Bangladesh J. Genet Pl. Breed., 20(2): 01-10, 2007

MOLECULAR CHARACTERIZATION AND GENETIC DIVERSITY IN FINE GRAIN AND AROMATIC LANDRACES OF RICE (*Oryza sativa* L.) USING MICROSATELLITE MARKERS

M. Z. Hossain[1], M. G. Rasul, M. S. Ali[2], K. M. Iftekharuddaula[2] and M. A. K. Mian

Department of Genetics and Plant Breeding
Bangabandhu Sheikh Mujibur Rahman Agricultural University
Gazipur 1706, Bangladesh

ABSTRACT

A total of thirty microsatellite molecular markers were used across 21 rice genotypes for their characterization and discrimination. The number of alleles per locus ranged from three (RM165, RM219, RM248, RM463, RM470 and RM517) to nine (RM223), with an average of 4.53 alleles across the 30 loci obtained in the study. The polymorphism information content (PIC) values ranged from 0.30 (RM219) to 0.84 (RM223) in all 30 loci. RM223 was found the best marker for the identification of 21 genotypes as revealed by PIC values. The frequency of the most common allele at each locus ranged from 24% (RM223 and RM334) to 81% (RM219). A two dimensional principal coordinate analysis (PCoA) with 21 genotypes showed that the genotypes Supper Basmoti, Basmati370, BasmatiD, Keora, Chinisakkora, Thakurbhog, Benaful, Kolgochi, Buchi, Awnedtapl and Kalijira-11 were found far away from centroid of the cluster than rest of the genotypes which placed around the centroid. The pair-wise genetic dissimilarity coefficients indicated that the highest genetic distance was obtained between Thakurbhog and Supper Basmoti (0.81) as well as between Benaful and Keora (0.81). Basmati (Basmoti D, Super Basmati, Basmati 370) and Kalijira (Kalijira 11, 12, 13, 14) genotypes had close similarity among them but showed wide dissimilarity with other local genotypes. Being grouped into distant clusters, SupperBasmoti, Thakurbhog, Keora, and Benaful could be utilized as potential parents for the improvement of fine grain aromatic rice varieties. Genotypes Kolgochi and Buchi (having zero dissimilairty) might be possessed same genetic background. The microsatellite marker based molecular fingerprinting could serve as a sound basis in the identification of genetically distant accessions as well as in the duplicate sorting of the morphologically close accessions.

Key words: Rice (*Oryza sativa* L), genetic diversity, microsatellite marker, SSR

INTRODUCTION

Consumers all over the world prefer aromatic rice due to its flavor and palatability. In Bangladesh, there are more than four thousand local landraces of rice,

[1]Seed Certification Agency, Ministry of Agriculture, Joydebpur, Gazipur 1701, [2] Biotech Division, Bangladesh Rice Research Institute, Joydebpur, Gazipur 1701, Bangladesh

adapted in different parts of the country, some of which have very nice quality, fineness, aroma, taste and protein contents (Kaul *et al.,* 1982). These are generally low yielding, cannot compete with modern varieties of rice. Research on these landraces was not emphasized in the past because of ever increasing demands of higher production to feed teeming millions. By now, HYV rice production technology has been well developed and adopted by farmers for maximal productivity of rice. It is proper time to emphasize on the improvements of productivity of local landraces of fine grain and aromatic types.

Genetic variability is the mainstay for the success of any breeding program hence assessing the level of genetic variation among the rice varieties or genotypes has been of great interest to the breeders. Rice taxonomists are interested in the rapid classification of different taxonomic groups, while breeders are concerned about the determination of usable agronomic variation in breeding programs (Zhou, 1995). Although large number of them is available, little analysis of the genetic diversity has been done at molecular level so far. Genetic diversity studies based on molecular markers generate information-base for more efficient use of the valuable genetic resources. Pre-breeding or genetic enhancement has become a necessary and planned part of all rice breeding activities

Molecular marker technologies can assist conventional breeding efforts and are valuable tools for the analysis of genetic relatedness and the identification and selection of desirable genotypes for crosses as well as for germplasm conservation in gene banks. Among many molecular techniques, simple sequence repeat (SSR) markers (microsatellites) are co-dominant, hypervariable, abundant and well distributed throughout the rice genome (Temnykh *et al.,* 2001). Abundance of microsatellite markers is now available through the published high-density linkage map (McCouch *et al.,* 2002, IRGSP, 2005) or public database.

Our study is involved with the aromatic and fine grain traditional landraces of rice locally cultivated by farmers those were either not used, or poorly used as parents in the Bangladeshi rice-breeding program. These germplasm can function as the source of alternative genetic pools to the improved varieties. Thus, 30 microsatellite markers were used against 21 fine grain aromatic genotypes to assess the extent of molecular diversity in the traditional aromatic rice genotype, to establish their genetic relationships and to evaluate their potential utility in the breeding programme.

MATERIALS AND METHODS

Plant materials

The 18 genotypes of aromatic and fine grain landraces of rice and three basmati types. (Table 1) were selected from the rice germplasm collection at the Bangladesh Rice Research Institute (BRRI) gene bank. The whole experiment was conducted at Biotechnology Division of BRRI during 2006-7. Five grams of germinated seed from each genotype was sown in the pot

Table 1. List of the 18 genotypes of aromatic and fine grain landraces of rice and 3 basmati types

Sl No.	Genotypes	BRRI Accession No.	Sl No.	Genotypes	BRRI Accession No.
G40	Benaful	981	G50	Keora	731
G41	Dakshahi	983	G69	Kalijira11	2501
G42	Thakurbhog	872	G70	Radhunipagol	2504
G43	Jirabhogfiner	4831	G71	Awnedtapl546	2940
G44	Agoli	2009	G72	Ovatapl600	2990
G45	Badshabhog	4355	G73	Kalijira12	4755
G46	Kolgochi	352	G74	Kalijira13	4817
G47	Buchi	369	G75	Kalijira14	4832
G48	Moloti	169	G76	BasmotiD	--
G49	Chinisakkora2	387	G77	Basmati370	--
			G78	SupperBasmoti	--

Genotyping protocol

DNA was extracted from young leaves of three-week-old plants following a simple and modified protocol to isolate total genomic DNA for PCR analysis which did not require liquid nitrogen and required only a very small amount of tissue samples as described by Zheng *et al.* (1995)

.

PCR was performed in 12.5 µl reaction containing 5-25 ng of DNA template, 1.25 µl of $MgCl_2$-free 10X PCR buffer (100 mM Tris-HCl pH 9.0 at 25°C, 500 mM KCl, 0.1% Triton® X-100 and H_2O), 1.5 µl of 25 mM $MgCl_2$, 0.25 µl of 10mM dNTP, 0.25 µl of 5 U/µl Taq polymerase enzyme, 0.625 µl each of 10 µM forward and reverse primers using in an MJ Research single 96-well thermal cycler. The mixer was overlaid with one drop of mineral oil to prevent evaporation.

After initial denaturation for 5 min at 94°C, each cycle comprised of 1 min denaturation at 94°C, 1 min annealing at 55°C, and 2 min extension at 72°C with a final extension for 7 min at 72°C at the end of 35 cycles. The PCR products were mixed with bromophenol blue gel loading dye and were analyzed by electrophoresis on 8% polyacrylamide gel using mini vertical polyacrylamide gels for high throughput manual genotyping (CBS Scientific Co. Inc., CA, USA). Two and half µl of amplification products were resolved by running gel in 1xTBE buffer for 2-2.5 hrs depending upon the allele size at around 75 volts and 180 mA current. The gels were stained in 0.5mg/ml ethidium bromide and were documented using UVPRO (Uvipro Platinum, EU) gel documentation unit. Microsatellite or Simple Sequence Sepeat (SSR) markers were used for selection (Temnykh *et al*. 2001; McCouch *et al*., 2002; IRGSP 2005). Thirty SSR markers (distributed in the 12 chromosomes) with clear amplifications were selected for genetic diversity analysis of 21 genotypes (Table 2).

Data analysis

Molecular weight for each amplified allele was measured in base pair using Alpha-EaseFC 5.0 software. The summary statistics including the number of alleles per locus, major allele frequency, gene diversity, polymorphism information content (PIC) values were determined using PowerMarker version 3.25 (Liu and Muse 2005). For the unrooted phylogenetic tree, genetic distance was calculated using the "C.S. Chord 1967" distance (Cavalli-Sfoza and Edwards 1967) followed by phylogeny reconstruction using neighbor-joining as implemented in PowerMarker with tree viewed using Treeview (Page, 1996). The allele frequency data from Powermarker was used to export the data in binary format (allele presence = "1" and allele absence = "0") for analysis with NTSYS-pc version 2.1 (Rohlf, 2002). A similarity matrix was calculated with the Simqual subprogram using the Dice coefficient, followed by cluster analysis with the SAHN subprogram using the UPGMA clustering method as implemented in NTSYS-pc. The similarity matrix was also used for principal coordinate analysis (PCoA) with the DCenter, Eigen, Output, and MXPlot subprograms in NTSYS-pc.

RESULT AND DISCUSSION

Overall SSR diversity

A total of 136 alleles were detected at the loci of 30 microsatellite markers across 21 rice genotypes. Similar number of microsatellite makers was previously used as subset for genetic diversity analysis of *O. sativa* (Garris *et al.*, 2005, Chakrabarthia and Naravaneni, 2006, Thomson *et al*, 2007). The number of alleles per locus ranged from 3 alleles (RM165, RM219, RM248, RM463, RM470 and RM517) to 9 alleles (RM223), with an average of 4.53 alleles across the 30 loci (Table 2). This vale is comparable to 1-8 allele per SSR locus with an average number of alleles of 4.58 per locus for various classes of microsattelite (Siwach *et al*, 2004) and 2-7 alleles per locus as reported by Chakrabarthia and Naravaneni (2006).

The frequency of the most common allele at each locus ranged from 24% (RM223 and RM334) to 81% (RM219), which is comparable with Thomson *et al* (2007). On an average, 49% of the 21 rice genotypes shared a common major allele at any given locus. The range of polymorphic information content (PIC) values was from 0.30 to 0.84 with an average 0.58. The PIC values observed, are comparable to three previous estimates of microsattelite analysis in rice viz 0.67-0.88 (Gohain *et al.*, 2006), 0.34-0.88 (Thomson *et a.l*, 2007), 0.20-.90 with an average.56 (Jain *et al*, 2003). The highest PIC value (0.84) was obtained for RM223 followed by for RM334 (0.80), RM224 (0.78), RM17 (0.74) and RM258 (0.71), respectively (Table 2). PIC value revealed that RM223 was considered as the best marker for 21 genotypes and this was reported to be linked to the gene for fragrance (*fgr*) (Gohain *et al.* 2006). Fig. 1 showed a gel picture of the microsatellite marker RM223 for all 21 genotypes.

Genetic distance-based analysis

The genetic distance-based results seen in the unrooted neighbor-joining tree revealed six groups in the 21 genotypes (Fig.2). The three *Basmati* type genotypes were clustered in the same group. Again, the three Kalijira genotypes clustered separately, but closer to the *Basmati* type. Two aromatic landraces (Kolgochi and Buchi) were found exactly same and clustered with Badshabhog. Furthermore, the two non-aromatic fine grain landraces viz. Moloti and Keora were clustered in the same group. Five aromatic fine grain landraces

(Benaful, Dakshahi, Thakurbhog, Jirabhogfiner and Agoli) were clustered distinctly in the same group. The four landraces (Awnedtapl546, Kalijira11, Radhunipagol, Ovaltapl600) with strong aroma were clustered in other group near the *Basmati* type *group*.

Table 2. Data on the number of alleles, allele size range, highest frequency allele and polymorphism information content (PIC) found among 21 rice genotypes for 30 microsatellite markers

Marker	Chr No	Position (cM)	Motif*	Allele no.	Size range (bp)	Highest frequency allele		PIC Value
						Size (bp)	Freq (%)	
RM17	12	106.1	(GA)21	5	164-190	170	28	0.7443
RM19	12	20.9	(ATC)10	4	222-244	232	38	0.6529
RM21	11	85.7	(GA)18	6	122-162	142	29	0.7452
RM107	9	82.4	(GA)7	6	167-205	186	.28	0.7452
RM122	5	0	(GA)7A(GA)2A(GA)11	4	222-236	230	38	0.6489
RM125	7	24.8	(GCT)8	4	117-136	122	62	0.5192
RM140	1	78.4	(CT)12	4	264-301	272	38	0.6221
RM154	2	4.8	(GA)21	5	178-202	186	52	0.6175
RM165	1	186.6-196.5	(CT)13	3	178-187	187	76	0.3360
RM206	11	102.9	(CT)21	4	121-151	140	57	0.5179
RM211	3	14.4	(TC)3A(TC)18	4	138-154	142	71	0.4070
RM215	9	99.4	(CT)18	4	145-155	145	48	0.5736
RM219	9	11.7	(CT)17	3	188-215	206	81	0.3026
RM223	8	80.5	(CT)25	9	142-160	155	24	0.8414
RM224	11	120.1	(AAG)8(AG)13	7	122-163	145	28	0.7809
RM229	11	77.8	(TC)11(CT)5C3(CT)5	4	96-105	101	47	0.5425
RM248	7	116.6	(CT)25	3	85-101	101	76	0.3360
RM258	10	70.8	(GA)21(GGA)3	5	114-140	128	38	0.7159
RM271	10	59.4	(GA)15	4	92-105	100	43	0.6030
RM307	4	0	(AT)14(GT)21	4	120-135	130	52	0.5334
RM312	1	71.6	(ATTT)4(GT)9	4	94-101	99	66	0.4732
RM334	5	141.8	(CTT)20	7	159-187	181	24	0.8016
RM447	8	124.6	(CTT)8	4	105-130	130	38	0.6529
RM463	12	75.5	(TTAT)5	3	178-207	195	57	0.5157
RM470	4	115.5	(CTT)14	3	78-100	90	66	0.4470
RM517	3	42.9	(CT)15	3	245-277	260	66	0.4470
RM522	1	33.9	(AAT)6	4	137-160	155	42	0.6030
RM541	6	75.5	(TC)16	5	156-193	177	38	0.6630
RM566	9	47.7	(AG)15	5	245-308	256	61	0.5355
RM580	1	68.2	(CTT)19	6	195-223	210	61	0.5519
Mean				4.5333			49	0.5825

- Motif of the SSR and number of repeats as previously published (http://www.gramene.org)

The genetic similarity analysis using UPGMA clustering agreed with the neighbor-joining data. The UPGMA clustering system also generated six genetic clusters with similarity coefficient of 40%. Here, the Basmati type group also clustered near Kalijira and strong aroma group (Fig. 3). Five aromatic fine grain landraces (Benaful, Dakshahi, Thakurbhog, Jirabhogfiner and Agoli) were clustered distinctly in the same group. UPGMA cluster revealed distinct concentration of Basmati type in the same cluster while most of the aromatic landraces are distributed in other clusters. It was also obtained that two aromatic landraces (Kolgochi and Buchi) were found exactly same (100% similar) and they were clustered with Badshabog. The results indicated that the genotypes viz. Kolgochi and Buchi might be of same genetic background which could be verified by using more markers. These cluster tree analysis agreed with the allelic diversity observed among Basmati and Non-Basmati long grain *indica r*ice varieties using microsattelite

markers (Siwach *et al*., 2004). DNA fingerprinting and phylogenetic analysis of Indian aromatic high quality rice germplasm also showed the same trend (Jain *et al*., 2003).

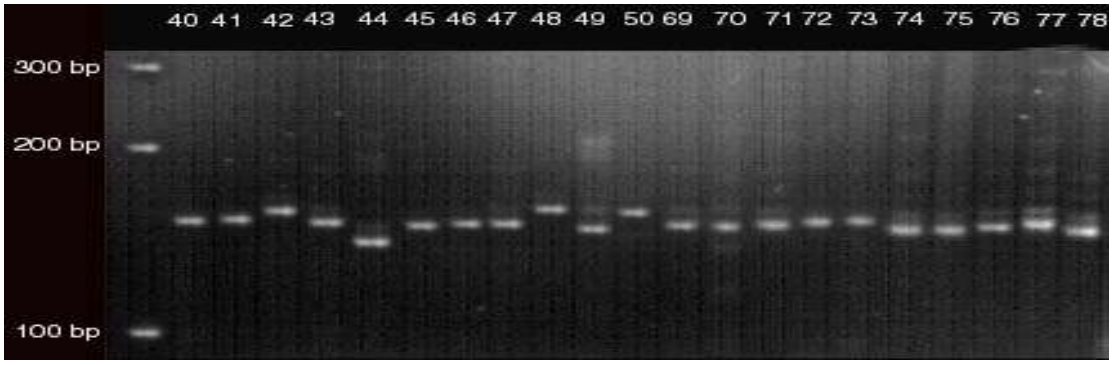

Fig.1. DNA profile of the 21 rice landraces with the microsatellite marker RM223

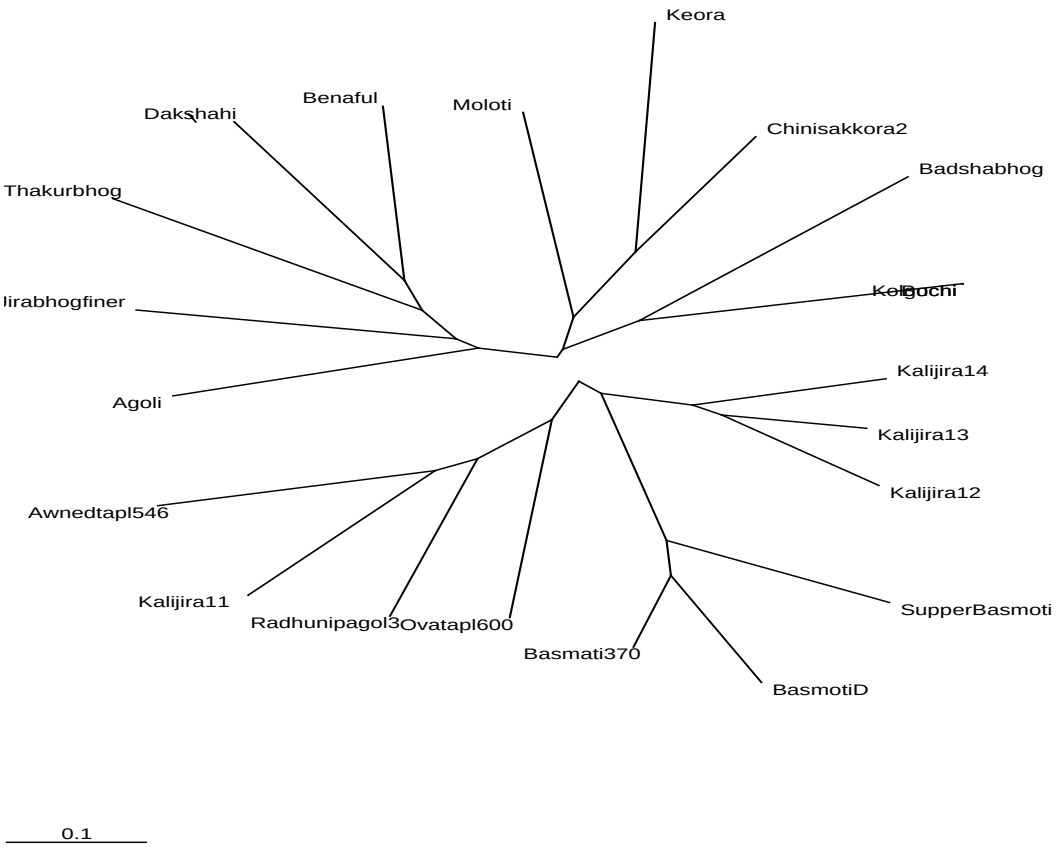

Fig.2. An unrooted neighbor-joining tree showing the genetic relationships among 21 aromatic and fine grain rice landraces of Bangladesh based on the alleles detected by 30 microsatellite markers

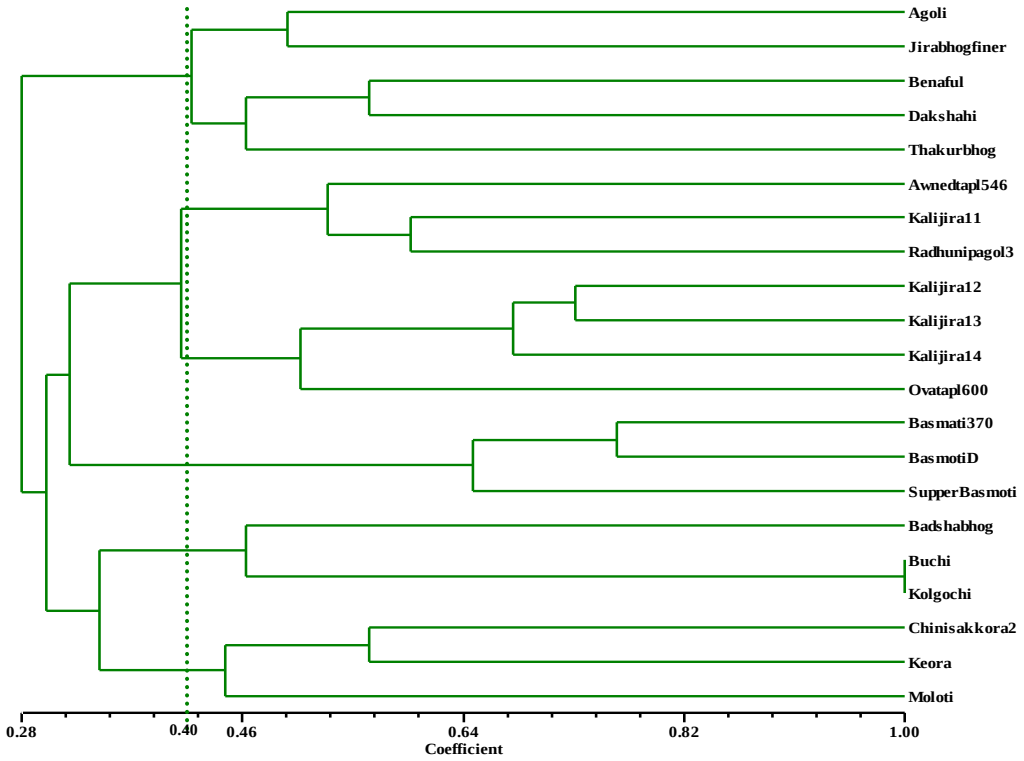

Fig.3. A UPGMA cluster dendrogram showing the genetic relationships between 21 aromatic and fine grain rice landraces of Bangladesh based on the alleles detected by 30 microsatellite markers

The two dimensional graphical view of Principal Coordinate Analysis (PCoA) showed the spatial distribution of the genotypes along the two principal axes. The genotypes viz. SupperBasmoti, Basmati370, BasmatiD, Keora, Chinisakkora, Thakurbhog, Benaful, Kolgochi, Buchi, Awnedtapl and Kalijira-11 were found placing far away from the centroid of the cluster and rest of the genotypes were placed more or less around the centroid (Fig. 4). The results indicated that the genotypes placed far away from the centroid were more genetically diverse while the genotypes placed near around the centroid possessed more or less similar genetic background. However, centroid may be defined as the vector representing the middle point of the cluster which contained at least one number for each variable. The connecting line between the each genotype and the centroid represented eigen vectors for the respective genotypes.

Table 3. Genetic dissimilarity index among 21 rice genotypes using 30 microsatellite markers

Genotypes	Agoli	Awned	Badsha bhog	Basmati 370	Basmoti D	Benaful	Buchi	Chini sakkora	Dak shahi	Jirabhog finer	Kalijira 11	Kalijira 12	Kalijira 13	Kalijira 14	Keora	Kolgochi	Moloti	Ovatapl 600	Radhuni pagol	Supper Basmoti	Thakur bhog
Agoli	0.00																				
Awned tapl546	0.63	0.000																			
Badsha bhog	0.630	0.600	0.000																		
Basmati 370	0.630	0.720	0.630	0.0000																	
BasmotiD	0.660	0.720	0.750	0.2101	0.000																
Benaful	0.540	0.660	0.690	0.7503	0.750	0.000															
Buchi	0.510	0.750	0.480	0.5702	0.660	0.600	0.000														
Chini sakkora2	0.540	0.720	0.600	0.5702	0.660	0.630	0.510	0.000													
Dakshahi	0.570	0.660	0.660	0.6602	0.690	0.390	0.660	0.720	0.000												
Jirabhog finer	0.450	0.630	0.660	0.6902	0.720	0.420	0.720	0.630	0.480	0.000											
Kalijira11	0.660	0.390	0.660	0.7203	0.660	0.630	0.720	0.630	0.600	0.630	0.000										
Kalijira12	0.660	0.660	0.540	0.4802	0.600	0.660	0.600	0.600	0.570	0.510	0.510	0.000									
Kalijira13	0.600	0.510	0.570	0.4802	0.540	0.660	0.540	0.540	0.570	0.570	0.450	0.240	0.000								
Kalijira14	0.570	0.570	0.510	0.4201	0.570	0.630	0.540	0.450	0.660	0.540	0.630	0.330	0.240	0.000							
Keora	0.660	0.720	0.600	0.6302	0.720	0.810	0.690	0.390	0.750	0.750	0.750	0.630	0.660	0.660	0.000						
Kolgochi	0.510	0.750	0.480	0.5702	0.660	0.600	0.000	0.510	0.660	0.720	0.720	0.600	0.540	0.540	0.690	0.000					
Moloti	0.480	0.570	0.600	0.6302	0.630	0.690	0.540	0.450	0.660	0.540	0.600	0.570	0.540	0.540	0.540	0.540	0.000				
Ovatapl600	0.600	0.510	0.630	0.5702	0.570	0.510	0.660	0.570	0.630	0.540	0.510	0.390	0.420	0.510	0.690	0.660	0.540	0.000			
Radhuni pagol3	0.600	0.450	0.570	0.6302	0.690	0.480	0.630	0.570	0.540	0.630	0.360	0.510	0.480	0.540	0.660	0.630	0.630	0.450	0.000		
Supper Basmoti	0.660	0.750	0.690	0.2701	0.360	0.780	0.660	0.690	0.690	0.690	0.750	0.570	0.480	0.570	0.720	0.660	0.630	0.570	0.720	0.000	
Thakurbhog	0.540	0.630	0.690	0.7203	0.720	0.420	0.630	0.690	0.540	0.570	0.690	0.750	0.660	0.600	0.720	0.630	0.630	0.600	0.630	0.810	0.000

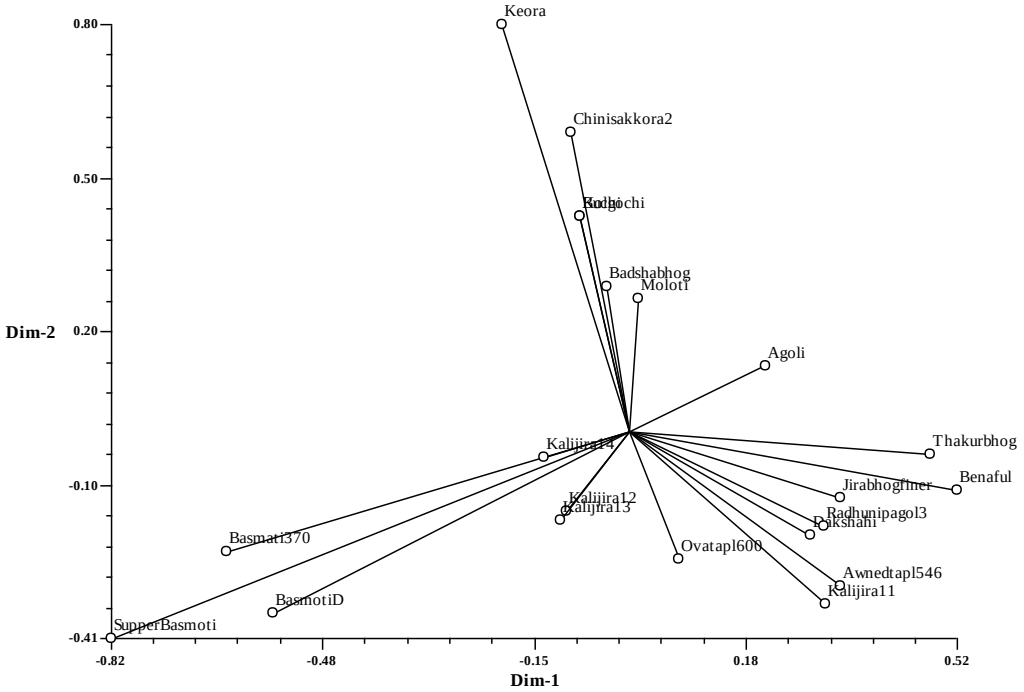

Fig 4. Two-dimensional view of Principal Coordinate Analysis (PCoA) with 30 microsatellite markers over 21 genotypes.

The pair-wise genetic dissimilarity coefficients indicated that the highest genetic distance was obtained between Thakurbhog and SupperBasmoti (0.81) as well as between Benaful and Keora (0.81) (Table 3) followed by between SupperBasmoti and Benaful (78%), Dakshahi, Jirabhogfiner and Kalijira11 with Keora (75%), Kolgochi and Buchi with Awnedtapl (75%), Basmati370 and BasmatiD with Benaful (75%). The results were in agreement with the findings of principal coordinate analysis and suggested that these genotypes were more genetically diverse. In crop improvement program these genetically diverse genotypes could be chosen as parents in the crossing program to create genetic variability. On the other hand BasmotiD, Basmoti 370 and SuperBasmoti were the most similar genotype (21%-27% dissimilar) followed by Kalijira 11, 12, 13, 14 (24% dissimilar). Though three Basmoti genotypes were more similar with each other but most dissimilar with other landraces. Likewise, Kalijira 11, 12, 13, 14 were more similar with each other but dissimilar with other landraces rice genotypes. It was also obtained that two aromatic landraces (Kolgochi and Buchi) were found exactly same in this analysis (0% dissimilar). Hence, microsatellite marker based molecular fingerprinting could serve as a potential basis in the identification of genetically distant accessions as well as in identification of the morphologically close accessions.

REFERENCES

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. American J Hum Genet. 19: 233-257.

Chakrabarthi, B K and R. Naravaneni. 2006. SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L). African J. Biotech. 5(9): 684-88.

Garris A.J., T. H. Tai, J. R. Coburn, S. Kresovich and S. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics. 169: 1631–1638.

Gohain, B., A. Talukdar and M. K. Modi. 2006. Mining of allele(s) for aroma in local aromatic rice (joha) germplasm of Assam using SSR marker. ABSTRACTS of the 26TH Int. Rice Conference and 2nd Int. Rice Congress, October 9-13, 2006, New Delhi, India. p74.

IRGSP, 2005. The map-based sequence of the rice genome. Nature. 436: 793-800.

Jain, S., S. E. Mitchell, R. K. Jain, S. Kresovich and S. R. McCouch, 2003. DNA fingerprinting and phylogenetic analysis of Indian aromatic high-quality rice germplasm using panels of fluorescent-labeled microsattelite markers. In: Advance in Rice Genetics ed. by Khush *et al*. IRRI, Philippine. 162-166.

Kual, A. K., M. R. I. Khan and K. M. Munir. 1982. Rice quality, a survey of Bangladesh germplasm. BRRI, Joydebpur, Bangladesh. P-178.

Liu, K. and S. V. Muse. 2005. PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics. 21 2128–2129.

McCouch, S. R., L. Teytelman, Y. Xu, K. B. Lobos, K. Clare, M. Walton, B. Fu, R. Maghirang, Z. Li and Y. Xing. 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res. 9: 199-207.

Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Mol. Biol. 12: 357–358.

Rohlf, F. 2002. NTSYS-pc: Numerical taxonomy and multivariate analysis system, 2.1 edn. Department of Ecology and Evolution, State University of NY, Stony Brook.

Siwach, P., S. Jain, N. Saini, V. K. Chowdhury and R. K. Jain. 2004. Allelic Diversity among Basmati and Non-Basmati long grain *indica* Rice Varieties using microsattelite markers. J. Pl. Biochem. Biotech. 13:25-32.

Temnykh, S., G. DeClerck, A. Lukashova, L. Lipovich, S. Cartinhour and S. R. McCouch. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposons associations and genetic marker potential. Genome Res. 11: 1441-1452.

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

Zheng, K., N. Huang, J. Bennet and G. S. Khus. 1995. PCR-based marker assisted selection in rice breeding. IRRI, Manila. pp 16-18.

Zhou, Z. 1995. Characterization of DNA minisatellite sequences from *Oryza sativa* L. and their application in DNA fingerprinting. PhD thesis. A dissertation presented to University of Columbia. p-9.