https://doi.org/10.17221/58/2019-CJGPB

Genetic diversity of released Malaysian rice varieties based on single nucleotide polymorphism markers

Shahril Ab Razak[1], Nor Helwa Ezzah Nor Azman[1], Rahiniza Kamaruzaman[2], Shamsul Amri Saidon[2], Muhammad Fairuz Mohd Yusof[1], Siti Norhayati Ismail[1], Mohd Azwan Jaafar[1], Norzihan Abdullah[1]*

[1]*Agri-Omic & Bioinformatic Program, Biotechnology and Nanotechnology Research Centre, MARDI Headquarter, Serdang, Selangor, Malaysia*

[2]*Rice Breeding Program, Rice Research Centre, MARDI Seberang Perai, Kepala Batas, Pulau Pinang, Malaysia*

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

Citation: Ab Razak S., Nor Azman N.H.E., Kamaruzaman R., Saidon S.A., Mohd Yusof M.F., Ismail S.N., Jaafar M.A., Abdullah N. (2020): Genetic diversity of released Malaysian rice varieties based on single nucleotide polymorphism markers. Czech J. Genet. Plant Breed., 56: 62–70.

Abstract: Understanding genetic diversity is a main key for crop improvement and genetic resource management. In this study, we aim to evaluate the genetic diversity of the released Malaysian rice varieties using single nucleotide polymorphism (SNP) markers. A total of 46 released Malaysian rice varieties were genotyped using 1536 SNP markers to evaluate their diversity. Out of 1536 SNPs, only 932 SNPs (60.7%) represented high quality alleles, whereas the remainder either failed to amplify or had low call rates across the samples. Analysis of the 932 SNPs revealed that a total of 16 SNPs were monomorphic. The analysis of the SNPs per chromosome revealed that the average of the polymorphic information content (PIC) value ranged from 0.173 for chromosome 12 to 0.259 for chromosome 11, with an average of 0.213 per locus. The genetic analysis of the 46 released Malaysian rice varieties using an unweighted pair group method with arithmetic mean (UPGMA) dendrogram revealed the presence of two major groups. The analysis was supported by the findings from the STRUCTURE analysis which indicated the ΔK value to be at the highest peak at K = 2, followed by K = 4. The pairwise genetic distance of the shared alleles showed that the value ranged from 0.000 (MR159↔MR167) to 0.723 (MRIA↔Setanjung), which suggested that MR159 and MR167 were identical, and that the highest dissimilarity was detected between MRIA 1 and Setanjung. The results of the study will be very useful for the variety identification, the proper management and conservation of the genetic resources, and the exploitation and utilisation in future breeding programmes.

Keywords: DNA marker; genetic relationship; *Oryza sativa*

Genetic diversity is crucial for the improvement of numerous crop plants, including rice. Breeders lean on the accessibility of the genetic diversity information to design their breeding programme in order to predict the genetic gain and heterosis level. A variety established with a broad genetic base can be useful to boost the plant performance, produce a high level of resistance against diseases and pests, and increase yield production under innumerable agro-climatic environments (Zhu et al. 2000). Studying the genetic variation is very crucial for the varietal characterisation and identification, for the appropriate seed purity management for conservation and breeding purposes, and is even useful for plant varietal protection (PVP) purposes. Pejic et al. (1998) state that the genetic diversity

Supported by the Malaysian government (NKEA EPP14).

https://doi.org/10.17221/58/2019-CJGPB

can be measured using several strategies such as a pedigree analysis, a morphological characterisation, and an analysis of the molecular markers. However, a pedigree analysis is impractical and unrealistic for measuring the genetic diversity of a target crop (Fufa et al. 2005). In addition, diversity based on the morphological characters produce unpredictable and inconsistent data since the morphological characters could be influenced by environmental conditions and agronomic practices (Marić et al. 2004).

DNA markers offer a preferred method to evaluate the genetic diversity and variation since the markers are not influenced by environmental factors or agronomic practices, are highly abundant in the genome, and are not obliged to previous lineage information (Bohn et al. 1999). Single nucleotide polymorphisms (SNPs) are one of the molecular markers that, in recent times, attained popularity owing to advancements in the field of genomics, particularly next-generation sequencing (NGS) platforms. NGS has revolutionised the molecular marker system where it has driven a shift from anonymous markers, such as random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR), to the direct evaluation of the sequence variation, including SNP genotyping. SNPs are glamorous markers in many ways, including their low rate of scoring errors, since no stutter bands appear in the SNPs like they do in SSRs. Furthermore, SNP markers are also amenable to high-throughput genotyping since the markers can be multiplexed depending on the genotyping platform, which leads to low operation costs per unit of data generated. Additionally, SNP high-throughput genotyping is also suitable for poor quality or degraded samples, such as deteriorated or ancient samples (Morin & McCarthy 2007; Helyar et al. 2011). Though showing superiority, in terms of informative level, the SNP marker was still below that of the SSR marker (Rosenberg et al. 2003). Despite this, more current studies demonstrated that SNPs can provide a lot of information for population structure analyses, outperforming the SSR markers (Liu et al. 2005). Even though the SSR markers were generally multiallelic at each locus, the individual SNPs were able to strongly segregate among the populations (Freamo et al. 2011).

In this study, we intend to resolve the diversity of the released Malaysian rice varieties using SNP markers. To date, more than 40 varieties have been released since 1952. The findings of this study will be very useful for breeders to strategize their breeding programme in order to predict the genetic gain. In addition, the results will also be very useful for conservation management programmes.

MATERIAL AND METHODS

Plant materials and DNA extraction. The seeds from each of the 46 released Malaysian rice varieties were germinated for approximately three weeks for DNA extraction purposes. The details of the varieties are summarised in Table 1. Young fresh leaves from three uniform individual plants of each variety were collected. Small fragments of the young leaves were transferred into a 96-well plate containing stainless steel beads and were instantly frozen at –80 °C for a minimum of one night. The frozen tissue was ground using a Tissue Lyser (Qiagen, Hilden, Germany) immediately after the addition of 600 µl of an extraction buffer (2% CTAB, 100 mM Tris-HCl pH8, 20 mM EDTA, 1.4 M NaCl, 0.05% β-mercaptoethanol). The genomic DNA was extracted following the protocol as described by Mace et al. (2003). The DNA integrity was assessed using 0.8% agarose gels and the DNA concentration was measured using a Fluoraskan Ascent (Thermo Fisher Scientific, Waltham, United States). The DNA was diluted as recommend for the Illumina Golden Gate Genotyping Technology (GGGT).

SNP selection and genotyping. SNPs were mined from the genomes of three varieties, namely Nipponbarre (Japonica), 93-11 (Indica), and Indica 1 (Habibuddin et al. 2013). A total of 1 536 SNPs were selected and used to construct an Illumina bead array based on the GoldenGate assay (Illumina Inc., California, United States). The SNPs were chosen based on three criteria: (1) they were evenly spaced in the rice genome with approximately 1 SNP/150 kb, (2) they were preferably in the genic regions, and (3) there were no other SNPs within 70 bp. Genotyping was performed using the Illumina Golden Gate Genotyping Technology (GGGT). The identified SNPs along with their 100 bp upstream and downstream flanking sequences were submitted to Illumina for construction of custom made beadchip assays.

The SNP genotyping was conducted follow the standard manufacturer's protocol (Illumina's BeadArray Express Reader). The hybridised custom-made beadchips were scanned using the Illumina iScan system. The generated raw data were loaded into GenomeStudio Software (Ver. 2011.1, Illumina) in

https://doi.org/10.17221/58/2019-CJGPB

Table 1. The list of the rice varieties used in this study and their respective information

No.	Variety	Release year	Parent	Type
1	Malinja	1964	Siam 29 × Pebifun	white rice
2	Mahsuri	1965	Mayang Ebos 80 × Taichung 65	white rice
3	Ria (Rebranding of IR8)	1966	Peta × Dee-Geo-Woo_Gen	white rice
4	Bahagia	1968	Peta × Tangkai Rotan	white rice
5	Murni	1972	Bahagia × Ria	white rice
6	Masria	1972	IR8 × Muey Nahng 62M	white glutinous
7	Jaya	1973	rebranding of C4-63	white rice
8	Sri Malaysia 1	1974	Peta × Tangkai	white rice
9	Sri Malaysia II	1974	Ria × Pankhari 203	white rice
10	Pulut Malaysia I	1974	Pulut Sutera × Ria	white glutinous
11	Setanjung	1979	IR22 × Pazudofusu	white rice
12	Sekencang	1979	Jaya × Tadukan	white rice
13	Sekembang	1979	Seribu Gantang × Ria 163	white rice
14	Kadaria	1981	(Seribu Gantang × TKM-6)///TKM-6	white rice
15	Pulut Siding	1981	Pulut Sutera × Ria	white glutinous
16	Manik	1984	(Radin × Tadukan)///Radin Goi	white rice
17	Muda	1984	RU 243 × BRJ51	white rice
18	Seberang	1984	MR 50 × IR 4215	white rice
19	Makmur	1985	Setanjung × Pongsu Seribu	white rice
20	MR84	1986	CR261-7039-236 × MR 50	white rice
21	MR81	1988	MR24 × IR36	white rice
22	MR103	1990	RU 1217-432 × RU 1378-24-4	white rice
23	MR106	1990	(MR71 × IR 21912-131)/MR71	white rice
24	PH9	1990	MR23 × PULUT HITAM SIAM	black glutinous
25	MR123	1991	Y776 × Y680	white rice
26	MR127	1991	Setanjung, Sekencang, Muda	white rice
27	MR159	1995	Y833 × IR5491	white rice
28	MR167	1995	Y978/PTB18//Muda	white rice
29	MR185	1995	Y1056 × MR133	white rice
30	MR211	1999	MR84 × Hoshiyutaka	white rice
31	MRQ50	1999	Q34 × KDML	white aromatic
32	MR219	2001	MR151 × MR137	white rice
33	MR220	2003	MR151 × MR137	white rice
34	MRQ74	2005	Q34 × KDML ///Kasturi	white aromatic
35	MR232	2006	W60 × Y1157	white rice
36	MR220CL1	2010	IMI-TR-1770 × MR220	clearfield white rice
37	MR220CL2	2010	IMI-TR-1770 × MR220	clearfield white rice
38	MR253	2010	PTB 33 × SPM 92	white rice
39	MR263	2010	SPM 156 × MR221	white rice
40	MRQ76	2012	Q72 × Cuicak Wangi	white aromatic
41	MR269	2012	P347 × Y1362	white rice
42	MR284	2015	ER3070 × MR220	white rice
43	Padi MARDI Siraj 297 (MR297)	2017	(MRQ76 × P446)/P446	white rice
44	MARDI Sempadan (MR303)	2018	(MR256 × MR253)/MR256	white rice
45	MARDI Sebernas (MR307)	2018	MR256 × P493	white rice
46	MRIA 1	2014	mutation of IR76569-259-1-2-1	aerobic rice

https://doi.org/10.17221/58/2019-CJGPB

order to call the allele at each locus. The cluster calls of each SNP were assigned to all the samples and were manually checked for an error before being rescored to designate the heterozygous and homozygous clusters. Then, the SNP reliability was assessed by analysing the call rates, the GenTrain and GenCall scores as generated by GenomeStudio software (Ver. 2011.1). As suggested by Illumina, a call rate of 95 was used as the threshold value for reliable SNPs, and a value of 0.4 was set as the threshold value for the GenTrain and GenCall scores.

Data analysis. The allele scoring data were used as the input for PowerMarker (Ver. 3.25; Liu & Muse 2005) in order to calculate the allele numbers, the polymorphic information content (PIC) values, the genetic diversity, the observed heterozygosities, and the maximum allele frequency of each SNP. Additionally, the pairwise genetic distance of each variety was also calculated using the same program. STRUCTURE (Ver. 2) software (Pritchard et al. 2000) was used to assign the subpopulations of the genotypes. The assignment involved the calculation of the K (population number) values by varying the K values from 1 to 10 with 20 independent runs per K value, with a 50 000 burn-in period and 100 000 Markov chain Monte Carlo (MCMC) repetitions. The optimal value of K was calculated using STRUCTURE HARVESTER (Earl 2012) using the formula as described by Evanno et al. (2005) ΔK = mean ([L"K])/sd[L(K)].

In addition, a UPGMA dendrogram tree was generated and visualised using MEGA7 (Kumar et al. 2016).

RESULTS AND DISCUSSION

SNP distribution and characterisation. The genotyping and analysis of 1 536 SNPs showed that only 932 SNPs (60.7%) gave high quality allele scores, whereas the remaining SNPs either failed to amplify or had low call rates across the samples. The distribution of the 932 SNPs across the twelve rice chromosomes ranged from 53 SNPs on chromosome 11 to 121 SNPs on chromosome 1. Analysis of the 932 SNPs revealed a total of 16 SNPs were monomorphic. None of the monomorphic SNPs were detected on chromosomes 1, 4, 8, 11, or 12. The details of the SNP numbers and distribution are summarised in Figure 1.

The analysis of the SNPs per chromosome revealed the mean PIC value ranged from 0.173 on chromosome 12 to 0.259 on chromosome 11, with an average of 0.213. A low PIC value was detected when the SNP was a bi-allelic marker, and the maximum PIC value reached 0.5. The PIC value provides information about the polymorphism level of the genotypes under study. Values near 1 indicate a high degree of genetic diversity and are associated with a high number of alleles, whereas values less than 0.5 indicate a low level of polymorphism (Becerra et al. 2015). The mean PIC value suggested a low level of polymorphism among the studied genotypes. It was also notable that the obtained PIC value was lower than those in a previous study using SNP markers conducted by Xu et al. (2016), who found a value of 0.375, but was close to the PIC value obtained by Chen et al. (2011), who found a value of 0.257. The mean allele numbers by chromosome ranged from 1.944 on chromosome 7 to 2.000 on chromosomes 1, 4, 8, 11, and 12, with an average of 1.982. Low levels of heterozygosity were detected since the value ranged from 0.007 on chromosome 7 to 0.026 on chromosome 10. The details of the SNPs characterised across the 12 rice chromosomes are summarised in Table 2.

The SNPs represent the most abundant markers in the plant genomes. Unlike the SSR markers, SNP analyses can be performed without separation of the allele size and, therefore, it is an amenable high-throughput genotyping assay. Even though the SNPs showed a relatively low level of polymorphism compared to the SSR markers, the bi-allelic nature of the SNPs offers the advantage of lower error rates in the allele calling and results in the reproducibility across the laboratories. These advantages have led the SNP markers becoming the favourable markers for the varietal identification and diversity studies

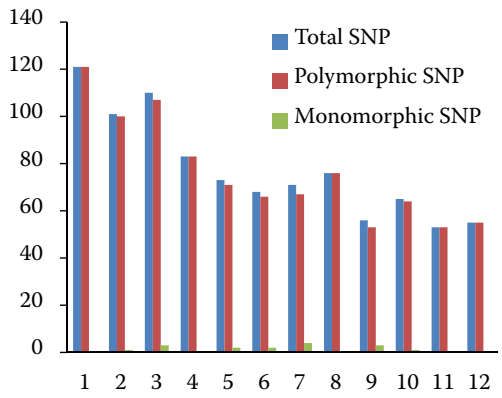

Figure 1. The single nucleotide polymorphisms (SNPs) distribution across the twelve chromosomes in rice

https://doi.org/10.17221/58/2019-CJGPB

Table 2. The single nucleotide polymorphism characterisation based on the 12 rice chromosomes

Chromosome	Major allele frequency	Allele No.	Gene diversity	Heterozygosity	PIC
1	0.844	2.000	0.237	0.008	0.200
2	0.816	1.990	0.259	0.010	0.213
3	0.854	1.973	0.225	0.019	0.191
4	0.803	2.000	0.271	0.008	0.220
5	0.804	1.973	0.268	0.009	0.217
6	0.801	1.971	0.269	0.012	0.218
7	0.860	1.944	0.208	0.007	0.175
8	0.791	2.000	0.299	0.017	0.245
9	0.802	1.946	0.273	0.024	0.223
10	0.791	1.985	0.283	0.026	0.227
11	0.756	2.000	0.324	0.020	0.259
12	0.870	2.000	0.203	0.019	0.173
Mean	0.816	1.982	0.260	0.015	0.213

PIC – polymorphic information content

in diverse crops such as rice (Reig-Valiente et al. 2016), barley (Soleimani et al. 2003), wheat (Ren et al. 2013), and maize (Dao et al. 2014).

Genetic diversity and population structure of the released Malaysian rice varieties. The genetic analysis of 46 released Malaysian rice varieties revealed the presence of two major groups (red and green) as revealed by the UPGMA dendrogram analysis and illustrated in Figure 2. The dendrogram tree showed that two varieties, namely MR269 and MRIA, were clustered in the red group, and the remaining varieties were clustered in the green group. The analysis was supported by the findings from the STRUCTURE analysis which indicated the ΔK value to be at the highest peak at K = 2 (as described by Evanno et al. 2005), followed by K = 4 (Figure 3). When K = 2, the varieties were similarly grouped to that in the dendrogram analysis where MRIA 1 and MR269 (the

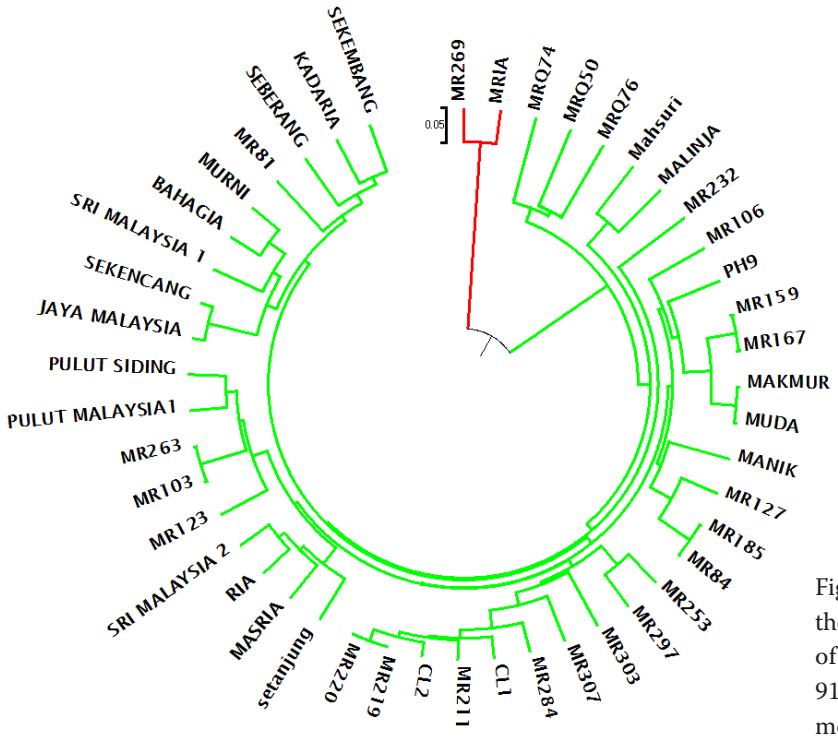

Figure 2. An unweighted pair group method with an arithmetic mean dendrogram of 46 Malaysian rice varieties based on 916 polymorphic single nucleotide polymorphism markers

https://doi.org/10.17221/58/2019-CJGPB

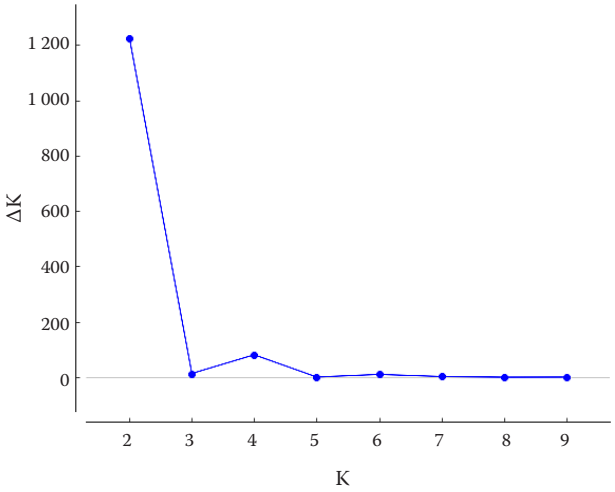

Figure 3. Prediction of the number of K (number of population structure), where the number of K represents the highest peak of ΔK, (ΔK = m (|L" (K)|)/sd[L(K)])

red group) were clustered separately from the other varieties (Figure 4A). Because MRIA 1 is the only aerobic rice variety used in this study, the variety was expected to be diverse from the other varieties. MRIA 1 is derived from a chemical mutation of the IR76569-259-1-2-1 variety, whereas MR269 is derived from the hybridisation of two advanced lines namely P347 × Y1362. The pedigree information revealed that one of the parents (Y1362) is of Indian-origin, Pankhari203, which might contribute to the high variability of MR269. When K = 4, 3 subclusters arose from the Green Cluster (subcluster II, III and IV) (Figure 4B). Interestingly, all the glutinous rice varieties (Masria, Pulut Malaysia 1, Pulut Siding and Pulut Hitam 9) were grouped in subcluster II. Meanwhile, all the aromatic rice varieties, namely MRQ50, MRQ74 and MRQ76, were clustered in subcluster III. This finding is in agreement with Roy et al. (2015) who also found all aromatic rice varieties being grouped in the same subcluster. In addition, most of the semi-dwarf high-yielding rice varieties were grouped in subcluster IV. As expected, the herbicide tolerant rice varieties, namely MR220-CL1 and MR220-CL2, were grouped together with their respective parent, namely MR220. The list of the varieties within the subclusters is described in Table 3. The pairwise genetic distance of the shared alleles (data not shown) showed that the value ranged from 0.000 (MR159↔ MR167) to 0.723 (MRIA↔ Setanjung), thus, suggesting that MR159 and MR167 are identical and that there was the highest dissimilarity between MRIA 1 and Setanjung according to the SNP dataset. Both MR159 and MR167 were introduced by the Malaysian government to Malaysian farmers in 1995, whereas Setanjung and MRIA 1 were introduced in 1979 and 2014, respectively. Setanjung is a lowland rainfed rice variety while MRIA 1 is an aerobic rice variety. The principle component analysis (PCoA) was generated to visualise the variety coordinates

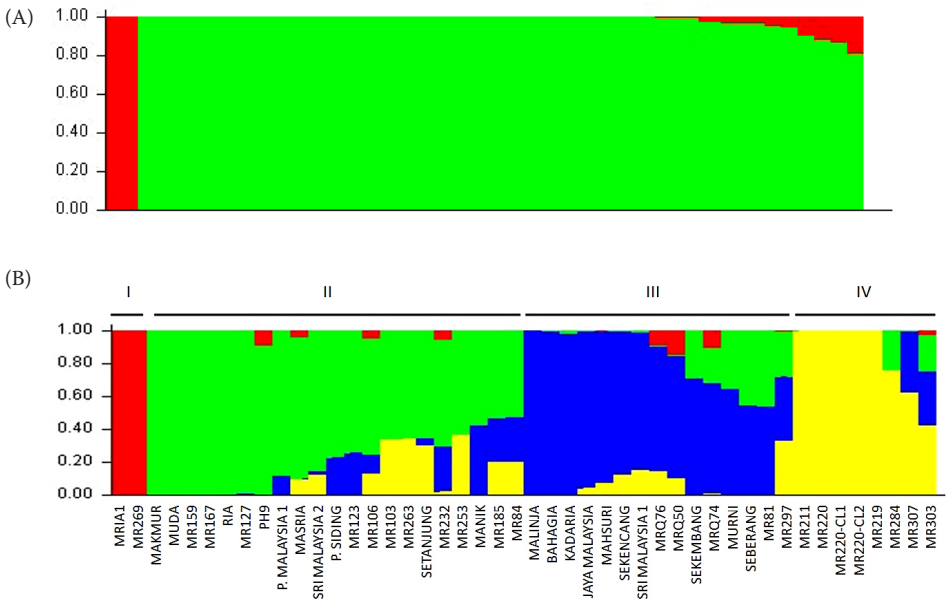

Figure 4. The genetic structure diagrammatic generated by the STRUCTURE program based on K = 2 (highest peak of ΔK) (A) and K = 4 (second highest of ΔK) (B) using 916 polymorphic single nucleotide polymorphisms

https://doi.org/10.17221/58/2019-CJGPB

Table 3. The list of the varieties in their respective subcluster

Subcluster	Variety
I	MRIA 1, MR269
II	Makmur, Muda, MR159, MR167, Ria, MR127, Pulut Hitam 9, Pulut Malaysia, Masria, Sri Malaysia 2, Pulut Siding, MR123, MR106, MR103, MR263, Setanjung, MR232, MR253, Manik, MR185, MR84
III	Malinja, Bahagia, Kadaria, Jaya Malaysia, Mahsuri, Sekencang, Sri Malaysia 1, MRQ76, MRQ50, Sekembang, MRQ74, Murni, Seberang, MR81, MR297
IV	MR211, MR220, MR220-CL1, MR220-CL2, MR219, MR284, MR307, MR303

based on the genetic distance matrices (Figure 5). The PCoA was performed using a model-based approach as obtained by the STRUCTURE analysis, which revealed the existence of a large genetic diversity in the Malaysian rice varieties. The first three axes explained 34.79% of the cumulative variation (Table 4). Both the green and red groups showed distinct grouping in the PCoA. An AMOVA was performed on these two populations using a model-based approach. Among the two populations, 69% variance was recorded; 29% variance was recorded among the individuals; and 2% variance was recorded within the individual (Table 5). The value suggests a high genetic variation between the differentiation of the two populations. The percentage of variance obtained in this study between the two population differentiations is quite similar to the percentage of variance obtained by Singh (2019). One possible reason for the low variance within the individual is the use of readily released rice varieties that have been adopted to high-selection pressure, instead of using the wild

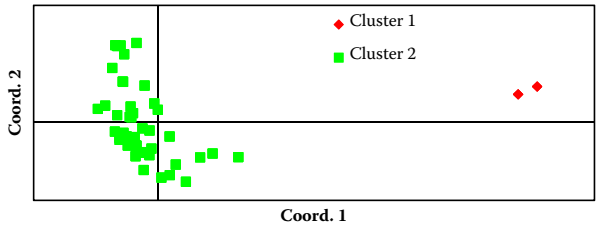

Figure 5. The Principle Component Analysis (PCoA) of the released Malaysian rice varieties based on 916 polymorphic SNPs using a model-based approach

Table 4. The percentage of the variation explained by the first three axes

Axis	1	2	3
%	17.98	9.39	7.39
Cumulative %	17.98	27.36	34.76

relative species of landraces cultivars (Thomson et al. 2007). Although there is a diversity between the two groups, the lack of varieties in the red group is a major concern. Selection based on consumer preferences has also led to the reduction in the genetic basis of the released Malaysian rice varieties. The evaluation of crop diversity offers some acumen for plant breeders to develop and improve the cultivars for the desired traits or characteristics through breeding programmes without losing the genetic variability. Exploitation of the natural genetic diversity for food requirements began in the early era of agriculture. Now, such exploitation has focused on improving food crops for the expanding population (Ahuja & Jain 2015). However, the modern varieties of most crops, especially rice, were developed mainly to produce high yields, which led to a reduction in the genetic variation. Hence, a forthcoming breeding programme to diversify high yielding rice varieties is required for rice sustainability.

CONCLUSION

Assessment of the diversity and population structure using SNP markers is very useful for assisting

Table 5. The summary of the AMOVA obtained by the STRUCTURE program and a model-based approach

Source	*df*	SS	MS	Est. var.	%
Among populations	1	5372.736	5372.736	225.471	69
Among individuals	136	26753.318	196.716	95.048	29
Within individuals	138	913.500	6.620	6.620	2
Total	275	33039.554		327.138	100

df – degrees of freedom; SS – sum of squares; MS – mean squares; Est. var. – estimate of variance

https://doi.org/10.17221/58/2019-CJGPB

in the classification and identification, the proper management and conservation of genetic resources, and for the utilisation of genetic resources in the forthcoming breeding programme. Additionally, a molecular based identification or characterisation must be applied to sustain the integrity and purity of the varieties, which will benefit Malaysian farmers through the cultivation of high-quality seeds for better rice production.

REFERENCES

Ahuja M.R., Jain S.M. (2015): Genetic Diversity and Erosion in Plants: Indicators and Prevention. Vol. 7. Cham, Springer International Publishing Switzerland.

Becerra V., Paredes M., Gutiérrez E., Rojo C. (2015): Genetic diversity, identification, and certification of Chilean rice varieties using molecular markers. Chilean Journal of Agricultural Research, 75: 267–274.

Bohn M., Utz H.F., Melchinger A.E. (1999): Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. Crop Science, 39: 228–237.

Chen H., He H., Zou Y., Chen W., Yu R., Liu X., Yang Y., Gao Y.M., Xu J.L., Fan L.M., Li Y. (2011): Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). Theoretical and Applied Genetics, 123: 869–879.

Dao A., Sanou J., Mitchell S.E., Gracen V., Danquah E.Y. (2014): Genetic diversity among INERA maize inbred lines with single nucleotide polymorphism (SNP) markers and their relationship with CIMMYT, IITA, and temperate lines. BMC Genetics, 15: 1–14.

Earl D.A. (2012): STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4: 359–361.

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.

Freamo H., O'Reilly P., Berg P.R., Lien S., Boulding E.G. (2011): Outlier SNPs show more genetic structure between two Bay of Fundy metapopulations of Atlantic salmon than do neutral SNPs. Molecular Ecology Resources, 11: 254–267.

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.

Habibuddin H., Shen Y., Rabiatul Adawiah A.Z., Khairun Hisam N., Hayati A., Shahril Firdaus A.R., Siti Norhayati I., Azwan J. (2013): Potensi penggunaan 1536 SNP dalam kajian-kajian genetik padi. In: Proc. Persidangan Padi Kebangsaan 2013, Pulai Pinang, Dec 10–12, 2013: 412–417. (in Malay)

Helyar S.J., Hemmer-Hansen J., Bekkevold D., Taylor M.I., Ogden R., Limborg M.T., Cariani A., Maes G.E., Diopere E., Carvalho G.R., Nielsen E.E. (2011): Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources, 11: 123–136.

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.

Liu N., Chen L., Wang S., Oh C., Zhao H. (2005): Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. BMC Genetics, 6: S26.

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, 21: 459–460.

Marić S., Bolarić S., Martinčić J., Pejić I., Kozumplik V. (2004): Genetic diversity of hexaploid wheat cultivars estimated by RAPD markers, morphological traits and coefficients of parentage. Plant Breeding, 123: 366–369.

Morin P.A., McCarthy M. (2007): Highly accurate SNP genotyping from historical and low-quality samples. Molecular Ecology Notes, 7: 937–946.

Pejic I., Ajmone-Marsan P., Morgante M., Kozumplick V., Castiglioni P., Taramino G., Motto M. (1998): Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theoretical and Applied Genetics, 97: 1248–1255.

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

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.

Ren J., Chen L., Sun D., You F.M., Wang J., Peng Y., Nevo E., Beiles A., Sun D., Luo M.C., Peng J. (2013): SNP-revealed genetic diversity in wild emmer wheat correlates with ecological factors. BMC Evolutionary Biology, 13: 1–15.

Rosenberg N.A., Li L.M., Ward R., Pritchard J.K. (2003): Informativeness of genetic markers for inference of ancestry. American Journal of Human Genetics, 73: 1402–1422.

https://doi.org/10.17221/58/2019-CJGPB

Roy S., Banerjee A., Mawkhlieng B., Misra A.K., Pattanayak A., Harish G.D., Singh S.K., Ngachan S.V. Bansal K.C. (2015): Genetic diversity and population structure in aromatic and quality rice (*Oryza sativa* L.) landraces from North-Eastern India. PLOS ONE, 10: 1–13.

Singh Y.T. (2019): Genetic Diversity and Population Structure in Upland Rice (*Oryza sativa* L.) of Mizoram, North East India as revealed by morphological, biochemical and molecular markers. Biochemical Genetics, 57: 421–442.

Soleimani V.D., Baum B.R., Johnson D.A. (2003): Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley. Plant Molecular Biology Reporter, 21: 281–288.

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. Theoretical and Applied Genetics, 114: 559–568.

Xu Q., Yuan X., Wang S., Feng Y., Yu H., Wang Y., Yang Y., Li X. (2016): The genetic diversity and structure of indica rice in China as detected by single nucleotide polymorphism analysis. BMC Genetics, 17: 1–8.

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.

Received: June 19, 2019
Accepted: October 7, 2019
Published online: January 22, 2020