Genetic characterization and analysis of the population of selected Sarawak rice cultivars using simple sequence repeat markers

SHAHRIL AB RAZAK *, SHAMSUL AMRI SAIDON, MUHAMMAD FAIRUZ MOHD YUSOF, SITI NORHAYATI ISMAIL

Malaysian Agricultural Research and Development Institute (MARDI), Malaysia

Abstract

Genetic characterization refers to variation detection using genetic markers. In this study, 28 simple sequence repeat (SSR) markers were used to characterize five selected Sarawak rice cultivars, namely Bali1, Bali2, Bali3, Bubok, and Mamut. An analysis based on these 28 polymorphic SSR markers revealed that the number of alleles ranged from 2 (RM312, RM510, and RM536) to 9 (RM307) with an average of 3.86. The polymorphism information content (PIC) values ranged from 0.0175 (RM489) to 0.8174 (RM19) with an average of 0.5179. A dendogram constructed on the basis of the genetic similarity to Nei's 1983 genetic distance showed a high similarity between Bali2 and Bali3, which was further supported by a principal coordinate analysis (PCoA). The population structure analysis illustrated the log likelihood peaking at $K = 2$, indicating that the entire population could be divided into two distinct groups. To develop an SSR fingerprinting panel, 10 SSR markers were selected on the basis of a high PIC value. As this panel provides a better discrimination power in cultivar identification, it could be an efficient and effective tool in cultivar identification, particularly for these selected Sarawak varieties, thus helping breeders to protect their intellectual property rights (IPR). Moreover, the markers could be used to evaluate the seed purity of the selected rice cultivars.

Key word: genetic characterization, Sarawak rice cultivar, simple sequence repeat

Introduction

As rice is one of the main staples and a strategic industry in Malaysia, its production is given the utmost attention by the Malaysian government. The rice industry contributes to economic growth and poverty alleviation socially, economically, and politically; therefore, the availability and the supply of rice are critical for national food security. Malaysia's rice industry production increased at a rate of 1.47% per annum from 2000 to 2013, rising from 2.14 to 2.63 metric tons over this period. While the areas of paddy fields declined from 698702 hectares in 1993 to 688207 hectares in 2013, a steady improvement of the average annual yield from 3.064 tons/hectare to 3.820 tons/hectare was attained between 2000 and 2013. Similarly, boosting the growth rate by 0.26% per annum between 2000 and 2013, Malaysia's production of rice for its own needs improved from 70% to 73.5%, while the total consumption increased as a consequence of the exponential population growth, prompting a considerable increase in rice import between 2000 and 2013 from 596200 tons to 876100 tons (Rosnani, 2015).

To prolong the supply of the national staple food, the Malaysian government has taken proactive measures to improve the rice industry. Entry Point Projects 11 (EPP11) under the National Key Economic Area for the agricultural sector has been implemented to establish large-scale rice cultivation in the non-granary areas. The primary aim of EPP11 is to improve productivity, pest and disease resistance, and drought tolerance of the five Sarawak traditional rice cultivars (Bali1, Bali2, Bali3, Bubok, and Mamut), which were selected because of their distinctive characteristics in terms of aroma, color, and tolerance to drought, hence the need to genetically characterize these five Sarawak traditional rice cultivars.

* Corresponding author: Malaysian Agricultural Research and Development Institute (MARDI), Malaysia; e-mail: shahrilf@mardi.gov.my

Although cultivar identification based on phenotypic traits is still essential for appropriate cultivar identification, identification only on the basis of morphological traits is inaccurate and inconsistent because of the effects of the environmental and phonological factors, and is often restricted by a number of discriminating traits (Arif et al., 2010a). Therefore, molecular markers are more dependable than morphological markers. Molecular markers include SSR or microsatellites: a short repetition of 1- to 6-nucleotide-long or base pair (bp) sequences that exhibit high allelic variations at a given locus (Tautz, 1989; Temnykh et al., 2000). Microsatellites are codominant genetic markers. As reported previously, the positions of the markers mapped in the rice genome displayed a high level of polymorphisms as compared to other plants (Temnykh et al., 2000; Garcia et al., 2004). In the genetic research on rice, microsatellites have been widely utilized because they are multiallelic, enabling the detection and analysis of high levels of allelic diversity (Thomson et al., 2007; Zhao et al., 2009). Many studies on rice fingerprinting have been conducted, including the studies by Singh et al. (1999) and Chakravarthi and Naravaneni (2006).

In this study, we used 28 SSR markers to characterize five selected Sarawak rice cultivars: Bali1, Bali2, Bali3, Bubok, and Mamut. Ultimately, efficient conservation and research management could be achieved by discerning their genetic diversity and population stratification. Finally, to improve cultivar identification and management, we developed an SSR marker panel for fingerprinting the rice cultivars, which could be used for IPR protection. The developed SSR fingerprinting panel could also be used to evaluate the seed purity of the selected cultivars to sustain rice productivity in Malaysia.

Materials and method

Plant materials

The seeds of rice cultivars were collected from Sri Aman, Sarawak. The seeds were planted and self-pollinated for two seasons at the Malaysian Agricultural Research and Development Institute station in Seberang Perai for uniformity purposes. The seeds of Bali1, Bali2, Bali3, Mamut, and Bubok were germinated for two weeks. Leaf samples from 23 individual plants were collected from each cultivar. The morphological characteristics of these plants are summarized in Table 1.

DNA extraction

DNA was extracted using a method described by Mace et al. (2003) with some modifications (in terms of the grinding and the incubation time). Approximately 1 g of the leaf samples was ground using Tissue Lyser (Qiagen, Netherlands) before being incubated at 65°C for 1 h in 600 μl of CTAB extraction buffer (2% CTAB, pH 8, 100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, 0.05% β-mercaptoethanol). The DNA was precipitated in an equal volume of cold isopropanol before being washed twice with 70% ethanol. The DNA pellet was air-dried and re-suspended in 50 μl of TE-RNase buffer. The DNA concentration was measured using Thermo Labsystems Fluoroskan Ascent™ (Thermo Scientific, USA), and a 0.8% agarose gel was used to assess the DNA integrity.

SSR selection and genotyping

The markers for the SSR panel were selected from GRAMENE (http://archive.gramene.org/markers/), consisting of 50 SSR markers evenly distributed throughout the 12 rice chromosomes. Detailed information regarding these selected markers is given in Table 2. Each marker was concatenated with a M13 sequence at its forward primers as described by Schuelke (2000). The PCR was performed in 10 μl reaction mixtures containing 10 × Invitrogen PCR Buffer, 2.5 mM MgCl$_2$, 1 μl genomic DNA, 2 μM dNTPs, 10 μM of each primer pair, and 1 U of Taq polymerase (Invitrogen, California). The amplification was performed using the Applied Biosystem GeneAmp (Thermo Fischer Scientific, California). The PCR profile was set with an initial denaturation at 95°C for 2 min followed by 35 cycles of 1) denaturation for 30 s at 94°C, 2) annealing for 30 s at 45–64°C, and 3) elongation for 45 s at 70°C. The reaction was terminated with a final elongation at 70°C for 5 min. Then, the PCR products were analyzed using an ABI3730xl DNA Analyzer capillary array (Thermo Fischer Scientific, California) and GeneScan™ 500 LIZ (Applied Biosystems, CA) as the DNA ladder.

Scoring and data analyses

The output files from ABI3730 DNA Analyzer were analyzed using GeneMapper5.0 (Applied Biosystems). GS500LIZ was used as a standard ladder. The allele peaks in the electropherograms were scored and analyzed as suggested by Arif et al. (2010b). The scored peaks were then imported into Microsoft Excel and

Table 1. Morpho-agronomic traits of five selected Sarawak rice cultivars

Characters (average)	Bali1	Bali2	Bali3	Bubok	Mamut
Culm height [cm]	172.60	168.60	156.52	131.76	150.77
Panicle length [cm]	34.38	36.10	26.56	34.4	30.63
Flag leaf length [cm]	55.54	56.60	31	58.12	70.23
Flag leaf width [cm]	2.24	2.0	1.32	1.92	1.74
Thousand grain weight [g]	17.2–18.6	17.5–19.0	22.4–24.8	20–23	23–25
Grain length [mm]	6.23 ± 0.25	5.92 ± 0.24	6.24 ± 0.25	5.43 ± 0.14	6.75 ± 0.23
Grain width	2.17 ± 0.08	2.26 ± 0.08	2.14 ± 0.12	1.80 ± 0.05	1.96 ± 0.05
L/W Ratio	2.87	2.62	2.92	3.02	3.44
Amylose [%]	13.9	14.4	14.4	19.5	18.1
Aromatic	strong	strong	strong	strong	nil or slight
Coloured grain	coloured	coloured	coloured	white	coloured

formatted using the appropriate data input files. POWER-MARKER (Liu and Muse, 2005) was used to calculate the allele number, heterozygosities, PIC value, and Nei's 1983 genetic distance. Then, Nei's 1983 genetic distance was used as an input to generate UPGMA dendograms using MEGA7 (Kumar et al., 2016). The Principal coordinate analysis (PCoA) and the F_{ST} values were calculated and analyzed using GenAlEx 6.501 (Peakall and Smouse, 2006). Structure 2.3.4 (Pritchard et al., 2000; Falush et al., 2003) was used to investigate the population structure of these selected cultivars on the basis of the Bayesian clustering model. To determine the exact K value, the range of K was set from 1 to 5; the run length was set with 100000 generations as burn-in, and followed by 100000 Markov-Chain Monte Carlo (MCMC) replicates. The true number of populations was predicted using two approaches: nonparametric test by comparing LnP(D) for each value of K as described by Pritchard et al. (2000), and calculation of $\Delta K (\Delta K = m([L " K])/s[L(K)])$ as described by Evanno et al. (2005).

Results and discussions

Microsatellite characterization

SSR markers proved to be reliable markers for the genetic study of a species, particularly for cultivar identification. After testing 50 SSR markers, we obtained 28 polymorphic markers; the remaining were monomorphic, were difficult to optimize, produced an unspecific peak pattern, or were unsuccessful at amplifying more than 80% of the samples. The average number of alleles per locus, ranging from 2 (RM312, RM510, and RM536) to 9 (RM307), was 3.86. The low number of alleles was expected as only five cultivars (Bali1, Bali2, Bali3, Bubok, and Mamut) were evaluated in this study. Fourteen SSR markers (50.0%) from the total of 28 amplified three alleles. Next, the PIC value was applied to evaluate the degree of information of each polymorphism in the studied cultivars. The PIC value ranged from 0.0175 (RM489) to 0.8174 (RM19) with an average of 0.5179. Botstein et al. (1980) categorized the PIC value for codominant markers into three categories: highly informative (PIC > 0.5), reasonably informative (0.5 > PIC > 0.25), and slightly informative (PIC < 0.25). These categories determined that from the 28 SSR markers giving positive results in this study, 14 (50%) were classified as highly informative, 13 (46.4%) were reasonably informative, and only one (3.6%) was slightly informative. As proposed by Botstein et al. (1980), the sites having a PIC value of close to 1.0 and producing many alleles are the most desirable. The heterozygosities and genetic diversity ranged from 0.2087 to 0.9912 and from 0.0176 to 0.8381 with an average of 0.2210 and 0.4874, respectively. The low value of heterozygosity and gene diversity suggested that the studied cultivars were not diverse. Moreover, the low number of cultivars studied contributed to the low values of heterozygosity and diversity. The characterization of the 28 SSR markers is summarized in Table 3.

Genotype characterization and population analyses

A dendogram (Fig. 1) generated on the basis of Nei's 1983 genetic distance (Table 4) showed that the culti-

Table 2. List and information of SSR markers used in this study.
The information was obtained from GRAMENE (http://archive.gramene.org/markers/)

SSR ID	Chromosomes	Primer sequence	Repeat motif	Expected size [bp]
RM495	1	F: aatccaaggtgcagagatgg R: caacgatgacgaacacaacc	(CTG)7	148–160
RM1	1	F: gcgaaaacacaatgcaaaaa R: gcgttggttggacctgac	(GA)26	67–119
RM283	1	F: gtctacatgtacccttgttggg R: cggcatgagagtctgtgatg	(GA)18	130–176
RM259	1	F: tggagtttgagaggaggg R: cttgttgcatggtgccatgt	(CT)17	133–186
RM312	1	F: gtatgcatatttgataagag R: aagtcaccgagtttaccttc	(ATTT)4(GT)9	86–106
RM5	1	F: tgcaacttctagctgctcga R: gcatccgatcttgatggg	(GA)14	94–138
RM237	1	F: caaatcccgactgctgtcc R: tgggaagagagcactacagc	(CT)18	105–153
RM431	1	F: tcctgcgaactgaagagttg R: agagcaaaaccctggttcac	(AG)16	233–261
RM154	2	F: accctctccgcctcgcctcctc R: ctcctcctcctgcgaccgctcc	(GA)21	148–230
RM452	2	F: ctgatcgagagcgttaagggg R: gggatcaaaccacgtttctg	(GTC)9	192–213
RM489	3	F: acttgagacgatcggacacc R: tcacccatggatgttgtcag	(ATA)8	223–289
OSR13	3	F: catttgtgcgtcacggagta R: agccacagcgcccatctctc	(GA)n	85–122
RM338	3	F: cacaggagcaggagaagagc R: ggcaaaccgatcactcagtc	(CTT)6	178–184
RM55	3	F: ccgtcgccgtagtagagaag R: tcccggttattttaaggcg	(GA)17	216–247
RM514	3	F: agattgatctcccattcccc R: cacgagcatattactagtgg	(AC)12	229–278
RM307	4	F: gtactaccgacctaccgttcac R: ctgctatgcatgaactgctc	(AT)14(GT)21	116–191
RM124	4	F: atcgtctgcgttgcggctgctg R: catggatcaccgagctcccccc	(TC)10	257–289
RM507	5	F: cttaagctccagccgaaatg R: ctcaccctcatcatcgcc	(AAGA)7	234–257
RM413	5	F: ggcgattcttggatgaagag R: tccccaccaatcttgtcttc	(AG)11	71–114
RM161	5	F: tgcagatgagaagcggcgcctc R: tgtgtcatcagacggcgctccg	(AG)20	154–187
RM178	5	F: tcgcgtgaaagataagcggcgc R: gatcaccgttccctccgcctgc	(GA)5(AG)8	112–131
RM334	5	F: gttcagtgttcagtgccacc R: gactttgatctttggtggacg	(CTT)20	119–207
RM133	6	F: ttggattgttttgctggctcgc R: ggaacacggggtcggaagcgac	(CT)8	226–237
RM510	6	F: aaccggattagtttctcgcc R: tgaggacgacgagcagattc	(GA)15	99–127

RM454	6	F: ctcaagcttagctgctgctg R: gtgatcagtgcaccatagcg	(GCT)8	249–292
RM162	6	F: gccagcaaaaccagggatccgg R: caaggtcttgtgcggcttgcgg	(AC)20	191–244
RM125	7	F: atcagcagccatggcagcgacc R: aggggatcatgtgccgaaggcc	(GCT)8	105–147
RM11	7	F: tctcctcttcccccgatc R: atagcgggcgaggcttag	(GA)17	118–151
RM455	7	F: aacaacccaccacctgtctc R: agaaggaaaagggctcgatc	(TTCT)5	127–144
RM118	7	F: ccaatcggagccaccggagagc R: cacatcctccagcgacgccgag	(GA)8	149–165
RM408	8	F: caacgagctaacttccgtcc R: actgctacttgggtagctgacc	(CT)13	112–128
RM152	8	F: gaaaccaccacacctcaccg R: ccgtagaccttcttgaagtag	(GGC)10	133–157
RM25	8	F: ggaaagaatgatcttttcatgg R: ctaccatcaaaaccaatgttc	(GA)18	121–159
RM44	8	F: acgggcaatccgaacaacc R: tcgggaaaacctaccctacc	(GA)16	82–132
RM284	8	F: atctctgatactccatccatcc R: cctgtacgttgatccgaagc	(GA)8	139–159
RM433	8	F: tgcgctgaactaaacacagc R: agacaaacctggccattcac	(AG)13	216–248
RM447	8	F: cccttgtgctgtctcctctc R: acgggcttcttctccttctc	(CTT)8	95–146
RM316	9	F: ctagttgggcatacgatggc R: acgcttatatgttacgtcaac	(GT)8–(TG)9(TTTG)4(TG)4	194–216
RM105	9	F: gtcgtcgacccatcggagccac R: tggtcgaggtggggatcgggtc	(CCT)6	100–141
RM215	9	F: caaaatggagcagcaagagc R: tgagcacctccttctctgtag	(CT)16	126–161
RM474	10	F: aagatgtacgggtggcattc R: tatgagctggtgagcaatgg	(AT)13	216–288
RM271	10	F: tcagatctacaattccatcc R: tcggtgagacctagagagcc	(GA)15	80–120
RM171	10	F: aacgcgaggacacgtacttac R: acgagatacgtacgcctttg	(GATG)5	307–347
RM484	10	F: tctccctcctcaccattgtc R: tgctgccctctctctctctc	(AT)9	286–298
RM552	11	F: cgcagttgtggatttcagtg R: tgctcaacgtttgactgtcc	(TAT)13	167–258
RM536	11	F: tctctcctcttgtttggctc R: acacaccaacacgaccacac	(CT)16	223–247
RM287	11	F: ttccctgttaagagagaaatc R: gtgtatttggtgaaagcaac	(GA)21	82–118
RM144	11	F: tgccctggcgcaaatttgatcc R:gctagaggagatcagatggtagtgcatg	(ATT)11	216–295
RM19	12	F: caaaaacagagcagatgac R: ctcaagatggacgccaaga	(ATC)10	192–250
RM277	12	F: cggtcaaatcatcacctgac R: caaggcttgcaagggaag	(GA)11	104–121

Table 3. Characteristics of 28 polymorphic SSR markers with respect to the selected Sarawak rice cultivars

No	Marker	Major allele freq	Number of aleles	Gene dversity	PIC
1	RM1	0.4000	4	0.6944	0.6379
2	RM283	0.5565	3	0.5284	0.4275
3	RM312	0.6000	2	0.4800	0.3648
4	RM5	0.4000	3	0.6400	0.5632
5	RM237	0.6000	3	0.5600	0.4992
6	RM154	0.4107	4	0.6666	0.6004
7	RM489	0.9912	3	0.0176	0.0175
8	RM55	0.3026	5	0.7211	0.6677
9	RM338	0.7544	3	0.3868	0.3366
10	RM307	0.5045	9	0.6591	0.6112
11	RM124	0.5217	3	0.5617	0.4693
12	RM334	0.5957	3	0.4852	0.3716
13	RM413	0.5826	3	0.5003	0.3913
14	RM178	0.5965	3	0.5628	0.5015
15	RM133	0.6842	6	0.4946	0.4594
16	RM510	0.6000	2	0.4800	0.3648
17	RM162	0.4130	3	0.6344	0.5567
18	RM125	0.4000	5	0.7217	0.6745
19	RM408	0.6000	3	0.5600	0.4992
20	RM25	0.4043	3	0.6575	0.5835
21	RM319	0.3000	6	0.7784	0.7442
22	RM215	0.3652	4	0.6981	0.6410
23	RM484	0.2719	6	0.8043	0.7758
24	RM474	0.6000	3	0.5600	0.4992
25	RM552	0.3174	4	0.7394	0.6912
26	RM536	0.6000	2	0.4800	0.3648
27	RM277	0.6000	3	0.4834	0.3707
28	RM19	0.2087	7	0.8381	0.8174
Mean		0.5065	3.86	0.5855	0.5179

Table 4. Nei's 1983 pairwise genetic distance of the selected Sarawak rice cultivars based on 28 polymorphic SSR markers

	Bali1	Bali2	Bali3	Bubok	Mamut
Bali1	0.0000				
Bali2	0.1875	0.0000			
Bali3	0.1960	0.1271	0.0000		
Bubok	0.8617	0.8635	0.8597	0.0000	
Mamut	0.9643	0.9651	0.9643	0.4755	0.0000
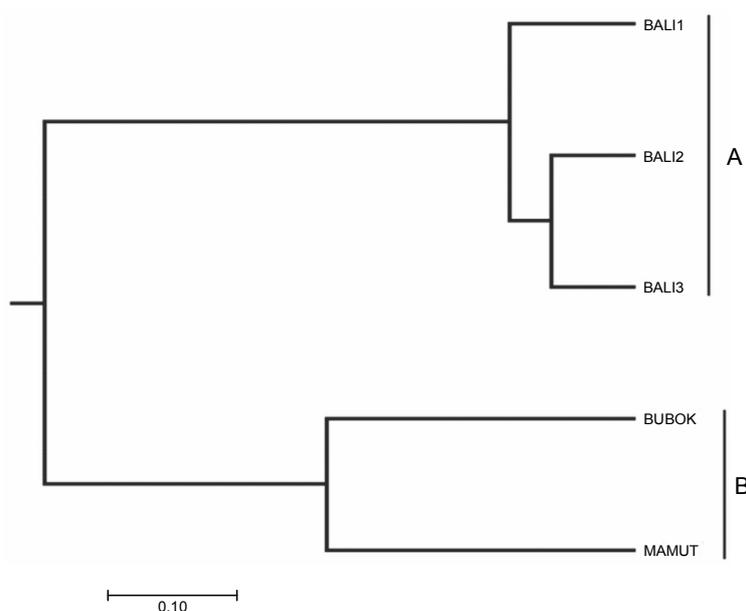
Fig. 1. Unweighted pair group method with an arithmetic mean (UPGMA) dendogram
based on the 28 polymorphic SSR markers
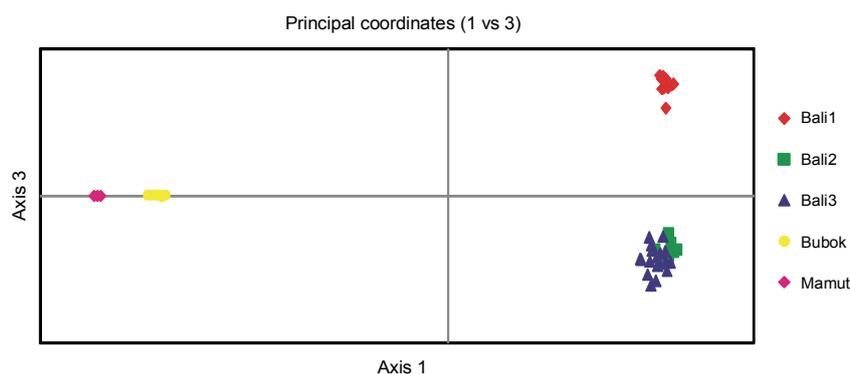
Fig. 2. PCoA of selected Sarawak rice cultivars by GenAlEx 6.501
based on 28 polymorphic SSR markers

vars could be grouped into two major groups, A and B. Group A consisted of Bali1, Bali2, and Bali3, while group B consisted of Bubok and Mamut. The pairwise genetic distance value also showed the lowest genetic distance between Bali2 and Bali3 with a value of 0.1271, suggesting a high allelic similarity between these two cultivars. Despite a high similarity between the cultivars, the SSR marker could differentiate between the two cultivars (Table 4). Morphologically, these two cultivars also showed a high similarity in agronomic traits; however, a morphology-based identification might be difficult to conduct and requires a considerable amount of time as the plant must be matured before the identification. The analysis also showed the highest genetic distance be-

tween Mamut and Bali2 with a value of 0.9561 even though both of them are colored rice. The results of the PCoA analysis (Fig. 2) supported the dendogram analysis in that the selected cultivars were clustered in two major groups. It also showed a high correlation between Bali2 and Bali3 (Fig. 2).

The population structure among the cultivars was assessed with the Bayesian approach using the STRUCTURE software. The log-likelihood values obtained from the STRUCTURE analysis showed that the exact number of K was 2, indicating that all the individuals could be grouped into two genotype groups (Fig. 3). The diagram presented in Figure 4 shows that all the Bali variants were grouped together. According to the Wright

A

B
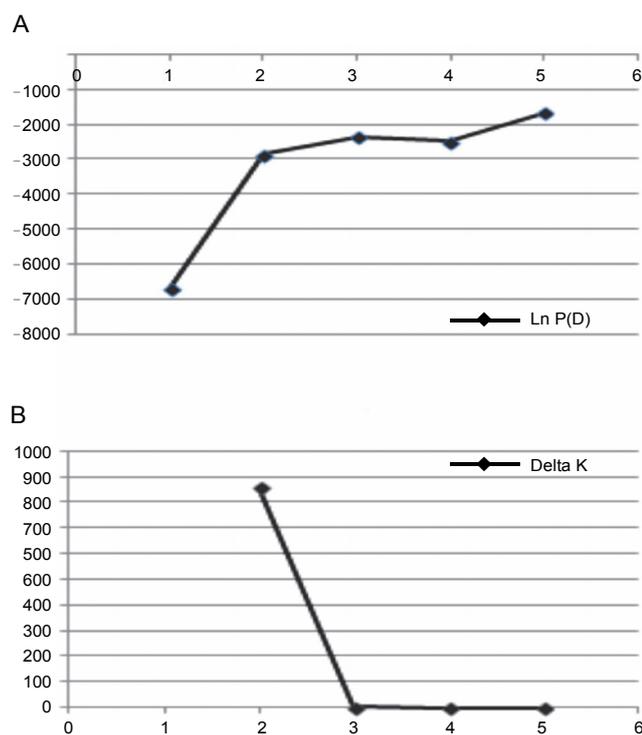
Fig. 3. Estimation of the number of populations, K where $K = 2$ based on A) LnP(D) value where K corresponds to the LnP(D) value when it starts to plateau and B) ΔK ($\Delta K = m([L \text{ “ } K])$ / $s[L(K)]$) where K refers to the highest value of ΔK
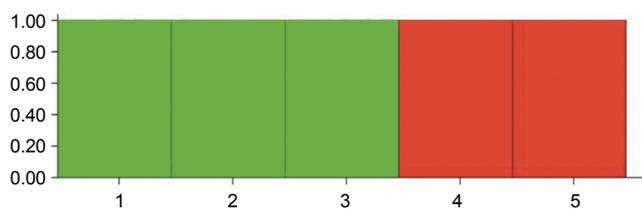
Fig. 4. Genetic structure obtained using the Structure program with $K = 2$ as determined by LnP(D) and ΔK based on 28 polymorphic SSR markers

Table 5. Pairwise F_{ST} value of the selected Sarawak rice cultivars calculated by Genalex 6.501 based on 28 polymorphic SSR markers

	Bali1	Bali2	Bali3	Bubok	Mamut
Bali1	0.000				
Bali2	0.259	0.000			
Bali3	0.263	0.198	0.000		
Bubok	0.739	0.707	0.694	0.000	
Mamut	0.871	0.832	0.822	0.686	0.000

Table 6. Nei's 1983 pairwise genetic distance of the selected Sarawak rice cultivars based on 28 polymorphic SSR markers

	Bali1	Bali2	Bali3	Bubok	Mamut
Bali1	0.0000				
Bali2	0.4186	0.0000			
Bali3	0.3436	0.2500	0.0000		
Bubok	0.9517	0.9504	0.9517	0.0000	
Mamut	1.0000	1.0000	1.0000	0.7041	0.0000

(Table 5). This result suggested a high gene flow between cultivars Bali1 and Mamut. The calculated F_{ST} also suggested that the analysis using the 28 polymorphic SSR markers could successfully differentiate the five tested cultivars.

Development of SSR fingerprinting panel

The purpose of developing an SSR fingerprinting panel was to provide an effective and efficient approach to cultivar identification at minimal cost. To determine the best SSR combinations for the panel development, the SSR markers were ordered according to their PIC value, and then, markers having a value of more than 0.6 were selected for the development of the SSR fingerprinting panel. In all, 10 SSR markers (RM1, RM19, RM55, RM125, RM154, RM215, RM307, RM319, RM484, and RM552) had a PIC value of more than 0.6 with RM19 having the highest PIC value of 0.8174. The genetic distance analysis using this set of markers proved that this panel had more discrimination power than the analysis using 28 polymorphic SSR markers. Using 28 SSR markers, we found that the pairwise genetic distance for Bali2 and Bali3 was 0.1271, but using the developed panel, we obtained the pairwise genetic

(1978) guideline for the interpretation of the Fst value for population differentiation, values ranging from 0 to 0.05, from 0.05 to 0.15, from 0.15 to 0.25, and 0.25 and above indicate low, moderate, high, and very high genetic differentiation between populations, respectively. In this study, the obtained pairwise Fst value describing the genetic population of the tested population ranged from 0.198 (Bali2–Bali3) to 0.871 (Bali1–Mamut). Hence, in this study, only the Bali2 and Bali3 populations had high genetic differentiation between them, while the other pairwise Fst values demonstrated a very high genetic differentiation between the populations

distance of 0.2500 for Bali2 and Bali3 (Table 6). Hence, this panel provided a better means for the cultivar identification of the selected Sarawak cultivars than the use of the 28 SSR markers. Moreover, it minimized the required cost and time. There have been reports of DNA fingerprinting in rice using SSR markers (Sarao et al., 2010; Zhu et al., 2012), but the set of plant material or cultivars was definitely different from that used in the present study. SSR markers have become the markers of choice because of their highly polymorphic, reproducible, multi-allelic, and co-dominant molecular marker characteristics (Miah et al., 2013; Vieira et al., 2016). These traits have enabled the use of SSR markers for the identification of very similar rice varieties.

Conclusions

An SSR fingerprinting panel for high-similarity rice varieties, which are very difficult to identify using morphological characteristics, was developed. This panel provided a better approach to rice cultivar identification as the identification process could be carried out even when the plant was in the seedling stage, than the morphology-based identification where the plant needs to be matured before its identification. Therefore, we believe that the developed SSR fingerprinting panel, which consists of high-PIC-value SSR markers, is the best approach to the identification of closely related cultivars and the evaluation of seed purity.

Acknowledgments

References

Arif I.A., Bakir M.A., Khan H.A., Al Farhan A.H., Al Homaidan A.A., Bahkali A.H., Al Sadoon M., Shobrak M. (2010a) *Application of RAPD for molecular characterization of plant species of medicinal value from an arid environment.* Genet. Mol. Res. 9: 2191–9198.

Arif I.A., Khan H.A., Shobrak M., Al Homaidan A.A., Al Sadoon M., Al Farhan A.H., Bahkali A.H. (2010b) *Interpretation of electrophoretograms of seven microsatellite loci to determine the genetic diversity of the Arabian Oryx.* Genet. Mol. Res. 9: 259–265.

Botstein D., White R.L., Skolnick M., Davis R.W. (1980) *Construction of a genetic linkage map in man using restriction fragment length polymorphisms.* Am. J. Hum. Genet. 32: 314–331.

Chakravarthi B. K., Naravaneni R. (2006) *SSR marker based DNA fingerprinting and diversity study in rice (Oryza sativa. L).* Afr. J. Biotechnol. 5: 684–688.

Evanno G., Regnaut S., Goudet J. (2005) *Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.* Mol. Ecol. 14: 2611–2620.

Falush D., Stephens M., Pritchard J.K. (2003) *Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies.* Genetics 164: 1567–1587.

Garcia A.A., Benchimol L.L., Barbosa A.M., Geraldi I.O., Souza Jr C.L., Souza A.P.D. (2004) *Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines.* Genet. Mol. Biol. 27: 579–588.

Kumar S., Stecher G., Tamura K. (2016) *MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets.* Mol. Biol. Evol. 33: 1870–1874.

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

Lu H., Redus M.A., Coburn J.R., Rutger J.N., McCouch S.R., Tai T.H. (2005) *Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis.* Crop. Sci. 45: 66–76.

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 Mol. Biol. Report. 21: 459–460.

Miah G., Rafii M.Y., Ismail M.R., Puteh A.B., Rahim H.A., Islam K.N., Latif M.A. (2013) *A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance.* Int. J. Mol. Sci., 14: 22499–22528.

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

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

Rosnani H. (2015) *Policies and economic development of rice production in Malaysia.* FFTC Agricultural Policy Articles. Taipei. Taiwan.

Sarao N.K., Vikal Y., Singh K., Joshi M.A., Sharma R.C. (2010) *SSR marker-based DNA fingerprinting and cultivar identification of rice (Oryza sativa L.) in Punjab state of India.* Plant Genetic Resour. 8: 42–44.

Tautz D. (1989) *Hypervariabflity of simple sequences as a general source for polymorphic DNA markers.* Nucl. Acids Res. 17: 6463–6471.

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 se-*

quences in rice (Oryza sativa L.). Theor. Appl. Genet. 100: 697–712.

Thomson M.J., Septiningsih E.M., Suwardjo F., Santoso T.J., Silitonga T.S., McCouch S.R. (2007) *Genetic diversity analysis of traditional and improved Indonesian rice (Oryza sativa L.) germplasm using microsatellite markers*. Theor. Appl. Genet. 114: 559–568.

Tu M., Lu B.R., Zhu Y., Wang Y. (2007) *Abundant within-varietal genetic diversity in rice germplasm from Yunnan Province of China revealed by SSR fingerprints*. Biochem. Genet. 45: 789–801.

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*. Genet. Mol. Biol. 39: 312–328.

Wright S. (1978) *Evolution and the genetics of population. Variability within and among natural populations*. The University of Chicago Press, Chicago.

Zhao W., Chung J.W., Ma K.H., Kim T.S., Kim S.M., Shin D.I., Kim C.H., Koo H.M., Park Y.J. (2009) *Analysis of genetic diversity and population structure of rice cultivars from Korea, China and Japan using SSR markers.* Genes Genom. 31: 283–292.

Zhu Y.F., Qin G.C., Hu J., Wang Y., Wang J.C., Zhu S.J. (2012) *Fingerprinting and variety identification of rice (Oryza sativa L.) based on simple sequence repeat markers*. Plant Omics. 5: 421–426.