Pak. J. Bot., 42(4): 2551-2564, 2010.

GENETIC ANALYSIS OF BASMATI AND NON-BASMATI PAKISTANI RICE (*ORYZA SATIVA* L.) CULTIVARS USING MICROSATELLITE MARKERS

MALIK ASHIQ RABBANI[1*], MUHAMMAD SHAHID MASOOD[1], ZABTA KHAN SHINWARI[2] AND KAZUKO YAMAGUCHI-SHINOZAKI[3]

[1]Institute of Agri-Biotechnology & Genetic Resources, National Agricultural Research Center Islamabad, Pakistan
[2]Biotechnology Department, Quaid-i-Azam University, Islamabad, Pakistan
[3]Biological Resources Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Japan
[*]Corresponding author E-mail: rabbani316@yahoo.com

Abstract

Information of genetic variability and relatedness among rice genotypes is essential for future breeding programmes and derivation of superior cultivars. The objective of the present study was to evaluate the genetic relationship among traditional and improved cultivars of Pakistani rice and to determine differences in the patterns of variation between two *indica* rice groups: basmati and non-basmati. Forty-one cultivars were evaluated by means of 30 microsatellite markers distributed over the whole rice genome. A total of 104 alleles were detected by 30 markers, all of them (100%) were polymorphic. The number of alleles generated by each marker ranged from 2 to 6 with an average of 3.5 alleles marker^{-1}. Polymorphism information content (PIC) varied from 0.259 to 0.782 with an average of 0.571. A significant positive correlation (r = 0.71**) was found between the number of alleles at SSR locus and the PIC values. Pair-wise Nei and Li's similarity coefficients ranged from 0.10 to 0.99. A dendrogram based on cluster analysis by microsatellite polymorphism grouped 41 rice cultivars into 2 major groups effectively differentiating the late maturing, tall and slender-grain basmati and other aromatic rice cultivars from the early, short statured, short bold and long bold grain non-aromatic cultivars. Higher level of genetic diversity between basmati and non-basmati support the concept that former had a long history of independent evolution and diverged from non-basmati rice a long time ago through human selection and patronage. Present investigation further indicated that genetically basmati rice is different from that of coarse *indica* and *japonica* type. The results suggested that microsatellite markers could efficiently be utilized for diversity analysis, and differentiation of basmati and non-basmati rice cultivars. In addition, marker-based identification of traditional basmati rice may help in maintaining the integrity of this high quality product to the benefit of both farmers and consumers.

Introduction

Rice (*Oryza sativa* L.) is one of the oldest domesticated crop species in the world and has fed more people throughout human history than any other food plant. It is also widely recognized an ideal model plant for the study of grass genetics and genome organization due to its diploid origin (2n = 24), relatively small genome size (430 Mb), availability of whole genome sequences, relative ease of transformation, development of several key genomic mapping resources, considerable level of genetic polymorphism, large amount of well conserved genetically diverse material and the availability of widely collected, compatible wild species. Rice occupies almost one-fifth of the total land area covered under cereals. It is cultivated over an area of 1.5 billion hectares of land having overall worldwide production of 645 million tons per annum. Rice is grown under diverse

cultural conditions and over wide geographical range. Most of the world's rice is cultivated and consumed in Asia, which constitutes more than half of the global population. Approximately 11% of the world's arable land is planted annually to rice, and it ranks next to wheat (Chakravarthi & Naravaneni, 2006).

In Pakistan, rice occupies about 10% of the total cultivated area, accounts for 6.1% of value added in agriculture and 1.3% in gross domestic product. Production of rice during 2005-06 was recorded as 5,547 thousand tons, 10.4% higher than previous year with 6.1% increase in yield per hectare (Anon., 2006). Basmati rice is economically more valuable and highly priced in the domestic as well as international markets due to superior quality of the grain and constitutes an important source of revenue for the country. Pakistan has become one of the leading exporters of aromatic rice in the global markets. Country's annual rice export stands at about 0.84 million tons of basmati and 2.85 million tons of non-basmati rice, contributing a total of 69325 million rupees to the Pakistan economy (Anon., 2006).

The study of genetic diversity in basmati gene pool and other rice types is necessary for varietal identification, proper purity maintenance, and basmati rice breeding. In view of the implementation of plant variety protection rights and export under WTO regulations, increasing attention is being paid towards comprehensive characterization of elite basmati quality rice germplasm, supplementing the existing morphological descriptors with reliable and repeatable DNA based marker profiles. Genetic analysis is important to ensure the export quality of basmati rice, for maintaining the distinctiveness' of basmati varieties and to differentiate between the various grades of basmati rice (Bligh, 2000; Nagaraju *et al*., 2002). In most of previous studies, basmati types clustered into a separate group distinct from that of *indica* and *japonica* rice varieties (Nagaraju *et al*., 2002; Jain *et al*., 2004). The results indicated that basmati germplasm may have a long, independent and complex pattern of evolution that distinguishes it from other groups within *Oryza sativa* (Jain *et al*., 2004). More than 40 rice cultivars have been developed and released for cultivation in Pakistan since early nineteen century. Most of the information we have so far is about common varieties, and our knowledge on traditional cultivars is still lacking. Management of the indigenous aromatic rice genetic resources by way of characterization and documentation helps in protection of these unique bio-resources in accordance with the provision laid out in the (1992) meet on Conservation of Biological Diversity (CBD).

Molecular markers have proven to be powerful tools in the assessment of genetic variation and in the elucidation of genetic relationships within and among species (Shinwari, 1995). Among PCR based markers, microsatellites are highly polymorphic, more reproducible, co-dominant and distributed throughout the genome. More than 2200 microsatellite markers have been mapped to specific locations in rice genome (McCouch *et al*., 2002). These markers have been utilized for many purposes including genome mapping, gene tagging, estimation of genetic diversity, varietal differentiation and purity testing (McCouch *et al*., 1997; Coburn *et al*., 2002; Nagaraju *et al*., 2002). It has been reported that the number of parental lines used in the breeding of many of the currently released Pakistani long-grain aromatic cultivars of rice is very limited and for this reason, it is important that any method used can readily detect variation between closely related cultivars (Rabbani *et al*., 2008). Our objectives were to use microsatellite markers for the assessment of genetic variability and relationships among traditional and improved cultivars of rice from Pakistan and to determine differences in the patterns of diversity within two *indica* rice groups: basmati (aromatic) and non-basmati (coarse).

Table 1. Basmati and non-basmati cultivars of Pakistani rice used in the present study.

Sr. No.	Cultivar name	Parentage	Varietal group	Release year	Breeding institute*
1.	Basmati-370	Local selection	Aromatic	1933	RRI, KSK, Lahore
2.	Basmati-C622	Basmati-370/Mushkan-7	Aromatic	1964	RRI, KSK, Lahore
3.	Basmati-Pak	CM7-6/Basmati-370	Aromatic	1968	RRI, KSK, Lahore
4.	Basmati-198	Basmati-370/TN1	Aromatic	1972	RRI, KSK, Lahore
5.	Basmati-385	Basmati-370/TN1	Aromatic	1988	RRI, KSK, Lahore
6.	Super-Basmati	Basmati-370/10486	Aromatic	1996	RRI, KSK, Lahore
7.	Basmati-2000	Basmai-385/Super-Basmati	Aromatic	2000	RRI, KSK, Lahore
8.	Kashmir-Basmati	Mutant of Basmati-370	Aromatic	1981	NIAB, Faisalabad
9.	Rachna-Basmati	Basmati-370 SC Variant	Aromatic	1996	NARC, Islamabad
10.	Jajai-77	Pure line selection	Aromatic	1932	RRI, Dokri, Sindh
11.	Sugdasi-Bengalo	Introduction from Bengal	Aromatic	1942	RRI, Dokri, Sindh
12.	Sugdasi-Sadagulab	Local selection	Aromatic	1945	RRI, Dokri, Sindh
13.	Sonahri-Sugdasi	Local selection	Aromatic	1952	RRI, Dokri, Sindh
14.	Sugdasi-Ratria	Local selection	Aromatic	1956	RRI, Dokri, Sindh
15.	Khushboo-95	Mutant of Jajai-77	Aromatic	1996	NIA, Tandojam, Sindh
16.	Jhona-349	Local selection	Non-aromatic	1933	RRI, KSK, Lahore
17.	Sathra	Local selection	Non-aromatic	1934	RRI, KSK, Lahore
18.	Palman-Suffaid	Local selection	Non-aromatic	1939	RRI, KSK, Lahore
19.	Pak23710	Advanced breeding line	Non-aromatic	-	RRI, KSK, Lahore
20.	KS-282	Basmati-370/IR95	Non-aromatic	1982	RRI, KSK, Lahore
21.	NIAB-IR9	Mutant of IR6	Non-aromatic	2001	NIAB, Faisalabad
22.	Pak23717	Advanced breeding line	Non-aromatic	-	SSRI, Pindi Bhatian

Table 1. (Cont'd.).

Sr. No.	Cultivar name	Parentage	Varietal group	Release year	Breeding institute*
23.	IR6	Siam29/DGWG	Non-aromatic	1971	RRI, Dokri, Sindh
24.	Kangni-27	Pure line selection	Non-aromatic	1932	RRI, Dokri, Sindh
25.	DR-82	BU1/CR-115	Non-aromatic	1982	RRI, Dokri, Sindh
26.	DR-83	IR1833/C4-63	Non-aromatic	1983	RRI, Dokri, Sindh
27.	Pak23725	Advanced breeding line	Non-aromatic	-	RRI, Dokri, Sindh
28.	Sada-Hayat	PTB33/IR4432-53	Non-aromatic	1988	RRI, Dokri, Sindh
29.	DR-92	IR8/IET1039-2-4	Non-aromatic	1993	RRI, Dokri, Sindh
30.	Kanwal-95	Selection from CN540	Non-aromatic	1998	RRI, Dokri, Sindh
31.	Shua-92	Mutant of IR8	Non-aromatic	1993	NIA, Tandojam, Sindh
32.	Shadab	Mutant of IR6	Non-aromatic	1999	NIA, Tandojam, Sindh
33.	Sarshar	Mutant of IR8	Non-aromatic	2001	NIA, Tandojam, Sindh
34.	Swat-1	JP5/YRL	Non-aromatic	1984	ARI, Mingora, Swat
35.	Swat-2	CR126-42-2/IR2061-2/3	Non-aromatic	1984	ARI, Mingora, Swat
36.	Pakhal	IR36/IR10154-23-3-3// IR9129-209-2-2-1	Non-aromatic	1993	ARS, Mansehra
37.	Dilrosh-97	Chaing-Nan-TSE/IR747B2-6-3//IR9129-209-2-2	Non-aromatic	1997	ARI, Mingora, Swat
38.	Fakhre-Malakand	Taichung Sen-10// IR29/ I-Geo-Gen//Taichung Sen-10	Non-aromatic	2006	ARI, Mingora, Swat
39.	IR36	IR2042/CR9413	Non-aromatic	1976	IRRI, Philippines
40.	Kinmaze	Ryosaku/AichnakateAsahi	Japonica	1948	NIAS, Japan
41.	Nipponbare	Yamabiko/Sachikaze	Japonica	1961	NIAS, Japan

*ARI= Agricultural Research Institute, Mingora, Swat, Pakistan; ARS= Agricultural Research Station, Dhodial, Mansehra, Pakistan; NARC= National Agricultural Research Center, Islamabad, Pakistan; NIA= Nuclear Institute of Agriculture, Tandojam, Sindh, Pakistan; NIAB= Nuclear Institute for Agriculture & Biology, Faisalabad, Pakistan; RRI= Rice Research Institute, Dokri, Sindh, Pakistan; RRI, KSK= Rice Research Institute, Kala Shah Kaku, Lahore, Pakistan; SSRI= Soil Salinity Research Institute, Pindi Bhattian, Hafizabad, Pakistan.

Table 2. Details of SSR markers used, their location on rice chromosomes, number of polymorphic alleles and polymorphism information content (PIC) values.

Marker	Chromosome	SSR motifs	Polymorphic alleles	PIC value
RM1	1	(GA)26	3	0.561
RM10	2	(GA)15	3	0.437
RM13	5	(GA)16	3	0.582
RM16	3	(GA)15	3	0.513
RM17	12	(GA)21	3	0.658
RM19	4	(ATC)10	3	0.535
RM44	8	(GA)16	3	0.580
RM55	5	(GA)17	3	0.652
RM60	3	(AATT)5AATCT(AATT)	2	0.408
RM72	8	(TAT)5C(ATT)15	6	**0.782**
RM84	1	(TCT)10	2	0.444
RM122	5	(GA)11	5	0.704
RM161	5	(AG)20	6	0.567
RM163	5	(GGAGA)4(GA)11C(GA)20	5	0.732
RM171	10	(GATG)5	2	0.500
RM182	7	(AT)16	3	**0.259**
RM201	9	(CT)17	3	0.557
RM202	11	(GA)30	3	0.592
RM222	10	(GA)18	2	0.497
RM223	8	(GA)25	3	0.543
RM224	11	(GA)13	4	0.658
RM234	7	(GA)25	3	0.515
RM241	4	(GA)31	3	0.461
RM242	9	(GA)26	3	0.501
RM252	4	(GA)19	4	0.598
RM253	6	(GA)25	3	0.632
RM257	9	(GA)24	6	0.743
RM302	1	(GT)30(AT)8	5	0.676
RM310	8	(GT)19	5	0.757
RM348	4	(CAG)7	2	0.482
Total	-	-	**104**	-
Average	-	-	**3.5**	**0.571**

Materials and Methods

Plant materials: A total of 41 cultivars of Pakistani rice were used in present investigation (Table 1). Among them 15 belong to basmati/aromatic, 24 belong to non-basmati coarse, while two *japonica* cultivars, Kinmaze and Nipponbare served as control for comparison.

Microsatellite marker analysis: Total DNA was extracted from dried seeds of 10 individual plants of each cultivar according to the method described by Kang *et al*., (1998) with slight modification. Purity and concentration of DNA of each individual was monitored spectrophotometrically at a wavelength of 260 and 280nm using NanoDrop ND-1000 Spectrophotometer. All individual DNA samples were diluted to a working

concentration of 20ng/µl with TE before use. An equal amount of genomic DNA from 10 individuals of each cultivar was mixed to make a bulk sample for PCR analysis.

Forty SSR primer pairs were selected as a subset of markers on the basis of published rice microsatellite framework (Temnykh *et al*., 2000; McCouch *et al*., 2002), well distributed on all the 12 chromosomes and previously used for genetic diversity analysis of rice. Preference was given to those SSRs that have been used greatly for the analysis of basmati rice. Individual PCR amplification was performed in a total volume of 20µl reaction mixture that consisted of 2µl of 10×buffer (10mM Tris HCl pH 8.3, 50mM KCl, 1.5mM $MgCl_2$), 200µM each of dNTPs, 20pmol of each forward and reverse primer, 1 unit Taq DNA polymerase (ExTaq, Takara Japan) and 20ng template DNA. Amplifications were carried out in a GeneAmp PCR System 9700 (PE Applied Biosystems USA). Thermal cycler was programmed to 1 cycle of 5 min., at 94oC as an initial hot start and strand separation step. This was followed by 35 cycles of 1 min., at 94oC for denaturation, 1 min., at 55oC for annealing and 2 min., at 72oC for primer elongation. Finally, 1 cycle of 7 min., at 72oC was used for final extension.

After amplification, a 12µl aliquot of the amplified microsatellite samples was combined with 3µl of a loading buffer (0.4% bromo-phenol blue, 0.4% xylene cyanole and 5 ml of glycerol) and was analyzed directly on 3% GTG NuSieve agarose gels in 1X TBE buffer (10mM Tris-Borate, 1mM EDTA) buffer containing 0.5µg ml^{-1} of Ethidium bromide. A 20bp DNA ladder (Takara Bio Inc., Japan) was used as a size marker to compare the molecular weights of amplified products. After electrophoresis, the gels were documented using a FluorChem FC2 Imaging System (Alpha Innotech Corporation, USA).

Data analysis: Amplified products from microsatellite analyses were scored qualitatively for presence and absence for each marker allele-genotype combination. Data were entered into a binary matrix as discrete variables, 1 for presence and 0 for absence of the character. Polymorphism information content (PIC) value of a marker was calculated according to the following formula:

$$PIC = 1 - \sum_{j=1}^{n} P_{ij}^2$$

where P_{ij} is the frequency of *j*th allele for the *i*th marker, and summed over *n* alleles. Pair-wise comparisons of the cultivars based on the proportion of unique and shared alleles were used to measure the genetic similarity. Estimates of genetic similarity were calculated between all pairs of the cultivars according to Nei & Li (1979) based on the formula: Similarity (F) = 2Nab/(Na + Nb) where Nab = number of fragments shared by individuals ‘a’ and ‘b’, Na = total number of fragments detected in individual ‘a’ and Nb = total number of fragments shown by individual ‘b’. The resultant similarity matrix data was employed to construct a dendrogram based on unweighted pair-group method with an arithmetic average (UPGMA) to infer genetic relationships among cultivars. All computations were carried out using the NTSYS-pc, Version 2.2 package (Rohlf, 2005).

Results

DNA amplification and cultivar identification: A total of 40 microsatellite (SSR) markers covering all 12 chromosomes were utilized to characterize and assess genetic diversity among 41 basmati and non-basmati cultivars of rice. Among them, 30 markers which showed clear and consistent banding patterns and amplification of each cultivar were ultimately chosen for examining the genetic diversity and relationship among the

cultivars used (Table 2). Amplification profile revealed by one of the polymorphic markers (RM-202) across rice cultivars is depicted in Fig. 1. A considerable level of variability was observed among different cultivars. In most of the cases, basmati and other aromatic cultivars exhibited similar banding patterns. The microsatellites exhibited several bands that were shared among the basmati and other fine cultivars, whereas a few bands were shared among 'basmati' and 'non-basmati' cultivars of rice. The cultivar 'Jhona-349' displayed unique bands in comparison with all other non-aromatic cultivars. Interestingly, many primers revealed characteristic fragments in this cultivar which were not produced in any of the other non-aromatic cultivars used. Two *japonica* cultivars 'Kinmaze' and 'Nipponbare' had unique as well as shared fragments with the other aromatic rice cultivars from Pakistan. The study showed that most of the basmati cultivars genetically resembled each other and were found to be closer to *japonica* types as compared to non-basmati coarse cultivars.

Number of alleles: The level of polymorphism among 41 rice cultivars was evaluated by calculating allele number and PIC values for each of the 30 SSR loci evaluated. Each of the 30 loci differed significantly in their ability to determine variability among the cultivars. A total of 104 polymorphic alleles were detected across 41 cultivars using 30 SSR markers (Table 2). The number of alleles per locus generated by each marker varied from two (at loci including RM60, RM84, RM171, etc.) to six (RM72, RM161 and RM257) with an average of 3.5 alleles per locus. As could be seen from Table 2, no correlation was observed between the number of alleles detected and the number of SSR repeats in the SSR loci. For example, the microsatellite loci containing the (GA) repeat motifs varying from (GA)11 to (GA)31 did not show any correlation with the number of alleles they revealed. Maximum number of polymorphic alleles i.e., 6 was obtained with the marker RM72, RM161 and RM257, while the minimum number of 2 alleles per marker was amplified with RM60, RM84 etc. The average number of polymorphic alleles per locus was 3.5. An advanced breeding line 'Pak23717' gave the highest numbers of alleles (44). It was followed by genotypes 'DR83' and 'Pak23725' producing 43 alleles each, while cultivars 'Basmati-Pak', 'DR92' and 'Pakhal' gave the least number of alleles (i.e., 31 alleles each).

Rare alleles: An allele that was observed in only one or two of the 41 cultivars was considered rare. A total of 14 (13%) rare alleles were observed at 10 of the 30 SSR loci, with an average of 1.4 rare alleles per locus. Maximum number of rare alleles (3 alleles) was observed at RM161, followed by RM257 and RM302 loci (2 alleles each). In general, markers detecting a greater number of alleles per locus detected more rare alleles. Thirteen (32%) of the rice cultivars had unique SSR alleles. Maximum of 3 rare alleles were found in 'Kinmaze', six cultivars (Jhona-349, Sathra, DR82, DR92, Kanwal-95 and Nipponbare) had 2 rare alleles each, while other six of the cultivars (Pak23710, DR83, Pak23725, Swat-1, Pakhal and IR36) had single rare allele at individual locus.

PIC values: Polymorphism information content (PIC) value is the reflection of allele diversity and frequency among the varieties, and varied greatly for all the SSR loci tested. The PIC values varied widely among loci and ranged from 0.259 (RM182 on chromosome 7) to 0.782 (RM72 on chromosome 8) with an average of 0.571 per locus (Table 2). PIC values showed a significant positive correlation with number of alleles at SSR locus ($r = 071^{**}$; $p<0.01$, analysis of variance).

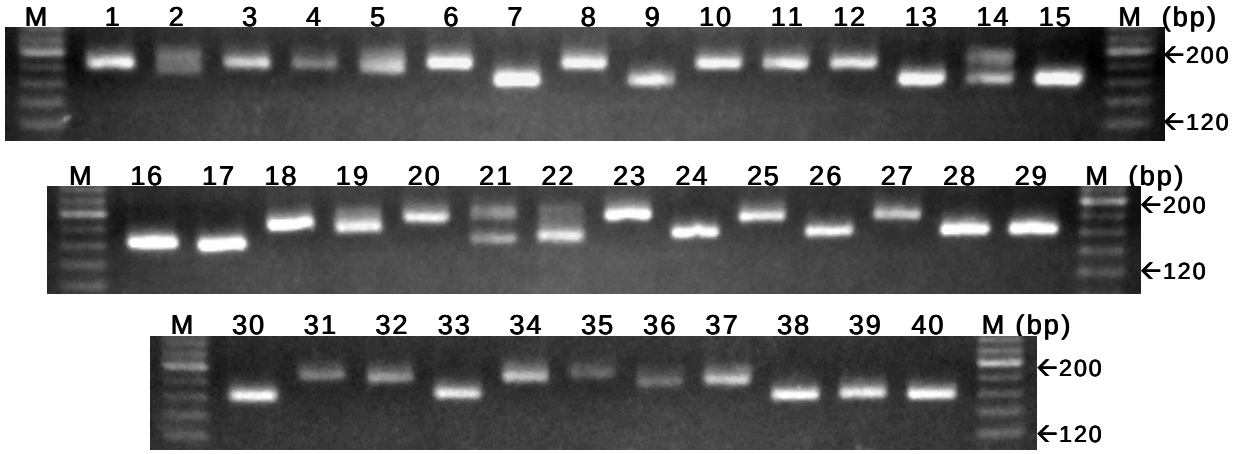

Fig. 1. SSR banding pattern of 40 basmati and non-basmati cultivars of Pakistani rice generated by primer pair RM202. The lanes represent M-20bp molecular size marker; 1) Basmati-370, 2) Jhona-349, 3) Sathra, 4) Palman-Sufaid, 5) Basmati-C622, 6) Basmati-Pak, 7) Pak23710, 8) KS-282, 9) Basmati-385, 10) Super-Basmati, 11) Basmati-2000, 12) Kashmir-Basmati, 13) NIAB-IR9, 14) Pak23717, 15) Rachna-Basmati, 16) IR6, 17) Sugdasi-Ratria, 18) Jajai-77, 19) Kangni-27, 20) DR-82, 21) DR-83, 22) Pak23725, 23) Sada-Hayat, 24) DR-92, 25) Kanwal-95, 26) Shua-92, 27) Khushboo-95, 28) Shadab, 29) Sarshar, 30) Swat-1, 31) Swat-2, 32) Pakhal, 33) Dilrosh-97, 34) Fakhre-Malakand, 35) IR36, 36) Kinmaze, 37) Nipponbare, 38) Sugdasi-Bengalo, 39) Sugdasi-Sadagulab, 40) Sonahri-Sugdasi.

Similarity matrix: A similarity matrix based on the proportion of shared SSR alleles was used to establish the level of relatedness between the cultivars surveyed. Pair-wise genetic similarity coefficients varied from 0.10 to 0.99, and the average similarity of 0.50. Two traditional aromatic cultivars 'Sugdasi-Sadagulab' and 'Sonahri-Sugdasi' were the closest genotypes having the highest similarity index of 99%. This was closely followed by 96% similarity between a pair of coarse cultivars 'NIAB-IR9' and 'Shadab'. The lowest level of genetic similarity (10%) was obtained between 'Super-Basmati' and 'Pakhal'. When basmati and other aromatic cultivars were compared with non-basmati coarse types, basmati types were found to be more similar to each other than to non-basmati group. Among the basmati and other aromatic cultivars, similarity coefficients ranged from 0.44 to 0.99, whereas non-basmati cultivars showed similarity coefficients of 0.19 to 0.96 among themselves. It indicated that the non-aromatic coarse cultivars were more diverse than the basmati group. The similarity coefficients of Japanese cultivars, 'Kinmaze' and 'Nipponbare' with aromatic cultivars ranged from 0.46 to 0.68, with non-aromatic coarse cultivars varied from 0.15 to 0.72, while both showed 0.88 similarity index with one another. Based on study the large range of similarity values for related cultivars using SSRs provides greater confidence for the assessment of genetic diversity and relationships.

Cluster analysis: The genetic relationship among the rice cultivars was assessed by a cluster analysis of the similarity matrix. A UPGMA cluster diagram grouped the 41 rice genotypes into two major clusters, I and II effectively differentiating the basmati and other aromatic and quality rice cultivars from the short and long bold non-aromatic coarse cultivars with additional sub-clusters (Fig. 2). Group-I consisted 15 long slender basmati and other aromatic and quality rice cultivars used in this study, 5 short bold non-

aromatic rice cultivars and 2 *japonica* cultivars, while group-II comprised of 19 short to long bold non-aromatic cultivars. Group I was more diverse and could be further divided into four subgroups; one comprising all the traditional and improved cultivars of basmati rice, 2nd containing Sugdasi (aromatic) rice and two coarse cultivars from Sindh, 3rd containing *japonica* rice varieties (Kinmaze, Nipponbare) and Swat-1, while 4th sub-group consisted of two traditional coarse cultivars from Punjab. Two *japonica* cultivars, 'Kinmaze' and 'Nipponbare' clustered into a separate sub-group, which were closer to the basmati cultivars than to the non-basmati *indica* group. Two traditional cultivar 'Jhona-349' and 'Sathra' formed an independent sub-cluster which showed significant difference from all other non-aromatic coarse cultivars. It might be due to very specific characteristics present in them but absent in remaining non-aromatic cultivars. Dendogram revealed that the cultivars that are derivatives of genetically similar type clustered more together. The cluster analysis also revealed that the traditional basmati varieties were closer to the *japonica* group than to the *indica* varieties. Cultivars in the same subgroup usually shared a high proportion of ancestry and/or agronomic characteristics such as height, maturity, quality traits etc.

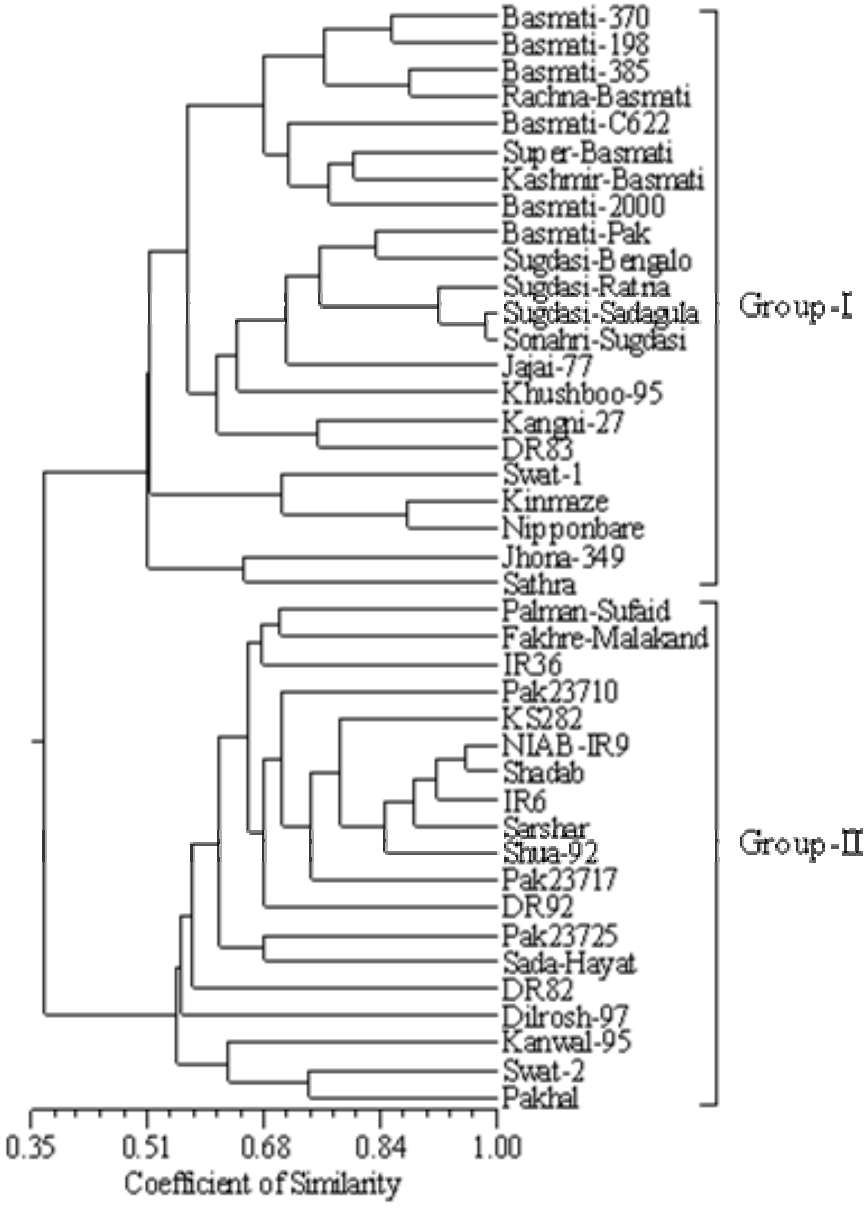

Fig. 2. UPGMA cluster analysis showing the diversity and relatedness among 41 basmati and non-basmati varieties of Pakistani rice using 104 alleles generated by 30 microsatellite markers.

Discussion

Agro-morphological and seed traits have long been the means of studying taxonomy and variability among populations and species. The study of genetic variation and structure has been greatly facilitated by the advent of DNA marker technology in the 1980s, which offered a large number of environmentally-insensitive genetic markers that could be generated to follow the inheritance of important agronomic traits. Microsatellites or simple sequence repeats (SSRs) are among the most commonly used DNA marker types for a wide range of purposes e.g., diversity, genome mapping, varietal identification. The first report of microsatellites in plants was made by Condit & Hubbel (1991) who suggested that they are abundant in plant systems. Later, Akkaya *et al*., (1992) reported length polymorphisms of SSRs in soybean, which opened up a new source of PCR-based molecular markers for other plant genomes. At present, SSRs are the most preferred marker types because they are highly polymorphic even between closely related lines, require low amounts of DNA, can be easily automated and allow high-throughput screening, can be exchanged between laboratories and are highly transferable between populations.

Although the use of molecular markers to study the genetic diversity and relationships among the different cultivars has been previously reported (Nagaraju *et al*., 2002; Singh *et al*., 2004), the relationship between Pakistani cultivars is very limited. In the present investigation, 30 SSR markers were used to assess the genetic diversity of 41 traditional varieties and improved cultivars of rice. This study was the first step for the characterization of the molecular diversity of Pakistani rice developed since early nineteen. The results indicated a considerable level of genetic variation in the cultivars used. Microsatellites produced a number of alleles that were shared among the basmati and other fine cultivars, whereas comparatively lower number of bands was common among basmati and non-basmati coarse cultivars of rice. Cultivar 'Jhona-349' shared limited number of fragments with all the other coarse cultivars, presenting its distant association with non-aromatic group. The microsatellite assay generated variety-specific alleles in some of the genotypes screened; these may be used as DNA fingerprints for variety identification. This would be of enormous assistance for the establishment of proprietary rights and the determination of cultivar purity.

The results from the genetic similarity-based analysis of 41 rice varieties with 30 microsatellite markers revealed a clear divide of Pakistani cultivars into aromatic and non-aromatic groups. The microsatellite data from the present study demonstrated the lack of genetic difference between 'Sugdasi-Sadagulab' and 'Sonehri-Sugdasi', and 'NIAB-IR9' and 'Shadab' which could be due to either extremely low difference between the two varieties at the DNA level and/or admixture of seed during multiplication. These two pairs of varieties were also very similar in morphological and agronomic traits that strengthen the former supposition rather than the latter. There was, however, sufficient amount of molecular variation among all other aromatic as well as non-aromatic varieties.

The number of alleles detected by microsatellite markers varied from 2 to 6 with an average of 3.5 alleles per locus. The number of alleles observed in the present study corresponded well with the earlier report of Siwach *et al*., (2004) among basmati and non-basmati rice varieties from India. The number of alleles detected in the present study was comparatively higher than the average number of alleles reported by Nagaraju *et al*., (2002) and Singh *et al*., (2004) in aromatic rice of India. This may be due to inclusion of both

indica and *japonica* varieties and inclusion of basmati as well as non-basmati cultivars in the present study. The proportion of alleles noticed in present study was relatively lower than that reported earlier by Jain *et al*., (2004), Brondani *et al*., (2006), Jayamani *et al*., (2007), and Thomson *et al*., (2007) who observed an average of 7.8, 14.6, 7.7 and 13 alleles locus^{-1} using Indian rice varieties, traditional varieties of Brazilian rice, a diverse collection of rice in Portugal and Indonesian rice germplasm, respectively. This discrepancy might be related to the genotypes used and selection of SSR primers with scorable alleles. The other reason could be the use of more diverse genotypes by these authors. SSR markers involving di-nucleotide repeats motifs particularly those with GA repeats amplified relatively more number of bands as reported earlier in rice (Saini *et al*., 2004).

The level of polymorphism, as assessed by PIC values, was quite high and varied considerably among SSR loci (0.259 to 0.782, average value 0.571). The 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). It was higher than the earlier observations of Singh *et al*., (2004), but lower than that previously reported by Brondani *et al*., (2006) who observed an average PIC value of 0.74 for traditional varieties of Brazilian rice. They included more diverse set of rice germplasm.

Similarity coefficients among various cultivars ranged from 0.10 to 0.99 in present investigations with an average of 0.50. Similar values of similarity coefficients were obtained among 18 basmati and non-basmati varieties using molecular markers (Saini *et al*., 2004). 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. Siwach *et al*., (2004) observed 0.67 to 0.91 genetic similarity among basmati and non-basmati long-grain *indica* rice varieties using microsatellite markers. Jain *et al*., (2004) obtained relatively higher level of similarity ranging from 79 to 99.6% with an average of 0.89 between Indian basmati/aromatic germplasm using panels of fluorescently-labeled microsatellite markers. One of the reasons for this high level of similarity recorded by these studies could be that the intra-specific variation in the germplasm used is narrow due to use of same ancestors and selection of similar traits as compared with our rice genotypes used.

Cluster analysis based on similarity coefficients classified 41 rice genotypes into 2 major groups. Basmati and other aromatic/scented cultivars fell into close sub-groups generally corresponding to pedigree, grain type, breeding institute etc., For example, all 15 aromatic cultivars grouped in the upper portion of the dendrogram which is consistent with grain quality (aromatic type) as a trait that differentiated cultivars into broad categories. Cluster analysis grouped most of the basmati cultivars from Punjab like ‘Basmati-370’, ‘Basmati-Pak’, ‘Basmati-385’, ‘Basmati-2000’ 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. Similarly aromatic rice from Sindh such as ‘Sonahri-Sugdasi’, ‘Sugdasi-Sadagulab’, ‘Sugdasi-Ratria’, ‘Sugdasi-Bengalo’ and ‘Jajai-77’ grouped together which is rational since all are local selections from Sugdasi rice, belong to Rice Research Institute, Dokri, Sindh and possibly have similar ancestors. In the similar type of study 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 rices fell

into different groups. These authors also reported that traditional and evolved basmati varieties shared a high degree of similarity using SSR markers.

Distinct distribution of long slender aromatic from short bold rice corresponded well with their grain characteristics as well as geographic distribution. The long slender and medium slender aromatic rice is largely grown in Kallar Tract and central parts of Punjab like Sialkot, Gujranwala, Gujrat, Hafizabad, Sheikhupura and Lahore and possess most desirable grain characteristics and cooking qualities which depend upon the combined effect of several physicochemical properties. However, the short bold rice cultivars are grown mostly in Sindh, southern parts of NWFP and northern cooler areas of Pakistan. These cultivars, although possess different kinds of aroma with varying strength, lack the more desirable basmati traits. This suggests a probability that the widely accepted long slender basmati types may have developed by natural mutation of some indigenous non-basmati types. 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 later being closer to *japonica* than to *indica*. The study shows that basmati rice varieties are genetically distinct from other groups within *O. sativa,* viz., *indica* and *japonica*. Higher levels of genetic diversity between basmati and non-basmati support the concept that former had a long history of independent evolution and diverged from non-basmati rice a long time ago through conscious selection and patronage (Nagaraju *et al*., 2002; Jain *et al*., 2004; Saini *et al*., 2004). 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, Glaszmann (1987) also reported that basmati genotypes are genetically distinct from the other groups. Although all of the Pakistani cultivars could be assigned as belonging to the either basmati or non-basmati *indica* groups, a few atypical cultivars were detected. Two traditional non-basmati coarse varieties viz., Jhona-349' and 'Sathra grouped separately from the main groups and fell between the aromatic and coarse groups. The unusual nature of these traditional cultivars may be due to admixture of seeds during post-harvest handling.

Conclusions

From these investigations it is concluded that the basmati rice cultivars in Pakistan had relatively narrow genetic base as compared to non-basmati coarse types. Lower level of polymorphism in basmati cultivars indicated that there is a basic similarity among the basmati genotypes used in this study, which is to be expected due to their same ancestors and selection for similar characteristics. Higher levels of genetic diversity between basmati and non-basmati support the concept that former had a long history of independent evolution and diverged from non-basmati rice a long time ago through conscious selection and patronage. Present investigation further indicated that genetically basmati rice is different from that of coarse *indica* and *japonica* type. It is suggested that microsatellite analysis can be efficiently utilized for diversity analysis and differentiation of basmati and non-basmati coarse rice cultivars. These markers can be effectively exploited for testing purity of the commercial seed lots and maintenance of seed purity at different levels of seed production of aromatic rice. In addition, marker-based identification and differentiation of traditional Basmati rice may help to maintain the integrity of this high quality product to the benefit of both farmers and consumers. These

studies may be followed by the recent advances in genetic engineering (Nakashima *et al*., 2000 & Narusaka *et al*., 2003).

Acknowledgments

We are indebted to various Rice Research Institutes/Programs of Pakistan for providing the seed material of their basmati and non-basmati 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 work was carried out under Agricultural Linkages Program. Authors are grateful to Pakistan Agricultural Research Council, Islamabad for the financial support for this work from the Agricultural Linkages Program under the AREF. We also gratefully acknowledge the financial support from ICRISAT under Generation Challenge Program and Japanese Society of Promotion of Science, Japan under Short-term JSPS Fellowship Program. We are grateful to Prof. Dr. S. I. Ali of KIBGE-Pakistan for going through earlier draft of the manuscript.

References

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan. 1992. Length polymorphism of simple sequence repeat DNA in soybean. *Genetics,* 132: 1131-1139.

Anonymous. 2006. *Agricultural Statistics of Pakistan*. Ministry of Food, Agriculture & Cooperative (MINFAL), Government of Pakistan.

Bligh, H.F.J. 2000. Detection of adulteration of Basmati rice with non-premium long grain rice. *Int. J. Food Sci. Tech*., 35: 257-265.

Brondani, C., T.C.O. Borba, P.H.N. Rangel and R.P.V. Brondani. 2006. Determination of genetic variability of traditional varieties of Brazilian rice using microsatellite markers. *Genet. Mol. Biol*., 29: 676-684.

Chakravarthi, B.K. and R. Naravaneni. 2006. SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L.). *Afr. J. Biotech*., 5: 684-688.

Coburn, J.R., S.V. Temnykh, E.M. Paul and S.R. McCouch. 2002. Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.). *Crop Sci*., 42: 2092-2099.

Condit, R. and S.P. Hubbell. 1991. Abundance and DNA sequence of two base repeat regions in tropical tree genomes. *Genome*, 34: 66-71.

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

Glaszmann, J.C. 1987. Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet*., 74: 21-30.

Jain, S., R.K. Jain and S.R. McCouch. 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., S. Negrão, M. Martins, B. Maçãs and M.M. Oliveira. 2007. Genetic relatedness of Portuguese rice accessions from diverse origins as assessed by microsatellite markers. *Crop Sci*., 47: 879-886.

Kang, H.W., Y.G. Cho, U.H. Yoon and M.Y. Eun. 1998. A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. *Plant Mol. Biol. Reporter*, 16: 1-9.

McCouch, S., L. Teytelman, Y. Xu, K. Lobos, K. Clare, M. Walton, B. Fu, R. Maghirang, Z. Li, Y. Xing, Q. Zhang, I. Kono, M. Yano, R. Fjellstrom, G. DeClerck, D. Schneider, S. Cartinhour, D. Ware and L. Stein. 2002. Development of 2,240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res*., 9: 199-207.

McCouch, S.R., X. Chen, O. Panaud, S. Temnykh, Y. Xu, Y.G. Cho, N. Huang, T. Ishii and M. Blair. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol. Biol*., 35: 89-99.

Nagaraju, J., M. Kathirvel, R.R. Kumar, E.A. Siddiq and S.E. Hasnain. 2002. Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. *Proc. Nat. Acad. Sci. USA*, 99: 5836-5841.

Nakashima, K., Z.K. Shinwari, S. Miura, Y. Sakuma, M. Seki, K. Yamaguchi-Shinozaki and K. Shinozaki. 2000 Structural organization, expression and promoter activity of an Arabidopsis gene family encoding DRE/CRT binding proteins involved in dehydration- and high salinity-responsive gene expression. *Plant Mol. Biol.*, 42: 657-665.

Narusaka, Y., K. Nakashima, Z.K Shinwari, Y. Sakuma, T. Furihata, H. Abe, M. Narusaka and K. Shinozaki and KY Shinozaki 2003. Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis* rd29A gene in response to dehydration and high salinity stresses. *The Plant J.,* 34: 137-149.

Nei, M. and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Nat. Acad. Sci. USA*, 76: 5269-5273.

Rabbani, M.A., Z.H. Pervaiz and M.S. Masood. 2008. Genetic diversity analysis of traditional and improved cultivars of Pakistani rice (*Oryza sativa* L.) using RAPD markers. *Electronic J. Biotech*., 11: 1-10.

Ravi, M., S. Geethanjali, F. Sameeyafarheen and M. Maheswaran. 2003. Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. *Euphytica*, 133: 243-252.

Rohlf, F.J. 2005. NTSYS-pc: *Numerical Taxonomy and Multivariate Analysis System*, version 2.2. Exeter Software, Applied Biostatistics Inc., New York, USA.

Saini, N., N. Jain, S. Jain and R.K. Jain. 2004. Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. *Euphytica*, 140: 133-146.

Shinwari, Z.K. 1995 Congruence between morphology and molecular phylogeneties in Prosartes (Liliaceae). *Pak. J. Bot*., 27: 361-369.

Singh, R.K., R.K. Sharma, A.K. Singh, V.P. Singh, N.K. Singh, S.P. Tiwari and T. Mohapatra. 2004. Suitability of mapped sequence tagged microatellite site markers for establishing distinctness, uniformaity and stability in aromatic rice. *Euphytica*, 135: 135-143.

Siwach, P., S. Jain, N. Sain, V.K. Chowdhury and R.K. Jain. 2004. Allelic diversity among Basmati and non-Basmati long-grain indica rice varieties using microsatellite markers. *J. Plant Bioch. Biotech*., 13: 25-32.

Temnykh, S., W.D. Park, N. Ayres, S. Cartinhour, N. Hauck, L. Lipovich, Y.G. Cho, T. Ishii and S.R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet*., 100: 697-712.

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

(Received for publication 23 June 2009)