Use of anchored $(AG)_n$ and $(GA)_n$ primers to assess genetic diversity of Indian landraces and varieties of rice

N. Sarla*, C. N. Neeraja and E. A. Siddiq

Directorate of Rice Research, Rajendranagar, Hyderabad 500 030, India

The potential of ISSR–PCR for diversity analysis was evaluated in 86 accessions of Indian rice using 14 anchored primers based on (AG) and (GA) repeats. The accessions included 50 landraces, 20 improved varieties and 16 other accessions for comparison. In all, 220 band positions (loci) and 5514 bands were generated and all loci were polymorphic. The 14 primers revealed genetic relationships among the 86 accessions adequately. The wild accessions, japonica varieties, improved indica varieties, hybrids and landraces were clearly delineated. The frequency of clustered GA repeats was more in varieties than in landraces. The polymorphism information content (PIC) and resolving power (Rp) values of each primer were different in landraces and varieties. Both PIC and Rp were negatively correlated to mean genetic similarity among accessions. Primers with 3′ dinucleotide anchor revealed a higher PIC than those with one nucleotide anchor. The most informative primer $(GA)_8$ YG alone could help distinguish 81 accessions. The average expected heterozygosity of landraces was higher than that of varieties. The four wild accessions of *Oryza nivara* and *O. rufipogon* were together more distant from the landraces than from the varieties. The landraces are quite distinct from the cultivated varieties and need to be conserved and used to widen the genetic base of rice varieties. $(AG)_n$ and $(GA)_n$ based primers are highly informative for cost-effective assessment of genetic diversity in rice germplasm.

Keywords: Fingerprinting, GA repeats, ISSR–PCR, landraces, polymorphism information content, resolving power.

RICE is a highly polymorphic crop species with wide geographic dispersal and ecogenetic differentiation. The large germplasm of rice includes many landraces and related wild species. India is a major centre of diversity, notably the mid-eastern part (Chattisgarh region) and the north-eastern hills. Except for a few earlier studies on classification of germplasm using isozymes[1], there have been few studies to assess the diversity of Indian landraces *vis-à-vis* varieties at the molecular level[2]. ISSR–PCR (Inter Simple Sequence Repeat–Polymerase Chain Reaction) is a microsatellite-based multilocus marker technique, which is simple and useful for estimating genetic diversity in several crop plants[3,4]. The technique has the advantages of RAPD (Random Amplified Polymorphic DNA) and in addition shows higher level of polymorphism, reproducibility and cost-effectiveness per polymorphism. ISSR–PCR has been used in genetic diversity studies in several crop plants. In rice, ISSR–PCR has been used for the analysis of microsatellite motif frequency and fingerprinting of varieties[5,6], determining phylogenetic relationships among *Oryza* species[7,8], in distinguishing basmati rice varieties[9] as markers for restorer genes[10], and in studies on comparing effectiveness of different molecular markers[11–13].

The usefulness of multi locus ISSR–PCR markers in determining diversity within landraces of rice has not been previously investigated. The few studies on comparison of genetic diversity in landraces and modern cultivars have used single locus microsatellite (SSR simple sequence repeat) markers[14–16] or AFLP (amplified fragment length polymorphism) markers[2]. The choice of primers (motif, repeat length and anchor) used in ISSR amplification is critical to obtaining high levels of polymorphism, which is also representative of the whole genome[5,6].

The value of a primer used in ISSR–PCR to reveal polymorphism and distinguish between genotypes has been measured by two indices – polymorphism information content (PIC)[17] and resolving power (Rp)[18]. While the former has been widely used to compare different kinds (diallelic and multi allelic) of molecular markers, the latter has been used to compare primers for ISSR–PCR in fingerprinting of potato[18] and assessing diversity in lupin[19] and barley cultivars[20]. The utility of the two indices has not been demonstrated in a large dataset, including landraces and varieties of rice earlier.

The objectives of this study were therefore: (i) to assess the extent of genetic diversity in Indian landraces of rice compared to improved cultivars and other accessions using $(AG)_n$ and $(GA)_n$ based primers for ISSR amplification; (ii) to compare the level of polymorphism revealed by each primer, and between landraces and varieties, and (iii) to develop a rational basis for the choice of informative primers using PIC, Rp and mean genetic similarity. This would help develop a practical and cost-effective approach for analysis

*For correspondence. (e-mail: sarlan@operamail.com)

Table 1. Accessions used, total number of bands amplified and mean genetic similarity based on data of 14 primers

Sl. no.	Accession	Group	No. of bands	Mean genetic similarity	Sl. no.	Accession	Group	No. of bands	Mean genetic similarity
1	INRC 9205	L	64	0.64	45	INRC 9864	L	69	0.59
2	INRC 9244	L	66	0.61	46	INRC 9882	L	60	0.57
3	INRC 9246	L	61	0.61	47	INRC 10189	L	62	0.56
4	INRC 9263	L	62	0.58	48	INRC 10683	L	56	0.57
5	INRC 9358	L	69	0.63	49	INRC 10763	L	66	0.6
6	INRC 9377	L	77	0.61	50	INRC 10808	L	66	0.62
7	INRC 9380	L	70	0.61	51	Anaikomban	AL	71	0.62
8	INRC 9382	L	63	0.63	52	Latisail	AL	72	0.63
9	INRC 9524	L	57	0.6	53	Mudgo	AL	73	0.57
10	INRC 9525	L	56	0.57	54	PTB18	AL	60	0.6
11	INRC 9834	L	63	0.6	55	PTB21	AL	66	0.59
12	INRC 9884	L	67	0.59	56	Eswarakora	AL	74	0.54
13	INRC 9887	L	63	0.62	57	Gharbharan	LV	61	0.59
14	INRC 9932	L	68	0.64	58	Hathipanjira	LV	66	0.63
15	INRC 9968	L	65	0.62	59	Nagpur 27	LV	66	0.57
16	INRC 9991	L	64	0.63	60	Sabita	LV	68	0.6
17	INRC 9996	L	67	0.62	61	Dular	LV	75	0.53
18	INRC 9999	L	58	0.6	62	IR 8	V	66	0.56
19	INRC 10059	L	61	0.63	63	TN 1	V	65	0.59
20	INRC 10060	L	62	0.62	64	Mahsuri	V	70	0.59
21	INRC 10062	L	59	0.62	65	Jaya	V	65	0.57
22	INRC 10063	L	59	0.62	66	TKM 6	V	61	0.58
23	INRC 10064	L	54	0.57	67	Basmati 370	LV	68	0.55
24	INRC 10066	L	67	0.67	68	T 141	V	67	0.61
25	INRC 10067	L	62	0.65	69	Lunisree	V	86	0.62
26	INRC 10068	L	63	0.64	70	Swarna	V	72	0.62
27	INRC 10069	L	58	0.61	71	Sambamahsuri	V	64	0.59
28	INRC 10192	L*	62	0.6	72	Utkalprabha	V	57	0.52
29	INRC 10194	L*	60	0.6	73	IR 64	V	72	0.57
30	INRC 10196	L*	70	0.57	74	MTU 1001	V	58	0.51
31	INRC 10199	L*	46	0.5	75	CORH 2	H	65	0.59
32	INRC 10204	L*	74	0.6	76	KRH 2	H	77	0.62
33	INRC 10349	L*	62	0.57	77	Sahayadri	H	76	0.6
34	INRC 10418	L	72	0.62	78	Pusa 1266	V	61	0.56
35	INRC 10422	L	58	0.61	79	NPT 6	J	73	0.59
36	INRC 10447	L	54	0.55	80	NPT 16	J	56	0.55
37	INRC 10663	L	48	0.5	81	Taipei 309	J	59	0.54
38	INRC 10719	L	64	0.64	82	Wu 10B	J	56	0.54
39	INRC 10722	L	65	0.59	83	IRGC 105308	R	66	0.47
40	INRC 10744	L	56	0.59	84	IRGC 105325	R	72	0.49
41	INRC 10760	L	64	0.58	85	IC 21009	N	64	0.49
42	INRC 10767	L	72	0.64	86	IC 21022	N	47	0.44
43	INRC 10833	L	52	0.53					
44	INRC 9216	L	56	0.56		Total no. of bands		5514	

AL, Ancestral landrace; H, Hybrid; IC, Indian collection; IRGC, International Rice Germplasm Collection; J, Japonica; L, Landrace (Raipur collection); L*, Landrace (Assam rice collection); LV, Landrace variety; V, Improved variety; R, *O. rufipogon*; N, *O. nivara*.

of diversity in landraces and varieties of rice using a few well-chosen primers in ISSR amplification.

Materials and methods

Plant material

The accessions used in this study (50 landraces, 20 high-yielding varieties (cultivars) and 16 other accessions) are listed in Table 1, and the agronomic importance of 36 of these accessions is given in Appendix 1. The 20 improved varieties (accessions 59 to 78) included three released three-line hybrids (H) and four popular landrace varieties. The 16 other accessions included six ancestral landraces, two old landrace varieties, two *japonica* varieties, two accessions of New Plant Type (NPT) and two accessions each of *O. nivara* and *O. rufipogon*, the wild progenitors of *O. sativa*. Ancestral landrace (AL) refers to landraces used by breeders for the development of improved varieties. Landrace variety (LV) refers to popular traditional varieties which are either landraces or cultivars based on pure line selection[21]. Variety (V) refers to high-yielding improved varieties developed by breeders. The two

Table 2. Number of bands, polymorphism information content (PIC), resolving power (Rp) and mean genetic similarity revealed by each of the 14 primers. Data are based on 5514 bands at 220 loci in 86 accessions

Primer number UBC	Sequence (5′–3′)	Annealing temperature (°C)	No. of band positions (loci)	Size range (bp)	Mean no. of bands/accession	Mean no. of bands/locus	PIC	Rp	Mean genetic similarity
$(AG)_n$ based primers									
807	$(AG)_8T$	45	11	400–1900	3.1	24.27	0.84	3.42	0.55
808	$(AG)_8C$	47	15	300–1300	5.91	33.87	0.74	5.63	0.62
809	$(AG)_8G$	47	11	500–2000	4.57	35.73	0.63	1.23	0.86
834	$(AG)_8YT$	48	17	470–1900	3.59	18.17	0.92	6.07	0.33
835	$(AG)_8YC$	50	19	350–2000	5.08	23.00	0.83	5.14	0.62
836	$(AG)_8YA$	50	15	400–1350	3.72	21.33	0.88	5.40	0.45
884	$HBH(AG)_7$	45.5	18	450–1500	6.52	31.17	0.78	6.95	0.59
Total			106		32.49	187.54	5.62	33.84	4.02
$(GA)_n$ based primers									
810	$(GA)_8T$	45	13	500–2000	4.25	28.15	0.81	4.79	0.56
811	$(GA)_8C$	47	12	560–2500	4.36	31.25	0.79	5.65	0.51
812	$(GA)_8A$	45	14	400–1350	4.71	28.93	0.82	6.95	0.51
840	$(GA)_8YT$	48	23	300–2000	3.91	14.61	0.89	3.07	0.62
841	$(GA)_8YC$	45	13	400–1400	4.39	29.08	0.80	5.21	0.58
842	$(GA)_8YG$	50	23	300–1300	5.15	19.26	0.89	7.33	0.47
885	$BHB(GA)_7$	46.5	16	300–1500	4.84	26.00	0.83	5.81	0.54
Total			114		31.61	177.28	5.83	38.81	3.79

Y, Any pyrimidine; H, A, T, C; B, G, C, T.

NPT accessions and four Indian wild rice accessions were obtained from IRRI, Philippines and the other 80 accessions were from the collection at Directorate of Rice Research, Hyderabad. The experiments were conducted during 1999–2000.

DNA extraction, PCR amplification and electrophoresis

A bulk sample of leaves from ten plants of each accession was collected and genomic DNA was extracted using the PVP method[22]. The quality and quantity of DNA was estimated using a UV spectrophotometer (DV 650, USA) and also checked visually by ethidium bromide staining of 0.8% agarose gels. The DNA samples were diluted to 10 ng/µl.

Fourteen primers based on (AG) and (GA) repeats were selected for the study (Table 2). These (set # 9 of 100 UBC primers) were obtained from the University of British Columbia, Vancouver, Canada. PCR reaction was carried out in a DNA thermal cycler (Gene Amp, PCR system 9700 PE, Applied Biosystems, USA) using a single primer in each reaction. Each 25 µl reaction mixture contained 10 mM tris-HCl pH9, 50 mM KCl, 1.5 mM $MgCl_2$, 1 unit of Taq DNA polymerase, 200 µM each of the four dNTPs (Bangalore Genei Pvt Ltd, India), 2% formamide, 0.5 mM spermidine, 0.2 µM primer and 30 ng template DNA. PCR amplification conditions were as follows: initial extended step of denaturation at 94°C for 5 min followed by 44 cycles of denaturing at 94°C for 1 min, primer annealing at 50°C for 45 s, elongation at 72°C for 2 min, followed by an extended elongation step at 72°C for 5 min. The annealing temperature was adjusted according to the Tm of each primer (Table 2). Reaction products were mixed with 2 µl 6X indicator dye (0.25% bromophenol blue and 40% sucrose, w/v) and resolved on 2% agarose gel containing 0.5 µg ethidium bromide using 0.5 X tris-borate buffer, pH 8.0. Gels were run until the indicator dye ran 10 cm from the well (5 h at 90 mA). The gels were photographed and documented using Image Master VDS (Amersham Pharmacia).

Scoring and data analysis

Only those bands that were clear and reproducible were scored for data analysis. The repeatability of the scored bands was ensured by running two PCR amplifications for the primers with 5′ anchors and for some accessions using 3′ anchors, and scoring of bands by two individuals. Bands were scored as diallelic markers, present (1) or absent (zero) at each band position (considered locus not in the strict sense) for the 86 samples for each primer. Molecular weight of the bands was estimated using λ *Eco*RI-*Hin*dIII double digest or 100 bp DNA ladder (Bangalore Genie) as standard. Bands with the same molecular weight and mobility were treated as identical fragments. The total number of bands, distribution of bands across accessions, number of polymorphic bands in a set of accessions, and average number of bands per primer were calculated. The value of each of the 14 primers was assessed using two indices; PIC, which is the same as diversity index (DI)[23,24], and Rp (Prevost and Wilkinson)[18].

PIC or DI was estimated as PIC = $\Sigma(1 - p_i^2)/n$, where n is the number of band positions analysed in the set of accessions, p_i is the frequency of the ith pattern (band position, since each band is considered as one diallelic locus in ISSRs which are dominant markers).

The ability of the 14 primers to distinguish between accessions was assessed by calculating their resolving power as, $Rp = \Sigma I_b$, where I_b is band informativeness, and $I_b = 1 - (2 \times 0.5 - p_i)$, where p_i is the proportion of accessions containing band i (Prevost and Wilkinson[18]).

The PIC and Rp values were calculated for the 86 accessions as a whole and also for the following five groups within them: (i) 50 landraces, (ii) 20 improved varieties, (iii) six ancestral landraces and two old landrace varieties, (iv) two *japonica* varieties and two NPTs, and (v) four wild accessions. The comparison was essentially between the landraces and varieties. Loci monomorphic in each set were excluded from the analysis. The proportion of polymorphic markers in each of the five sets of accessions was calculated. The H_{av} (p) values were used to compare expected heterozygosity within each of the five sets. A correction for sample size was introduced by multiplying $H_{av}(p)$ by $n/n - 1$, where n is the sample size[25].

Similarity matrix and cluster analysis

All the numerical taxonomic analyses were conducted using the software package NTSyS-pc 2.0 (Numerical Taxonomy and Multivariate Analysis System)[26]. Fourteen similarity matrices were constructed for data from each primer using Dice Similarity Coefficient values for each of the 3655 pairwise comparisons between accessions using the routine SIMQUAL (similarity for qualitative data). The average similarity coefficient (= mean genetic similarity) for each primer was calculated from each of the 14 matrices. Pairwise correlation coefficient between any two of the five characteristics – mean number of bands per accession, mean number of bands per marker (in all accessions), PIC, Rp and mean genetic similarity was calculated using CORREL function in Excel.

The 0/1 matrix data from all the 14 primers was pooled to get one similarity matrix used for deriving relationships among and within accessions. The matrix of similarity coefficients was subjected to UPGMA (unweighted pair-group method with arithmetic averages) analysis to generate a dendrogram using SAHN (Sequential, Agglomerative, Heirarchical and Nested clustering) and Tree plot routines of NTSyS-pc. The similarity matrices derived from data of different subsets (AG-based primers vs GA-based primers; 3′ one nucleotide anchored primers (3′ N) vs 3′ dinucleotide anchored primers (3′ NN) vs 5′ anchored primers (5′ NNN)) were compared using the Mantel matrix correspondence test (MxComp module of NTSyS). Bootstrap analysis was performed using 1000 permutations in Winboot. Bootstrap values over 50 are considered significant and mentioned on the dendrogram.

Results

In all, 5514 ISSR bands were scored in all accessions at 220 band positions amplified (Tables 1 and 2). It was only the very faint bands that were inconsistent in replicates. The primers with 3′ anchors gave clear bands with most of the accessions. The size of the amplified bands ranged from 300 to 2500 bp, with more bands around 1000 bp. The mean number of bands per primer and the mean number of bands amplified per accession by each primer are also given in Table 2. At 46 of the 220 band positions, bands were present in 34 to 69 of the accessions. The primers 808, 835 and 884 among the $(AG)_n$ primers and 812, 842 and 885 among the $(GA)_n$ primers, each generated at least five of the 46 informative band positions. There were 18 band positions where bands were present in more than 70 accessions.

Unique and rare loci

Among the 220 band positions amplified, 20 were unique (present in only one accession out of 86). Thirteen landraces, including three from Assam rice collection (ARC), five varieties (Mudgo, Dular, Mahsuri, IR 64 and Pusa 1266) and one *O. nivara* accession had unique bands. INRC 10663 had two unique bands. Positions at which bands were present in only two out of 86 accessions were termed rare. There were 22 such rare band positions distributed in 28 accessions, of which 15 were landraces, including four from ARC, ten varieties (Mudgo, Eswarakora, Gharbharan, Sabita, IR 64, Dular, Sahayadri, Pusa 1266, Mahsuri and NPT 16) and three wild accessions. Unique bands were amplified by 11 of the 14 primers used. Primers 807, 884 and 840 did not amplify a unique band. Similarly, rare bands were amplified by 13 of the 14 primers (841 was the exception). Primer 835 amplified two unique and four rare bands. Among the 22 rare bands, six were present only in landraces, six were shared between landraces and varieties, four were present only in varieties, three were shared between wild species and varieties, and three were shared between wild species and landraces.

Taken together, the seven primers based on $(AG)_n$ generated about the same number of loci (106) as those generated (114) by the seven primers based on $(GA)_n$. Among the 3′ anchored primers, those with a dinucleotide anchor produced more markers (mean 18.3) than those with one nucleotide as anchor (mean 12.6). The band profile obtained using a dinucleotide anchored primer 835 $(AG)_8YA$ in 29 landraces and 29 other accessions is shown in Figure 1. The 5′ anchored primers generated 17 band positions on an average. If all the primers are considered together, the *japonica* group had the least total number of bands/accession (61), the varieties and ancestral landraces the highest (67.9 and 67.8 respectively), and landraces and wild accessions were intermediate.

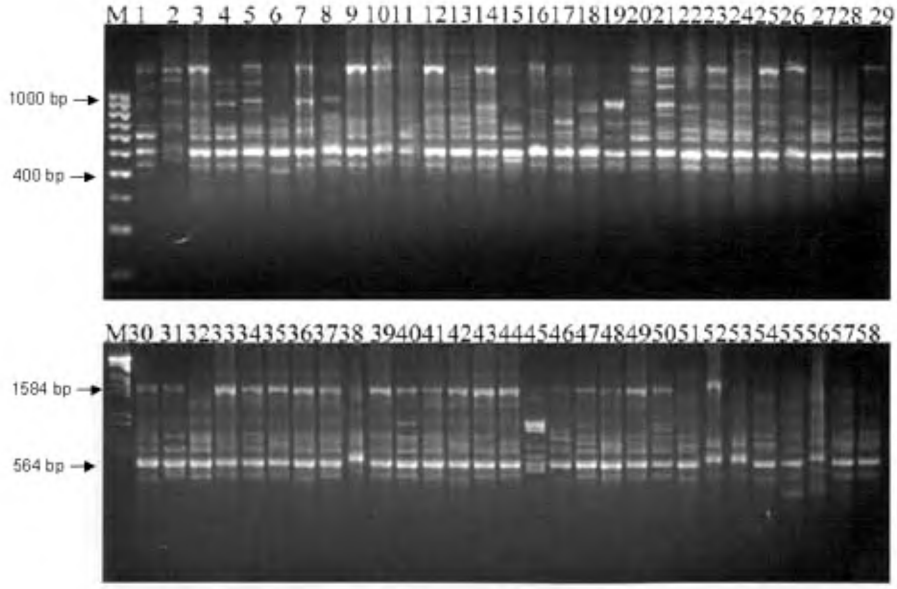

Figure 1. ISSR–PCR amplification profiles of 58 accessions using primer 835. Lanes 1–29 correspond to accessions 1–29 (landraces) respectively, of Table 1 and lanes 30–58 correspond to accessions 58 to 86 (varieties and wild species) respectively, of Table 1. Marker lane at extreme left is 100 base pair DNA ladder (top panel) and lambda DNA/*Eco*R1 + *Hin*dIII double digest (bottom panel).

Comparative ability of different primers to detect diversity in 86 accessions

The mean PIC value of all 14 primers was 0.82. Primer 809 showed the lowest PIC value (0.63). Primer 834 had the highest PIC value (0.92) and therefore detects the highest level of polymorphism (Table 2). The mean PIC value of the six primers with a dinucleotide anchor (0.87) was higher than that of the six primers with a single nucleotide as anchor (0.77). The mean Rp value for the 14 primers was 5.2. Primer 809 had the lowest Rp (1.23) and primer 842 the highest Rp (7.33). In general, $(GA)_n$ primers had a higher resolving power than $(AG)_n$ primers. The mean Rp value of the seven $(GA)_n$ primers was 5.54 compared to 4.83 of $(AG)_n$ primers. Primer 834 had the highest values of PIC and Rp among the $(AG)_n$ primers and 842 had the highest values of PIC and Rp among the $(GA)_n$ primers. Thus these two primers can be considered highly useful in terms of polymorphism level revealed as well as resolving power. The mean genetic similarity values among the 86 accessions revealed by these two primers were also the lowest (0.33 and 0.47 respectively).

Both PIC and Rp were negatively correlated to the mean genetic similarity, PIC being more highly correlated than Rp (Table 3). The mean number of bands per band position was highly negatively correlated to PIC and positively to mean genetic similarity. The mean genetic similarity value was the least with primer 834, which showed the highest PIC value and the mean genetic similarity was highest with the least informative primer 809 which had the lowest PIC and Rp (Table 2).

Table 3. Correlation between mean number of bands, PIC, Rp and mean genetic similarity

	Mean no. of bands/ accession	Mean no. of bands/ locus	PIC	Rp	Mean genetic similarity
	a	b	c	d	e
a	1				
b	0.48	1			
c	–0.40	–0.9**	1		
d	0.44	–0.11	0.45	1	
e	0.29	0.55*	–0.83**	–0.69**	1

*Significant at 5%; **Significant at 1%.

Comparison of informativeness of primers in landraces and varieties

The PIC and Rp values of each primer in each of the five sets into which the 86 accessions were grouped, are listed in Table 4. Primers which showed a high PIC and Rp in landraces showed much lower values in varieties. Primers 834 and 835 among the AG-based primers and 840 and 842 among the GA-based primers had the highest Rp and PIC value for landraces. However, these four primers did not show high values when the group of improved varieties was considered. Instead, in varieties the highest PIC and Rp values were shown by 808 and 884 respectively among the AG-based primers, and 811 and 842 respectively, among GA-based primers. Based on PIC and Rp values, the four most informative primers useful for studying diversity in landraces were 834, 842, 835 and 812. Significantly, the four most informative primers for varieties

Table 4. PIC and Rp (in parenthesis) of each primer in five groups of accessions

Primer	Landrace 50	Variety 20	Ancestral landrace 8	*Japonica* 4	Wild accession 4
807	0.78 (1.68)	0.64 (3.4)	0.72 (5.0)	0.27 (2.5)	0.10 (1.0)
808	0.53 (4.84)	0.78 (5.6)	0.41 (5.0)	0.26 (1.5)	0.37 (4.0)
809	0.53 (1.04)	0.39 (0.8)	0.37 (2.7)	0 (0)	0.09 (1.0)
834	0.84 (6.76)	0.51 (3.2)	0.24 (2.33)	0.23 (2.5)	0.12 (2.5)
835	0.89 (5.16)	0.36 (3.6)	0.30 (3.33)	0.27 (3.5)	0.24 (3.0)
836	0.70 (5.04)	0.43 (4.1)	0.49 (4.01)	0.27 (2.0)	0.32 (4.0)
884	0.70 (4.6)	0.64 (7.0)	0.40 (6.0)	0.43 (4.5)	0.68 (9.5)
810	0.41 (3.72)	0.41 (4.1)	0.30 (4.0)	0.32 (4.5)	0.57 (5.0)
811	0.62 (4.12)	0.68 (5.5)	0.55 (6.33)	0.46 (5.0)	0.18 (2.0)
812	0.60 (4.88)	0.46 (3.4)	0.36 (5.0)	0.43 (4.5)	0.48 (6.0)
840	0.76 (2.88)	0.32 (3.1)	0 (0)	0 (0)	0.31 (5.0)
841	0.60 (3.88)	0.59 (3.4)	0.32 (3.0)	0.17 (1.0)	0.31 (2.5)
842	0.80 (6.72)	0.54 (4.6)	0.34 (5.0)	0.22 (3.5)	0.41 (9.5)
885	0.70 (4.0)	0.30 (4.1)	0.79 (7.66)	0.33 (5.0)	0.27 (4.5)

Table 5. Mean genetic similarity and range (in parenthesis) among and within five groups of accessions

Accession group	No. of accessions	Landrace	Improved variety	Ancestral landrace	*Japonica*	Wild accession
Landrace	50	**0.64** (0.49–0.84)				
Variety	20	**0.54** (0.38–0.70)	**0.67** (0.52–0.88)			
Ancestral landrace	8	**0.61** (0.41–0.79)	**0.57** (0.42–0.70)	**0.67** (0.61– 0.78)		
Japonica	4	**0.52** (0.39–0.66)	**0.62** (0.46–0.75)	**0.52** (0.42–0.63)	**0.72** (0.67–0.77)	
Wild accession	4	**0.44** (0.28–0.55)	**0.52** (0.42–0.62)	**0.45** (0.33–0.53)	**0.58** (0.49–0.66)	**0.6** (0.44–0.74)

were different – 808, 884, 811 and 807. It is noteworthy that among the seven $(GA)_n$ primers used in varieties, primer 811 generated the least number of bands but had the highest Rp and PIC values. This was also true for the *japonica* group. Primer 885 showed the highest values for both PIC and Rp in the ancestral landraces group and 884 in the wild accessions group.

Diversity among accessions based on genetic similarity

The Dice Similarity Coefficient among all 3655 pairwise comparisons of 86 accessions based on data of all the 14 primers ranged from 0.28 to 0.90, with a mean genetic similarity of 0.59. The similarity matrix is available on request. The mean genetic similarity of each of the 86 accessions is listed in Table 1. The range and mean similarity value of all pairwise comparisons within and between five groups of accessions is shown in Table 5. The landraces had a mean genetic similarity of 0.64 and the varieties 0.67. If the landrace varieties Sabita, Nagpur 27, Dular and Basmati 370 were excluded from the group of improved varieties, the mean genetic similarity of the remaining varieties rose to 0.71. The mean genetic similarity of the four wild accessions was the least (0.44), indicating their high diversity and that among *japonicas* the highest (0.72), indicating low diversity. The average expected heterozygosity of landraces was 0.78 compared to 0.65 of varieties.

The landrace accessions INRC 10833 and 10199 were the most distant from the 20 varieties considered together or from Jaya and IR64, each considered separately. Accessions 10062, 10066 and 10192 which were moderately distant from Jaya and IR64 have been used in crosses to help identify, map and introgress novel QTL alleles for yield, and recombinant inbred lines are being developed (Directorate of Rice Research, Annual Reports 2000–02).

Clustering among accessions

A dendrogram derived from the whole data showed two distinct clusters (Figure 2). Cluster 1 consisted of all *O. sativa* accessions and cluster 2 the two wild species, *O. nivara* and *O. rufipogon*. Two landraces, 10199 from ARC and 10663 from Raipur Collection clustered together, but

separately from the other *O. sativa* accessions. Cluster 1 forked into two sub-clusters – one consisting of all landraces, including six ancestral landraces and two old landrace varieties and the other consisting of all the 20 improved varieties and the two accessions each of *japonica* and NPT. The two landrace accessions 10066 and 10067 could be considered very closely related or duplicates, as they showed more than 90% similarity. It is interesting to note that four out of six accessions from ARC clustered together, indicating a geographical bias for genetic similarity.

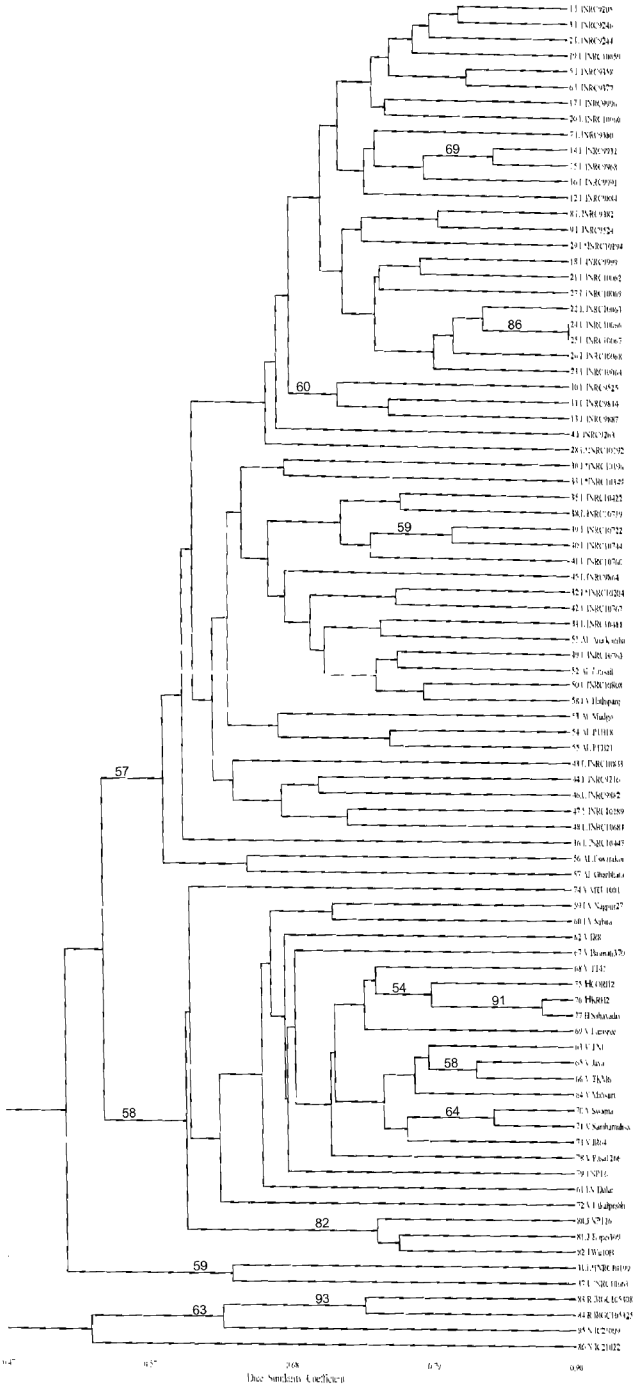

Figure 2. UPGMA dendrogram showing clustering of 86 rice accessions. It is derived from Dice coefficient matrix of 3655 pairwise comparisons using ISSR–PCR data (5514 bands at 220 band positions) from 14 primers. Significant bootstrap values are mentioned on the nodes.

Among the varieties, differentiation of *indica* and *japonica* type was clear, they segregated into two clusters at about 60% similarity value. Popular varieties such as Jaya, Mahsuri, IR 64, Sambamahsuri and Swarna showed more than 75% similarity. The three hybrids (H) clustered together very closely, KRH2 and Sahyadri being 88% similar. The two varieties, Swarna and Sambamahsuri also showed a high similarity of 83%. One of the NPT clustered with the two *japonica* varieties at 0.74 similarity. The four popular landrace varieties which are currently cultivated (Dular, Nagpur 27, Sabita and Basmati 370) and NPT-6 clustered loosely with the other high-yielding improved varieties. It is noteworthy that Dular belongs to isozyme group II and Basmati 370 to group V of Glaszmann's classification, while most *indica* varieties are considered to belong to group I.

Fingerprinting

Dendrogram derived from data of primer 842 alone could distinguish 76 out of 86 accessions uniquely and in addition, members of each of the five pairs could not be distinguished from each other (tree not shown). Thus, if there was only one of each pair, 81 accessions could be uniquely identified. If data from primer 835 were also included, the tree separated each of the 86 accessions except KRH2 and Sahyadri, which have a common parent IR 58025A (tree not shown). Thus, it is possible to fingerprint landraces and varieties using just one or two informative primers, such as 835 and 842 in ISSR–PCR amplification.

The similarity matrix generated by the use of seven $(AG)_n$-based primers was not significantly correlated ($r = 0.37$) with that based on seven $(GA)_n$-based primers (data not shown). A dendrogram based on GA data clearly delineated the two wild species first (separated at 0.40), then varieties and landraces, followed by *indica* and *japonica*, as did the whole data. On the other hand, the dendrogram from AG data delineated a few landraces first, then two groups, one containing varieties clubbed with wild species and the other containing landraces. Unlike traditional classification, the wild species *O. rufipogon* and *O. sativa* subsp-*japonica* clustered together. Along with the *O. nivara* cluster, they then joined the cluster of varieties. The AG dendrogram however delineated Basmati 370 (GrV) and Dular (Gr II) from the other varieties. The landrace INRC 10199 was distinct from the other landraces in both the dendrograms. The similarity matrices based on data of 3′ N-anchored primers, 3′ NN-anchored primers and 5′ NNN-anchored primers were also not significantly correlated with each other. The similarity matrices and dendrograms are available on request.

Discussion

Genetic relationship among accessions

ISSR polymorphism obtained using anchored $(AG)_n$ and $(GA)_n$ primers reflected the conventional genetic and taxonomic relationship of a large set of rice germplasm. The clear delineation of wild species, landraces and *indica* and *japonica* varieties indicates that the markers were associated with genomic regions where specific and sub-specific differentiation has occurred. The average heterozygosity of landraces was higher than varieties as reported also in earlier studies using RAPD and SSR markers[12,14,16,27]. A geographical bias in clustering was also reported for *O. nivara*[8]. The robustness of clustering was evident as the three hybrids KRH2, CORH2 and Sahayadri, all of which have a common parent IR 58025A (a widely used cms line), grouped together at about 80% similarity. Likewise, Swarna and Sambamahsuri grouped together, Mahsuri being a common parent. The two wild species grouped separately as expected. In our study, NPT6 and NPT16 were considered as *japonica*, as they have a *japonica* plant-type, but these have been developed using introgressions from *indica*. This explains the grouping of NPT6 with other *indica* varieties. Basmati 370 (isozyme grpV) was reported to group separately in earlier studies also[6,8]. The mean genetic similarity among the 20 varieties was 0.71, when only the popularly grown semi-dwarf high-yielding varieties were considered. Davierwala *et al.*[13] and Singh *et al.*[28] reported values of 0.70 and 0.80 respectively, in Indian elite varieties. There is thus a need to widen the genetic base of Indian cultivars.

Choice of primers

The choice of primers is crucial in studies on diversity using ISSR amplification[6,18,29]. A primer should amplify regions spanning the whole genome fairly evenly, to be representative of the whole genome and be useful for diversity analysis. In addition, it should target regions where polymorphisms are most likely to exist. Different SSR motifs vary in their frequency and number of repeats in different species[30,31].

Both AG and GA repeats proved informative in studies on different *Oryza* species[7,32], elite rice varieties[13], in other genera[18,29,33–35], and even when they were not used as primers but were the amplified products as in SSR amplification in rice[15,16,36,37]. AG repeats are more often perfect repeats, i.e. have no intervening repeats[30]. They are more abundant in plants than in animals[31] and the highest polymorphism was obtained with (AG) repeats[29]. GA repeats are the most abundant SSRs in rice[38,39] and show maximum diversity[40]. GA repeats (GAGA elements) control gene expression in animals and plants[41–43]. The GA motif confers a remarkable conformational polymorphism on DNA[44] and enhances homologous DNA recombinantion in SV 40 minichromosomes[45].

The primer 884 $HBH(AG)_7$ generated the highest number of bands (561) in the set of 86 accessions as a whole. The 5′ trinucleotide anchor is degenerate. Such anchoring ensures that the primer binds to many sites of 5′ (CT) repeats, which may have different sequences at their 3′-end. This explains the high number of bands obtained. It is interesting to note that this primer generated the highest number of bands in four out of the five sets of accessions. However, the most informative primer was 842 $(GA)_8YG$. The similarity matrices from data of 3′- and 5′-anchored primers were not significantly correlated, contrary to an earlier report[6]. Primers 809 and 840 detected less polymorphism, but amplified unique and rare bands. Unique ISSR amplicons have been used to design species/accession-specific primers[4].

GA repeats are more in varieties than in landraces

The number of bands produced by an ISSR primer with a given microsatellite repeat provides an estimate of the motif abundance and clustering in a genome[6]. The presence of more bands in varieties than in landraces only when $(GA)_n$-based primers were used, clearly indicates that clustered GA repeats occur more often in varieties than in landraces. This has not been reported earlier. It is interesting to note that in *Arabidopsis* $(GA/TC)_8$ repeats occur particularly within 1500 bp upstream of gene start codons included in some homeodomain genes of different classes, and a transcription factor that binds to $(GA)_8$ has been reported in barley[43]. Ramakrishna *et al.*[37] compared wild and cultivated species and reported that the expansion of GA repeats at one locus (RM 122) was associated with domestication.

PIC, Rp and mean genetic similarity

PIC is an index of diversity in a given set of accessions as deciphered by a given primer. The primer 834 $(AG)_8YT$, which showed the highest PIC (0.92) in this study, was also the only primer out of 100 which could distinguish a maintainer line from a restorer line in rice[10], amplified many putative species-specific bands in *Oryza*[7], distinguished citrus cultivars[46] and along with 841, it could help distinguish 34 potato genotypes[18]. Primer 834 also had a high Rp value of 6.07.

Rp of a primer is strongly correlated with its ability to distinguish genotypes, but provides no information on the ability of a primer to reflect genetic or taxonomic relationship of a group of genotypes[18,20]. The present work shows that PIC revealed by a primer in a given set of germplasm is also important. Both PIC and Rp are important in determining the value of a primer to distinguish genotypes.

Appendix 1. Details of agronomic importance of 36 accessions (51–86 in Table 1)

Anaikomban	Tamil Nadu, high quality rice, res. to blast, WBPH
Latisail	West Bengal, wide adaptability, res. to RTV, parent of Peta, good quality rice
Mudgo	Res. to BPH, a parent of IR varieties
Ptb-18	(Eravapandi) Kerala, high quality rice, res. to stem borer, GM, RTV, BPH, GLH, Isozyme Gr. I
Ptb-21	(Thekkan) Kerala, high quality rice, res. to stem borer, RTV, GM, BPH, Isozyme Gr. I
Eswarakora	Res. to GM
Gharbharan	Uttar Pradesh, adapted to rainfed, upland, sturdy culm, deep root, high biomass
Hathipanjira	Wide adaptability, photoinsensitive
Nagpur 27	High number of spikelets
Sabita	1986, West Bengal, flood-tolerant, 30–60 cm water depth, pure line selection from Boyan, long slender grain, national check for semi-deep water
Dular	Wide compatible variety, Isozyme Gr. II
IR 8	1966 (Peta/DGWG) first semi-dwarf, high-yielding var., Isozyme Gr. I, long bold grain
TN 1	(DGWG/TYC), main source of semi-dwarfism
Mahsuri	1972, (Mayang Ebos 80/Taichung 65) *indica–japonica* hybrid derivative widely grown variety in several countries, high quality rice, medium slender grain, photoperiod insensitive
Jaya	1968 var. from All India Coordinated Rice Improvement Project (TN1/T141), long, bold grain, res. to blast, national check for irrigated medium
TKM6	Tamil Nadu (GEB 24/Co18), high quality rice, res. to yellow stem borer, leaf folder, bacterial blight (xa 4), stripe borer, RTV, source of resistance in IRRI vars IR 20 to IR 36
Basmati 370	Pre-1960 var. still grown, high quality rice, aromatic, long slender grain, Isozyme Gr. V
T 141	Pre-1960 var. (Soruchinamalai), still grown, good agronomic base, high quality rice, photosensitive
Lunishree	West Bengal, salt-tolerant, mutant of Nonasail, drought-resistant, delayed senescence
Swarna	1982 (Vasishtha/Mahsuri), rainfed shallow land, responds to low N, medium slender grain
Sambamahsuri	1986 (GEB 24/TN1/Mahsuri), rainfed shallow land, medium slender grain, resistant to blast
Utkalprabha	1983 Orissa (Waikokku/CR1014), semideep and deepwater, medium slender grain
IR 64	1992 widely grown rainfed shallow land, national check for slender grain, Isozyme Gr. I
MTU 1001	(Vijeta), Andhra Pradesh, grown in coastal areas, bold grain, preferred for parboiled rice
CORH 2	1998 hybrid, Tamil Nadu (IR 58025A/C20R), medium bold grain
KRH 2	1996 hybrid, Karnataka (IR58025A/KMR3), long bold grain, high yield at many locations
Sahayadri	1998 hybrid, Maharashtra (IR58025A/BR827-35-3-1-IR), long slender grain
Pusa 1266	New plant type variety from India, semi-dwarf, high yield, *japonicas* used in its development
NPT 6	New plant type from IRRI (IR 65600-38-1-2-1), *japonica*-like
NPT 16	New plant type from IRRI (IR 66160-5-2-3-2), *japonica*-like
T 309	*Japonica* variety Taipei 309
Wu 10b	*Japonica* variety
R 64	*O. rufipogon* (IRGC acc no. 105308), Kerala, India, local name Vranellu
R 65	*O. rufipogon* (IRGC acc no. 105325), Kerala, India
W 107	*O. nivara* (IC 21009), India
W 106	*O. nivara* (IC 21022), India

BPH, Brown plant hopper; GLH, Green leaf hopper; GM, Gall midge; Res., Resistant; RTV, Rice tungro virus; WBPH, White backed plant hopper.

If mean genetic similarity value among all accessions as determined by a primer is high, it implies that the primer cannot distinguish accessions as well as a primer which gives a lower mean genetic similarity value. This is well illustrated in our studies using two examples of primers whose PIC was the same but Rp varied (840, 842) or whose Rp was same but PIC varied (884, 812). This was correlated to mean genetic similarity value determined by each primer. In the first case both primers, 840$(GA)_8$YT and 842 $(GA)_8$YG, generated 23 bands each, detected the same level of polymorphism (0.89), but the Rp of primer 840 was 3.07 and that of primer 842 was 7.33, the highest Rp among all the 14 primers studied (Table 2). The mean genetic similarity value revealed by primer 840 was accordingly much higher (0.62) than that of 842 (0.47), indicating that resolving power of a primer also determines the extent of genetic similarity revealed among accessions. Accordingly, the dendrogram based on data of primer 840 could resolve only 52 genotypes compared to 81 distinguished in the dendrogram based on data of primer 842. Similarly, in the second case, though the Rp values of 884 and 812 were the same, the mean genetic similarity shown by 812 was lower (0.51) because PIC revealed by it was higher.

The utility of the index Rp in determining informativeness of ISSR–PCR primers is demonstrated in a large set of accessions. This study emphasizes that the value of primers for diversity analysis needs to be estimated based on both PIC and Rp values. More importantly, it provides evidence that the values of both these indices vary with markedly different germplasm, i.e. landraces or varieties or wild accessions. Also, the distinctness of landraces

from varieties is shown at the molecular level. It is thus demonstrated that using a few, well chosen, highly informative $(AG)_n$ and $(GA)_n$ based primers, the genetic diversity among large germplasm collections of rice can be determined easily in a cost-effective and robust manner. Tremendous scope exists to widen the genetic base of popular varieties using landraces. Accessions 10062, 10066, 10192, which were moderately distant from Jaya and IR64 were used in crosses to identify, map and introgress novel QTL alleles for yield and related traits. Recombinant inbred lines are being developed from these crosses (Directorate of Rice Research, Annual Reports 2000–02).

1. Glaszmann, J. C., Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.*, 1987, **74**, 21–30.
2. Prashanth, S. R., Parani, M., Mohanty, B. P., Talame, V., Tuberosa, R. and Parida, A., Genetic diversity in cultivars and landraces of *Oryza sativa* subsp. *indica* as revealed by AFLP markers. *Genome*, 2002, **45**, 451–459.
3. Zietkiewicz, E., Rafalski, A. and Labuda, D., Genome fingerprinting by simple sequence repeat (SSR) – anchored polymerase chain reaction amplification. *Genomics*, 1994, **20**, 176–183.
4. Reddy, M. P., Sarla, N. and Siddiq, E. A., Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica*, 2002, **128**, 9–17.
5. Parsons, B. J., Newbury, H. J., Jackson, M. T. and Ford-Lloyd, B. V., Contrasting genetic diversity relationships are revealed in rice (*Oryza sativa* L.) using different marker types. *Mol. Breed.*, 1997, **3**, 115–125.
6. Blair, M. W., Panaud, O. and McCouch, S. R., Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). *Theor. Appl. Genet.*, 1999, **98**, 780–792.
7. Joshi, S. P., Gupta, V. S., Aggarwal, R. K., Ranjekar, P. K. and Brar, D. S., Genetic diversity and phylogenetic relationship as revealed by inter-simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theor. Appl. Genet.*, 2000, **100**, 1311–1320.
8. Sarla, N., Sunil Babu and Siddiq, E. A., ISSR and SSR markers based on AG and GA repeats delineate geographically diverse *Oryza nivara* accessions and reveal rare alleles. *Curr. Sci.*, 2003, **84**, 683–690.
9. Nagaraju, J., Kathirvel, M., Ramesh Kumar, R., Siddiq, E. A. and Hasnain, S. E., Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR–PCR and SSR markers. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 5836–5841.
10. Akagi, H., Yokozeki, Y., Inagaki, A., Nakamura, A. and Fujimura, T., A co-dominant DNA marker closely linked to the rice nuclear restorer gene *Rf-1*, identified with inter-SSR fingerprinting. *Genome*, 1996, **39**, 1205–1209.
11. Fuentes, J. L., Escobar, F., Alvarez, A., Gallego, G., Duque, M. C., Dews, J. E. and Tohme, J. M., Analysis of genetic diversity in Cuban rice varieties using isozyme, RAPD and AFLP markers. *Euphytica*, 1999, **109**, 107–115.
12. Virk, P. S., Zhu, J., Newbury, H. J., Bryan, G. J., Jackson, M. T. and Ford-Lloyd, B. V., Effectiveness of different classes of molecular marker for classifying and revealing variation in rice (*Oryza sativa*) germplasm. *Euphytica*, 2000, **112**, 275–284.
13. Davierwala, A. P., Chowdari, K. V., Shiv Kumar, Reddy, A. P. K., Ranjekar, P. K. and Gupta, V. S., Use of three different marker systems to estimate genetic diversity of Indian elite rice varieties. *Genetica*, 2000, **108**, 269–284.
14. Yang, G. P., Saghai Maroof, M. A., Xu, C. G., Zhang, Q. and Biyashev, R. M., Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars. *Mol. Gen. Genet.*, 1994, **245**, 187–194.
15. Olufowote, J. O. *et al.*, Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. *Genome*, 1997, **40**, 370–378.
16. Sebastian, L. S., Hipolito, L. R. and Garcia, J. S., Molecular marker analysis of the contribution made by landraces to modern Philippine rice cultivars. *SABRAO J. Breed. Genet.*, 1998, **30**, 73–81.
17. Powell, W., Morgante, M., Andre, C., Managey, M., Vogel, J., Tingey, S. and Rafalski, A., The utility of RFLP, RAPD, AFLP and SSLP (microsatellite) markers for germplasm analysis. *Mol. Breed.*, 1996, **2**, 225–238.
18. Prevost, A. and Wilkinson, M. J., A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor. Appl. Genet.*, 1999, **98**, 107–112.
19. Gilbert, J. E., Lewis, R. V., Wilkinson, M. J. and Caligari, P. D. S., Developing an appropriate strategy to assess genetic variability in plant germplasm collections. *Theor. Appl. Genet.*, 1999, **98**, 1125–1131.
20. Fernandez, M. E., Figueiras, A. M. and Benito, C., The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. *Theor. Appl. Genet.*, 2002, **104**, 845–851.
21. Zeven, A. C., Landraces: A review of definitions and classifications. *Euphytica*, 1998, **104**, 127–139.
22. Dellaporta, S. L., Wood, J. and Hick, B., A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.*, 1983, **1**, 19–21.
23. Botstein, D., White, R. L., Skolnick, M. H. and Davies, R. W., Construction of a genetic map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.*, 1980, **32**, 314–331.
24. Milbourne, D. *et al.*, Comparison of PCR based marker systems for the analysis of genetic relationship in cultivated potato. *Mol. Breed.*, 1997, **3**, 127–136.
25. Gregorius, H. R., The relationship between the concepts of genetic diversity and differentiation. *Theor. Appl. Genet.*, 1987, **74**, 397–401.
26. Rohlf, F. J., NTSYS-pc version 2.0, State University of New York, Exeter Software, Setauket, New York, 1993.
27. Neeraja, C. N., Sarla, N. and Siddiq, E. A., RAPD analysis of genetic diversity in Indian landraces of rice (*Oryza sativa* L.). *J. Plant Biochem. Biotechnol.*, 2002, **11**, 93–97.
28. Singh, K. N., Nandi, R., Shanmugasundram, P., Sadasivam, S., Huang, N., Brar, D. S. and Khush, G. S., High resolution DNA fingerprinting of Indian rice (*Oryza sativa* L.) varieties by amplified fragment length polymorphism. *Genet. Res. Crop Evol.*, 1999, **46**, 427–433.
29. Tsumura, Y., Ohba, K. and Strauss, S. H., Diversity and inheritance of inter-simple sequence repeat polymorphisms in Douglas fir (*Pseudotsuga menziesii*) and sugi (*Cryptomeria japonica*). *Theor. Appl. Genet.*, 1996, **92**, 40–45.
30. Morgante, M. and Olivieri, A. M., PCR amplifed microsatellites as markers in plant genetics. *Plant J.*, 1993, **3**, 175–182.
31. Lagercrantz, U., Ellegren, H. and Andersson, L., The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.*, 1993, **21**, 1111–1115.
32. Sarla, N., Neeraja, C. N. and Siddiq, E. A., Determining genetic diversity in Indian landraces of rice using inter-simple sequence repeat (ISSR) polymorphism. In Proceedings of the Fourth International Rice Genetics Symposium, IRRI, Philippines, 22–27 October 2000, p. 217.
33. Ajibade, S. R., Weeden, N. F. and Chite, S. M., Inter simple sequence repeat analysis of genetic relationships in the genus *Vigna*. *Euphytica*, 2000, **111**, 47–55.
34. Charters, Y. M., Robertson, A., Wilkinson, M. J. and Ramsay, G., PCR analysis of oilseed rape cultivars (*Brassica napus* L. spp.

oleifera) using 5′-anchored simple sequence repeat (SSR) primers. *Theor. Appl. Genet.*, 2000, **92**, 780–792.
35. Charters, Y. M. and Wilkinson, M. J., The use of self-pollinated progenies as 'in-groups' for the genetic characterization of cocoa germplasm. *Theor. Appl. Genet.*, 2000, **100**, 160–166.
36. Garland, S. H., Lewin, L., Abedinia, M., Henry, R. and Blakeney, A., The use of microsatellite polymorphisms for the identification of Australian breeding lines of rice (*Oryza sativa* L.). *Euphytica*, 1999, **108**, 53–63.
37. Ramakrishna, W., Davierwala, A. P., Gupta, V. S. and Ranjekar, P. K., Expansion of a (GA) dinucleotide at a microsatellite locus associated with domestication in rice. *Biochem. Genet.*, 1998, **36**, 323–327.
38. Wu, K. S. and Tanksley, S. D., Abundance, polymorphism and genetic mapping of microsatellites in rice. *Mol. Gen. Genet.*, 1993, **241**, 225–235.
39. McCouch, S. R. *et al.*, Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.*, 2002, **9**, 199–207.
40. Cho, Y. G. *et al.*, Diversity of microsatellites derived from genetic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet.*, 2000, **100**, 713–722.
41. Sangwan, I. and Brian, M. R. O., Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA. *Plant Physiol.*, 2002, **129**, 1788–1794.
42. van Steensel, B., Delrow, J. and Bussemaker, H. J., Genomewide analysis of *Drosophila* GAGA factor target genes reveals context-dependent DNA binding. *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 2580–2585.
43. Santi, L. *et al.*, The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*. *Plant J.*, 2003, **34**, 813–826.
44. Vorlickova, M., Kejnovska, I., Kovanda, J. and Kypr, J., Conformational properties of DNA strands containing guanine–adenine and thymine–adenine repeats. *Nucleic Acids Res.*, 1998, **26**, 1509–1514.
45. Benet, A., Molla, G. and Azorin, F., d(GA × TC)(*n*) microsatellite DNA sequences enhance homologous DNA recombination in SV 40 minichromosomes. *Nucleic Acids Res.*, 2000, **28**, 4617–4622.
46. Fang, D. Q. and Roose, M. L., Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theor. Appl. Genet.*, 1997, **95**, 211–219.

ACKNOWLEDGEMENTS. We are grateful to the Indian Council of Agricultural Research, New Delhi for financial support (National Professor Project to E.A.S.), Directorate of Rice Research for facilities, and Lakshmi Raman and A. V. V. Lakshmi for technical assistance.

Received 27 January 2005; revised accepted 11 August 2005