https://doi.org/10.17221/133/2025-CJGPB

Genetic diversity of selected Malaysian rice accessions using microsatellite markers

Shahril Ab Razak[1], Alny Marlynni Abd Majid[2], Rahiniza Kamaruzaman[1], Norliza Abu Bakar[2], Rabiatul Adawiah Zainal Abidin[2], Yun Shin Sew[2], Norfarhan Mohd-Assaad[3,4], Asmuni Mohd Ikmal[1], Noraziyah Abd Aziz Shamsudin[1]*

[1]*Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia*
[2]*Agri-omic and Bioinformatic Programme, Biotechnology & Nanotechnology Research Centre, MARDI Headquarters, Serdang, Malaysia*
[3]*Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia*
[4]*Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, Bangi, Malaysia*

**Corresponding author: shahrilf@mardi.gov.my*

Citation: Ab Razak S., Abd Majid A.M., Kamaruzaman R., Abu Bakar N., Zainal Abidin R.A., Sew Y.S., Mohd-Assaad N., Mohd Ikmal A., Abd Aziz Shamsudin N. (2026): Genetic diversity of selected Malaysian rice accessions using microsatellite markers. Czech J. Genet. Plant Breed., 62: 146–156.

Abstract: Genetic diversity of plant genetic resources provides the foundation for breeding programmes aimed at developing high-yielding rice varieties with tolerance to biotic and abiotic stresses. Given the abundance of available genetic resources, efficient approaches for their characterisation are essential. In this study, 182 Malaysian rice accessions representing different maturity groups were characterised using 20 polymorphic simple sequence repeat (SSR) markers. The analysis identified 183 alleles, ranging from two (RM507) to 22 (RM154), with an average of 9.15 alleles per locus. Observed and expected heterozygosity ranged from 0.000 to 0.506 and 0.319 to 0.864, respectively. Polymorphism information content (PIC) values ranged from 0.2744 (RM495) to 0.8475 (RM154), with an average of 0.6216 per locus. Unweighted pair group method with arithmetic mean (UPGMA) analysis revealed two major groups. In general, the accessions were clustered according to their adaptive ecosystem type, with most lowland varieties, including lowland breeding lines (92.4%), assigned to Group I, whereas most upland varieties (86.7%) belonged to Group II. This grouping pattern was supported by STRUCTURE analysis, which identified $K = 2$ as the optimal number of clusters, indicating that the studied accessions were structured into two major genetic groups. Principal coordinate analysis (PCoA) further supported this grouping pattern, with the first three axes explaining 39.63% of the total variation. Analysis of molecular variance (AMOVA) showed that 31% of the total variation occurred among populations, 63% among accessions, and 6% within accessions. The results also indicated the possible presence of duplicate accessions within the collection. This study provides valuable insights for future breeding programmes aimed at developing high-yielding rice varieties with a broad genetic base and supports the effective management and conservation of rice genetic resources.

Keywords: genetic variability; population structure; SSR marker; traditional rice

Supported by the Government of Malaysia under RMK12 (PRB502) and KRB-307.

https://doi.org/10.17221/133/2025-CJGPB

Rice (*Oryza sativa* L.) is the most significant staple crop in the world. It is crucial to develop superior rice varieties with the desired traits (high quality, tolerance to biotic and abiotic stresses) to meet the demand of the growing global population. Since improving rice cultivars in terms of productivity and quality remains essential, the breeding programme for crop improvement should focus on widening the genetic sources of breeding materials (Mudhale et al. 2024). Notably, only a meagre fraction of rice germplasm collection has been used for breeding initiatives despite the vast landrace rice genetic resources. This shortcoming has limited the genetic base for breeding lines and improving rice varieties. Apart from its pathogen and environmental equilibrium, landrace rice has been proven to be genetically dynamic (Beena et al. 2025). However, several high-yield varieties have substituted and replaced landrace rice to cater to the increasing food demands. The genetic variability found within landraces provides genetic flexibility, which leads to the potential ability to adapt to local field conditions, and they can adapt to changing environments and farming practices. Besides, landrace rice is also a perfect candidate for rice improvement due to its high genetic variations to withstand biotic and abiotic stresses, although its lower yield productivity is a setback compared to commercial high-yield rice varieties (Hanamaratti et al. 2008). Therefore, it is crucial to develop high-yield and superior quality rice varieties while at the same time conserving the landrace rice to ensure the sustenance of the agroecological functions.

Rice varieties with a wide genetic base would be robust in terms of their resistance to pests and diseases, increased yield under diverse agro-climatic conditions, and capacity to boost plant performance (Zhu et al. 2000). Therefore, information regarding genetic diversity is essential for breeders in order for them to develop efficient breeding programmes that can determine the level of heterosis and genetic gain. Determining genetic variation is an integral step for plant varietal protection, cultivar identification and characterization, and effective seed purity management in order to facilitate breeding and conservation efforts. According to Rocha et al. (2020), molecular markers, pedigree, and morphological characterisation analyses can be deployed to identify genetic diversity. Nonetheless, morphological and pedigree analysis have been proven unreliable and impractical to measure genetic diversity in crops (Fufa et al. 2005). Diversity data derived from morphological characters are non-viable and inaccurate due to highly influential agronomic practices and environmental settings (Chesnokov et al. 2020).

DNA markers are currently the method of choice for determining genetic diversity because of the following reasons: they are not influenced by agronomic practices or climate conditions; prior lineage data is not required; they have a high level of polymorphism; and they are both rich and plentiful in the genome (Venkateshwarlu et al. 2025). Notably, highly sensitive microsatellite markers have been widely applied in a variety of plant genomic studies due to their transferability across genotypes, requirement of simple analytical and technical processes, highly abundant presence in the genome, high polymorphisms and multi-allelic characteristic (Vieira et al. 2016). Although single-nucleotide polymorphism (SNP)-based platforms and genome-wide datasets are increasingly used for diversity analysis, simple sequence repeat (SSR) markers remain relevant for germplasm characterisation because they are highly informative per locus, cost-effective for moderate sample sizes, and widely comparable with earlier studies and reference datasets (Zurn et al. 2020). In addition, SSR-based diversity analysis can complement SNP-based studies by providing a practical baseline for collections that are not yet fully characterised using genome-wide data, as both marker systems have been shown to reveal broadly consistent patterns of relationship and diversity in crop germplasm studies (Singh et al. 2018). In Malaysia, molecular diversity studies on rice have been reported, including SNP-based evaluation of released Malaysian rice varieties (Ab Razak et al. 2020). However, broader characterisation of Malaysian rice germplasm encompassing landraces, breeding materials, and accessions from contrasting adaptive ecosystems remains limited. This is particularly important because Malaysian rice germplasm includes both upland and lowland materials with potentially distinct genetic backgrounds and breeding value. Therefore, this study aimed to assess the diversity of selected Malaysian rice accessions representing contrasting adaptive ecosystems using a set of microsatellite markers. The findings of this study may provide useful direction for future conservation management programmes as well as for breeders seeking to strategize their breeding operations with the goal of predicting genetic gain.

https://doi.org/10.17221/133/2025-CJGPB

MATERIAL AND METHODS

Plant material and genomic DNA isolation. A total of 182 rice accessions consisting of 162 landraces and 20 improved varieties were evaluated in the present study. Among these accessions, seven originated outside Malaysia, whereas the remaining accessions originated from Malaysia, including Peninsular Malaysia, Sabah, Sarawak, and breeding materials developed locally. The accessions were selected with a maturity period of less than 140 days to minimise the inclusion of strongly photoperiod-sensitive materials, as this population will be used in future association analysis. For ecological interpretation, the accessions were first classified into upland and lowland rice based on passport and varietal background records. The lowland group comprised both traditional lowland landraces and lowland breeding lines. The seeds of these rice accessions were obtained from the Malaysian Agricultural Research and Development Institute (MARDI) Rice Genebank, MARDI Seberang Perai, which is located in the northern region of Peninsular Malaysia. All the rice accessions were propagated and multiplied at MARDI Headquarters, Selangor, Malaysia. The details of the accessions are summarised in Table S1 in the Electronic Supplementary Material (ESM). Small pieces of the young leaf tissue were placed together with stainless steel beads into a 96-well plate and were frozen at –80 °C overnight. The extraction buffer (100 mM Tris-HCl [pH8.0], 20 mM ethylenediaminetetraacetic acid (EDTA), 1.4 M NaCl, 2% cetyltrimethylammonium bromide (CTAB) and 0.05% β-mercaptoethanol) was placed in each well containing the frozen tissues before being ground using Tissue Lyser (Qiagen, Germany). The genomic DNA was extracted using the protocol described by Mace et al. (2003) with some modifications in terms of incubation time. The integrity and quantity of the extracted DNA were measured on 0.8% agarose gel and Fluoraskan Ascent (Thermo Fisher Scientific, United States), respectively.

Microsatellite amplification and genotyping. A total of 24 microsatellite markers that covered the 12 rice chromosomes (Table 1) were selected from the GRAMENE database. The polymerase chain reaction (PCR) was conducted as described by Schuelke (2000) using the M13-tailed primer method. The M13 sequence tail (TGTAAAACGACGGCCAGT) was added to the 5' end of the forward primer, while the reverse primer remained sequence-specific. A universal M13 primer labelled with VIC, FAM, PET, or NED was included in the reaction. During PCR amplification, the fluorescently labelled universal M13 primer annealed to the M13 tail and generated fluorescently labelled PCR products. The PCR master mix was prepared to a final volume of 10 µL, containing 1× PCR buffer, 10 µM each of tailed-forward and reverse primers, 5 µM fluorescently labelled universal M13 primer, 2 mM dNTP (Invitrogen, United States), and 1 U of *Taq* polymerase (Invitrogen, United States). The amplification was done using GeneAmp® PCR System 9700 (Applied Biosystems, United States) with the following parameters: Pre-denaturation at 94 °C for 5 min, followed by 34 cycles of denaturation at 94 °C for 1 min, primer annealing at 41–65 °C for 1 min (Table 1), and extension at 72 °C for 1 min, followed by final extension at 72 °C for 5 min. Then, the amplified PCR products were multiplexed with up to four primer pairs per set using a combination of four different fluorescent dyes (VIC, FAM, PET and NED) in one multiplex group. Then, the multiplexed PCR products were resolved using ABI3730xL Genetic Analyser (Applied Biosystems, United States), and GeneScan 500 LIZ (Applied Biosystems, United States) was used as a standard molecular weight ladder. A blank well (H12) was included in each 96-well plate as a negative control for genotyping quality control.

Data analysis. The allele of each sample was scored using GeneMapper Version 5 (Thermo Fisher Scientific, United States) using automatic binning for allele size calling, followed by manual inspection of each sample to ensure accurate scoring. The electropherograms of alleles were scored as suggested by Arif et al. (2010). Markers with call rates below 90% across all samples or those producing multiple banding patterns were excluded from analysis. In addition, MICRO-CHECKER (Van Oosterhout et al. 2004) was used to examine the presence of null alleles and stutter bands. Then, the data was sorted and formatted in Microsoft Excel before being used in PowerMarker (Liu & Muse 2005), STRUCTURE (Pritchard et al. 2000), PopGene32 (Yeh et al. 2000) and GenAlEx (Peakall & Smouse 2006). The number of effective alleles for each locus was calculated using PopGene32. PowerMarker was used to calculate the number of alleles, major allele frequency, polymorphism information content (PIC) and gene diversity of each SSR marker (Liu & Muse 2005). The same software was also used to generate the shared-allele genetic distance to describe the genetic relatedness among all varieties. Shared-allele genetic distance values were interpreted as indicators of relative ge-

https://doi.org/10.17221/133/2025-CJGPB

netic divergence among accessions, where higher values represented greater genetic dissimilarity. These calculated genetic distance matrices were provided as an input to generate the dendrogram based on the unweighted pair group method with arithmetic mean (UPGMA) using MEGA software (Ver. 7) (Kumar et al. 2016). Bootstrap support for the dendrogram was estimated using 1 000 replicates. GenAlEx (Ver. 6.502) was used to calculate the analysis of molecular variance (AMOVA), pairwise Fst values among the ecosystem-based groups, and to visualise the principal coordinate analysis (PCoA). The significance of AMOVA and pairwise Fst values was tested using 999 permutations. The STRUCTURE (Ver. 2.3) (Pritchard et al. 2000) was used to designate the varieties to their respective structure group in the absence of knowledge of their population affinity using Bayesian clustering. An admixture model with correlated allele frequencies was applied. The number of groups (K) was assessed from one to eight, with 15 independent runs per K value, with a run length of 50 000 Markov chain Monte Carlo (MCMC) repetitions and 100 000 generations as burn-in period. The optimal value of K was calculated using STRUCTURE SELECTOR based on both LnP(K) and ΔK statistics (Li & Liu 2018). The highest ΔK value was used to infer the major population structure. Cluster assignment patterns were examined across the 15 independent runs for each K value, and the best run at the optimal K was selected for visualisation. Individuals with membership coefficients (Q values) of ≥ 0.80 were assigned to a specific genetic cluster, whereas those with Q values of < 0.80 were considered admixed.

RESULTS AND DISCUSSION

SSR marker characterisation. Genotyping of 24 SSR markers across 182 rice accessions showed that only 20 SSRs were suitable for further analyses, while four SSRs were excluded because they either failed to amplify across 90% of the total samples

Table 1. Characterisation of 20 microsatellite markers using 182 rice accessions

Marker	Chr	Major allele frequency	Allele No.	Gene diversity	PIC	Effective No. of alleles (N_e)	Observed heterozygosity (H_o)	Expected heterozygosity (H_e)
RM495	1	0.802	4.000	0.321	0.274	1.467	0.017	0.319
RM1	1	0.305	17.000	0.841	0.825	6.278	0.006	0.843
RM154	2	0.261	22.000	0.828	0.848	5.726	0.006	0.828
RM452	2	0.487	4.000	0.527	0.415	2.114	0.022	0.528
RM55	3	0.432	6.000	0.656	0.591	2.915	0.017	0.659
RM514	3	0.284	7.000	0.796	0.766	4.880	0.017	0.797
RM124	4	0.501	5.000	0.601	0.524	2.482	0.017	0.599
RM507	5	0.523	2.000	0.499	0.375	1.996	0.017	0.500
RM413	5	0.588	10.000	0.597	0.557	2.484	0.022	0.599
RM133	6	0.579	9.000	0.590	0.537	2.432	0.029	0.591
RM510	6	0.469	6.000	0.610	0.534	2.570	0.017	0.613
RM11	7	0.249	12.000	0.861	0.846	7.220	0.018	0.864
RM433	8	0.571	8.000	0.596	0.544	2.477	0.166	0.598
RM447	8	0.407	8.000	0.658	0.594	2.896	0.028	0.657
RM316	9	0.423	11.000	0.721	0.680	3.605	0.506	0.725
RM215	9	0.442	6.000	0.717	0.677	3.532	0.011	0.719
RM474	10	0.285	20.000	0.858	0.845	6.986	0.000	0.860
RM536	11	0.287	10.000	0.809	0.784	5.309	0.028	0.814
RM277	12	0.491	5.000	0.570	0.477	2.322	0.006	0.571
RM19	12	0.288	11.000	0.808	0.782	5.100	0.033	0.806
Mean		0.434	9.150	0.673	0.622	3.739	0.049	0.674

Chr – chromosome; Allele No. – number of alleles detected; PIC – polymorphism information content

https://doi.org/10.17221/133/2025-CJGPB

or produced multiple banding patterns. All chromosomes were represented by two SSR markers except chromosome 4, 7, 10 and 11, where each had only one representative SSR marker. Analysis of the 20 SSRs revealed 183 alleles, ranging from two (RM507) to twenty-two (RM154), with an average of 9.150 alleles per locus. The PIC values ranged from 0.274 (RM495) to 0.848 (RM154), with an average of 0.622 per locus. The major allele frequency ranged from 0.249 (RM11) to 0.802 (RM495), with an average of 0.434. Gene diversity values ranged from 0.321 (RM495) to 0.861 (RM11), with an average of 0.673 per locus. Details of the SSR marker characteristics are summarised in Table 1.

The importance of genetic variability in supporting plant breeding programmes and conservation management is well established (Salgotra & Chauhan 2023). Among DNA-based marker systems, SSRs remain widely applied due to their high polymorphism, multiallelic nature, co-dominant inheritance, reproducibility, and ease of detection by PCR. In rice, SSR resources are broadly accessible through databases such as Gramene and RAP, making them a practical and cost-effective platform for genetic diversity analysis.

In this study, the detection of 183 alleles with an average of 9.15 alleles per locus is higher than that reported for Indian rice varieties by Singh et al. (2016) but slightly lower than that reported for East and Southeast Asian rice by Dang et al. (2015), indicating that Malaysian rice germplasm possesses substantial genetic variability. The high PIC values support the discriminatory power of the selected markers, with 80% of SSRs being highly informative. Based on the classification proposed by Botstein et al. (1980), markers with PIC values greater than 0.50 are considered highly informative, indicating that the majority of SSR loci used in this study were effective for diversity analysis and varietal discrimination. Consistent with previous studies (Thomson et al. 2009; Comertpay et al. 2016), RM154 showed the highest PIC value, reflecting its hypervariable dinucleotide repeat motif. However, dinucleotide repeats can be more difficult to score due to a higher tendency to produce stutter peaks (Merritt et al. 2015). Dinucleotide repeats are also more prone to replication slippage, contributing to higher mutation rates and scoring complexity (Chakraborty et al. 1997; Villa et al. 2005). These technical constraints can limit SSR resolution for distinguishing very closely related genotypes. The low mean observed heterozygosity (0.049) compared with the expected heterozygosity (0.674) is consistent with the predominantly self-pollinating nature of rice. This pattern may also reflect repeated selfing during seed maintenance and cultivation of landraces, as well as line development in breeding materials, which together contribute to reduced heterozygosity and increased homozygosity across loci. Although chromosome 4 and chromosome 11 were each represented by only one SSR locus after quality filtering, the final marker set still provided coverage across all 12 rice chromosomes and was sufficiently informative to resolve the major diversity patterns among the studied accessions.

Genetic diversity and population structure of selected rice accessions. Genetic analysis of 182 rice accessions based on 20 polymorphic SSR markers revealed two major groups (Group I and Group II) in the UPGMA dendrogram (Figure 1). Accessions were generally grouped according to their adaptive ecosystem, with most lowland accessions, including lowland breeding lines (92.4%), clustered in Group I, whereas most upland varieties (86.7%) clustered in Group II. Pairwise genetic distance for each accession is provided in Table S2 in ESM. STRUCTURE analysis supported the dendrogram results. The ΔK method described by Evanno et al. (2005) showed a peak at $K = 2$ (Figure 2), indicating that the accessions clustered into two major genetic groups. The sharp peak at $K = 2$ (1 692.021) suggests the presence of two major genetic groups within the studied germplasm. The mean alpha value across the 15 runs at $K = 2$ was 0.044, indicating a low level of admixture among the studied accessions and supporting the presence of two relatively distinct genetic clusters. This concordance between UPGMA and STRUCTURE indicates that the major genetic structure of the studied rice accessions is primarily associated with the adaptive ecosystem, particularly the separation between lowland and upland rice. Because the lowland group comprised both traditional lowland landraces and lowland breeding lines, the clustering pattern at $K = 3$ was examined only as an additional visualisation of subdivision within the broader lowland group, rather than as the principal population structure (Figure 3). A dendrogram constructed based on three groups (lowland breeding lines, lowland landrace, and upland landrace) showed higher similarity between lowland breeding lines and lowland landrace compared to lowland breeding lines and upland landrace (Figure 4). Pairwise genetic distance values ranged from 0.4773 (lowland

https://doi.org/10.17221/133/2025-CJGPB

breeding lines ↔lowland landrace) to 0.7496 (lowland breeding lines ↔ upland landrace), indicating that lowland breeding lines were more similar to lowland landraces than to upland landraces (Table 2).

Ecological structuring of rice germplasm showed clear separation between upland and lowland gene pools. In comparison, Zhang et al. (2013) reported that ecological differentiation between upland and lowland landraces was pronounced only within the japonica subspecies, whereas indica accessions did not show such a clear upland–lowland differentiation. Dao et al. (2024) observed an intermediate pattern, where traditional lowland varieties formed exclusive clusters, while upland rainfed interspecific accessions (*O. sativa*, *O. glaberrima*, and *O. longistaminata*) clustered together with several lowland landraces, reflecting genetic admixture across ecosystems. Collectively, these patterns suggest that the strength of upland–lowland differentiation depends on subspecies composition, interspecific introgression, and farmer-mediated germplasm exchange. In the Malaysian context, this separation is likely to reflect long-term adaptation to contrasting agro-ecological conditions, where upland rice is typically exposed to rainfed, drier, and more variable environments, whereas lowland rice is cultivated under wetter and more controlled paddy conditions. Such ecological contrast may have contributed to divergence in adaptive traits and selec-

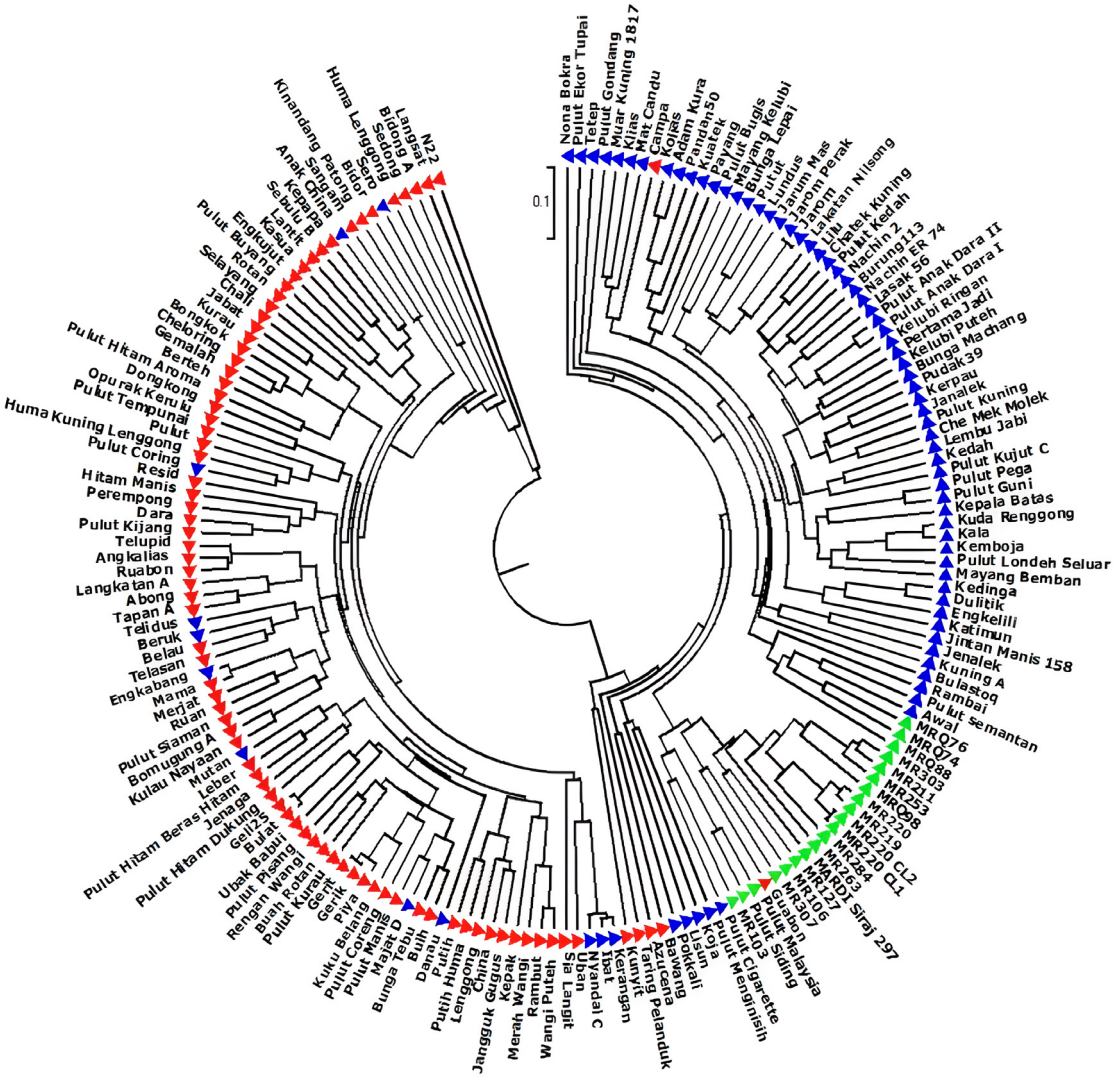

Figure 1. The unweighted pair group method with arithmetic mean (UPGMA) dendrogram of 182 rice accessions based on 20 microsatellite markers

Red indicates upland landraces, blue indicates lowland landraces, and green indicates lowland breeding lines

https://doi.org/10.17221/133/2025-CJGPB

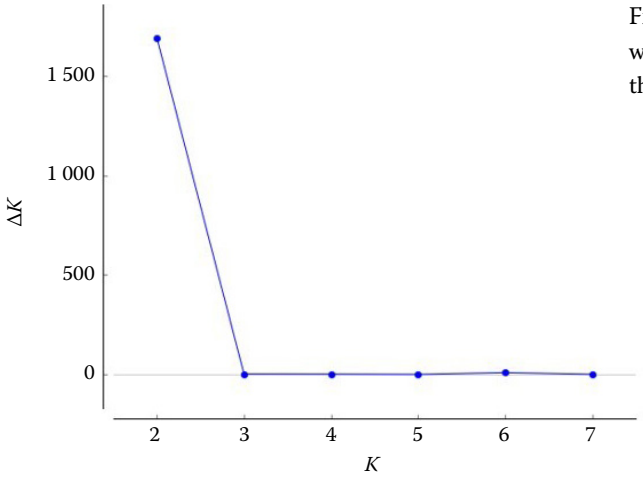

Figure 2. Estimation of the optimal number of *K* clusters where *K* = 2 based on f Δ*K*, (Δ*K* = mean (|L''(*K*)|)/sd[L(*K*)]), where the number of *K* represents the highest peak of ΔK

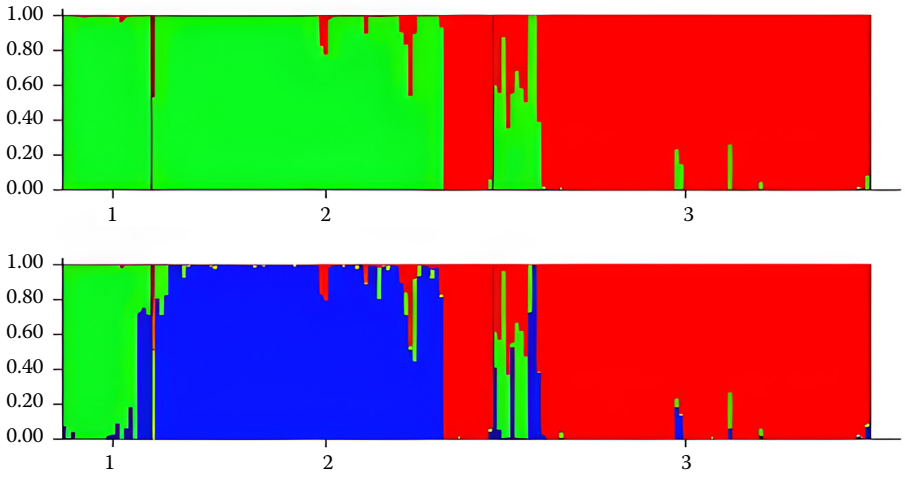

Figure 3. Population structure of 182 rice accessions inferred by STRUCTURE analysis at *K* = 2 (above) and *K* = 3 (below) using 20 polymorphic microsatellite markers; the pattern at *K* = 2 represents the major population structure, whereas *K* = 3 illustrates an early subdivision within the lowland group
1 – lowland breeding lines; 2 – lowland landraces; 3 – upland landraces

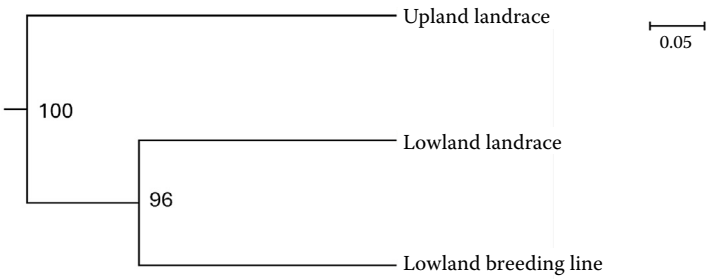

Figure 4. The unweighted pair group method with arithmetic mean (UPGMA) dendrogram showing the relationship among three rice accession groups: lowland breeding lines, lowland landraces, and upland landraces

tion history, thereby shaping the observed two-group genetic structure. In the present study, the analyses indicate that lowland breeding lines are genetically closer to lowland landraces than to upland landraces, suggesting underutilisation of upland germplasm in current breeding programmes and supporting the strategic incorporation of genetically distant upland landraces to broaden the genetic base.

https://doi.org/10.17221/133/2025-CJGPB

Table 2. Pairwise genetic distance among lowland breeding lines, lowland landraces, and upland landraces

OTU	Breeding line lowland	Landrace lowland	Landrace upland
Breeding line lowland	0.0000	0.4773	0.7496
Landrace lowland	0.4773	0.0000	0.6212
Landrace upland	0.7496	0.6212	0.0000

OTU – operational taxonomic unit

PCoA based on genetic distance matrices was used to visualise variety coordinates and support dendrogram and STRUCTURE analyses (Figure 5). The PCoA showed a broad separation between lowland and upland accessions, while lowland breeding lines were positioned closer to lowland landraces than to upland landraces, in agreement with the UPGMA and STRUCTURE results. The first three axes explained 39.63% of the cumulative variation (Table S3 in ESM), which was slightly higher than reported for breeding lines (34.79%) and higher than reported by Singh et al. (2016) for selected Indian rice varieties (15.9%). AMOVA based on the two major ecosystem groups recorded 31% variance among populations, 63% variance among individuals, and 6% variance within individuals (Table 3), indicating that a substantial proportion of the total genetic variation was attributable to differences between lowland and upland groups. The AMOVA components were significant based on 999 permutations, supporting the conclusion that ecosystem-related grouping contributed significantly to population differentiation. Pairwise Fst values ranged from 0.22 (lowland landrace ↔lowland breeding lines) to 0.44 (upland landrace ↔lowland breeding lines) (Figure 6). The lower Fst value between lowland landraces and lowland breeding lines indicates weaker differentiation and is consistent with the likelihood that current breeding materials were derived largely from the lowland gene pool. In contrast, the higher Fst value between upland landraces and lowland breeding lines suggests stronger differentiation and highlights the potential value of upland germplasm as a source of novel alleles for broadening the breeding base.

Implications for germplasm management and breeding. Landraces represent early forms of cultivated crops and are generally genetically diverse, locally adapted, and shaped by farmer-driven selection under low-input environments (Zeven 1998; Villa et al. 2005). In contrast, modern breeding lines often show reduced diversity due to intensive selection for high yield (Ab Razak et al. 2020). In Malaysia, reliance on narrowly diverse varieties has been associated with disease outbreaks such as rice blast, highlighting

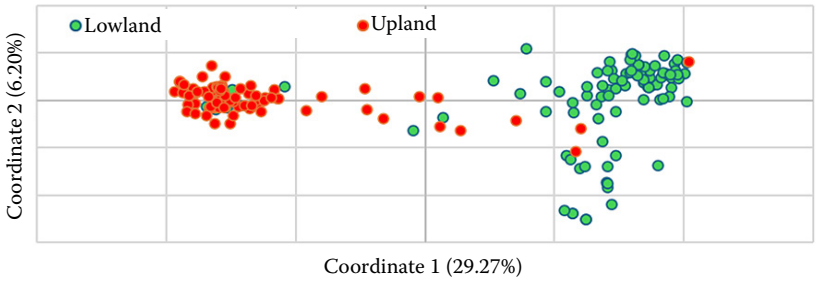

Figure 5. The principal coordinate analysis (PCoA) of 182 rice accessions based on 20 microsatellite markers

Table 3. AMOVA based on two major ecosystem groups using 20 microsatellite markers

Source	*df*	SS	MS	Est. Var.	%
Among populations	1	459.356	459.356	2.476	31
Among individuals	180	1 934.521	10.747	5.131	63
Within individuals	182	88.500	0.486	0.486	6
Total	363	2 482.376		8.092	100

AMOVA – analysis of molecular variance; *df* – degrees of freedom; SS – sum of squares; MS – mean squares; Est. Var. – estimated variance component; % – percentage of total variance

https://doi.org/10.17221/133/2025-CJGPB

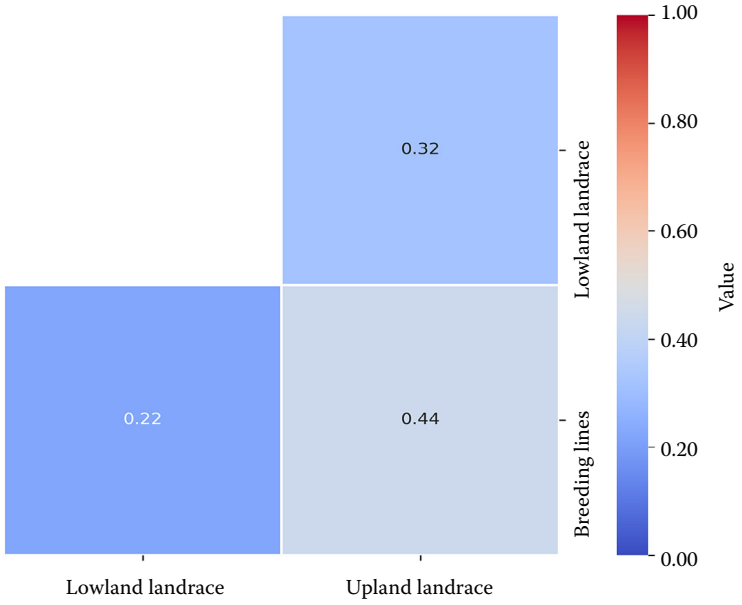

Figure 6. Heatmap of pairwise Fst values among lowland breeding lines, lowland landraces, and upland landraces

risks linked to genetic uniformity (Misman & Zakaria 2019). The present results, therefore, have direct implications for breeding strategy, particularly the need to diversify parental selection beyond closely related lowland materials. The clear differentiation between upland and lowland germplasm suggests that upland landraces could serve as valuable donors of adaptive alleles to enhance resilience under climate-related stresses, including drought-prone or more variable stress environments. At the same time, the closer relationship between lowland breeding lines and lowland landraces indicates that current breeding programmes may still be operating within a relatively narrow lowland genetic background. A broader crossing strategy involving genetically distant upland and lowland materials could therefore help increase allelic diversity and improve long-term breeding gain.

Germplasm duplication also presents challenges in genebank management. In this study, putatively identical genetic profiles across the 20 SSR loci were observed for two pairs of accessions (Gerit–Gerik and Jarom Perak–Jarom), likely reflecting different names for the same genetic material collected from different localities. Because these accessions were collected in Peninsular Malaysia, their identical profiles may reflect farmer-mediated seed exchange, movement of planting materials between neighbouring localities, or different local names assigned to the same genetic material. Similarly, MR220_CL1 and MR220_CL2 were indistinguishable using SSR markers despite previous differentiation using high-density SNP markers (Ab Razak et al. 2020), underscoring the limitation of SSR markers in resolving very closely related breeding lines and the potential need for additional SSR loci or SNP-based platforms for precise discrimination. Nevertheless, SSR markers remain practical and cost-effective for broad diversity assessment, preliminary fingerprinting, and routine germplasm management, whereas SNP platforms are more suitable when higher resolution is required for distinguishing closely related breeding materials. These findings highlight the complementary roles of SSR and SNP markers in germplasm management: SSRs remain useful for broad diversity assessment and preliminary duplicate detection, whereas higher-density SNP platforms are more suitable for confirming identity among closely related breeding materials.

CONCLUSION

Genetic diversity has been the cornerstone of plant breeding since the early era of agriculture. It provides the starting point for plant breeders to develop new cultivars to meet farmers' needs, adapt to climate change, and meet the increasing global demand for food. Various crop genetic resources, breeding instruments, and techniques are employed by plant breeders to incorporate genetic diversity into commercial cultivars. Plant genetic resources are the biological basis of food security and sustainability. As in this study, we have managed to resolve the genetic diversity of selected landrace rice acces-

https://doi.org/10.17221/133/2025-CJGPB

sions using a total of 20 SSR markers. Our work has provided insights for future sustainable breeding strategies to develop new rice varieties. The strategy should be aimed at extending the genetic core of the rice crop to ensure food sustainability and security for the future. Moreover, this study could also benefit conservation management programmes as we identified several duplications in the germplasm collection of landrace rice accessions, where detecting the duplicate accessions is essential to minimise maintenance and facilitate ideal germplasm banks.

Acknowledgements. The authors would like to express their sincere appreciation to all individuals who directly or indirectly contributed to this project.

REFERENCES

Ab Razak S., Azman N., Kamaruzaman R., Saidon S.A., Yusof M.F.M., Ismail S.N., Jaafar M.A., Abdullah N. (2020): Genetic diversity of released Malaysian rice varieties based on single nucleotide polymorphism markers. Czech Journal of Genetics and Plant Breeding, 56: 62–70.

Arif I.A., Khan H.A., Shobrak M., Al Homaidan A.A., Al Sadoon M., Al Farhan A.H., Bahkali A.H. (2010): Interpretation of electrophoretograms of seven microsatellite loci to determine the genetic diversity of the Arabian Oryx. Genetic and Molecular Research, 9: 259–265.

Beena R., Sowmiya S., Visakh R.L., Shelvy S., Sasmita B., Sah R.P. (2025): Unravelling genetic mechanisms for heat tolerance in rice landraces of Kerala, India. Euphytica, 221: 51.

Botstein D., White R.L., Skolnick M., Davis R.W. (1980): Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32: 314.

Chakraborty R., Kimmel M., Stivers D.N., Davison L.J., Deka R. (1997): Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences, 94: 1041–1046.

Chesnokov Y.V., Kosolapov V.M., Savchenko I.V. (2020): Morphological genetic markers in plants. Russian Journal of Genetics, 56: 1406–1415.

Comertpay G., Baloch F.S., Derya M., Andeden E.E., Alsaleh A., Surek H., Özkan H. (2016): Population structure of rice varieties used in Turkish rice breeding programs determined using simple sequence repeat and inter primer binding site retrotransposon data. Genetic and Molecular Research, 15: gmr.15017158.

Dang X., Giang Tran Thi T., Mawuli Edzesi W., Liang L., Liu Q., Liu E., Wang Y., Qiang S., Liu L., Hong D. (2015): Population genetic structure of *Oryza sativa* in East and Southeast Asia and the discovery of elite alleles for grain traits. Scientific Reports, 5: 1–13.

Dao S., Traoré D., Teketé C., Kassogué A., Coulibaly M., Ouattara O., Goïta O. (2024): Genetic diversity assessment of lowland and upland rice varieties of Mali using microsatellite markers. International Journal of Genetics and Molecular Biology, 16: 1–11.

Evanno G., Regnaut S., Goudet J. (2005): Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 14: 2611–2620.

Fufa H., Baenziger P.S., Beecher B.S., Dweikat I., Graybosch R.A., Eskridge K.M. (2005): Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. Euphytica, 145: 133–146.

Hanamaratti N.G., Prashanthi S.K., Salimath P.M., Hanchinal R.R., Mohankumar H.D., Parameshwarappa K.G. (2008): Traditional landraces of rice in Karnataka: Reservoirs of valuable traits. Current Science, 94: 242–247.

Kumar S., Stecher G., Tamura K. (2016): MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, 33: 1870–1874.

Li Y.L., Liu J.X. (2018): StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. Molecular Ecology Resources, 18: 176–177.

Liu K., Muse S.V. (2005): PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics, 21: 2128–2129.

Mace E.S., Buhariwalla K.K., Buhariwalla H.K., Crouch J.H. (2003): A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Molecular Biology Reporter, 21: 459–460.

Merritt B.J., Culley T.M., Avanesyan A., Stokes R., Brzyski J. (2015): An empirical review: Characteristics of plant microsatellite markers that confer higher levels of genetic variation. Applications in Plant Sciences, 3: 1500025.

Misman S.N., Zakaria L. (2019): Pathotype identification of rice blast pathogen, *Pyricularia oryzae* using differential varieties in Peninsular Malaysia. Tropical Life Sciences Research, 30: 181–190.

Mudhale A., Sar P., Kumar J., Bhowmick P.K., Banerjee A., Chakraborty K., Roy S. (2024): Characterization of rice (*Oryza sativa* L.) landraces from Majuli and surrounding riverine ecologies in Assam: Assessment of morphogenetic variability and submergence tolerance. Plant Breeding, 143: 469–480.

Peakall R.O.D., Smouse P.E. (2006): GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6: 288–295.

https://doi.org/10.17221/133/2025-CJGPB

Pritchard J.K., Stephens M., Donnelly P. (2000): Inference of population structure using multilocus genotype data. Genetics, 155: 945–959.

Rocha S.S., Londe L.C.N., Pimenta S., Cardoso M.M., Goncalves N.P., Gome W.S., Calaes J.G. (2020): Congruence between morphological and molecular markers for genetic diversity analysis applied to forage palm genotypes propagated via bioreactors. Industrial Crops and Products, 147: 112230.

Salgotra R.K., Chauhan B.S. (2023): Genetic diversity, conservation, and utilization of plant genetic resources. Genes, 14: 174.

Schuelke M. (2000): An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18: 233–234.

Singh N., Choudhury D.R., Tiwari G., Singh A.K., Kumar S., Srinivasan K., Tyagi R.K., Sharma A.D., Singh N.K., Singh R. (2016): Genetic diversity trend in Indian rice varieties: An analysis using SSR markers. BMC Genetics, 17: 1–13.

Singh S., Gupta S.K., Thudi M., Das R.R., Vemula A., Garg V., Yadav D.V. (2018): Genetic diversity patterns and heterosis prediction based on SSRs and SNPs in hybrid parents of pearl millet. Crop Science, 58: 2379–2390.

Thomson M.J., Polato N.R., Prasetiyono J., Trijatmiko K.R., Silitonga T.S., McCouch S.R. (2009): Genetic diversity of isolated populations of Indonesian landraces of rice (*Oryza sativa* L.) collected in East Kalimantan on the island of Borneo. Rice, 2: 80–92.

Van Oosterhout C., Hutchinson W.F., Wills D.P., Shipley P. (2004): MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4: 535–538.

Venkateshwarlu C., Kole P.C., Singh A.K., Paul P.J., Sinha P., Singh V.K., Kumar A. (2025): Mapping genomic regions for reproductive stage drought tolerance in rice from exotic landrace-derived population. Frontiers in Plant Science, 15: 1495241.

Vieira M.L.C., Santini L., Diniz A.L., Munhoz C.D.F. (2016): Microsatellite markers: What they mean and why they are so useful. Genetics and Molecular Biology, 39: 312–328.

Villa T.C.C., Maxted N., Scholten M., Ford-Lloyd B. (2005): Defining and identifying crop landraces. Plant Genetic Resources, 3: 373–384.

Yeh F.C., Yang R.C., Boyle T.J., Ye Z., Xiyan J.M. (2000): PopGene32. Microsoft Windows-based free software for population genetic analysis. Release 1.32. Available at https://www.scienceopen.com/book?vid=2d45ad78-b140-4b66-b80f-2c9f513ec997

Zeven A.C. (1998): Landraces: A review of definitions and classifications. Euphytica, 104: 127–139.

Zhang L.N., Cao G.L., Han L.Z. (2013): Genetic diversity of rice landraces from lowland and upland accessions of China. Rice Science, 20: 259–266.

Zhu Y., Chen H., Fan J., Wang Y., Li Y., Chen J., Fan J., Yang S., Hu L., Leung H., Mew T.W. (2000): Genetic diversity and disease control in rice. Nature, 406: 718–722.

Zurn J.D., Nyberg A., Montanari S., Postman J., Neale D., Bassil N. (2020): A new SSR fingerprinting set and its comparison to existing SSR-and SNP-based genotyping platforms to manage *Pyrus* germplasm resources. Tree Genetics & Genomes, 16: 72.

Received: December 29, 2025
Accepted: April 28, 2026
Published online: June 10, 2026