Volume: 2, Issue:2, (December, 2020), pp 31-55
www.rjt.ac.lk/agri/sljae

SRI LANKAN JOURNAL OF AGRICULTURE AND ECOSYSTEMS *e*ISSN: 2673-1401

SLJAE

ORIGINAL ARTICLE

Establishment of a Scheme of DNA Fingerprinting and Sequencing to Distinguish the Caryopsis Samples of the Exportable Rice Cultivars from a Panel of Mega-Varieties

G. K. S. Ananda[1#], W. W. M. U. K. Wijesundara[1#], H. M. V. P. K. Herath[2], S. I. Karunrathne[1], S. K. Kannangara[1], L. T. Ranaweera[1], P. G. R. G. Rathnayake[1], S. P. Abeysundara[2], U. A. K. S. Udawela[3], W. M. W. Weerakoon[4], C. K. Weebadde[5], S. D. S. S. Sooriyapathirana[1*]

[1]*Department of Molecular Biology and Biotechnology, Faculty of Science, University of Peradeniya, Peradeniya (20400), Sri Lanka.*

[2]*Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya (20400), Sri Lanka.*

[3]*Rice Research and Development Institute, Bathalagoda, Ibbagamuwa, Sri Lanka.*

[4]*Department of Agriculture, Peradeniya (20400), Sri Lanka.*

[5]*Department of Plant, Soil and Microbial Sciences, College of Agriculture and Natural Resources, Michigan State University, East Lansing, MI 48824, USA.*

#These authors have contributed equally.

Correspondence:
[*]*sunethuop@gmail.com*,
https://orcid.org/0000-0002-5592-1742

Abstract

The Rice Research and Development Institute (RRDI) of Sri Lanka has developed five exportable rice cultivars; At 362, Bg 94-1, Bg 360, Bg 1165-2, and Bw-Bs-1-2-31.The present study was conducted to establish the identities of these five cultivars at caryopsis level in comparison compared to those of mega rice cultivars (Bg 352, Bg 300, Bg 358, Bg 359, Bg 357, Bg 379-2, and At 353) in Sri Lanka using DNA fingerprinting and sequencing. These rice-cultivars were grown in a greenhouse and a field using breeder-seeds. The seeds were harvested and subjected to a morphometric analysis using decision tree algorithms based on the size and colour of seeds and caryopses. The algorithms estimated the percentage accuracy of detection based on morphometric analysis ranged from 3.13-84.38 %. Similar seed and caryopsis combinations were grouped and exposed them to a panel of human subjects to discriminate the samples in each combination and subjected the data to calculate Kappa (K) and inter-rater reliability (IRR) statistics. The K was always 0.00, and IRR was 27% implying the inability of accurate visual differentiation. In the DNA fingerprinting analysis, a set of six SSR markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475*, and *RM23744*) were selected that can establish the cultivar identity. In addition, the combined analysis of DNA sequencing of 12 cultivars with three selected loci, (*Seq 7-8*, *HvSSR12-34* and *RM23744*) authenticated the varietal identities.

Keywords: *Rice exporting in Sri Lanka*, *Rice Varietal Identity, Rice varietal ownership, Varietal identity of rice caryopses*,

DOI: http://doi.org/10.4038/sljae.v2i2.37

Date of Submission: 29-01-2020 Date of Acceptance: 21-10.2020

1. Introduction

The Rice Research and Development Institute (RRDI) has recently identified five exportable Sri Lankan rice cultivars. They consist of three released varieties namely, At 362, Bg 94-1, and Bg 360, and unreleased lines Bg 1165-2, and Bw-Bs-1-2-31. The RRDI must claim a defined percentage of the revenue from the export market to fund the breeding programs currently that depend on the limited government funds.

The traits used to determine the novelty of a newly bred variety can be either physiological, morphological, or other characteristics which are mostly the field based observations (Cooke and Reeves 2003) provided by the guidelines of the International Convention for the Protection of New Varieties of Plants (UPOV) (Barton 1982; Kjeldgaard and Marsh 1994; Williams 1984). The varietal discrimination for the identification of novel varieties and diversity analysis based on the morphological parameters have been carried out in Sri Lanka (Suriyagoda et al. 2011; Wijayawardhana et al. 2015) as well as worldwide (Caldo et al. 1996) for the varieties with significant morphological variations. However, due to the narrow genetic diversity, the morphological characteristics in improved varieties, which are closely related to each other, have limited applicability for varietal discrimination to award Intellectual Property Rights (IPRs) (Rahman et al. 2009). Also, most of the morphological descriptors especially at the caryopsis level used in varietal identification are quantitative. It makes the morphological descriptors are not distinct enough to be employed in varietal identification as the environmental factors influence the expression of quantitative traits (Weising et al. 2005). The genetic relationships between the novel varieties and their relatives and the subjectivity in the data collection also obscure the varietal discrimination based on morphological variations (Zhu et al. 2012; Nybom et al. 2014).

Comparatively, molecular markers define differences in their nucleotide sequences (Kwon et al. 2005). Thereby the molecular markers allow fast and precise varietal identification (Singh et al. 2013). The PCR based molecular markers, such as Simple Sequence Repeat (SSR) markers, are preferred due to their higher reproducibility, simplicity, reliability, polymorphism and, co-dominant nature (Salgotra et al. 2015). Many scientists used SSR markers for successful cultivar identification of apple (Moria et al. 2011), almond (Dangl et al. 2009), potato (Coomb et al. 2004), pineapple (Shoda et al. 2012), soybean (Rongwen et al. 1995), and many other crop species. Also, because of higher abundance and distribution across the

genome (McCouch et al. 1997), SSR markers are being readily utilized in diversity analyses (Choudhury et al. 2001; Jain et al. 2004) and varietal identification of rice (Zhu et al. 2012).

The variations in a DNA sequence is the basis for genetic diversity accounting for a significant fraction of observed differences among plant varieties. Naturally occurring genetic variations consist of small insertions, deletions as well as base substitutions which are difficult to be recognized by SSR-marker-based DNA fingerprinting. However, the variations in DNA can be identified with DNA sequencing (McNally et al. 2006) which makes it a feasible genetic tool for varietal identification. So far, DNA sequencing has been fruitfully utilized for identification of the plant varieties such as *Arabidopsis thaliana* (Kim et al. 2007) and *Cucumis melo* (Deleu et al. 2009). The literature suggests that DNA sequencing is a valuable technique to differentiate closely related rice varieties with their well informative genomic differences based on intensely distributed single nucleotide variations (Shirasawa et al. 2004; Sato et al. 2002).

The present study was conducted to accomplish three objectives. Firstly, the study focused on testing the applicability of morphometric trait analysis of the caryopsis samples as the preliminary step of cultivar identification for RRDI. Then, we aimed at establishing a DNA fingerprinting protocol with SSR markers to discriminate caryopsis samples of the rice cultivars identified for exporting. Finally, we targeted introducing DNA sequencing of SSR loci to put forward DNA barcodes to set up a standard procedure for varietal identification.

2. Materials and Methods

Growing plants for phenotypic measurements

Twelve rice cultivars were assessed including the five exportable and seven local mega rice cultivars for their identity (Table 1) by growing ten seedlings from authenticated breeder-seeds of each cultivar collected from RRDI, Sri Lanka under field and greenhouse conditions according to Completely Randomized Design (CRD) layouts in the *Maha* Season (October – February) of 2015/2016, *Yala* Season (April-July) of 2016 and *Maha* Season of 2016/2017 (*Yala* and *Maha* are the two crop growing seasons in Sri Lanka). Table 2 provides the mean rainfall, temperature, relative humidity, and day-length conditions of the seasons with the GPS coordinates of the field and greenhouse locations. The rice plants were managed according to the crop recommendations given by the Department of Agriculture, Sri Lanka and, harvested the panicles at the right maturity stage and

processed to obtain seeds and caryopses (i.e., edible part of rice) for the analyses.

Analyses of the morphometric data of seeds and caryopses for cultivar discrimination

The length, width, and 100-unit weight of seeds and caryopses, together with the red, green, and blue (RGB) values of the seed and caryopses samples were measured and recorded. The morphological parameters of cultivars were classified using; length, width, 100 unit weight and, RGB values of both seeds and caryopses of each cultivar collected in each season by executing decision tree algorithms to differentiate rice cultivars using the following equation (Kennedy et al. 1998).

Let ϕ (s|t) be a measure of the "goodness" of a candidate split *s* at node *t*, where

$$\phi(s|t) = 2P_L P_R \sum_{j=1}^{\#\,Classes} |P(j|t_L) - P(j|t_R)|$$

Where, the optimal split is whichever split maximizes this measure ϕ (s|t) overall possible splits at node t.

t_L = left child node of node t

t_R = right child node of node t

$$P_L = \frac{\text{number of records at } t_L}{\text{number of records in training set}}$$

$$P_R = \frac{\text{number of records at } t_R}{\text{number of records in training set}}$$

$$P_{(j|t_L)} = \frac{\text{number of class j records at } t_L}{\text{number of records at t}}$$

$$P_{(j|t_R)} = \frac{\text{number of class j records at } t_R}{\text{number of records at t}}$$

First, we determined the tree topology using decision tree algorithm, and Recursive Partitioning (R-PART) (Therneau and Atkinson 1997) by using 70% data as the training data and the remaining 30% as the testing data. There, the tree topology was determined by a binary grouping of variables. The algorithm implements as a two-stage procedure where it initially finds the best variable which can group the dataset into two subsets. This cycle was applied via each variable until the best topology was received. Further, we used the C5.0 Classification model (Kuhn 2013) to build a decision tree for comparison. In this model, the grouping was carried out by separating the groups which gave the maximum information gain. The algorithm was executed until the splitting of the subsampled dataset finished. Finally, we implemented the two decision tree algorithms on a final combined data set including all seed, caryopsis, and RGB values to discover the ability to classify the 12 cultivars using the morphological traits.

Visual differentiation of seed and caryopsis samples of rice cultivars

We grouped the 12 rice cultivars based on the sizes of seeds and caryopses (Table 3). In the grouping, four cultivar combinations each based on similar seed sizes (A1-A4 in Table 3) and caryopsis sizes (B1-B4 in Table 3) were identified. We employed 30 well-

experienced human subjects (i.e., raters) and they were given seeds and caryopses samples as the combinations given in Table 3. The representative images of these seed and caryopsis combinations are given in Plate 1. The raters expressed their opinion on whether they could differentiate all into individual cultivars, groups of cultivars or the inability of discrimination, and the data were recorded. We subjected these response data to calculate the Kappa value (K) and the inter-rater reliability (IRR) as given in the following equations (McHugh 2012) using the Statistical Package Minitab 16 (Minitab Inc., USA, 2018).

$$\text{K (Kappa value)} = \frac{[\text{Pr(a)} - \text{Pr(e)}]}{1 - \text{Pr(e)}}$$

Where, Pr (a) : actual observed agreement; Pr (e): chance agreement; IRR: Inter – rater reliability

$$\text{IRR} = \frac{\text{No. of raters with correct answer}}{\text{Total No. of raters}} x100$$

DNA fingerprinting

The leaf DNA was initially used to establish the protocol. The DNA was extracted from immature leaf samples using Wizard® Genomic DNA Purification Kit (Cat. No.: A1120, www.worldwide.promega.com; Promega Kit) and stored the extracted DNA at -20 °C. Then using the extracted DNA samples, duplex PCR was carried out with *K20*; a monomorphic DNA marker linked with *Pup1* [a quantitative trait locus (QTL) in rice genome associated with phosphorous uptake] (Chin et al. 2010), as the standard marker, for 19 simple sequence repeat (SSR) markers. Also, simplex PCR was carried out with four SSR markers due to their different annealing temperatures and poor amplification with *K20* marker when duplexed (Table 4). We used PCR mixtures (15 µL) each comprised of 1.5 µL of template DNA (50 ng – 80 ng), 7.5 µL of 2× GoTaq Green® Master Mix, and 0.5 µL of each primer and 4 µL of nuclease-free water. The PCR was performed in a Thermal Cycler (Takara, Otsu Shiga, Japan) using the conditions; 5 mins initial denaturation at 94 °C, followed by 35 cycles of 30 sec of denaturation at 94 °C, 1 min at annealing temperatures (T_a) (Table 4), 2 mins at 72 °C and a final extension at 72 °C for 10 mins. Then we extracted genomic DNA from rice caryopsis samples, and PCR amplified using the same 23 SSR markers. Also, before the above amplification procedure, we diluted each DNA sample with autoclaved distilled water in a 1:30 ratio to meet the appropriate template concentrations for PCR. Finally, all the amplified fragments were resolved by 2.5% agarose gel electrophoresis.

Analysis of DNA fingerprinting data

We identified a minimum set of SSR markers which can define the identities of rice cultivars tested by constructing a dendrogram based on the polymorphic bands observed for the selected markers

Table 1: Important characteristics of the selected rice cultivars.

Cultivar	Key characteristics
At 362*	Red pericarp; moderately resistant to brown planthopper and bacterial blight (Aluwihare et al. 2016); salt tolerant (Pradheeban et al. 2015); good eating quality (Rajkumar et al. 2016)
Bg 94-1*	White pericarp; phosphorous deficiency tolerant (Kekulandara et al. 2017); high yielding; moderately susceptible to rice blast, iron toxicity, and thrips; moderately resistant to bacterial blight; susceptible to rice gall midge (Biotype 1 and Biotype 2) and brown planthopper; ideal as parboiled rice (Aluwihare et al. 2016)
Bg 360*	White pericarp; highly salt sensitive (Pradheeban et al. 2015); resistant to rice gall midge (Biotype 1 and Biotype 2), brown planthopper, rice blast, and bacterial blight; moderately resistant to iron toxicity; excellent eating quality; very small caryopsis size (DOA 2017)
Bg 1165-2*	White pericarp
Bw-Bs-1-2-31*	White pericarp
Bg 352	White pericarp; salt sensitive (Pradheeban et al. 2015); susceptible to rice gall midge (Biotype 2); moderately susceptible to thrips; resistant to rice blast, bacterial blight, rice gall midge (Biotype 1), brown plant hopper and, iron toxicity; intermediate bold type caryopsis; wide adaptability (DOA 2017; Rajkumar et al. 2016)
Bg 300	White pericarp; resistant to rice gall midge (Biotype 1), bacterial blight, rice blast, and brown planthopper; moderately resistant to green leafhopper; high yielding (DOA 2017; Aluwihare et al. 2016; Rajkumar et al. 2016)
Bg 358	White pericarp; small caryopsis; high yielding; resistant to bacterial blight; rice blast and brown planthopper; moderately tolerant to iron toxicity; high yielding (DOA 2017; Aluwihare et al. 2016; Rajkumar et al. 2016)
Bg 359	White pericarp; small caryopsis; resistant to brown planthopper, rice gall midge (Biotype 1 and Biotype 2) and bacterial blight; moderately resistant to thrips, iron toxicity and low temperature; higher grain weight; higher yield (DOA 2017; Kekulandara et al. 2017)
Bg 357	White pericarp; resistant to rice gall midge (Biotype 1 and Biotype 2), rice blast, bacterial blight, and brown plat hopper; moderately resistant to iron toxicity, high amylose content; intermediate gelatinization temperature; higher yielding (DOA, 2017; Aluwihare et al. 2016; Kekulandara et al. 2017)
Bg 379-2	White pericarp; resistant to brown planthopper and bacterial blight; moderately resistant to green leafhopper and rice blast; higher caryopsis quality and high yielding (DOA 2017; Aluwihare et al. 2016; Rajkumar et al. 2016)
At 353	Red pericarp; salt tolerant (Pradheeban et al. 2015); moderately resistant to bacterial blight and rice blast; ideal for potential acid/saline conditions; ideal as parboiled rice; phosphorous deficiency tolerant (Kekulandara et al. 2017)

*Exportable rice cultivars

The detailed cultivar information can be found in RRDI (2018).

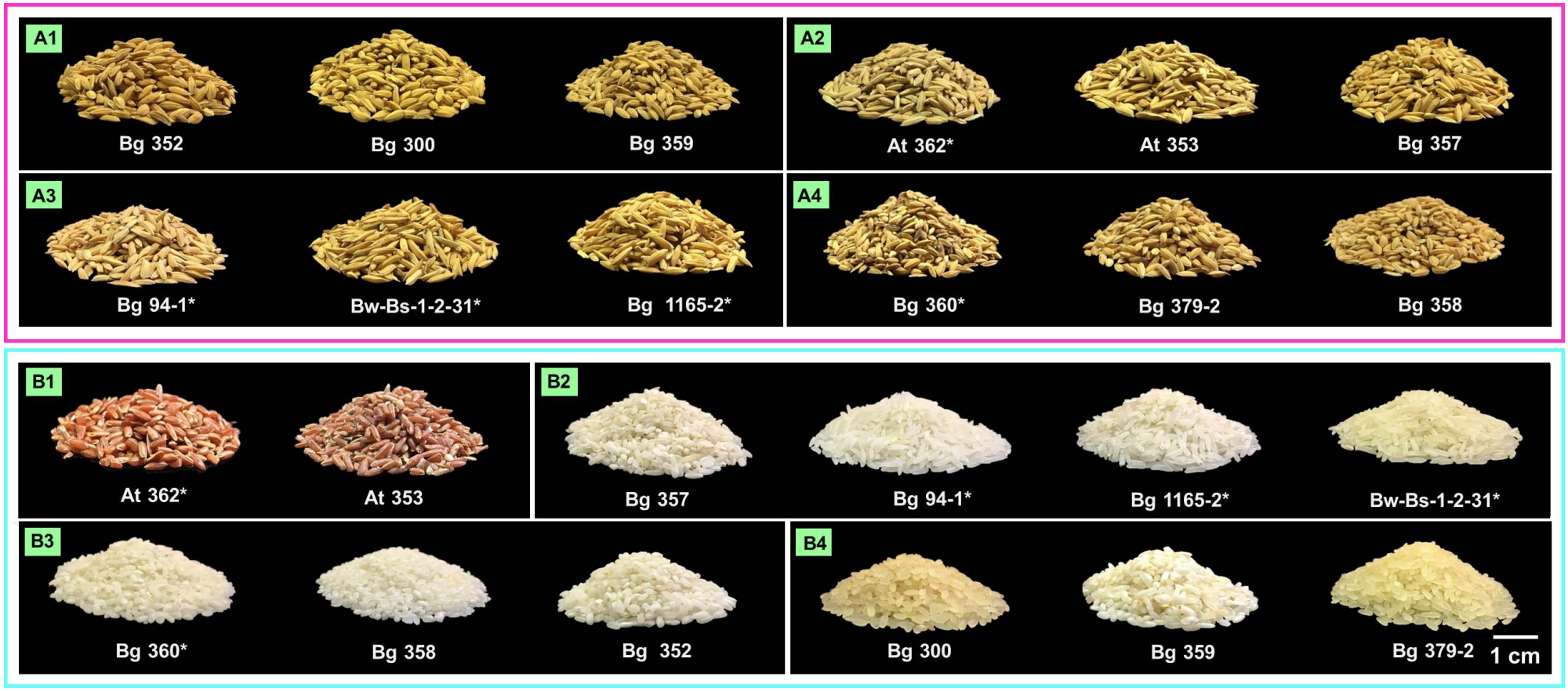

Plate 1: Seeds caryopses of rice cultivars grouped into respective combinations (Table 3) based on their similar visual characteristics. A1-A4: Combinations using morphology of seeds; B1-B4: Combinations using morphology of caryopses. The sizes of the seeds and caryopses are indicated by the scale bar on the upper left corner of the Figure. Exportable rice cultivars are marked with *.

Table 2: The mean weather conditions of the growing seasons (*Maha* and *Yala*) at field and greenhouse locations.

Season, Year (months)	Field/ Greenhouse (GH)	Location (District, GPS coordinates)	Mean Monthly Temperature (°C)	Mean Monthly Rainfall (mm)	Mean Relative Humidity (%)	Mean Day length (hrs)
Maha, December, 2015 - March, 2016	Field	*Kurunegala* 7.531502 °N, 80.435510°E	27	365.69	83.00	11.40
Maha, December, 2015 - March, 2016	GH	Kandy 7.258704 °N, 80.597150 °E	27	269.96*	80.00	11.40
Yala, June - September, 2016	Field	*Kurunegala* (same)	26	47.70	79.00	12.30
Yala, June - September, 2016	GH	Kandy (same)	26	26.64*	77.00	12.30
Maha, December, 2016 - April, 2017	GH	Kandy (same)	26	46.10*	74.00	11.40

Sources: World Weather Online (2018) and timeanddate.com (2018), *Not affected directly on the GH.

(The conventional *Maha* season and *Yala* season proceed from September to March and May to August respectively).

for the cultivars using Complete Linkage and Euclidean Distance methods in Minitab 16. The polymorphism of each marker was recorded by considering the bands detected on agarose gels as alleles and the Polymorphism Information Content (PIC) value of each SSR marker was calculated according to the following equation.

$$\text{PIC} = 1 - \sum(p_i^2)$$

Where p_i is the proportion of the genotypes containing the allele in all the samples analyzed.

DNA sequencing

We performed DNA sequencing for 11 DNA markers for all 12 cultivars were performed (Table 5) and selected three SSR markers, *HvSSR12-34* (Singh et al. 2010), *Seq 7-8* (Lu et al. 2012), and *RM23744* (Mukherjee et al. 2013). Next, we amplified DNA samples from each of the 12 rice cultivars using simplex PCR for the selected SSR markers. The PCR cycle consisted of initial denaturation at 94 °C for 5 mins, followed by 35 cycles including 30 sec of denaturation at 94 °C, 1 min annealing at an appropriate temperature (Table 5) and 2 mins extension at 72 °C followed by final extension of 10 mins at 72 °C. We separated and visualized PCR products in 1 % agarose gel electrophoresis and purified the PCR products using QIAquick PCR purification kit

Table 3: The statistics showing the inability to discriminate rice cultivars by the respondents.

Combination ID	Cultivars in the Combination	Kappa value	Inter-rater reliability (IRR) (%)
A1	Bg 352, Bg 300 and Bg 359	0.00	0.67
A2	At 362*, At 353 and Bg 357	0.00	0.17
A3	Bg 94-1*, Bw-Bs-1-2-31* and Bg 1165-2*	0.00	0.20
A4	Bg 360*, Bg 379-2 and Bg 358	0.00	0.40
B1	At 362* and At 353	0.00	0.13
B2	Bg 357, Bg 94-1*, Bg 1165-2* and Bw-Bs-1-2-31*	0.00	0.27
B3	Bg 360*, Bg 358 and Bg 352	0.00	0.67
B4	Bg 300, Bg 359 and Bg 379-2	0.00	0.23
	Mean IRR		0.27

*Exportable rice cultivars

(Qiagen, Hilden, Germany) and cycle-sequenced the PCR products at Macrogen Inc., (Seoul, Korea).

Analysis of DNA sequencing data

Initially, we constructed three separate alignments for the markers *HvSSR12-34*, S*eq 7-8,* and *RM23744* using MEGA V7 software (Kumar et al. 2016) by manually checking the reading frame of the contig sequences. Implementing uncorrected pairwise distances among the sequences, we constructed three separate dendrograms using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm. Furthermore, we combined three data sets of 12 cultivars and re-analyzed using the UPGMA method. Finally, we modified all the resulted trees using FigTree v1.4.3 (Rambaut 2014). All 36 DNA sequences generated in the present study were deposited in GenBank under the accession numbers; MK264379-MK264390 (*HvSSR12-34*), MK293964 - MK293975 (*Seq 7-8*) and MK293976-MK293987 (*RM23744*) (Table 6).

3. Results

Analysis based on Kappa and inter-rater reliability values

Table 3 shows the kappa and the IRR values for each question in the questionnaire. According to them, all the combinations had the same kappa value (0.00). The IRR was ranging from 0.07-0.67, and the mean IRR was 0.27 indicating that the observations made by the panelists were not in the same direction.

Analysis based on R-PART and C5.0 classification models

The R-PART and C5.0 classification models revealed that the discrimination of 12 cultivars from each other based on 100 seed or caryopsis weight, length, and width of seeds and caryopses the accuracy percentage was ranging from 25.00 - 81.81% (Table 7). The same analysis using

Table 4: The details of the markers used and the detected polymorphism.

Marker	Sequences of the forward and reverse primers	T_a (°C)	References	Chromosome	No. of alleles observed	Allele size/s (bp)	PIC Value
*K20**	5'TCAGGTGATGGGAATCATTG3',5'TGTTCCAACCAAACAACCTG3'	55	Chin et al. (2010)	12	1	245	0.0000
RM144	5'TGCCCTGGCGCAAATTTGATCC3',5'GCTAGAGGAGATCAGATGGTAGTGCATG3'	55	Khush et al. (2003)	11	1	225	0.0000
RM153	5'GCCTCGAGCATCATCATCAG3', 5'ATCAACCTGCACTTGCCTGG3'	55	Rahman et al. (2009)	5	1	205	0.0000
RM154	5'ACCCTCTCCGCCTCGCCTCCTC3', 5'CTCCTCCTCCTGCGACCGCTCC3'	61	Temnykh et al. (2000)	2	3	300, 275,225	0.3633
RM161	5'TGCAGATGAGAAGCGGCGCCTC3', 5'TGTGTCATCAGACGGCGCTCCG3'	61	Temnykh et al. (2000)	5	1	175	0.0000
RM162	5'GCCAGCAAAACCAGGGATCCGG3', 5'CAAGGTCTTGTGCGGCTTGCGG3'	61	Temnykh et al. (2000)	6	3	340, 185,170	0.5433
RM202	5'CAGATTGGAGATGAAGTCCTCC3', 5'CCAGCAAGCATGTCAATGTA3'	55	Chen et al. (1997)	11	2	200,175	0.2392
RM206	5'CCCATGCGTTTAACTATTCT3', 5'CGTTCCATCGATCCGTATGG3'	55	Rahman et al. (2010)	2	2	175,150	0.3457
RM224	5'ATCGATCGATCTTCACGAGG3', 5'TGCTATAAAAGGCATTCGGG3'	55	Khush et al. (2003)	11	2	160,130	0.2392
RM246	5'GAGCTCCATCAGCCATTCAG3', 5'CTGAGTGCTGCTGCGACT3'	55	Chen et al. (1997)	1	2	120,90	0.3043
RM251	5'GAATGGCAATGGCGCTAG3', 5'ATGCGGTTCAAGATTCGATC3'	55	Rahman et al. (2010)	3	2	150,105	0.3750
RM307	5'GTACTACCGACCTACCGTTCAC3', 5'CTGCTATGCATGAACTGCTC3'	55	Rahman et al. (2010)	4	2	185,125	-
RM316	5'CTAGTTGGGCATACGATGGC3', 5'ACGCTTATATGTTACGTCAAC3'	55	Temnykh et al. (2000)	9	2	200,215	0.3680
RM333	5'GTACGACTACGAGTGTCACCAA3', 5'GTCTTCGCGATCACTCGC3'	55	Rahman et al. (2010)	10	2	195-175	0.3680
RM334	5'GTTCAGTGTTCAGTGCCACC3', 5'GACTTTGATCTTTGGTGGACG3'	55	Temnykh et al. (2000)	5	2	190-175	0.3047
RM335	5'GTACACACCCACATCGAGAAG3', 5'GCTCTATGCGAGTATCCATGG3'	55	Rahman et al. (2010)	4	2	150-110	0.3680
RM336	5'CTTACAGAGAAACGGCATCG3', 5'GCTGGTTTGTTTCAGGTTCG3'	55	Temnykh et al. (2000)	7	3	200,170,160	0.5926
RM475	5'CCTCACGATTTTCCTCCAAC3', 5'ACGGTGGGATTAGACTGTGC3'	55	Rahman et al. (2010)	2	2	190,175	0.3750
RM489	5'ACTTGAGACGATCGGACACC3', 5'TCACCCATGGATGTTGTCAG3'	55	Jamil et al. (2013)	3	1	235	0.0000
RM552	5'CGCAGTTGTGGATTTCAGTG3', 5'TGCTCAACGTTTGACTGTCC3'	55	Luther et al. (2017)	11	2	245,175	-
RM1369	5'AACCTGAGAGTGCCAATTGG3', 5'TCCCCTAGTAAAGCGGATTC3'	55	Mukherjee et al. (2013)	6	2	120,80	-
RM5479	5'AACTCCTGATGCCTCCTAAG, 5'TCCATAGAAACAATTTGTGC3'	55	Mukherjee et al. (2013)	2	1	200	0.0000
RM25181	5'AAAGAGCTTCCCTAATGGCTTCG 3',5'GAGAGAATGACCTCTCCCAAGACC3'	55	Mukherjee et al. (2013)	10	2	150,140	0.2392
RM23744	5'CTTAATACTCCGACGTAACAGTGG3',5' CCTGACTAAATGGAGCTTCTTCC3'	55	Mukherjee et al. (2013)	9	2	300-290	0.3457

* Monomorphic marker used for duplex PCR

Table 5: The details of the SSR markers used for the DNA sequencing of rice cultivars.

Marker	Sequences of forward and reverse primers	T_a (°C)	Band size (bp)	Quality of sequencing reaction
HvSSR03-02	5'TAGCGGAGTTGGAATAACAC3', 5'CTGCACTGCATACCTCATAA3'	55	228	unsuccessful
*HvSSR12-34**	5'ATGACCATAATCCCAACAAA 3',5'GTCGTGGTGTATTCTTGGT3'	56	300	successful
K20	5'TCAGGTGATGGGAATCATTG3' 5'TGTTCCAACCAAACAACCTG3'	55	245	Not enough polymorphism
K46-K1	5'TGAGATAGCCGTCAAGATGCT3', 5'TGAGCCAGTAGAATGTTTTGAGG3'	55	523	unsuccessful
RM154	5'ACCCTCTCCGCCTCGCCTCCTC3', 5'CTCCTCCTCCTGCGACCGCTCC3'	61	300-225	unsuccessful
RM206	5'CCCATGCGTTTAACTATTCT3', 5'CGTTCCATCGATCCGTATGG3'	55	150-175	unsuccessful
RM246	5'GAGCTCCATCAGCCATTCAG3', 5'CTGAGTGCTGCTGCGACT3'	55	90-120	unsuccessful
RM336	5'CTTACAGAGAAACGGCATCG3', 5'GCTGGTTTGTTTCAGGTTCG3'	55	200-160	unsuccessful
RM472	5'CCATGGCCTGAGAGAGAGAG3', 5'AGCTAAATGGCCATACGGTG3'	55	300	Partially successful
RM493	5'TAGCTCCAACAGGATCGACC3', 5'GTACGTAAACGCGGAAGGTG3'	55	210	unsuccessful
*Seq 7-8**	5'CATACGGATCCAGCCTCTGT3', 5'TTGCAATGATGCGTATTCAC3'	54	900	successful
*RM23744**	5'CTTAATACTCCGACGTAACAGTGG3', 5' CCTGACTAAATGGAGCTTCTTCC3'	55	290-300	successful

RGB parameters yielded even less percentage accuracy values ranging from 2.78 – 34.48 % (Table 8). Furthermore, the combined analysis of size and colour data revealed that the percentage accuracy was ranging from 13.79 – 84.38 % (Table 9).

Plate 2 shows the external appearance of rice seeds and caryopses. The red rice cultivars, At 362 and At 353, look similar in size, shape as well as in colour. Then the slender-grain cultivars, Bg 94-1, Bg 1165-2, Bw-Bs-1-2-31 look alike in shape and size. Also, small sized varieties Bg 358 and Bg 360 show similarities in size and shape while medium-sized varieties Bg 352, Bg 300, Bg 359, Bg 379-2, and Bg 357 have similar features.

DNA fingerprinting

Table 4 represents the numbers of alleles, band sizes of each allele and PIC values respectively. Based on the banding patterns observed, we selected a minimum set containing six SSR markers; *RM206*, *RM246*, *RM251*, *RM335*, *RM475*, and *RM23744* out of 23 markers that can differentiate exportable cultivars from the mega varieties. The banding patterns generated by each of the six selected markers duplexed with *K20* are given in Fig.1A. The UPGMA dendrogram drawn for the polymorphic bands observed for six markers is presented in Fig.1B, showing the cultivar identity based on the bands observed (Fig.1A). The polymorphic banding patterns obtained for six markers were verified using four plants per cultivar (Fig.2) and the DNA extracted from caryopses (Fig.3).

Table 6: The details of the GenBank accession numbers of the three SSR markers and no. of SNPs and no: of INDELS of three sequence alignments.

Cultivar	*Marker*								
	HvSSR12-34			***Seq7-8***			***RM23744***		
	Accession Number	**No: of SNPs**	**No: of INDELs**	**Accession Number**	**No: of SNPs**	**No: of INDELs**	**Accession Number**	**No: of SNPs**	**No: of INDELs**
At 362*	MK264379	5	7	MK293964	20	2	MK293976	7	5
Bg 94-1*	MK264380			MK293965			MK293977		
Bg 360*	MK264381			MK293966			MK293978		
Bg 1165-2*	MK264382			MK293967			MK293979		
Bw-Bs-1-2-31*	MK264383			MK293968			MK293980		
Bg 352	MK264384			MK293969			MK293981		
Bg 300	MK264385			MK293970			MK293982		
Bg 358	MK264386			MK293971			MK293983		
Bg 359	MK264387			MK293972			MK293984		
Bg 357	MK264388			MK293973			MK293985		
Bg 379-2	MK264389			MK293974			MK293986		
At 353	MK264390			MK293975			MK293987		

*Exportable rice cultivars

Table 7: The accuracy of discrimination of the rice cultivars based on the size traits of seeds and caryopses, estimated using Recursive Partitioning (R-PART) and C5.0 Classification models.

Sample origin (location and season)	% accuracy of discriminating 12 cultivars from each other based on size traits (100-unit weight, length and width)			
	Seeds		Caryopses	
	R-PART	C5.0	R-PART	C5.0
Field-*Maha*, 2015/2016	50.00	55.56	41.67	61.10
GH-*Maha*, 2015/2016	36.67	60.00	30.00	73.33
GH-*Yala*, 2016	42.42	81.81	42.42	75.76
Field-*Yala*, 2016	47.20	72.20	41.67	66.67
GH-*Maha*, 2016/2017	28.57	32.14	25.00	42.86
Samples combined	**52.12**	**60.74**	**52.76**	**58.28**

Table 8: The accuracy of discrimination of the rice cultivars based on the colour metrics of seeds and caryopses estimated using Recursive Partitioning (R-PART) and C5.0 Classification models.

Sample origin (location and season)	% accuracy of discriminating 12 cultivars from each other based on colour metrics (R, G and B)			
	Seeds		Caryopses	
	R-PART	C5.0	R-PART	C5.0
Field-*Maha*, 2015/2016	08.34	16.67	11.10	36.11
GH-*Maha*, 2015/2016	12.50	15.63	08.34	19.40
GH-*Yala*, 2016	03.13	12.50	03.13	12.50
Field-*Yala*, 2016	05.56	02.78	16.67	16.67
GH-*Maha*, 2016/2017	03.45	03.45	13.79	34.48
Samples combined	**12.65**	**06.63**	**14.46**	**16.67**

Table 9: The accuracy of discrimination of the rice cultivars based on the size traits and colour metrics of seeds and caryopses, estimated using Recursive Partitioning (R-PART) and C5.0 Classification models.

Sample origin (location and season)	% accuracy of discriminating 12 cultivars from each other based on size traits (100-unit weight, length and width) and colour metrics (R, G and B)			
	Seeds		Caryopses	
	R-PART	C5.0	R-PART	C5.0
Field-*Maha*, 2015/2016	30.56	58.33	30.56	61.11
GH-*Maha*, 2015/2016	40.00	80.00	43.33	63.33
GH-*Yala*, 2016	34.30	71.90	34.36	84.38
Field-*Yala*, 2016	27.78	63.89	36.11	77.78
GH-*Maha*, 2016/2017	13.79	31.03	24.14	48.28
Samples combined	**57.06**	**58.90**	**61.96**	**60.74**

Plate 2: Variation in morphological appearance of the seeds and caryopses of rice cultivars. The green circle on the left side contains the rice seeds, and the yellow circle on the right side contains the caryopses after dehulling for each cultivar. The sizes of the seeds and caryopses are indicated by the scale bar on the lower right corner of the Figure. Around 40 seeds and caryopses per rice cultivars were used in this Figure. Exportable rice cultivars are marked with *.

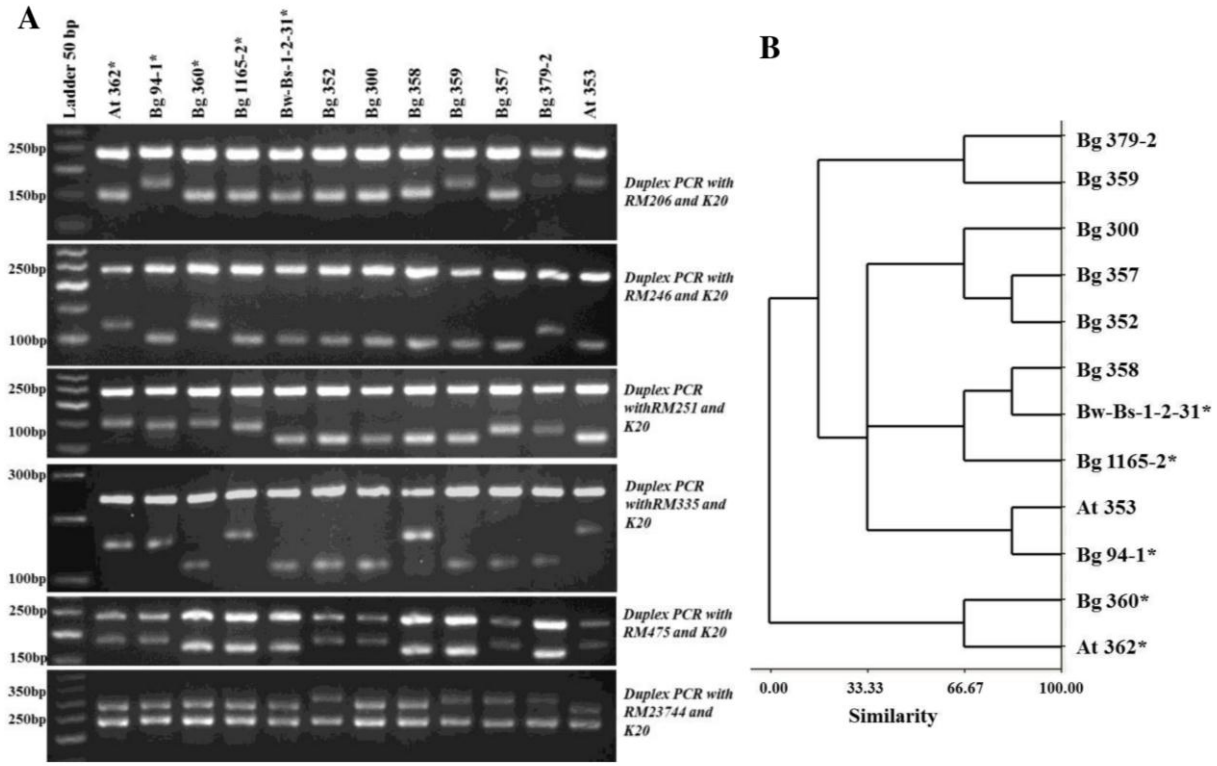

Figure 1: The banding patterns of the selected set of markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475*, and *RM23744*) for DNA fingerprinting and the dendrogram developed to depict the polymorphism. A: The composite agarose gel image for all tested markers. The names of SSR markers are indicated on the right side of the Figure, and the band sizes are shown on the left. Names of the rice cultivars are given on the top. Marker *K20* was used as a standard marker in the duplex PCR, and the 245 bp size band represents the amplified band for *K20* in duplex PCR while the rest of the bands represent the amplified bands for the other SSR marker. B: Dendrogram drawn using Complete Linkage and Euclidean Distance in Minitab 16. The exportable rice cultivars are marked with *.

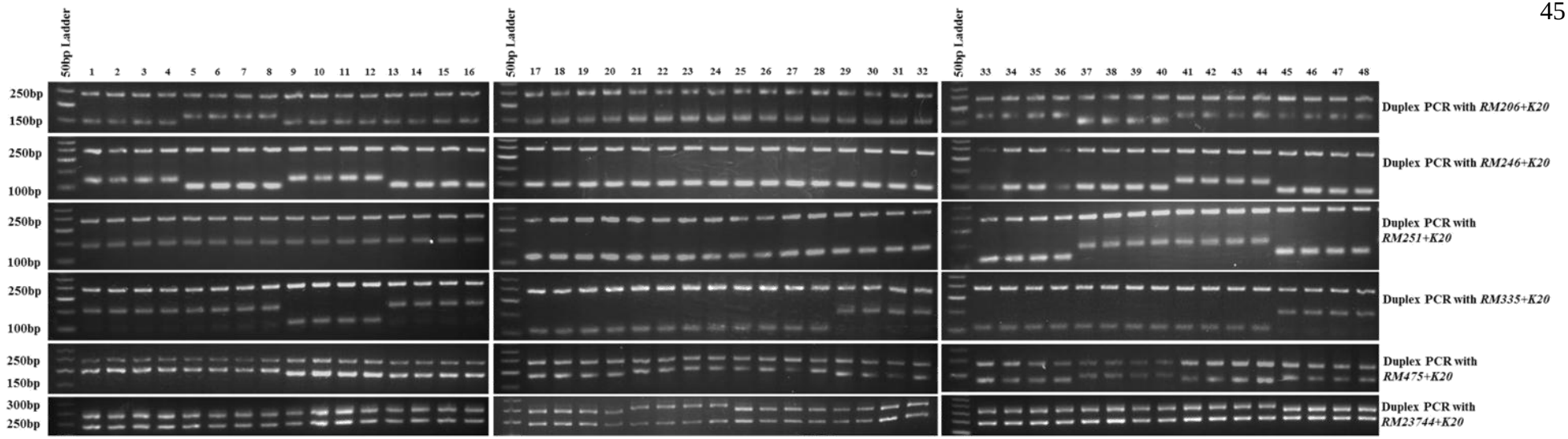

Figure 2: Composite gel image of the selected set of markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475* and, *RM23744*) which can be used in DNA fingerprinting (2.5 % agarose gel electrophoresis). The names of the markers are shown at the right side and corresponding band sizes are shown at the left side of the composite image. Four samples from each rice cultivar.1-4: At 362*, 5-8: Bg 94-1*, 9-12: Bg 360*, 13-16: Bg 1165-2*, 17-20: Bw-Bs-1-2-31 , 21-24: Bg 352, 25-28: Bg 300, 29-32: Bg 358, 33-36: Bg 359, 37-40: Bg 357, 41-44: Bg 379-2, 45-48: At 353. Exportable rice cultivars are marked with *.

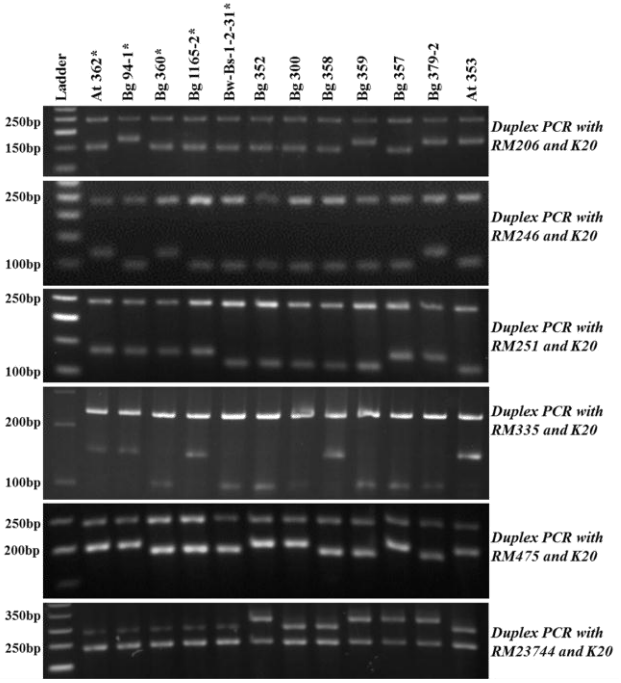

Figure 3: Composite gel image of duplex PCR for the DNA samples extracted from rice caryopsis amplified with the selected set of markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475*, and *RM23744*) duplexed with the marker *K20* (2.5 % agarose gel electrophoresis). The names of SSR markers are indicated on the right side of the Figure, and the band sizes are indicated on the left side. Names of the rice cultivars are given in the top. The exportable rice cultivars marked with a *. One kb ladder was used in gel electrophoresis. Marker *K20* was used as a monomorphic marker in the duplex PCR, and the 245 bp size band represents the amplified band for *K20* in duplex PCR while the rest of the bands represent the amplified bands for SSR marker in each duplex event.

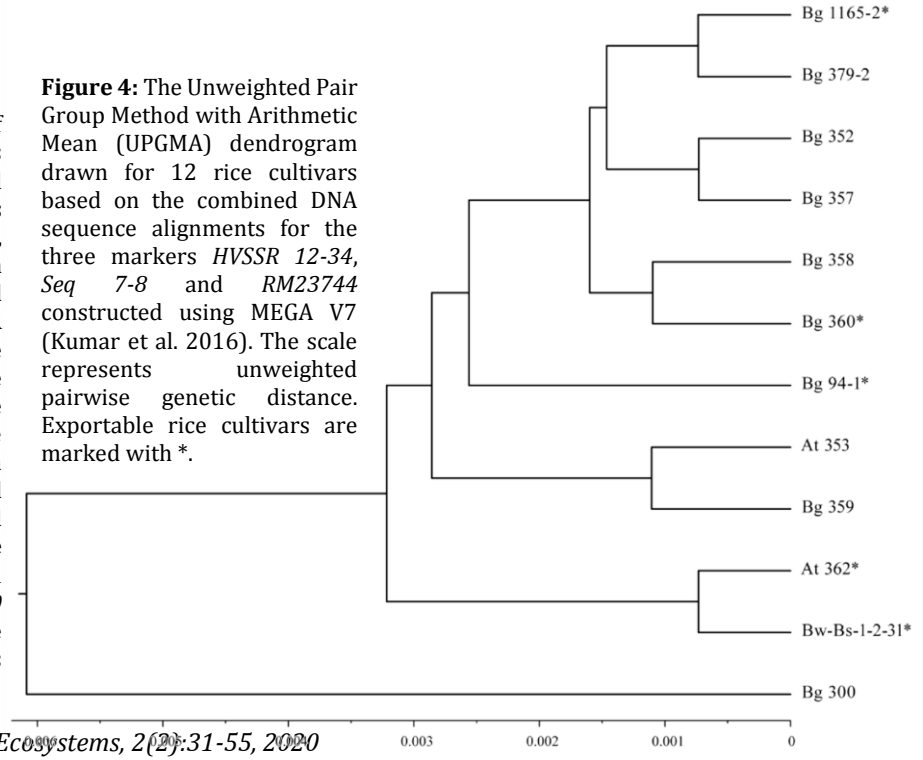

Figure 4: The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrogram drawn for 12 rice cultivars based on the combined DNA sequence alignments for the three markers *HVSSR 12-34*, *Seq 7-8* and *RM23744* constructed using MEGA V7 (Kumar et al. 2016). The scale represents unweighted pairwise genetic distance. Exportable rice cultivars are marked with *.

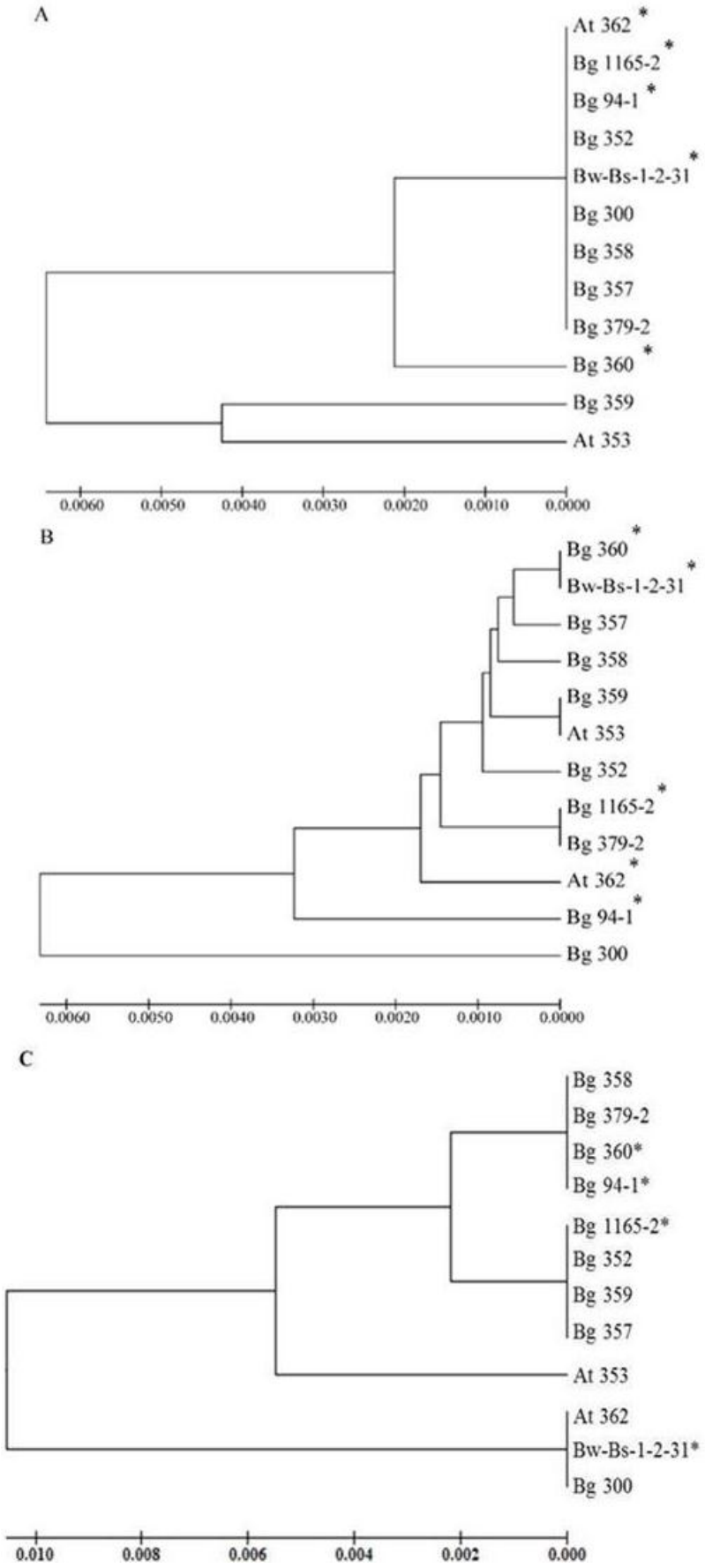

Figure 5: The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendograms drawn for 12 rice cultivars based on DNA sequence alignments separately for the three markers A: *HVSSR 12-34,* B: *Seq 7-8* and C: *RM23744* constructed using in MEGA V7 (Kumar et al. 2016). The scale represents unweighted pairwise genetic distance. Exportable rice cultivars are marked with *.

DNA sequencing

We screened 12 loci of the rice genome to be used in DNA sequencing and selected three loci based on the success of obtaining unambiguous sequence reads. The DNA sequences of the 12 cultivars were obtained for the *HvSSR12-34*, *Seq 7-8*, and *RM23744* markers to establish the varietal identities. When the sequences of all three loci were combined and subjected to UPGMA analysis, the cultivars were completely separated (Fig.4). However, the sequences of the three loci did not resolve the varietal identities when separately subjected to UPGMA dendrogram construction (Fig.5 A-C). The details of the sequence polymorphism of the three loci among 12 cultivars are given in Table 5.

4. Discussion

The precise definition of the plant breeders' rights (PBR) for newly bred varieties is essential to stimulate the breeding of new plant varieties and fund the breeders to employ high tech molecular breeding facilities in crop improvement programs. The policies to award PBR are firmly established in many countries (Barton 1982; Ginarte and Park 1997). However, in Sri Lanka, currently, legislators are drafting an act (Protection of the New Plant Varieties Act) to establish the required laws to award PBR. In line with this mission, in the present study, we are proposing a molecular strategy to define the uniqueness of five exportable rice cultivars at the caryopsis level in comparison to the mostly grown rice cultivars (i.e., mega varieties) in Sri Lanka. The present study processes the utmost significance as it is the first research project to employ DNA based varietal identification for rice in Sri Lanka. There are more than 2,000 rice accessions including landraces, wild types, and improved cultivars present in the country (Rathnathunga et al. 2016). However, there are only 84 cultivars released by the RRDI. Out of these 84 cultivars, only seven cultivars are considered as mega varieties (Table 1). The RRDI breeders reported that the mega varieties can get changed from time to time, however, only seven has been listed at the time of sampling. Therefore, we focused on the five exportable rice cultivars in comparison to only the seven mega varieties because seeds or caryopses of other accessions or cultivars are not coming to the export market.

We used the seeds and caryopses produced under greenhouse and field conditions and under *Yala* and *Maha* seasons of Sri Lanka to cover any variation caused by the environment in the morphometric analysis (Table 2). We assessed the applicability of morphometric parameters; size and colour traits, to differentiate the seed and caryopsis

samples of the cultivars. The R-PART and C5.0 decision tree algorithms used to analyze the morphometric data; only yielded an accuracy ranging from 25.00 - 81.81% to discriminate seeds and caryopses based on size traits (Table 7). The R-PART and C5.0 decision tree algorithms only yielded the accuracy range of 2.78 - 36.11 % to discriminate seeds and caryopses based on colour traits (Table 8). However, when we attempted to differentiate seeds and caryopses based on both size and colour traits, the percentage accuracy varied from 13.79 - 84.38% (Table 9). If R-PART and C5.0 algorithms yielded 100 % accuracy, the morphometric trait-based assessment could have been employed to differentiate rice cultivars and define their uniqueness. However, the percentage accuracy was less than 84.38 % leaving an error of 15.62 % causing recurrent ambiguity in defining the cultivar identities using morphological traits. The rice cultivar combinations subjected to discriminatory observations by human subjects (Table 3) estimated the K value is always 0.0 and mean IRR of 0.27. If the human subjects could differentiate seeds and caryopses samples of these combinations without any ambiguity, both K and IRR must be equal to 1.00. The mean IRR of 0.27 implies that the identification of the uniqueness of a rice cultivar based on the visual observations of seeds and caryopses has associated an error of 73.00%. Thereby we proved that the morphometric analysis could not be used to detect the uniqueness of rice cultivars based on seed and caryopsis appearances. The present study is the first time of using K and IRR statistics to differentiate the rice cultivars to define their uniqueness. However, there are reported studies using K and IRR on the identification of crop performance and field status to take management decisions (Peña-Barragán et al. 2011).

We screened SSR markers and identified six of them to define the uniqueness of five exportable rice cultivars from the seven mega varieties. In this study, the duplex PCR approach was followed by mixing the polymorphic SSR primer pairs with the primer pair of a monomorphic marker (*K20*) to enable the positive selection. The simplex PCR is doubtful because the absence of a band could also be a PCR failure. With the duplex PCR approach, the band for monomorphic marker should be present in all the cultivars assessed leading to the elimination of the confusions caused by PCR failures. The selected set of six SSR markers provided the required variability as revealed by the PIC values (Table 4). The gel images and the dendrogram given in Fig.1 define the uniqueness of the assessed rice cultivars. Although rice is a self-pollinated crop and all the cultivars tested were pure lines; we confirmed the bands obtained

using four replicates per cultivar and verified that there is no intra-cultivar DNA variability present for the assessed loci (Fig.2). When the DNA fingerprinting approach is employed in defining the uniqueness of the rice cultivars, the DNA extracted from caryopses must be used as the template. Therefore, after identifying six SSR markers and their required polymorphism using leaf DNA, we verified the banding patterns obtained using the DNA extracted from the caryopses (Fig.3).

The UPGMA dendrogram; developed based on the sequence polymorphism of the three loci selected based on the sufficient template lengths for sequencing, clear amplification and positive sequencing results (Hossain et al. 2015) (Table 5), revealed the identity of each cultivar (Fig.4). The legal authorities should provide the samples of exporting bulks to a DNA fingerprinting and sequencing laboratory whenever there is a need to detect the cultivar identities. The DNA fingerprinting method suggested in the present study is adequate to define the identities of exportable cultivars. However, if an independent verification is required, DNA sequencing of the detected loci followed by sequencing can be used. The K value indicating the accuracy precise detection was zero, implying that morphometric differentiation by human subjects is impossible which is further supported by the 73 % error detected in IRR statistic. The six SSR markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475,* and *RM23744*) and three sequenced loci (*Seq 7-8*, *HvSS12-34* and *RM23744*) could differentiate five exportable rice cultivars from seven mega varieties at caryopsis level.

5. Conclusions

The identity of the rice cultivars developed by RRDI (At 362, Bg 94-1, Bg 360, Bg 1165-2, and Bw-Bs-1-2-31) could not be established using morphometric trait analysis at caryopsis level in comparison to seven mega cultivars grown in Sri Lanka (Bg 352, Bg 300, Bg 358, Bg 359, Bg 357, Bg 379-2, and At 353). In the DNA fingerprinting analysis, a set of six SSR markers (*RM206*, *RM246*, *RM251*, *RM335*, *RM475*, and *RM23744*) that were selected out of 23 SSR markers differentiates all 12 rice cultivars and DNA sequencing of 12 cultivars with the selected three loci, (*Seq 7-8*, *HvSSR12-34* and *RM23744*) further authenticates the varietal distinctiveness.

6. Acknowledgments

University of Peradeniya, Sri Lanka (2016) Research Grant (No.: URG/2016/59/S).
The Director, Rice Research and Development Institute (RRDI), Bathalagoda, Sri Lanka.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

7. **References**

Aluwihare Y C, Ishan M, Chamikara M D M, Weebadde C K, Sirisena D N, Samarasinghe W L G, Sooriyapathirana S D S S (2016) Characterization and selection of phosphorus deficiency tolerant rice cultivars in Sri Lanka. Rice Sci 23(4): 184-195. DOI: 10.1016/j.rsci.2015.10.001.

Barton J H (1982) The international breeders' rights system and crop plant innovation. Science 216(4550): 1071-1075. DOI: 10.1126/science.216.4550.1071.

Caldo R, Sebastian L, Hernandez J, (1996) Morphology-based genetic diversity analysis of ancestral lines of Philippine rice cultivars. Philipp J Crop Sci 21(3): 86-92. ISSN: 0115-463X.

Chen X, Temnykh S, Xu Y, Cho Y G, McCouch S R (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor Appl Genet 95(4): 553-567. DOI: 10.1007/s001220050596.

Chin J H, Lu X, Haefele H, Gamuyao R, Ismail A, Wissuwa M, Heuer S (2010) Development and application of gene-based markers for the major rice QTL *Phosphorus uptake 1*. Theor Appl Genet 120: 1073-1086. DOI: 10.1007/s00122-009-1235-7.

Choudhury P R, Kohli S, Srinivasan K, Mohapatra T, Sharma R P (2001) Identification and classification of aromatic rice based on DNA fingerprinting. Euphytica 118(3): 243-251. DOI: 10.1023/A: 1017554600145.

Cooke R J, Reeves J C (2003) Plant genetic resources and molecular markers: variety registration in a new era. Plant Genet Resour Newsl 1(2-3): 81-87. DOI: 10.1079/PGR200312.

Coomb J J, Frank L M, Douches D S (2004) An applied fingerprinting system for potato using simple sequence repeats. Am Potato J 81(4): 243-250. DOI: 10.1007/ BF02871765.

Dangl G S, Yang J, Golino D A, Gradziel T (2009) A practical method for almond cultivar identification and parental analysis using simple sequence repeat markers. Euphytica 168(1): 41-48. DOI: 10.1007/s10681-008-9877-0.

Deleu W, Esteras C, Roig C, González-To M, Fernández-Silva I, Gonzalez-Ibeas D, Blanca J, Aranda M A, Arus P, Nuez F, Monforte A J (2009) A set of EST-SNPs for map saturation

and cultivar identification in Melon. BMC Plant Biol 9(1): 90-99. DOI: 10.1186/1471-2229-9-90.

DOA (Department of Agriculture), Sri Lanka (2017) Rice cultivation. Available at: www.doa.gov.lk/. Accessed on 12 November 2017.

Ginarte J C, Park W G (1997) Determinants of patent rights: A cross national study. Res Policy 26(3): 283-301. DOI: 10.1016/S0048-7333(97)00022-X.

Hossain H, Rahman M A, Alam M S, Singh R K (2015) Mapping of Quantitative Trait Loci (QTL) associated with reproductive-stage salt tolerance in rice. J Agron Crop Sci 201(1): 17-31. DOI: 10.1111/jac.12086.

Jain S, Jain R K, MacCouch S R (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently labeled microsatellite markers. Theor Appl Genet 109(5): 965-977. DOI: 10.1007/s00122-004-1700-2.

Jamil M, Rana I A, Ali Z, Awan F S, Shahzad Z, Khan A S (2013) Estimation of genetic diversity in rice (*Oryza sativa* L.) cultivars using Simple Sequence Repeats. Mol Plant Breed 4. DOI: 10.5376/mpb.2013.04.0036.

Kekulandara D S, Bandaranayake P C G, Sirisena D N, Samarasinghe W L G, Suriyagoda L D B (2017) Temporal tillering behavior of Sri Lankan elite rice varieties in response to phosphorus availability. J Trop Agric 28(2): 133-143.

Kennedy R L, Lee Y, Van Roy B, Reed C D, Lippman R P (1998) Solving Data Mining Problems Through Pattern Recognition. Prentice-Hall.

Khush G S, Brar D S, Hardy B (2003) Advances in rice genetics 1, International Rice Research Institute.

Kim S, Plagnol V, Hu T T, Toomajian C, Clark R M, Ossowski S, Ecker J R, Weigel D Nordberg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nat Genet 39(9): 1151-1155. DOI: 10.1038/ng2115.

Kjeldgaard R H, Marsh D R (1994) Intellectual property rights for plants. Plant Cell 6(11): 1524. DOI: 10.1105/tpc.6.11.1524.

Kuhn M (2013) Classification using C5.0 Use R! Groton CT: Pfizer Global R D Jiawei Han, Data Mining: Concepts and techniques.

Kumar S, Strecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics

Analysis Version 7.0 for bigger datasets. Mol Biol Evol 33(7): 1870-1878. DOI: 10.1093/molbev/msw054.

Kwon Y S, Lee J M, Yi G B, Yi S I, Kim K M, Soh E H, Bae K M, Park E K, Song I H, Kim B D (2005) Use of SSR markers to complement tests of distinctiveness, uniformity, and stability (DUS) of pepper (*Capsicum annuum* L.) cultivars. Molecules and Cells 9(3): 428-435. PMID: 15995361.

Lu L, Yan W, Xue W, Shao D, Xing Y (2012) Evolution and association analysis of *Ghd7* in rice. PLoS ONE 7(5): e34021. DOI:10.1371/journal.pone.0034021.

Luther Z, Akromah R, Nyadanu D, Tokpah D P, Page Z, Voor V M, Kwaloe A D (2017) Evaluation of genetic diversity in rice (*Oryza sativa* and *Oryza glaberrima*) germplasm from Liberia and Ghana using simple sequence repeat (SSR) markers. Afr J Biotechnol 16(41):1990-1996. DOI: 10.5897/AJB2017.16212.

McCouch S R, Chen X, Panaud O, Temnykh S, Xu Y, Che Y G, Huang N, Ishi T, Blair M (1997) Microsatellite marker mapping, development and applications in rice genetics and breeding. Plant Mol Biol 35(1-2): 9-99. DOI: 10.1023/A: 1005711431474.

McHugh M L (2012) Interrater reliability: the kappa statistic. Biochem Med 22(3): 276-282. PMID: 23092060.

McNally K L, Bruskiewich R, Mackill D, Buell C R, Leach J E, Leung H (2006) Sequencing multiple and diverse rice varieties; connecting whole-genome variation with phenotypes. Plant Physiol 141(1): 26-31. DOI: 10.1104/pp.106.077313.

Moria S, Iwanami H, Okada K, Yamamoto T, Abe K (2011) A practical method for apple cultivar identification and parent-offspring analysis using simple sequence repeat markers. Euphytica 177(1): 135-150. DOI: 10.1007/s10681-010-0295-8.

Mukherjee B, Das P, Alam Q, Nath D, Dasgupta T, (2013) Molecular characterization of rice cultivars using microsatellite markers. Oryza 50(1): 35-40. DOI: 10.5958/0975-928X.2017.00072.2.

Nybom H, Weising K, Rotter B (2014) DNA fingerprinting in botany: past, present, future. Investig Genet 5: 10. DOI: 10.1186/2041-2223-5-1.

Peña-Barragán J M, Ngugi M K, Plant R E, Six J (2011) Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens Enviro 115(6): 1301-1316. DOI: 10.1016/j.rse.2011.01.009.

Pradheeban L, Nissanka S P, Suriyagoda L D B (2015) Screening commonly cultivated rice cultivars in Sri Lanka with special reference to Jaffna for salt tolerance at seedling stage under hydroponics. IJAAR 7(5)1-13.

Rahman M S, Molla M R, Alam M S, Lutfur R. (2009) DNA fingerprinting of rice (*Oryza sativa* L.) cultivars using microsatellite markers. Aust J Crop Sci 3(3): 122-128. ISSN: 1835-2707.

Rahman M S, Sohag M K H, Rahman L (2010) Microsatellite based DNA fingerprinting of 28 local rice (*Oryza sativa* L.) varieties of Bangladesh. JBAU 8(1): 7-17. DOI: 10.3329/jbau. v8i1.6391.

Rajkumar G, Weerasena J, Silva R, Fernando K (2016) Genomic diversity of Sri Lankan new improved rice varieties revealed by AFLP markers. Int J Appl Sci Biotechnol 4(1): 32-38. DOI: 10.3126/ijasbt. v4i1.13768.

Rambaut A (2014) FigTree, a graphical viewer of phylogenetic trees. Available at: http://tree.bio.ed.ac.uk/software/figtree. Accessed on 27 January 2018.

Rathnathunga E U U, Senanayake G, Dissanayake N, Seneweera S. Geekiyanage S (2016) Development of a mini core collection from Sri Lankan traditional rice for flowering time variation. Aust J Crop Sci 10(9)1357-1367.

Rongwen J, Akkaya M S, Bhagwat A A, Lavi U, Cregan P B (1995) The use of microsatellite DNA markers for soybean cultivar identification. Theor Appl Genet 90(1): 43-48. DOI: 10.1007/BF00220994.

RRDI (2018). Recommended Rice Varieties in Sri Lanka (1958-2016). (Research and Development Institute, Department of Agriculture, Bathalagoda, Ibbagamuwa, Sri Lanka.

Salgotra R K, Gupta B B, Bhat J A, Sharma S (2015) Genetic diversity and population structure of basmati rice (*Oryza sativa* L.) germplasm collected from North Western Himalayas using trait linked SSR markers. PLoS ONE 10(7): e0131858. DOI: 10.1371/journal.pone.0131858.

Sato H, Suzuki Y, Sakai M, Imbe T (2002) Molecular characterization of *Wx-mq*, a novel mutant gene for low-amylose content in endosperm of rice (*Oryza sativa* L.). Breed Sci 52(2): 131-135. DOI: 10.1270/jsbbs.52.131.

Shirasawa K, Monna L, Kishitani S, Nishio T (2004) Single nucleotide polymorphisms in

randomly selected genes among japonica rice (*Oryza sativa* L.) varieties identified by PCR-RF-SSCP. DNA Res 11(4): 275-283. DOI: 10.1093/dnares/11.4.275.

Shoda M, Urasaki N, Sakiyama S, Terakami S, Hosaka F, Shigrta N, Nishitani C, Yamamoto T (2012) DNA profiling of pineapple cultivars in Japan discriminated by SSR markers. Breed Sci 62(4): 352-359. DOI: 10.1270/jsbbs.62.352.

Singh H, Deshmukh R K, Singh A, Singh A K, Gaikwad K, Sharma T R, Mahapatra T, Singh N T (2010) Highly variable SSR markers suitable rice genotyping using agarose gels. Mol Breed 25: 359-364. DOI: 10.1007/s11032-009-9328-1.

Singh N, Choudhury D R, Singh A K, Kumar S, Srinivasan K, Tyagi T K, Singh N K, Singh R (2013) Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. PLoS ONE 8(12): e84136. DOI: 10.1371/journal.pone.0084136.

Suriyagoda L D B, Thilakarathne R M M S, Nissanka S P, Samita S (2011) Morphological variation in selected rice (*Oryza sativa* L.) germplasm of Sri Lanka. J Natl Sci Found 39(2): 35-43. DOI: 10.4038/jnsfsr.v39i2.3173.

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.). Theor Appl Genet 100(5): 697–712. DOI: 10.1007/ s001220051342.

Therneau T, Atkinson E, (1997) An introduction to recursive partitioning using the RPART routines.

timeanddate.com (2018). Available at: https://www.timeanddate.com/sun/ @1230553?month=12 year=2016, Accessed on 19 June 2018.

Weising K, Nybom H, Pfenninger M, Wolff K, Kahl G (2005) DNA Fingerprinting in Plants; Principles, Methods, and Applications. 2nd edition, pp 2-245. CRC Press Taylor and Francis Group, 6000, Broken Sound Parkway NW, Boca Raton, FL 33487-2742.

Wijayawardhana H C D, Herath H M V G, Weerasinghe P A, Herath H M D A K (2015) Morphological variation in selected Sri Lankan rice (*Oryza sativa* L.) accessions in relation to the vegetative parameters. J Trop Agric 26(2): 380-389. DOI: 10.4038/tar.v26i2.8100.

Williams S B (1984) Protection of plant varieties and parts as intellectual property.

Science 225(4657): 18-23. DOI: 10.1126/science.225.4657.18.

World Weather Online (2018) Available at:https://www.worldweatheronline.com, Accessed on 19 June 2018.

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(4): 421-426. ISSN: 1836-3644.