

Molecular Profiling of Major Indian Rice Cultivars Using a Set of Eight Hypervariable Microsatellite Markers

L.R. VEMIREDDY*, N. RANJITHKUMAR, A. VIPPARLA, M. SURAPANENI, G. CHOUDHARY,
K.V. SUDHAKARRAO and E.A. SIDDIQ

Institute of Biotechnology, Acharya NG Ranga Agricultural University, Rajendranagar,
Hyderabad-500030, AP, India

(Received 27 December 2013; Accepted 6 June 2014;
Communicated by H. Bürstmayr)

India bred high yielding rice varieties have enriched to a great extent the global rice germplasm since the mid-sixties. Systematic research efforts for development of cultivar-specific DNA fingerprints of major Indian rice cultivars, however, have not received due attention. The present investigation was aimed at development of DNA fingerprints for 90 high yielding rice varieties using hypervariable microsatellite (hvRM) markers. A panel of eight markers, viz. RM11313, RM13584, RM15004, RM5844, RM22250, RM22565, RM24260 and RM8207 was chosen from 52 polymorphic markers based on their highly polymorphic nature, SSR repeat type and number and ability to distinguish genotypes, in order to develop DNA fingerprints of 90 varieties. The remaining high polymorphic hvRM markers could be of immense value in future to distinguish new cultivars, in case they could not be distinguished by the 8 marker panel. Four of the 8 markers, viz. RM22250, RM13584, RM24260 and RM5844 were located in expressed genes and could be of value in DUS (Distinctness, Uniformity and Stability) testing. Thus we suggested, that this set of 8 loci could be used as standard for DNA fingerprinting of Indian rice cultivars.

Keywords: molecular profiling, hypervariable microsatellite markers, genetic diversity, rice

Introduction

Being on the periphery of the centre of origin of rice, India is endowed with rich genetic diversity. Subject to natural and manual selection since its domestication, the Asian rice (*Oryza sativa*) differentiated into three major geographical groups, viz. *indica*, *japonica* and *javanica* (Siddiq et al. 2012). As a result of Green Revolution in the mid-1960s, the development and release of rice varieties with dwarf plant stature gained momentum in India. As of now, more than 900 high yielding dwarf and semidwarf varieties have been released for general cultivation for irrigated, rainfed upland and lowland rice ecologies. In every ecology, a few varieties remain popular with farmers and consumers for decades

* Corresponding author; E-mail: vlnreddi@gmail.com; Phone:+91-9440440433; Fax-91-40-24022695

since their release. For instance, Swarna, Savithri, BPT 5204, etc., are popular in lowland ecologies, whereas Pusa Basmati-1, Pusa-44, etc., were in irrigated ecology. Further, many improved varieties released worldwide are derived from the donor sources from India only. For instance, the donor of recently released submergence tolerant variety Swarna Sub-1 is FR13A, a deep-water landrace from India (Xu et al. 2006). Likewise, GEB-24, a high quality rice from Southern India, was the source to develop as many as 31 varieties from IRRI and elsewhere, while Eswarakara of Andhra Pradesh state has been the source for many of the gallmidge resistant varieties (Roy et al. 1969).

These examples underscore the importance of Indian rice varieties as valuable gene sources. Establishment of the genetic identity of high yielding varieties/donor sources through development of cultivar specific molecular profiles is important for their registration under the Protection of Plant Varieties and Farmers Rights Act, so as to assert plant breeders' right as well as to prevent unauthorized use on commercial scale. Traditionally, morphological traits and isozyme markers were used to establish the identity of a variety (Glaszmann 1987; Patra and Chawla 2010). Such traits and markers are, due to their limited numbers, stage specificity and environmental influence of expression, not efficient enough in nature to characterize reliably and precisely the germplasm. On the other hand, the highly heritable, abundant and polymorphic DNA markers are powerful enough to discriminate not only distantly related but also closely related genotypes. For this reasons, DNA fingerprinting is considered and widely used as potential for establishing genetic identity and relatedness of genotypes in crop breeding research and germplasm management.

Among DNA based markers, microsatellites or SSRs (Simple Sequence Repeat) are found to be of value in gene mapping and tagging of traits of interest, genetic diversity analysis and establishment of genetic purity (Garris et al. 2005; Marri et al. 2005; Vemireddy et al. 2007; Sundaram et al. 2008; Srividhya et al. 2011). After whole genome sequencing, more than 50,000 microsatellite markers are available now (Zhang et al. 2007), of which 18,828 were Class I SSRs, i.e. SSRs containing perfect repeats (= 20 nucleotides in length) and also known as hypervariable SSRs. Recently, evenly distributed hypervariable SSRs have been shortlisted and demonstrated to be quite handy in studies of genetic diversity, population structure analysis, DNA fingerprinting and marker-assisted selection/breeding in rice (Narshimulu et al. 2011). Also, very recently, hypervariable SSR markers have been used to estimate the genetic diversity of major Indian rice cultivars released from 1970 to 2010 (Choudhary et al. 2013). On the strength of this background, the present study was initiated to develop cultivar specific allelic profiles for major Indian rice varieties using a set of 8 hypervariable SSR markers.

Materials and Methods

Plant material, genomic DNA isolation and microsatellite genotyping

A total of 90 genotypes comprising most popular high yielding rice varieties released in India were included in the study (Table S1*). DNA was isolated from leaf samples of

* Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

20 days old seedlings of the 90 rice genotypes grown in Petri dishes under room temperature following the modified CTAB method (Murray and Thompson 1980). Fifty-two hypervariable microsatellite markers distributed over the 12 chromosomes were chosen for the study. PCR amplification was performed in 10 µl reaction containing 20 ng DNA, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 0.5 unit of Taq polymerase, 50 µM of dNTPs and 0.1 µM each of forward and reverse primers using Veriti (ABI, Foster City, CA, USA) thermal cycler. Conditions for thermal cycler reaction used were as follows: after an initial denaturation at 94 °C for 10 min, the PCR mix was cycled 35 times at 94 °C of denaturation for 45 s, 55–60 °C of annealing for 45 s (according to the optimal temperature of the primers) and 72 °C of extension for 60 s followed by a final extension step at 72 °C for 10 min. The samples were then mixed with bromophenol blue and run on a 3% agarose gel along with the 50-bp ladder (MBI Fermentas, Canada) for an hour in 0.5× Tris-Acetic acid-EDTA (TAE) buffer. The resolved PCR bands were documented using Bio-Rad Molecular Imager Gel Doc XR System. If PCR products were not amplified, the reactions were re-amplified.

PCR with tailed primers

The allele sizes of the eight primers, viz. RM11313, RM13584, RM15004, RM5844, RM22250, RM22565, RM24260 and RM8270, were confirmed by screening all 90 genotypes with tailed primers. The primer sequences can be obtained from the gramene site (www.gramene.org). PCR reaction was carried out with 1 picomole of forward primer tailed at 5' end with 19 base sequence (5' CACGACGTTGTAAAACGAC3'), 2 picomole of reverse primers, 1 picomole of M13 fluorescent label (6-FAM) were added along with all the PCR components. The amplified PCR samples were run in ABI3130XL capillary electrophoresis and allele sizes were estimated using Genemapper v 3.7 software (Applied Biosystems, Foster City, USA).

Data analysis

Only clear and unambiguous bands of SSR markers were scored. The sizes of the amplified fragments were estimated with the help of Alpha image software by Gel documentation system using 50-bp DNA ladders (MBI Fermentas, Canada) as size standards. Markers were scored for the presence (1) or absence (0) of the corresponding band among the genotypes. A data matrix comprising '1' and '0' was formed depending on the character. The data matrix was subjected to unweighted pair group method with arithmetic averages algorithm (UPGMA) cluster analysis using the program NTSYSpc 2.02i to estimate the genetic relatedness among genotypes. To measure the informativeness of the markers, the polymorphism information content (PIC) for each SSR marker was calculated according to the formula $PIC = 1 - \sum p_i^2$ where 'i' is the total number of alleles detected for SSR marker and 'Pi' is the frequency of the *i*th allele in the set of 90 genotypes investigated.

DNA fingerprinting

In our previous study, we have shortlisted 840 hypervariable rice microsatellite markers (hvRM) from the publicly available 18,882 Class I SSR markers (Narshimulu et al. 2011). Of the 840 hvRM markers, 52 were chosen for the present study based on their distribution and evenness over all the 12 chromosomes to screen the 90 genotypes. Of 52 markers, eight hvRM markers were picked out based on their high PIC values and repeat type and ability to discriminate the cultivars unambiguously. Using these eight markers cultivar specific allelic profile was developed, by which one can differentiate all the cultivars. Usefulness of these markers have been validated in a set of survey samples of BPT3291 (Sona Mahsuri), a popular variety, collected from 24 villages in Andhra Pradesh, India.

Results

Hypervariable microsatellite marker analysis

Fifty-two hvRM markers, distributed evenly on all the 12 chromosomes, were chosen to generate molecular profiles and assess thereby genetic variability among the 90 genotypes. A total of 184 alleles was identified from amplification of the 52 polymorphic SSR loci with product sizes ranging from 110 to 650 bp (Fig. 1). Average number of alleles per locus was 3.53, ranging from 2 (RM16416) to 7 (RM8207). All parameters of polymorphic markers were given in Table 1.

Cluster analysis

The cluster analysis based on 52 primers revealed similarity coefficients ranging from 0.61 to 0.89, while it was 0.55 to 0.97 with eight SSR markers used. The dendrogram grouped the 90 genotypes into four major clusters in the former case and into six clusters in the latter. Samba Mahsuri and Improved Samba Mahsuri fell into two different clusters with a genetic similarity of 62% in the dendrogram using 52 SSRs, while in the other case both the genotypes grouped in the same cluster exhibiting a similarity of 69%. Similarly, Swarna and Swarna Sub1 fell in two different clusters in both cases, showing a similarity of 63% for all 52 markers used and 56% for eight markers used. The Basmati varieties, viz. Basmati370 and Basmati386 along with Swarna formed a separate cluster showing genetic similarities ranging from 70 to 86% with 8 primers. Tikkana and Pinakini together formed a sub-cluster in both cases with 85% similarity. Likewise, 'Pusa44' and 'Ghanteswari' also grouped together as a sub-cluster with a similarity of 91%. Comparison of cluster analysis using 52 and 8 markers revealed different clustering patterns. A possible explanation is that the 8 markers chosen for molecular profiling did not cover all the 12 chromosomes (Fig. 2).

Development of cultivar specific allele profiles

For screening of all the genotypes, a subset of eight markers, viz. RM11313 (PIC-0.86), RM13584 (PIC-0.88), RM15004 (PIC-0.91), RM5844 (PIC-0.94), RM22250 (PIC-0.94),

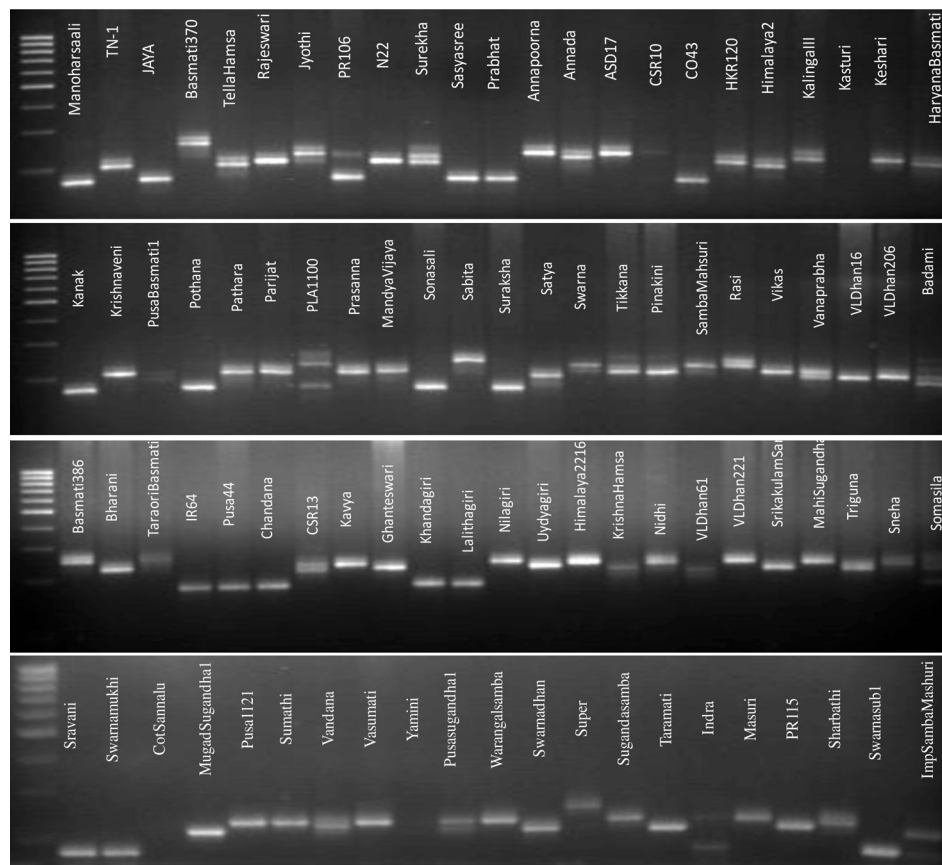


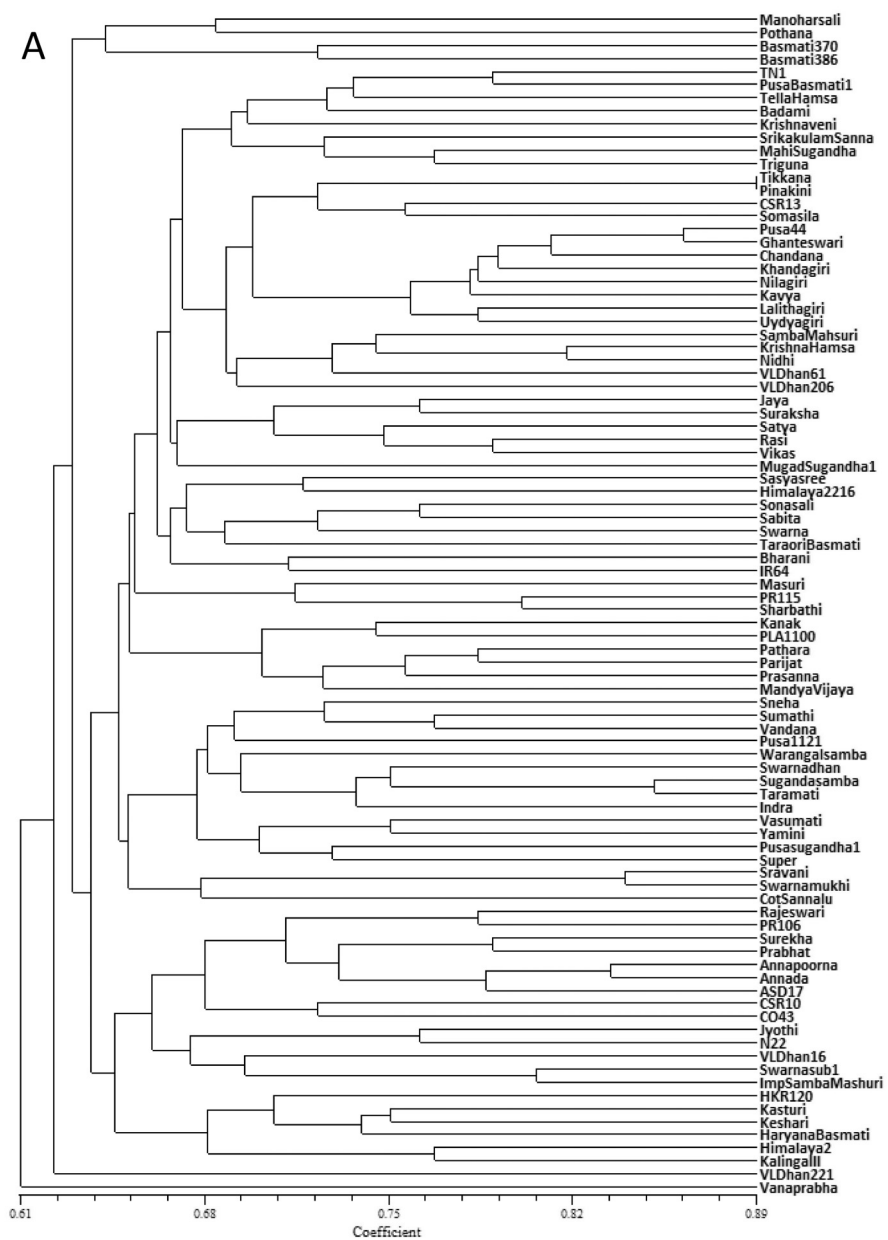
Figure 1. Amplification profile of 90 rice genotypes using RM22250 primer

Table 1. SSR locus, chromosome location (Chr.), number of alleles (Na), SSR motifs, SSR allele size (bp), PIC values for 52 microsatellite markers used in the study

S. No.	SSR locus	Chr.	Na	SSR motif	Allele size (bp)	PIC
1	RM562	1	5	(AAG)13	110–140	0.94
2	RM11313	1	3	(AAT)23	380–400	0.86
3	RM11340	1	4	(AT)42	280–300	0.88
4	RM11356	1	3	(AT)31	360–390	0.84
5	RM11597	1	3	(AT)42	390–420	0.86
6	RM12031	1	3	(AG)40	300–320	0.59
7	RM12253	1	4	(AAT)23	340–370	0.90
8	RM12292	1	3	(AT)38	450–470	0.92
9	RM12353	2	3	(AAT)32	460–480	0.81
10	RM12548	2	3	(AT)46	340–360	0.88

Table 1 (cont.)

S. No.	SSR locus	Chr.	Na	SSR motif	Allele size (bp)	PIC
11	RM13131	2	3	(AT)32	290–310	0.86
12	RM13584	2	3	(AATC)5	390–420	0.88
13	RM14270	3	4	(AT)46	380–400	0.87
14	RM14735	3	5	(AT)42	290–320	0.97
15	RM14778	3	3	(AT)37	190–210	0.88
16	RM15004	3	4	(AAT)38	410–440	0.89
17	RM15580	3	3	(AT)50	260–290	0.86
18	RM16416	4	2	(AAT)12	250–260	0.63
19	RM16577	4	3	(AT)29	480–500	0.85
20	RM17405	4	3	(AAT)36	520–540	0.85
21	RM17669	4	3	(AT)30	190–210	0.86
22	RM5693	5	4	(AAT)18	360–380	0.74
23	RM5844	5	5	(AAT)20	190–220	0.91
24	RM5907	5	3	(AAT)19	260–280	0.84
25	RM18384	5	5	(AAG)22	290–320	0.85
26	RM18639	5	4	(AAT)17	310–380	0.74
27	RM19545	6	3	(AAT)21	460–490	0.88
28	RM20037	6	3	(AT)38	260–290	0.86
29	RM20710	6	5	(AT)46	260–290	0.95
30	RM21693	7	5	(AT)44	450–490	0.90
31	RM21941	7	3	(AAT)25	210–230	0.86
32	RM22250	8	5	(AAT)30	250–300	0.93
33	RM22554	8	4	(AAT)20	160–200	0.86
34	RM22565	8	4	(ACAT)15	260–300	0.90
35	RM22688	8	2	(AAT)28	420–440	0.69
36	RM22273	8	4	(AT)35	270–290	0.84
37	RM23017	8	3	(AAT)18	190–200	0.88
38	RM23036	8	4	(AGAT)15	240–280	0.89
39	RM23362	8	5	(AAG)19	170–190	0.92
40	RM23741	9	4	(AAT)28	440–470	0.74
41	RM24015	9	4	(AGAT)9	180–200	0.84
42	RM24044	9	3	(AAG)11	190–220	0.88
43	RM24260	9	3	(AAT)31	240–270	0.89
44	RM5708	10	4	(AAT)22	260–290	0.83
45	RM8207	10	5	(AAG)23	140–190	0.91
46	RM25262	10	2	(AAT)38	380–390	0.74
47	RM25969	11	3	(AAG)18	340–360	0.81
48	RM26190	11	3	(AGAT)13	240–270	0.86
49	RM26632	11	3	(AAAG)9	440–470	0.86
50	RM27840	12	3	(AAT)37	510–530	0.87
51	RM28279	12	3	(AATC)8	310–330	0.81
52	RM6965	7	3	(AAG)15	630–650	0.93
Mean			3.53			0.85



(See Part B on the next page.)

Figure 2. Comparison of dendrograms based on 52 (A) and 8 (B) hvRM markers showing phylogenetic relationship among 90 rice genotypes

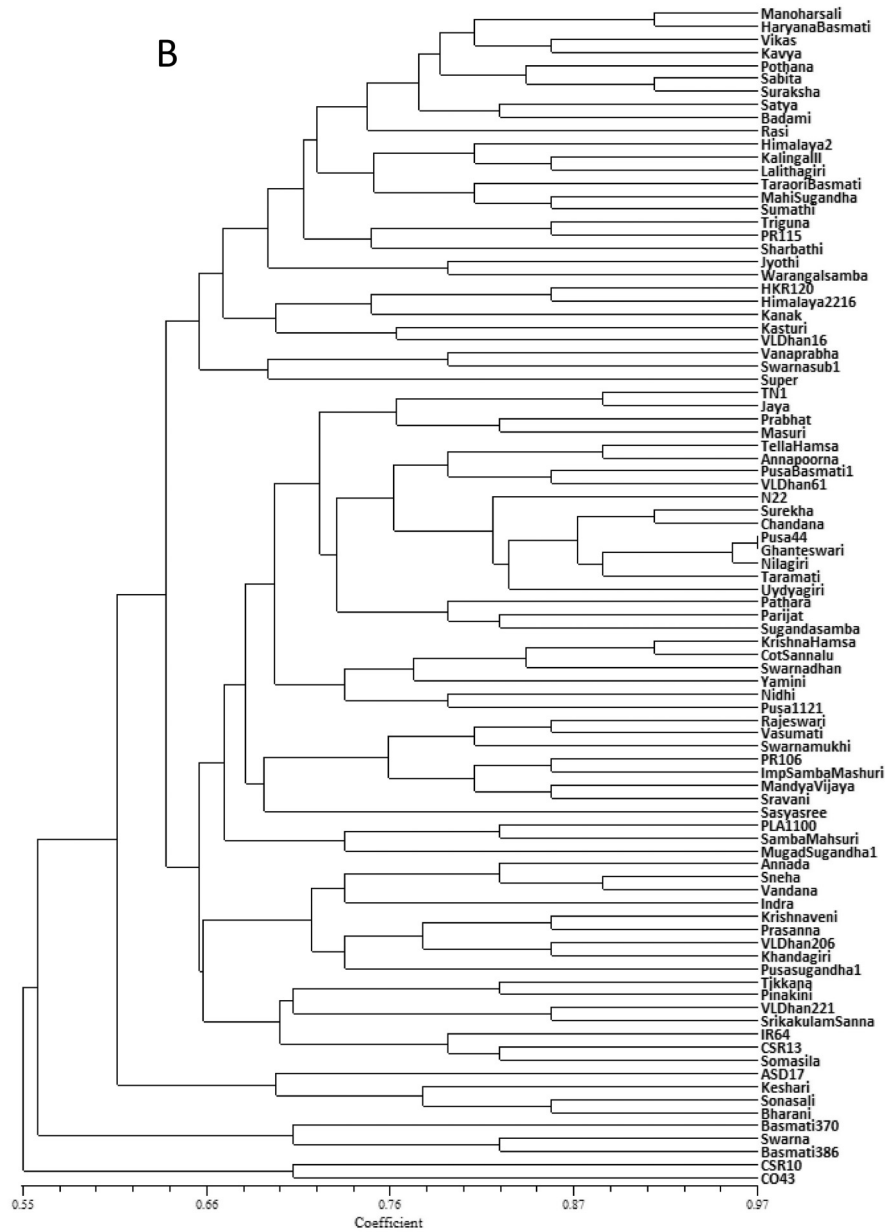


Figure 2 (cont.)

RM22565 (PIC-0.9), RM24260 (PIC-0.88) and RM8207 (0.95) was chosen based on their high PIC values (ranging from 0.86 to 0.95) and microsatellite repeat numbers and types and their ability to distinguish the genotypes. These markers have been found to be sufficient and efficient for unequivocal identification of the test genotypes (Table 2). The eight markers were distributed on the chromosomes 1, 2, 3, 5, 8, 9 and 10. The type of microsatellite repeats of these eight markers used in the present study are two tetra- and six tri-nucleotide types of 17 di-, 28 tri- and 7 tetra-types. Since di-nucleotide repeat containing microsatellite markers cause significant noise in the capillary electrophoresis (CE) due to slippage, they were precluded from the shortlisted loci chosen for cultivar discrimination. The average number of SSR repeats for dinucleotide hvRMs (39) is higher than for tri- (23) or tetra-types (11). The latter, however, exhibit almost the same number of alleles (*Na*) and PIC values as those of the di-nucleotide hvRMs, indicating that neither number nor type of SSR repeats correlate directly with *Na* and PIC. For instance, tri- and tetra-nucleotide loci RM562 (0.94) and RM26190 (0.92) with high PIC did not have a higher cultivar discriminating power compared to low PIC exhibiting loci, such as RM11313 (0.86), RM13584 (0.88) and RM24260 (0.88). Likewise, the loci, such as RM5844 ((AAT)20), RM11313 ((AAT)23) and RM8207 ((AAG)23) having less microsatellite repeats, proved to be more effective in cultivar discrimination as compared to high numbers of repeat containing loci, like RM25262 ((AAT)38), RM27840 ((AAT)37), etc.

For distinguishing cultivars from one another easily, alleles were given codes, i.e. A, B, C, D, and E based on their allele sizes in ascending order. The patterns of allele codes of these eight markers were found unique for each of the cultivars studied. Differences in the patterns with respect of even a single allele code could distinguish one variety from the other. For instance, the variety Manoharsali had allele code (AAAAACAB), which is different from TN1 (BAABACBA) suggesting that the allele codes can be used for distinguishing varieties from each other. We have developed a specific code for all 90 rice varieties. Similar kind of allele codes had been developed for all notified traditional, evolved and non-Basmati varieties earlier

Samba Mahsuri and Swarna sub1A, introgressed with bacterial blight resistance and submergence tolerance genes, from their respective original varieties Samba Mahsuri (BPT5204) and Swarna, though they were claimed to be more than 95% similar to their original varieties. Thus, on the basis of unique allele code patterns, 90 cultivars could be distinguished using the subset of eight markers.

Discussion

The present study was aimed to develop cultivar specific allele profiles for major Indian high yielding cultivars employing a set of hypervariable SSR markers. Of the 52 hvRMs, eight, viz. RM11313, RM13584, RM15004, RM5844, RM22250, RM22565, RM24260 and RM8270 have been found to be sufficient for unambiguously distinguishing all the representative major rice cultivars of India. The eight markers distributed on as many as seven chromosomes (1, 2, 3, 5, 8, 9 and 10) covering uniformly and nearly the entire rice genome, satisfied the standards proposed by the UPOV (<http://www.upov.int>) with

Table 2. Microsatellite allele profile of major Indian rice cultivars

S. No.	Variety	RM22250	RM22565	RM13584	RM8207	RM24260	RM5844	RM15004	RM11313	Allele code
1	Manoharsali	A	A	A	A	A	C	A	B	AAAAACAB
2	TN1	B	A	A	B	A	C	B	A	BAABACBA
3	Jaya	A	A	A	B	A	C	A	B	AAABACAB
4	Basmati370	E	C	A	A	B	D	B	C	ECAABDBC
5	Tella Hamsa	B	A	A	B	C	B	B	B	BAABCBBB
6	Rajeswari	C	D	B	C	B	D	B	A	CDBCDBBA
7	Jyothi	D	B	B	B	B	D	B	B	DBBBDBBB
8	PR106	A	A	B	C	B	D	A	A	AAABCBA
9	N22	C	A	B	C	A	B	B	B	CABCABBB
10	Surekha	B	A	B	D	B	B	C	B	BABDBBCB
11	Prabhat	A	A	B	C	B	D	A	B	AAABCDBA
12	Swarnadhan	C	C	A	B	C	C	B	A	CCABCBA
13	Masuri	D	B	B	B	A	A	B	A	DBBBAABA
14	Annapooma	D	A	B	C	C	B	A	B	DABCCBAB
15	Annada	D	D	B	C	B	D	A	B	DDBCBDAB
16	ASD17	D	C	B	C	C	B	A	B	DCBCCBAB
17	CSR10	D	B	B	C	C	D	C	B	DBBDCDCB
18	CO43	A	A	B	D	C	D	C	B	AABDCDCB
19	HKR120	C	B	B	B	A	C	C	A	CBBACCA
20	Himalaya2	B	A	B	D	B	C	B	B	BABDBCBB
21	Kalinga III	C	A	B	D	B	B	A	B	CABDBBAB
22	Kasturi	C	D	B	A	B	D	B	B	CDBABDBB
23	Keshari	B	C	B	C	C	B	B	C	BCBCCBBC
24	Kanak	A	B	A	A	A	C	A	B	ABABACAB
25	Pusa Basmati1	B	A	B	A	A	B	B	A	BABAABBA
26	Pothana	A	A	A	B	B	D	A	B	AAABBDAB
27	Pathara	B	C	A	B	B	B	C	B	BCABBBBCB
28	Parijat	B	A	A	C	B	C	C	B	BAACBCCB
29	PLA1100	A	D	A	A	A	B	C	B	ADAAAACB
30	Prasanna	B	A	B	C	A	C	B	B	BABCBCBB
31	Mandya Vijaya	B	C	A	A	B	D	B	B	BCAABDBB
32	Sasya Sree	A	B	A	C	B	D	B	B	ABACBDBB

Table 2 (cont.)

S. No.	Variety	RM22250	RM22565	RM13584	RM8207	RM24260	RM5844	RM15004	RM11313	Allele code
33	Sonali	A	B	A	C	A	B	B	B	ABACABBB
34	Sabita	C	B	A	B	B	B	B	A	CBABBBBA
35	Suraksha	A	A	A	B	A	B	A	B	AAABABAB
36	Satya	B	A	A	B	C	B	B	B	BAABCBBB
37	Swarna	C	C	A	A	B	D	C	C	CCAABDCC
38	Tikkana	B	D	B	B	B	B	B	A	BDBBCBA
39	Pinakini	B	A	B	B	B	C	C	A	BABBBCCA
40	Samba Mahsuri	C	A	B	A	B	B	C	A	CABABBCA
41	Rasi	D	A	B	B	C	C	C	A	DABBBCCA
42	Vikas	B	A	B	B	C	C	B	A	BABBBCCA
43	Vanaprabha	B	A	B	D	A	C	B	B	BABDACBB
44	VLDhan16	C	D	B	C	B	B	A	B	CDBCBBAB
45	VLDhan206	C	B	B	C	A	C	A	B	CBBCACAB
46	Haryana Basmati	B	B	B	A	A	C	B	B	BBBAACBB
47	Krishnaveni	B	A	B	C	A	C	B	B	BABCACBB
48	Badami	B	A	A	B	B	B	A	A	BAABBBAB
49	Basmati386	C	C	A	A	A	B	C	C	CCAAAACB
50	Taraori Basmati	C	C	A	B	A	C	A	A	CCABACAB
51	IR64	A	B	A	C	A	C	B	A	ABBCACBA
52	Pusa44	A	A	A	B	A	B	B	B	AAABABBA
53	Chandana	C	A	A	B	A	B	C	C	CAABABCA
54	CSR13	B	D	A	C	B	C	C	C	BDACBCCA
55	Kavya	C	A	A	B	A	C	B	B	CAABACBA
56	Ghanteswari	C	A	A	B	A	B	B	B	CAABABBA
57	Khandagiri	A	B	A	B	A	C	B	B	ABABACBA
58	Lalithagiri	C	A	A	C	B	C	B	B	CAACBCBB
59	Nilagiri	B	A	B	B	A	B	A	A	BABBABAA
60	Udyagiri	C	B	B	C	C	B	B	B	CBBCBBBA
61	Himalaya2216	C	B	B	B	B	C	C	C	CBBBBCBB
62	Krishna Hamsa	C	C	B	B	C	C	B	B	CCBBCCBA
63	Nidhi	C	A	B	C	B	C	B	B	CABCBCBA
64	VLDhan61	B	A	B	C	A	B	B	B	BABCABBB

Table 2 (cont.)

S. No.	Variety	RM22250	RM22565	RM13584	RM8207	RM24260	RM5844	RM15004	RM11313	Allele code
65	VLDhan221	C	D	D	B	A	C	A	A	CDDBACAA
66	Mahi Sugandha	A	C	A	B	B	C	A	A	ACABBCAB
67	Triguna	C	A	A	B	B	B	B	B	CAABBBBA
68	Sneha	C	D	A	C	B	C	A	A	CDACBCAA
69	Swarnamukhi	A	C	A	B	A	D	A	A	ACABADAA
70	Bharani	B	B	B	C	A	B	A	A	BBBCABAA
71	Somasila	A	D	B	C	B	C	A	A	ADBCBCAA
72	Sravani	A	C	A	C	A	D	A	A	ACACADAA
73	Srikakulam Sannalu	B	A	A	B	B	C	B	B	BAABBCBA
74	Cottondora Sannalu	A	B	B	B	A	C	A	A	ABBBACAA
75	Mugad Sugandha I	C	A	A	B	A	B	A	A	CAABABAA
76	Pusa 1121	C	A	A	C	B	C	C	C	CAACBCCB
77	Sumathi	C	D	A	B	B	C	A	A	CDABBCAA
78	Vandana	C	A	A	C	B	C	A	A	CAACBCAA
79	Vasumati	C	A	A	B	B	D	B	B	CAABDDBB
80	Yamini	C	C	A	A	C	D	B	B	CCAACDDB
81	Pusa sugandha	B	C	A	A	C	C	C	C	BCAACCCA
82	WGL32100	D	A	A	B	A	D	B	B	DAABCDDBA
83	Super	C	A	A	A	A	C	A	A	CAAAACAC
84	Suganda samba	D	A	A	C	C	B	C	C	DAACCBCA
85	Taramati	C	A	A	B	B	B	B	B	CAABBBBA
86	Indra	A	A	A	B	B	B	C	C	AAABBBBCA
87	PR115	C	B	B	B	B	A	C	C	CBBBBACA
88	Sharbathi	D	B	B	B	A	C	B	B	DBBBACBA
89	Swarna Sub1	C	A	B	D	B	A	A	A	CABDBAAA
90	Imp. Samba Mashuri	B	A	B	A	B	A	B	B	BABABABA
Allele size code key										
A		250	260	390	140	240	190	410	380	
B		270	270	400	160	260	200	420	390	
C		290	280	420	170	270	210	440	400	
D		300	300		190		220			
E		310								

respect to genome distribution of molecular markers for DNA profiling of cultivars. Hence, these markers could be of great value not only in cultivar identification but also in DUS (Distinctness, Uniformity and Stability) testing for genetic identity and uniformity as well as for detection and quantification of adulteration in high quality rices like basmati. In spite of the fact, that all the major cultivars are derived from a few foundation genotypes, still it was possible to discriminate and establish them as unique cultivars employing the subset of hypervariable markers. Also, the possibility of identifying as many as 40,320 combinations of allele profiles with the aid of this subset of markers, suggested its potential in distinguishing more and new varieties to come in future. Characterization of crop cultivars such as soybean (Giancola et al. 2002), pepper (Yong-Sham et al. 2005), corn (Gunjaca et al. 2008) and rice (Bonow et al. 2009) using molecular profiles enabled to grant and protect the rights granted to breeders.

The allele codes, developed in the course of the present study, for establishing the distinctiveness of rice cultivars, could be considered as an official DNA signature to distinguish all future rice genotypes from the existing cultivars with accuracy and thereby protecting breeders rights. The marker profiles could be a powerful tool for the characterization of rice accessions, avoiding redundancy and duplicates in the germplasm being collected and conserved in the gene banks. The evenly distributed eight hypervariable SSR markers exhibiting higher PIC values and higher allele numbers (0.87 and 3.63) compared to the randomly selected markers (0.77 and 2.48) could be used as potential background markers in marker assisted selection as reported earlier (Narshimulu et al. 2011).

The overall PIC values using hvRM markers are substantially higher than what was reported by earlier workers in rice using larger numbers of accessions (Garris et al. 2005; Giarrocco et al. 2007). The shortlisted eight marker panel, found to be able to distinguish the traditional, evolved and non-basmati rice varieties clearly, is quite valuable not only to identify the highly prized traditional basmati rice from relatively low priced basmati like non-basmati varieties, but also to accurately detect and quantify adulteration. Earlier, a CE-based assay was developed for detection as well as quantification of adulteration in basmati rice employing eight microsatellite markers, i.e. RM1, RM171, RM202, RM241, RM348, RM44, RM55 and RM72 (Archak et al. 2007; Vemireddy et al. 2007). Out of the eight microsatellite markers, however, five, i.e. RM1, RM202, RM241, RM44 and RM55 were of di-nucleotide repeats.

Since di-nucleotide repeat microsatellite markers produce stutter bands when run in capillary electrophoresis, it is difficult to determine the major peak of the allele, thereby causing an erroneous quantification of level of adulteration in export consignments. This is less of a problem with the use of tri- and tetra-nucleotide repeat based SSRs. These results reaffirm, that the ideal set of markers for use in molecular characterization of cultivars should consist of either tri- or tetra-nucleotide repeat based markers.

Applicability of eight marker profile in survey samples

The marker profiles developed in the present investigation will be more reliable for distinguishing rice cultivars as well as detecting possible adulterations. The power of the eight hypervariable markers, was validated in seed samples of BPT3291 (Sona Mashuri) col-

lected from 24 villages in Vizianagaram district of Andhra Pradesh, India, wherein the level of adulteration was found to range between 10 and 12% as compared to the pure sample of the variety (data not shown). Also, an eight marker keys could be used to resolve differences among cultivars bearing the same name. For instance, the popular fine grain varieties of South India, i.e. BPT3291 (Sona Mahsuri) and similar ones are marketed as highly consumer preferred Samba Mahsuri (BPT5204) and Kurnool Mahsuri, though they are genetically different varieties. They could be clearly differentiated using the eight marker keys developed in the present study. Likewise, the key has been found quite efficient in distinguishing improved versions of the old popular varieties as well. For instance, improved varieties of BPT5204 and Swarna, viz. improved BPT5204 and Swarna-Sub1, respectively, could be distinguished from their original parents using this eight marker keys, though they were reported to be over 95% similar to their parental varieties.

A recent study (Bonow et al. 2009) suggested that SSR markers identified in and around the introgressed gene(s) of interest were useful for identification of Brazilian rice varieties as well as for DUS testing. Also, they have suggested that microsatellite markers linked to expressed genes could greatly supplement the morphological descriptors in establishing distinctiveness of new varieties as recommended by the International Union for the Protection of Plant Varieties (UPOV). Of the eight marker panels chosen for the cultivar identification in the present study four markers, i.e. RM22250, RM13584, RM24260 and RM5844 have been found to be located in expressed genes (LOC_Os04g02030.1 (Protein coding), LOC_Os02g39570 (acetolactate synthase/amino acid binding protein), LOC_Os09g23720.1 (expressed protein) and LOC_Os05g16180 (NB-ARC domain containing protein). These four markers were found to differentiate more than 50% of the varieties used in the present study indicating their potential use in DUS testing.

In summary, this is the first systematic study, to our knowledge, to develop cultivar specific allele profiles for major Indian rice varieties. The shortlisted eight hvRM loci are useful for Indian rice breeders for choosing genetically divergent parents for breeding of complexly inherited traits like yield, etc., exploiting the available genetic variation, and precisely classifying new varieties. The remaining high PIC hvRM markers could be of immense value in the future to conserve new cultivars, in case they could not be distinguished by the eight marker panels. Four of the eight markers located in genes could be of value in DUS testing by supplementing the morphological descriptors. This panel of eight markers, therefore, may be considered as reliable and authentic “DNA fingerprint” of Indian rice varieties for unambiguous identification.

References

- Archak, S., Lakshminarayana Reddy, V., Nagaraju, J. 2007. High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in basmati rice (*Oryza sativa*). *Electrophoresis* **28**:2396–2405.
- Bonow, S., Von Pinho, E.V.R., Vieira, M.G.C., Vosman, B. 2009. Microsatellite markers in and around rice genes. *Crop Sci.* **49**:880–886.

- Choudhary, G., Ranjitkumar, N., Surapaneni, M., Deborah, D.A., Vipparla, A., Anuradha, G., Siddiq, E.A., Vemireddy, L.R. 2013. Molecular genetic diversity of major indian rice cultivars over decadal periods. *PLoS one* **8**:e66197.
- Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., McCouch, S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**:1631–1638.
- Giancola, S., Marcucci Poltri, S., Lacaze, P., Hopp, H.E. 2002. Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. *Euphytica* **127**:95–113.
- Giarocco, L.E., Marassi, M.A., Salerno, G.L. 2007. Assessment of the genetic diversity in Argentine rice cultivars with SSR. *Crop Sci.* **47**:853–858.
- Glaszmann, J.C. 1987. Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* **74**:21–30.
- Gunjaca, J., Buhinicek, I., Jukic, M., Sarcevic, H., Vragolovic, A., Kozic, Z., Jambrovic, A., Pejic, I. 2008. Discriminating maize inbred lines using molecular and DUS data. *Euphytica* **161**:165–172.
- Marri, P., Sarla, N., Reddy, L., Siddiq, E.A. 2005. Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. *BMC Genetics* **6**:33.
- Murray, M.G., Thompson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **8**:4321–4325.
- Narshimulu, G., Jamaloddin, M., Vemireddy, L.R., Anuradha, G., Siddiq, E.A. 2011. Potentiality of evenly distributed hypervariable microsatellite markers in marker-assisted breeding of rice. *Plant Breeding* **130**:314–320.
- Patra, N., Chawla, H.S. 2010. Biochemical and RAPD molecular markers for establishing distinctiveness of basmati rice (*Oryza sativa* L.) varieties as additional descriptors for plant variety protection. *Indian J. Biotech.* **9**:371–377.
- Roy, J.K., Israel, P., Panwar, M.S. 1969. Breeding for insect resistance. *Oryza* **6**:38–44.
- Siddiq, E.A., Vemireddy, L.R., Nagaraju, J. 2012. Basmati rices: Genetics, breeding and trade. *Agric. Res.* **1**:25–36.
- Srividhya, A., Vemireddy, L., Sridhar, S., Jayaprada, M., Ramanarao, P., Hariprasad, A., Reddy, H., Anuradha, G., Siddiq, E. 2011. Molecular mapping of QTLs for yield and its components under two water supply conditions in rice (*Oryza sativa* L.). *J. Crop Sci. Biotechnol.* **14**:45–56.
- Sundaram, R.M., Naveenkumar, B., Biradar, S.K., Balachandran, S.M., Mishra, B., IlyasAhmed, M., Viraktamath, B.C., Ramesha, M.S., Sarma, N.P. 2008. Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica* **163**:215–224.
- Vemireddy, L.R., Archak, S., Nagaraju, J. 2007. Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). *J Agric Food Chem.* **55**:8112–8117.
- Xu, K., Xu, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., Ismail, A.M., Bailey-Serres, J., Ronald, P.C., Mackill, D.J. 2006. Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**:705–708.
- Yong-Sham, K., Je-Min, L., Gi-Bum, Y., Seung-In, Y., Kyung-Min, K., Eun-Hee, S., Kyung-Mi, B., Eun-Kyung, P., In-Ho, S., Byung-Dong, K. 2005. Use of SSR markers to complement tests of distinctiveness, uniformity, and stability (DUS) of pepper (*Capsicum annum* L.) varieties. *Mol. Cells* **19**:428–435.
- Zhang, Z., Deng, Y., Tan, J., Hu, S., Yu, J., Xue, Q., 2007. A genome-wide microsatellite polymorphism database for the indica and japonica rice. *DNA Research* **14**:37–45.

Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at <http://www.akademai.com/content/120427/>

Electronic Supplementary Table S1. Rice genotypes included in the study