International Journal of Bio-resource and Stress Management https://ojs.pphouse.org/index.php/IJBSM

IJBSM June 2024, 15(6): 01-07

Article AR5312a

Research Article Natural Resource Management

DOI: HTTPS://DOI.ORG/10.23910/1.2024.5312a

Genetic Diversity Analysis in Rice Maintainer Lines Using K-means Clustering for Yield and Yield Attributes

B. Edukondalu[1], V. Ram Reddy[1], T. Shobha Rani[1], CH. Aruna Kumari[2] and B. Soundharya[3]

[1]Dept. of Genetics and Plant Breeding, Agriculture College, Jagtial, PJTSAU, Telangana (505 529), India
[2]Dept. of Crop Physiology, Agriculture College, Jagtial, PJTSAU, Telangana (505 529), India
[3]Dept. of Genetics and Plant Breeding, Regional Sugarcane and Rice Research Station, PJTSAU, Rudrur, Nizamabad, Telangana (503 188), India

Open Access

Corresponding edukondalu0208@gmail.com

0000-0002-1528-9690

ABSTRACT

The experiment was conducted during the *kharif* season (June–October, 2016) at the Regional Agricultural Research Station, Jagtial, Telangana, India to evaluate 40 maintainer lines for genetic diversity using K-means cluster analysis for grain yield and its yield attributes of rice. Multivariate analysis was performed in Randomized Block Design (RBD) with two replications to comprehend the nature and magnitude of genetic divergence among the 40 maintainer lines of rice. Based on the K-means cluster analysis, the optimal number of clusters formed was seven. The highest number of genotypes were grouped in cluster I (27), followed by cluster IV (7). Intra-cluster distances varied among the clusters with the lowest recorded in cluster I (8.64) followed by cluster III (33.64) and cluster IV (75.69), while inter-cluster distances were maximum between cluster II and VI (142.32) followed by cluster VI and VII (113.42) indicating greater genetic diversity between certain clusters. This diversity suggests potential for maximizing heterosis through hybridization between genotypes from the most divergent clusters. Cluster mean analysis revealed significant differences in traits among clusters with some clusters exhibiting consistently higher or lower values for specific characteristics. Kernel length, days to flowering, hulling percentage, and plant height are major contributors to genetic divergence. Certain maintainer lines, namely, B18, CMS46B, B42, and B88 from different clusters with suitable trait means, can be selected for use in crossing programs to create desirable variations in the breeding material for yield improvement.

KEYWORDS: Genetic diversity, K-means clustering, maintainer lines, rice

***Citation* (VANCOUVER):** Edukondalu et al., Genetic Diversity Analysis in Rice Maintainer Lines Using K-means Clustering for Yield and Yield Attributes. *International Journal of Bio-resource and Stress Management,* 2024; 15(6), 01-07. HTTPS://DOI.ORG/10.23910/1.2024.5312a.

Data Availability Statement: Legal restrictions are imposed on the public sharing of raw data. However, authors have full right to transfer or share the data in raw form upon request subject to either meeting the conditions of the original consents and the original research study. Further, access of data needs to meet whether the user complies with the ethical and legal obligations as data controllers to allow for secondary use of the data outside of the original study.

Conflict of interests: The authors have declared that no conflict of interest exists.

RECEIVED on 18th March 2024 RECEIVED in revised form on 29th May 2024 ACCEPTED in final form on 20th June 2024 PUBLISHED on 29th June 2024

1. INTRODUCTION

Rice (*Oryza sativa* L.) stands as one of the primary food crops and makes a significant contribution to global economic expansion. Nearly two-thirds of the world's population relies on it as their main crop, placing it third globally in production after maize and wheat (Bhandari et al., 2021; Ghazy et al., 2021). India holds the top position in rice cultivation, covering an area of 45.7 mha and ranking second in production with 124.36 mt in the year 2020–2021 Soujanya et al., 2023. It is estimated that 70% to 100% increase in the cereal food supply is required by 2050 to feed the predicted world population of 9.8 billion people Godfray et al., 2010. Rapid population growth may pose a threat of severe food shortage around the world, including India, in the near future due to shrinking of natural resources. These demands increase in production of staple cereal crops like rice in the realm of dwindling land and water resources and a rapidly changing climate Sruthi et al., 2019. Another concern is the lack of sufficient genetic diversity among the parental lines of hybrids, which could be a contributing factor to the limitations encountered in breeding programs and the adoption of hybrid rice technology in the tropics Xie et al., 2012. Hybrid rice technology offers a great promise to produce 15 to 20% more yield in comparison to pure line varieties Soujanya et al., 2023. The adoption of hybrid rice technology has been slower than expected in the country due to certain challenges like moderate levels of heterosis, susceptibility of the hybrids to biotic and abiotic stresses, grain quality issues, higher seed cost of rice hybrids and the lack of parental lines with specific desirable traits for developing heterotic rice hybrids Ponnuswamy et al., 2020. Narrow genetic base of hybrid rice parental lines is the major constraint to develop heterotic hybrids in the tropics Xiao et al., 1996a, chapter 10. Improvement of parental lines of hybrid rice based on diversity we can find most diverse parents that will help developing heterotic hybrid combinations. To enhance heterosis levels and broaden the genetic diversity among parents, Hari Prasad et al., 2018. Therefore, knowledge of genetic diversity of parental lines is essential for judicious use of germplasm and to select best heterotic parents for hybridization. To breed heterotic rice hybrids, parental lines should be genetically diverse with high yielding ability and they must be good combiners. Careful selection of maintainer and restorer lines on the basis of their genetic diversity may lead to the development of hybrids with higher yield potential than parents and standard check varieties Julfiquar et al., 1985.

The development of new high-yielding and high-quality rice varieties or hybrids, surpassing the existing ones, heavily relies on the level of genetic variability in the population. Therefore, conducting genetic divergence studies becomes crucial for selecting suitable parents in hybridization programs. Assessing genetic diversity in terms of the nature and extent of variation in grain yield, its components, and quality traits aids plant breeders in identifying appropriate parent combinations for hybridization efforts and devising effective breeding strategies aimed at creating high-yield varieties with desirable grain quality. Mahalanobis D2 statistics serves as a valuable tool for quantifying the extent of genetic divergence, enabling breeders to estimate the genetic divergence in the population for use in plant breeding programmes.

Cluster analysis is a multivariate analysis that have the function of minimizing differences within clusters and maximizing differences between clusters Oliveira et al., 2016. Cluster analysis can be used to classify genotypes and determine the best cluster Kozak et al., 2008. Cluster analysis had been done for selection of potential genotypes by Rashmi et al. (2017) based on estimation of genetic diversity among traditional landraces. Various studies have been reported based on cluster analysis for estimation of genetic diversity in rice (Roy and Sharma, 2014, Pillai et al., 2020, Naik et al., 2021, Nagamani et al., 2022). Keeping in view, the present investigation was conducted to evaluate 40 maintainer lines for genetic diversity using K-means cluster analysis for grain yield and its yield attributes of rice.

2. MATERIALS AND METHODS

The experiment was conducted at the Regional Agricultural Research Station, Jagtial, and Telangana, India, during *kharif*, 2016 (June–October). 40 genotypes (Table 1) were laid in Randomized Block Design (RBD) with two replications and a spacing of 20×15 cm^2 twenty-eight days old seedlings were transplanted in the main field and all the necessary package of practices were followed

Table 1: 40 List of maintainer lines used in the study and their source

Sl. No.	Genotype	Source
1.	B1	RARS, Jagtial
2.	B2	RARS, Jagtial
3.	B3	RARS, Jagtial
4.	B4	RARS, Jagtial
5.	B5	RARS, Jagtial
6.	B6	RARS, Jagtial
7.	B7	RARS, Jagtial
8.	B8	RARS, Jagtial
9.	B9	RARS, Jagtial
10.	B11	RARS, Jagtial
11.	B12	RARS, Jagtial
12.	B13	RARS, Jagtial

Table 1: Continue...

Sl. No.	Genotype	Source
13.	B15	RARS, Jagtial
14.	B16	RARS, Jagtial
15.	B17	RARS, Jagtial
16.	B18	RARS, Jagtial
17.	B19	RARS, Jagtial
18.	B20	RARS, Jagtial
19.	B22	RARS, Jagtial
20.	B23	RARS, Jagtial
21.	B24	RARS, Jagtial
22.	B35	RARS, Jagtial
23.	B42	RARS, Jagtial
24.	B86	RARS, Jagtial
25.	B88	RARS, Jagtial
26.	JMS11B	RARS, Jagtial
27.	JMS13B	RARS, Jagtial
28.	JMS14B	RARS, Jagtial
29.	JMS17B	RARS, Jagtial
30.	JMS18B	RARS, Jagtial
31.	JMS19B	RARS, Jagtial
32.	JMS20B	RARS, Jagtial
33.	JMS21B	RARS, Jagtial
34.	CMS11B	IRRI, Philippines
35.	CMS14B	IRRI, Philippines
36.	CMS23B	IRRI, Philippines
37.	CMS46B	IRRI, Philippines
38.	CMS52B	IRRI, Philippines
39.	CMS59B	IRRI, Philippines
40.	CMS64B	IRRI, Philippines

to raise a healthy crop. Observations were recorded for yield, yield attributing characters and quality traits on five randomly selected competitive plants for each entry in each replication for 15 characters viz., days to 50% flowering (DFF), days to maturity (DM), plant height (PH), number of tillers plant^{-1} (NTPP), panicle length (PL), number of grains panicle^{-1} (NGPP), kernel length (KL), kernel breadth (KB) L/B ratio (L/B), 1000 grain weight (TW), hulling percentage (HL), milling percentage (ML), head rice recovery percentage (HRR), grain yield plant^{-1} (GYP), bran oil percentage (BO). The 40 maintainer lines were classified following model-based k-means clustering approach as suggested by Mac queen, 1967. cluster analysis was carried out in R 4.1.2 using the "Nbclust", "factoextra", "ggplot2" Wickham (2016) and "circlize" Gu et al., 2014.

Statistical analysis using k-means clustering 40 maintainer lines were classified following 'k-means clustering' model as explained by Macqueen (1967) and Forgy (1965) to unravel organization of variability using R 4.1.2 software.

$$J= \sum m_{i=k}\sum^{k}_{k=1}\boldsymbol{\omega}_{ik} II\ X^{2}- \mu_{k}\ II^{2} \ldots\ldots\ldots\ldots (1)$$

Where, $IIX^{2}- \mu_{k}II^{2}$- Indicator of distance between n data point from their respective cluster center.

xi: Number of data points.

μk: Number of cluster centre

3. RESULTS AND DISCUSSION

3.1. Genetic diversity study using k-means clustering

K-means clustering is an iterative algorithm that intends to partition the dataset into k distinct non-overlapping clusters where each data point belongs to the cluster with nearest mean. K-means is a centroid-based clustering algorithm. 'K' represents the number of clusters, and it is also an input parameter Kanavi et al., 2020. In K-means clustering, the optimal clustering is the one with the smallest amount of variation within clusters, which is calculated using the within-clusters sum of squares. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster. This method produces exactly k different clusters of greatest possible distinction (Meirmans and Van Tienderen, 2004).

3.2. Grouping of genotypes into various clusters

Clustering of 40 maintainer lines using K-means cluster analysis based Mahalanobis genetic distance resulted in seven distinct clusters. The genotypes within each cluster were closer to each other than the genotypes belonging to different clusters. Highest number of maintainers were included in cluster I (27 maintainer lines) followed by cluster IV (seven maintainer lines), cluster III (two maintainer lines), II, V, VI and VII each cluster had one maintainer line in (Table 2 and Figure 1). Similar clusters reported Sreedhar et al., 2018. The mode of distribution of genotypes from different geographical regions into various clusters was at random indicating that the genotypes originating from different agro-climatic/geographical regions grouped together into different clusters showing no parallelism between genetic diversity and geographical distribution. Similar kind of results were reported by (Pandey and Kaushik, 2011, Rukmini Devi et al., 2020 and Nagamani et al., 2022). Shanmugam and Rangasamy, 1982 reported that grouping of material of same geographical origin into different clusters was an indication of the broad genetic base of genotypes belonging to that origin.

Table 2: Clustering pattern among 40 maintainer lines of rice

Clusters	Frequency	Genotypes
I	27	B1,B2,B3,B4,B5,B6,B7,B8,B9,B11, B12, B13, B15, B16, B17, B19, B20, B22, B24 B86, JMS11B, JMS21B, CMS11B, CMS14B, CMS23B, CMS52B, CMS59B
II	1	B18
III	2	B23, CMS64B
IV	7	B35, JMS13B, JMS14B JMS17B, JMS18B, JMS19B, JMS20B
V	1	B42
VI	1	B88
VII	1	CMS46B

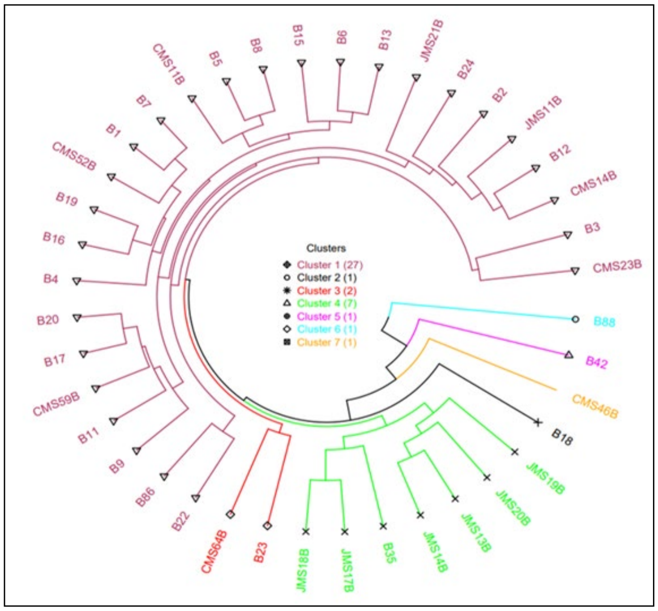

Figure 1: Clustering of 40 maintainer lines (each cluster denoted by different color)

3.3. *Average intra and inter cluster distances*

The average intra- and inter-cluster distances, are presented in Table 3. Intra-cluster distances varied among the clusters, with the lowest recorded in cluster I (8.64), followed by cluster III (33.64) and cluster IV (75.69). Cluster I exhibiting the lowest distance, suggesting a higher degree of genetic similarity within that cluster. This might be unfavorable for hybrid breeding as it limits the potential for heterosis. If cluster I lines possess desirable traits like high yield or disease resistance, they can be strategically crossed with maintainer lines from other clusters with wider genetic divergence. This can introduce the beneficial traits from cluster I while maintaining genetic diversity in the resulting hybrids.

Table 3: Average inter-cluster values for fifteen characters in 40 maintainer line of rice

Clusters	I	II	III	IV	V	VI	VII
I	8.64	21.43	13.69	13.10	45.56	68.06	37.08
II		0	45.42	34.45	57.15	142.32	48.86
III			33.64	19.00	79.92	74.82	48.02
IV				75.69	57.60	84.27	38.56
V					0	82.08	47.19
VI						0	113.42
VII							0

In contrast, inter-cluster distances exhibited greater variation, ranging from 13.10 to 142.32 between clusters IV and VI. The inter cluster distances was maximum observed between cluster II and VI (142.32) followed by cluster VI and VII (113.42), IV and VI (84.27) and V and VI (82.08). The greater the distance between two clusters, the more genetic diversity exists between genotypes. Genotypes in the most divergent clusters might be capable of maximizing heterosis. This suggests that hybridization between the maintainer lines of cluster II (B18) and cluster VII (B88), cluster VI (B88) and cluster VII (CMS 46B), cluster IV (B35, JMS13B, JMS14B JMS17B, JMS18B, JMS19B, and JMS20B), cluster VI (B88) and cluster V (B42) and cluster VI (B88) may be used as parents in a hybridization programme to provide breeding material with significant genetic diversity. Similarly reported by Manasa et al., 2023. The genotypes from distant clusters exhibit wide diversity Falconer, 1981. Maintainer lines act as a bridge between diverse restorer lines and the CMS lines used for hybrid development. Breeders can strategically select maintainer lines from distant clusters exhibiting high inter-cluster distances. These lines could be enabling crosses with restorer lines from different clusters, further amplifying genetic variation in the resulting hybrids.

3.4. *Cluster means of the character*

Each cluster has an average value for each measured trait, known as the cluster mean. If a cluster exhibits consistently higher or lower values for desirable traits like yield or disease resistance, it becomes a potential source for selecting parent varieties. The differences in cluster means for specific traits highlight which characteristics vary most significantly between groups, guiding breeders towards traits with the highest potential for improvement.

Clusters mean values for all characters are presented in (Table 4), highest scoring for days to 50% flowering was observed in cluster VII, while the lowest was in cluster VI. Days to maturity showed the highest reading in cluster VII and the lowest in cluster I. Plant height exhibited its highest

Table 4: Clusters means of 40 maintainer lines for 15 traits

	DF	DM	PH	NTPP	PL	NGPP	KL	KB	LB	GYPP	TW	HU %	ML %	HRR %	BO %
I	83.67	117.76	99.86	12.44	24.69	124.93	6.89	1.93	3.58	24.7	20.84	83.66	75.69	63.7	14.42
II	87.5	122.5	117.5	16.7	27.9	85.5	7.45	1.85	4.03	29.6	20.85	85.92	78.4	51.73	14.29
III	85.75	121.5	88.45	10.95	21.03	130.2	7.25	1.65	4.4	22.38	17.03	84.44	73.91	64.07	16.43
IV	91.21	125.93	101.36	13.29	23.89	246.87	6.12	1.75	3.51	26.19	14.85	83.93	78.59	65.53	15.84
V	98.5	131.5	151	14.1	25.8	76.5	6.65	2.55	2.61	26.75	24.07	83.8	71.75	69.87	18.00
VI	74.5	118.5	93.3	7.7	17.5	84.73	5.55	2.7	2.06	19	19.87	75.11	69.75	68.69	18.95
VII	118.5	149.5	96.6	13.3	25.2	121.8	7.15	2.05	3.49	26.1	21.75	84.89	71.14	64.81	14.00

reading in cluster V and it's lowest in cluster III. Number of tillers^{-1} highest reading in cluster II and the lowest in cluster VI. Panicle length was highest in cluster II and lowest in cluster VI. The highest number of grains^{-1} was recorded in cluster IV, while the lowest was in cluster V. Kernel length was highest in cluster II and lowest in cluster VI. Kernel width showed its highest reading in cluster VI and its lowest in cluster III. 1000 grain weight was recorded highest in cluster IV and lowest in cluster V. L/B ratio highest values in cluster III and their lowest in cluster VI. Grain yield per plant was highest in cluster II and lowest in cluster VI. Test weight was highest in cluster V and lowest in cluster IV. Hulling percentage was highest in cluster II and lowest in cluster VI. Milling percentage was highest in cluster IV and lowest in cluster VI. Head rice recovery percentage Test weight was highest in cluster V and lowest in cluster II. Bran oil percentage was highest in cluster VI and lowest in cluster VII.

The diversity among genotypes was also supported by the appreciable amount of variation among the cluster means for different characters. Cluster II with single genotype B18 represented highest mean values for number of tillers per plant, panicle length, kernal length, grain yield per plant and hulling percentage. Cluster VII with single genotype CMS46B represented high mean values for character days to 50% flowering and days to maturity. Cluster IV with (B35, JMS13B, JMS14B JMS17B, JMS18B, JMS19B and JMS20B) represented high mean values for characters number of grains per panicle and milling percentage. Cluster V with single genotype B42 represented high mean values for characters plant height, test weight and head rice recovery. Cluster VI with Genotype B88 recorded lowest cluster mean for days to 50% flowering, number of tillers per plant, panicle length, kernal length, L/B ratio, grain yield per plant, hulling percentage and milling percentage. Hence the genotypes from these clusters with suitable traits means can be selected to use in crossing programme to create desirable variations in the breeding material for yield improvement.

Similar findings reported Dalsaniya et al., 2009 reported that genotypes of cluster-X which had higher cluster mean values for yield and other desired characters like leaf area, ten pods weight, number of pods per plant and green pod yield per plant etc. CA Manoj et al., 2021 showed genotypes grouped under cluster I and II consisting of five and three genotypes respectively showed high mean values for days to first flowering, days to percent flowering and days to maturity. Nagamani et al., 2022 reported Clusters VI and VIII contains traditional red rice genotypes recorded lowest mean values for days to 50% flowering (91.00 and 92.00 days respectively) while genotypes of same clusters recorded highest mean value for plant height. Manasa et al., 2023 reported days to 50% flowering scoring was highest in cluster IV (101.00) and lowest in cluster V (75.00), days to maturity reading recorded highest in cluster I (129.03) and lowest in cluster V (110.00), plant height reading recorded highest in cluster IV(141.20) and lowest in cluster I(81.58)

3.5. Relative contribution of characters towards genetic divergence

The results showed (Table 5) that the contribution of kernel length was highest towards genetic divergence (30.90%), followed by days to flowering (18.46%), hulling percentage (16.54%), plant height (8.85%), grain yield per plant (4.49%), milling percentage (3.97%), head rice recovery and 1000-grain weight (3.85%), bran oil percentage (3.72%), number of grains panicle^{-1} (2.95%), days to maturity (1.15%), number of tillers plant^{-1} (0.64%), L/B ratio (0.51%), and kernel breadth (0.13%). The kernel length, days to 50% flowering, hulling percentage and plant height together contributed 74.75% towards total divergence. Therefore, these characters should be given importance during selection and hybridization program. Similar findings reported (Sreedhar et al., 2018, Nagamani et al., 2022).

Table 5: Relative contribution of different characters towards total genetic divergence

Sl. No.	Character	No of times ranked first	Contribution (%)
1.	Days to flowering	144	18.46
2.	Days to maturity	9	1.15
3.	Plant height (cm)	69	8.85
4.	No. of tillers plant^{-1}	5	0.64
5.	Panicle length (cm)	0	0
6.	No. of grains panicle^{-1}	23	2.95
7.	Kernel length (mm)	241	30.9
8.	Kernel breadth (mm)	1	0.13
9.	L/B ratio	4	0.51
10.	Grain yield plant^{-1} (g)	35	4.49
11.	1000-grain weight (g)	30	3.85
12.	Hulling (%)	129	16.54
13.	Milling (%)	31	3.97
14.	Head rice recovery (%)	30	3.85
15.	Bran oil (%)	29	3.72

4. CONCLUSION

The study utilized K-means clustering to investigate genetic diversity among genotypes. The clustering analysis resulted in the formation of distinct groups, with varying levels of intra-cluster genetic similarity and inter-cluster genetic diversity. Clusters with greater inter-cluster distances were identified, indicating significant genetic diversity between genotypes. Additionally, analysis of cluster means for various traits highlighted potential parent varieties with desirable characteristics. Kernel length, days to flowering, hulling percentage, and plant height were identified as major contributors to genetic divergence.

5. ACKNOWLEDGMENT

Professor Jayashankar Telangana State Agricultural University (PJTSAU) provided significant materials and funding for this study, which the authors gratefully acknowledge.

6. REFERENCES

Bhandari, D., Bhattarai, C., Dahal, B.R., Gautam, A., Joshi, P.P., 2021. Determinants of seed replacement rate (SRR) of rice in Kanchanpur district of Nepal. Nepalese Journal of Agricultural Sciences 20, 59–69.

Dalsaniya, S.B., Poshiya, V.K., Savaliya, J.J., Pansuriya, A.G., Davada, B.K., 2009. Genetic divergence in cowpea [*Vigna unguiculata* (L.) Walp.]. Legume Research-An International Journal 32(4), 250–254.

Devi, R.K., Venkanna, V., Hari, Y., Chandra, S.B., Lingaiah, N., Rajendra, P.K., 2020. Studies on genetic diversity and variability for yield and quality traits in promising germplasm lines in rice (*Oryza sativa* L.). The Pharma Innovation Journal 9(1), 391–399.

Falconer, D.S. (1981) Introduction to Quantitative Genetics. 2nd Edition, Longman Group Ltd., London, 1–133.

Forgy, E.W., 1965. Cluster analysis of multivariate data efficiency vs. interpretability of classifications. Biometrics 21, 768–769.

Ghazy, M.I., Salem, K.F.M., Sallam, A., 2021. Utilization of genetic diversity and markertrait to improve drought tolerance in rice (*Oryza sativa* L). Molecular Biology Reports 48(1), 157–170

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. Science 327(5967), 812–818.

Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B., 2014. "Circlize" implements and enhances circular visualization in. Bioinformatics 30(19), 2811–2812.

Julfiquar, R.A.W., Virmani, S.S., Carpena, A.L., 1985. Genetic divergence among some maintainer and restorer lines in relation to hybrid breeding in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 70, 671–678.

Kanavi, M.S.P., Prakash, K., Somu, G., Marappa, N., 2020. Genetic diversity study through k-means clustering in germplasm accessions of green gram [*Vigna radiata* (L.)] under drought condition. International Journal of Bio-resource and Stress Management 11(2), 138–147.

Kozak, M., Bocianowski, J., Rybinski, W., 2008. Selection of promising genotypes based on path and cluster analyses. Journal of Agricultural Science 146(1), 85–92.

MacQueen, J., 1967. Classification and analysis of multivariate observations. In proceedings of the 5th Berkeley Symposium on mathematical statistics and probability. Berkeley, University of California Press, 1, 281–297.

Manasa, S., Reddy, S.M., Murthy, K., Meena, A., Prasad, M.R., 2023. Genetic diversity studies in rice landraces (*Oryza sativa* L.) based on Mahalanobis D2 distance. International Journal of Environment and Climate Change 13(2), 28–34.

Manoj, C.A., Keerthi, N., Patil, A., Ramesh, A., Naveen, D.V., Kanavi, M.S.P., Rao, G.E., Venkataravana, P., Savithramma, D.L., 2022. Genetic divergence assessment through k-means clustering and principal component analysis for seed yield, zinc, iron and protein content in *Vigna unguiculata* L. Walp.

Meirmans, P.G., Van Tienderen, P.H., 2004. Genotype and genodive: Two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Resources 4(4), 792–794.

Nagamani, V., Shivakumar, N., Nagaraja, T.E., 2022. Genetic diversity analysis of red kernel rice genotypes using k-means clustering for grain yield, its components and cooking quality traits. International Journal of Bio-resource and Stress Management 13(8), 845–853.

Naik, M.V., Pillai, M.A., Saravanan, S., 2021. Genetic diversity analysis for yield associated and quality traits in promising rice varieties of Tamil Nadu. International Journal of Bio-resource and Stress Management 12(4), 361–369.

Oliveira, G.H.F., Amaral, C.B., Silva, F.A.M., Dutra, S.M.F., Marconato, M.B., Moro, G.V., 2016. Mixed models and multivariate analysis for selection of superior maize genotypes. Chilean Journal of Agricultural Research 76(4), 427–431.

Pandey, D.P., Kaushik, R.P., 2011. Genetic divergence for yield and quality traits in rice (*Oryza sativa* L.). Journal of Rice Research 4(1&2), 1–5.

Pillai, C., Faseela, K.V., Harikumaran, T., 2020. Nutritional composition of selected traditional rice varieties of Kerala. Journal of Tropical Agriculture 58(1), 33–43.

Ponnuswamy, R., Singh, A.K., Raman, M.S., Subbarao, L.V., Cn, N., 2020. Conversion of partial restorer Swarna into restorer by transferring fertility restorer Rf gene (s) through marker assisted backcross breeding (MABB) in rice. Scientific Reports 10(1), 1101.

Prasad, A.H., Senguttuvel, P., Revathi, P., Kemparaju, K.B., Sruthi, K., Sundaram, R.M., Seshu Madhav, M., Prasad, M.S., Laha, G.S., 2018. Breeding strategies for hybrid rice parental line improvement. Oryza 55, 38–41.

Rashmi, D., Bisen, P., Saha, S., Loitongbam, B., Singh, S., Singh, P.K., 2017. Genetic diversity analysis in rice (*Oryza sativa* L.) accessions using SSR markers. International Journal of Agriculture, Environment and Biotechnology 10(4), 457–467.

Roy, S.C., Sharma, B.D., 2014. Assessment of genetic diversity in rice (*Oryza sativa* L.) germplasm based on agro-morphology traits and zinc-iron content for crop improvement. Physiology and Molecular Biology of Plants 20(2), 209–240.

Shanmugam, A.S., Rangasamy, S.R.S., 1982. Genetic diversity for quantitative characters in green gram (*Vigna radiate* L.). Madras Agricultural Journal 69(10), 631–636.

Soujanya, T., Madamsetty, S.P., Hemalatha, V., KN, Y., Bandela, E., AS, H.P., Raman, M.S. Ponnuswamy, R., 2023. Improving hybrid rice parental lines for blast resistance by introgression of broad-spectrum resistance genes Pi54 and Pi9 by marker-assisted selection. Plant Breeding 142 (3), 300–311.

Sreedhar, S., 2018. Studies on variability, heritability, genetic advance and divergence for earliness, yield and yield components in rice (*Oryza sativa* L.) genotypes. International Journal of Bio-resource and Stress Management 9(1), 062–068.

Sruthi, K., Divya, B., Senguttuvel, P., Revathi, P., Kemparaju, K.B., Koteswararao, P., Sundaram, R.M., Singh, V.J., Ranjith Kumar, E., Bhowmick, P.K., Vinod, K.K., 2020. Evaluation of genetic diversity of parental lines for development of heterotic groups in hybrid rice (*Oryza sativa* L.). Journal of Plant Biochemistry and Biotechnology 29, 236–252.

Wickham, H., 2016. GGPLOT2: Elegant Graphics for Data Analysis. Springer-Verlag New York.https://doi.org/10.1007/978-3-319-24277-4.

Xiao, C.L., Virmani, S.S., Viraktamath, B.C., 1996a. Improving parental lines to increase efficiency of hybrid rice breeding: some new approaches-Chapter 10 in Advances in hybrid rice technology. In: Virmani, S.S., Siddiq, E.A., Muralidharan, K. (Eds.), Proceedings of the 3rd international symposium on Hybrid Rice 14–16 November, Hyderabad, India.

Xie, F.M., Guo, L.B., Ren, G.J., Hu, P.S., Wang, F., Xu, J.L., Li, X.Q., Qiu, F.L., dela Paz, M.A., 2012. Genetic diversity and structure of indica rice varieties from two heterotic pools of southern China and IRRI. Plant Genetic Resources: Characterisation and Utilisation 10, 186–193.