

*Full Length Research Paper*

# Determination of genetic variability of Asian rice (*Oryza sativa* L.) varieties using microsatellite markers

Zahida H. Pervaiz<sup>1\*</sup>, Malik A. Rabbani<sup>2</sup>, Stephen R. Pearce<sup>3</sup> and Salman A. Malik<sup>1</sup>

<sup>1</sup>Department of Biochemistry, Quaid-i-Azam University, Islamabad, Pakistan.

<sup>2</sup>Institute of Agri-Biotechnology and Genetic Resources, National Agricultural Research Center, Park Road, Islamabad, Pakistan.

<sup>3</sup>School of life sciences, University of Sussex, Brighton, UK.

Accepted 31 August, 2009

The microsatellite or simple sequence repeat (SSR) marker analysis was done to determine the allelic diversity and relationship among thirty-five Asian cultivars of rice including 19 aromatic, 13 non-aromatic and 3 japonica type cultivars. A total of 144 alleles were detected at the 32 SSR loci, of which 141 (98%) were polymorphic. The number of alleles generated by each marker ranged from 2 to 13 with an average of 4.5 alleles per marker. The size of smallest and largest allele ranged from 8 (RM122) to as high as 71 (RM302). Polymorphism information content (PIC) values ranged between 0.157 (RM19, RM55) and 0.897 (RM70), with an average of 0.603 per marker. Basmati rice varieties amplified different alleles at 15 of the SSR loci than those in the *japonica* and/ or *indica* rice varieties. A number of SSRs were identified that could be utilized to differentiate between basmati and other non-basmati rice varieties. The RM252 and RM310 showed a clear differentiation of *japonica* cultivars from other ones. Pair-wise Nei and Li's similarity coefficients ranged from 0.19 to 0.90. The dendrogram based on the cluster analysis by microsatellite polymorphism, grouped 35 rice cultivars into two major groups effectively differentiating the tall, late maturing and slender aromatic cultivars from the short statured, early, short bold and long bold non-aromatic cultivars. These results could be useful for monitoring purity, genotype identification and for plant variety protection.

**Key words:** Genetic diversity, *Oryza sativa*, Asian rice, microsatellites.

## INTRODUCTION

Rice is the main source of nutrition for 50% of the world population and is a major crop in many Asian countries (Sasaki and Burr, 2000). Among Asian rice growing countries, Pakistan is a major producer of many rice varieties such as aromatic rices and old landraces. In the context of global biodiversity loss, Pakistan missed several rice varieties. Therefore, it is highly necessary not only to conserve the landrace genotypes but also to investigate the gene-pool of aromatic rice for breeding purposes of high yielding varieties in the country (Rabbani et al., 2008). For this purpose, identification of different genotypes

at molecular level is imperative. The DNA-based markers are promising and effective tools for measuring genetic diversity in plants germplasm and elucidating their evolutionary relationships. Amongst the DNA markers, the microsatellites [also known as simple sequence repeats (SSRs)] are useful as genetic markers because they detect high levels of allelic diversity. These are co-dominant, easily and economically assayed by PCR and automated. Thousands of microsatellites have been identified in rice, and over 2,500 have been developed as molecular markers (Temnykh et al., 2000; McCouch et al., 2002). The advent of genomic sequences in rice offers new opportunities to enhance the density of locus-specific and polymorphic markers for high-resolution genetic analysis. Owing to technical efficiency and multiplex potential, these markers are preferable for many forms of high throughput mapping, genetic analysis and marker assisted plant improvement strategies (Coburn et

\*Corresponding author. Email: [zahidahasan82pk@hotmail.com](mailto:zahidahasan82pk@hotmail.com).

**Table 1.** Rice cultivars used in present study.

No.	Variety Name	Parentage	Origin	Type	Seed Source
1	Basmati-370(P)	Local selection	Pakistan	Aromatic	RRI, KSK, Lahore
2	Mushkan	Local selection	Pakistan	Aromatic	RRI, KSK, Lahore
3	Basmati-Pak	CM7-6/Basmati-370	Pakistan	Aromatic	RRI, KSK, Lahore
4	IR8	Peta/Dee-gee-woo-gen	Philippines	Coarse	RRI, KSK, Lahore
5	Basmati-198	Basmati-370/TN1	Pakistan	Aromatic	RRI, KSK, Lahore
6	Super-Basmati	Basmati-370/10486	Pakistan	Aromatic	RRI, KSK, Lahore
7	JP5	Local selection	Pakistan	Coarse	ARI, Mingora, Swat
8	Mahlar-346	Local selection	Pakistan	Coarse	RRI, KSK, Lahore
9	Purple Marker	N/A	Pakistan	Coarse	RRI, Dokri, Sindh
10	Dokri-Basmati	Local selection	Pakistan	Aromatic	RRI, Dokri, Sindh
11	IR36	IR2042/CR9413	Philippines	Coarse	NIAS, Tsukuba
12	Kinmaze	Ryosaku/AichnakateAsahi	Japan	Japonica	NIAS, Tsukuba
13	KangniXTorh	Kangni x Torh	Pakistan	Coarse	RRI, Dokri, Sindh
14	Sonahri-Kangni	Kangni-27 x Colombia	Pakistan	Coarse	RRI, Dokri, Sindh
15	Kharai-Ganga	N/A	Pakistan	Coarse	RRI, Dokri, Sindh
16	KSK133	KS282//4321(PK198-12-3-2-1-1)	Pakistan	Coarse	RRI, KSK, Lahore
17	Shahkar	IR13423-10-2-3// IR13419-113-3	Pakistan	Coarse	RRI, Dokri, Sindh
18	Shandar	Mutant of IR6	Pakistan	Coarse	NIA, Tandojam
19	Mehak	Mutant of Basmati-370	Pakistan	Aromatic	NIA, Tandojam
20	Azucena	Traditional japonica cv	Philippines	Japonica	NIAS, Tsukuba
21	Kasalath	Traditional cultivar	India	Coarse	NIAS, Tsukuba
22	Basmati-370(I)	Local selection	India	Aromatic	IRRI, Philippines
23	Dheradun-Basmati	Traditional basmati cv	Nepal	Aromatic	IRRI, Philippines
24	Khao-Dawk-Mali-105	Traditional scented cv	Thailand	Aromatic	IRRI, Philippines
25	Basmati-217	Traditional basmati cv	India	Aromatic	IRRI, Philippines
26	Dehradun-Basmati	Traditional basmati cv	India	Aromatic	IRRI, Philippines
27	Punjab-Basmati-1	Sona/Basmati-370	India	Aromatic	IRRI, Philippines
28	Pusa-Basmati-1	Pusa 150/Karnal local	India	Aromatic	IRRI, Philippines
29	Chini-Sakkor	N/A	Bangladesh	Aromatic	IRRI, Philippines
30	Ranbir-Basmati	Selection from Basmati-370-90-95	India	Aromatic	IRRI, Philippines
31	Jasmine-Scented	N/A	Thailand	Aromatic	IRRI, Philippines
32	Niaw-Hawn-Mali	N/A	Thailand	Aromatic	IRRI, Philippines
33	Khao-Jao-Hawm	N/A	Thailand	Aromatic	IRRI, Philippines
34	PK386	N/A	Pakistan	Coarse	Used as adulterant
35	TN1	Dee-gee-woo-gen/Tasi Younli	Taiwan	Japonica	IRRI, Philippines

al., 2002).

Present study was conducted to determine the level of genetic variation and relatedness among some selected Asian rice varieties by using microsatellite markers. In the investigation reported here, we used thirty-two SSRs to fingerprint thirty-five diverse varieties. The varieties included basmati and other aromatic/quality rice collections from different parts of the Asia as well as non-basmati (*indica* and *japonica*) cultivars. The SSR data was used to evaluate the level of genetic diversity within basmati rice, to assess genetic relationships among the varieties, and to determine whether existing SSR markers provide adequate power of resolution to discriminate between high- and low-quality aromatic varieties for use in

commercial evaluation. This study could be useful in the identification of rice varieties and their preservation as well.

## MATERIALS AND METHODS

### Plant material

Thirty-five Asian rice varieties were selected for this study. Amongst them, 18 cultivars were from Pakistan, 2 from International Rice Research Institute, Philippines, 7 from India, 4 from Thailand, 1 from Bangladesh, 1 from Nepal, 1 from Japan, and 1 from Taiwan. The detailed description of the materials used in this study is described in Table 1.

### Genomic DNA isolation

The total DNA was extracted from dry seed samples of each cultivar by following Kang et al. (1998) with minor modifications. The purity and concentration of DNA was determined spectrophotometrically at 260 and 280 nm by using the NanoDrop ND-1000 Spectrophotometer. The DNA samples were diluted to a concentration of 20 ng/ $\mu$ l with TE (Tris-EDTA) for SSR analysis.

### Selection of primers

Thirty five microsatellite primer pairs covering all twelve chromosomes were selected for the genetic diversity analysis on the basis of published rice microsatellite framework map. Three primers exhibited monomorphic fragments and were excluded. The original source, repeat motifs, primer sequences and chromosomal positions for these markers can be found in the rice genome database (<http://www.gramene.org>). Microsatellite primer pairs were obtained from Hokkaido System Science, Japan.

### Microsatellite marker analysis

The microsatellite analysis was performed following Ravi et al. (2003) with minor modifications. PCR amplification reactions were done in a total volume of 20  $\mu$ l containing 10 mM Tris HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M each of deoxynucleotide triphosphate (dNTP), 0.2  $\mu$ M of each forward and reverse primer, 1 unit Taq DNA polymerase (Fermentas Life Sciences) and 20 ng of template DNA. The PCR amplifications were carried out using a Peltier Thermal Cycler (DNA Engine PTC-200). The thermal cycler was set at 1 cycle per 5 min at 94°C as an initial hot start and strand separation step. This was followed by 35 cycles of 1 min at 94°C for denaturation, 1 min for annealing temperature depending on the marker used (55 - 65°C) and 2 min at 72°C for primer elongation. Finally, 1 cycle of 7 min at 72°C was used for final extension. Amplified products were stored at -20°C until further use. The reproducibility of amplification products was confirmed twice for each of the primer used in this study.

### Electrophoresis of amplified products

After amplification, a 15  $\mu$ l aliquot of the amplified microsatellite samples was combined with 3  $\mu$ l of a loading buffer (0.4% bromophenol blue, 0.4% xylene cyanole and 5 ml of glycerol) and was analyzed directly on 3% Gene Choice High Resolution Agarose (CLP, USA) gels in 1XTBE buffer (10 mM Tris-Borate, 1 mM EDTA) containing 0.5  $\mu$ g per ml of ethidium bromide. A 25 bp DNA ladder (Biolabs, New England, UK) was used as a size marker to compare the molecular weights of amplified products. After electrophoresis, the gels were documented using an UVIdoc Gel Documentation System (UVITEC, Cambridge, UK).

### Allele scoring and data analysis

Ethidium bromide staining of agarose gels generally revealed a multiple number of bands. The size of the most intensively amplified band for each microsatellite marker was determined based on its electrophoretic mobility relative to molecular weight marker (25 bp). Amplified products from microsatellite analyses were scored qualitatively for presence and absence of each marker allele-genotype combination. The SSR bands amplified by using given primers were further treated as a unit character. The data were entered into a binary matrix as discrete variables (e.g., 1 for presence and 0 for absence of character). The relevant primers were

employed on the basis of polymorphism extent. The Polymorphic Information Content (PIC) value of the used marker was calculated by following a method developed by Anderson et al. (1993). The pair-wise comparisons of the cultivars based on the proportion of unique and shared amplification products (alleles) were used to measure the genetic similarity by Dice coefficients. The Dice coefficients were employed by using Simqual sub-program in similarity routine of software NTSYS-pc version 2.2 (Exeter Software, Setauket, NY, U.S.A.) software package (Rohlf, 2005). The estimation of genetic similarity (F) were calculated by following the method described elsewhere (Nei and Li, 1979). The resultant similarity matrix data was employed to construct a dendrogram by using Sequential Agglomerative Hierarchical Nesting (SAHN) based on unweighted pair-group method with an arithmetic average (UPGMA) to infer genetic relationships and phylogeny among cultivars. The principal component analysis (PCA) was also done by using the subroutine EIGEN. All computations were done by using the NTSYS-pc, Version 2.2 package (Rohlf, 2005). Further details are available elsewhere (Rabbani et al., 2008).

## RESULTS

### DNA amplification and cultivar identification

A total of thirty-two microsatellite or SSR markers (Table 2) covering all 12 chromosomes were utilized to characterize and assess genetic diversity among thirty-five rice varieties from Asian region. Amplification profile as revealed by RM1 across a number of cultivars used is depicted in Figure 1. In this study, a considerable variability was found among different cultivars. In most of the cases, Basmati and other aromatic cultivars exhibited similar banding patterns. The used markers showed several bands, which were shared among the basmati and other fine cultivars. Contrarily, the aromatic and non-aromatic cultivars of rice shared a few bands. The cultivar 'KangniXTorh' displayed unique bands in comparison with all other non-aromatic genotypes. Many primers showed characteristic fragments in this cultivar, which were not produced in any of the other non-aromatic cultivars used. Three *japonica* type cultivars 'Azucena', 'Kinmaze' and 'TN1' showed different as well as shared fragments with the other rice cultivars from Asian regions. The aromatic varieties could be differentiated from the non-aromatic varieties by a number of SSR loci, while many of the basmati specific alleles were shared by other aromatic and *japonica* varieties. The DNA fingerprint database developed in this study showed enough polymorphism to differentiate basmati and coarse rice varieties.

### Number of alleles and allele size

The level of polymorphism among the rice cultivars was evaluated by calculating allele number and PIC values for each of the 32 SSR loci evaluated. Each of the primer pairs differed significantly in their ability to determine variability among the cultivars (Table 2). Some of the primers generated several markers but some generated

**Table 2.** Details of SSR markers used, indicating their location on rice chromosomes, number of alleles detected, allele size range and polymorphism information content (PIC).

Marker	Chromosome	SSR motifs	Total alleles	Polymorphic alleles	Size range (bp)	Difference	PIC value
RM1	1	(GA)26	5	5	75 - 135	60	0.753
RM10	2	(GA)15	2	2	142 - 159	17	0.202
RM13	5	(GA)16	2	2	130 - 150	20	0.353
RM16	3	(GA)15	3	3	165 - 187	22	0.486
RM17	12	(GA)21	3	3	160 - 185	15	0.442
RM19	4	(ATC)10	2	2	192 - 252	60	<b>0.157</b>
RM44	8	(GA)16	4	2	100 - 125	25	0.552
RM55	5	(GA)17	2	2	216 - 247	31	<b>0.157</b>
RM60	3	(AATT)5AATCT(AATT)	2	2	162 - 176	14	0.291
RM70	7	(ATT)33	13	13	128 - 167	39	<b>0.897</b>
RM72	8	(TAT)5C(ATT)15	6	6	151 - 200	49	0.756
RM-110	2	(GA)15	5	5	138 - 159	21	0.790
RM122	5	(GA)11	5	5	229 - 237	8	0.720
RM163	5	(GGAGA)4(GA)11C(GA)20	9	9	130 - 175	45	0.770
RM170	6	(CCT)7	5	5	99 - 119	20	0.680
RM182	7	(AT)16	8	8	328 - 349	21	0.828
RM201	9	(CT)17	4	4	144 - 159	15	0.707
RM202	11	(GA)30	5	5	161 - 190	29	0.769
RM222	10	(CT)18	3	3	199 - 225	26	0.640
RM223	8	(GA)25	4	4	140 - 170	30	0.576
RM224	11	(GA)13	4	4	118 - 157	39	0.696
RM234	7	(GA)25	3	3	133 - 163	30	0.640
RM241	4	(CT)31	3	3	104 - 149	45	0.554
RM242	9	(CT)26	3	3	197 - 255	58	0.485
RM252	4	(GA)19	9	9	194 - 262	68	0.744
RM253	6	(GA)25	4	3	117 - 146	29	0.648
RM257	9	(CT)24	3	3	132 - 147	15	0.647
RM263	2	(CT)34	5	5	162 - 199	37	0.633
RM302	1	(GT)30(AT)8	6	6	120 - 191	71	0.778
RM310	8	(GT)19	5	5	87 - 123	36	0.738
RM333	10		4	4	166 - 196	30	0.637
RM348	4	(CAG)7	3	3	131 - 143	12	0.560
Total			144	141			
Average			4.5	4.4			0.603

**Figure 1.** SSR banding pattern of 32 of basmati and aromatic/scented and other non-aromatic *indica* and *japonica* varieties of Asian rice generated by primer pair RM1. The lanes represent (M) 25 bp molecular size marker; (1) Basmati-370(P), (2) Mushkan, (3) Basmati-Pak, (4) IR8, (5) Basmati-198, (6) Super-basmati, (7) JP5, (8) Mahlar-346, (9) Purple-marker, (10) Dokri-basmati, (11) IR36, (12) Kinmaze, (13) KangniXTorh, (14) Sonahri-kangni, (15) Kharai-ganga, (16) KSK-133, (17) Shahkar, (18) Shandar, (19) Mehak, (20) Azucena, (21) Kasalath, (22) Basmati-370(I), (23) Dehradun-basmati(I), (24) KDML105, (25) Basmati-217, (26) Dehradun-basmati(N), (27) Punjab-basmati, (28) Pusa-basmati, (29) Chini-Sakkor, (30) Ranbir-basmati, (31) Jasmine-scented, (32) Niaw-hawn-mali.

only few. Among the polymorphic markers, 5 amplified two alleles each, 8 produced three alleles each, 6 generated four alleles each, 7 produced five alleles each, 2 markers gave six alleles, 1 marker gave seven alleles, 1 gave eight alleles, 2 produced nine alleles each and 1 marker produced thirteen alleles. A total of 144 alleles were detected across 35 cultivars using 32 SSR markers (Table 2). The number of alleles per locus generated by each marker varied from 2 (RM10, RM13, RM19, etc) to 13 (RM70) with an average of 4.5 alleles per locus. The overall size of the amplified product varied from 75 bp (RM1) to 349 bp (RM182). The size difference between the smallest and largest allele at a given SSR locus varied from 8 (RM122) to 71 (RM302). Of the 144 alleles scored, 141 (98%) were found to be polymorphic. Maximum number of polymorphic alleles, that is 13, was obtained with the marker RM70, while the minimum number of alleles (e.g., 2) were amplified by using RM10 and RM13 markers etc. The average number of polymorphic alleles per marker was 4.41.

As may be seen from Table 3, there appears to be a correlation between the number of alleles detected and the number of repeats in the SSR loci. For example, the microsatellite loci containing the (AT) repeat motifs varying from (AT)16 to (ATT)33 in RM70 and RM182 respectively, show a correlation with the number of alleles they revealed (8 alleles for AT(16) and 13 alleles for ATT(33)).

### Rare alleles

An allele that was observed in only one or two of the thirty-five cultivars was considered rare (defined as a frequency  $\leq 0.05$  in the total sample). A total of 32 (22%) rare alleles were observed at 12 of the 32 SSR loci, with an average of 2.7 rare alleles per locus. Maximum number of rare alleles were observed at RM163 locus (6 alleles), followed by RM70 and RM252 loci (5 alleles each). In general, markers detecting a greater number of alleles per locus detected more rare alleles. Twenty-one (60%) of the rice cultivars had rare SSR alleles. Maximum number of rare alleles were present in two cultivars, IR8 and Kasalath (3 alleles each), which can be attributed to their diverse origin as compared to other cultivars. Two rare alleles were found in each of 'Mushkan', 'Basmati-198', 'Kinmaze', 'Sonahri-Kangni', 'Kharai-Ganga', 'Pusa-Basmati', and 'PK386', while 12 cultivars had single rare allele at individual locus. A variety was assigned null allele for a microsatellite locus whenever an amplification product could not be detected for a particular genotype-marker combination. In the set of 40 cultivars, 2 loci showed null alleles. Null allele was observed in 'Super-basmati' at RM241 locus and, in 'Jasmine-scented' and 'Niaw-Hawn-Mali' at RM60 locus.

### PIC values

Polymorphism information content (PIC) value is a reflection

of allele diversity and frequency among the varieties. PIC value of each marker can be evaluated on the basis of its alleles and it varied greatly for all the SSR loci tested. The level of polymorphism among the 35 varieties was evaluated by calculating PIC values for each of the 32 SSR loci. The PIC values varied widely among loci and ranged from 0.157 (RM19 and RM55) to 0.897 (RM70) with an average of 0.603 per locus (Table 2). PIC values showed a significant positive linear correlation with number of alleles at SSR locus ( $r = 0.73$ ;  $P < 0.01$ , analysis of variance). Cultivar 'Basmati-198' gave the highest numbers of alleles (42). It was followed by 'Dokri-basmati' and 'Basmati-217' producing 40 alleles each, while cultivars 'JP5', 'Chini-Sakkor', 'Jasmine-Scented' and 'Niaw-Hawn-Mali' gave the least number of alleles (that is 34 alleles each). Interestingly, aromatic cultivars were found to be more nearer to japonica types as compared to non-aromatic cultivars.

### Similarity matrix

The SSR-derived data were subjected to calculate the genetic similarity (Table 3). The similarity matrix was used to determine the level of relatedness among the cultivars studied. Pair-wise estimates of similarity ranged from 0.19 to 0.90 (Table 3) and the average similarity among all 35 cultivars was 0.41. Two basmati cultivars 'Basmati-370' (No.1) and 'Basmati-Pak' (No.3) were found close to each other having a similarity index of 90%. This was followed by 75% similarity between two pairs of basmati cultivars 'Basmati-370' (No.1) and 'Super-basmati' (No.6), and 'Basmati-370' (No.1) and 'Ranbir-basmati' (No.30). The lowest level of similarity (19%) was obtained between two pairs of cultivars, 'IR8' (No.4) and 'Super-basmati' (No.6), and 'Kinmaze' (No.12) and 'PK386' (No.34). As expected, genetic similarities between the aromatic cultivars were comparatively high. While comparing aromatic and non-aromatic cultivars, the aromatic cultivars tend to be more similar to each other than to non-aromatic group. The similarity coefficients among the aromatic cultivars ranged from 0.22 to 0.90, but, these values varied from 0.19 to 0.72 in non-aromatic cultivars. The comparison of aromatic cultivars showed that these were closely related but were not exactly similar to any other cultivar. Amongst the analyzed cultivars, 'IR8' (No.4) and 'IR36' (No.11) showed strong dissimilarity to all other cultivars used in this study. The similarity coefficients of 'IR8' (No.4) with all the other cultivars ranged from 0.19 to 0.49, while 'IR36' (No.11) showed similarity coefficients of 0.20 to 0.54 with all the genotypes used. Similarly, the similarity coefficients of 'Kinmaze' (No.12) with other cultivars ranged from 0.19 to 0.54.

### Cluster and principal component analysis

The genetic relationship among the rice cultivars was assessed by a UPGMA cluster analysis of the similarity

**Table 3.** Nei and Li coefficients of similarity among pairs of Asian rice cultivars.

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00																
2	0.44	1.00															
3	0.91	0.43	1.00														
4	0.31	0.34	0.29	1.00													
5	0.71	0.35	0.70	0.36	1.00												
6	0.78	0.41	0.76	0.27	0.71	1.00											
7	0.56	0.49	0.56	0.33	0.40	0.52	1.00										
8	0.70	0.43	0.68	0.32	0.61	0.77	0.49	1.00									
9	0.38	0.46	0.41	0.52	0.42	0.36	0.36	0.33	1.00								
10	0.64	0.38	0.67	0.40	0.62	0.60	0.49	0.48	0.48	1.00							
11	0.30	0.50	0.33	0.44	0.30	0.28	0.44	0.28	0.53	0.43	1.00						
12	0.49	0.49	0.52	0.34	0.41	0.40	0.53	0.48	0.43	0.40	0.37	1.00					
13	0.41	0.69	0.44	0.43	0.39	0.38	0.41	0.43	0.58	0.42	0.57	0.46	1.00				
14	0.54	0.59	0.54	0.46	0.48	0.45	0.43	0.55	0.58	0.45	0.52	0.59	0.63	1.00			
15	0.42	0.45	0.43	0.44	0.49	0.43	0.38	0.46	0.58	0.45	0.45	0.47	0.57	0.62	1.00		
16	0.39	0.44	0.40	0.46	0.41	0.30	0.33	0.30	0.53	0.45	0.49	0.39	0.41	0.54	0.54	1.00	
17	0.39	0.44	0.39	0.45	0.36	0.35	0.40	0.32	0.62	0.44	0.41	0.43	0.43	0.46	0.51	0.65	1.00
18	0.39	0.37	0.40	0.51	0.37	0.35	0.46	0.33	0.53	0.52	0.37	0.39	0.37	0.46	0.54	0.71	0.75
19	0.58	0.48	0.56	0.35	0.45	0.54	0.55	0.56	0.31	0.48	0.28	0.50	0.43	0.45	0.35	0.38	0.42
20	0.63	0.47	0.62	0.43	0.55	0.58	0.53	0.63	0.43	0.54	0.37	0.46	0.44	0.49	0.42	0.34	0.29
21	0.53	0.46	0.51	0.40	0.42	0.41	0.42	0.54	0.33	0.46	0.38	0.40	0.45	0.50	0.38	0.38	0.32
22	0.71	0.42	0.64	0.29	0.57	0.58	0.43	0.55	0.38	0.56	0.32	0.37	0.39	0.41	0.35	0.34	0.39
23	0.57	0.55	0.58	0.45	0.52	0.46	0.49	0.51	0.41	0.53	0.53	0.48	0.62	0.67	0.51	0.45	0.40
24	0.53	0.39	0.51	0.38	0.55	0.47	0.48	0.54	0.40	0.51	0.46	0.41	0.46	0.48	0.46	0.39	0.36
25	0.73	0.45	0.71	0.35	0.69	0.65	0.46	0.58	0.43	0.57	0.36	0.54	0.52	0.52	0.43	0.42	0.44
26	0.61	0.40	0.62	0.29	0.67	0.53	0.41	0.55	0.40	0.54	0.35	0.44	0.37	0.49	0.42	0.41	0.36
27	0.63	0.46	0.61	0.27	0.64	0.62	0.44	0.59	0.41	0.46	0.46	0.48	0.48	0.50	0.48	0.33	0.40
28	0.53	0.49	0.49	0.45	0.55	0.40	0.38	0.32	0.54	0.47	0.37	0.43	0.48	0.55	0.49	0.43	0.45
29	0.53	0.33	0.56	0.38	0.48	0.55	0.53	0.55	0.42	0.54	0.36	0.38	0.41	0.48	0.46	0.28	0.35
30	0.77	0.49	0.76	0.31	0.68	0.64	0.48	0.64	0.47	0.53	0.37	0.53	0.46	0.53	0.49	0.39	0.43
31	0.51	0.38	0.44	0.33	0.31	0.42	0.42	0.44	0.29	0.44	0.38	0.35	0.41	0.43	0.33	0.28	0.38
32	0.41	0.44	0.38	0.35	0.38	0.39	0.39	0.39	0.39	0.49	0.44	0.35	0.41	0.46	0.41	0.41	0.38
33	0.37	0.44	0.32	0.41	0.32	0.30	0.46	0.35	0.55	0.40	0.57	0.34	0.49	0.46	0.42	0.44	0.39
34	0.38	0.46	0.35	0.54	0.40	0.33	0.34	0.33	0.56	0.48	0.58	0.28	0.53	0.50	0.48	0.53	0.49
35	0.41	0.51	0.41	0.45	0.41	0.40	0.38	0.42	0.47	0.47	0.41	0.39	0.48	0.46	0.46	0.41	0.43

matrix shown in Table 3. A UPGMA cluster diagram grouped the 35 rice genotypes into two major clusters, I and II, differentiating the aromatic rice cultivars from non-aromatic cultivars having additional sub-clusters within the both clusters (Figure 2). Group-I consisted 16 long slender basmati and other aromatic and quality rice cultivars used in this study and 2 coarse as well as 3 *japonica* type cultivars, while group-II comprised of a total of 14 cultivars most of them belonged to non-aromatic coarse group with few exception. The cluster analysis placed most of the tall, long-grained aromatic cultivars together, which shows a high level of genetic relatedness

among these cultivars. The aromatic cultivars formed distinct clusters from those in case of non-aromatic group. Three *japonica* cultivars were more scattered and closer to the aromatic cultivars than to the non-aromatic *indica* group. The dendrogram showed that the genotypes that were derivatives of genetically similar type, clustered together. The genotypes were closely related with each other within the first aromatic group, whereas relatively diverse within the second non-aromatic group. The cluster analysis also revealed that the basmati varieties were nearer to the *japonica* group than to the *indica* varieties. The cultivars in the same subgroup mostly

Table 3. contd.

#	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
18	1.00																	
19	0.40	1.00																
20	0.34	0.55	1.00															
21	0.35	0.41	0.50	1.00														
22	0.39	0.53	0.61	0.45	1.00													
23	0.40	0.44	0.50	0.49	0.48	1.00												
24	0.43	0.49	0.55	0.52	0.51	0.52	1.00											
25	0.47	0.48	0.52	0.48	0.66	0.51	0.65	1.00										
26	0.41	0.48	0.51	0.48	0.59	0.48	0.53	0.66	1.00									
27	0.33	0.49	0.50	0.44	0.50	0.49	0.62	0.67	0.70	1.00								
28	0.41	0.42	0.41	0.37	0.43	0.49	0.40	0.56	0.53	0.52	1.00							
29	0.43	0.52	0.53	0.60	0.53	0.52	0.63	0.54	0.46	0.44	0.43	1.00						
30	0.31	0.52	0.60	0.52	0.58	0.52	0.60	0.67	0.58	0.67	0.57	0.55	1.00					
31	0.35	0.52	0.43	0.36	0.43	0.49	0.48	0.41	0.33	0.47	0.43	0.45	0.48	1.00				
32	0.30	0.34	0.46	0.49	0.38	0.49	0.50	0.49	0.41	0.47	0.38	0.50	0.55	0.50	1.00			
33	0.37	0.38	0.44	0.45	0.46	0.40	0.51	0.42	0.46	0.43	0.41	0.43	0.34	0.43	0.41	1.00		
34	0.60	0.33	0.35	0.36	0.40	0.44	0.44	0.46	0.30	0.36	0.44	0.34	0.37	0.42	0.36	0.55	1.00	
35	0.41	0.44	0.53	0.47	0.36	0.49	0.52	0.42	0.34	0.49	0.45	0.40	0.48	0.55	0.50	0.48	0.57	1.00

shared a high proportion of ancestry and/or agronomic characteristics (e.g., height, maturity, quality traits, etc). In this study, a number of traditional and improved cultivars originally from various regions, did not form distinct groups. These were interspersed with each another in the cluster analysis, which confirmed no association between the used cultivars SSR patterns and their geographic origin. The current study did not resolve the eco-geographical specific differentiation and the varieties from different origins were mixed and distributed throughout the two main groups.

The PCA was also done to determine the genetic relationships among basmati and other scented/quality rice varieties from different regions of Asia (Figure 3). The groupings identified by PCA were very similar to those identified by the UPGMA cluster analysis. Of 35 genotypes, 13 aromatic and 4 coarse/japonica rice cultivars formed a distinct cluster in the right side of the diagram, separated from the main indica cluster. This cluster corresponded to the 21 varieties that appear as a major group in Figure 2. A second group of 16 genotypes included the 11 coarse and 3 aromatic cultivars corresponded well with the second group of varieties in the dendrogram. The remaining 4 cultivars were interspersed within the two clusters. Three coarse genotypes and two japonica cultivars were located close to the aromatic cluster in the PCA, suggesting close relationship between these cultivars.

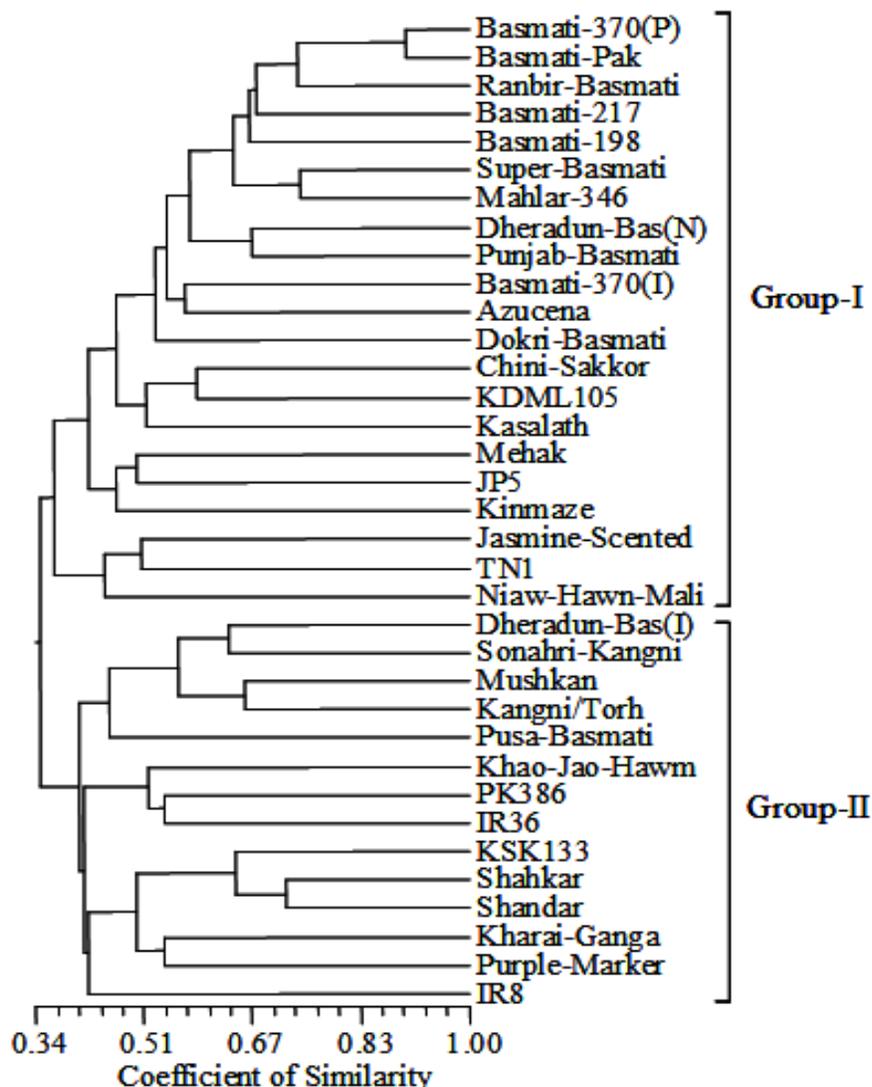
## DISCUSSION

Agro-morphological and seed traits have long been the

means of studying the taxonomy and variability among plant species. The microsatellites or SSRs are amongst the most widely used DNA marker types for many purposes such as diversity, genome mapping and varietal identification etc.

The use of these markers to investigate the genotypic variations among the different cultivars is previously reported by some researchers (Joshi and Behera, 2006; Nagaraju et al., 2002; Singh et al., 2004). The implication of these molecular techniques to investigate genetics of Pakistani cultivars is much limited. We used 32 SSR markers to determine the genetic diversity of 35 cultivars of the Asian rice. The results indicated a high level of genetic variation in the cultivars used. The microsatellites showed a number of alleles shared by basmati and other fine rice cultivars, while relatively less number of bands was shared among aromatic and non-aromatic rice cultivars. The cultivar 'IR8', 'IR36' and 'Kinmaze' were distinct and showed a limited number of shared alleles with the other non-aromatic cultivars. The microsatellite assay generated variety-specific alleles in some of the genotypes screened; these could be used for DNA fingerprints to identify different varieties.

The results based on 32 microsatellite markers analysis showed a clear division of cultivars into aromatic and non-aromatic groups. A close relationship between 'Basmati-370' and 'Basmati-Pak' could be due to small differences at the DNA level between the two varieties. In addition, these varieties were morphologically similar in agronomic traits, which supported a close relationship between them. However, there exist sufficient variations at molecular level among all rice varieties.

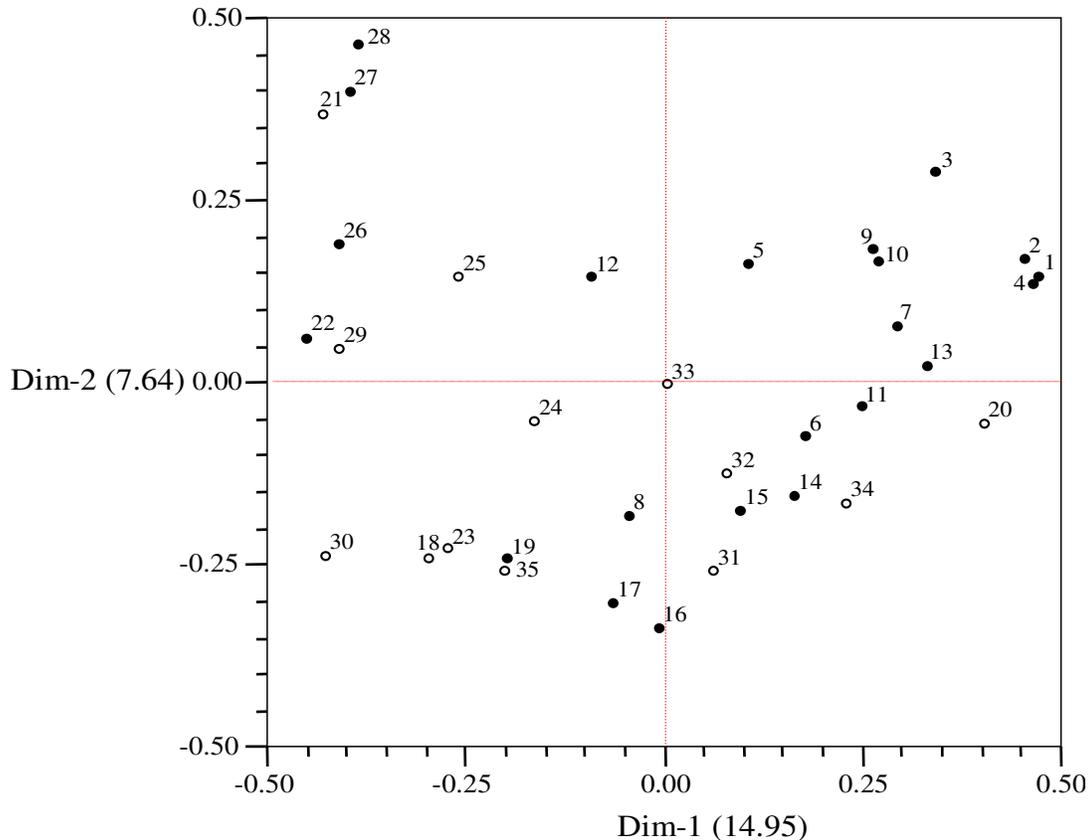


**Figure 2.** UPGMA cluster analysis showing the diversity and relationship among 35 Asian rice cultivars based on 144 alleles generated by 32 SSR markers.

The number of alleles detected by microsatellite markers varied from 2 to 13 with an average of 4.5 alleles per locus. The numbers of alleles observed in the present study correspond well to some earlier reports on the Indian rice varieties (Siwach et al., 2004; Neeraja et al., 2005). Our results were also comparable to 2.0 - 5.5 alleles per SSR locus for various classes of microsatellites reported by Cho et al. (2000), who used a different set of rice germplasm. Contrarily the number of alleles detected in the present study was significantly higher than the average number of alleles in aromatic rice of India (Nagaraju et al. 2002; Singh et al. 2004; Joshi and Behera, 2006). This may be due to inclusion of both *indica* and *japonica* varieties and inclusion of basmati as well as non-basmati cultivars of diverse origin in the present study. Similarly the proportion of alleles noted in this study was relatively lower than that reported earlier

reported by Ni et al. (2002), Jain et al. (2004), Xu et al. (2004), Lu et al. (2005), Brondani et al. (2006), Jayamani et al. (2007) and Thomson et al. (2007), who reported an average of 6.8, 7.8, 11.9, 6.57, 14.6, 7.7 and 13 alleles per locus using rice subspecies, Indian quality rice germplasm, US rice genetic resources, traditional varieties of Brazilian rice, a diverse collection of Portuguese rice and Indonesian rice germplasm, respectively. This inconsistency might be due to the genotypes used and selection of SSR primers with scorable alleles. The SSR markers involving di-nucleotide repeats motifs particularly those with GA repeats amplified relatively small number of bands as reported earlier in rice (Cho et al., 2000; Saini et al., 2004).

The level of polymorphism determined by the PIC values, was quite high and varied (range 0.157 to 0.897, average value 0.603) considerably among SSR loci. The



**Figure 3.** The two-dimensional scaling resulting from PCA analysis (NTSYS-PC) of 35 genotypes using genetic diversity data for 144 alleles at 32 SSR loci. Filled circles represent Basmati and other scented/quality rice varieties, while open circles denote coarse indica/japonica cultivars of rice.

PIC values observed in our study were comparable to previous estimates of microsatellite marker analysis in rice (Jain et al., 2004; Saini et al., 2004; Siwach et al., 2004; Lu et al., 2005; Jayamani et al., 2007; Thomson et al., 2007). The PIC value was higher than the earlier observations (Singh et al., 2004; Joshi and Behera 2006), but lower than that previously reported by Xu et al. (2004) and Brondani et al. (2006), who observed an average PIC value of 0.73 and 0.74 for the world collection and traditional varieties of Brazilian rice, respectively. They included a more diverse set of rice germplasm.

Similarity coefficients among various cultivars ranged from 0.19 to 0.90 in present investigations with an average of 0.41. Saini and colleagues (2004) also reported almost similar values of similarity co-efficients among 18 basmati and non-basmati varieties using molecular markers. Likewise, similarity coefficients ranging from 0.24 to 0.92 were observed in eight basmati accessions originating from Pakistan and one solitary *indica* accession for the SSR analysis (Jayamani et al., 2007). Ravi et al. (2003) obtained an average genetic similarity of 0.79 between 40 cultivated varieties and five wild relatives of rice with SSR markers. Jain et al. (2004) obtained relatively high degree of genetic similarity ranging

from 79 to 99.6% with an average of 89% between Indian basmati/aromatic rice germplasm using panels of fluorescently-labeled microsatellite markers. Similarly, Siwach et al. (2004) also observed higher level of similarity ranging from 0.67 to 0.91 among basmati and non-basmati long-grain *indica* rice varieties using microsatellite markers. One of the reasons for this high level of similarity recorded by these studies could be due to intra-specific variation in the germplasm used.

Cluster analysis based on similarity coefficients placed 35 rice genotypes into 2 major groups. Most of the aromatic and japonica cultivars fell into close sub-groups. For example, 15 of the 19 aromatic cultivars were placed in the upper region of dendrogram. This classification was further confirmed on the basis of their morphological traits (for example, grain quality and aromatic type etc). Cluster analysis grouped most of the basmati cultivars from Punjab like 'Basmati-370', 'Basmati-Pak', 'Basmati-198' and 'Super-Basmati' together indicating that they are genetically similar with each other and have common ancestors. These rice cultivars share 'Basmati-370' as one of the parents in their pedigree. Similar studies conducted by Nagaraju et al. (2002) and Saini et al. (2004) using SSR markers, long-grain basmati cultivars

were grouped together, whereas the other short-grained non-aromatic rice fell into different groups. These authors also reported that traditional and evolved basmati varieties shared a high degree of similarity using SSR markers.

Though, the number of basmati and *japonica* rice genotypes analyzed in our study were low, our results show a clear distinction between *indica* and traditional basmati rice varieties with the latter being closer to *japonica* than to *indica*. The study shows that basmati rice varieties are genetically distinct from other groups such as *O. sativa*, namely *indica* and *japonica* etc. Higher levels of genetic diversity between basmati and non-basmati supports the concept that former had a long history of independent evolution and diverged from non-basmati rices a long time ago through conscious selection and patronage (Nagaraju et al., 2002; Jain et al., 2004; Saini et al., 2004). The placement of basmati and aromatic varieties closer to *japonica* than *indica*, is in conformity with earlier studies using SSR markers (Jain et al., 2004; Saini et al., 2004; Siwach et al., 2004; Garris et al., 2005; Thomson et al., 2007). Based on isozyme analysis, Glazemann (1987) also described that the basmati genotypes were genetically distinct from the other groups. The SSR markers used in this study are well distributed on 12 chromosomes, and are located in both coding and non-coding segments of the genome (Cho et al., 2000; Temnykh et al., 2000). Although most of the Asia varieties could be assigned as belonging to either the aromatic or non-aromatic *indica* groups, a few varieties were detected. Three traditional varieties fell outside of the larger clusters: 'Pusa-basmati' and 'Dheradun-basmati' (aromatic) was closer to the non-aromatic cultivars, while the 'Mahlar-346' (coarse) clustered with the aromatic and *japonica* types. The unusual nature of these traditional cultivars may be due to admixture of seeds during post-harvest handling.

From this study, it might be concluded that microsatellite analysis could be efficiently utilized for diversity analysis and differentiation of aromatic and non-aromatic rice cultivars. In addition, marker-based identification and differentiation of Basmati rices could be helpful to preserve the integrity of the high quality rice varieties such as basmati and aromatic varieties, etc, to benefit both farmers and consumers communities.

## ACKNOWLEDGMENTS

We are indebted to various Rice Research Institutes/Programs of Pakistan for providing the seed material of their traditional and improved cultivars of rice. We would also like to acknowledge the seed of japonica cultivars of rice kindly provided by the National Institute of Agrobiological Sciences, Tsukuba, Japan. This is the part of Ph.D. thesis work of the first author. We gratefully acknowledge the financial support for this work from the Agricultural Linkages Program, PARC under the AREP.

Financial assistance from the Higher Education Commission, Islamabad, Pakistan, is also gratefully acknowledged in form of scholarship to the first author. Authors are also very grateful to Dr. M. Shahid Masood, Sr. Director, PGRI, NARC, Islamabad for providing laboratory facilitation for this work.

## REFERENCES

- Brondani C, Borba TCO, Rangel PHN, Brondani RPV (2006). Determination of genetic variability of traditional varieties of Brazilian rice using microsatellite markers. *Genet. Mol. Biol.* 29: 676-684.
- Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayers N, Cartinhour S (2000). Diversity of microsatellites derived from genomic libraries and Gene Bank sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 100: 713-722.
- Coburn JR, Temnykh SV, Paul EM, McCouch SR (2002). Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.). *Crop Sci.* 42: 2092-2099.
- Garris AJ, Thomas HT, Coburn J, Kresovich S, McCouch S (2005). Genetic structure and diversity in *Oryza sativa* L. *Genet.* 169: 1631-1638.
- Glazemann JC (1987). Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* 74: 21-30.
- Jain S, Jain RK, McCouch SR (2004). Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theor. Appl. Genet.* 109: 965-977.
- Jayamani P, Negrao S, Martins M, Macas B, Oiveira MM (2007). Genetic relatedness of Portuguese rice accessions from diverse origins as assessed by microsatellite markers. *Crop Sci.* 47: 879-886.
- Joshi RK, Behrea L (2006). Identification and differentiation of indigenous non-Basmati aromatic rice genotypes of India using microsatellite markers. *Afr. J. Biotechnol.* 6: 348-354.
- Kang HW, Cho YG, Yoon UH, Eun MY (1998). A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. *Plant Mol. Biol. Report*, 16: 1-9.
- Lu H, Redus MA, Coburn JR, Rutger JN, McCouch SR, Ti TH (2005). Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. *Crop Sci.* 45: 66-76.
- McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, Declerck G, Schneider D, Cartinhour S, Ware D, Stein L (2002). Development of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* 9: 199-207.
- Nagaraju J, Kathirvel M, Kumar R, Siddiq EA, Hasnain, SE (2002). Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. *Proc. Natl. Acad. Sci. USA*, 99: 5836-5841.
- Neeraja CN, Hariparshad AS, Malathi S, Siddiq EA (2005). Characterization of tall landraces of rice (*Oryza sativa* L.) using gene-derived simple sequence repeats. *Curr. Sci.* 88: 149-152.
- Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA*, 76: 5269-5273.
- Ni J, Colowit PM, Mackill DJ (2002). Evaluation of genetic diversity in rice sub species using microsatellite markers. *Crop Sci.* 42: 601-607.
- Rabbani MA, Pervaiz ZH, Masood MS (2008). Genetic diversity analysis of traditional and improved cultivars of Pakistani rice (*Oryza sativa* L.) using RAPD markers. *Electron. J. Biotech.* 11: 1-10.
- Ravi M, Geethanjali S, Sameeyafaheen F, Maheswaran M (2003). Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. *Euphytica*, 133: 243-252.
- Rohlf FJ (2005). NTSYSpc (Numerical Taxonomy and Multivariate Analysis System). Version 2.2 Exeter Software, Appl Biostat. Inc., New York, USA.
- Saini N, Jain N, Jain S, Jin R (2004). Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. *Euphytica*, 140: 133-146.

- Sasaki T, Burr B (2000). International rice genome sequence project: the effort to complete the sequence of rice genome. *Curr. Opin. Plant Biol.* 3: 138-141.
- Singh RK, Sharma RK, Singh AK, Singh VP, Singh NK, Tiwari SP, Mohapatara T (2004). Suitability of mapped sequence tagged microsatellite site markers for establishing distinctness, uniformity and stability in aromatic rice. *Euphytica*, 135: 135-143.
- Siwach P, Jain S, Saini N, Choudary VK, Jain RK (2004). Allelic diversity among Basmati and non-Basmati long-grain indica rice varieties using microsatellite markers. *J. Plant Biochem. Biotechnol.* 13: 25-32.
- Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 100: 697-712.
- Thomson MJ, Septininsihin EM, Suwardjio F, Santoso TJ, Silintonga TS, McCouch SR (2007). Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* 114: 559-568. [www.gramene.org](http://www.gramene.org).
- Xu Y, Beachel H, McCouch SR (2004). A marker-based approach to broadening the genetic base of rice in the USA. *Crop Sci.* 44: 1947-1959.