

## Allelic Diversity Among Basmati and Non-Basmati Long-grain *Indica* Rice Varieties using Microsatellite Markers

Priyanka Siwach<sup>1</sup>, Sunita Jain<sup>2</sup>, Navinder Saini<sup>1</sup>, Vijay K Chowdhury<sup>1</sup> and Rajinder K Jain<sup>1\*</sup>

<sup>1</sup>Department of Biotechnology and Molecular Biology, CCS Haryana Agricultural University, Hisar 125 004, India

<sup>2</sup>Department of Biochemistry, CCS Haryana Agricultural University, Hisar 125 004, India

Microsatellite or simple sequence repeat (SSR) marker analysis was carried out to assess allelic diversity and prepare a DNA fingerprint database of 24 rice genotypes including three premium traditional Basmati, 9 cross-bred Basmati, a local scented selection, eight *indica* and three *japonica* rice varieties. A total of 229 alleles were detected at the 50 SSR loci and 49 alleles were in fact present in only one of the 24 varieties. The size difference between the smallest and largest allele varied from 1 (RM333) to as high as 82 (RM206). Multiple alleles were observed at 13 loci. Polymorphism information content (PIC) values ranged between 0.0 (RM167) to 0.78 (RM170), with an average of 0.62 per marker. At 15 of the SSR loci, traditional and cross-bred Basmati rice varieties amplified different alleles than those in the *indica* and/or *japonica* rice varieties. A number of SSRs have been identified, which can be used to differentiate among the traditional Basmati varieties and between traditional Basmati and other cross-bred Basmati or long grain, non-Basmati rice varieties. Genetic relationships among rice genotypes as determined by UPGMA cluster analysis and three-dimensional scaling based on Principal Component Analysis showed that the three traditional Basmati rice varieties are closely related and have varying degree of similarity with other cross-bred Basmati rice varieties. Further implications of these results in genotype identification, monitoring purity and adulteration, and plant variety protection are discussed.

**Key words:** Basmati rice, DNA fingerprinting, genetic diversity, microsatellite markers, *Oryza sativa*.

There are thousands of rice varieties and landraces, which differ with respect to plant and grain characteristics. Of these, Basmati rices constitute a small but special group, with superior aroma and grain quality (1-3). Basmati rice occupies a prime position in the Indian subcontinent and is becoming increasingly popular in Middle East, Europe, USA and even in non-traditional rice-growing countries such as Australia (4). High-quality, traditional Basmati rice varieties command premium prices, more than three times that of non-Basmati rices in the Indian as well as world market (4). Basmati rice has its origin in the foot-hills of the Himalayas (Indian states of Uttaranchal, Uttar Pradesh, Bihar, etc.) and spread to other states and neighbouring countries (5). Basmati rice varieties are genetically differentiated from *indica* (varietal group I) and *japonica* (varietal group VI) rice varieties on the basis of isozyme polymorphism data and clustered in a distinct isozyme group (varietal group V) (6). Hundreds of locally adapted genotypes of Basmati rice are grown in the Indian subcontinent, which have evolved as a result of both natural and human selection. Greater appreciation of genetic diversity contained in Basmati gene pool is necessary for

classification, proper maintenance, conservation and effective utilization of this traditional rice germplasm.

Breeding efforts have been made to develop short-duration, short-statured and high-yielding varieties with intact Basmati grain-quality traits but the progress has been slow (1, 3). The genetics of each of the Basmati grain quality components (intermediate amylose content, intermediate gel consistency, gelatinization temperature and kernel elongation) except aroma is quite complex and it has been difficult to find a desirable segregant possessing all these components alongwith high yield potential (2). A few semi-dwarf high yielding Basmati rice varieties such as Pusa Basmati 1 and Super Basmati have been developed from *indica* x Basmati rice crosses but traditional Basmati rice varieties still have edge in quality, which is reflected in price differences between them (3). Recently, molecular marker analysis using PCR-based markers has been widely used for varietal identification and to identify cases or even level of adulteration in Basmati rice supplies (7, 8). This is especially important in the case of Basmati rice to ensure the export quality, for maintaining the 'distinctiveness' of Basmati varieties, and to differentiate between the various grades of Basmati rice (8, 9).

\*Corresponding author. E-mail: rkjain@hau.nic.in

**Table 1.** A brief description of rice varieties used for SSR analysis

Rice types	Varieties	Abbreviation	Source, Accession no.	Remarks
Traditional Basmati	Basmati 370	Bas370	HAU-Kaul, 1176/838	Premium traditional Basmati rice variety
	HBC19	HBC19	HAU-Kaul, 1882/992	Premium traditional Basmati rice variety, a pure line selection from Taraori Basmati
	Taraori Basmati	Taraori	HAU-Kaul	Premium traditional Basmati rice variety
Local scented	Dehraduni Basmati 284	DB284	DRR, 1320/138	An aromatic rice-selection from Taraoregion
Cross-bred Basmati	Pusa Basmati 1	PB1	HAU-Kaul	Selection from Pusa-167*/Karnal local Basmati
	HKR228	HKR228	HAU-Kaul	Selection Sona*/Basmati 370
	Improved Sabarmati	ImpSabar	GBPU	Selection from T(N)-1*/Basmati 370
	CSR30	CSR30	CSSRI	Salt tolerant, obtained from BR4-10*/ Pakistan Basmati
	Kasturi	Kasturi	GBPU	Selection from CR-88-17-1-5*/Basmati 370
	Sabarmati	Sabar	GBPU	Selection from T (N) 1*/Basmati 370
	Super Basmati	Super	IRRI	Selection from IR-661*/Basmati 370
	Texmati	Texmati	USA	Selection from American dwarf <i>indica</i> rice variety/ Pakistani Basmati; RiceTec, USA
	Jasmati	Jasmati	USA	Selection from crosses between Indian and Thailand aromatic rice varieties
<i>Indica</i>	HKR120	HKR120	HAU-Kaul, 1870/975	High yielding <i>indica</i> variety
	IR64	IR64	IRRI	<i>Indica</i> rice variety developed at IRRI
	IR36	IR36	IRRI	--do--
	IR24	IR24	IRRI	--do--
	CSR10	CSR10	CSSRI	Salt tolerant, cross-bred, <i>indica</i> rice variety
	Pokkali	Pokk	CSSRI	Salt tolerant landrace ( <i>indica</i> ) from South India (Kerala)
	Sharbati	Sharbati	GBPU	A local selection from Uttar Pradesh
<i>Japonica</i>	New Plant Type II	NPTII	IRRI, IR68552-100-1-2-2	Tropical <i>japonica</i> rice accession developed at IRRI
	Nipponbare Azucena	Nippon Azu	IARI CCS HAU, 1014/141	<i>Japonica</i> rice variety Aromatic, temperate <i>japonica</i> rice variety from Philippines

\**Indica*; DRR, Directorate of Rice Research, Rajender Nagar, Hyderabad - 500030, India; HAU Kaul, CCS Haryana Agricultural University Rice Research Station, Kaul - 132021, India; CSSRI, Central Soil Salinity Research Institute, Karnal 132 001, India; IRRI, International Rice Research Institute, Los Banos, Philippines; GBPU, G.B. Pant University of Agriculture and Technology, Pantnagar 261 145, Uttaranchal, India; IARI, Indian Agricultural Research Institute, New Delhi, India.

Microsatellites, also known as simple sequence repeats (SSRs), are simple tandemly repeated di- to tetra-NT sequence motifs flanked by unique sequences (10). These markers are often polymorphic and have already been demonstrated to be a powerful tool in genotype identification and variety protection, seed-purity evaluation and germplasm characterization, diversity studies, gene and quantitative trait locus (QTL) analysis, pedigree analysis and marker assisted breeding (7, 8, 10-12). The genetic profile produced by using SSRs can be used in conjunction with pedigrees and agronomic data to help document ownership and protect intellectual property rights. This is important especially for protection of proprietary germplasm.

In this paper, we report the development of a DNA fingerprint database of commercially important Basmati and non-Basmati rice varieties for 50 microsatellite DNA

markers (SSRs). The data have been used to evaluate levels of genetic diversity within Basmati group and to assess genetic relationships among various rice types.

## Materials and Methods

**Plant material** — A total of 24 rice genotypes including three premium traditional Basmati (Taraori, HBC19 and Basmati 370), 9 cross-bred Basmati (Super Basmati, Texmati, CSR30, HKR228, Jasmati, Pusa Basmati 1, Sabarmati, Improved Sabarmati, Kasturi), a local, scented selection (Dehraduni Basmati 284), eight *indica* (HKR120, CSR10, Calrose76, Sharbati, IR36, IR24, IR64, and Pokkali) and three *japonica* (Nipponbare, Azucena, New Plant Type-II) rice varieties were evaluated; the characteristics, origin, source and accession numbers of these genotypes are given in Table 1. Rice varieties, Azucena, IR36, and Nipponbare, also served as controls for determining allele

molecular weight because they had previously been assayed at the same SSR loci (13, 14). Nurseries of all these genotypes were raised in a net house at CCS Haryana Agricultural University, Hisar, India.

**Selection of primers** — Fifty microsatellite primer pairs covering all the twelve chromosomes were selected for the genetic diversity analysis on the basis of the published rice microsatellite framework map. The original source, repeat motifs, primer sequences and chromosomal position for these markers can be found in Temnykh *et al* (14) and in the RiceGenes database ([http://www.gramene.org/microsat/RM\\_primers.html](http://www.gramene.org/microsat/RM_primers.html)). Microsatellite primer pairs were obtained from Research Genetics, Inc. (Huntsville, AL, USA).

**Microsatellite marker analysis** — Genomic DNA from most of the rice genotypes (except Texmati and Jasmati) was isolated from bulked leaf samples (~20 mg each) from five plants using CTAB method of Saghai Maroof *et al* (15). DNA from Texmati and Jasmati was isolated from milled rice samples using a modified SDS mini-prep method (16). SSR amplifications were performed with PTC-100 TM 96V thermocycler (MJ Research Inc., Watertown, MA, USA) and Taq ampligold DNA polymerase (Promega Inc., Madison, WI, USA). The PCR reaction was conducted in a reaction volume of 20 µl containing: 1x PCR buffer, 200 µM dNTPs, 0.4 µM of each primer, 1.2 mM MgCl<sub>2</sub>, 1 unit Taq DNA polymerase and 20 ng template DNA. The PCR amplification was performed with a hot start of 94 °C for 5 min and then 35 cycles of 1 min denaturation at 94 °C, 1 min annealing at 55 °C and 2 min extension at 72 °C, and 5 min at 72 °C for the final product extension. Amplification products were stored at -20 °C till further use.

Amplification products were denatured at 95 °C for 2 min and resolved on 4% denatured polyacrylamide gel, using Aluminum Backed sequencing System Model # 535 (Owl Scientific Inc., Woburn, MA, USA) essentially as described by Chen *et al* (13). Electrophoresis was performed at constant power of 100 Watt for 3.5 h including a 1 h pre-run to warm the gel to 50-60 °C. Following electrophoresis, DNA bands were visualized using the Silver Sequence™ DNA Sequencing System (Promega Technical Manual No. 023).

**Allele scoring** — Silver staining of the polyacrylamide gels generally revealed a cluster of 2 to 5 discrete bands (stutters) or blurred stutters for most markers. The size (in nucleotide base pairs) of the most intensely amplified band for each microsatellite marker was determined based on

its migration relative to a molecular-weight size marker (10 bp DNA ladder from Gibco BRL, Gaithersburg, Md.). IR-36 (*indica*) and Azucena (*japonica*) were also used as molecular weight references because sequence based estimates of allele size for these varieties are available (14, [http://www.gramene.org/microsat/RM\\_primers.html](http://www.gramene.org/microsat/RM_primers.html)).

**Diversity analysis** — The frequency of microsatellite polymorphism was calculated based on presence or absence of common bands (17). The polymorphism information content (PIC) value described by Anderson *et al* (18) was calculated as follows:

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

Where, P<sub>ij</sub> is the frequency of the j<sup>th</sup> allele for i<sup>th</sup> marker and summation extends over n alleles.

The genetic associations among varieties were evaluated by calculating the Jaccard similarity coefficient for pair-wise comparisons based on the proportion of shared bands (alleles) produced by primers. Similarity matrices were generated using 'Simqual' sub-program of software NTSYS-PC (19). Similarity coefficients were used for cluster analysis of varieties performed using the 'Sahn' sub-program of NTSYS-PC and dendograms were built by the un-weighted pair-group method with arithmetic average (UPGMA) subprogram of NTSYS-PC.

## Results

A microsatellite fingerprint database (available at <http://hau.nic.in/basmati.htm>), has been generated with 50 SSR markers for 24 rice genotypes, and used for diversity analysis and determination of genetic relationships. Salient features of the microsatellite marker based allelic diversity data of 24 rice varieties are given below.

**Number of alleles** — A total of 229 alleles for 50 SSR loci were detected among 24 rice varieties (Table 2). The number of alleles per locus ranged between 1 (RM167) and 8 (RM204 and RM234) with an average of 4.58 alleles per locus. Loci with tri-nucleotide repeat motifs were somewhat more polymorphic (mean number of alleles 5.2, n=5) than those with di-nucleotide repeat motifs (mean number of alleles 4.5, n=22) or tetra nucleotide repeat motifs (RM60 with 3 alleles).

**Allele size range** — The overall size of PCR products amplified using 50 SSR primer pairs ranged from 72 to 260 bp (Table 2). The molecular size difference between the

**Table 2.** Data on the number of alleles, number of rare alleles, number of genotypes with multiple alleles, allele size range, highest frequency allele and polymorphism information content (PIC) found among 24 rice genotypes for 50 microsatellite markers

Marker	Chromosome location	No. of alleles	No. of rare alleles	No. of genotypes with multiple alleles	Size range (bp)	Difference (bp)	Highest frequency allele		PIC Value
							Size (bp)	Frequency (%)	
RM1	1	5	4	0	77-113	36	111	54.2	0.629
RM2	7	4	0	0	147-153	6	149,151	29.2	0.743
RM3	6	4	2	0	118-146	28	146	54.2	0.592
RM5	1	5	2	0	110-118	8	114	54.2	0.612
RM13	5	5	3	0	131-147	16	131	45.8	0.612
RM16	3	3	1	0	168-186	16	168,184	45.8	0.573
RM17	12	3	0	1	161-186	25	186	33.3	0.588
RM18	7	5	1	0	151-163	12	157	37.5	0.743
RM21	11	6	3	0	133-164	3	135	41.7	0.743
RM22	3	4	1	0	187-197	10	193	62.5	0.681
RM25	8	3	1	2	131-146	15	146	75.0	0.462
RM26	5	3	1	2	121-126	5	123	37.5	0.368
RM31	5	4	1	0	141-153	12	147	62.5	0.698
RM38	8	3	0	0	245-260	15	250	58.3	0.538
RM44	8	5	2	1	103-131	28	103	45.8	0.585
RM53	2	7	5	0	92-220	28	198	75.0	0.719
RM60	3	3	1	0	169-173	4	169	50.0	0.392
RM70	7	7	4	1	130-170	40	159	41.7	0.697
RM84	1	4	1	0	111-130	19	114	62.5	0.663
RM122	5	2	0	0	229-232	3	229	100.0	0.469
RM167	11	1	0	0	126	-	126	29.2	0.0
RM170	6	6	2	0	99-119	20	105,116	45.8	0.778
RM174	2	4	1	0	207-222	15	222	29.2	0.649
RM204	6	8	4	1	146-174	28	168,174	41.7	0.772
RM206	11	5	2	0	130-212	82	140	45.8	0.717
RM207	2	7	5	0	123-142	19	129	29.2	0.691
RM209	11	3	0	0	127-151	24	135	66.7	0.497
RM210	8	6	3	1	136-15	19	143	37.5	0.717
RM212	1	4	0	0	112-1345	22	116	45.8	0.684
RM215	9	4	0	1	147-151	4	151	37.5	0.693
RM222	10	6	1	0	199-225	26	215	33.3	0.774
RM223	8	5	3	1	145-167	22	149	45.8	0.606
RM224	11	7	5	0	121-156	35	134	45.8	0.691
RM225	6	4	1	0	123-145	22	123	50.0	0.642
RM226	1	7	4	0	195-256	61	220	41.7	0.736
RM229	11	7	4	0	106-137	3	117	33.3	0.771
RM231	3	4	2	3	169-191	22	185	79.2	0.321
RM234	7	8	6	1	133-163	30	139,154	37.5	0.703
RM235	12	6	3	0	90-132	42	100	45.8	0.692
RM240	2	3	0	0	113-135	22	113	50.0	0.622

Contd...

Table 2. contd...

RM248	7	6	4	0	72-96	24	98	54.2	0.615
RM253	6	6	3	0	117-146	29	142	45.8	0.719
RM255	4	3	1	0	145-151	6	145	62.5	0.517
RM258	10	4	0	0	133-150	17	138,150	33.3	0.722
RM261	4	2	0	0	124-126	2	124	54.2	0.497
RM282	3	3	0	0	127-135	8	127	45.8	0.622
RM286	11	4	1	1	98-125	27	112	45.8	0.599
RM323	1	2	0	0	241-242	1	241	54.2	0.497
RM333	10	7	4	1	163-194	31	163	41.7	0.742
RM337	8	2	0	0	156-186	30	156	66.7	0.445
Total		229	93						
Mean		4.58	1.86	0.34				49.3	0.617

smallest and the largest allele for a given SSR locus varied from one (RM323 with CAT repeat motif) to 82 bp (RM206 with GA repeat motif). Maximum variation in allele size was observed for RM markers with GA repeat motifs (RM206, 82 bp; RM1, 36 bp; RM224, 34 bp; RM234, 30 bp; RM235, 42 bp), AT motifs (RM226, 61 bp), ATT motifs (RM70, 40 bp), and CTT motifs (RM333, 31 bp; RM337, 30 bp).

**Rare alleles** — An allele that was observed in only one or two of the 24 genotypes was considered rare. A total of 93 rare alleles were observed at 36 of the 50 SSR loci (Table 2). Of these, 49 were in fact present in only one variety. Fifteen rice genotypes (Azucena, Nipponbare, NPTII, IR24, IR36, IR64, HKR120, Pokkali, CSR10, Calrose 76, Dehradun Basmati 284, Pusa Basmati 1, Super Basmati, Sabarmati, CSR30) displayed one or more of such rare alleles. Most of these rare alleles were present in three *japonica* rice varieties, i.e. NPT-II (10), Nipponbare (9 alleles) and Azucena (6); which can be attributed to the low number of *japonica* rice varieties used in this study. Higher number rare alleles were observed at RM234 locus (6 alleles), followed by RM53, RM207, RM224 loci (5 alleles each). Notably, two SSR markers, RM3 (chromosome 6, 134 bp) and RM210 (chr. 8, 136 bp), amplified specific alleles only in the two salt-tolerant varieties, Pokkali and CSR10. Of the three traditional Basmati varieties (Taraori, HBC19 and Basmati 370), only Basmati 370 had a unique allele (135 bp) at RM207 locus.

**Multiple and null alleles** — Multiple alleles were inferred whenever a given marker for an individual genotype produced two cluster of bands. Multiple alleles, were observed at 13 of the 50 loci with an average frequency of 0.34 per locus (Table 2). Varieties showing multiple alleles included the Improved Sabarmati, CSR30, NPT-II, Sharbati,

Azucena, IR64, Kasturi and Pusa Basmati I. No null allele (absence of amplification product at a particular locus) was observed for any of the SSR marker in 24 rice genotypes.

**High frequency alleles** — On average, 49.3% of the varieties shared a common allele at any given locus. There was considerable range in allele frequency, as illustrated by the fact that at RM2, RM170 and RM204 loci, the highest frequency allele was present in only 29.2% of the genotypes while RM167 was monomorphic.

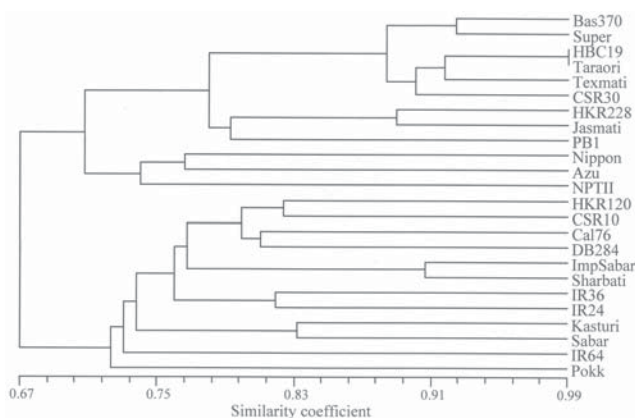
**PIC values** — The level of polymorphism among the 24 varieties was evaluated by calculating PIC values for each of the 50 SSR loci (Table 2). PIC values varied widely among loci and ranged from 0.0 (RM 167) to 0.8 (RM 170), with an average value of 0.62 per locus. PIC values showed a positive correlation of 0.5 with the number of alleles ( $P < 0.01$ , analysis of variance), of 0.1 with allele size range ( $P < 0.02$ ) and of 0.17 with the number of unique alleles ( $P < 0.01$ ) at a SSR locus.

**Differentiation between Basmati and non-Basmati rice varieties** — The SSR marker analysis revealed ample difference between the Basmati and non-Basmati rice varieties. Of the 50 SSR markers, 15 (RM1, RM21, RM22, RM38, RM44, RM60, RM70, RRM84, RM170, RM207, RM210, RM226, RM229, RM253 and RM333) amplified different alleles in traditional and cross-bred Basmati rice varieties than in the *indica* and/or *japonica* rice varieties. For seven of these markers (RM1, RM21, RM38, RM70, RM210, RM226 and RM229), the three traditional Basmati varieties (Basmati 370, Taraori, HBC19) shared a common allele.

Eight markers (RM215, RM207, RM44, RM222, RM22, RM2, RM3 and RM60) were identified which could

distinguish between Basmati 370, Taraori Basmati and HBC19 (a pure line selection from Taraori). As expected, Taraori Basmati and HBC19 had a very high similarity coefficient of 0.99; the two accessions showed polymorphism at only one (RM2) of 50 SSR loci. Three of the nine cross-bred varieties (Sabarmati, Kasturi and improved Sabarmati) and DB284 (a local scented collection) shared more alleles with *indica* varieties than with the traditional Basmati varieties. The remaining six cross-bred Basmati varieties shared more alleles with traditional Basmati types.

**Genetic relationships between the Basmati and non-Basmati varieties** — The UPGMA cluster tree analysis led to the grouping of 24 varieties in two major groups, which were further divided into several small groups (Fig. 1). Group I is comprised of *indica* varieties (HKR120, CSR10, Sharbati, IR24, IR36, IR64, Pokkali), some of the



**Fig. 1.** Dendrogram of 24 rice genotypes based on UPGMA cluster analysis using allelic diversity data for 50 SSR markers (229 alleles)

cross-bred Basmati varieties (Kasturi derived from crosses and backcrosses Bas370 x CR-88-17-1-5, Sabarmati and Improved Sabarmati derived from TN1 x Bas370), aromatic rice variety Calrose 76 from United States and a local scented rice Dehradun Basmati 284 collected from Uttar Pradesh. Group II could be further divided into two sub-groups; one comprising of *japonica* rice varieties (Azucena, Nipponbare and NPT-II) and the other containing traditional Basmati rice varieties (Taraori, HBC19 and Basmati 370) and the remaining cross-bred Basmati rice varieties (Super Basmati, Texmati, CSR30, HKR228, Jasmati and Pusa Basmati 1). The three traditional Basmati rice varieties had a higher similarity coefficients mean ( $=0.964$ ) compared to that observed for cross-bred Basmati varieties (0.612), *indica* varieties (0.754) or *japonica* varieties (0.748). The

cross-bred Basmati rice varieties had varying degree of similarities with the traditional Basmati varieties; Super Basmati was closer to Basmati 370; Texmati and CSR30 were closer to the Taraori Basmati accessions. The three cross-bred Basmati varieties, HKR228, Jasmati and Pusa Basmati 1, grouped together with the traditional group at 0.79 similarity level. The remaining cross-bred Basmati varieties (Kasturi, Sabarmati, Improved Sabarmati and Dehradun Basmati 284), merged with the Group II at a similarity level of 0.67. The cluster analysis also revealed that the traditional Basmati group is closer to the *japonica* group than to the *indica* varieties.

## Discussion

Traditional Basmati rice varieties command premium prices, about three times higher than the cross-bred Basmati or long-grain *indica* rice varieties, in both domestic and international markets (4). Consequently, the letter may be used to adulterate premium traditional Basmati rice supplies. A database of DNA fingerprints of premium Basmati, and other rice varieties should help protection of consumers' interest by providing regulators and quality assurance personnel with methods for authenticating labeling claims and purity.

In this study, a microsatellite fingerprint database has been generated for 24 rice genotypes using SSR primers (available at <http://hau.nic.in/basmati.htm>). High levels of polymorphism were observed among the different rice genotypes. A total of 229 alleles were detected with an average number of alleles of 4.58 per locus (range 1-8 per locus). This value is lower than that reported earlier (7.4 alleles per locus; range 3-17) by Olufowote *et al* (20) but quite comparable to 2.0-5.5 alleles per SSR locus for various classes of microsatellites reported by Cho *et al* (21) using a different set of rice germplasm.

The size difference between the smallest and the largest allele for a given SSR locus varied from 1 to 82 bp. Maximum variation in allele size was observed with the RM markers with GA, AT, ATT and CTT repeat motifs. In this respect, our results are comparable to published SSR diversity data in rice, which have reported high variation in allelic size for markers with GA, AT, ATT and CTT repeat motifs (13, 14, 20, 22-24). SSR loci with greater size difference between the alleles would be relatively better for fingerprinting and diversity analysis. Further, no null

allele (absence of amplification product at a particular locus) was observed. Cho *et al* (21) reported null alleles in 1 to 6 varieties out of 13 studied at 17% of the total (323) microsatellite loci. While the variation in the size of amplified product at microsatellite loci generally results from a change in the number of repeats, null alleles are the consequence of polymorphism in the primer binding site(s). The absence of null alleles may be attributed to the small sample size as well as absence of wild rice genotypes in this study. Null alleles have been most frequent in genetically distant wild *Oryza* species, because of the greater divergence of the primer annealing sequences (14, 22).

Multiple alleles were observed at 13 of the 50 loci analyzed but only in some of the rice varieties including Improved Sabarmati, CSR30, NPT-II, Sharbati, Azucena, IR64, Kasturi, Pusa Basmati 1. Blair *et al* (11) also reported the presence of multiple alleles at six out of thirty microsatellite loci evaluated for 59 genotypes. In this study, most varieties showing multiple alleles are ones developed after several crosses, backcrosses, and selfing involving diverse parental types. Also at 5 of 13 loci that showed the presence of two alleles, one of the alleles was in fact a rare allele. While occurrence of such rare alleles in cross-bred varieties remains to be explained, it may have resulted from un-equal crossing over, translocation, other types of mutations, or residual heterozygosity. Since the varieties were genotyped based on a bulk DNA sample, multiple alleles could also result from intra-varietal heterogeneity (11).

Further, the level of polymorphism, as assessed by the PIC values, was quite high and varied (range 0.0 to 0.8, average value 0.62) considerably among SSR loci. The PIC values observed, are comparable to five previous estimates of microsatellite marker analysis in rice (20-22, 24, 25) but lower than that reported by Blair *et al* (11), who included a more diverse set of rice germplasm. Blair *et al* (11) also reported the presence of rare alleles at a higher frequency (41%).

Microsatellite markers have also been used to detect the adulteration of Basmati rice, supporting their utility in quality assurance (7, 8). In our study, several microsatellite markers (e.g. RM1, RM21, RM38, RM170, RM210, RM226 and RM229) were identified that could readily distinguish the Basmati rice varieties from non-Basmati rice varieties. Even more specifically, eight markers (RM215, RM207, RM44, RM222, RM22, RM2, RM3 and RM60) were

identified that distinguish the closely related traditional Basmati rice varieties, such as Basmati 370, from Taraori Basmati. Thus, fingerprinting technology using SSRs should be a powerful tool to help detection of adulteration of varieties, thus safeguarding the interests of both consumers and producers of traditional Basmati rices (7, 8, 12). It will also be interesting to see if these results can be extrapolated to other traditional Basmati rice and related types. Further, SSRs can be used to prepare a DNA fingerprint database of newly developed rice varieties, which can be used for varietal registration and protection of plant breeders' rights.

Classification and assessment of genetic diversity in Asian cultivated rice is important for Basmati rice breeding. Earlier studies have primarily focused mainly on the differentiation and evolution of *indica* and *japonica* rices (5, 26-28). The *indica* group, has been found to consist of many diverse types as characterized on the basis of morphological traits, origin, isolation barriers ( $F_1$  sterility) and isozyme/molecular marker polymorphism (5, 6, 29, 30). Glaszmann (6) used isozyme data for the classification of the Asian rice varieties and classified them into six major groups. Group I and Group VI corresponded to *indica* and *japonica* rice varieties, respectively. Group II, III, IV and V were atypical and had been classified as *indica* varieties in earlier conventional classifications. Aromatic rice varieties including Basmati types were assigned a separate distinct group V, which was closer to *japonica* than to the typical Group I *indica*. Khush (5) suggested that these different groups may have been domesticated from different populations of *Oryza nivara* at different locations and on different time scales. 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 allelic polymorphism data also suggest a long and independent evolution for Basmati rices in India.

In this study, the grouping of cultivars based on SSR polymorphism data corresponds well to their known pedigree data. For example, three traditional Basmati rice varieties had very high degree of similarity between them. The cross-bred varieties, which had been derived from crosses and back-crosses of Indica x Basmati rice showed varying levels of similarity with the traditional Basmati rice group. These cross-bred varieties may have differing levels of genetic content contributed from their respective *indica*

parents. Other *indica* x Basmati cross-bred varieties, such as Kasturi, Sabarmati, and Improved Sabarmati, clustered with *indica* rice group, suggesting a higher level of genetic content from *indica* rice parent(s). Dehradun Basmati 284, a local scented collection from Uttaranchal/ Uttar Pradesh (India), did not closely resemble the traditional Basmati rice group and clustered with *indica*. Sharbati, which is commonly used for adulteration of Basmati rice supplies, could be easily detected by virtue of allelic differences at the RM26, RM38, RM70, RM223, RM229, RM235 and RM261 loci.

### Acknowledgements

This research was supported by a grant from the Rockefeller Foundation, New York, USA. We also thank Susan McCouch and Stephen Kresovich (Department of Plant Breeding, Cornell University, Ithaca, NY 14853-1901, USA) for providing technical know-how for SSR marker analysis and Post Doctoral Fellowship to one of us (SJ).

Received 24 February, 2003; revised 5 June, 2003.

### References

- 1 **Khush GS and dela Cruz N**, In *Speciality rices of the world: Breeding, production and marketing*, (R Duffy, Editor), Science Pub, Inc, Enfield, USA (2001) pp 15-18.
- 2 **Singh RK, Singh US & Khush GS & Rohilla R**, In *Aromatic rices*, (RK Singh, US Singh, GS Khush, Editors), Oxford & IBH Pub Co Pvt Ltd, New Delhi, India (2000) pp 47-69.
- 3 **Singh RK, Khush GS, Singh US, Singh AK & Singh S**, In *Aromatic rices*, (RK Singh, US Singh, GS Khush, Editors), Oxford & IBH Pub Co Pvt Ltd, New Delhi, India (2000) pp 71-105.
- 4 **Bhasin VK**, In *Aromatic rices*, (RK Singh, US Singh, GS Khush, Editors), Oxford & IBH Pub Co Pvt Ltd, New Delhi, India (2000) pp 252-276.
- 5 **Khush GS**, In *Aromatic rices*, (RK Singh, US Singh, GS Khush, Editors), Oxford & IBH Pub Co Pvt Ltd, New Delhi, India, (2000) pp 5-13.
- 6 **Glaszmann JC**, *Theor Appl Genet*, **74** (1987) 21.
- 7 **Bligh HFJ**, *Intl J Food Sci Technol*, **35** (2000) 257.
- 8 **Bligh HFJ, Blackhall NW, Edwards KJ & McClung AM**, *Crop Sci*, **39** (1999) 1715.
- 9 **Jain S, Mitchel S, Jain RK, Kresovich S & McCouch SR**, In *Rice Genetic IV Supplement*, (GS Khush, DS Brar, B Hardy, Editors), IRRI, Los Baños, Laguna, Philippines (2003) (In press).
- 10 **McCouch SR, Temnykh S, Lukashova A, Coburn J, DeClerck G, Cartinhour S, Harrington S, Thomson M, Septiningsih E, Semon M, Moncada P & Li J**, In *Rice Genetics IV* (GS Khush, DS Brar, B Hardy, Editors), IRRI, Los Baños, Philippines, Science Publishers, Inc, New Delhi, India (2001) pp 117-135.
- 11 **Blair MW, Hedetale V & McCouch SR**, *Theor Appl Genet*, **105** (2002) 449.
- 12 **Nagaraju J, Kathirvel M, Kumar RR, Siddiq EA & Hasnain SE**, *Proc Natl Sci Acad, USA*, **99** (2002) 5836.
- 13 **Chen X, Temnykh S, Xu Y, Cho YG & McCouch SR**, *Theor Appl Genet*, **95** (1997) 553.
- 14 **Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T & McCouch SR**, *Theor Appl Genet*, **100** (2000) 697.
- 15 **Saghai-Marooof MA, Soliman KM, Jorgensen RA & Allard RW**, *Proc Natl Acad Sci, USA*, **81** (1984) 8014.
- 16 **Pal S, Jain S, Saini N & Jain RK**, *Rice Genet Newslett*, **18** (2001) 94.
- 17 **Ghosh S, Karanjawala ZE, Hauser ER, Ally D, Knapp JI, Rayman JB, Musick A, Tannenbaum J, Te C, Shapiro S, Elrridge W, Musick T, Martin C, Smith JR, Carpten, Brownstein MJ, Powell JI, Whiten R, Chines P, Nylund SJ, Magnuson VL, Boehnke M & Collins FS**, *Genome Res*, **7** (1997)165.
- 18 **Anderson JA, Churchill GA, Autrique JE, Tanksley SD & Sorrells ME**, *Genome*, **36** (1993) 181.
- 19 **Rohlf FJ**, *NTSYS-PC: Numerical Taxonomy and Multivariate Analysis System Version 18*, Exeter Software, Setauket, New York (1993)
- 20 **Olufowote JO, Xu Y, Chen X, Park WO, Beachell HM, Dilday RH, Goto M & McCouch SR**, *Genome*, **40** (1997) 370.
- 21 **Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L & McCouch SR**, *Theor Appl Genet*, **100** (2000) 713.
- 22 **Panaud O, Chen X & McCouch SR**, *Mol Gen Genet*, **252** (1996) 597.
- 23 **Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, & McCouch SR**, *Genome Res*, **11** (2001) 1441.
- 24 **Wu KS & Tanksley SD**, *Mol Genet*, **241** (1993) 225.
- 25 **Ni J, Colowit PM & Mackill DJ**, *Crop Sci*, **42** (2002) 601.
- 26 **Kato S, Kosaka H & Hara S**, *Bull Sci Fac Agric Kyushu Univ, Fukuoka, Japan*, **3** (1928) 132
- 27 **Matuso T**, *Bull Natl Institute of Agril Sci, Japan*, **3** (1952) 1 (in Japanese).
- 28 **Ting Y**, *Acta Agron Sinica*, **8** (3) (1957) 243 (in Chinese).
- 29 **Aggarwal RK, Brar DS, Nandi S, Huang N & Khush GS**, *Theor Appl Genet*, **98** (1999) 1320.
- 30 **Ishii T, Brar DS, Second G, Tsunewaki K & Khush GS**, *Jap J Genet*, **70** (1995) 643.