



Research Article

GENOMIC DIVERSITY OF SRI LANKAN NEW IMPROVED RICE VARIETIES REVEALED BY AFLP MARKERS

Gowri Rajkumar^{1,2*}, Jagathpriya Weerasena², Rangika Silva³ and Kumudu Fernando⁴

¹ Department of Botany, Faculty of Science, University of Jaffna, Sri Lanka

² Institute of Biochemistry Molecular Biology and Biotechnology, University of Colombo, Sri Lanka

³ Department of Computer Science, Faculty of Applied Science, University of Sri Jeyawardenapura, Sri Lanka

⁴ Agricultural Biotechnology Centre, Faculty of Agriculture, University of Peradeniya, Sri Lanka

*Corresponding author's email: gowrir@jfn.ac.lk

Abstract

Genetic relationships among 28 new improved rice varieties were established using Amplified Fragment Length Polymorphism (AFLP) markers. Cultivars were analyzed with 10 *EcoRI* and *MseI* primer combinations. A total of 517 fluorescent AFLP markers were generated and analyzed. Of these 480 fragments were polymorphic (92.84%) and 37 (7.16%) fragments were monomorphic. The Jaccard's similarity indices (J) based on the AFLP profiles of the 28 varieties were computed and Unweighted Pair Group Method with Arithmetic mean (UPGMA) based dendrogram was constructed. The dendrogram separated varieties into three major clusters. Outliers used in the study were uniquely separated from the rest confirming the reliability of data and analysis. The Cophenetic correlation with 0.862 strongly supported the clustering pattern of UPGMA dendrogram. Principal Coordinate analysis and the unrooted tree also confirmed the clustering pattern of the UPGMA dendrogram. Rice varieties in the same cluster showed similar characteristic features (Eg. Grain colour, life span etc). Therefore this genetic diversity data at molecular level will provide detailed estimates of the genetic variation among Sri Lankan new improved rice varieties and also useful in *ex situ* and *in situ* genetic conservation, utilization and exchange of genetic material.

Keywords: AFLP; rice; polymorphism; improved varieties

Introduction

Rice is the staple food of 20 million Sri Lankans and the livelihood of more than 1.8 million farmer families (Sandika and Dushani, 2009). Rice accounts for about 45% of the per capita calories and 40% of the per capita protein consumption in the average Sri Lankan diet and per capita consumption of milled rice estimated 106 kg per head per year (Department of Census and Statistics, Sri Lanka 2005). Sri Lanka is considered as one of the secondary diversity centre for rice genetic resources (Ikeda and Vaughan 1991). The varieties cultivated from ancient time to middle of last century were all traditional types. In ancient times, farmers cultivated traditional rice varieties, because of their adaptability to Sri Lankan soil types, climate, geography and harsh environmental conditions even though the grain yield potential of those varieties were very low (Rajkumar *et al.*, 2011). Therefore high yielding new improved varieties have been bred from crosses between traditional as well as exotic genotypes to gain better yield. Almost all rice varieties grown today in Sri Lanka are new improved varieties. Total rice production in Sri Lanka has gradually increased during the past 30 years. The 1970s and 1980s were dedicated to developing high yielding varieties

resistance to a host of major pest and diseases prevalent in Sri Lankan rice fields such as brown planthopper, rice gall midge, rice blast and bacterial leaf blight. Sri Lankan scientists have developed more than 60 New Improved Varieties (NIV) during the past 50 years and some of them are now widely adopted not only in Sri Lanka but released in other Asian, African and Latin American countries (IRRI report). The major attributes of NIVs are improved photosynthetic efficiency, enhanced fertilizer responsiveness, resistance to lodging and short maturity duration and resistance /tolerance to stress factors which are either biotic or abiotic such as salinity, iron toxicity, water stress. In Sri Lanka, currently 95 % of the rice extent is under NIV. Sri Lanka at present is self-sufficient in rice production as a result of these achievements.

In recent years numerous molecular phylogeny markers revealed the genetic diversity/similarity between closely related organisms (Loh *et al.*, 1999). In this study, Fluorescent AFLP analysis (FAFLP) was performed with several primer pairs and found sufficient variation to draw conclusions about the genetic relationship within and between Sri Lankan new improved varieties. Among various marker systems AFLP is widely accepted as an effective tool for identifying genomic differences.

Furthermore, fluorescent labeling in FAFLP replaces the radioactive labeling and increases the throughput by enabling automated detection and scoring of fragments generated in AFLP without prior sequence knowledge of genomic DNA.

Objectives of this study were to assess the genetic diversity among some most common improved rice varieties available in Sri Lanka by AFLP markers and provide more direct and reliable genetic information for selecting suitable parents in rice breeding programs and enhance *ex situ* and *in situ* genetic conservation, utilization and exchange of genetic material.

Materials and Methods

Plant Materials

The 28 rice genotypes and two outliers were used in the study are listed in the Table 1. All the genotypes are under the category of New Improved Rice varieties. Seeds were

obtained from the seed bank of Plant Genetic Resources Centre (PGRC) at Gannoruwa, Peredeniya, Sri Lanka. The seeds were planted in pots of same size filled with soil collected from paddy field and grown under greenhouse conditions without adding fertilizer. After two weeks the young tender leaves were harvested from randomly selected plants from each variety and surface sterilized according to the procedure described by Zhang *et al.*, (1997) then stored at -80 °C until extraction of DNA.

DNA Extraction and Quantification

Genomic DNA was extracted from rice leaves using modified CTAB based procedure described by Chen and Ronald (1999). The extracted DNA was quantified by agarose gel electrophoresis followed by visualization under the UV transilluminator with standard DNA concentration markers. Then concentrations of all DNA samples were adjusted to 300ng/μl.

Table 1: List of new improved rice varieties used in this study.

Selection No	Accession no	Name	Age (months)	Yield (t/ha)	Grain pericarp colour	Attributes
1	2828	Bg 11-11	4 - 4 1/2	6.5	white	
2	2830	Bg 379-2	4 - 4 1/2	8.5	white	Resistant to Bph and BB
3	2836	Bg 380	4 - 4 1/2	10.0	white	Resistant to GM-1
4	2837	Bg 350	3 1/2	6.0	red	Resistant to GM-1
5	2840	Bg 300	3	8.0	white	Resistant to GM-1, BL,BB ,Bph
6	2846	Bg 745	5- 6	6.0	white	PS
7	4010	Bg 3-5	5- 6		white	
8	4017	Bg 90-2	4 - 4 1/2	6.5	white	
9	8919	Bg 357	3 1/2	10.5	white	Resistant to GM-1& 2, BL, BB, Bph.
10	8923	Bg 359	3 1/2	7.0	white	Resistant to GM 1 & 2, BL, BB, Bph,
11	11630	At 306	3	6.5	white	Rsistant to BL, BB, Bph and susceptible to GM.
12	6311	Bw267-3	3 1/2	4.5	white	Resistant
13	8920	Bg 360	3 1/2	7.0	white	Resistant to GM-1 ,GM-2, BL,Bph
14	10325	At 362	3 1/2	8.0	red	
15	10591	Ld 356	3 1/2	4.5	red	resistant to iron toxicity
16	10935	Bw 363	3 1/2	8.0	white	
17	8682	Bg 403	4 - 4 1/2	8.0	white	Resistant to BB, BL and Bph
18	8540	Bg 304	3	7.5	white	Resistant to GM 1&2, BL, BB, ,Bph
19	5310	Bw 351	3 1/2	5.0	red	resistant to sheath blight
20	9102	At 303	3	5.0	red	Resistant to BL
21	7182	Bg 352	3 1/2	8.5	white	Resistant to BL,BB & GM-1,Bph
22	9103	At 353	3 1/2	6.5	red	resistant to BL and BB
23	9055	Bw452	4 - 4 1/2	5.0	red	
24	2843	Bg301	3	6.0	red	Tolerant to drought and resistant to BL and BB
25	10446	Bw361	3 1/2	8.0	red	Resistant to GM and Bph, BL
26		At402	4 - 4 1/2	7.5	red	
27	2844	Bg400-1	4 - 4 1/2	8.5	white	Resistant to GM-1 and BB
28	9478	Bg358	3 1/2	9.0	white	High yielding samba variety

BB: Bacterial leaf blight BL: Rice blast disease GM-1: Biotype one of rice gall midge

GM-2: Biotype two of rice gall midge

Bph: Brown plant hopper PS: Photo period sensitive

(Source: Jeyawardena *et al.*, 2010, RRD, Batalagoda, Department of Agriculture, Sri Lanka)

AFLP Analysis

Fluorescent AFLP analysis was performed as described by Vos *et al.* (1995) with minor modifications in amount of initial genomic DNA, PCR components and thermal cyclic

conditions. Briefly, 1 μg of genomic DNA was digested with 5 units of *EcoRI* and 5 units of *MseI* enzymes by incubating at 37 °C for 3 1/2 hours. After the double digestion of DNA, Specific synthetic double stranded

adapters, *EcoRI* adaptor (10 pmol/ μ l), *MseI* adaptor (10 pmol/ μ l) for each restriction site were ligated by using 5U of T4 DNA ligase (New England Biolabs, USA) to generate a DNA template for amplification. The ligation mixture was incubated at 37°C in a water bath for overnight (~16 h). Then 2 μ l of digested/ligated DNAs were preamplified in 25 μ l of reaction containing 20 pmol of each preamplification primers, 0.5 mM dNTPs, 1U of Taq DNA polymerase (5U/ μ l, Genscript USA) and 1X PCR buffer containing 1.5mM MgCl₂ (Genscript USA). The PCR amplification was performed on thermal cycler (Eppendorf® Master cycler gradient) following temperature profile: 30 cycles of denaturation at 94 °C for 30 sec, annealing at 56 °C for 30 sec, extension at 72 °C for 60 sec. The preamplification product was then diluted 20 times with sterile distilled water and used as a template for selective amplification.

The selective amplification reaction was conducted in 20 μ l reaction containing 5 μ l of diluted preamplification product, fluorescently labeled (HEX, TMR or FAM) *EcoRI* primer (5.6 pmol), *MseI* primer (5.6 pmol), 0.5 mM dNTPs, 1U of Taq DNA polymerase (5U/ μ l, Genscript USA) and 1X PCR buffer containing 1.5mM MgCl₂ (Genscript USA). The PCR amplification was carried out as follows on thermal cycler, Denaturation at 94 °C for 30 sec, annealing at 58 °C for 30 sec. in the first cycle and then the annealing temperature was reduced by 0.7 °C per cycle for the next 12 cycles. Thereafter the cycles were repeated 23 more times with the annealing temperature of 49 °C. Extension was at 72 °C for 1 min for all cycles. 10 different primer combinations with different fluorescent labeled *EcoRI* primers were used for selective amplifications.

Selective PCR amplification products were purified by ethanol precipitation and resuspended in sterile water.

Capillary Electrophoresis

Purified PCR product (2.5 μ l) was mixed with ET550-ROX size standard (GE Healthcare Life Sciences UK) and deionized formamide according to the manufactures instructions. Then the samples were denatured at 95 °C for 2 minutes and capillary electrophoresis was performed on automated MegaBACE 1000 DNA sequencer (GE Healthcare Life Sciences UK) by injecting samples at 3kv for 45sec and run at 10 kv for 75 min. The AFLP finger printing for each sample was repeated to ensure the reproducibility.

Data analysis

Peaks in the range of 30-550 bp of the electropherogram were analyzed and compared by using MegaBACE Genetic Profiler software version 2.2 (GE Healthcare Life Sciences UK). Each sample was analyzed and compared in duplicates and unambiguous, reproducible peaks in the electropherogram were scored as AFLP markers and inconsistent fragments in duplicates were not included in the analysis. A tolerance limit of 1.0 base and a minimum

peak height of 100 were consistently maintained while scoring the peaks. Each differently sized fragment was treated as a unique character and converted to binary data as presence (1) and absence (0). This binary data was used to compute Jaccard's similarity coefficients (Jaccard, 1901) and similarity coefficient matrix was generated to assess the genetic resemblances among varieties by using Multi Variate Statistical Package (MVSP 3.1) (Kovach 1998). Then the similarity matrix was used for cluster analysis by Unweighted Pair Group Method with Arithmetic mean (UPGMA) as suggested by Sneath and Sokal (1973) and the dendrogram was generated by MVSP 3.1 software. Average genetic similarity between all accession/varieties was used to establish a cut off value for cluster analysis for the dendrogram.

Nei and Li's similarity coefficients (Nei and Li 1979) were also computed and UPGMA cluster analysis was performed using NEIGHBOR program in PHYLIP version 3.69 (Felsenstein, 2009). An unrooted tree was generated by TREEVIEW program (Roderic, 1996) by using tree file generated by PHYLIP software.

Principal Coordinate Analysis (PCoA) was performed between the accessions using Gower General similarity coefficient (Gower and Legendre 1986) to find out possible relations that could not be visualized in cluster analysis. In PCoA analysis, data were plotted on first two dimensions (PCo1 and PCo2) using MVSP software (Kovach, 1998). A cophenetic correlation coefficient (Sokal and Rohlf, 1962) which measures the best fit model to the data (cluster analysis with the associated similarity matrix) was assessed by Mantel test (Mantel, 1967) as describe in Abhijit *et al.* (2004).

Results

Level of Polymorphism

A total of 517 fluorescent AFLP markers were generated from ten primer combination. Of these 480 fragments were polymorphic (92.84%) and 37 (7.16%) fragments were monomorphic. The high level of polymorphism observed for each primer combination. Sixteen *EcoRI* and *MseI* selective amplification primer combinations were initially screened and only ten pairs of primers which generated good amplification and distinct polymorphism with all 28 improved rice varieties were selected for data analysis. The number of bands and the degree of polymorphism revealed by each primer combination is given in Table 2. The number of amplification products generated by individual pair of primer, ranged from 27 (E-AA $\times$ M-G) to 94 (E-AT $\times$ M-G) with an average of 51.7 fragments per pair of primers. Percentage of polymorphism ranged from 88.29 (E-AT $\times$ M-G) to 96.15 (E-AG $\times$ M-T) with an average of 92.84% polymorphism.

Genetic relationship and cluster analysis

The results obtained by the Jaccard's similarity coefficients (Fig 1) showed that the genetic similarity varies from 0.231 to 0.643. Varieties Bg360 and Bw267-3 showed highest genetic similarity value (0.643) whereas varieties Bg379-2 and Bg357 showed lowest genetic similarity value (0.231). The UPGMA dendrogram (Fig. 1) separated the accessions into three main clusters at similarity coefficient of 0.310. Cluster I and IV contains only one accession namely *Hygroryza* and TN(1) respectively. Cluster II encloses 7 varieties. Cluster III is subdivided into two clusters III A and III B at the similarity coefficient value of 0.321. Cluster III A encloses 14 varieties whereas cluster III B contains 7 varieties. *Hygroryza aristata* was found to be the most divergent line and clustered in Cluster I by separating from

all the other accessions at the similarity coefficient value of 0.281.

Cophenetic correlation coefficient(*r*) from the comparison between the dendrogram and the similarity matrix was 0.862 (without bootstrap replicates) and it was 0.819 with 1000 bootstrap replicates.

The results of the principal coordinate (PCoA) analysis supported the clustering pattern of the UPGMA dendrogram generated. Based on PCoA analysis, three well defined groups were identified (Fig. 2).

Unrooted neighbor-joining tree generated on the basis of Nei and Li similarity coefficient by using TREEVIEW program of PHYLIP software also revealed three branches among Sri Lankan new improved rice varieties (Fig. 3).

Table 2: Selective amplification primer combinations and percentage of polymorphism for each primer combination.

Primer combination	Total No of bands	No of Polymorphic bands	Percentage of polymorphism
E-AT × M-G	94	83	88.29
E-CA × M-G	61	55	90.16
E-AA × M-G	27	25	92.59
E-AC × M-T	46	43	93.47
E-AG × M-T	52	50	96.15
E-AT × M-C	50	48	96.00
E-AT × M-T	62	59	95.16
E-AA × M-C	42	40	95.23
E-AG × M-A	49	46	93.87
E-AT × M-A	34	31	91.17
Total	517	480	92.84

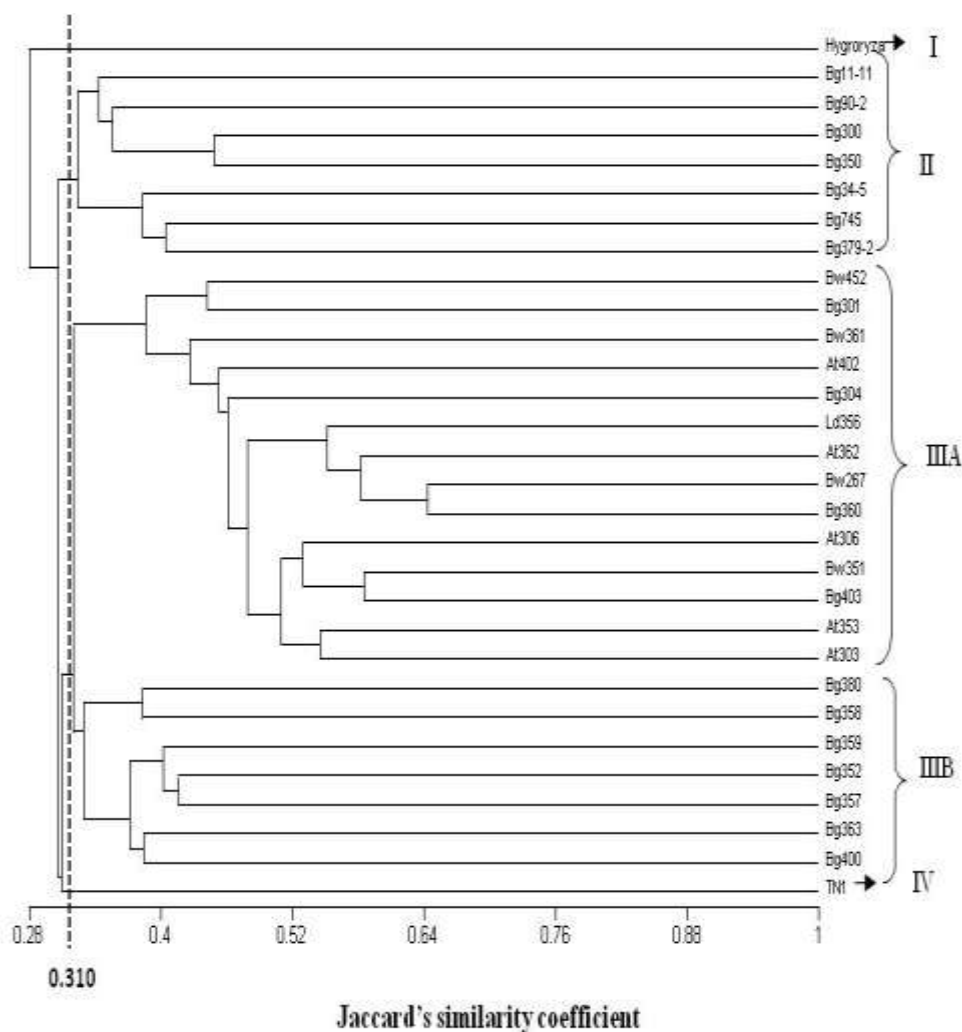


Fig. 1: UPGMA dendrogram based on Jaccard similarity coefficient of Sri Lankan improved rice varieties.

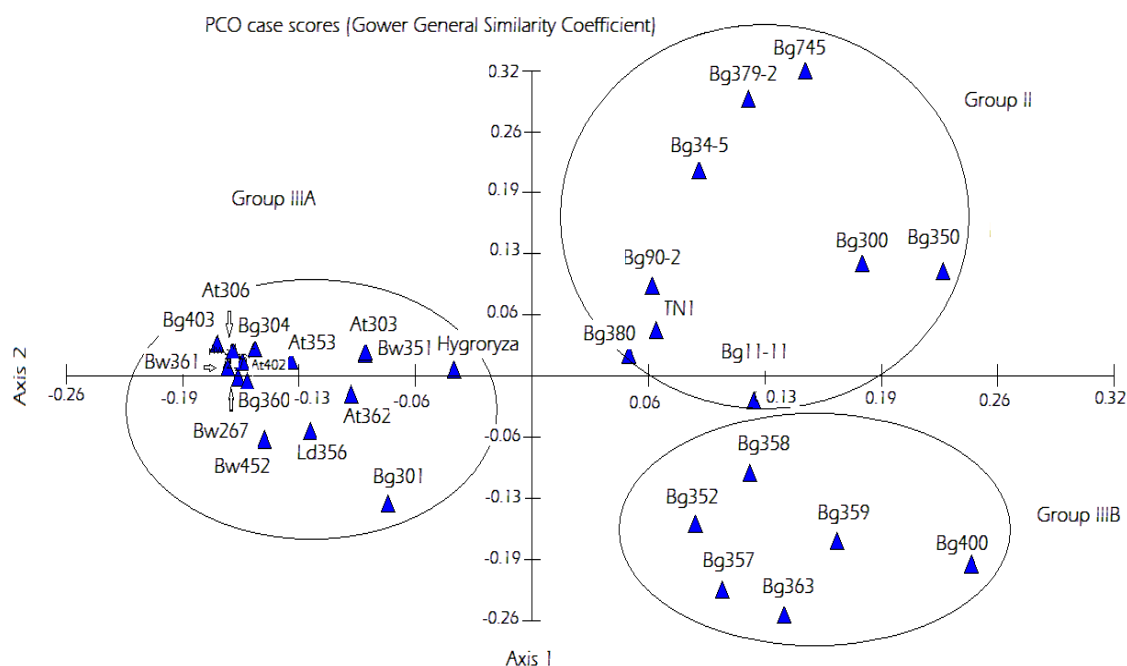


Fig. 2: Scatter diagram of first two coordinates (PCo1and PCo2) revealed three well defined groups of varieties.

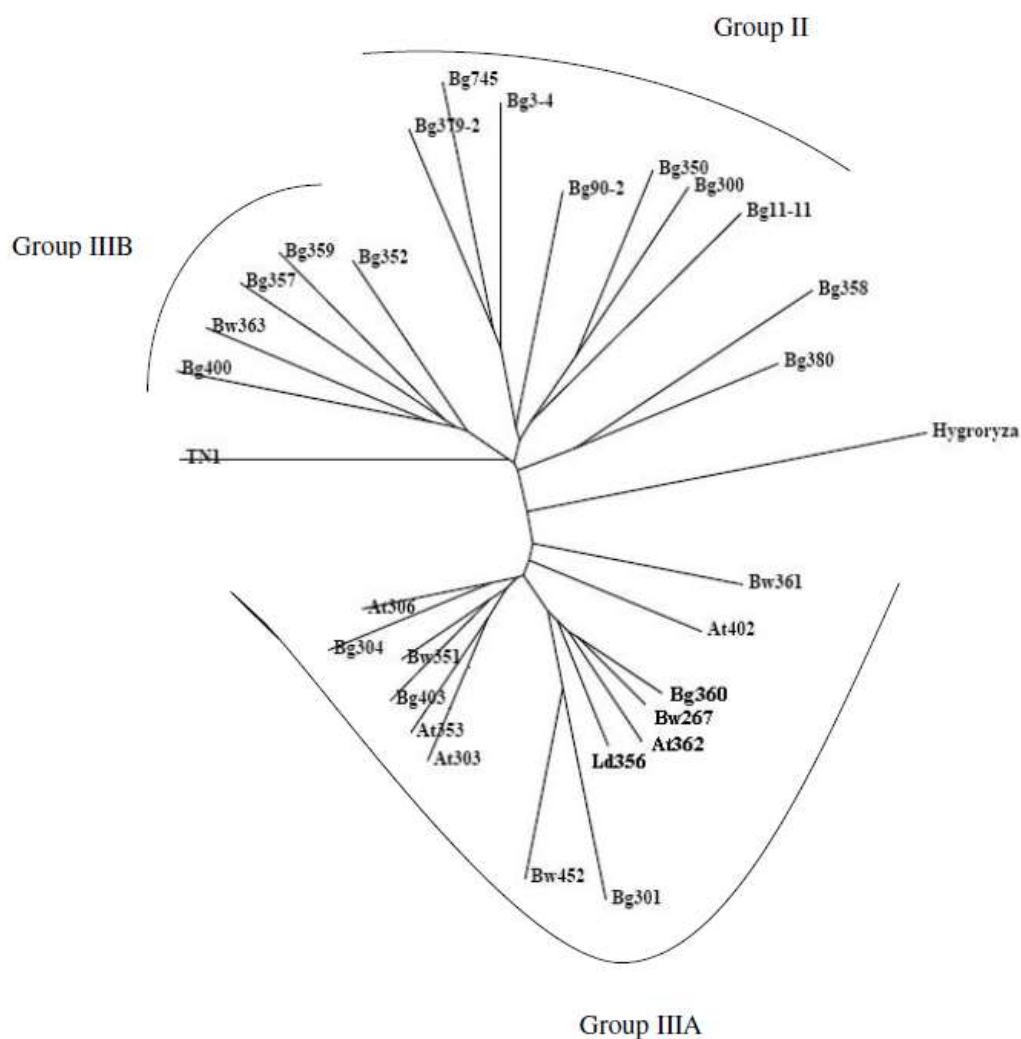


Fig. 3: Unrooted neighbor-joining tree generated on the basis of Nei and Li similarity coefficient of Sri Lankan new improved rice varieties.

Discussion

This is the first report on genetic variability and relationship among Sri Lankan new improved rice varieties based on FAFLP markers. Previous research was conducted on genetic diversity of Sri Lankan traditional and wild rice species (Rajkumar *et al.*, 2011). The high level of polymorphism observed (92.84%) using different primer combinations indicates a remarkable amount of intraspecific genetic variation among Sri Lankan rice varieties. Therefore these high level of polymorphism allow for the selection of parents and the identification of true hybrids in breeding programs (Garcia *et al.*, 2004).

UPGMA dendrogram clustered the rice varieties into three (Figure 1). Cluster I contains only one sample from genus *Hygroryza*. It is more closely related to genus *Oryza*. Which was previously classified under genus *Oryza* but now it is separated into different genus *Hygroryza*. But both genus *Hygroryza* and *Oryza* are under same tripe. Cluster IV represents one modern rice variety Taichung Native 1(abbr. TN-1), which is the first *Indica* variety carrying the well-known Dee-geo-woo-gene (semi dwarfing gene). Unique

separation of these two outliers confirming the reliability of our data used in this study and analysis methods performed.

Varieties found in cluster II were high yield potential but their life spans are more than 4 months except Bg300 and Bg350. Therefore those were currently removed from cultivation. Varieties Bg11-11 and Bg90-2 were released by Sri Lankan scientist during 1970s as those were resistance to abiotic stresses but later they became susceptible to disease like brown plant hopper, blast and bacterial blight disease and removed from cultivation. These varieties were introduced to some other countries from Sri Lanka and those were more resistance to biotic stresses, high yield and most popular in those countries (China).

Eleven varieties out of fourteen in cluster IIIA are red grain varieties. Any red grains variety is not grouped in other clusters. In Sri Lanka red grain varieties are more admired by public because of its high nutritional value. Life span of varieties found in this cluster were 3,1/2 months except Bw452, Bg301 and At402.

All varieties in cluster III B are recommended and cultivated varieties as those yield potential are very high (more than 7.5 tons/ha) and life span of varieties found in this cluster is 3,1/2 months. Therefore most rice accessions found in same cluster showed similar characteristics. All the varieties in IIIB are resistance to GM1 and GM2. Most of the varieties in cluster IIIA are resistance to BB (At306, Bg403, Bg304, At353 and Bg301).

According to the genetic ancestry map released from Rice Research and Development Institute (RRDI) of Sri Lanka, varieties 379-2, Bg 745, Bg 90-2 and Bg 350 were derived from the same parental ancestors and Those varieties are clustered together (Cluster II). However, ancestry map for all rice varieties are not available.

The high values of cophenetic correlation coefficient strongly support the clustering pattern of dendrograms. Principal coordinate analysis (PCoA) displayed the relationship among Sri Lankan improved rice varieties on two coordinate axes (PCoA1 and PCoA2). Grouping pattern of the PCoA supported the results of the UPGMA cluster analysis. Group II of PCoA represents the UPGMA cluster II, group IIIA represents the UPGMA cluster IIIA as well as group IIIB represents the cluster IIIB.

Branching patterns of the unrooted tree generated from PHYLIP package also confirms the PCoA and UPGMA analysis. Therefore all methods of analysis produced similar results and confirm the data and analysis methods.

Conclusion

The results obtained from this study estimates the high level of genetic variation among Sri Lankan new improved rice varieties. This genetic diversity information will helpful to distinguish breeding lines for varietal development by estimates of genetic similarities useful in choosing widely divergent parents with desirable traits. Furthermore the genetic diversity data at molecular level will be useful in categorizing the accessions in core collections of Gene banks and also helpful in *ex situ* and *in situ* genetic conservation, utilization and exchange of genetic material.

Acknowledgement

The authors greatly appreciate the financial support given by the National Research Council of Sri Lanka (Grant No. 05-61). *Technical assistance* of Ms. Anoma Jayasoma is gratefully acknowledged. We thank Gene Bank of PGRC for providing seeds samples and Green house facilities.

References

Abhijit SD Kamlesh J Jose MG Vyankatesh JP Milind SPDilip RR and Yogesh SS (2004) Comparison of 16S rRNA gene sequences of genus *Methanobrevibacter*. *BMC Microbiology* **4**: 20. DOI: 10.1186/1471-2180-4-20

Chen DH Ronald PC (1999) A Rapid Miniprep Method Suitable for AFLP and PCR Applications. *Plant Molecular Biology Reporter* **17**: 53-57. DOI: 10.1023/A:1007585532036

Felsenstein J (2009) PHYLIP Version 3.69. University of Washington Seattle.

Garcia AAF, Benchimol LL, Barbosa AMM, Geraldi IO, Jr CLS and Souza AP (2004) Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology*. **27**:579-588. DOI: 10.1590/S1415-47572004000400019

Gower JC and Legendre P (1986) Metric and Euclidean properties of dissimilarity coefficients.

Ikeda R and Duncan AV (1991). The distribution of resistance genes to the brown planthopper in rice germplasm. *Rice Genetics Newsletter*. 8:125

Jaccard P (1901) Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vandoise Sci Nat* **37**,547-579. *Journal of Classification* **3**: 5-48. DOI: 10.1007/BF01896809

Kovach WL (1998) *MVSP - A Multivariate Statistical Package for Windows, ver. 3.1*. Kovach Computing Services, Pentraeth, Wales, U.K.

Loh JP, Kiew R, Kee A, Gan L and Gan Y (1999) Amplified Fragment Length polymorphism (AFLP) provides Molecular Markers for the Identification of *Caladium bicolor* Cultivars. *Annals of Botany* **84**: 155- 161. DOI: 10.1006/anbo.1999.0903

Mantel NA (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209-220.

Nei M and Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*. **76**: 5269-5273

Rajkumar G, Weerasena J. Fernando K, Liyanage A and Silva R (2011) Genetic differentiation of Sri Lankan traditional rice (*Oryza sativa*) varieties and Wild rice species by AFLP markers. *Nordic Journal of Botany*. **29** (2): 238–243. DOI: 10.1111/j.1756-1051.2011.00998.x

Roderic DM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in BioSciences* **12**: 357-358. DOI: 10.1093/bioinformatics/12.4.357

Sandka AL and Dushani SN (2009) Growth performance in rice sector: the presence scenario in Sri Lanka. *Tropical Agricultural Research and Extension* **12**(2): 71-76. DOI: 10.4038/tare.v12i2.2793

Sneath PHA and Sokal RR (1973) *Numerical Taxonomy*. Freeman, San Francisco.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acid Research* **23**: 4407-4414. DOI: 10.1093/nar/23.21.4407