GMR
The Original

EXPLORING GENETIC VARIABILITY THROUGH MORPHOLOGICAL AND MOLECULAR APPROACHES IN LOWLAND RICE (ORYZA SATIVA L.) OF BIHAR

Nitesh Kushwaha[1,3*], Ravi Kant[2*], Rajesh Kumar[3], Prem Chand Gyani[4] Tushar Arun Mohanty[5], Kanchan Kumari Gupta[6], Digvijay Singh[7], Amrita Thomas[1], H.M. Anil[1], Bobochand Singh Oinam[1],

[1*] Division of Genetics, ICAR-Indian Agricultural Research Institute, Pusa, New Delhi.
[2.] Madhopur Research Station, Department of Plant Breeding and Genetics, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar
[3.] Department of Plant Breeding and Genetics, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar
[4.] Division of Genetics and Tree Improvement, ICFRE- Rainforest Research Institute, Jorhat, Assam.
[5.] Department of Genetics and Plant Breeding, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu.
[6.] Division of Vegetable Science, Assam Agricultural University, Jorhat, Assam
[7.] Department of Seed Science and Technology, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya.
*Corresponding author Email: farmernk.998@gmail.com

ABSTRACT

To evaluate the genetic diversity and phylogenetic affinities between 22 distinct rice (*Oryza sativa* L.) genotypes based morphological traits and SSR based molecular analysis, research work was conducted at *Rice Breeding Section(for DUS trait evaluation) and Rice Molecular Lab (for SSR marker analysis), Faculty of Agriculture, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur*. Through DUS based characterization, important differences were found in characteristics like auricle pigmentation, basal sheath color, stigma color, culm attitude, and Attitude of flag. A few genotypes displayed anthocyanin pigment in their leaf collars and purple auricles, which suggests trait-specific diversity. The range of grain morphology was short to medium in length and narrow to wide in breadth, therefore highlighting the phenotypic variation among genotypes. SSR based assessment of genetic diversity using 15 SSR markers revealed overall genotypic variance ranging from 0. 00 to 0. 994, the polymorphism information content (PIC) values averaged about 0. 75, showing that the markers were quite informative. Shannon's Diversity Index (H = 3. 091), Simpson's Index (λ = 0. 9545), and evenness (E. 5 = 1. 0) stating considerable allelic richness and diversity. Principal Coordinate Analysis (PCoA) clarified 42. 8% of total variation and exposed different spatial clusters; UPGMA clustering grouped the genotypes into four main clusters at approximately 60% similarity, hence showing significant genetic variance. The coexistence of both frequent and rare alleles as well as varied degrees of genetic distance helps to highlight the significance of these genotypes for germplasm preservation and breeding techniques. Combining morphological and molecular data offers a solid foundation for parent selection in rice improvement initiatives targeting yield, stress resistance, and genetic improvement.

KEYWORDS: Rice, genetic diversity, SSR markers, morphological traits, PIC, PCoA, clustering, breeding.

INTRODUCTION

Rice stands as one of the most indispensable food crops worldwide, serving as the primary nutritional source for over half of the global population (Mohidem et al., 2022). Its universal significance has earned it the “Global Grain” (Debsharma et al., 2022). By 2050, the projected demand for rice will reach approximately 943.60 million tons, requiring an annual increase of 5.80 million tons from current production levels (FAO, 2017). However, meeting this demand is challenging due to the reduction of cultivable land by over 1% due to industrial, residential, and infrastructural development (Mohanty et al., 2025). Currently, over 800 million people worldwide experience hunger and food insecurity. Increasing rice productivity has become inevitable due to limited arable land for cultivation and the expanding global population (Mohanty et al., 2025). In the 2023–24 crop year, India’s rice production reached 137.82 million tonnes, cultivated across 47.6 million hectares (Chaturvedi et al., 2023). With an ever-growing demand, continuous efforts are being made to enhance per-hectare productivity.

One of rice’s most defining characteristics is its extraordinary genetic diversity, enabling it to thrive across various ecological conditions. This extensive intra-species variation plays a crucial role in global food security, sustainable agriculture, and genetic improvement (Kimwemwe et al., 2023). Rice germplasm serves as a vast reservoir of valuable genes and gene combinations, making it a vital resource for breeding programs aimed at enhancing desirable traits (Gaikwad et al., 2021). Given its significance, the agro-morphological characterization of rice accessions is an essential step in conserving and utilizing this diversity effectively.

Characterization involves the systematic documentation of key descriptors that aid in differentiating accessions, registering new varieties, and streamlining data retrieval for breeding programs (Lusty et al., 2021). To standardize

this process, many countries have adopted Distinctness, Uniformity, and Stability (DUS) testing guidelines, which provide a structured approach to characterizing varieties, landraces, and cultivars. These guidelines, formulated based on the principles of the International Union for the Protection of New Varieties of Plants (UPOV), help establish clear distinctions between different varieties (Gayathri et al., 2023).

Recognizing India's unique agricultural and socio-economic landscape, the government introduced the Protection of Plant Varieties and Farmers' Rights Act (PPV&FRA) in 2001 as an alternative to the UPOV model (Brahmi et al., 2004). This legislation not only protects the intellectual property rights of plant breeders but also safeguards the interests of farmers while promoting genetic conservation and crop improvement.

Despite their effectiveness, agro-morphological descriptors are inherently influenced by environmental factors, making conventional characterization methods prone to inconsistencies. Additionally, the growing number of reference varieties presents challenges in achieving precise identification using only traditional descriptors (Ansari et al., 2023). To overcome these limitations, molecular markers have emerged as a highly reliable and efficient alternative, enabling more specific and accurate characterization of rice germplasm.

Early molecular marker systems, such as Restriction Fragment Length Polymorphism (RFLP) and Random Amplified Polymorphic DNA (RAPD), marked the beginning of genetic characterization in rice (Amiteye, 2021). However, advancements in molecular biology have led to the development of more precise and widely distributed markers, such as Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphisms (SNPs). Compared to conventional morpho-physiological traits, molecular markers provide superior tools for assessing genetic variation and understanding genetic relationships within and among species (Özbek, 2024).

Among these, SSR markers have gained preference due to their high polymorphism, co-dominant inheritance, abundance across the genome, and strong reproducibility. Additionally, their cost-effectiveness and time efficiency make them an ideal choice for genetic characterization (Xu et al., 2024). By integrating agro-morphological and molecular characterization techniques, researchers can achieve a more comprehensive and reliable assessment of germplasm diversity, leading to more precise breeding strategies.

Flooding remains one of the most significant abiotic stressors affecting rice cultivation in Bihar, threatening crop productivity and farmer livelihoods. This study focuses on the characterization of lowland rice cultivars grown in flood-prone regions of Bihar, utilizing a combination of DUS-based descriptors and SSR markers. By adopting an integrated approach, this research aims to enhance genetic understanding and contribute to the development of more resilient rice varieties suited for challenging environmental conditions.

MATERIAL AND METHODS

Plant materials

In this study, twenty-two lowland rice cultivars, viz. *Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Janaki, Barogar, Madhukar, Singhara, Ujala Dhusaria, Sagar Samba, Jagannath Ballava, Brasali, Meghnad, Silhat, Shyamala, BPT-5204-Sub-1, Kishori* were used. This list includes lowland improved varieties and as well as landraces grown in Bihar *(Table 1)*

Experimental site:

The present study was conducted during the kharif season of 2019 at the Rice Breeding Section, Pusa Farm, under Dr. Rajendra Prasad Central Agricultural University (RPCAU), Pusa, Samastipur, Bihar. The research site is geographically positioned at 25.98°N latitude and 85.67°E longitude, with an elevation of 51.8 meters above mean sea level. Located within the agro-climatic zone of the Middle Gangetic Plains, this region provides an ideal environment for rice cultivation and agricultural research. The kharif season, spanning from June to October, coincides with the monsoon period, offering optimal conditions for rice growth and field experimentation.

Bihar, situated in eastern India, extends from 24°20'10" N to 27°31'15" N latitude and from 83°19'50" E to 88°17'40" E longitude. Its subtropical climate is characterized by extreme seasonal variations, with summer temperatures soaring to 45–47°C, while winter temperatures can plummet to as low as 2–3°C in December and January. The state experiences an average annual temperature of 25.5°C, with an annual rainfall of approximately 1,097 mm, primarily influenced by the southwest monsoon.

For agro-morphological characterization, the genotypes were sown in a randomized block design (RBD) with three replications, maintaining a standard of 20 × 15 cm. The evaluation was carried out using DUS descriptors in accordance with the guidelines issued by the Directorate of Rice Research, Hyderabad (Table: 4). Thirty-six traits were taken under study, that *are Coleoptile color, Leaf: intensity of green colour, Culm attitude, attitude of flag leaf (early observation), Presence of auricles on leaf, Presence of pubescence on leaf blade surface, Anthocyanin coloration of auricles, Presence of collar on leaves, Anthocyanin coloration of collar, Presence of ligule on leaf, Shape of ligule, Length of leaf blade, Attitude of flag leaf (Late observation), Leaf senescence, Anthocyanin coloration of nodes, Intensity of anthocyanin coloration of nodes, Anthocyanin coloration of internodes, Color of stigma, Curvature of main axis of panicle, Density of pubescence of lemma, Anthocyanin coloration on tip of lemma,*

Anthocyanin coloration of area below apex of lemma, Anthocyanin coloration of keel, Presence of secondary branching on panicle Distribution of awns, Exertion of panicle, Color of lemma and palea, Awns, Color of awns, Distribution of awns, Presence of anthocyanin in leaf, Grain length, Grain width, Decorticated grain aroma. All the twenty-two genotypes were evaluated for these traits. Throughout the experiment, standard crop management practices were diligently followed to ensure optimal growth and accurate assessments

Molecular characterization:

DNA Amplification

For molecular characterization, DNA was extracted using the standard CTAB method from 200 mg of rice leaf tissue collected at the seedling stage. PCR amplification was performed using fifteen SSR primer pairs (Table 2) sourced from the Gramene database. The final PCR reaction was set up in a 25.0 µl volume, consisting of: 1.0 µl (20 ng/µl) DNA template, 2.0 µl dNTPs (2.5 mM each), 0.5 µl Taq polymerase (3U/µl), 2.5 µl reaction buffer (10X), 1.0 µl forward primer (20 µM), 1.0 µl reverse primer (20 µM), and 17.0 µl nuclease-free water. The thermal cycling conditions included an initial denaturation at 94°C for 3 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 1 minute. A final extension was performed at 72°C for 7 minutes, and the PCR products were stored at 4°C until further analysis.

Gel Electrophoresis and Band Visualization

Following PCR amplification, the DNA samples were subjected to agarose gel electrophoresis using a 2% agarose gel prepared in 1X TBE buffer. Ethidium bromide (10 µl) was incorporated into 200 µl of TBE buffer for staining. The electrophoresis was carried out at 110 volts for 1 hour and 10 minutes. After separation, the gel was visualized under a UV transilluminator in a gel documentation system, and images were captured using an integrated camera. The banding patterns generated by each primer set were analyzed separately, and fragment sizes were determined by comparing the migration distances against a 50 base pair (bp) DNA ladder. The presence of a band at a specific position was recorded as '1,' while its absence was noted as '0' for subsequent data analysis.

Statistical Analysis.

Genetic diversity analysis was performed on 22 rice genotypes using 36 qualitative morphological traits. The dataset, stored in CSV format, was imported and processed in R (version 4.4.3). All traits were treated as categorical and converted to factors. Multiple Correspondence Analysis (MCA) was employed using the FactoMineR and factoextra packages to visualize trait and genotype associations in reduced dimensions. Two key plots were generated: MCA variable plot and genotype plot, colored by contribution and cos^2 values, respectively.To assess genetic relationships, pairwise dissimilarities were calculated using Gower's distance, suitable for qualitative data, through the cluster package. Hierarchical clustering was performed using Ward's method (hclust), and a dendrogram was visualized with branch colors representing four major clusters using the dendextend package. A dissimilarity heatmap was generated using gplots::heatmap.2, applying the "viridis" color palette to highlight trait-based differentiation among genotypes. Cluster validation was conducted using silhouette analysis (fviz_silhouette) to evaluate intra-cluster homogeneity and inter-cluster separation.

Genetic diversity analysis was conducted using R software, employing multiple specialized packages. SSR marker data were imported using the **readxl** package and formatted with **dplyr**. Genotypic data were structured as a *genind* object using **adegenet**, assuming diploid, codominant markers. Genetic distances were computed, and a dendrogram was generated using the UPGMA method via the **ape** package. **Principal Coordinate Analysis (PCoA)** was performed to visualize genotype relationships. Population structure was inferred using the *find.clusters* function from **adegenet**, employing K-means clustering. A strata data frame was constructed and integrated into the genclone object with the **poppr** package. **Analysis of Molecular Variance (AMOVA)** was conducted using **poppr** to assess genetic differentiation within and among inferred populations. The **pegas** package supported distance-based analyses. All analyses used default parameters unless otherwise specified, and graphical and tabular outputs were saved automatically to organized result directories.

RESULT AND DISCUSSION

DUS Characterization:

This study meticulously characterized lowland rice cultivars of Bihar using 36 DUS descriptors, adhering to the guidelines of the Directorate of Rice Research, Hyderabad. The findings, detailed in *Table 3 &4,* reveal significant variations in key morphological traits among the 22 genotypes.

Hierarchical cluster analysis (*Figure 2*)of 22 rice genotypes based on 36 qualitative characters revealed distinct genetic relationships among the studied cultivars. The dendrogram generated using agglomerative clustering partitioned the genotypes into four major clusters at approximately 0.8 height threshold, Cluster 1 comprised only two genotypes *(Janaki and Meghnad),* indicating their close genetic similarity and distinct separation from other cultivars. Cluster 2

formed the largest group with seven genotypes *(Rajshree, Madhukar, Sarisa, Shyamala, Singhara, Sudha, And Vaidehi)*, suggesting shared morphological characteristics among these varieties. Cluster 3 contained a single genotype *(Radha),* representing its unique qualitative trait profile compared to other studied genotypes. Cluster 4 was the most diverse, encompassing twelve genotypes including *Santosh, Pankaj, Kanak, Satyam, Sagar Samba, Swathi Ballava, BPT-5204-Sublm, Swarna, Brasali, Kishori, Barogar, and Silhat.* Figure(), shows dendrogram showing grouping of genotypes.

Molecular characterization

1. Assessment of Genetic Variability Among Rice Genotypes

The genetic variability (*Table 6*) among the 22 rice genotypes was assessed using multilocus genotyping data. Each genotype displayed a unique multilocus genotype (MLG = 22), which signifies the absence of redundancy and highlights the high level of genotypic uniqueness within the germplasm collection. Furthermore, the expected number of multilocus genotypes (eMLG) was also 22, and the standard error was zero (SE = 0), indicating consistency and uniformity in the dataset.

Diversity parameters revealed a Shannon's Diversity Index (H) value of 3.091, suggesting a broad distribution of allelic richness across loci. Similarly, the Simpson's Diversity Index (λ) was 0.9545, indicating a very high probability that two randomly selected genotypes would be genetically different. The expected heterozygosity (H exp) was calculated as 0.298, representing a moderate level of gene diversity, which is consistent with the self-pollinated nature of rice.

Additionally, the evenness index (E.5 = 1.0) revealed an equal contribution of all genotypes to the observed genetic diversity. The Index of Association (Ia) and the standardized index ($\hat{h}D$) were recorded as 1.005 and 0.030, respectively, both of which are low and suggest minimal linkage disequilibrium, thus supporting the assumption of independent segregation of markers.

2. Distribution of Allele Frequencies and Marker Informativeness

The distribution of allele frequencies (*Table 8 and fig. 5*) across all loci indicated the presence of both common and rare alleles. While several alleles were found to be widespread among the genotypes, a significant proportion of alleles were rare and occurred at low frequencies. These rare alleles may play a crucial role in carrying novel genetic variation, which could be of potential interest in breeding programs targeting stress tolerance and adaptation.

Polymorphism Information Content (PIC) (*Table 8 and fig. 6*) analysis showed that majority of markers had PIC values above 0.5, reflecting their high informativeness. The visual distribution of PIC values exhibited a skew toward higher values, confirming the utility of the selected marker set for effective genotypic discrimination and diversity estimation.

3. Genotype-Specific Allelic Statistics

The analysis of genotype-specific allele profiles *(Table 7)* showed that the total number of alleles per genotype varied between 14 and 16. The number of unique alleles identified per genotype ranged from 0 to 2. Genotypes such as *Santosh, Kanak*, and *Pankaj* did not contribute any unique alleles, suggesting these genotypes share most of their alleles with others. In contrast, *Radha, Satyam, Vaidehi, and Brasali* each possessed one unique allele, indicating some degree of genetic distinction. Notably, *Swarna and Silhat* stood out by possessing two unique alleles each, highlighting their potential as distinct sources of novel alleles.

4. Locus-Specific Diversity Parameters

A total of 40 loci were analyzed for their contribution to allelic diversity. The PIC *(Table 8*) values among these loci ranged from 0.00 to 0.994, with an average value of approximately 0.75, indicating high overall informativeness. Certain loci, including RM 70_m1 and MRG-2805-HAU-2805_m2, showed the highest PIC values of 0.994, while RM 521_m1/m3 followed closely with a PIC of 0.978. These loci are highly polymorphic and are therefore particularly useful for distinguishing closely related genotypes.

Conversely, some loci such as MRG 2894 IRRI2894_m1 and RM321_m1 were found to be monomorphic with a PIC of 0.000, indicating a lack of variation and limited use in assessing diversity. Multiple loci, such as RM 566, RM 520, and RM 70, displayed more than three allelic forms, contributing significantly to the resolution power of the marker panel. Several rare alleles with frequencies less than or equal to 0.09 were detected in loci such as RM 555_m3 and RM 70_m1. The number of genotypes sharing a specific allele varied widely, with some alleles found in only a single genotype and others present in all 22 genotypes.

5. Genetic Relationships and Clustering Pattern

The UPGMA-based dendrogram *(Figure 9)* illustrated the genetic relationships among the 22 rice genotypes. Based on a similarity threshold of approximately 60%, the genotypes were grouped into four major clusters. The genetic similarity coefficients among the genotypes ranged from 0.20 to 1.00, indicating a wide range of genetic diversity.

Cluster I included genotypes *Kanaka, Radhasa, and Santosh*, which showed high genetic similarity. Cluster II comprised Pankaj, Rajshree, and BPT-5204 *Sub-1*, indicating a moderate level of relatedness. Cluster III contained *Kishori, Singhara, Sudha, and Vaidehi*, which were closely related to each other. Cluster IV was the most diverse,

comprising genotypes *Satyam, Shyamala, Swarna, Barogar, Madhukar, Brasali, Jagannath Ballava, Silhat, Meghnad, Sagar, Samba, Janaki, and Ujala Dhusaris*, reflecting substantial genetic divergence.

6. Genetic Distance Analysis

The genetic distance heatmap (*Figure 7*) illustrated the pairwise dissimilarity between genotypes, with values ranging from 0.0 to 0.6. Genotype pairs such as *Santosh and Radha, Kanak and Shyamala* shared low genetic distances, suggesting high genetic similarity. Conversely, pairs such as *Santosh and Dhusarisa, and BPT-5204-Sub1* and *Jagannath Ballava*, displayed high dissimilarity, indicating significant genetic divergence.

Certain genotypes such as *Brasali, Madhukar, and Vaidehi* exhibited moderate genetic distances with other genotypes, suggesting they occupy an intermediate position in terms of genetic relatedness and may serve as potential bridge parents in breeding programs.

7. Principal Coordinate Analysis (PCoA)

The Principal Coordinate Analysis (PCoA) ***(Figure 8)*** helped visualize the genetic structure among the 22 rice genotypes. The first principal coordinate (PCoA1) accounted for 24.9% of the genetic variation, while the second coordinate (PCoA2) accounted for 17.9%, together explaining a cumulative 42.8% of the total variance.

The plot revealed a clear spatial separation among the genotypes. Genotypes such as *Santosh, Radha, And Kanak* were closely grouped, indicating high genetic similarity. On the other hand, *Ujala Dhusarisa* and *BPT-5204-Sub-1* were positioned distantly, reflecting substantial genetic divergence. Genotypes like *Brasali, Shyamala, and Madhukar* appeared between the main clusters, indicating moderate divergence and potential for use in recombination breeding.

8. PIC Value Distribution

The distribution of Polymorphism Information Content (PIC) values across the molecular markers ranged from 0.0 to approximately 0.9. Most markers showed PIC values between 0.4 and 0.8, which is considered indicative of moderate to high informativeness. Approximately 10 to 12 markers exhibited PIC values in the range of 0.5 to 0.7. A small number of markers had PIC values below 0.2, suggesting limited effectiveness in differentiating genotypes.

These results confirm that the SSR markers used in this study were effective in capturing allelic variation and in supporting genetic diversity assessments.

9. Shared Alleles Among Genotypes

The analysis of shared alleles (Figure 8) using a heatmap revealed that the number of shared alleles between genotype pairs ranged from 6 to 16. Pairs such as *Santosh and Radha*, as well as *Kanak and Satyam*, shared a high number of alleles (14–16), indicating a strong genetic similarity. Conversely, genotypes like *BPT-5204-Sub-1* and *Jagannath Ballava*, and *Ujala Dhusarisa* and *Pankaj*, shared relatively few alleles (6–8), reflecting significant genetic divergence.

The observed variation in the number of shared alleles supports the classification of genotypes into distinct clusters and highlights the presence of both closely related and genetically distant genotypes. This information is valuable for selecting genetically diverse parents to maximize heterosis in rice breeding programs.

DISCUSSION

DUS Characterization

A striking uniformity was observed in coleoptile color, which remained consistently white across all genotypes. Auricles and leaf collars were present in every variety; however, four genotypes Ujala Dhusaria, Meghnad, Janaki, and Shyamala exhibited distinctive purple auricles. Similarly, anthocyanin pigmentation in leaf collars was noted in Rajshree, Madhukar, Janaki, Meghnad, and Ujala Dhusaria. Leaf length was universally long, and all genotypes displayed split ligules without exception. Variation in basal leaf sheath coloration was evident, with genotypes falling into three distinct categories: green (19 genotypes), light purple (2 genotypes), and uniform purple (1 genotype). Leaf color intensity also varied, ranging from light to medium and dark green. Most genotypes exhibited an erect culm, though Janaki and Meghnad had a semi-erect posture, while Singhara showed a spreading growth pattern. The stigma was predominantly white, though purple stigma coloration was observed in Janaki, Madhukar, Meghnad, and Shyamala.

Anthocyanin pigmentation was largely absent in nodes and internodes, except in Janaki, Madhukar, and Meghnad, where slight pigmentation was present. Strong anthocyanin coloration of the keel was a defining trait of Janaki. The tip of the lemma varied, with red coloration in Madhukar and Ujala Dhusaria, while Janaki, Shyamala, and Meghnad exhibited purple lemma tips. Flag leaf attitude in the early stages of growth was primarily erect, with two genotypes showing a semi-erect stance. In the later stages, most genotypes retained an erect or semi-erect flag leaf, while Jagannath Ballava and Singhara displayed a horizontal orientation. None of the genotypes in this study were aromatic, as they failed to produce any noticeable fragrance during cooking. Panicle exertion was well-exerted or mostly exerted in all cases, except for Singhara, which was the only awned genotype, bearing long white awns. Secondary branching was consistently present across all genotypes.

Grain morphology revealed diverse structural traits. Grain length varied from short to medium, while width ranged from narrow to medium. Among the genotypes, BPT-5204-Sub-1 had the narrowest grains, whereas Barogar exhibited

the broadest grains. The clustering pattern indicates varying degrees of genetic diversity among the rice genotypes, with some showing close relationships while others demonstrate distinct characteristics. This genetic variability information can be valuable for rice breeding programs, cultivar selection, and conservation strategies. The dendrogram effectively illustrates the phylogenetic relationships and genetic distances between genotypes based on Their qualitative morphological traits. These results align with previous findings from studies conducted by Gayathri et al. (2023), Lavanya et al. (2021), Kalyan et al. (2017), Prasanna et al. (2024), Ansari et al. (2023), Gaikwad et al. (2021), Pourbed et al., 2015 and Mondal et al. (2014), reinforcing the reliability of the observed traits.

Molecular characterization

The comprehensive molecular characterization of 22 rice genotypes using SSR markers revealed substantial genetic variation, underscoring the importance of conserving and utilizing genetic resources in breeding programs. The high Shannon's and Simpson's diversity indices observed in this study confirm the broad genetic base of the evaluated germplasm, which is essential for sustainable crop improvement (Huang et al., 2023; Rao et al., 2021; Sar et al., 2024) The presence of unique multilocus genotypes (MLGs) for all entries highlights the distinctiveness of each genotype and the absence of redundancy within the panel (Soro et al., 2024; Mohanty et al., 2024; Bui et al., 2025). The moderate expected heterozygosity and high evenness index are consistent with the autogamous nature of rice, which generally limits cross-pollination and recombination events (Tin et al., 2001). Nonetheless, the observed diversity metrics suggest that even within self-pollinating crops, considerable allelic variability can exist when carefully selected genotypes are included (Begana, 2021; Hussian et al., 2021)

Allele frequency distribution indicated the presence of rare alleles, which are often overlooked yet may hold key alleles for traits such as stress tolerance or grain quality (Raj & Nadarajah, 2022; Ali et al., 20217; Lee et al., 2024; He et al., 2024). The high Polymorphism Information Content (PIC) values observed for most markers validated the efficiency of the marker set for assessing genetic diversity and differentiating closely related accessions (Al-Daej et al., 2023; Singh et al., 2024)

The detection of both unique and shared alleles among genotypes provides further insight into the genetic relationships within the panel. Genotypes like SWARNA and SILHAT, which contributed unique alleles, represent valuable genetic resources for introducing novel variation into breeding populations. (Allier et al., 2020; Salgotra, & Chauhan, 2023)

Locus-specific diversity analysis showed that several SSR markers were highly informative, with PIC values approaching 0.99, while a few markers remained monomorphic (Singh et al., 2013; Bajracharya, 2023; Zewodu et al., 2024). The presence of multiple allelic forms in several loci suggests the effectiveness of these markers in capturing subtle genetic variation across genotypes (Kaur et al., 2015; Singh et al., 2024)

Cluster analysis based on UPGMA revealed four major genetic groups, indicating structured diversity and underlying genetic divergence (Zhang et al., 2022; Darkwa, et al., 2020). The presence of highly diverse genotypes in Cluster IV, comprising entries such as MADHUKAR and SHYAMALA, highlights the availability of broad variability, which can be exploited in parent selection for hybrid breeding. Similarly, the PCoA plot confirmed both convergence and dispersion among genotypes, emphasizing the multi-dimensional structure of the genetic relationships (Singh et al., 2024; Siddique et al., 2016; Tao et al., 2018).

Genetic distance estimates and shared allele matrices reinforced the clustering patterns and revealed contrasting genotypic combinations (Fatamatuzzohora et al., 2024; Nachimuthu et al., 2015). Genotypes with low shared alleles and high genetic distance, such as *BPT-5204-Sub-1* and *Jagannath Ballava* could be targeted for crosses aimed at maximizing heterosis.

Overall, the integration of multiple analyses—allele frequencies, diversity indices, genetic distance, and clustering—provides a robust framework for understanding molecular diversity in rice. The findings from this study offer valuable guidance for selecting genetically divergent parents, conserving unique alleles, and designing effective crossing strategies in rice improvement programs.

Acknowledgment

The authors are thankful to H.O.D., Department of Plant Breeding & Genetics, DRPCAU, Pusa, Samastipur, and Scientists and Staff of Rice Breeding Section, Pusa, for their kind support during investigation on the evaluation DUS and molecular characterization of all twenty-two rice cultivars.

Author Contribution

NK, TAM, DS, R, N and RK were involved in planning and designing of experiment, TAM, DS carried out the research work and PCG and KKG helped in data analysis, NK, DS and TAM, PCG, KKG were involved in preparation and R, RK and N reviewed the manuscript.

Statements & Declarations:

Funding

The authors did not receive support from any organization for the submitted work.

Data Availability
All data generated or analyzed during this study are included in this article and its supplementary information files. and its supplementary information files.

Ethical Approval and Consent to Participate
Not applicable.

Consent for Publication
Not applicable.

Competing Interests
The authors have no relevant financial or non-financial interests to disclose.

REFERENCES
1. Al-Daej, M. I., Rezk, A. A., El-Malky, M. M., Shalaby, T. A., & Ismail, M. (2023). Comparative genetic diversity assessment and marker–trait association using two DNA marker systems in rice (Oryza sativa L.). *Agronomy*, *13*(2), 329.
2. Ali, J., Xu, J. L., Gao, Y. M., Ma, X. F., Meng, L. J., Wang, Y., ... & Li, Z. K. (2017). Harnessing the hidden genetic diversity for improving multiple abiotic stress tolerance in rice (Oryza sativa L.). *PLoS One*, *12*(3), e0172515.
3. Allier, A., Teyssèdre, S., Lehermeier, C., Moreau, L., & Charcosset, A. (2020). Optimized breeding strategies to harness genetic resources with different performance levels. *BMC genomics*, *21*, 1-16.
4. Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon*, *7*(10).
5. Ansari, A., Pranesti, A., Telaumbanua, M., Alam, T., Wulandari, R. A., & Nugroho, B. D. A. (2023). Evaluating the effect of climate change on rice production in Indonesia using multimodelling approach. *Heliyon*, *9*(9).
6. Arunakumari, K., Durgarani, C. V., Satturu, V., Sarikonda, K. R., Chittoor, P. D. R., Vutukuri, B., ... & Sundaram, R. M. (2016). Marker-assisted pyramiding of genes conferring resistance against bacterial blight and blast diseases into Indian rice variety MTU1010. *Rice Science*, *23*(6), 306-316.
7. Bajracharya, J. (2003). *Genetic diversity study in landraces of rice (Oryza sativa L.) by agro-morphological characters and microsatellite DNA markers*. Bangor University (United Kingdom).
8. Begna, T. (2021). Conventional breeding methods widely used to improve self-pollinated crops. *International Journal of Research*, *7*(1), 1-16.
9. Brahmi, P., Saxena, S., & Dhillon, B. S. (2004). The protection of plant varieties and Farmers' Rights Act of India. *Current Science*, *86*(3), 392-398.
10. Bui, T. L., Truong, N. T., Luong, B. D., Nguyen, T. T., Tran, N. Q., Tran, G. H., ... & Tran, T. T. K. (2025). Multivariate analysis of agro-morphological traits and GBSS1 gene variation in Mekong Delta mega rice varieties. *Journal of Crop Science and Biotechnology*, *28*(2), 227-238.
11. Chaturvedi, P., Nahatakar, S. B., & Rajput, (2023). A. Growth of rice production in India. Multilogic in Science, 48(13), 1049-1052.
12. Coşkun, Ö. F., Toprak, S., & Mavi, K. (2025). Molecular characterization and population structure analysis of luffa (Luffa aegyptiaca Mill.) germplasm using inter-Primer Binding Site (iPBS) retrotransposon markers. *Genetic Resources and Crop Evolution*, 1-12.
13. Darkwa, K., Agre, P., Olasanmi, B., Iseki, K., Matsumoto, R., Powell, A., ... & Asfaw, A. (2020). Comparative assessment of genetic diversity matrices and clustering methods in white Guinea yam (Dioscorea rotundata) based on morphological and molecular markers. *Scientific Reports*, *10*(1), 13191.
14. Debsharma, S. K., Syed, M. A., Ali, M. H., Maniruzzaman, S., Roy, P. R., Brestic, M., ... & Hossain, A. (2022). Harnessing on genetic variability and diversity of rice (Oryza sativa L.) genotypes based on quantitative and qualitative traits for desirable crossing materials. *Genes*, *14*(1), 10.
15. Dwivedi, P., Ramawat, N., Dhawan, G., Gopala Krishnan, S., Vinod, K. K., Singh, M. P., ... & Singh, A. K. (2021). Drought tolerant near isogenic lines (NILs) of Pusa 44 developed through marker assisted introgression of qDTY2. 1 and qDTY3. 1 enhances yield under reproductive stage drought stress. *Agriculture*, *11*(1), 64.
16. Fatamatuzzohora, M., Islam, M. S., Rabbee, M. F., Hossain, M. S., Kang, S. G., & Matin, M. N. (2024). Genetic variation, population structure, and cluster analysis of rice (Oryza sativa L.) genotypes using morphological characteristics and molecular markers. *Cereal Research Communications*, *52*(4), 1751-1768.
17. Food and Agricultural Organization (FAO). (2017). Rice Market Monitor; Food and Agricultural Organization (FAO): Bangkok, Thailand; pp. 1–42.

18. Gaikwad, K. B., Singh, N., Kaur, P., Rani, S., Babu H, P., & Singh, K. (2021). Deployment of wild relatives for genetic improvement in rice (Oryza sativa). *Plant Breeding*, *140*(1), 23-52.
19. Gayathri, N. K., Srujana, Y., & Venkateswarlu, N. C. (2023). DUS characterization of rice (Oryza sativa L.) germplasm. *Electronic Journal of Plant Breeding*, *14*(1), 314-322.
20. Hannum, S., & Indriyani, R. (2018, December). Genetic diversity of local rice (Oryza sativa L.) from North Sumatera using simple sequence repeat (SSR). In *Journal of physics: conference series* (Vol. 1116, No. 5, p. 052025). IOP Publishing.
21. He, H., Leng, Y., Cao, X., Zhu, Y., Li, X., Yuan, Q., ... & Shang, L. (2024). The pan-tandem repeat map highlights multiallelic variants underlying gene expression and agronomic traits in rice. *Nature Communications*, *15*(1), 7291.
22. Hoque, M. I., Islam, M. M., Begum, S. N., Yasmine, F., & Khanom, M. S. R. (2021). Genetic diversity analysis of rice (Oryzae sativa L.) landraces using SSR markers in Bangladesh. *SAARC Journal of Agriculture*, *19*(2), 13-25.
23. Huang, Y. H., Ku, H. M., Wang, C. A., Chen, L. Y., He, S. S., Chen, S., ... & Kao, C. F. (2022). A multiple phenotype imputation method for genetic diversity and core collection in Taiwanese vegetable soybean. *Frontiers in Plant Science*, *13*, 948349.
24. Hussain, A., Arshad, K., Abdullah, J., Aslam, A., Azam, A., Bilal, M., ... & Abdullah, M. (2021). A comprehensive review on breeding technologies and selection methods of self-pollinated and cross-pollinated crops. *Asian Journal of Biotechnology and Genetic Engineering*, *4*(3), 35-47.
25. Kalyan, B., K.V. Radha Krishna and Subba Rao, L.V. 2017. DUS Characterization for Germplasm of Rice. *International Journal of Current Microbiology and Applied Sciences*. 6(10): 3480-3487.
26. Kaur, S., Panesar, P. S., Bera, M. B., & Kaur, V. (2015). Simple sequence repeat markers in genetic divergence and marker-assisted selection of rice cultivars: a review. *Critical Reviews in Food Science and Nutrition*, *55*(1), 41-49.
27. Kimwemwe, P. K., Bukomarhe, C. B., Mamati, E. G., Githiri, S. M., Civava, R. M., Mignouna, J., ... & Fofana, M. (2023). Population structure and genetic diversity of Rice (Oryza sativa L.) germplasm from the Democratic Republic of Congo (DRC) using DArTseq-Derived single nucleotide polymorphism (SNP). *Agronomy*, *13*(7), 1906.
28. Lavanya, K., Chiranjeevi, M., Surender, R., Yadav PAS, Fiyaz AR, Sudhakar, Rao LS. (2021). Characterization of rice traditional varieties (*Oryza sativa* L.) based on DUS descriptors' Pharma Innovation Journal. ;10(3):760-771.
29. Lee, S. Y., Jeung, J. U., & Mo, Y. (2024). Allelic combinations of Hd1, Hd16, and Ghd7 exhibit pleiotropic effects on agronomic traits in rice. *G3: Genes, Genomes, Genetics*, *14*(3), jkad300.
30. Lusty, C., Sackville Hamilton, R., Guarino, L., Richards, C., Jamora, N., & Hawtin, G. (2021). Envisaging an effective global long-term agrobiodiversity conservation system that promotes and facilitates use. *Plants*, *10*(12), 2764.
31. Mohanty T A, Dharmalingam K, Swaminathan M, Ramalingam S, Narayanan M B, Natarajan S. Multi-trait selection indices for identifying elite rice genotypes in rice breeding programs. Plant Science Today (Early Access). https://doi.org/10.14719/pst.6215.
32. Mohanty, S. P., Khan, A., Patra, S., Behera, S., Nayak, A. K., Upadhyaya, S., ... & Sah, R. P. (2024). Unraveling the genetic diversity in selected rice cultivars released in the last 60 years using gene-based yield-related markers. *Genetic Resources and Crop Evolution*, 1-15.
33. Mohanty, T. A., Kumaresan, D., Manonmani, S., Ramalingam, S., Boopathi, N. M., & Natarajan, S. (2025). Enhancing rice breeding through two-line hybrids: integrative analysis of combining ability, heterosis, MGIDI, and grain quality traits. *Euphytica*, *221*(3), 29.
34. Mohidem, N. A., Hashim, N., Shamsudin, R., & Che Man, H. (2022). Rice for food security: Revisiting its production, diversity, rice milling process and nutrient content. *Agriculture*, *12*(6), 741.
35. Mondal, B., Singh, S. P., & Joshi, D. C. (2014). DUS Characterization of Rice (Oryza Sativa L.) Using Morphological Descriptors and Quality Parameters. Outlook on Agriculture, 43(2), 131-137. https://doi.org/10.5367/oa.2014.0167.
36. Mukta, S., Bappy, M. N. I., Bhuiyan, J., Zohora, F. T., & Afrin, D. (2024). Assessment of genetic diversity in Bangladeshi rice (Oryza sativa L.) varieties utilizing SSR markers. *Gene Reports*, *37*, 102051.
37. Nachimuthu, V. V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, B. A., Ponniah, G.,& Sabariappan, R. (2015). Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in Oryza sativa. *Rice*, *8*, 1-25.
38. Özbek, Ö. (2024). Molecular Markers Used to Reveal Genetic Diversity and Phylogenetic Relationships in Crop Plants. *OBM Genetics*, *8*(4), 1-25.
39. Pourabed, Ehsan, Jazayeri Noushabadi, Mohammad Reza, Jamali, Seyed Hossein, Moheb Alipour, Naser, Zareyan, Abbas, Sadeghi, Leila, Identification and DUS Testing of Rice Varieties through Microsatellite Markers, *International Journal of Plant Genomics*, 2015, 965073, 7 pages, 2015

40. Prasanna, G. S., Joshi, J. L., & Muraleedharan, A. (2024). Distinctiveness, Uniformity and Stability (DUS) Characterization in Twenty-Five Landraces of Rice (Oryza sativa L.). *Journal of Advances in Biology & Biotechnology*, *27*(3), 77-84.
41. Raj, S. R. G., & Nadarajah, K. (2022). QTL and candidate genes: techniques and advancement in abiotic stress resistance breeding of major cereals. *International Journal of Molecular Sciences*, *24*(1), 6.
42. Rao, N. N. D., Roja, V., Tushara, M., Rao, V. S., & Rao, V. S. (2021). Characterization and diversity analysis of rice germplasm. In *Biological Forum–An Int. J* (Vol. 13, No. 3a, pp. 170-179).
43. Salgotra, R. K., & Chauhan, B. S. (2023). Genetic diversity, conservation, and utilization of plant genetic resources. *Genes*, *14*(1), 174.
44. Sar, P., Gupta, S., Behera, M., Chakraborty, K., Ngangkham, U., Verma, B. C., ... & Roy, S. (2024). Exploring genetic diversity within aus rice germplasm: insights into the variations in agro-morphological traits. *Rice*, *17*(1), 20.
45. Siddique, M. A., Khalequzzaman, M., Islam, M. Z., Rashid, E. S. M. H., Baktiar, M. H. K., & Chowdhury, M. A. Z. (2016). Genetic diversity in Aus rice genotypes using ILP markers. *Bangladesh Rice Journal*, *20*(2), 13-19.
46. Singh, A. K., Kumar, D., Gemmati, D., Ellur, R. K., Singh, A., Tisato, V., ... & Singh, A. V. (2024). Investigating genetic diversity and population structure in rice breeding from association mapping of 116 accessions using 64 polymorphic SSR markers. *Crops*, *4*(2), 180-194.
47. Singh, A. K., Kumar, D., Gemmati, D., Ellur, R. K., Singh, A., Tisato, V., ... & Singh, A. V. (2024). Investigating genetic diversity and population structure in rice breeding from association mapping of 116 accessions using 64 polymorphic SSR markers. *Crops*, *4*(2), 180-194.
48. Singh, N., Choudhury, D. R., Singh, A. K., Kumar, S., Srinivasan, K., Tyagi, R. K., ... & Singh, R. (2013). Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PloS one*, *8*(12), e84136.
49. Soro, M., Zida, S. M. F. W. P., Somé, K., Tiendrébéogo, F., Otron, D. H., Pita, J. S., ... & Koné, D. (2024). Estimation of genetic diversity and number of unique genotypes of Cassava Germplasm from Burkina Faso using microsatellite markers. *Genes*, *15*(1), 73.
50. Tao, R., Zhang, J., Bian, Y., Dong, R., Liu, X., Jin, C., ... & Li, C. (2018). Investigation of 12 X-STR loci in Mongolian and Eastern Han populations of China with comparison to other populations. *Scientific Reports*, *8*(1), 4287.
51. Tin, H. Q., Berg, T., & Bjørnstad, Å. (2001). Diversity and adaptation in rice varieties under static (ex situ) and dynamic (in situ) management. *Euphytica*, *122*, 491-502.
52. Xu, J., Du, R., Wu, K., & Chen, J. (2024). Development of SSR markers related to agarwood production and genetic diversity of Aquilaria sinensis (Lour.) Spreng wild populations. *Journal of Applied Research on Medicinal and Aromatic Plants*, *42*, 100565.
53. Zewodu, A., Mohammed, W., & Shiferaw, E. (2024). Analysis of genetic diversity and population structure of some Ethiopian barley (Hordeum vulgare L.) accessions using SSR markers. *Plos one*, *19*(6), e0305945.
54. Zhang, Y., He, Q., Zhou, X., Zheng, S., Wang, Y., Li, P., & Wang, Y. (2022). Genetic diversity and population structure of 93 rice cultivars (lines)(Oryza sativa Xian group) in Qinba in China by 3 types of genetic markers. *BMC genomics*, *23*(1), 550.

Table 1: List of twenty-two genotypes

SN0	GENOTYPE	SOURCE
1	SANTOSH	R.P.C.A.U., PUSA
2	KANAK	R.P.C.A.U., PUSA
3	RADHA	R.P.C.A.U., PUSA
4	PANKAJ	O.U.A.T., ORISSA
5	SATYAM	R.P.C.A.U., PUSA
6	RAJSHREE	R.P.C.A.U., PUSA
7	SWARNA	N.R.R.I, CUTTACK
8	SUDHA	R.P.C.A.U., PUSA
9	VAIDEHI	R.P.C.A.U., PUSA
10	JANAKI	R.P.C.A.U., PUSA
11	BAROGAR	R.P.C.A.U., PUSA
12	MADHUKAR	N.D.A.U.T, U.P.
13	SINGHARA	R.P.C.A.U., PUSA
14	UJALA DHUSARISA	LAND RACE, BIHAR
15	SAGAR SAMBA	O.U.A.T., ORISSA
16	JAGANNATH BALLAVA	O.U.A.T., ORISSA.
17	BRASALI	N.D.A.U.T, U.P.
18	MEGHNAD	N.D.A.U.T, U.P.
19	SILHAT	N.D.A.U.T, U.P.
20	SHYAMALA	I.G.K.V, RAIPUR
21	BPT-5204-SUB 1	N.R.R.I, CUTTACK
22	KISHORI	R.P.C.A.U., PUSA

Table 2: List of SSR markers and primer sequences.

S.NO	PRIMER	FORWARD SEQUENCE	REVERSE SEQUENCE	ANNEALING TEMPERATURE	PRODUCT SIZE(bp)	REPEAT MOTIFS
1.	RM566	CCCAACTACGATCAGCTCG	CTCCAGGAACACGCTCTTTC	55ºC	239	(AG)15
2.	RM520	AGGAGCAAGAAAAGTTCCCC	GCCAATGTGTGACGCAATAG	55ºC	247	(AG)10
3.	RM555	TTGGATCAGCCAAAGGAGAC	CAGCATTGTGGCATGGATAC	55ºC	223	(AG)11
4.	RM324	CTGATTCCACACACTTGTGC	GATTCCACGTCAGGATCTTC	55ºC	175	(CAT)21
5.	RM521	TTCCCTTATTCCTGCTCTCC	GGGATTTGCAGTGAGCTAGC	55ºC	260	(TC)14
6.	RM431	TCCTGCGAACTGAAGAGTTG	AGAGCAAAACCCTGGTTCAC	55ºC	251	(AG)16
7.	RM28166	TGCTTGCAAACATTGCTTCTGG	ACTGATGTACTGAACACGGGAAGG	55ºC	194	(CT)12
8.	RM5791	GAAGCAGAATACGCTTCGC	ACGTCACAAGGGTTCTTGC	55ºC	103	(AGC)8
9.	RM319	ATCAAGGTACCTAGACCACCAC	TCCTGGTGCAGCTATGTCTG	55ºC	134	(GT)10
10.	RM416	GGGAGTTAGGGTTTTGGAGC	TCCAGTTTCACACTGCTTCG	55ºC	114	(GA)9
11.	MRG 2894 IRRI 2894	TATGCTCTCTCCTTCAGGCC	CTTACCAACTCCGCACTTGC	55ºC	NA	NA
12.	MRG2805. HAU 2805	AGAGGAAGAAGCCAAGGAGG	CATCAACGTACCAACCATGG	55ºC	NA	NA
13.	RM321	CCAACACTGCCACTCTGTTC	GAGGATGGACACCTTGATCG	55ºC	200	(CAT)5
14.	RM286	GGCTTCATCTTTGGCGAC	CCGGATTCACGAGATAAACTC	55ºC	110	(GA)16
15.	RM70	GTGGACTTCATTTCAACTCG	GATGTATAAGATAGTCCC	55ºC	130	(ATT)33

Table3: Tabular characterization of 22 genotypes based on the codes of DUS guidelines.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
SANTOSH	1	1	3	3	1	1	1	1	1	1	1	3	3	1	3	1	0	1	1	1	1	1	1	1	9	7	1	1	0	0	0	1	1	5	3	1
KANAK	1	1	3	3	1	1	1	1	1	1	1	3	3	1	5	1	0	1	1	1	1	1	1	1	9	5	1	1	0	0	0	1	1	5	5	1
RADHA	1	1	3	3	1	1	1	1	1	1	1	3	3	1	3	1	0	1	1	1	1	1	1	1	9	7	1	1	0	0	0	1	1	3	5	1
PANKAJ	1	1	3	3	1	1	1	1	1	1	1	3	3	1	3	1	0	1	1	1	1	1	1	1	9	7	1	1	0	0	0	1	1	5	5	1
SATYAM	1	1	3	3	1	1	1	1	1	1	1	3	3	3	3	1	0	1	1	1	7	1	1	1	9	5	1	1	0	0	0	1	1	5	5	1
RAJSHREE	1	1	3	1	1	1	1	1	1	9	1	3	3	3	5	9	7	1	1	1	5	1	1	1	9	5	1	1	0	0	0	1	1	3	5	1
SWARNA	1	1	7	1	1	3	1	1	1	1	1	3	3	1	5	1	0	1	1	1	5	1	1	1	9	5	1	1	0	0	0	1	1	3	5	1

SUDHA	1	1	3	7	1	1	1	1	1	1	1	3	3	1	7	1	0	1	1	1	5	1	1	1	9	7	1	1	0	0	0	1	1	5	5	1
VAIDEHI	1	1	3	7	1	1	1	1	1	1	1	3	3	1	7	1	0	1	1	1	5	1	1	1	9	5	1	1	0	0	0	1	1	5	5	1
JANAKI	1	2	5	7	3	1	1	1	3	9	1	3	3	3	3	9	7	9	9	1	5	9	7	7	9	5	4	1	0	0	0	1	1	5	3	1
BAROGAR	1	1	3	1	1	1	1	1	1	1	1	3	3	1	7	1	0	1	1	1	5	1	1	1	9	5	1	1	0	0	0	1	1	3	3	1
MADHUKAR	1	1	5	7	1	1	1	1	1	9	1	3	3	3	5	9	7	1	1	1	7	7	1	1	9	5	1	1	0	0	0	1	1	5	5	1
SINGHARA	1	1	5	7	7	3	1	1	1	1	1	3	3	5	3	1	0	1	1	1	7	1	1	1	9	7	4	9	1	1	7	1	1	3	3	1
UJALADHUSARISA	1	2	3	1	1	1	1	1	3	9	1	3	3	1	5	1	0	1	9	1	7	7	1	1	9	5	1	1	0	0	0	1	1	5	5	1
SAGARSAMBA	1	1	5	1	1	1	1	1	1	1	1	3	3	3	5	1	0	1	1	1	5	1	1	1	9	7	1	1	0	0	0	1	1	5	3	1
JAGANNATHBALLAVA	1	1	5	1	1	1	1	1	1	1	1	3	3	5	5	1	0	1	1	1	1	1	1	1	9	7	1	1	0	0	0	1	1	3	3	1
BRASALI	1	1	7	1	1	1	1	1	1	1	1	3	3	1	5	1	0	1	1	1	9	1	1	1	9	7	1	1	0	0	0	1	1	5	3	1
MEGHNAD	1	2	5	7	3	1	1	1	3	9	1	3	3	3	3	9	7	9	9	1	5	9	7	1	9	7	4	1	0	0	0	1	1	5	5	1
SILHAT	1	1	3	1	1	1	1	1	1	1	1	3	3	1	5	1	0	1	1	1	7	1	1	1	9	5	1	1	0	0	0	1	1	5	3	1
SHYAMALA	1	4	3	1	1	1	1	1	3	1	1	3	3	3	5	9	0	1	9	1	9	9	1	1	9	5	1	1	0	0	0	1	9	5	3	1
BPT-5204-SUB1	1	1	7	1	1	1	1	1	1	1	1	3	3	1	7	1	0	1	1	1	1	1	1	1	9	7	1	1	0	0	0	1	1	3	1	1
KISHORI	1	1	7	1	1	1	1	1	1	1	1	3	3	1	5	1	0	1	1	1	9	1	1	1	9	5	1	1	0	0	0	1	1	3	3	1

Table 4: Classification of Genotypes into Distinct Phenotypic classes of 36 DUS based descriptors.

SL No.	Characteristics	States	Code	State	Genotypes	frequency
1.	Coleoptile color	Colorless	1	Colorless	All	100
		Green	2			
		Purple	3			
2	Basal Leaf sheath color	Green	1	Green	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Madhukar Barogar, Singhara, Sagar Samba, Jagannath Ballava, BPT-5204-Sub-1, Kishori, Silhat, Brasali	81.81
		Light purple	2	Light Purple	Janki, Ujala Dhusaria, Meghnad	13.6
		Purple lines	3	Uniform purple	Shyamala	4.5
		Uniform purple	4			
3	Leaf: intensity of green colour	Light	3	Light	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Sudha, Vaidehi , Ujala Dhusaria, Barogar, Silhat	50
		Medium	5	Medium	Madhukar, Janki, Singhara , Sagar Samba, Jagannath Ballava , Meghnad	27.27
		Dark	7	Dark	Swarna, Kishori, BPT-5204-Sub-1, Brasali	18.18
4	Presence of pubescence on leaf blade surface:	Absent	1	Weak	Rajshree, Swarna, Ujala Dhusaria, Sagar Samba, Kishori, Brasali	27.27
		Weak	3	Medium	Santosh, Kanak, Radha, Pankaj, Satyam, Jagannath Ballava, Shyamala, BPT-5204-Sub-1, Barogar, Silhat	45.45
		Medium	5	Strong	Sudha, Vaidehi, Janki, Madhukar, Singhara, Meghnad	27.27
		Strong	7			
		Very Strong	9			
5.	Culm attitude	Erect	1	Erect	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Ujala Dhusaria , Sagar Samba, Jagannath Ballava, Barogar, Silhat, Madhukar, Brasali, BPT-5204-Sub-1, Kishori, Shyamala	86.36
		Semi-erect	3	Semi erect	Janki, meghnad	9.09
		Open	5			
		Spreading	7	Spreading	Singhara	4.5
6	Attitude of flag leaf (early observation)	Erect	1	Erect	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Sudha, Vaidehi, Janki,	91

					Madhukar, Ujala Dhusaria, Sagar Samba, Jagannath Ballava, Meghnad, Kishori, BPT-5204-Sub-1, Brasali, Shyamala, Silhat, Barogar	
		Semi erect	3	Semi erect	Swarna, Singhara	9.09
		Horizontal	5			
		Drooping	7			
7.	Presence of auricles on leaf	Absent	1	Present	All	100
		Present	9			
8.	Anthocyanin coloration of auricles	Colorless	1	Colorless	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Sudha, Vaidehi, Madhukar, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Brasali, Barogar, Silhat, Singhara, Swarna	81.18
		Light purple	2			
		Purple	3	Purple	Ujala Dhusaria, Meghnad, Janki, Shyamala	18.18
9.	Presence of collar on leaves	Absent	1	Present	All	100
		Present	9			
10.	Anthocyanin coloration of collar:	Absent	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori	77.27
		Present	9	Present	Rajshree, Madhukar, Janaki, Meghnad, Ujala Dhusaria	22.27
11.	Presence of ligule on leaf:	Absent	1	Present	All	100
		Present	9			
12.	Shape of ligule:	Truncate	1	Split	All	100
		Acute	2			
		Split	3			
13.	Length of leaf blade	Short (<30 cm)	3	Long	All	100
		Medium (30-45 cm)	5			
		Long (>45 cm)	7			
14.	Attitude of flag leaf (Late observation)	Erect	1	Erect	Santosh, Pankaj, Vaidehi, Barogar, Silhat, Kishori, Shayamla, Kanak, Sudha, BPT-5402-Sub-1, Ujala dhusaria , Brasali	54.54
		Semi-erect	3	Semi erect	Janaki, Madhukar, Meghnad, Satyam, Sagar Samba, Rajshree, Shyamala , kanak	36.36
		Horizontal	5	Horizontal	Jagannath Ballava, Singhara	9.09
		Deflexed	7			
15.	Leaf senescence:	Early	3	Early	Santosh, Janaki, Meghnad, Satyam, Pankaj, Radha, Singhara	31.81
		Medium	5	Medium	Madhukar, Ujala Dhusaria , Jagannath Ballava, Brasali.	50

					Silhat, Kishori, Shyamala, Kanak, Swarna, Sagar Samba, Rajshree	
		Late	7	Late	Barogar, Sudha, Vaidehi, BPT-5402-Sub-1	18.18
16.	Anthocyanin coloration of nodes	Absent	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba,Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori	77.27
		Present	9	Present	Rajshree, Madhukar, Meghnad, Janaki, Shyamala	22.7
17.	Intensity of anthocyanin coloration of nodes	Weak	3	Strong	Rajshree, Madhukar, Meghnad, Janaki, Shyamala	22.7
		Medium	5	Strong		
		Strong	7	Strong		
18.	Anthocyanin coloration of internodes:	Absent	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori, Shyamala	90.90
		Present	9	Present	Janki, Meghnad	9.09
19.	Color of stigma:	White	1	White	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori	81.81
		Light green	2	Purple	Janki, Ujala Dhusaria, Meghnad, Shyamala	18.18
		Yellow	3			
		Light purple	4			
		Purple	5			
20.	Curvature of main axis of panicle:	Straight	1	Straight	Janki, Barogar, Madhukar, Singhra Meghnad, BPT-5204-Sub-1, Rajshree	31.8
		Semi-straight	3	Semi-straight	Santosh, Ujala Dhusaria ,Silhat, Shyamala, Sagar Samba	22.7
		Deflexed	5	Deflexed	Brasali, Radha	9.09
		Dropping	7	Dropping	Kanak Pankaj, Satyam, Kishori, Swarna, Sudha, Jagannath Ballava, Vaidehi	36.36
21	Density of pubescence of lemma	Absent	1	Weak	Jagannath Ballava, Santosh, Kanak, Radha, BPT-5204-Sub-1, Pankaj	27.27
		Weak	3	Medium	Janki, Meghnad, Swarna, Vaidehi, Sudha, Rajshree, Sagar Samba, Barogar	36.36

		Medium	5	Strong	Madhukar, Singhara, Ujala Dhusaria, Silhat, Satyam	22.72
		Strong	7	Very strong	Shyamala, Kishori, Brasali	13.6
		Very strong	9			
22	Anthocyanin coloration on tip of lemma: ,	White	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori,	77.27
		Yellow	3	Red	Madhukar, Ujala Dhusaria	9.09
		Brown	5	Purple	Janaki, Meghnad, Shyamala	13.6
		Red	7			
		Purple	9			
		Black	11			
23.	Anthocyanin coloration of area below apex of lemma:	Absent	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Singhara, Sagar Samba, Jagannath Ballava, Kishori, BPT-5204-Sub-1, Silhat, Brasali, Kishori, Madhukar, Ujala Dhusaria,	90.90
		Weak	3	Strong	Janki, Meghnad	9.09
		Medium	5			
		Strong	7			
		Very strong	9			
24.	Anthocyanin coloration of keel	Absent	1	Absent	Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Madhukar, Singhara, Ujala Dhusaria, Sagar Samba, Jagannath Ballava, Brasali Meghnad , silhat, Shyamala, BPT-5204-Sub-1, Kishori	95.45
		Weak	3	Strong	Janaki	4.5
		Medium	5			
		Strong	7			
		Very strong	9			
25. .	Presence of secondary branching on panicles	Absent	1	Present	All	100
		Present	9			
26.	Exertion of panicle	Partly exserted	3	Mostly exserted	Barogar, Madhukar, Ujala Dhusaria, Silhat, Kishori, Kanak Shyamala, Swarna, Satyam, Janaki, Rajshree, Vaidehi	54.54
		Mostly exserted	5			
		Well exserted	7	Well exerted	Jagannath Ballava, Brasali, Meghnad, Singhara, Pankaj, Radha, Santosh, Sudha, Sagar Samba, BPT-5204-Sub-1	45.45
27.		Straw	1	Straw		81.81

	Color of lemma and palea	Gold and gold	2		Santosh, Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Vaidehi, Barogar, Ujala Dhusaria, Sagar Samba, Jagannath Ballava, Brasali , Silhat, Shyamala, BPT-5204-Sub-1, Kishori	
		Brown spots on straw furrows on straw background	3			
		Brown furrows on straw	4	Brown spot on straw	Janaki, Singhara, Meghnad	13.6
		Brown (tawny)	5			
		Reddish to light purple	6	Reddish to light	Madhukar	4.5
		Purple spots / furrows on straw	7			
		Purple	8			
		Black	9			
28.	Awns	Absent	1	Absent	21 genotypes	95.45
		Present	9	Present	Singhara	4.5
29	Color of awns	White	1	White	Singhara	4.5
		Yellowsh Brown	2			
		Brown	3			
		Reddish brown	4			
		Light red	5			
		Red	6			
30.	Distribution of awns	Tip only	1	Tip only	Singhara	4.5
		Upper half only	3			
		Whole length	5			
31.	Length of awns	Very short	1	Long	Singhara	4.5
		Short	3			
		Medium	5			
		Long	7			
		Very long	9			
32.	Color of sterile lemma:	Straw	1	Straw	All	100
		Gold	2			
		Red	3			
		Purple	4			
33.	Presence of anthocyanin in leaf:	Absent	1	Absent	21 genotypes	95.45
		Present	9	Present	Shyamala	4.5
34.	Grain length	Very short (<6.0 mm)	1	Short	Radha, Pankaj, Rajshree, Swarna, Barogar, Singhara, Jagannath Ballava, BPT-5204-SUB-1, Kishori	40.90
		Short (6.1-8.5 mm)	3			

		Medium (8.6-10.5 mm)	5	Medium	Santosh, Kanak, Satyam, Sudha, Vaidehi, Janaki, Madhukar, Ujala Dhusaria, Sagar Samba, Brasali, Meghnad, Silhat, Shyamala.	59.09
		Long (10.6-12.5 mm)	7			
		Very long (>12.5 mm)	9			
35	Grain width	Very narrow (<2mm)	1	Very narrow (<2mm)	BPT-5204-SUB-1	4.5
		Narrow (2.1-2.5mm)	3	Narrow (2.1-2.5mm)	Santosh, Vaidehi, Singhara, Sagar Samba, Jagannath Ballava, Brasali, Silhat, Shyamala, Kishori.	45.5
		Medium (2.6-3mm)	5	Medium (2.6-3mm)	Kanak, Radha, Pankaj, Satyam, Rajshree, Swarna, Sudha, Janaki, Madhukar, Ujala Dhusaria, Meghnad,	50
		Broad (3.1-3.5mm)	7			
		Very broad (>3.5)	9			
36	Decorticated grain aroma	Absent	1	Absent	All	100
		Present	9			

Table 5 : Representative DUS based images taken during characterization

			
Colorless coleoptile color	Light purple basal leaf sheath colour.	Purple basal leaf sheath color	Spreading culm attitude
			
Semi erect flag leaf(early observation)	Presence of auricles	Purple anthocyanin coloration of auricles	Presence of collar on leaves

			
Purple anthocyanin coloration of collar	Presence of ligule and split ligule	Horizontal flag leaf (Late observation)	Early leaf senescence
			
Anthocyanin coloration of nodes	Anthocyanin coloration of internodes	Purple stigma	Deflexed curvature of main axis.
			
Purple anthocyanin pigmentation of Tip of Lemma and strong a**nthocyanin coloration of area below apex of lemma:**	Red anthocyanin pigmentation of Tip of Lemma	Presence of secondary branches	Mostly exerted panicle
			
Reddish to light color of lemma and palea	Presence of awns	Anthocyanin coloration of leaf	Medium width grain size.

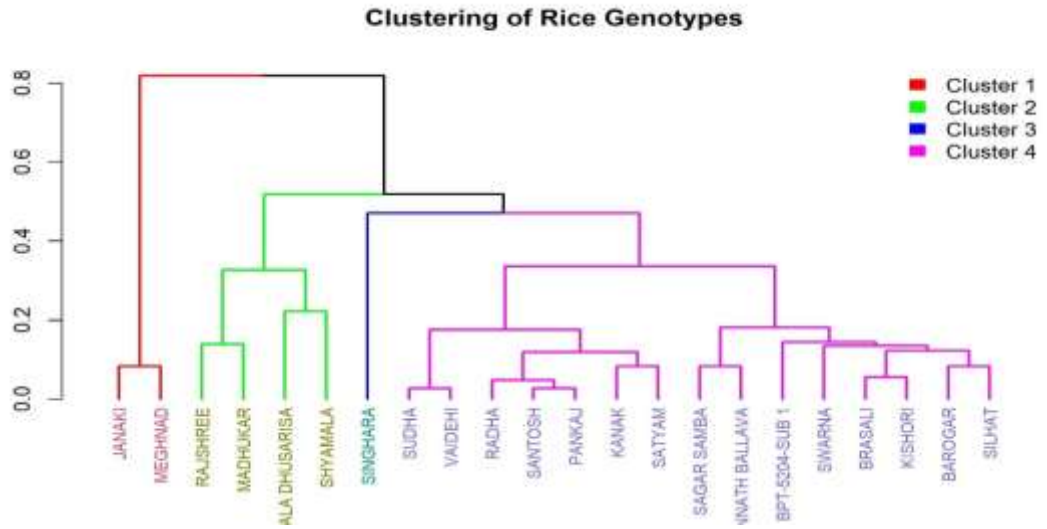

Figure 2: Hierarchical cluster dendrogram of 22 rice genotypes based on 36 qualitative morphological characters

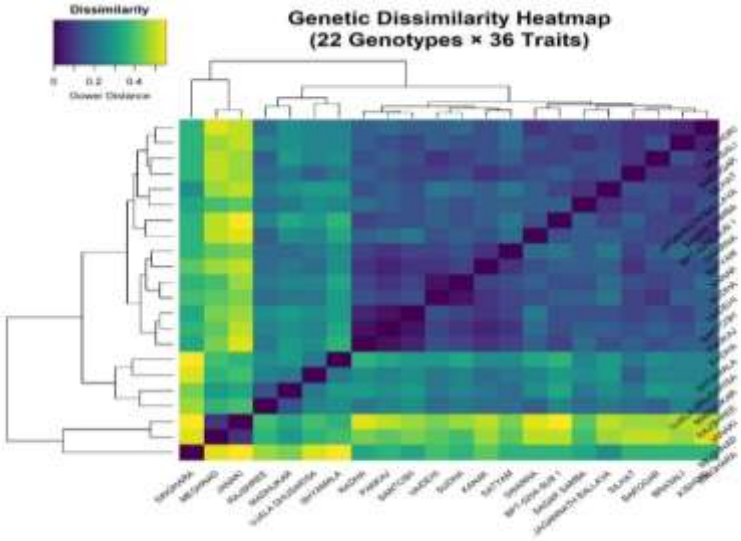

Figure 3: Heat map showing Genetic Dissimilarity

GP1=SANTOSH, GP2=KANAK, GP3=RADHA, GP4=PANKAJ, GP5=SATYAM, GP6=RAJSHREE, GP7=SWARNA, GP8=SUDHA, GP9=VAIDEHI, GP10=JANAKI, GP11=BAROGAR, GP12=MADHUKAR, GP13=SINGHARA, GP14= UJALA DHUSARIA, GP15= SAGAR SAMBA, GP16= JAGANNATH BALLAVA, GP17= BRASALI, GP18= MEGHNAD, GP19= SILHAT, GP20= SHYAMALA, GP21= BPT-5204-SUB-1, GP22= KISHORI

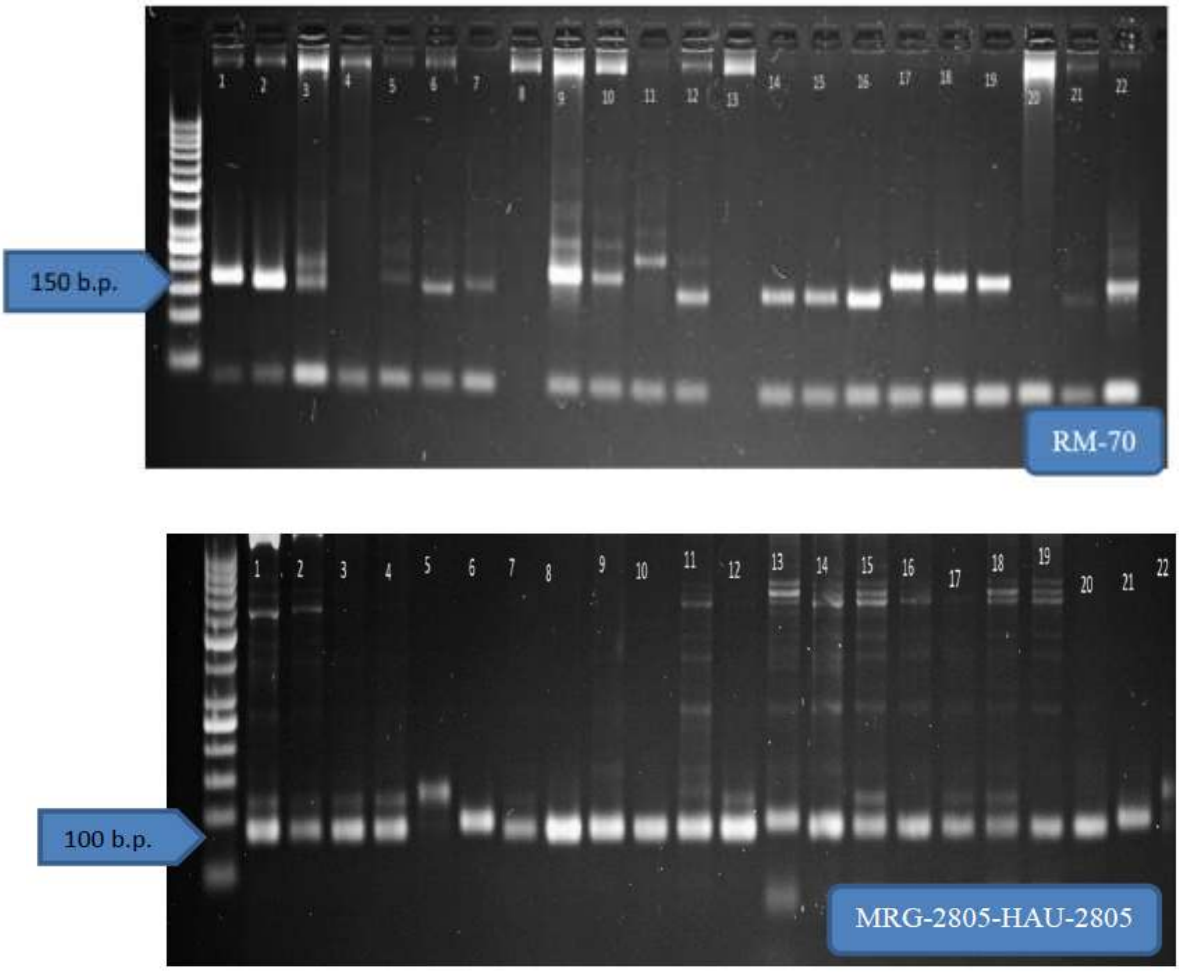

Figure 4: Representative gel image of PCR amplification using SRR markers

Table 6: Genetic Diversity Statistic based on SSR marker data

Pop	No of genotypes	MLG	eMLG	SE	H	G	lambda	E.5	Hexp	Ia	rbarD
	22	22	22	0	3.091042	22	0.954545	1	0.298454	1.005011	0.030474

Table 7: Table showing total alleles and unique alleles

Genotype	Total Alleles	Unique Alleles
SANTOSH	16	0
KANAK	14	0
RADHA	15	1
PANKAJ	15	0
SATYAM	16	1
RAJSHREE	16	0
SWARNA	15	0
SUDHA	14	0
VAIDEHI	16	0
JANAKI	15	0
BAROGAR	16	2
MADHUKAR	14	0
SINGHARA	14	0
UJALA DHUSARISA	15	0
SAGAR SAMBA	15	0
JAGANNATH BALLAVA	16	0
BRASALI	16	0
MEGHNAD	15	0
SILHAT	16	0
SHYAMALA	14	0
BPT-5204-SUB 1	15	0
KISHORI	15	1

Table 8: Table Showing locus wise allelic frequency, PIC value and number of genotypes

Locus	Allele Frequency	PIC	No. of Genotypes
RM 566_m1	0.136	0.954	3
RM 566_m2	0.727	0.392	16
RM 566_m3	0.091	0.978	2
RM 520_m1	0.091	0.978	2
RM 520_m2	0.864	0.226	19
RM 520_m3	0.091	0.978	2
RM 319_m1	0.364	0.761	8
RM 319_m2	0.636	0.488	14
RM 416_m1	0.409	0.716	9
RM 416_m2	0.591	0.534	13
MRG 2894 IRRI2894 m1	1.000	0.000	22
RM-324_m1	0.182	0.923	4
RM-324_m2	0.500	0.625	11
RM-324_m3	0.364	0.761	8
RM 521_m1	0.091	0.978	2
RM 521_m2	0.955	0.085	21
RM 521_m3	0.091	0.978	2
MRG-2805-HAU-2805_m1	0.955	0.085	21
MRG-2805-HAU-2805_m2	0.045	0.994	1
MRG-2805-HAU-2805_m3	0.136	0.954	3
RM431_m1	0.091	0.978	2

RM431_m2	0.909	0.160	20
RM70_m1	0.045	0.994	1
RM70_m2	0.273	0.847	6
RM70_m3	0.091	0.978	2
RM70_m4	0.182	0.923	4
RM321_m1	1.000	0.000	22
RM28166_m1	0.455	0.670	10
RM28166_m2	0.591	0.534	13
RM28166_m3	0.182	0.923	4
RM5791_m1	0.273	0.847	6
RM5791_m2	0.727	0.392	16
RM-555_m1	0.091	0.978	2
RM-555_m2	0.864	0.226	19
RM-555_m3	0.045	0.994	1
RM286_m1	0.727	0.392	16
RM286_m2	0.273	0.847	6

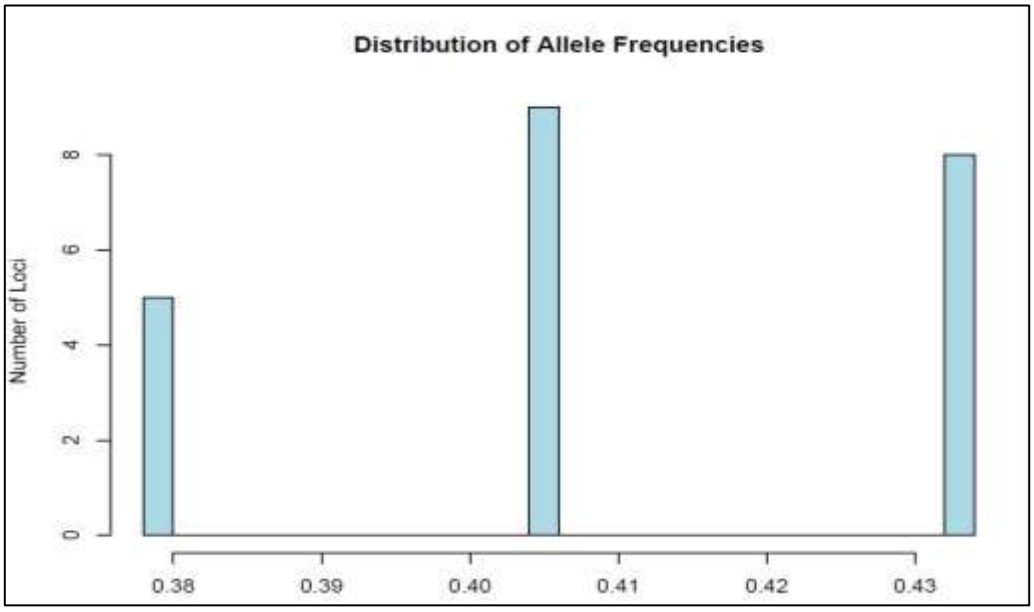

Figure 5: Graphical representation of distribution of allelic frequency based on SSR marker analysis.

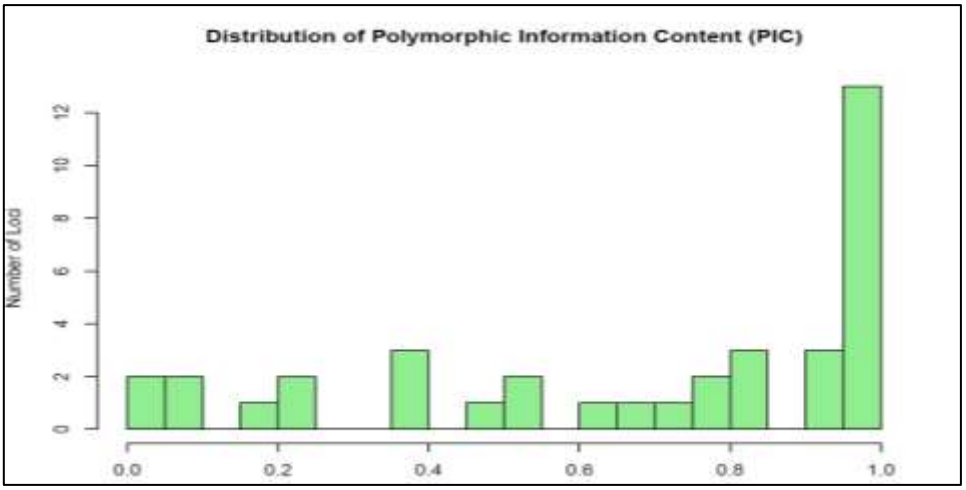

Figure 6: Graphical representation of distribution of PIC value.

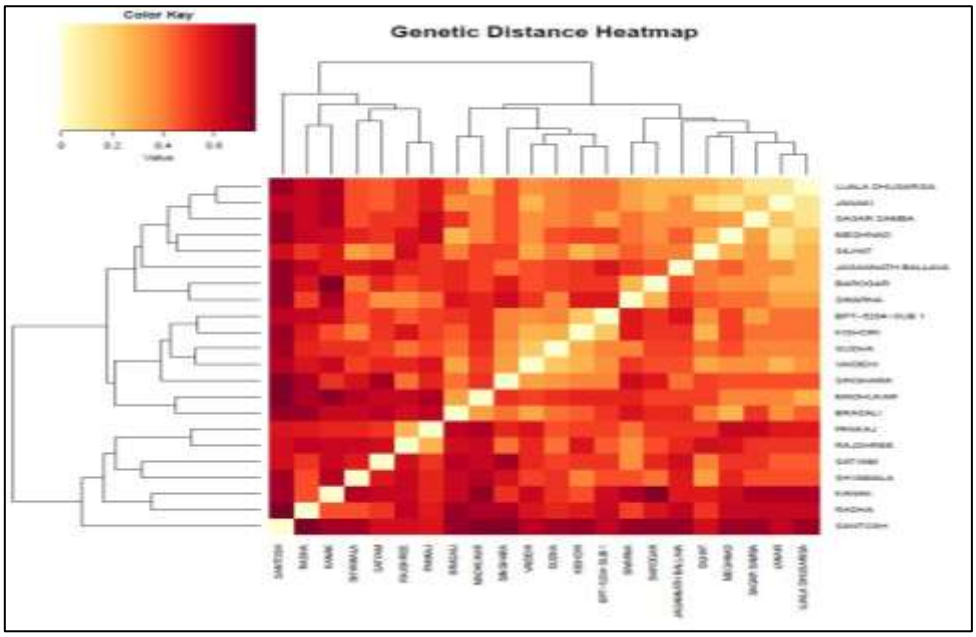

Figure 7: Heat map showing Genetic Distance among 22 genotypes based on SSR marker analysis.

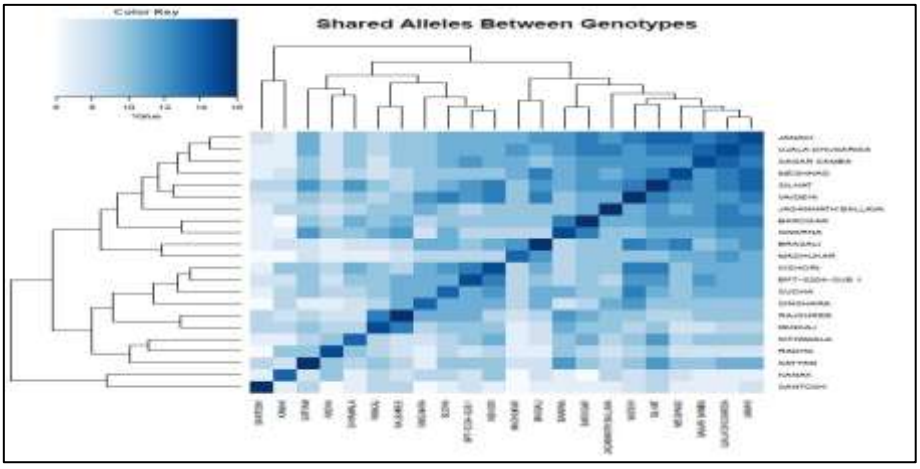

Figure 8: heat map showing shared alleles among 22 genotypes

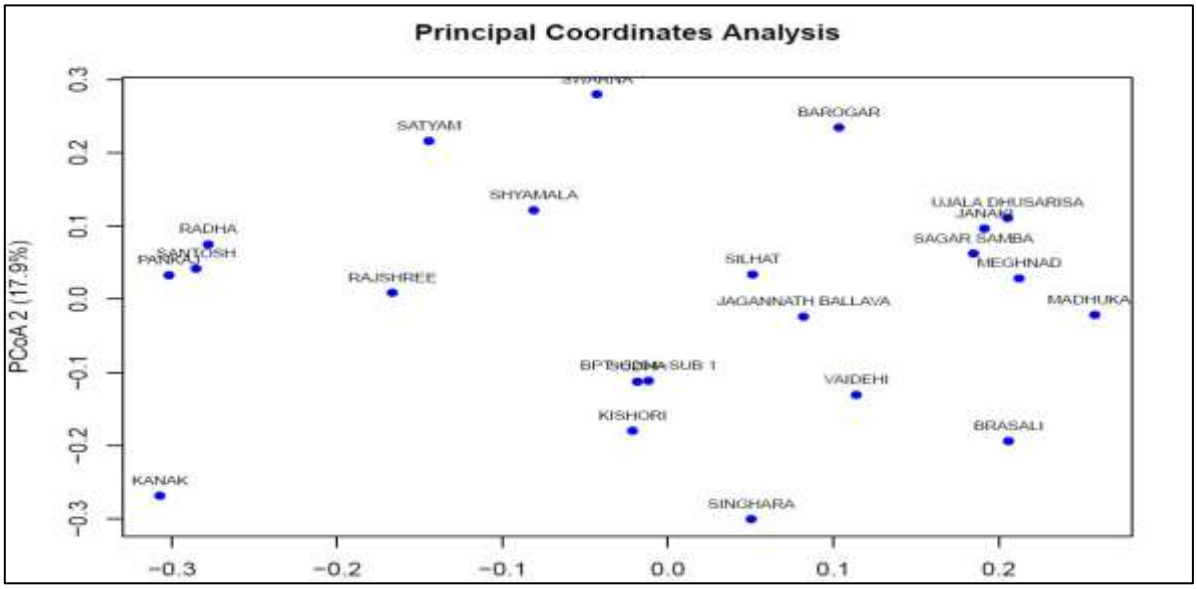

Figure 8: Scatter plot of Different Genotypes based Principal Coordinates analysis.

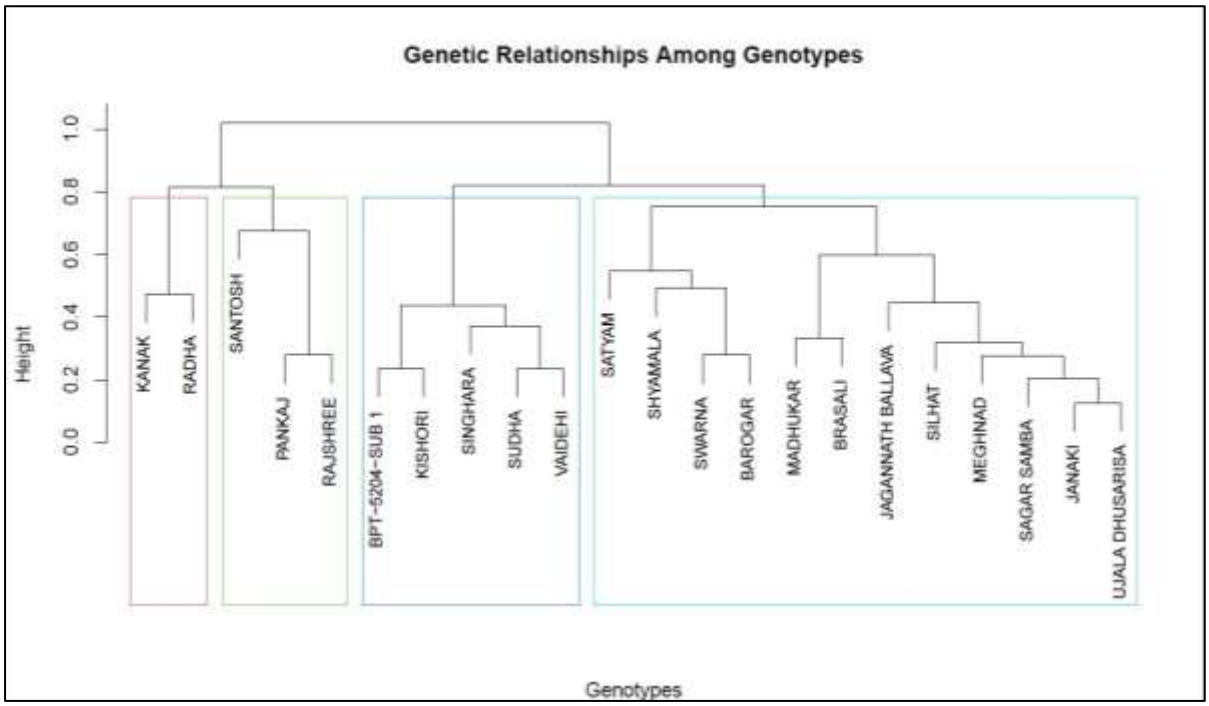

Figure 9: Dendrogram showing clustering and genetic relationship of 22 genotypes