Exploiting genetic diversity in aromatic rice germplasm for parent selection using cluster analysis

ATUL KUMAR PACHAURI*

School of Agriculture, Sanjeev Agrawal Global Educational University
Bhopal 462022 Madhya Pradesh, India
*Email for correspondence: atul.pachauri@sageuniversity.edu.in

© Society for Advancement of Human and Nature (SADHNA) *Received: 05.07.2026/Accepted: 19.08.2026*

ABSTRACT

The present study was undertaken to assess genetic diversity among 42 rice germplasm accessions, including two checks, based on 18 yield and grain quality traits. The experiment was conducted during Kharif 2024-25 at Sanjeev Agrawal Global Educational University, Bhopal, Madhya Pradesh using a randomized block design with three replications. Mahalanobis' D^2 statistics and UPGMA cluster analysis grouped the germplasm into eight clusters, indicating considerable genetic diversity. The highest inter-cluster distance was observed between clusters V and VIII (92.416), while grain yield per plant was highest in cluster II (43.32 g) and lowest in cluster VIII (20.37 g). Based on genetic divergence and cluster performance, IC459148 from cluster V and IR64 from cluster VIII were identified as promising parents for hybridization. The identified diverse germplasm can be effectively utilized for developing high-yielding rice varieties with desirable grain and cooking quality.

Keywords: Rice germplasm; Genetic diversity; Mahalanobis' D^2; Cluster analysis; UPGMA; Genetic divergence; Heterosis; Grain quality; Hybridization

INTRODUCTION

Oryza sativa L serves as the staple food for over 3.5 billion people and sustains rural economies through household food security, employment and income (Anon 2013, Luis et al 2015). Driven by demographic shifts and changing diets, global rice production is projected to increase by 26 per cent by 2035 (Firdaus et al 2020, Mboyerwa et al 2021).

Meeting these growing demands requires high-yielding, resilient crop cultivars, which plant breeders develop by tapping into the natural variation within genetic resources. While genetic diversity represents the overall range of genetic characteristics across a crop or species, genetic variation reflects specific differences between individual plants for a given trait (Begna and Teressa 2024). Selecting genetically divergent parents is essential in hybridization programmes, as it increases the chances of producing superior hybrids and desirable transgressive segregants (Burman et al 2021).

To guide parental selection, breeders rely on biometrical tools like Mahalanobis' D^2 statistic to evaluate population structure by quantifying intra- and inter-cluster divergence (Singh and Narayanam 2015). This technique reveals trait contributions to overall diversity, grouping genetically similar genotypes together while separating divergent ones (Santosh et al 2019). Its utility in identifying promising parents for hybridization is well established in cereal crops (Bhatt 1973), where genetic distance serves as a reliable guide for breeding strategies (Arunachalam 1981).

Evaluating large germplasm collections requires effective analytical methods like Mahalanobis D^2 clustering and principal component analysis (Ali et al 2026). Assessing rice germplasm across yield and quality traits using these techniques helps breeders reliably identify optimal parental combinations for variety development.

Despite the recognized importance of genetic divergence in rice improvement, comprehensive

assessments integrating both yield attributes and grain quality parameters in aromatic accessions remain vital for regional breeding programmes. Therefore, the present study evaluated genetic diversity among 42 rice accessions across 18 yield and quality traits during kharif 2024-25 at Sanjeev Agrawal Global Educational University, Bhopal, Madhya Pradesh, using Mahalanobis' D^2 statistics and UPGMA cluster analysis. The objective was to characterize the germplasm, quantify intra- and inter-cluster distances and identify potential, genetically divergent parents to facilitate targeted hybridization programmes aimed at developing high-yielding rice cultivars with superior grain and cooking quality.

MATERIAL and METHODS

The present investigations were carried out at the Sanjeev Agrawal Global Educational University, Bhopal, Madhya Pradesh during Kharif 2024-25. Forty two diverse type of rice germplasm accessions of rice including two checks were planted in randomized block design with three replications. Data were collected on different yield contributing traits like days to 50 per cent flowering, plant height, net effective tillers, panicle length, grain length-breadth ratio,1000-grain weight, hulling percentage, milling percentage, head rice recovery, decorticated grain length, decorticated grain width, decorticated grain length-breadth ratio, kernel length after cooking, kernel breadth after cooking, kernel length breadth after cooking, elongation ratio and gel consistency (Jasmin and Rajendran 2026). The collected data were subjected to statistical analysis by D^2 statistics technique (Mahalanobis 1928).

RESULTS and DISCUSSION

Cluster analysis

The genetic diversity analyses achieved by UPGMA (unweighted pair-group method with arithmetic average) dissimilarity measure grouped 42 promising rice accessions based on 18 traits using Euclidean distance (Table 1).

The highest number of genotypes (9) was grouped in cluster VII, followed by clusters V and VI, each comprising 7 genotypes. Cluster VI also included one check, TN-1. Cluster IV comprised 4 genotypes, whereas, clusters II and III contained 3 genotypes each. Cluster VIII consisted of five varieties, namely Annada, IR64, Pusa Basmati-1, Swarna and PTB 33, indicating their close genetic relatedness and similarity with the check genotypes. The grouping of these varieties also suggests a degree of genetic relatedness among the check genotypes themselves.

Inter- and intra-cluster distances

The inter- and intra-cluster distances for all the traits represented by UPGMA values have been presented in Table 2 (Fig 1). The average intra-cluster distance ranged from 12.171 (cluster IV) to 26.129 (cluster VIII).

The highest inter-cluster distance (92.416) was observed between clusters VIII and V, indicating substantial genetic divergence between the genotypes belonging to these clusters. Therefore, genotypes from these clusters may be used as potential parents to exploit heterosis and generate superior recombinants. Based on the genetic divergence, IC459148 from cluster V and IR64 from cluster VIII may be selected as promising parents for hybridization to develop superior hybrids. Shanmugam et al (2023) similarly evaluated rice landraces using cluster analysis to group accessions based on yield traits and identified genetically divergent parents from widely separated clusters for use in hybridization programmes.

The lowest intra-cluster distance was observed in cluster VIII (26.129), indicating a high degree of genetic similarity among the genotypes within this cluster. Genotypes belonging to this cluster may, therefore, be useful in breeding programmes aimed at exploiting transgressive segregation. In the present study, Annada and IR64, representing cluster VIII, were selected as potential parents for hybridization, as their combination may generate desirable transgressive segregants with extreme phenotypic expressions. Kunta et al (2021) and Tiwari et al (2022) similarly emphasized that evaluating multivariate genetic divergence in rice allows for the identification of widely separated, genetically distant parental lines capable of generating superior hybrids and desirable transgressive segregants for yield and quality traits.

Cluster IV recorded the highest cluster mean for days to 50 per cent flowering (109.68 days), whereas, cluster VIII had the lowest mean value (93.80 days). For plant height, cluster V exhibited the highest cluster mean (168.40 cm), while cluster VIII recorded the lowest (88.19 cm) (Table 3). The highest mean number of net effective tillers was observed in cluster I (11.72), whereas, the lowest was recorded in cluster VIII (8.83). Similarly, cluster V exhibited the highest

Table 1. Grouping of rice germplasm into clusters

Cluster	Genotype with IC/EC number	Number of accessions
I	IC453695, IC46174 X, IC126596, IC206606	4
II	IC454040, IC114326, IC301736	3
III	IC454377, IC267449, IC300226	3
IV	IC454223X, IC26472, IC207311, IC124028	4
V	IC459148, IR79538, IC459797, IC124743, IC124989, IC206699, IC301732	7
VI	IC459199, IC460308, IC124536, IC135425IC214133, IC86313, TN-1	7
VII	IC114796, IC461616, IC124546, IC124735, IC124763, IC207184, IC125139, IC301734, IC386429	9
VIII	Annada, IR64, Pusa Basmati-1, Swarna, PTB 33	5

Table 2. Estimation of intra- (diagonal and bold) and inter-cluster distances among eight clusters

Cluster	I	II	III	IV	V	VI	VII	VIII
I	**21.513**	80.654	45.564	54.290	65.951	42.759	38.105	52.678
II		**25.394**	43.423	45.583	53.973	75.013	67.904	77.684
III			**17.597**	40.587	43.619	56.662	52.475	52.501
IV				**12.171**	34.262	33.518	32.025	75.637
V					**22.027**	46.352	58.518	92.416
VI						**20.638**	26.060	84.055
VII							**17.471**	68.928
VIII								**26.129**

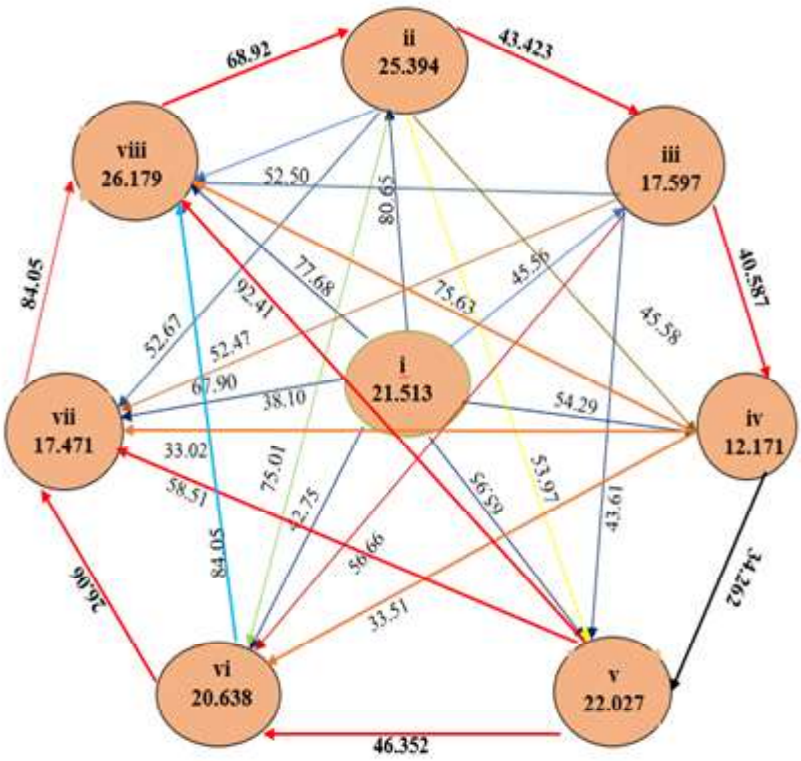

Fig 1. Diagrammatic representation of intra- and inter-cluster distance

Table 3. Cluster mean values of eight clusters for 18 characters

Character	Cluster mean value							
	I	II	III	IV	V	VI	VII	VIII
Days to 50% flowering	96.63	104.33	98.83	109.68	103.14	102.14	102.78	93.80
Plant height (cm)	115.96	151.39	128.17	158.79	168.40	157.62	143.47	88.19
Net effective tillers	11.72	10.28	10.11	9.29	11.48	9.37	9.71	8.83
Panicle length (cm)	23.59	25.21	24.60	24.17	25.85	24.53	22.57	23.43
Grain length-breadth ratio	4.31	4.59	3.64	3.69	3.71	3.75	3.57	4.16
1,000-grain weight (g)	2.66	2.42	2.40	2.66	2.76	2.19	2.63	2.05
Hulling percentage	74.26	55.69	73.11	74.44	77.12	74.52	73.58	70.76
Milling percentage	69.42	66.34	70.27	65.86	72.55	72.66	66.12	57.36
Head rice recovery (%)	59.41	18.90	46.70	36.74	61.53	55.68	35.56	29.14
Decorticated grain length (mm)	5.58	5.45	4.85	5.66	5.31	5.54	5.12	5.95
Decorticated grain width (mm)	2.20	1.82	2.07	2.10	2.26	2.18	1.98	1.55
Decorticated grain length-breadth ratio	2.65	3.06	2.50	2.61	2.38	2.63	2.58	3.77
Kernel length after cooking (mm)	9.65	8.47	8.68	8.68	9.29	9.09	8.94	8.79
Kernel bredth after cooking (mm)	3.54	3.03	3.40	3.10	3.49	3.14	3.11	2.84
Kernel length-breadth ratio after cooking	2.76	2.82	2.78	2.81	2.67	2.92	2.87	3.20
Elongation ratio	1.77	1.56	1.80	1.62	1.81	1.69	1.77	1.47
Gel consistency	44.00	100.17	85.83	64.19	82.79	38.50	37.40	68.80
Grain yield per plant (g)	37.53	43.32	39.21	38.15	42.40	37.72	42.68	20.37

mean panicle length (25.85 cm), while cluster VII recorded the lowest (22.57 cm). For grain length-to-breadth ratio, cluster II recorded the highest cluster mean (4.59), whereas cluster VII had the lowest (3.57). For 1,000-grain weight, the highest cluster mean was observed in cluster V (2.76 g), while the lowest was recorded in cluster VIII (2.05 g). Hulling percentage was highest in cluster V (77.12%) and lowest in cluster II (55.69%). These results indicate considerable variation among clusters for yield and quality traits, providing opportunities for the selection of desirable genotypes for breeding programmes.

Cluster VI recorded the highest cluster mean for milling percentage (72.66%), whereas, cluster VIII exhibited the lowest mean value (57.36%). For decorticated grain length, the highest cluster mean was observed in cluster VIII (5.95 mm), while the lowest was recorded in cluster III (4.85 mm). In contrast, decorticated grain breadth was highest in cluster V (2.26 mm) and lowest in cluster VIII (1.55 mm). The substantial variation among clusters for milling and grain quality traits indicates the potential of these genotypes for use in breeding programmes aimed at improving cooking and grain quality. Deepika et al (2026), Kesh et al (2022), Krishna et al (2022) and Kumar et al (2024) similarly observed wide genetic divergence and significant variations across rice germplasm clusters for milling recovery and physical grain dimension traits, demonstrating the effectiveness of multivariate cluster analysis in identifying specialized donor lines for grain quality improvement breeding programmes.

Cluster VIII recorded the highest cluster mean for decorticated grain length-to-breadth ratio (3.77), whereas, cluster V exhibited the lowest mean value (2.38). For kernel length after cooking, the highest cluster mean was observed in cluster I (9.65 mm), while the lowest was recorded in cluster II (8.47 mm). Similarly, kernel breadth after cooking was highest in cluster I (3.54 mm) and lowest in cluster VIII (2.84 mm), indicating considerable variation among clusters for these cooking quality traits.

For kernel length-to-breadth ratio after cooking, cluster VIII recorded the highest cluster mean (3.20), whereas cluster V had the lowest (2.67). Thang (2022) recorded that cluster VIII achieved the maximum cluster mean (3.20) for the cooked kernel length-to-breadth ratio, while Cluster V exhibited the minimum cluster mean (2.67) for this trait. The highest mean values for elongation ratio and gel consistency

were observed in clusters V (1.81) and II (100.17) respectively. In contrast, clusters VIII and VII recorded the lowest mean values for these traits, with values of 1.47 and 37.40 respectively. The observed variation among clusters for cooking quality traits highlights the potential of these genotypes for improving rice grain and cooking quality through targeted breeding programmes. Chandra et al (2026) and Prasad et al (2026) similarly observed wide cluster-wise variation for grain and cooking quality attributes in rice, emphasizing that such divergence across distinct clusters provides valuable donor lines for targeted quality improvement breeding programmes.

Grain yield per plant, one of the most important yield-related traits, recorded the highest cluster mean in cluster III (43.32 g), whereas, the lowest mean was observed in cluster VIII (20.37 g). This indicates that cluster III comprised relatively high-yielding genotypes. A considerable range of variation was observed among the 42 rice germplasms evaluated, which were grouped into eight distinct clusters based on Mahalanobis' D^2 analysis. The selection of parents for hybridization should be based on the magnitude of genetic divergence, the contribution of individual traits to total genetic divergence and the cluster mean performance for important traits, with the objective of exploiting heterosis.

Accordingly, crosses between genotypes belonging to clusters III and VIII and clusters VI and IV may exhibit considerable heterosis and generate favourable gene combinations in subsequent segregating generations. Amarnath and Ravi Kumar (2025) also observed that hybridization between genotypes from widely divergent clusters produces high heterosis and generates desirable transgressive segregants with valuable gene combinations in segregating generations.

CONCLUSION

The present study revealed substantial genetic diversity among 42 rice accessions evaluated for yield and grain quality traits, grouping them into eight distinct clusters based on variation in composition, cluster distances and mean performance. The highest inter-cluster distance was recorded between clusters V and VIII (92.416), highlighting high divergence and strong potential for exploiting heterosis using promising parents like IC459148 (cluster V) and IR64 (cluster VIII). Cluster II exhibited the highest mean grain yield per plant (43.32 g), whereas, cluster VIII had the lowest (20.37 g), alongside considerable variation for milling and cooking quality attributes. Consequently, strategic crosses between genetically divergent clusters, particularly III × VIII and VI × IV, offer significant opportunities for generating superior transgressive segregants and desirable gene combinations in future breeding programmes.

REFERENCES

Ali Z, Naeem M, Rafiq I, Nusrat AT, Usman U, Bahrun AH and Anshori MF 2026. Assessing genetic variability and trait associations in rice (*Oryza sativa* L): a study on yield components and grain quality. Brazilian Journal of Biology **86:** e305829; doi: 10.1590/1519-6984.305829.

Amarnath K and Ravi Kumar BNVSR 2025. Elucidation of rice advanced breeding lines (F_6 generation) for genetic diversity through principal component and hierarchical clustering analysis. Journal of Research ANGRAU **53(4):** 31-41; doi: 10.58537/jorangrau.2025.53.4.04.

Anonymous 2013. World rice statistics 2013. International Rice Research Institute, Los Banos, the Philippines.

Arunachalam V 1981. Genetic distance in plant breeding. Indian Journal of Genetics and Plant Breeding **41:** 226-236.

Begna T and Teressa T 2024. Genetic variability and its benefits in crop improvement: a review. Middle East Journal of Agriculture Research **13(1):** 128-136; doi: 10.36632/mejar/2024.13.1.6.

Bhatt GM 1973. Comparison of various methods of selecting parents for hybridization in common bread wheat (*Triticum aestivum* L). Australian Journal of Agricultural Research **24(4):** 457-464; doi: 10.1071/AR9730457.

Burman M, Nair SK and Sarawgi AK 2021. Principal component analysis for yield and its attributing traits in aromatic landraces of rice (*Oryza sativa* L). International Journal of Bio-resource and Stress Management **12(5):** 303-308; doi: 10.23910/1.2021.2348a.

Chandra BS, Hari Y, Neelima G, Kumar RS, Raju CD and Reddy RU 2026. Unravelling genetic divergence and trait dynamics in rice (*Oryza sativa* L) genotypes through multivariate analysis. Journal of Experimental Agriculture International **48(4):** 593-604; doi: 10.9734/jeai/2026/v48i44188.

Deepika K, Manimaran R, Pushpa R, Bama KS, Umamageswari C and Suresh R 2026. Evaluation of genetic variability, correlation and path analysis for yield and yield components in local rice landraces of Tamil Nadu, India. Plant Cell Biotechnology and Molecular Biology **27(5-6):** 151-158; doi: 10.56557/pcbmb/2026/v27i5-610602.

Firdaus RBR, Tan ML, Rahmat SR and Gunaratne MS 2020. Paddy, rice and food security in Malaysia: a review of climate change impacts. Cogent Social Sciences **6(1):** 1818373; doi: 10.1080/23311886.2020.1818373.

Jasmin A and Rajendran K 2026. Identification of rice genotypes for reproductive stage heat tolerance and yield stability through multi-trait and multivariate analysis. Frontiers in Plant Science **17:** 1853773; doi: 10.3389/fpls.2026.1853773.

Kesh H, Battan KR, Khan M and Yadav S 2022. Genetic diversity analysis of Basmati rice (*Oryza sativa*) genotypes for grain yield and quality traits. Indian Journal of Agricultural Sciences **92(7):** 862-865; doi: 10.56093/ijas.v92i7.110736.

Krishna K, Mohan YC, Krishna L, Parimala G and Jagadeeshwar R 2022. Multivariate analysis-based prediction of phenotypic diversity associated with yield and yield component traits in germplasm lines of rice (*Oryza sativa* L). Electronic Journal of Plant Breeding **13(3):** 764-771; doi: 10.37992/2022.1303.129.

Kumar S, Vimal SC, Meena RP, Prasad L, Luthra S, Srikanth B, Kumar A and Pal RK 2024. Multivariate analysis to study genetic diversity for yield and its attributing traits in rice (*Oryza sativa* L). International Journal of Environment and Climate Change **14(1):** 788-795; doi: 10.9734/ijecc/2024/v14i13897.

Kunta M, Shivani D, Jabeen F, Chaithanya K, Chiranjeevi M, Rao LVS, Sundaram RM and Rahaman AF 2021. Multivariate analysis and character association for agro-morphological traits in elite rice germplasm. Journal of Cereal Research **13(3):** 270-278; doi: 10.25174/2582-2675/2022/113997.

Luis JS, Paris TR, Rola-Rubzen MF and Pede VO 2015. Rural labor outmigration and gender dimension in an assessment of farm technical efficiency: a case study in selected rice villages in the Philippines. Asian Journal of Agriculture and Development **12(1):** 53-65.

Mahalanobis PC 1928. A statistical study at Chinese head measurement. Journal of Asiatic Society, Bangladesh **25:** 301-337.

Mboyerwa PA, Kibret K, Mtakwa PW and Aschalew A 2021. Evaluation of growth, yield and water productivity of paddy rice with water-saving irrigation and optimization of nitrogen fertilization. Agronomy **11(8):** 1629; doi: 10.3390/agronomy11081629.

Prasad GS, Sridhar V, Lingaiah N and Kylash KS 2026. Genetic variability, correlation and diversity analysis for yield and related traits among advanced breeding lines of rice (*Oryza sativa* L). International Journal of Advanced Biochemistry Research **10(4S):** 405-409; doi:10.33545/26174693.2026.v10.i4Se.8132.

Santosh, Jaiswal JP, Singh A and Gahatyari NC 2019. Genetic diversity analysis in bread wheat {*Triticum aestivum* (L) em Thell} for yield and physiological traits. International Journal of Current Microbiology and Applied Sciences **8(2):** 3059-3068; doi: 10.13140/RG.2.2.10738.89280.

Shanmugam A, Suresh R, Ramanathan A, Anandhi P and Sassikumar D 2023. Unravelling genetic diversity of south Indian rice landraces based on yield and its components. Electronic Journal of Plant Breeding **14(1):** 160-169; doi: 10.37992/2023.1401.007.

Singh P and Narayanam SS 2015. Biometrical techniques in plant breeding. Kalyani Publishers, New Delhi, India, pp 38-42.

Thang NB 2022. Genetic divergence of cultivated rice varieties in north Vietnam for grain quality traits using D^2 cluster analysis. Vietnam Journal of Agricultural Sciences **5(2):** 1435-1444; doi: 10.31817/vjas.2022.5.2.01.

Tiwari S, Singh Y, Upadhyay PK and Koutu GK 2022. Principal component analysis and genetic divergence studies for yield and quality-related attributes of rice restorer lines. Indian Journal of Genetics and Plant Breeding **82(1):** 94-98; doi: 10.31742/IJGPB.82.1.13.

How to cite this article: Pachauri AK 2026. Exploiting genetic diversity in aromatic rice germplasm for parent selection using cluster analysis. Int J Farm Sci 16(4): 28-33; doi: 10.5958/2250-0499.2026.00030.3.