HAYATI
Journal of Biosciences

Vol. 29 No. 6, November 2022 749-761
DOI:10.4308/hjb.29.6.749-761
ISSN: 1978-3019
EISSN: 2086-4094

Diversity and Population Structure of Local Rice Varieties from Indonesia Revealed by SSR Markers

Rerenstradika Tizar Terryana[1*], Puji Lestari[1], I Gusti Komang Dana Arsana[1], Kristianto Nugroho[1], I Made Jana Mejaya[1], Priatna Sasmita[2], Yudi Sastro[3], Karden Mulya[1], Dwinita Wikan Utami[1], Mastur[4]

[1]*National Research and Innovation Agency, Jakarta, Indonesia*
[2]*Indonesian Center for Food Crops Research and Development-IAARD, Bogor, Indonesia*
[3]*Indonesian Rice Research Center-IAARD, Subang, Indonesia*
[4]*Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development-IAARD, Bogor, Indonesia*

ARTICLE INFO

Article history:
Received May 21, 2020
Received in revised form April 6, 2022
Accepted June 2, 2022

KEYWORDS:
Crop improvement,
Genetic structure,
Molecular markers,
Oryza sativa

ABSTRACT

Indonesia's local rice varieties (*Oryza sativa* L.) have a wide range of diversity that can be valuable sources for crop improvement with molecular markers. This study investigated local rice varieties' genetic diversity and population structure using simple sequence repeat (SSR) markers. The SSR markers demonstrated their informativeness for genotypic characterization as represented by the gene diversity indices and polymorphic information content. The UPGMA dendrogram divided 63 varieties into two distinct clusters with different levels of sub-grouping and the tendency according to their origins, as supported by PCoA. In contrast, PCA of these varieties according to agro-morphological traits was scattered in all quadrants. Thus, DNA level variation analyzed by SSR seems to complement the phenotypic traits, which were not well structured and revealed significant genetic diversity among varieties, within, and among populations (P<0.01). The pattern of grouping structure analysis of total varieties into two subpopulations is similar to the dendrogram according to SSR markers but better resembles the pedigree information of the set local varieties. These findings have implied their importance in rice breeding for genetic improvement, maintenance, and management of local genetic resources in Indonesia.

1. Introduction

Indonesia's dependency on rice as the main energy source of carbohydrates is undeniable (Iskandar *et al.* 2018), as reflected by a high consumption rate of up to 150 kg per capita in 2017 (Ministry of Agriculture 2018). With 264 million people and a 1.27% annual population growth rate in 2018, Indonesia has to provide an enormous amount of rice food availability in the future (Ministry of Agriculture 2018). The local rice varieties play a significant role in the food security and sustainable development of agriculture, indicating their significant resource for rice genetic improvement (Iskandar *et al.* 2018). A breeding program is commonly targeted for the most desirable traits of rice, such as increased yield (Fuller 2011; Gross and Zhao 2014) and grain quality to suffice the high consumption of rice (Permana *et al.* 2018). Since genetic improvement in rice is crucial, this program should also broaden the breeding material stocks' genetic base.

Although improved rice varieties have been grown predominantly in many Indonesian areas, local rice varieties are still prevalent in several regions due to their unique characteristics and adaptability to specific environmental and climate conditions (Nurhasanah *et al.* 2017). For many generations, the local rice varieties have been cultivated and preserved by indigenous farmers on specific agroecology to the commercial variety. Most local rice varieties have good adaptability, good grain quality, and taste quality that meet the local consumer's preference. Surprisingly, a few traditional varieties have still been grown on the mountainsides and are highly regarded as adapted well to abiotic and biotic stresses that naturally occur in the environment (Thomson *et al.* 2009). Thus, the genetic characterization of this local germplasm, a snapshot of a dynamic gene pool of local varieties adapted to the local environment

* Corresponding Author
E-mail Address: rerenstradika@gmail.com
These authors contributed equally to this work

and in harmony with traditional culture preferences, represents an important basis for conserving this precious genetic diversity and improving rice varieties for future generations. Despite the potential for crop improvement, only a small proportion of them has been utilized in rice breeding programs (Reig-valiente *et al.* 2016). Consequently, proper conservation of local varieties in parallel with broadening the gene pool of rice varieties becomes urgent for rice breeding in Indonesia (Nurhasanah *et al.* 2016).

The foundation of genetic improvement or rice generally depends on the genetic diversity level. Genetic diversity can be estimated using morphological traits, isozymes, and molecular markers. Molecular marker technology based on DNA is a powerful tool for estimating the genetic variation of germplasm. Contrary to morphological traits, the molecular marker can reveal many differences among genotypes at the DNA level, providing a more proper, reliable, and efficient tool for germplasm characterization without being influenced by the environment. Simple sequence repeat (SSR) markers have been generally used in genetic diversity studies and applied in plant breeding (Carvalho *et al.* 2020; Gonçalves-Vidigal *et al.* 2011), including rice, due to the advantages of informativeness and stability (Reig-valiente *et al.* 2016; Singh *et al.* 2016). SSR markers were used to assess the genetic diversity level of local rice varieties (Kristamtini *et al.* 2014; Thomson *et al.* 2009) and the population structure (Suvi *et al.* 2019), which are helpful as basic information for breeding. Parallel to breeding, genetic diversity information gained from local variety would benefit the sustainable conservation of plant genetic resources for food and agriculture (Hue *et al.* 2018). This study aimed to evaluate the genetic diversity among 63 Indonesian local rice varieties and explore their genetic structure revealed by SSR markers complement agro-morphological information for better understanding and utilization as germplasm resources for breeding.

2. Materials and Methods

2.1. Genetic Materials

This study was conducted in the Molecular Biology Laboratory of the Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development, the Indonesian Agency for Agricultural Research and Development, from May to September 2019. As many as 63 local rice varieties were collected from seven provinces in Indonesia, namely East Java (17 varieties), Bali (22 varieties), South Sulawesi (14 varieties), East Kalimantan (2 varieties), West Kalimantan (2 varieties), Central Kalimantan (1 variety) and West Java (5 varieties) were used in this study (Table 1). These materials were obtained from the seed bank collection of the Assessment Institute for Agricultural Technology (AIAT) of Bali and AIAT of East Java.

2.2. Procedures

2.2.1. Genomic DNA Extraction and PCR Amplification

The genomic DNA was extracted from rice grain tissue by grinding it in a sterile mortar using 500 µl extraction buffer (100 mM Tris-HCL pH 8.0, 1.4 NaCl, 20 mM EDTA pH 8.0, 2% (w/v) CTAB (cetyl trimethyl ammonium bromide) for one sample per variety according to the previous method (Doyle and Doyle 1990), modified by adding 2% (w/v) PVP (polyvinylpyrrolidone) and 0.38% (w/v) sodium disulfide. The samples were put on a 2 ml microtube, followed by the addition of the extraction buffer adjusted up to 1 ml. The samples were incubated at 65°C for 15 minutes, extracted twice using chloroform: isoamyl alcohol solution (24:1), and centrifuged at 12,000 rpm for 10 minutes at 20°C. The supernatant was transferred to the new microtube. Furthermore, 3M sodium acetate pH 5.2 was added to as many as 1/10 of supernatant volume, followed by cold isopropanol as much as one supernatant volume. After incubating the mixture at -20°C for one hour, it was centrifuged at 12,000 rpm for 10 minutes at 20°C. The DNA pellets were then washed using 70% ethanol and dried using DNA Speed Vac Concentrator (ThermoScientific, USA). Dry pellets were dissolved in 100 µl TE solution (10 mM Tris pH 8.0 and 1 mM EDTA) and diluted to a final concentration of 10 ng/µl to enable polymerase chain reactions.

DNA amplification parameters are strongly influenced by the components of the Polymerase chain reaction (PCR) reaction and by thermal cycling conditions (Caetano-Anoles and Brant 1991). Therefore, careful optimization of reaction components and conditions would ultimately result in more reproducible and efficient amplification. A total of 15 SSR markers and their sequences and motif used for diversity and structure population analysis were obtained from several references (Table 2). These total markers were considered sufficient and adequate to assess the genetic diversity of 63 Indonesian local rice varieties due to high abundance, multiallelic, high reproducibility, and ease of assessing their size variation using PCR (Song *et al.* 2019). DNA analysis using a molecular marker is trustable when the DNA bands appear

Table 1. A list of 63 local rice varieties originating from seven provinces in Indonesia was used in this study

Local rice varieties	Grain color	Province
Baban (Daun Bendera Lebar)	White	East Java
Merah Wangi (Pct)	Red	East Java
Gogoniti-1	Black	East Java
Danau Putih	White	East Java
Blambangan A3	Red	East Java
Saman 6	Black	East Java
Kemanggisan	Black	East Java
Merah Lonjong (MLG)	Red	East Java
Srikandi (D) (MLG)	White	East Java
Pariwangi (MLG)	White	East Java
Sengkut 2 (MLG)	Red	East Java
Susu 3 (MLG)	White	East Java
Cemani	Red	East Java
Sempol	Red	East Java
Kali Ayah	White	East Java
Kali Goro	Red	East Java
Srijaya	Red	East Java
Padi Taun Ijo Gading	White	Bali
Padi Injin	Black	Bali
Mansur Tabanan	White	Bali
Beras Taun Ijo Gading Tabanan	White	Bali
Cicih Merah	Red	Bali
Padi Injin Sambangan Buleleng	Black	Bali
Padi Bali Sambangan Buleleng	White	Bali
Padi Taun Merah Pupuan	Red	Bali
Ingse Putih Selat	White	Bali
Padi Taun Putih Pupuan	White	Bali
Ingse Barak Selat	Red	Bali
Merah Cendana Urangaya	Red	Bali
Padi Ketan Taun Pupuan	White	Bali
Cicih Suar Tabanan	White	Bali
Merah Cendana	Red	Bali
Padi Merah Pupuan	Red	Bali
Padi Cicih Merah	Red	Bali
Padi Mansur Pupuan	White	Bali
Padi Sodaji Merah	Red	Bali
Padi Krotok	White	Bali
Padi Injin Buleleng	Black	Bali
Ketan Sudaji Buleleng	White	Bali
Pare Barri	White	South Sulawesi
Pare Solo	White	South Sulawesi
Pare Mansyur	White	South Sulawesi
Madakko	Red	South Sulawesi
Pare Bau'	White	South Sulawesi
Pulu' Pinjan	White	South Sulawesi
Pare Lallodo	Black	South Sulawesi
Pare Jagong	White	South Sulawesi
Pare Kaloko	White	South Sulawesi
Palu' Tau Padang	White	South Sulawesi
Pare Lambau	White	South Sulawesi
Pulu' Kombong	White	South Sulawesi
Pulu' Bolong	Black	South Sulawesi
Pare Kate	White	South Sulawesi
Padi Pulut Ayang	White	East Kalimantan
Ketan Maronto	White	East Kalimantan
Padi Hitam A	Black	West Kalimantan
Gadabung A	White	Central Kalimantan
Lima	White	West Kalimantan
Batanghari	White	West Java
Widas	White	West Java
Gogo	White	West Java
Segon Benggala	White	West Java
Sereh A	White	West Java

Table 2. List of SSR markers used in this study

Markers	Sequence	SSR motif	Chr	Tm	Product size	References
RM3894	F: TATGCTCTCTCCTTCAGGCC R: CTTACCAACTCCGCACTTGC	(GT)15	3	55	201	McCouch *et al.* (2002)
RM5	F: TGCAACTTCTAGCTGCTCGA R: GCATCCGATCTTGATGGG	(GA)14	1	55	113	Panaud *et al.* (1996)
RM11	F: TCTCCTCTTCCCCCGATC R: ATAGCGGGCGAGGCTTAG	(GA)17	7	55	140	Panaud *et al.* (1996)
RM19	F: CAAAAACAGAGCAGATGAC R: CTCAAGATGGACGCCAAGA	(ATC)10	12	55	226	Panaud *et al.* (1996)
RM72	F: CCGGCGATAAAACAATGAG R: GCATCGGTCCTAACTAAGGG	(TAT)5C(ATT)15	8	55	166	Temnykh *et al.* (2000)
RM201	F: CTCGTTTATTACCTACAGTACC R: CTACCTCCTTTCTAGACCGATA	(CT)17	9	55	158	Chen *et al.* (1997)
RM223	F: GAGTGAGCTTGGGCTGAAAC R: GAAGGCAAGTCTTGGCACTG	(CT)25	8	55	165	Chen *et al.* (1997)
RM228	F: CTGGCCATTAGTCCTTGG R: GCTTGCGGCTCTGCTTAC	(CA)6(GA)36	10	55	154	Chen *et al.* (1997)
RM263	F: CCCAGGCTAGCTCATGAACC R: GCTACGTTTGAGCTACCACG	(CT)34	2	55	199	Chen *et al.* (1997)

Table 2. Continued

Markers	Sequence	SSR motif	Chr	Tm	Product size	References
RM287	F: TTCCCTGTTAAGAGAGAAATC R: GTGTATTTGGTGAAAGCAAC	(GA)21	11	55	118	Temnykh *et al.* (2000)
RM324	F: CTGATTCCACACACTTGTGC R: GATTCCACGTCAGGATCTTC	(CAT)21	2	55	175	Temnykh *et al.* (2000)
RM328	F: CATAGTGGAGTATGCAGCTGC R: CCTTCTCCCAGTCGTATCTG	(CAT)5	9	55	172	Temnykh *et al.* (2000)
RM416	F: GGGAGTTAGGGTTTTGGAGC R: TCCAGTTTCACACTGCTTCG	(GA)9	3	55	114	Temnykh *et al.* (2000)
RM518	F: CTCTTCACTCACTCACCATGG R: ATCCATCTGGAGCAAGCAAC	(TC)15	4	55	171	Temnykh *et al.* (2000)
RM541	F:TATAACCGACCTCAGTGCCC R:CCTTACTCCCATGCCATGAG	(TC)16	6	55	158	Temnykh *et al.* (2000)

and are readable as expected; thus, only one sample was amplified for each variety in this study. The PCR reaction mixture of 10 µl containing 10 ng/µl DNA template; 2x MyTaq HS (Bioline, UK); 10 pmoles of forward and reverse primer, and Milli-Q water. The PCR reaction was performed in T1 Thermocycler (Biometra, Germany) using the following program: 95°C for 5 minutes' initial denaturation, 35 cycles of 94°C denaturation for 30 seconds, 55°C annealing for 1 minute, 72°C for 1-minute extension with a final extension of 60°C for 15 minutes.

2.2.2. Polyacrylamide Gel Electrophoresis and Allele Scoring

Amplified products were size separated in electrophoresis using 6% (w/v) polyacrylamide gel in a vertical electrophoresis tank contained with 1x TBE (Tris Borate EDTA) at 80 V for 1.5 hours. The PCR products were detected by staining them using ethidium bromide and visualized under UV light using Transilluminator (Biorad, USA).

Based on the expected product size given in the references (Table 2), the size of the most intensely amplified bands around the expected product size for each SSR marker was identified using standard molecular weight size markers (100 bp DNA ladder). Then, the allele score was given based on the presence of a particular size allele in each sample using GelAnalyzer 2010a software. In contrast, the samples that did not produce alleles were given a question mark symbol (?) which is considered missing data.

2.2.3. Agro-Morphological Characters

This study collected six important agro-morphological characteristics, including plant height, the number of productive tillers, days to harvest, 1000 grain weight, grain color, and yield potential of 63 local rice varieties. These selected agro-morphological data were observed of each variety when requesting its registration for plant protection.

2.2.4. Data Analysis

A matrix was constructed based on the molecular weight size of each presence allele for the set of 15 SSR markers. These SSR genotypic data were analyzed for genetic diversity and population structure. Genetic diversity parameters such as major allele frequency, the number of alleles per locus, gene diversity, heterozygosity, and the polymorphic information content (PIC) for a set of local varieties were estimated by POWERMARKER V3.25 (Liu and Muse 2005). Allele frequency represents the frequency of a particular allele for each marker. Heterozygosity is the proportion of heterozygous individuals in the population. Polymorphic information content representing the amount of polymorphism within a population was estimated (Botstein *et al.* 1980).

Unweighted pair group method with arithmetic mean (UPGMA) method based on Nei's genetic distances among genotypes (Nei 1978) using POWERMARKER V3.25 (Liu and Muse 2005) and MEGA 6.0 software (Tamura *et al.* 2013) was used for cluster analysis of 63 local rice varieties. The presence of molecular variance within and between hierarchical population structures estimated by Structure was assessed via analysis of molecular variance (AMOVA) by GenAlEx 6.5 (Peakall and Smouse 2012). AMOVA and principal coordinate analysis (PCoA) of the 63 local rice varieties according to agro-morphological characters were performed based on Nei's (Nei 1978) distance matrix using GenAlEx 6.5 (Peakall

and Smouse 2012). In contrast, principal component analysis (PCA) was generated using XLSTAT (Iyai *et al.* 2008).

The model-based program Structure v2.3.4 (Earl and vonHoldt 2012) analyzed the population structure of total varieties originating from 7 provinces by using 15 SSR primers. Ten independent simulations for each K (the number of populations) ranged from 1 to 10. For each simulation, 10,000 iterations before a burn-in length of 50,000 Markov Chain Monte Carlo (MCMC) replications were performed to select admixture and related frequency models. The plot of mean posterior probability (LnP(D)) values and optimal K-value was estimated using Evanno's 1K method (Evanno *et al.* 2005) with the online tool Structure Harvester (Earl and vonHoldt 2012).

3. Results

3.1. SSR Polymorphism

This present study estimated the genetic diversity of 63 local rice varieties originating from seven provinces in Indonesia. These local varieties predominantly were from Bali, followed by East Java and South Sulawesi, with a minor number of varieties from West Java and three provinces from Kalimantan. Fifteen SSR markers were polymorphic on 63 local rice varieties from Indonesia. They presented their examples of the banding patterns of DNA produced by some SSR markers (Figure 1), which amplified 456 alleles with a mean of 30.4 alleles. The mean major allele frequency was 0.125, and the RM324 marker exhibited the highest value. The number of alleles per loci varied from 16 (RM324) to 50 (RM72), with an average of 30 alleles per locus. The statistical summary of 15 SSR markers polymorphism is presented in Table 3.

The gene diversity indices reflecting polymorphism level (Nei 1978) varied from 0.852 (RM324) to 0.970 (RM72 and RM263), with an average of 0.937. As part of gene diversity indices, heterozygosity was observed in these local varieties as identified by nine SSR primers. Heterozygosity was very high (>0.5) as presented by some SSR markers (RM3894, RM5, RM11, RM19, RM72, RM201, RM223, RM228), indicating the heterogeneous nature of these local varieties. PIC value represents the relative informativeness of each marker with an average of 0.933. The highest genetic diversity is explained by the local germplasm observed with the mean PIC value for markers was 0.93 with a range of 0.83 (RM324) to 0.96 (RM72 and RM263).

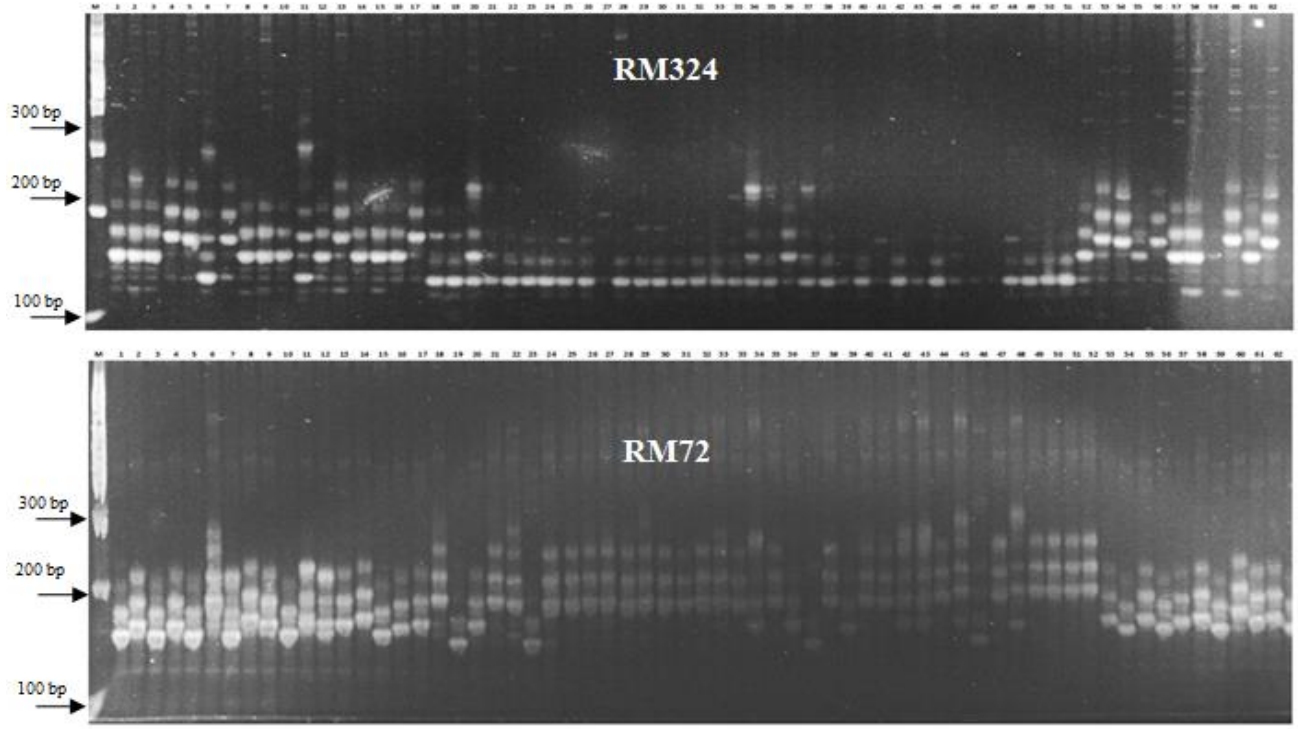


Figure 1. DNA banding pattern of RM324 and RM72 SSR markers obtained in 63 local rice

Table 3. Details information of SSR polymorphism observed in 63 local rice varieties

Marker	Major allele frequency	Allele number	Gene diversity indices	Heterozygosity	PIC
RM3894	0.103	24	0.937	1.000	0.934
RM5	0.145	18	0.920	0.710	0.914
RM11	0.115	32	0.950	0.902	0.948
RM19	0.111	44	0.950	0.984	0.948
RM72	0.079	50	0.970	0.968	0.969
RM201	0.079	34	0.955	0.968	0.954
RM223	0.069	34	0.958	0.845	0.957
RM228	0.125	43	0.959	0.900	0.958
RM263	0.051	47	0.970	0.492	0.969
RM287	0.129	20	0.921	0.000	0.916
RM324	0.254	16	0.852	0.000	0.837
RM328	0.119	25	0.945	0.000	0.942
RM416	0.175	25	0.922	0.000	0.918
RM518	0.190	21	0.904	0.000	0.896
RM541	0.127	23	0.941	0.000	0.938
Mean	0.125	30.4	0.937	0.518	0.933

3.2. Genetic Diversity

A phylogenetic tree based on Nei's coefficient grouped 63 local rice varieties into two distinct clusters (Figure 2). Cluster I consisted of 33 local rice varieties originating from Bali and the majority from South Sulawesi. The remaining 30 varieties, most of East Java, West Java, and Kalimantan origin, comprised cluster II, indicating their close genetic distance in the same group. Interestingly, varieties from South Sulawesi that were identified virtually in cluster I showed only one variety (Pare Kate) was close to others in cluster II. Cluster II was found to be lacking in the Bali population.

Analysis of molecular variance (AMOVA) of the total genetic variation among the total observed varieties indicated that variance components were significant ($P<0.01$). The result of AMOVA with seven populations of 63 varieties based on their origins revealed that the majority of genetic diversity was allocated within local varieties and accounted for 52% of the total variation, whereas 8% and 39% of the variation were attributed to differences among populations and varieties within populations, respectively (Table 4). While the average value of genetic distance was 0.794 (range 0.551-0.966), indicating the presence of a wide range of genetic diversity among total varieties. The most significant genetic distance (0.966) was between Bali and Central Kalimantan populations. Local rice varieties originated from East Kalimantan, and West Kalimantan was very similar to the Bali population. The genetic similarity and genetic distance of 63 local rice varieties are shown in Table 5.

3.3. PCoA and PCA

The PCoA showed two main groups of 63 local rice varieties similar to those detected by the UPGMA dendrogram based on SSR polymorphism. These clusters were specifically distributed along with Coordinates 1 and 2 in the PCoA plot (Figure 3). In this plot, Group 1 was mainly concentrated in quadrants 1 and 4, and about 2/3 of the Group 2 was located in quadrant 2. Both groups showed intermixing between each other varieties. The first two principal coordinates explained 27.78% and 17.44% of the molecular variance. Overall, about 45.22% of the total variation was described by the first two principal coordinates.

The PCA was carried out based on six important agro-morphological characters (plant height, number of productive tillers, days to harvest, 1000 grain weight, grains color, and yield potential) to understand the genetic diversity among local rice varieties used in this study. The result of PCA has shown that two out of six principal components with an eigenvalue above 1.0 contributed to 51.59% of total variations. The coefficients are provided for this purpose because these might show the correlations between observed and derived variables. The first and second principal components contributed 31.45% and 20.14% to the total variance, respectively (Figure 4). The first principal component, which accounted for the highest variability, mainly was related to agro-morphological characters. There are grains color (0.20), days to harvest (0.56), number of productive tillers (0.46), and yield potential (0.49).

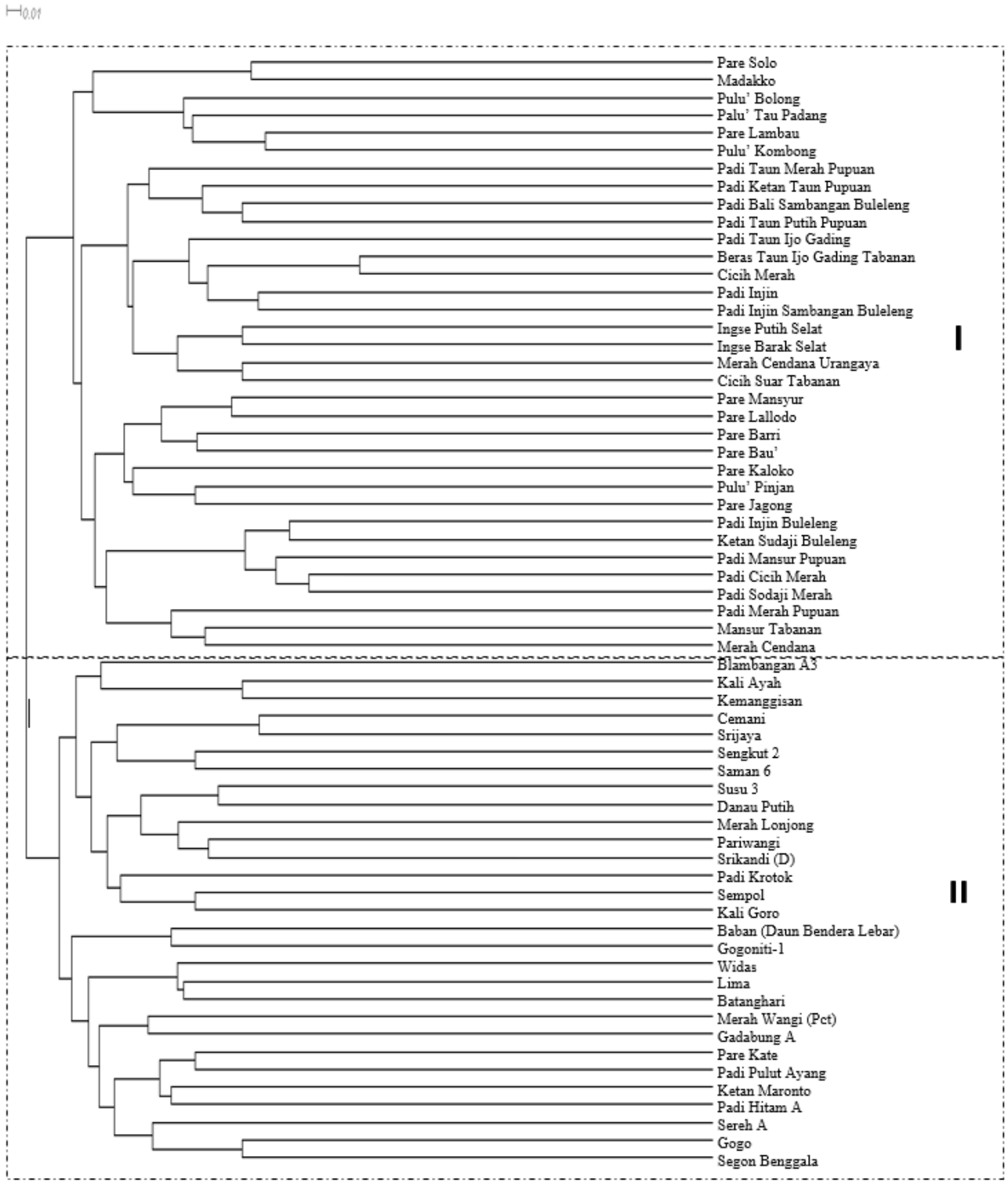


Figure 2. Dendrogram of genetic diversity of 63 local rice varieties using 15 SSR markers (Nei's coefficient)

Table 4. Results of AMOVA analysis and Nei's unbiased measures of genetic distance and genetic similarity of 63 local rice varieties based on 15 SSR markers

Source	df	Sum of squares	Variance component estimates	Percentage of the total variance (%)	P
Among populations	6	116.144	0.613	0.613	<0.01
Among varieties	57	541.262	2.849	2.849	<0.01
Within population	64	243.000	3.797	3.797	<0.01
Total	127	900.406	7.259	7.259	

Table 5. Genetic distance between local rice varieties according to their origin based on SSR polymorphism

Populations	East Java	Bali	South Sulawesi	East Kalimantan	West Kalimantan	Central Kalimantan	West Java
East Java	1						
Bali	0.738	1					
South Sulawesi	0.778	0.550	1				
East Kalimantan	0.713	0.894	0.852	1			
West Kalimantan	0.772	0.894	0.838	0.741	1		
Central Kalimantan	0.837	0.966	0.860	0.815	0.829	1	
West Java	0.696	0.865	0.805	0.787	0.656	0.789	1

Nei's genetic distance

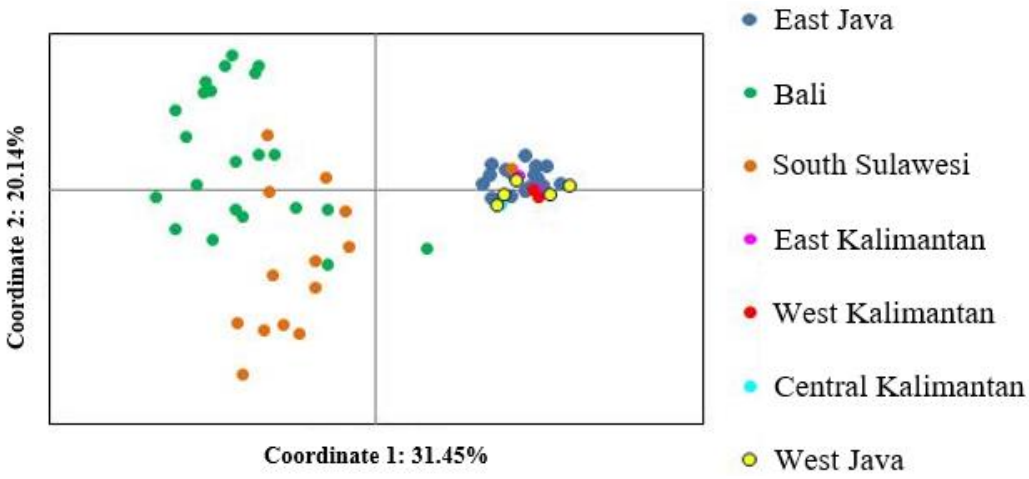
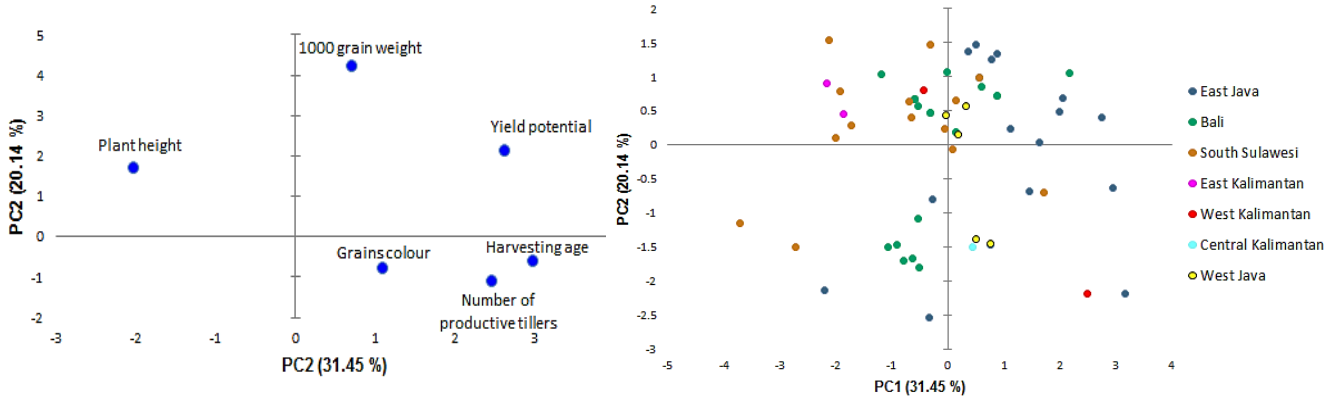


Figure 3. Principal coordinate analysis showing the spatial distribution of 63 local rice varieties based on SSR polymorphism



A B

Figure 4. The principal scatter plot depicts the contribution of agro-morphological in the first and second principal components in 63 local rice varieties (A) and the genetic diversity among 63 local rice varieties as shown by the first and second principal components (B)

The second principal component accounted for 20.14% of the total variance. Agro-morphological characters highly and positively correlated with the second PC are plant height (0.32) and 1000 grain weight (0.80) without considering grain color.

The distribution of 63 local rice varieties based on the first and second principal components exhibited the morphological character variation among the population. It explained how these were widely dispersed along both axes (Figure 4). The total local rice varieties evaluated through PCA were grouped into different clusters with more agro-morphological similarities among local varieties within the cluster. The distribution of local rice varieties into other groups revealed that genetic diversity existed among these local rice varieties. The scattered plot of the principal component showed that 63 local rice varieties were scattered in all the quadrants, which is also a representative that a high level of genetic variability is present among the local varieties used.

3.4. Population Structure

The population structure of 63 local rice varieties was analyzed using 15 SSR markers using a Bayesian approach. Each population group is based on the likelihood for each number of groups (K). The estimated membership fractions of the 63 local rice varieties for different k values ranged from 2 to 7. The log-likelihood revealed by the structure showed that the highest k value was 2 (K = 2), followed by a peak at K = 5. More mixing varieties in more subpopulations/clusters appeared in four and five subpopulations at K = 4 to K = 5, respectively. The best ΔK was 2, which indicated that the entire population evaluated could be clustered into two major groups with a high admixture among the varieties. A color denoted each variety presented by a single color line and each cluster (Figure 5).

This structure analysis in our study supported the phylogenetic tree by depicting the estimated membership of each local variety in each subpopulation (from K = 2 to K = 5), allowing us to

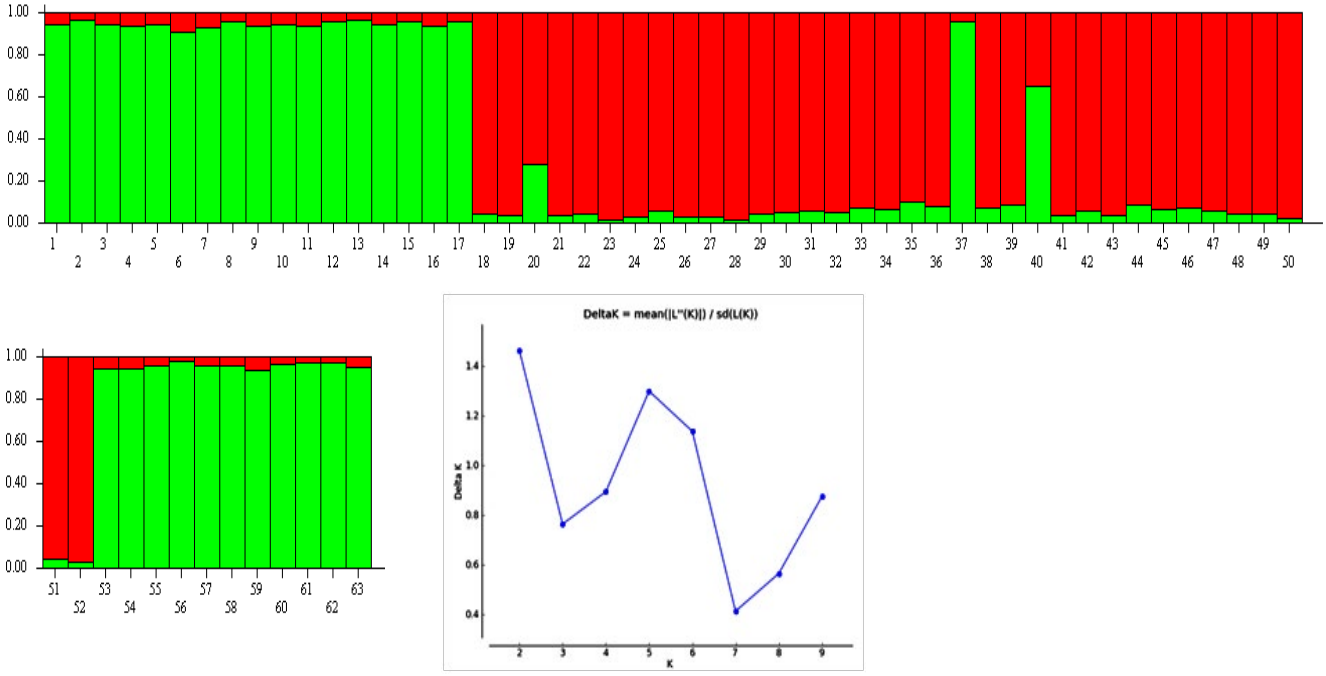


Figure 5. Sixty-three local rice varieties population structure at K = 2 and graph of estimated membership fraction at K = 2. The maximum K value was determined by structure harvest, which determined that the population can be grouped into two subgroups, 3, 4, and 5. The clusters' membership coefficients (y-axis) were determined using the STRUCTURE program based on 10,000 iterations. Bar lengths represent the membership probability of accessions belonging to different groups

identify the admixtures quickly. At the K = 2, the classified two groups with a green line as cluster 1 and a red line as cluster 2 comprised local rice varieties from seven provinces in Indonesia. Subpopulations 1 and 2 comprised 30 and 33 varieties, respectively. In this population structure, 17 varieties originating from East Java and five from West Java were in the same subpopulation together with five varieties from Kalimantan, one variety from Bali (Padi Krotok), and one from South Sulawesi (Pare Barri). Interestingly, unlike other local varieties in subpopulation 1, Pare Barri from South Sulawesi was found to have mixed ancestors in the pedigree. Similarly, Mansur Tabanan from Bali structured in subpopulation 2 showed to be influenced by genes or ancestors from subpopulation 1 (Figure 5). Notably, the grouping pattern into two subpopulations is similar to UPGMA dendrogram according to SSR markers.

4. Discussion

The genetic diversity estimated in 63 local rice varieties originating from seven provinces in Indonesia in this study could be essential to constitute the backbone and as a basis for breeding and crop improvement programs. The SSR markers proved their robustness in this study, which are in good agreement with their usage for rice germplasm characterization as part of pre-breeding (Hashimoto *et al.* 2016). The parameters of genetic diversity revealed relatively comparable values with others studies. The number of alleles detected in this study's rice varieties was superior to the 110 alleles reported by Shakil *et al.* (2015) or the other 166 alleles reported by Thomson *et al.* (2009) but lower than the 823 alleles observed in other rice collections (Wang *et al.* 2014). All local rice varieties had higher gene diversity indices than 0.49 across 183 local varieties collected from Kalimantan reported by Thomson *et al.* (2009). A value of 0.68 resulted from 309 local rice varieties collected from 20 provinces in Indonesia. The high gene diversity indices were found in these local varieties from seven provinces, including a minor proportion from Kalimantan, representing their highly diverse. This is important to the extent of available variability of genetic materials for rice improvement that is in good agreement with the previous report (Kumbhar *et al.* 2015).

Moreover, the high heterozygosity in our study suggests the heterogeneous nature of these local varieties, which is relevant to the previous study (Aljumaili *et al.* 2018). Thus, the population size is attributable to genetic differentiation and positively related to heterozygosity (Das *et al.* 2013) in the germplasm of rice genetic resources. In particular, the population's effective size is influenced by the magnitude of the process and the degree of genetic characteristics of a population or genetic resources (Woolfit 2009).

The high PIC value found in these local Indonesian varieties was much higher than those of rice landraces observed by Kumbhar *et al.* (2015) and Das *et al.* (2013), even compared to other landraces/accessions from Malaysia, India, and Pakistan (Aljumaili *et al.* 2018; Behera *et al.* 2012; Wang 2005). Notably, all markers used in this study were highly informative (Botstein *et al.* 1980), suggesting their suitability to differentiate among rice varieties/accessions that would be potentially used to explore the genetic diversity of other local rice varieties in Indonesia.

Genetic diversity analysis in germplasm is important for identifying potentially valuable genetic materials (Wang *et al.* 2014), as demonstrated in this study. The two main clusters generated from 63 local rice varieties revealed a relatively comparable number of varieties. This study demonstrated the distinct cluster of local varieties with SSR markers, relevant to previous studies on rice grouped into two main clusters (Singh *et al.* 2016; Upadhyay *et al.* 2011) and three clusters (Herrera *et al.* 2008) or more clusters. Additionally, this present study suggests that the tendency of local varieties is grouped according to their geographical origin even though there is high divergence among varieties. The geographical corresponded to the varieties with clustering patterns, similar to the previous report (Aljumaili *et al.* 2018).

More respective parameters were determined to explore the genetic diversity of these local Indonesian varieties. Genetic distance and genetic similarity are two parameters that are applied to assess the similarity among two different populations accurately and were estimated by unbiased measures. The conclusion of our study could meet the expectation that the high variation within the population compared to among populations and varieties represents a high magnitude of genetic differentiation that probably can further strengthen the divergence of the local rice population in Indonesia. The genetic differentiation among regions was not significant

in Indonesia. This is in good agreement with the previous report that reproductive biology could be the main factor in determining the genetic structure of plant populations demonstrated by these local varieties from each province with a specific environment. This high proportion of the genetic variability within populations presented in Indonesia rice local varieties indicates that small numbers of populations would support effective conservation compared with island endemics (Sheng *et al.* 2005). Genetic variation among and within populations is correlated with the life history of characters as well as the population history of a species, along with the environments (Hamrick and Godt 1989). Notably, high genetic differentiation is essential in the germplasm for developing new varieties possessing heterotic characteristics in breeding populations (Alam *et al.* 2015). Thus, genetic diversity analysis is useful to provide the basis for multiple purposes such as conservation strategy, utilization, and establishment of breeding and improvement for rice based on local varieties.

Multivariate statistical techniques could be used to analyze multiple measurements on each individual studied simultaneously and are widely used to analyze genetic diversity irrespective of whether it is based on the morphological or molecular markers (Ranjith *et al.* 2019). Cluster analysis, principal component analysis (PCA), and principal coordinate analysis (PCoA) are exceptional techniques among all multivariate techniques. The PCoA of these 63 local varieties showed two main groups are similar to those detected by the UPGMA dendrogram according to SSR polymorphism in this study. This PCoA revealed that high genetic diversity existed in Indonesian rice local varieties from seven provinces, which is in good agreement with the report of several local black rice from other regions in Indonesia (Kristamtini *et al.* 2014). To complement these Indonesian local rice populations, Indian rice germplasm also represented a similar pattern of rice germplasm variation with PCoA analysis to estimate the genetic variation (Choudhury *et al.* 2013).

The PCA in our study also agrees with the previous report that morphological characters with high variability are recommended to provide a high level of gene transfer during plant breeding programs (Gana *et al.* 2013; Nachimuthu *et al.* 2014). In agreement with this study, Maji and Shaibu (2012) studied 123 rice germplasm collected from thirteen villages in Nigeria and reported that the first two principal components together exerted 78% of the total variations observed among rice germplasm through plant height, grain weight, number of productive tillers, days to 50% flowering and number of grains per panicle. A study by Ranjith *et al.* (2019) on 36 rice genotypes also reported that the first two principal components contributed 64.28% of the total variations through productive tiller per hill, plant height, 1000-grain weight, and harvest index from the total genotypic variations observed.

The scattered plot of total varieties in all the quadrants in this study represented the high level of genetic variability among the Indonesian local varieties observed and provides a resource for developing improved varieties in the breeding program. Unlike SSR analysis, which was able to group local varieties mostly from the same regions, this PCA of agro-morphological traits did not show structuralization according to geographical origin, similar to traditional Philippine pigmented rice results (Gaby *et al.* 2019). Thus, it is notable that these agro-morphological characters in our study were useful for studying the genetic variability of local rice varieties as reflected by phenotypic expression. However, it might not always reflect the real genetic variation due to environmental effects, interaction, and unknown genetic control (Smith and Smith 1992). DNA level variation using SSR seems to complement the phenotypic traits and revealed a significant genetic diversity among varieties, within, and among populations.

The genetic structures of populations were reported in rice, including landraces or local varieties from different countries (Gaby *et al.* 2019; Kumbhar *et al.* 2015; Reigh-valiente *et al.* 2016; Upadhyay *et al.* 2012), which exhibited varied clusters. Previous studies revealed that population structure could explain a more apparent grouping than dendrogram with molecular markers. The greater area of color that an individual variety expected, the greater the chance of the individual belonging to the matching group (Tahir *et al.* 2019). Our study showed that the pattern of grouping into two subpopulations is similar to UPGMA dendrogram according to SSR markers. Interestingly, this UPGMA using molecular markers corresponds to one another mainly to their pedigree relation.

In contrast, this population structure can explain relationships better with a high degree

of simulation (Upadhyay *et al.* 2012). This study indicates that varieties in the admixture class may have partial ancestry in breeding history and domestication (Garris *et al.* 2003). These local varieties likely had an intricate history involving inter-crossing (Tahir *et al.* 2019). The relative high of pure varieties at K = 2 in both subpopulations suggests low natural interbreeding. Additionally, the genetic materials in the seed collection are linked to each other; therefore, having a blend between two subpopulations is possible during management.

Acknowledgements

The authors would like to express sincere thanks to the Indonesian Agency for Agricultural Research and Development under the Ministry of Agriculture for financially supporting this research project in the fiscal year 2019.

References

Alam, M., Juraimi, A., Rafii, M., Hamid, A., Arolu, I., Latif, M., 2015. Application of EST-SSR marker in detection of genetic variation among purslane (*Portulaca oleracea* L.) accessions. *Revista Brasileira de Botanica.* 38, 119–129. https://doi.org/10.1007/s40415-014-0103-0

Aljumaili, S.J., Rafii, M.Y., Latif, M.A., Sakimin, S.Z., Arolu, I.W., Miah, G., 2018. Genetic diversity of aromatic rice germplasm revealed by SSR markers. *BioMed Research International.* 1, 1–11. https://doi.org/10.1155/2018/7658032

Behera, L., Patra, B.C., Sahu, R.K., Nanda, A., Sahu, S.C., Patnaik, A., Rao, G.J.N., Singh, O.N., 2012. Assessment of genetic diversity in medicinal rice using microsatellite markers. *Australian Journal of Crop Science.* 6, 1369–1376.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic lingkage map in man using restriction fragment length polymorphism. *American Journal.* 32, 314–331.

Caetano-Anoles, G., Brant, B., 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Nature Biotechnology*. 9, 553–557. https://doi.org/10.1038/nbt0691-553

Carvalho, M.S., Ferreira, M.F.S., Oliveira, W.B.S., Marcal, T.S., Guilhen, J.H.S., Mengarda, L.H.G., Ferreira, A., 2020. Genetic diversity and population structure of *Euterpe edulis* by REML/BLUP analysis of fruit morphology and microsatellite markers. *Crop Breeding and Applied Biotechnology*. 20, e31662048. https://doi.org/10.1590/1984-70332020v20n4a61

Chen, X., Temnykh, S., Xu, Y., Cho, Y.G., McCouch, S.R., 1997. Development of microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theoritical and Applied Genetics.* 95, 553–567. https://doi.org/10.1007/s001220050596

Choudhury, , B., Khan, M., Dayanandan, S., 2013. Genetic structure and diversity of indigenous rice (*Oryza sativa* L.) varieties in the Eastern Himalyan region of Northeast India. *Springer Plus.* 2, 228–238. https://doi.org/10.1186/2193-1801-2-228

Das, B., Sengupta, S., Parida, S.K., Roy, B., Ghosh, M., Prasad, M., Ghose, T.K., 2013. Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. *BMC Genetics*. 14, 1–14. https://doi.org/10.1186/1471-2156-14-71

Doyle, J.J., Doyle, J.L., 1990. No isolation of plant DNA from fresh tissue. *Focus.* 12, 13–15. https://doi.org/10.2307/2419362

Earl, D.A., vonHoldt, B.M., 2012. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7

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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fuller, D.Q., 2011. Pathways to Asian civilizations: tracing the origins and spread of rice and rice cultures. *Rice.* 4, 78–92. https://doi.org/10.1007/s12284-011-9078-7

Gaby, A.N.M, Jones, H., Caguiat, X.G.I, Carandang, S., Ignacio, J.C., Ferrer, M.C., Boyd, L.A., Kretzschmar, T., 2019. Exploring the genetic diversity within traditional Philippine pigmented rice. *Rice*. 12, 1–18. https://doi.org/10.1186/s12284-019-0281-2

Gana, A.S., Shaba, S.Z., Tsado, E.K., 2013. Principal component analysis of morphological traits in thirty-nine accessions of rice (*Oryza sativa* L.) grown in a rainfed lowland ecology of Nigeria. *Journal of Plant Breeding and Crop Science*. 5, 120–126. https://doi.org/10.5897/JPBCS12.065

Garris, A.J., McCouch, S.R., Kresovich, S., 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (*Oryza sativa* L.). *Genetics*. 165, 759–769. https://doi.org/10.1093/genetics/165.2.759

Gross, B.L., Zhao, Z., 2014. Archaeological and genetic insights into the origins of domesticated rice. *PNAS.* 111, 6190–6197. https://doi.org/10.1073/pnas.1308942110

Gonçalves-Vidigal, M.C., Rubiano, L.B., 2011. Development and application of microsatellite in plant breeding. *Crop Breeding and Applied Biotechnology*. 11, 66-72. https://doi.org/10.1590/S1984-70332011000500010

Hamrick, J.L., Godt, M.J.W., 1989. Allozyme diversity in plant species. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L., Weir, B.S. (Eds.). Plant Population Genetics, Breeding, And Genetic Resources. Sunderland: Sinauer. pp. 43-63.

Hashimoto, M., Neriya, Y., Keima, T., Iwabuchi, N., Koinuma, H., Hagiwara-Komoda, Y., Ishikawa, K., Himeno, M., Maejima, K., Yamaji, Y., Namba, S., 2016. EXA1, a GYF domain protein, is responsible for loss-of-susceptibility to plantago asiatica mosaic virus in *Arabidopsis thaliana*. *Plant Journal*. 88, 120–131. https://doi.org/10.1111/tpj.13265

Herrera, T.G., Duque, D.P., Almeida, I.P., Nunez, G.T., Pieters, A.J., Martinez, C.P., Tohme, J.M., 2008. Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeats markers. *Electronic Journal of Biotechnology*. 11, 1–14. https://doi.org/10.2225/vol11-issue5-fulltext-6

Hue, H.T., Nghia, L.T., Minh, H.T., Anh, L.H., Trang, L.T.T., Khan, T.D., 2018. Evaluation of genetic diversity of local -colored rice landraces using SSR markers. *International Letters of Natural Sciences*. 67, 23–34. https://doi.org/10.18052/www.scipress.com/ILNS.67.24

Iskandar, J., Iskandar, B.S., Partasasmita, R., 2018. Review: The impact of social and economic change on domesticated plant diversity with special reference to wet rice field and home-garden farming of West Java, Indonesia. *Biodiversitas*. 19, 565–577. https://doi.org/10.13057/biodiv/d190227

Iyai, D.A., Woran, D., Sumpe, I., 2008. Clustering and principal component analyses of constraints in smallholding pig keeping systems in Manokwari, Indonesia. *Animal Production*. 12, 199–206.

Kristamtini, Taryono, Basunanda, P., Murti, R.H., 2014. keragaman genetik kultivar padi beras hitam lokal berdasarkan penanda mikrosatelit. *AgroBiogen*. 10, 69–76. https://doi.org/10.21082/jbio.v10n2.2014.p69-76

Kumbhar, S.D., Kulwal, P.L., Patil, J.V., Sarawate, C.D., Gaikwad, A.P., Jadhav, A.S., 2015. Genetic diversity and population structure in landraces and improved rice varieties from India. *Rice Science*. 22, 99–107. https://doi.org/10.1016/j.rsci.2015.05.013

Liu, K, Muse, S.V., 2005. PowerMaker: an integrated analysis environment for genetic maker analysis. *Bioinformatics*. 21, 2128–2129. https://doi.org/10.1016/j.rsci.2015.05.013

Maji, A.T., Shaibu, A.A., 2012. Application of principal component analysis for rice germplasm characterization and evaluation. *Journal of Plant Breeding and Crop Science*. 4, 87–93. https://doi.org/10.5897/JPBCS11.093

McCouch, S., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D., Stein, L., 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research*. 9, 257–279. https://doi.org/10.1093/dnares/9.6.257

Ministry of Agriculture. 2018. Optimis produksi beras 2018, kementan pastikan harga beras stabil. Available at: http://www.pertanian.go.id/home/?show=news&act=view&id=. [Date accessed: 24 January 2019]

Nachimuthu, V.V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S., Manonmani, S., 2014. Evaluation of rice genetic diversity and variability in a population panel by principal component analysis. *Indian Journal of Science and Technology*. 7, 1555–1562. https://doi.org/10.17485/ijst/2014/v7i10.14

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. 89, 583–590. https://doi.org/10.1093/genetics/89.3.583

Nurhasanah, Pratama, A.N., Sunaryo, W., 2016. Anther culture of local upland rice varieties from East Kalimantan: effect of panicle cold pre-treatment and putrescine enriched medium. *Biodiversitas*. 17, 148–153. https://doi.org/10.13057/biodiv/d170122

Nurhasanah, Sadaruddin, Sunaryo, W., 2017. Yield-related traits characterization of local upland rice cultivars originated from east and North Kalimantan, Indonesia. *Biodiversitas*. 18, 1165–1172. https://doi.org/10.13057/biodiv/d180339

Panaud, O., Chen, X., McCouch, S.R., 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Molecular and General Genetics. 252, 597–607. https://doi.org/10.1007/BF02172406

Peakall, R., Smouse, P.E., 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*. 28, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Permana, S., 2018. NaSJ Hamlet, Traditional Ecological Knowledge and Conservation of Plant Biodiversity. Plantaxia, Yogyakarta.

Ranjith, P., Sahu, S., Dash, S.K., Bastia, D.N., Pradhan, B.D., 2019. Principal component analysis of rice varieties under bacterial leaf blight incidence. *International Journal of Chemical Studies*. 7, 4184–4187.

Reig-valiente, J.L., Viruel, J., Sales, E., Marqués, L., Terol, J., Gut, M., Derdak, S., Talón, M., Domingo, C., 2016. Genetic diversity and population structure of rice varieties cultivated in temperate regions. *Rice*. 9, 1–12. https://doi.org/10.1186/s12284-016-0130-5

Shakil, S.K., Sultana, S., Hasan, M.M., Hossain, M.M., Ali, M.S., Prodhan, S.H., 2015. SSR marker based genetic diversity analysis of modern rice varieties and coastal landraces in Bangladesh. *Indian Journal of Biotechnology*. 14, 33–41.

Sheng, Y., Zheng, W., Pei, K., Keeping, M., 2005. Genetic variation within and among populations of a dominant desert tree Haloxylon ammodendron (Amaranthaceae) in China. *Annals of Botany*. 2, 245–252. https://doi.org/10.1093/aob/mci171

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. https://doi.org/10.1186/s12863-016-0437-7

Smith, J., Smith, O., 1992. Fingerprinting crop varieties. *Advances in Agronomy*. 47, 85–140. https://doi.org/10.1016/S0065-2113(08)60489-7

Song, S., Tian, D., Zhang, Z., Hu, S., Yu, J., 2019. Rice genomics: over the past two decades and into the future. *Genomics Proteomics Bioinformatics*. 16, 397–404. https://doi.org/10.1016/j.gpb.2019.01.001

Suvi, W.T., Shimelis, H., Mark, L., Mathew, I., Shayanowako, A.I.T., 2019. Assessment of the genetic diversity and population structure of rice genotypes using SSR markers. *Acta Agriculturae Scandinavica*. 70, 76–86. https://doi.org/10.1080/09064710.2019.1670859

Tahir, N., Omar, D., Lateef, D., Ahmad, D., Salih, S., Hiwa, K.L., 2019. Diversity and population structure analysis of faba bean (*Vicia faba* L.) accessions using SSR markers. *Journal of Agricultural Science and Technology*. 21, 463–474.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Temnykh, S., Park, W.D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T., McCouch, S.R., 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*. 100, 697–712. https://doi.org/10.1007/s001220051342

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 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. https://doi.org/10.1007/s12284-009-9023-1

Upadhyay, P., Neeraja, C.N., Kole, C., Singh, V.K., 2012. Population structure and genetic diversity in popular rice varieties of India as evidenced from SSR analysis. *Biochemical Genetics*. 50, 770–783. https://doi.org/10.1007/s10528-012-9519-z

Upadhyay, P., Singh, V.K., Neeraja, C.N., 2011. Identification of genotype specific alleles and molecular diversity assessment of popular rice (*Oryza sativa* L.) varieties of India. *International Journal of Plant Breeding and Genetics*. 5, 130–140. https://doi.org/10.3923/ijpbg.2011.130.140

Wang, J., Jiang, T., Zou, D., Zhao, H., Li, Q., Liu, H., Zhou, C., 2014. Genetic diversity and genetic relationships of Japonica rice varieties in Northeast Asia based on SSR markers. *Biotechnology and Biotechnological Equipment*. 28, 230–237. https://doi.org/10.1080/13102818.2014.908019

Wang, J., 2005. Estimation of effective population sizes from data on genetic markers. *Biological Sciences*. 360, 1395–1409. ttps://doi.org/10.1098/rstb.2005.1682

Woolfit, M. 2009. Effective population size and the rate and pattern of nucleotide substitutions. *Biology Letters*. 5, 417–420. https://doi.org/10.1098/rsbl.2009.0155