Glenn KC. Genetic diversity is indispensable for plant breeding to improve crops. Crop Science. 2021;61(2):839-52.
6. Thompson KA. Experimental hybridization studies suggest that pleiotropic alleles commonly underlie adaptive divergence between natural populations. The American Naturalist. 2020;196(1):E16-22.
7. Sinha N, Singh D, Mohanty TA. Study of genetic divergence in lowland rice genotypes of Bihar. Int. J. Curr. Microbiol. App. Sci. 2020;9(2):88-94.
8. Pratap N, Singh PK, Shekhar R, Soni SK, Mall AK. Genetic variability, character association and diversity analyses for economic traits in rice (Oryza sativa L.). SAARC Journal of Agriculture. 2012;10(2):83-94.
9. Rao CR. 1952. Advanced statistical methods in biometric research. New York: John Wiley & Sons.
10. Pandey SK, Dasgupta T, Rathore A, Vemula A. Relationship of parental genetic distance with heterosis and specific combining ability in sesame (*Sesamum indicum* L.) based on phenotypic and molecular marker analysis. Biochemical genetics. 2018;56:188-209.
11. Mahalanobis PC. On the generalized distance in statistics. Sankhyā: The Indian Journal of Statistics, Series A (2008). 2018;80:S1-7.
12. Sruthi K, Divya B, Senguttuvel P, Revathi P, Kemparaju KB, Koteswararao P, Sundaram RM, Singh VJ, Ranjith Kumar E, Bhowmick PK, Vinod KK. Evaluation of genetic diversity of parental lines for development of heterotic groups in hybrid rice (*Oryza sativa* L.). Journal of Plant Biochemistry and Biotechnology. 2020;29:236-52.
13. Beale EM. Euclidean cluster analysis. Scientific Control Systems Limited; 1969.

14. Sparks DN. Euclidean cluster analysis. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1973;22(1):126-30.
15. Ashok S, Jyothula DP, Ratnababu D. Genetic divergence studies for yield, yield components and grain quality parameters in rice (Oryza sativa L.). Electronic Journal of Plant Breeding. 2017;8(4):1240-6.
16. Prasad RKK, Suneetha Y, Srinivas T. Genetic diversity studies in rice (*Oryza sativa* L.). Electronic Journal of Plant Breeding. 2018; 9(4): 1335-1341.
17. Devi PS, Veni BK, Raja DS. Genetic Diversity for Yield Components, Physical and Biochemical Quality Parameters in Colored Rice (*Oryza sativa* L.). Int. J. Curr. Microbiol. App. Sci. 2019;8(11):18-26.
18. Sudeepthi K, Srinivas T, Ravikumar BNVSR, Jyothula DPB, Nafeez Umar SK. Genetic divergence studies for yield and yield component traits in rice (*Oryza sativa* L.). Multilogic in science. 2020b; 9:415-418.
19. Behera PP, Singh SK, Singh DK, Reddy YS, Habde S, Khaire A, Ashrutha MA. Genetic diversity analysis of rice (*Oryza sativa* L.) genotypes with high grain zinc content for yield and yield traits. Journal of Pharmacognosy and Phytochemistry. 2018;7(4):1319-23.
20. Singh KS, Suneetha Y, Kumar GV, Rao VS, Raja DS, Srinivas T. Genetic divergence studies for yield and quality traits in coloured rice. Journal of Pharmacognosy and Phytochemistry. 2020;9(4):1234-40.
21. Beevi HA, Venkatesan M. Genetic divergence studies in rice (*Oryza sativa* L.) genotypes under natural saline condition. Journal of the Andaman Science Association. 2015; 20(1):35-38.
22. Archana RS, Sudha Rani M, Vishnu Vardhan KM and Fareeda G. Genetic diversity studies among rice (*Oryza sativa* L.) Genotypes for grain yield, yield components and nutritional traits in rice International Journal of Chemical Studies. 2018; 6(6): 134-137.
23. Heera PK, Ram M, Kumar R, Murali S, Kumar AS. Analysis of Genetic Variability, Heritability and Genetic Advance for Yield and Yield Associated traits in Rice (*Oryza sativa* L.). Eco. Env. & Cons. 2023;29 (October Suppl. Issue): 160-163