Plant Archives Vol. 26, Supplement 1, 2026 pp. 967-978 e-ISSN:2581-6063 (online), ISSN:0972-5210

Plant Archives

Journal homepage: http://www.plantarchives.org

DOI Url : https://doi.org/10.51470/PLANTARCHIVES.2026.v26.supplement-1.129

EXPLORING GENETIC DIVERSITY ANALYSIS OF ADVANCED RICE LINES FOR SELECTION AND GENERATION OF HIGH-YIELDING GENOTYPES USING ENHANCED BREEDING STRATEGIES

K. Madhu Kumar, G. Sreenivas, P. Udayababu, M. Srinivasa Rao and D. Manojkumar*
ANGRAU- Agricultural Research Station, Ragolu, Acharya N G. Ranga Agricultural University, Andhra Pradesh, India
*Corresponding author E-mail: manojgene7@gmail.com

(Date of Receiving : 27-09-2025; Date of Acceptance : 09-12-2025)

ABSTRACT

Genetic diversity among advanced rice (*Oryza sativa* L.) breeding lines was assessed using multivariate statistical methods to evaluate the set of 68 advanced rice breeding lines representing diverse genetic backgrounds was evaluated during Kharif 2024 at the Agricultural Research Station, Ragolu, to assess the extent of phenotypic and genetic divergence. The experiment was conducted in an Alpha Lattice Design with three replications, and eight key traits days to 50% flowering, plant height, ear-bearing tillers, panicle length, filled and ill-filled grains per panicle, test weight and grain yield per plant were recorded from randomly selected plants at maturity. Multivariate statistical approaches, including Mahalanobis D^2 analysis, Tocher's clustering, factor analysis and principal component analysis (PCA), were applied to quantify variability and identify the traits contributing most to divergence. Significant multivariate differences among the genotypes confirmed the presence of substantial diversity. Days to 50% flowering contributed the highest proportion to total divergence, followed by test weight, emphasizing the importance of maturity duration and grain weight in differentiating the lines. Tocher's method grouped the genotypes into ten clusters that varied widely in size, including several singleton clusters representing highly distinct entries. The largest inter-cluster distances were observed between clusters with contrasting phenological and yield-related attributes, indicating their potential for generating wide variability in hybridization programs. Cluster mean analysis revealed groups with specific advantages such as high grain yield, greater spikelet fertility, bold grain type or early maturity. Factor analysis and PCA further supported the clustering pattern, with the first few components explaining the majority of variation and providing clear separation among genotypes. Overall, the study demonstrated rich and structured genetic diversity, offering valuable opportunities for selecting suitable parental combinations to enhance yield, quality and maturity traits in rice improvement programs.

Keywords**:** Crop improvement, Cluster analysis, Diversity analysis, Principal component analysis, Rice and Selection.

Introduction

Rice (*Oryza sativa* L.) is one of the most important food crops globally, serving as the primary staple food for more than half of the world's population, production sustainability challenged by multiple biotic and abiotic factors (Udayababu *et al.*, 2025). With rising population growth, climate change and the increasing demand for high-quality and high-yielding cultivars, rice breeding programs have become more reliant on the effective utilization of genetic diversity (Satyanarayana *et al.*, 2023). The availability of genetic variation not only enhances selection gains but also safeguards long-term breeding progress by mitigating vulnerability to biotic and abiotic stresses (Duppala *et al.*, 2023). The strategic assessment of diversity among advanced breeding lines is therefore a critical step toward generating novel recombinants capable of meeting modern production challenges (Kumar *et al.*, 2023).

Genetic diversity in rice is shaped by its extensive evolutionary history, domestication events and subsequent breeding interventions. Over time, intensive selection for specific agronomic traits has narrowed the genetic base of many rice cultivars, raising concerns over reduced resilience and adaptability (Fornasiero *et al.,* 2022). Several studies demonstrated a decline in allelic richness among Indian rice varieties over breeding decades, underscoring the need for continuous introduction of diverse parental sources (Singh *et al.,* 2016). Likewise, Xu *et al.* (2016) reported distinct genetic structuring within Chinese *indica* rice, suggesting that exploiting divergence among sub-populations can lead to greater heterotic potential. These findings emphasize that assessing diversity at both phenotypic and molecular levels is crucial to developing productive and climate-resilient rice cultivars.

Phenotypic evaluation remains central to plant breeding due to its direct relevance to agronomic performance. Multivariate analysis allows researchers to examine complex trait relationships, reduce dimensionality and identify underlying patterns of variation that are difficult to detect through univariate methods (Madhukumar *et al.,* 2023). Earlier rice studies have successfully employed multivariate statistics to reveal diversity patterns, identify superior genotypes and assist in parent selection. Chemutai *et al.* (2016), using principal coordinate analysis, demonstrated substantial trait-based variation among rice accessions, correlating physicochemical traits with agronomic divergence. Similarly, Choudhary *et al.* (2022) used PCA and cluster analysis to dissect morphological diversity in rice breeding lines and identified trait combinations contributing most to divergence (Sreenivas *et al.,* 2023). These studies confirm the value of multivariate approaches in interpreting high-dimensional phenotypic data in rice research.

Cluster analysis, in particular, facilitates grouping of genotypes into distinct clusters based on overall similarity or dissimilarity among traits (Kumar *et al.,* 2017). Genotypes positioned in widely separated clusters are typically considered suitable for hybridization, as the resulting crosses are expected to generate broad genetic variability and potentially high heterosis (Manoj *et al.,* 2022). The Mahalanobis D^2 statistic remains one of the most widely used methods for quantifying multivariate divergence due to its robustness and ability to incorporate covariance information among traits. Recent studies by Sar and Kole (2023) and Allam *et al.* (2017) have demonstrated the effectiveness of D^2-based clustering in identifying divergent parents for rice yield improvement and stress tolerance. These findings reaffirm that exploiting inter-cluster divergence is a practical strategy to broaden genetic bases in breeding programs.

Principal component analysis (PCA) and factor analysis complement cluster analysis by providing deeper insights into trait interrelationships and identifying the major components governing variation. PCA reduces complex trait data into a set of orthogonal components, enabling the visualization of genotype distribution across major axes of variation (Manojkumar *et al.,* 2018). PCA-based biplots allow breeders to distinguish genotypes with desirable combinations of traits such as early maturity and high yield. Studies by Jeong *et al.* (2017) and Khatun *et al.* (2023) have shown that the first few PCs in rice often reflect phenology, plant stature, panicle architecture and grain-yield traits, demonstrating the biological relevance of PCA “trait dimensions.” Factor analysis, on the other hand, uncovers latent structures underlying observed variables, helping to identify trait groups that are biologically or physiologically related. Such information is valuable for trait-based selection and ideotype development. Studies combining both approaches of pehontypic selection with multivariate analysis helps in reinforcing the utility of these tools in guiding parent selection. The integration of multivariate statistics into classical breeding pipelines ultimately supports more strategic, informed and efficient hybridization programs.

Given these advancements, assessing diversity among advanced breeding lines is particularly important because these lines represent selections with potential for varietal release (Swarup *et al.*, 2021). Understanding how they differ in key agronomic traits, flowering time, plant height, tillering ability, panicle structure, grain filling and grain yield, enables breeders to identify complementary parents for crossing. Traits such as days to 50% flowering and plant height often reflect maturity and plant architecture, whereas test weight and grain yield per plant relate directly to productivity and grain quality. Exploring the structure of diversity across these traits helps ensure that selection targets are aligned with current breeding objectives such as higher yield, climate adaptation and improved grain quality (Sinha *et al.*, 2023). The present study employed multivariate statistical techniques to comprehensively evaluate genetic diversity among advanced rice breeding lines. With the objectives of quantify the magnitude of genetic divergence using Mahalanobis D^2 and cluster analysis; identify the major factors governing trait variation using factor analysis; determine the principal components underlying agronomic differences among genotypes and integrate the findings to propose suitable parental combinations for future rice improvement. Such a multivariate approach provides a robust framework for understanding trait-based diversity and supports strategic decision-making in rice breeding programs aimed at enhancing yield potential, stability and adaptability in changing agroecosystems.

Materials and Methods

A set of 68 advanced rice breeding lines (genotypes) were evaluated at Agricultural Research Station, Ragolu, of Acharya N G Ranga Agricultural University, situated at 83.240° E longitude and 18.240° N latitude, with an altitude of 27 meters above mean sea level. The site, representative of the North Coastal Zone of Andhra Pradesh, typically receives an annual rainfall of about 1111 mm (Udayababu *et al.,* 2024), during the Kharif season of 2024. The entries represent diverse genetic backgrounds, drawn from different breeding programs to maximize variability. The experiment was laid out in a Alpha Lattice Design (ALD) with three replications. Each plot comprised plants spaced at [spacing, 20 cm × 15 cm] and cultural practices followed standard agronomic recommendations for rice. At maturity, ten plants per plot were randomly selected to record eight agronomic traits: days to 50% flowering (DFF), plant height (PH, in cm), ear-bearing tillers (EBT), panicle length (PL, in cm), filled grains per panicle (FGP), ill-filled grains per panicle (IFGP), test weight (TW, e.g., thousand-grain weight in g) and grain yield per plant (GYP).

Multivariate divergence among the genotypes was assessed using Mahalanobis's generalized distance (D^2), computed from standardized trait data and the resulting genetic distances were used to group the genotypes into distinct, non-overlapping clusters through Tocher's clustering method. The relative contribution of each trait to total genetic divergence was estimated following the procedure of Singh and Chaudhary (1979), enabling identification of the most influential discriminating traits. To further understand the structure of variability, factor analysis with varimax rotation was performed on the correlation matrix, retaining factors with eigenvalues greater than one and interpreting trait loadings to elucidate the latent dimensions governing trait variation. Additionally, principal component analysis (PCA) was applied to the standardized dataset, with the major principal components used to describe the proportion of total variability explained and to generate biplots illustrating genotype distribution across principal component space. All statistical analysis were executed using [Indostat 9.3.0].

Results and Discussion

The assessment of genetic diversity among the advanced rice breeding lines using a combination of multivariate tools, namely cluster analysis, factor analysis and principal component analysis (PCA), provided a comprehensive picture of the phenotypic variability present in the experimental material. The study utilized eight quantitative traits recognized as highly influential in rice productivity and adaptation: days to fifty per cent flowering (DFF), plant height (PH), ear-bearing tillers per plant (EBT), panicle length (PL), filled grains per panicle (FGP), ill-filled grains per panicle (IFGP), test weight (TW) and grain yield per plant (GYP). These traits represent major morphological, reproductive and grain-quality parameters widely used in rice improvement programs. The highly significant Wilks' Lambda test confirmed that the genotype set exhibited strong multivariate divergence, establishing a solid reason for further analysis (Table 1). The diversity patterns observed across these traits indicate the presence of substantial genetic variability, which can be effectively exploited for heterosis breeding and trait recombination. Recent studies in rice germplasm (Jha *et al*., 2024; Nivedha *et al*., 2024) have similarly emphasized the importance of integrated multivariate approaches to characterize complex relationships among traits, particularly when dealing with large sets of advanced breeding lines. The initial evaluation of trait contributions showed that DFF was the most influential trait, contributing nearly 69 per cent to the total divergence. This suggests that flowering time is a key driver of differentiation among the genotypes, which aligns with the physiological importance of DFF in determining the duration of vegetative and reproductive phases. Flowering time is highly sensitive to photoperiod, temperature and genetic background; thus, even minor genetic differences lead to large phenotypic shifts. TW also contributed substantially (≈17 per cent), indicating that grain weight plays an important role in distinguishing genotypes. Grain weight is known to be a complex trait influenced by grain size, grain density, endosperm development and assimilate allocation. Several studies conducted after 2016 (Fujino *et al*., 2020 and Kim *et al.,* 2024) have described the strong discriminatory capacity of DFF and TW in multivariate analysis of rice, thereby supporting the outcomes of the present study.

Cluster analysis grouped the genotypes into ten distinct clusters, indicating a wide spread of variability (Table 3 and 4). The formation of multiple clusters suggests that the evaluated lines represent diverse genetic backgrounds, possibly derived from multiple breeding pipelines or hybridization strategies. The distribution pattern of genotypes across clusters revealed that some clusters contained large numbers of genotypes, suggesting greater internal similarity, whereas others contained only one or two lines, reflecting their uniqueness. Inter-cluster distance analysis showed that the largest divergence occurred between Cluster 8 and Cluster 2, with a Mahalanobis distance exceeding 1300. This substantial magnitude implies that these clusters harbor highly differentiated genotypes suitable for maximizing heterotic gains when used as parents in hybridization. Other cluster pairs, such as Cluster 8 with Cluster 1 and Cluster 5 with Cluster 9, also exhibited high divergence, reflecting the broad genetic base within the study material. Conversely, the smallest distances were observed between clusters such as Cluster 1 and Cluster 7, suggesting close similarity and minimal chances of producing wide variability through crossing. Several researchers, including Allam *et al.* (2017), Sar and Kole (2023) and Chandraker *et al.* (2024), have reported the utility of Mahalanobis distance in identifying promising and genetically distant parents for rice improvement. The patterns observed in the present analysis strongly align with these findings.

The Tocher clustering method grouped the evaluated genotypes into ten distinct clusters, reflecting considerable genetic divergence within the material (Figure 1 and 2). Cluster sizes varied markedly, ranging from large, heterogeneous groups (Cluster I and Cluster II) to several single-genotype clusters (Clusters V, VI, VII, IX and X). Such variation in cluster composition indicates the presence of substantial diversity that can be exploited in a breeding program. Cluster I, containing the highest number of genotypes, included several MTU, BPT and RGL lines that shared similarity for multiple agronomic traits. The large size of this cluster suggests that these genotypes possess a comparatively narrow genetic base or similar phenotypic profiles. While these lines may be valuable for trait stabilization and selection within a homogeneous pool, they may contribute limited new variation when intercrossed among themselves. Cluster II also accommodated a large group of genotypes but differed in composition from Cluster I by including a mix of MTU, RGL, BPT and NLR lines. The moderate size of this cluster, combined with its diverse pedigree representation, indicates a broader genetic background. Genotypes from this cluster may serve as useful intermediate donors when aiming for recombinants with balanced performance.

Clusters III and IV each consisted of a sizeable collection of genotypes that were clearly separated from those in the first two clusters. The presence of multiple breeding lines such as MTU 2631-88-1-2-2, BPT 3500 and NLR 3895 in Cluster III and MTU 2744-96-1-1-1 and RNR 15048 in Cluster IV, points to unique trait combinations that distinguish these clusters from the rest. These groups likely represent sources of specific agronomic strengths such as stress tolerance, grain quality, or yield-related traits. Clusters V to X each contained only one genotype, indicating their pronounced genetic distinctiveness. Genotypes such as MTU 2675-2-2-2-1, BPT 3463, RGL 7038, RGL 7050 and KNM 1638 likely possess rare allele combinations or extreme trait values that set them apart from the other entries. Such unique genotypes are especially valuable in hybridization programs, as crossing them with members of the larger clusters can maximize heterosis and broaden the genetic base of future breeding populations.

The clear divergence among clusters provides a practical framework for parent selection. In general, crossing genetically distant clusters tends to produce greater heterosis and higher chances of identifying superior segregants. Based on the observed cluster pattern, Cluster I × Cluster V/VI/VII/IX/X combinations are expected to yield considerable variability due to the high divergence between the very large cluster and genetically isolated singletons. Cluster II or III × singleton clusters may also be suitable when targeting specific trait improvements such as grain quality or stress tolerance. Crosses within Cluster I or Cluster II are less likely to generate substantial variability but may be useful for refining traits under stabilizing selection.

The examination of cluster means offered additional clarity on the specific trait advantages embedded in each group (Table 4). Cluster 7 emerged as the most promising cluster for yield improvement owing to its superior GYP mean (≈60 g). This cluster also displayed advantageous values for EBT and PH, indicating that its genotypes combine high productive tillering with tall stature, which can enhance biomass production and assimilate availability during grain filling. Clusters 3 and 4 showed elevated FGP values, making them attractive sources for enhancing spikelet fertility. Meanwhile, Cluster 9 recorded the highest TW, supporting its utility as a donor for improving grain boldness and milling quality. Other clusters revealed strengths in PL, DFF or reduced IFGP, each offering unique breeding value depending on the targeted trait combination. Clusters 8 and 10 exhibited low DFF values, representing early-flowering genotypes. Such genotypes are particularly useful in regions where early maturity allows escape from terminal drought or late-season heat, or where multiple cropping systems require short-duration varieties. The distinctiveness of early and high-yield clusters implies opportunities to combine earliness with productivity, a breeding goal highlighted in several post-2016 studies (Thakur & Sarma, 2023; Kumar *et al*., 2021). The diversity revealed through clustering highlights the potential of the evaluated genotypes for breeding programs aimed at yield enhancement and trait improvement. The presence of several distinct clusters, including multiple singleton genotypes, ensures ample scope for developing high-performing segregating populations. Rational parent selection guided by cluster distance can therefore accelerate progress in varietal development and broaden the genetic base of future breeding lines.

Factor analysis elucidated the underlying structure of trait relationships by identifying three principal factors with eigenvalues greater than one (Table 5). The first factor encapsulated a phenology–biomass–fertility gradient, with positive loadings for DFF, PH and FGP and a negative loading for EBT. This suggests that taller, late-flowering genotypes exhibit stronger grain filling but tend to have fewer productive tillers. This trait combination is particularly relevant in varieties where high biomass contributes to assimilate partitioning during the reproductive stage, enhancing grain filling. However, the negative association with EBT indicates a trade-off typical in many inbred rice varieties, where genotypes invest either in tiller production or vegetative growth and panicle fertility. Factor 2 was dominated by high positive loadings for GYP and TW and a strong negative loading for IFGP, revealing that high yield is generally associated with larger grain size and reduced spikelet sterility. The complementarity of these traits is well recognized in rice breeding literature, where increasing grain weight and reducing sterility are major targets for yield enhancement. Factor 3 was defined by negative loadings for PL and IFGP, suggesting that shorter panicles tend to have fewer unproductive spikelets. Although this component explained a smaller proportion of variation, it represents a morpho-reproductive axis important for panicle efficiency. The nearly complete communalities for all traits affirm the robustness of the factor model, a result consistent with multivariate trait analysis reported in recent diversity studies (Nguyen *et al*., 2019; Khatun *et al*., 2023).

Principal Component Analysis, which is widely employed to identify major sources of orthogonal variation, revealed that four PCs accounted for over 87 per cent of the total variability (Table 6 and 7). This indicates that the eight traits used in the study captured most of the underlying differences among the genotypes. The first principal component (PC1) explained 43.72 per cent of the variation and was heavily influenced by DFF, PH and FGP, with a negative contribution from EBT. This axis represents the overall growth and reproductive maturity profile of genotypes, separating late, tall, high-fertility lines from early, short, tiller-rich types. The second component (PC2), which explained 16.55 per cent of the variance, was shaped by high contributions from GYP, TW and PL and a negative contribution from IFGP. PC2 can therefore be interpreted as a yield and grain-quality axis, where long panicles, bold grains and high yields cluster together. The third component (PC3) explained an additional 16.23 per cent of the variation and distinguished genotypes based on panicle morphology and sterility-related traits. PC4 (10.61 per cent) captured a trade-off between grain weight (positive loading) and grain yield (negative loading), reflecting resource allocation differences among genotypes. This compensation mechanism has been widely discussed in modern rice phenomics, especially in studies focusing on grain-size QTLs and source–sink balance (Choudhary *et al*., 2022; Jeong *et al*., 2017).

The scatter plots derived from PCA scores demonstrated the clear spatial separation of genotypes into well-defined groups (Figure 3 and 4). High PC1 scorers were typically found among late maturing, tall and highly fertile genotypes. High PC2 scorers represented superior yielders with strong grain weight and long panicles, making them highly valuable for breeding programs focused on improving productivity. Genotypes with strong PC3 contributions exhibited variation in reproductive efficiency and spikelet sterility, indicating potential for improving panicle structure and sink capacity. PC4-dominated genotypes possessed distinct grain-size characteristics irrespective of yield, suggesting their suitability for grain-quality breeding. These PCA-based genotype assignments closely aligned with the clusters identified in the Tocher method, showing strong consistency across independent multivariate techniques.

The integration of the three multivariate analytical approaches offers a robust framework for identifying promising parental combinations. Crosses among highly divergent clusters,such as Cluster 8 × Cluster 2 or Cluster 9 × Cluster 7, are likely to generate broad variability essential for selecting transgressive segregants. Similarly, pairing high PC2 genotypes (yield and quality) with high PC3 genotypes (panicle architecture) could help combine complementary traits, addressing both productivity and reproductive efficiency. Early-maturing clusters may be crossed with high-yielding groups to combine earliness with superior grain productivity. Such strategic parent selection aligns with contemporary breeding recommendations emphasizing the importance of combining divergence and complementarity (Kim *et al.*, 2024; Chandraker *et al.*, 2024; Sar & Kole, 2023).

In summary, the analysis demonstrate that the advanced rice lines possess extensive and structured variability suitable for improving multiple traits simultaneously. The combined use of cluster analysis, factor analysis and PCA enabled a clear understanding of the diversity structure, trait complexes and potential parental groups. These findings not only confirm the richness of the evaluated germplasm but also provide a practical roadmap for breeding programs targeting higher grain yield, improved grain quality, stable fertility and desirable maturity profiles. The study further emphasizes the importance of integrating phenotypic multivariate analysis with genomic information in future research to enhance precision and accelerate variety development.

Conclusion

This study revealed clear, structured genetic diversity among advanced rice breeding lines through the combined use of cluster analysis, factor analysis and PCA. The identification of highly divergent clusters (Cluster 8 vs. Cluster 2) underscores the potential of selected parent combinations for heterosis breeding. The factor analysis clarified underlying trait dimensions, such as the trade-off between maturity and tillering or the association of yield with grain quality, that can guide ideotype design. PCA complemented these insights by highlighting genotypes with superior profiles for early maturity, high yield, or grain weight. By integrating multivariate tools, we offer a rational and data-driven strategy for parent selection and cross planning. The selected genotypes from divergent clusters can be used in future hybridization schemes to simultaneously improve yield, adaptability and grain quality. For greater precision, subsequent work should incorporate genomic data and possibly genomic selection to validate and refine these parent choices. Overall, our findings provide a strong foundation for targeted breeding of rice cultivars that align with modern goals of productivity and resilience.

Table 1 : Contribution of different characters towards genetic divergence in advanced rice lines

S.No.	Trait	Times Ranked 1st	Contribution (%)
1	Days to 50 per cent flowering	1572	69.01
2	Plant height	38	1.67
3	Ear bearing tillers	6	0.26
4	Panicle length	167	7.33
5	Filled grains per panicle	48	2.11
6	Ill filled grains per panicle	57	2.5
7	Test weight	385	16.9
8	Grain yield	5	0.22

Table 2 : Clustering pattern of diversity in the advanced rice lines by Tocher's method

Cluster	Genotypes	Number of Genotypes
1	BPT 5204, BPT 3354, MTU 7029, MTU 2869-17-2-1-1, RGL 7044, MTU 1239, RGL 7051, MTU 2878-13-1-1, MTU2776-29-1-1-1-2, RGL 7047, RGL7045, MTU 2851-19-1-1-1, MTU 2837-56-1-1-1, RGL 7036, BPT 3456, MTU 2861-1-1-1-1-1, MTU2861-1-1-2, MTU 2854-9-1-1-1-1, BPT3278, RGL 7046, 14 MTU 1121, 1MTU 2969-19-2-1	22
2	BPT 3451, MTU 2626-3-3-1-1-1, RGL 7049, RGL 7048, RGL 7034, MTU 1310, BPT 3468, RGL 7039, MTU 1318, MTU 2823-23-1-1, MTU 1262, BPT 3485, NLR 3881, MTU 2855-6-2-1-1, RGL 2537	15
3	MTU 2631-88-1-2-2, MTU 2716, BPT 3500, BPT 3250, MTU 2761-29-1-1-1-1, MTU 2868-143-1-2-1, MTU 2944-83-1-1-1, MTU 2949-33-1-1, MTU 2701-1-1-1-1, RGL 2538, NLR 3893, NLR 3895, JMP 150	13
4	MTU 2744-96-1-1-1, NLR 3648, MTU 1224, RNR 15048, NLR 3869, NDLR 31, MTU 2879-19-2-1, MTU 2774-96-2-1-1	8
5	MTU 2675-2-2-2-1	1
6	BPT 3463	1
7	RGL 7038	1
8	PR 126, MTU 1153, NLR 3897, MTU 2745-101-1-1-1, RGL 1880	5
9	RGL 7050	1
10	KNM 1638	1

Table 3 : Average intra and inter-cluster distances (D^2 values) among ten clusters (obtained by Tocher's method).

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Cluster 6**	**Cluster 7**	**Cluster 8**	**Cluster 9**	**Cluster 10**
Cluster 1	**58.83**	214.98	153.59	286.90	143.00	436.78	90.52	632.05	182.21	578.70
Cluster 2	214.98	**66.05**	539.09	765.48	104.21	1069.26	160.53	1357.41	101.71	1290.60
Cluster 3	153.59	539.09	**75.30**	172.31	355.83	198.35	268.34	280.66	452.54	309.63
Cluster 4	286.90	765.48	172.31	**92.14**	674.53	207.18	344.67	336.84	609.03	152.75
Cluster 5	143.00	104.21	355.83	674.53	**0.00**	823.44	191.89	1004.27	168.25	1108.41
Cluster 6	436.78	1069.26	198.35	207.18	823.44	**0.00**	574.94	199.43	886.72	105.16
Cluster 7	90.52	160.53	268.34	344.67	191.89	574.94	0.00	**879.18**	108.08	698.83
Cluster 8	632.05	1357.41	280.66	336.84	1004.27	199.43	879.18	**101.18**	1218.96	239.34
Cluster 9	182.21	101.71	452.54	609.03	168.25	886.72	108.08	1218.96	**0.00**	1072.27
Cluster 10	578.70	1290.60	309.63	152.75	1108.41	105.16	698.83	239.34	1072.27	**0.00**

Table 4 : Mean values of ten clusters estimated by Tocher's method in the advanced rice lines.

	DFF	**PH**	**EBT**	**PL**	**FGP**	**IFGP**	**TW**	**GYLD**
Cluster 1	104.27	131.45	268.45	28.40	270.83	44.58	17.27	6.53
Cluster 2	115.57	143.23	278.03	28.98	308.12	54.73	16.26	7.02
Cluster 3	96.38	127.98	280.85	27.50	213.38	42.49	20.01	6.06
Cluster 4	91.69	112.11	294.13	24.48	201.27	37.04	14.32	5.97
Cluster 5	112.50	152.10	264.50	30.70	260.67	27.84	21.09	7.44
Cluster 6	87.50	124.70	245.50	32.05	247.00	77.17	17.46	4.43
Cluster 7	105.50	154.40	306.00	26.95	333.00	60.00	13.76	6.05
Cluster 8	84.10	119.02	271.60	25.16	161.67	17.47	24.21	6.06
Cluster 9	113.00	140.10	259.50	27.19	195.00	139.83	15.20	5.76
Cluster 10	83.00	110.90	222.50	25.77	182.17	33.17	13.69	6.37

Table 5 : Divergence explaining by factor analysis in the advanced rice lines

Factor Analysis			
Variable	**1 Factor**	**2 Factor**	**3 Factor**
Root	1.00	1.00	1.00
σ^2 % Explained	12.50	12.50	12.50
Σ σ^2 Cum Explained	33.33	66.67	100.00
Days to 50 per cent flowering	0.49	0.02	0.16
Plant height	0.50	-0.03	-0.30
Ear bearing tillers	-0.50	-0.01	0.09
Panicle length	-0.02	0.07	-0.75
Filled grains per panicle	0.49	0.00	0.01
Ill filled grains per panicle	-0.03	-0.64	-0.42
Test weight	-0.11	-0.47	-0.35
Grain yield	0.06	0.61	-0.10

Table 6 : Principal Component Analysis in the advanced rice lines

	1 Vector	**2 Vector**	**3 Vector**
Eigen Value (Root)	3.50	1.32	1.30
% Var. Exp.	43.72	16.55	16.23
Cum Var. Exp.	43.72	60.27	76.50
DFF	0.51	0.03	0.08
PM	0.43	0.20	-0.35
EST	-0.46	-0.14	0.15
PL	-0.18	0.42	-0.60
FGP	0.50	0.03	0.03
IFGP	-0.07	-0.36	-0.67
TW	-0.22	0.55	-0.05
GYP	-0.01	0.58	0.21

Table 7 : Factor loadings on principal component axis by the advanced rice lines

Genotype	PCAI (X Vector)	PCAII (Y Vector)	PCA III (Z Vector)
MTU 2774-96-2-1-1	61.41	23.39	-2.01
PR 126	52.33	24.10	0.51
MTU 2701-1-1-1-1	60.36	27.49	-1.87
MTU 2744-96-1-1-1	62.87	17.75	2.66
NLR 3648	62.55	17.00	4.48
MTU 2868-143-1-2-1	66.89	23.08	-1.35
NDLR 31	57.92	14.47	3.81
MTU 2949-33-1-1	63.12	24.58	-0.23
RGL 1580	53.96	33.95	-4.20
MTU 2879-19-2-1	60.95	22.76	-3.86
MTU 2761-29-1-1-1-1	65.49	24.59	-1.48
MTU 2745-101-1-1-1	56.54	30.47	-3.62
MTU 1153	50.88	26.86	-2.51
MTU1121	66.79	26.03	0.07
JMP 150	59.13	24.52	-5.23
MTU 2969-19-2-1	70.16	21.49	0.77
KNM 1638	56.24	21.62	-2.16
BPT 3463	57.83	24.27	-8.63
RNR 15048	64.74	16.97	-2.49
NLR 3869	59.58	21.72	3.97
MTU 2944-83-1-1-1	63.29	24.22	0.18
NLR 3897	54.18	24.37	0.50
RGL 7047	72.86	21.37	0.25
BPT 3468	80.02	19.61	2.57
MTU 2869-17-2-1-1	74.32	20.72	3.79
RGL 7034	82.79	20.42	-0.27
MTU 1239	74.21	25.93	1.00
BPT 3456	73.79	24.38	-3.52
MTU2861-1-1-2	71.13	26.72	-1.57
MTU2776-29-1-1-1-2	75.98	24.90	0.58
RGL7045	70.79	23.56	-1.71
RGL 7048	81.53	19.69	2.84
RGL 2538	61.61	24.67	2.46
MTU 2854-9-1-1-1-1	71.17	25.67	-2.91
MTU 2855-6-2-1-1	79.53	25.47	-1.50
MTU 2851-1-1-1-1-1	72.87	27.68	-1.51
MTU 2837-56-1-1-1	70.71	25.86	-1.07
BPT 3451	84.56	21.08	0.81
MTU 2851-19-1-1-1	74.85	23.22	-0.98
RGL 7036	73.17	24.66	-2.91
NLR 3895	64.36	22.46	1.82
RGL 7038	77.01	19.47	-1.16
MTU 1224	66.33	19.46	4.43
RGL 7046	69.50	21.10	-3.99
MTU 2626-3-3-1-1-1	84.06	22.47	0.83
BPT 3500	68.01	23.93	-5.90
MTU 2631-88-1-2-2	67.19	28.76	-3.54
BPT 3250	67.25	24.57	-4.68
NLR 3893	65.92	22.39	2.08
MTU 1262	86.02	22.78	3.14
MTU 7029	70.65	23.13	4.00
NLR 3881	78.04	23.39	0.92
MTU 1318	82.14	23.28	3.01
BPT3278	72.24	20.95	-3.92
BPT 5204	72.66	22.19	5.14

MTU 2823-23-1-1	82.92	25.38	3.64
RGL 7044	72.83	25.61	1.56
RGL 7051	74.81	23.38	0.07
RGL 7049	84.73	21.14	-1.97
MTU 2878-13-1-1	75.47	25.21	0.21
RGL 7050	79.41	18.46	-2.59
MTU 2675-2-2-2-1	78.21	29.07	-0.69
RGL 7039	80.94	24.50	0.54
MTU 1310	84.06	23.92	2.99
BPT 3485	80.32	27.06	1.40
BPT 3354	71.29	22.09	4.04
MTU2716	66.13	29.73	-2.34
RGL 2537	82.74	27.58	0.85

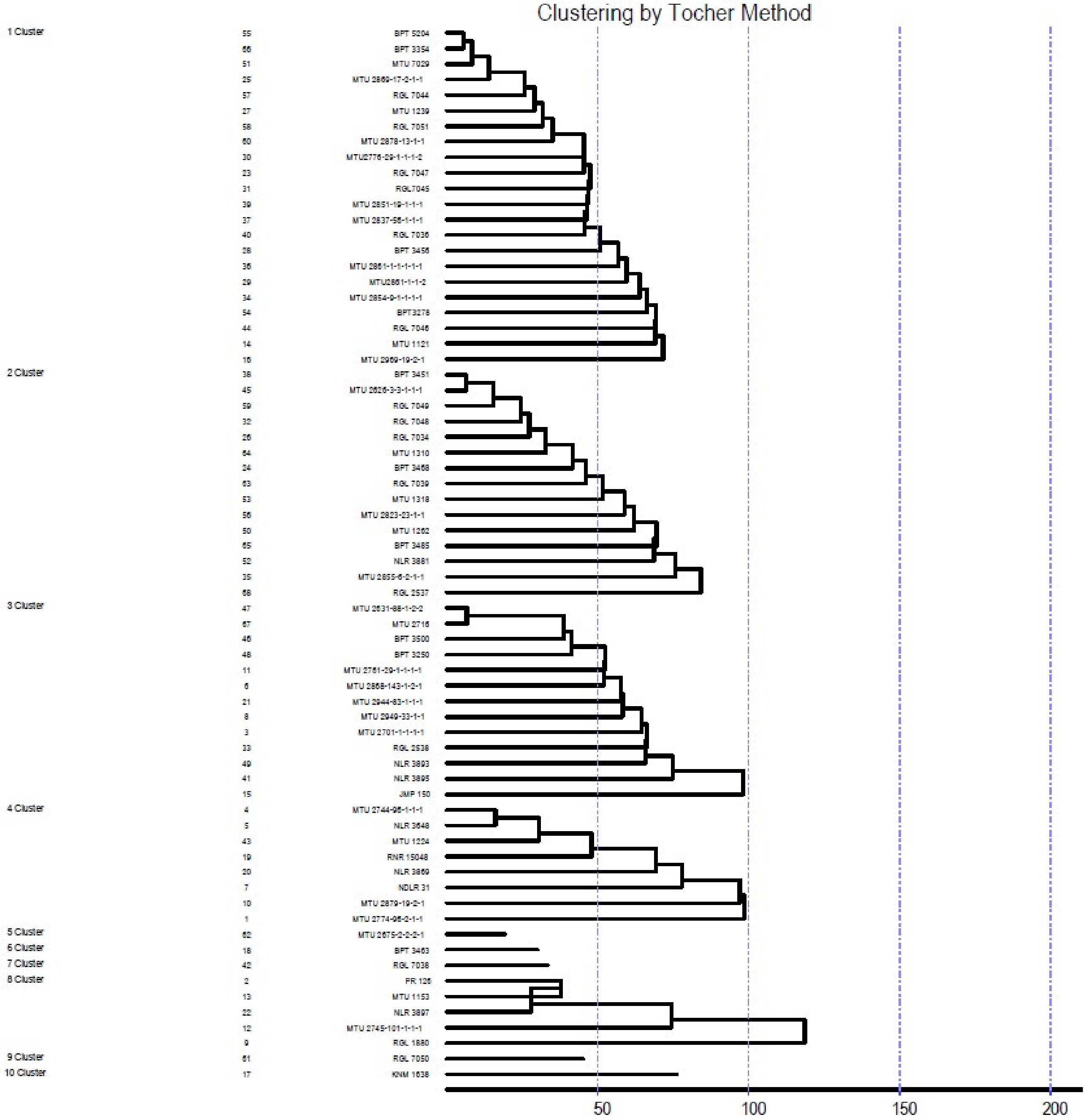

Fig. 1 : Dendrogram depicting the diversity among advanced lines in different clusters

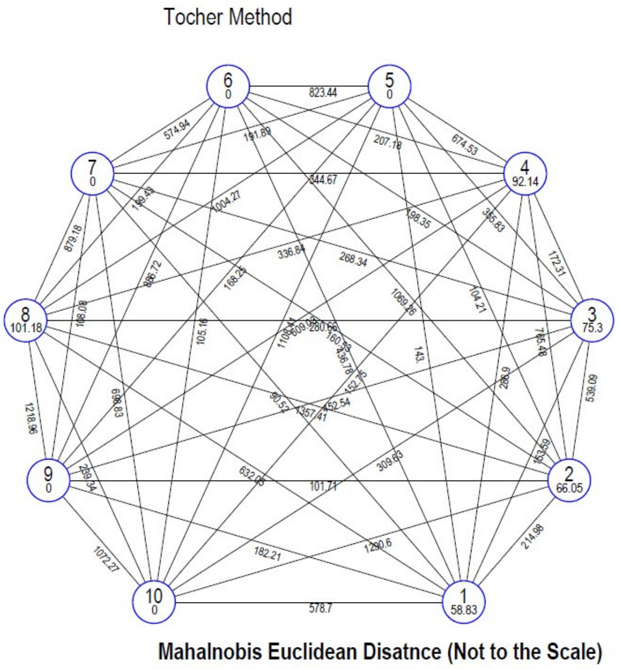

Fig. 2 : Intra and inter-cluster distance of genotypes in ten clusters based on Tocher's method

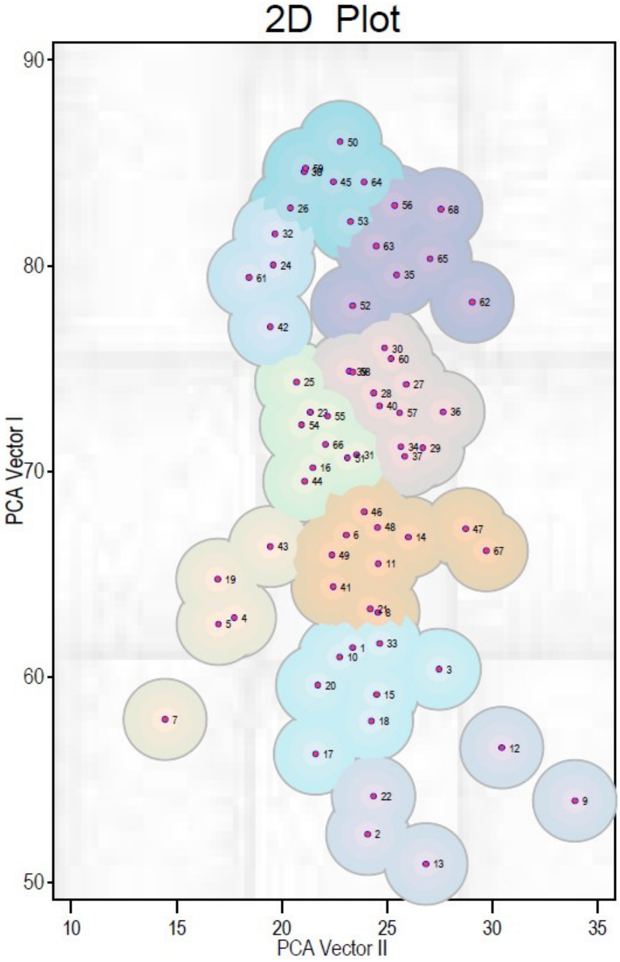

Fig. 3 : PCA 2D Biplot explaining the distribution of genotypes on different axis.

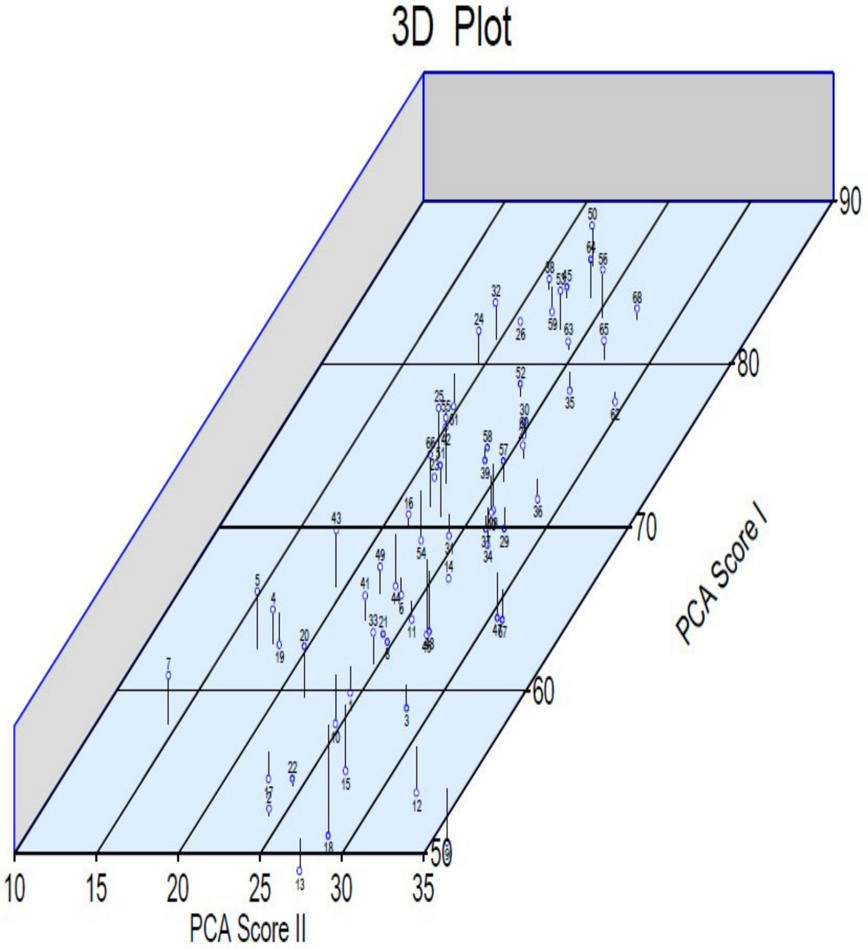

Fig. 4 : PCA 3D Biplot depicting the diversity among genotypes on different axis.

References

Allam, C.R., Raju, C.D., & Vindhiyavarman, P. (2017). Genetic divergence studies in rice (*Oryza sativa* L.) for yield and yield components. *International Journal of Current Microbiology and Applied Sciences,* **6**(11), 453–460.

Chandraker, P., Verulkar, S. B., Chandel, G., & Sharma, D. (2024). Multivariate analyses for identifying promising rice genotypes for yield and adaptability. *Journal of Crop Improvement*, **38**(2), 175–192.

Chemutai, L. R., Musyoki, M. A., Kioko, W. F., Mwenda, N. S., & Muriira, K. G. (2016). Genetic diversity studies on selected rice (*Oryza sativa* L.) genotypes based on gel consistency and alkali digestion. *Journal of Rice Research,* **4**(3), 172. https://doi.org/10.4172/2375-4338.1000172

Choudhary, M., Singh, A. K., Kumari, P., & Jaiswal, J. P. (2022). Genetic diversity and trait association analysis in rice using multivariate approaches. *Physiology and Molecular Biology of Plants,* **28**(6), 1259–1270.

Duppala, M.K., Srinivas, T., Rao, L.S., Suneetha, Y., Sundaram, R.M., Kumari, V.P., Bhuvaneswari, V and Ganesh, B. 2023. Screening of recombinant inbred lines for resistance to bacterial leaf blight pathotypes in rice (*Oryza sativa* L.). *Plant Science Today*. **10**(3): 343–353.

Fornasiero, A., Wing, R. A., & Ronald, P. (2022). Rice domestication. *Current Biology*, **32**(1), R20-R24.

Fujino, K., Matsuda, Y., & Sekiguchi, H. (2020). Genetic dissection of heading date using diversity panels in rice. *Breeding Science,* **70**(2), 233–242.

Jeong, S., Kim, Y. J., Lee, J., & Ha, S. (2017). Principal component analysis to evaluate genetic diversity in rice germplasm. *Korean Journal of Crop Science,* **62**(3), 217–225.

Khatun, F., Akter, A., & Rasul, M.G. (2023). Genetic variability and multivariate analysis in rice (*Oryza sativa* L.). *Bangladesh Journal of Agricultural Research,* **48**(1), 15–28.

Kim, J. H., Park, S., & Lee, G. (2024). Genome-enabled selection and diversity assessment in rice breeding populations. *Plant Breeding*, **143**(1), 55–68.

Kumar, A., Singh, R., & Patel, S. (2021). Integration of earliness and yield traits for improving crop productivity: Advances in breeding strategies. *Journal of Crop Improvement*, **35**(4), 512–528.

Kumar, M.D., Beena, N and Gowtham, K.S. 2017. Evaluation of Advanced Lines of Mustard Spp. Using Hierarchical Cluster Analysis. *The Bioscan*. **10** (1): 299-307.

Kumar, M.K., Satyanarayana, P.V., Udayababu, P., Sreenivas, G and Manojkumar, D. 2023. Principal Component Analysis for Yield in Black Gram Genotypes (*Vigna mungo* L. Hepper) under Rice Fallow Pulse Cropping

System. *Biological Forum – An International Journal.* **15**(3): 930-934.

Madhukumar, K., Satyanarayana, P.V., Udayababu, P., Sreenivas, G and Manojkumar, D. 2023. Principal Component Analysis for Yield in Green Gram Genotypes [*Vigna radiata* (L.) Wilczek] under Rice Fallow Pulses Conditions. *Frontiers in Crop Improvement.* 11(Special Issue-III): 1966-1970.

Manoj, K.D., Srinivas, T., Rao, L.S., Suneetha, Y., Sundaram, R.M. and Kumari, V.P., 2022. Genetic Variability and trait association analysis in F_3 population of YH3 x AKDRMS 21-54 cross. *The Andhra Agricultural Journal.* **69** (1): 46-57.

Manojkumar, D., Beena, N., Gowtham, K.S and Rathod, A. 2018. Evaluation of Advanced Breeding Lines of Indian Mustard Using Principal Component Analysis. *Journal of Oilseed Brassica*. **9**(1): 45-48.

Nguyen, T. T., Lee, J. H., & Park, Y. J. (2019). Multivariate analysis for evaluating genetic diversity in cultivated crops: Insights from morphological and yield-related traits. *Crop Science and Biotechnology*, **22**(3), 247–258.

Sanyal, A., Joshi, M. A., Tomar, B. S., & Aggarwal, D. (2016). Phenotypic divergence for agro-morphological traits among extant rice (*Oryza sativa*) varieties. *Journal of Applied and Natural Science,* **8**(1), 96–101.

Sar, A., & Kole, P. (2023). Evaluation of genetic divergence in rice using multivariate analyses. *Oryza,* **60**(2), 235–243.

Satyanarayana, P.V., Madhukumar, K., Udayababu, P., Srinivas, T and Manojkumar, D. 2023. Association Studies for Identifying the Selection Criteria Among Early varieties of Rice in North Coastal Zone of Andhra Pradesh. *Biological Forum – An International Journal,* **15**(11): 329-335.

Singh, N., Choudhury, D. R., Tiwari, G., Singh, A. K., *et al.* (2016). Genetic diversity trend in Indian rice varieties: An analysis using SSR markers. *BMC Genetics,* **17**, 127. https://doi.org/10.1186/s12863-016-0437-7

Sinha, D., Maurya, A. K., Abdi, G., Majeed, M., Agarwal, R., Mukherjee, R., ... & Chen, J. T. (2023). Integrated genomic selection for accelerating breeding programs of climate-smart cereals. *Genes*, **14**(7), 1484.

Swarup, S., Cargill, E. J., Crosby, K., Flagel, L., Kniskern, J., & Glenn, K. C. (2021). Genetic diversity is indispensable for plant breeding to improve crops. *Crop Science*, **61**(2), 839-852.

Thakur, P., & Sarma, D. (2023). Genetic approaches for enhancing early maturity and yield potential in field crops: Recent perspectives. *Plant Breeding Research*, **29**(2), 145–158.

Udayababu, P., Satyanarayana, P.V., Madhu Kumar, K and Manojkumar, D. 2024. Efficacy of Insecticides against Sucking Pests of Rice Fallow Blackgram in the Srikakulam District of North Coastal Andhra Pradesh. *Progressive Research : An International Journal.* **19** (Special-1): 232-237.

Udayababu, P., Satyanarayana, P.V., Madhu Kumar, K., Srinivasa Rao, M., Sreenivas, G and Manojkumar, D. 2025. Evaluation of the Bio-efficacy of Granular Insecticides in Managing Paddy Pest Infestations. *Journal of Advances in Biology & Biotechnology*, **28** (5); 70-83.

Xu, Q., Yuan, X., Wang, S., Zhao, C. *et al*. (2016). The genetic diversity and structure of indica rice in China as detected by single nucleotide polymorphism analysis. *BMC Genetics,* **17**, 53. https://doi.org/10.1186/s12863-016-0361-x.