

Research article

Potential of (GATA)_n microsatellites from rice for inter- and intra-specific variability studies

Armaity P Davierwala¹, Wusirika Ramakrishna², Venkat Chowdari³, Prabhakar K Ranjekar¹ and Vidya S Gupta*¹

Address: ¹Plant Molecular Biology Unit, Division of Biochemical Sciences, National Chemical Laboratory, Pune 411008, India, ²Department of Biology, H339 Hansen Building, Purdue University, West Lafayette, Indiana 47907, USA and ³Department of Psychiatry, WPIC 1451, University of Pittsburgh, Pittsburgh, PA 15213, USA

E-mail: Armaity P Davierwala - adavierwata@yahoo.com; Wusirika Ramakrishna - wusirika@hotmail.com; Venkat Chowdari - kchowdari@hotmail.com; Prabhakar K Ranjekar - prkr@ems.ncl.res.in; Vidya S Gupta* - vidya@ems.ncl.res.in

*Corresponding author

Published: 16 October 2001

Received: 22 June 2001

BMC Evolutionary Biology 2001, 1:7

Accepted: 16 October 2001

This article is available from: <http://www.biomedcentral.com/1471-2148/1/7>

© 2001 Davierwala et al; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any non-commercial purpose, provided this notice is preserved along with the article's original URL. For commercial use, contact info@biomedcentral.com

Abstract

Background: The microsatellite, (GATA)_n has been frequently used for DNA fingerprinting. However, very few attempts have been made to analyze (GATA)_n-containing loci in rice.

Results: Three polymorphic (GATA)_n-harboring loci viz. OSIA6, OSIH10 and OS2E7, containing 7–13 repeat motifs were identified from a genomic library of a cultivated rice, *Oryza sativa* var. Basmati-370 using oligonucleotide probe (GATA)₄. When (GATA)_n flanking primers were used to screen 26 wilds (representing different genomes of rice), 16 cultivars, 47 Indian elite rice varieties and 37 lines resistant/susceptible to bacterial blight, up to 22 alleles were obtained at an individual locus. Also, interestingly the bacterial blight resistant lines clustered into a separate group from the remaining rice genotypes, when a dendrogram was constructed based on the polymorphism obtained at the three loci. This may be due to the partial homology of the clones OSIH10 and OS2E7 to regions encoding *O. longistaminata* receptor kinase-like protein and pathogenesis-related protein. The ability of these *O. sativa* flanking primers to amplify DNA of maize, wheat, barley and oat indicates that these (GATA)_n-containing loci are conserved across different cereal genera.

Conclusions: The large allele number obtained reveals the potential of (GATA)_n-containing loci as powerful tools to detect simple sequence length polymorphism (SSLP). The (GATA)_n-flanking primers were not only useful in distinguishing between closely related genotypes, but could also be used for cross-species amplification and are also conserved across different cereal genera. These loci could also cluster the bacterial blight resistant/susceptible lines into different groups based on the resistance genes present in them.

Background

Microsatellites [1] or simple sequence repeats (SSRs) [2,3], consist of simple tandemly repeated di- to penta-nucleotide sequence motifs. Since they are abundant, hy-

pervariable, multi-allelic and evenly distributed throughout nuclear genomes of eukaryotes [4–8] as well as some prokaryotic genomes [9], they provide a valuable source of polymorphism making them an important

class of genetic markers. The exceptionally high levels of polymorphism detected by SSRs, referred to as SSLP (simple sequence length polymorphism) is due to the variability in the number of tandem repeats at a particular locus [10].

SSLP can be easily, accurately, rapidly and economically assayed by PCR with primers specific to the unique flanking sequences of the SSR [6,7,11,12]. These second generation markers are somatically stable and inherited in a co-dominant Mendelian manner and can, therefore, distinguish between heterozygotes and homozygotes [13–15]. All these factors together have made SSRs ideal markers for plant breeding, genetic linkage analysis, gene mapping, paternity analysis, forensic identification, conservation biology and population genetics [16–19].

SSLPs have been demonstrated to be extremely useful in rice, for DNA fingerprinting and variety protection [20], studies relating to germplasm conservation and utilization [20–23], gene and QTL analysis [22,24,25]; pedigree analysis and marker assisted breeding [22,26] and screening of large insert libraries [27].

In rice, about 50% of the genome consists of repetitive DNA sequences [28], with approximately 5,700 to 10,000 SSRs [27]. Totally only 323 SSRs have been mapped on the 12 chromosomes of rice, with an average distance of one SSLP per 6 cM [29,30]. Most of the SSR loci studied so far mainly contain dinucleotide motifs and few contain trinucleotide motifs while very few tetra-nucleotide-containing microsatellite loci have been identified.

Our laboratory has earlier reported the usefulness of microsatellites and inter-simple sequence repeats in DNA fingerprinting of different rice genotypes [31–34]. In addition, repeat expansion in cultivated rice compared to wild rices has been shown at a $(GA)_n$ -containing microsatellite locus, RM122 [35] and organization of a complex microsatellite-containing locus, OS1E6, comprising $(GATA)_n$ and $(AC)_n$ repeats has been studied [36]. Our laboratory has not only reported the usefulness of $(GATA)_4$ in DNA fingerprinting of different rice genotypes but has also revealed that this oligonucleotide is important in fingerprinting cultivars of pearl millet [37,38] and isolates of its pathogen, *Sclerospora graminicola* [39] and in detecting sex specific differences in papaya [40]. In the present work, we have cloned and sequenced three $(GATA)_n$ -containing loci from *Oryza sativa* var. Basmati-370, and used these $(GATA)_n$ locus specific microsatellite markers to study the allelic variation revealing the intra- and inter-specific relationships in rice.

Results

Isolation and cloning of loci containing $(GATA)_n$ repeats

Screening of a genomic library of *Oryza sativa* var. Basmati-370 in Charon-40 vector with the probe $(GATA)_4$ resulted in 150 positive signals in the first round of screening. Of these approximately 40 strongly hybridizing plaques were selected for subsequent screening and six clones that gave strong signals on hybridization with $(GATA)_4$ were randomly selected for further analysis. Since the six putative positive phage clones contained inserts greater than 10 kb in size, further subcloning into pUC18 or pBluescript vector was performed, in order to obtain subclones for sequencing.

DNA sequence analysis

Sequence analysis revealed that three subclones viz. OS1A6, OS1H10 and OS2E7 contained $(GATA)_n$ repeats and hence were deposited in GenBank, with accession numbers, AF294726 (BankIt 351397), AF285877 (BankIt 345999) and AF 284424 (BankIt 345056), respectively. A close observation of the sequences of these three subclones revealed that the GATA repeat regions were of varying complexity and length, with 7 to 13 repeats which may be perfect or imperfect, according to the definition of Weber [16]. The subclone OS1A6, of the phage clone GATA 7.1, contained seven perfect GATA repeats, while the subclone OS1H10, of the phage clone GATA 3.1, contained seven imperfect repeats, with interruption of the repeat motif by TA (Table 1). The subclone OS2E7 of the phage clone GATA 8.1 contained 13 imperfect repeats, with interruptions within the repeat motif (Table 1). In addition to GATA repeats, the subclone OS2E7 showed the presence of GATG, TATA and AATA repeat motifs, which may represent degenerated GATA sequences. Most of the degeneration can be explained by mutations of a single nucleotide within the GATA core sequence in the process of evolution. The OS2E7 subclone thus contained seven GATA motifs, four GATA-like repeats, with one mismatch and two AGTA motifs.

The remaining three subclones did not contain GATA or its complementary CTAT repeat. Although dissociation temperature ($T_d = T_m - 7.6^\circ\text{C}$) [41] is useful to predict the appropriate washing temperature, it was found to be least reliable for motifs with a high AT or GC content [42]. In rice, Panaud *et al* [42] observed that 11 out of 20 putative CGG-containing clones and 3 out of 20 TCT-containing clones did not contain the respective repeat motifs, although they did contain CG and AT rich regions, respectively. In our study too the three putative GATA clones contained AT rich regions, which could explain the positive hybridization signals at $(T_m - 5)^\circ\text{C}$ at these loci.

Table 1: Analysis of the three (GATA)_n-containing subclones with respect to repeat number, allele length and number of alleles

No.		OS1A6	OS1H10	OS2E7
1	Original phage clone number from which derived	GATA 7.1	GATA 3.1	GATA 8.1
2	Restriction endonuclease used to digest phage DNA	<i>MspI</i>	<i>TaqI</i>	<i>TaqI</i>
3	Subcloning vector	pUC-18	pUC-18	pUC-18
4	Cloning site	<i>AccI</i>	<i>AccI</i>	<i>AccI</i>
5	GenBank accession number	AF294726	AF285877	AF284424
6	Number of GATA motifs	7 imperfect GATATA (GATA) ₆	7 perfect (GATA) ₇	13 imperfect AATA (GATA) ₃ GAA (GATA) ₄ AGTA GAAA GATG AGTA
7	Primer sequences used for PCR amplification	F 5'GCTGCGCAGTGTATGCGTCT3' R 5'ACCATTTAAACAGCTAATCGA3'	F 5'GAGTGGTCTCAGTTTGGTGC3' R 5'TGGCGGCTGGAGAGCGTTTG3'	F 5'AAGTGGCGACGTGAGAGCG3' R 5'GTTAAGACAGTCCCCAGTG3'
8	Annealing temperature	55°C	67°C with 2% formamide	57°C with 2% formamide
9	Size of expected PCR product (from genomic subclone sequence)	184 bp	263 bp	141 bp
10	Number and size range of alleles detected in			
	(A) Wild and cultivated rice genotypes	7 (184 bp–208 bp)	22 (116–521 bp)	Unscorable ladder-like pattern obtained
	(B) Indian elite rice varieties	3 (174–197 bp)	10 (218–292 bp)	4 (157–179 bp)
	(C) Bacterial blight resistant and	5 (171–198 bp)	19 (140–535 bp)	6 (117–143 bp)
	(D) Cereal genotypes	1 (184 bp)	6 (168–270 bp)	-
11	Maximum number of bands per lane in			
	(A) Wild and cultivated rice genotypes	2	7	-
	(B) Indian elite rice varieties	3	3	3
	(C) Bacterial blight resistant and susceptible lines	2	9	4
	(D) Cereal genotypes	1	3	-

Sequence similarity search using BLAST [43] through the NCBI, BLAST email server revealed that two of the subclones, OS1H10 and OS2E7, showed significant but quite short homology to nucleotide sequences in the GenBank database. The subclones, OS1H10 and OS2E7 had 94 to 95% homology to the *Oryza longistaminata* receptor kinase-like protein, family member D (Acc #U72726) and 84 to 90% homology to the *Oryza sativa* pathogenesis-related protein class 1 (PR-1) gene (Acc #U89895). In both the subclones, the segments showing homology to genes for receptor kinase-like protein and pathogenesis-related protein were found to be overlapping each other, indicating similar regions within these two genes. The BLAST searches were based on full-length sequences from which the primer pairs were designed and were not just from the amplified sequences. When the sequences of these two subclones were compared, only a 66 bp region towards the end of subclone

OS1H10 was similar to a 66 bp region in the middle of subclone OS2E7. Since the remaining sequence of both these subclones was different and the region showing similarity was in the middle of OS2E7 subclone, these two subclones might be entirely different. Primers flanking the (GATA)_n repeat were synthesized (Table 1) and used to amplify various wild and cultivated genotypes of rice, Indian elite rice varieties and bacterial blight resistant and susceptible lines of rice.

Germplasm analysis using primers flanking the (GATA)_n repeat region

When different rice genotypes were amplified using the primers flanking the three different (GATA)_n containing subclones, it was observed that these primer pairs generated large amount of polymorphism even in closely related elite rice germplasm of India. Amplification was performed at 55°C, 67°C with 2% formamide and 57°C

Table 2: List of wild and cultivated rice genotypes used in our study

No.	List of samples used	Accession	Genome	Origin	Number of alleles obtained with			Total number of alleles
					OS1 A6	OS1 H10	OS2 E7	
1	<i>O. glaberrima</i>	TOG 6216	AA	Africa Via Warda	1	2	-	3
2	<i>O. glaberrima</i>	TOG-6229	AA	Africa Via Warda	1	3	-	4
3	<i>O. longistaminata</i>	-	AA	Africa Via CRR1	2	2	-	4
4	<i>O. perennis</i>	104823	AA	Thailand	1	2	-	3
5	<i>O. rufipogon</i>	106424	AA	Vietnam	2	1	-	3
6	<i>O. rufipogon</i>	105908	AA	Thailand	1	1	-	2
7	<i>O. punctata</i> (2n)	105980	BB	Cameroon	1	2	-	3
8	<i>O. punctata</i> (4n)	100884	BBCC	India	1	2	-	3
9	<i>O. minuta</i>	101141	BBCC	Philippines	1	3	-	4
10	<i>O. minuta</i>	101125	BBCC	Philippines	1	1	-	2
11	<i>O. rhizomatis</i>	105432	CC	Sri-Lanka	1	2	-	3
12	<i>O. officinalis</i>	100896	CC	Thailand	1	1	-	2
13	<i>O. officinalis</i>	101399	CC	Vietnam	1	0	-	1
14	<i>O. alta</i>	105143	CCDD	Guyana	1	1	-	2
15	<i>O. grandiglumis</i>	105669	CCDD	Brazil	1	1	-	2
16	<i>O. latifolia</i>	100167	CCDD	Costa Rica	1	0	-	1
17	<i>O. latifolia</i>	100965	CCDD	Costa Rica	1	2	-	3
18	<i>O. australiensis</i>	100882	EE	Australia Via CRR1	1	0	-	1
19	<i>O. australiensis</i>	T-1434	EE	Australia Via CRR1	1	2	-	3
20	<i>O. brachyantha</i>	101232	FF	Sierra Leone	1	0	-	1
21	<i>O. brachyantha</i>	B98-8025	FF	Africa	1	0	-	1
22	<i>O. granulata</i>	104986	GG	Via CRR1 India	1	1	-	2
23	<i>O. granulata</i>	106448	GG	Nepal	1	7	-	8
24	<i>O. longiglumis</i>	105148	HHJJ	Indonesia	0	0	-	0
25	<i>O. ridleyi</i>	100821	HHJJ	Thailand	1	1	-	2
26	<i>O. ridleyi</i>	101453	HHJJ	Malaysia	1	2	-	3
27	<i>O. nivara</i>	-	AA	India	2	2	-	4
28	<i>O. malampuzhaensis</i>	-	BBCC	India	1	0	-	1
29	<i>O. sativa</i> cultivars	-	-	-	-	-	-	-
30	IR 36	Indica	AA	Cultivated variety	1	2	1	3
31	IR 64	Indica	AA	Cultivated variety	1	1	-	2
32	BG-90-2	Indica	AA	Cultivated variety	1	2	1	3
33	Basmati 370	Aromatic rice	AA	Cultivated variety	1	1	1	2
34	Azucena	Japonica	AA	Cultivated variety	1	0	-	1
35	Indrayani	Indica	AA	Cultivated variety	1	2	-	3
36	Ambemohr	Indica	AA	Cultivated variety	1	1	3	2
37	Intan	Indonesian	AA	Cultivated variety	1	1	2	2
38	Khadkya	Indica	AA	Maharashtra India	2	1	4	3
39	Waseas	Japonica	AA	Cultivated variety	1	0	-	1
40	Hakkoda	Japonica	AA	Cultivated variety	1	1	-	2
41	<i>Porteresia coarctata</i>	Related	Unknown	Bangladesh	1	3	4	4
42	<i>Leersia</i>	Related	Unknown	Madagascar	1	0	-	1
43	<i>Rhynchoriza subulata</i>	Related	-	Argentina	1	1	-	2
44	<i>Zea diploperensis</i>	-	-	-	1	2	-	3
45	<i>Triticum durum</i> landrace Narsingarh III	1633	-	-	-	2	-	2
46	<i>Hordeum marinum</i>	PI 200341	-	-	-	1	-	1
47	<i>Avena vaviloviana</i>	PI 412733	-	-	-	3	-	3

with 2% formamide using primers flanking the sub-clones, OS1A6, OS1H10 and OS2E7, respectively.

Analysis of distantly related germplasm

Using OS1A6 flanking primers, seven alleles ranging in size from 184 bp to 208 bp were observed in the wild and cultivated rice germplasm set (Table 2), suggesting a difference of six tetranucleotide repeats between the smallest and the largest alleles (Table 1). 22 alleles (116 bp – 521 bp) were detected with OS1H10 flanking primers on the above germplasm (Table 1). On the other hand, OS2E7 flanking primers detected a ladder-like pattern with large number of bands, suggesting the presence of many alleles, which differ in the number of GATA repeats. When representative cereal genotypes of oat, barley, maize and wheat were studied, amplification was obtained in all the cereals using OS1H10 flanking primers while only the maize genotype, *Zea diploperensis*, gave amplification with OS1A6 flanking primers. When the wild and cultivated germplasm of rice and other cereals were considered together (Table 2), an average of 15.5 alleles were detected over the two scorable loci, OS1E6 and OS1H10. However, when the rice genotypes were considered alone, an average of 14.5 alleles were detected, while an average of 3.5 alleles was observed in the cereal genotypes studied, considering these two loci.

Analysis of closely related Indian elite rice varieties

When the (GATA)_n flanking primers were used to amplify 47 Indian elite rice varieties (Table 3), 3 (174 bp – 197 bp), 10 (218 bp – 292 bp) and 4 (157 bp – 179 bp) alleles were detected at the OS1A6, OS1H10 and OS2E7 loci, respectively (Table 1). Thus, even in the case of closely related germplasm, the three microsatellite loci were found to be quite polymorphic with an average of 5.67 alleles per locus. Although the 47 Indian elite rice varieties have a narrow genetic base as evaluated from molecular marker and parentage data [44], the three microsatellite loci, when used together, could distinguish about three-fourths of the elite rice cultivars unambiguously.

Analysis of bacterial blight resistant / susceptible germplasm

Thirty-seven bacterial blight resistant and susceptible lines (Table 4) resulted in 5, 19 and 6 alleles ranging in size from 171–198 bp, 140–535 bp and 117–143 bp, respectively, using primers flanking the OS1A6, OS1H10 and OS2E7 loci, respectively (Table 1), with an average of 10 alleles per locus. In all the cases, the number of alleles detected was excluding null alleles.

Cluster analysis based on the (GATA)_n-containing loci

A dendrogram was constructed based on the polymorphism data obtained at the three (GATA)_n microsatellite loci in few representative genotypes of the three sets of

Table 3: Indian elite rice cultivars used in the present study

No.	Name	Number of alleles obtained with			Total number of alleles obtained
		OS1A6	OS1H10	OS2E7	
1	Hema	1	1	2	4
2	Tellahamsa	1	1	1	3
3	Rambha	1	1	1	3
4	Jal Lahari	3	1	2	6
5	Phalguna	2	3	3	8
6	Intan	1	-	-	1
7	Rajeshwari	1	1	1	3
8	Jaya	1	-	1	2
9	Indrayani	1	1	1	3
10	Bhawani	1	-	-	1
11	Rasi	1	1	3	5
12	Lalat	2	1	1	4
13	Prakash	1	-	2	3
14	Gauri	1	1	1	3
15	Sankar	1	1	1	3
16	Ananga	1	-	2	3
17	Akashi	1	-	3	4
18	Sugandha	1	1	1	3
19	Madhukar	1	1	1	3
20	Subhadra	1	1	1	3
21	Vibhava	1	1	1	3
22	Rudra	1	1	2	4
23	Mangala	1	1	3	5
24	Varsha	1	1	2	4
25	Jalprika	1	1	1	3
26	Pothana	1	1	3	5
27	Ratna	1	-	0	1
28	Jalnidhi	1	1	1	3
29	Salivahana	1	1	1	3
30	Surekha	1	1	1	3
31	Daya	1	1	1	3
32	Narmada	1	1	1	3
33	Parijat	1	1	2	4
34	Jalmagna	1	1	2	4
35	Vani	1	2	1	4
36	Vikas	1	1	1	3
37	Mandya-vani	1	1	1	3
38	Chandana	1	2	1	4
39	Prasanna	1	1	3	5
40	Karjat-1	1	1	3	5
41	Karjat-2	1	1	1	3
42	Karjat-18	1	1	1	3
43	Karjat-23	1	2	2	5
44	Karjat-35-3	1	1	1	3
45	Karjat-184	1	-	1	2
46	Ratnagiri-1	1	1	1	3
47	Ratnagiri-2	1	1	2	4

germplasm analyzed. The dendrogram revealed that all the bacterial blight resistant lines clustered together, while the remaining genotypes clustered separately (data not shown). Since the bacterial blight resistant lines were found clustered in a separate group, a dendrogram was constructed using the patterns obtained at the three microsatellite loci for all the 37 lines either resistant or susceptible to bacterial blight. The dendrogram thus obtained (Figure 1) revealed the presence of six clusters, wherein the lines grouped together based on the resistance genes present in them. In cluster 1, with the exception of IR22, the remaining three lines were near isogenic to IR-24. The first subgroup of cluster 2 also contained all near-isogenic lines with the exception of Rasi, which joined the cluster separately. The second and third subgroups of the second cluster were heterogeneous, containing lines with *Xa4* and *xa5* resistance genes and one line BPT-5204, which has no resistance genes. Except Swarna and Nigeria-5, in clusters 3 and 6, respectively, all the lines in the third and sixth clusters contained the *xa5* gene. All the *Xa4* containing lines with the exception of T(N)1 in cluster 4 and Kuntalan in cluster 5, grouped together in clusters 4 and 5. In cluster 5, in addition to *Xa4* gene, the lines CRRI-1, CRRI-3 and IET-14444 also contained the *xa5* gene. Interestingly, Jaya, which contains no resistance genes totally outgrouped from the remaining lines. The possible reason due to which the bacterial blight resistant lines showed distinct clustering as above could be due the partial homology of two of these $(GATA)_n$ -containing loci, namely, OS1H10 and OS2E7 to regions encoding for receptor kinase-like protein and pathogenesis-related protein.

Discussion

Tetranucleotide $(GATA)_n$: a source of new STMS markers in rice

Rice has one of the smallest genome sizes among plants with a lower frequency of repeated sequences which necessitates the use of a variety of different SSRs to completely saturate the map. Three hundred and twenty three microsatellite markers identified by library screening [6,29,30,45-49] and GenBank searches of rice sequences [6,29,30,47,50] have been mapped on the rice genetic map. Here, most of the markers contain dinucleotides and trinucleotides and only seven loci contain tetranucleotide repeats such as $(AATT)_n$, $(TTCC)_n$, $(GATG)_n$, $(ATGT)_n$, $(GTAT)_n$, $(ATTT)_n$, $(TTTG)_n$ [6,29,30,47-49]. $(GATA)_n$ sequences which have not been mapped so far in rice, are the most frequent among tetranucleotide repeats with 270 poly $(GATA)_n$ motifs in the entire rice genome [42]. The degree of polymorphism shown by microsatellites rather than their abundance in the genome contributes to their usefulness in genome analysis. For fingerprinting plant genomes, Epplen [51] has shown that oligonucleotides based on -GATA- or -

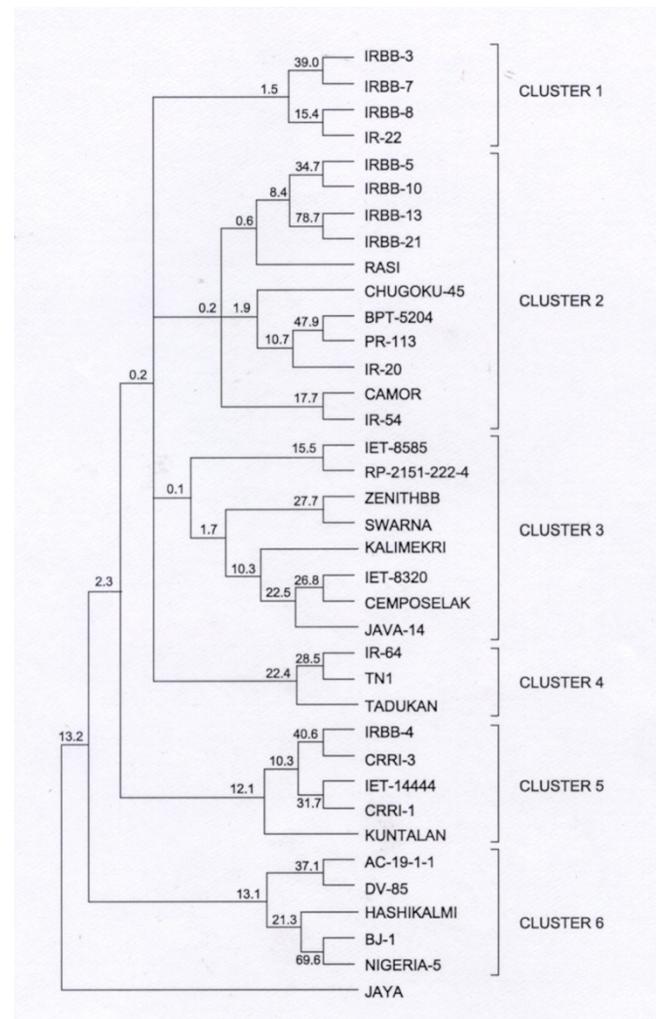


Figure 1

Dendrogram showing the genetic relationships between different near-isogenic lines, bacterial blight resistant and susceptible lines based on the polymorphism data obtained at the three $(GATA)_n$ -containing microsatellite loci. The numbers at the forks indicate the confidence limits (bootstrap values) for the grouping of those species, which are to the right of that fork. The numbers on the branches of the tree indicate the percentage of times the genotypes to the right of the branching point were grouped together during bootstrapping analysis.

GACA-microsatellite motifs are amongst the most frequently used for fingerprinting plant genomes. We, therefore, thought that it would be interesting to isolate these $(GATA)_n$ -containing regions from rice and convert them into PCR amplifiable microsatellites, to determine variations at these highly polymorphic loci.

Table 4: BLB resistant and susceptible rice lines used for analysis

No.	Lines analyzed		Number of alleles obtained with			Total number of alleles
			OSIA6	OSIH10	OS2E7	
1	IRBB-3	Near isogenic line	1	-	2	3
2	IRBB-4	Near isogenic line	1	1	1	3
3	IRBB-5	Near isogenic line	1	1	4	6
4	IRBB-7	Near isogenic line	1	1	1	3
5	IRBB-8	Near isogenic line	1	1	6	8
6	IRBB-10	Near isogenic line	1	1	6	8
7	IRBB-13	Near isogenic line	1	1	6	8
8	IRBB-21	Near isogenic line	1	1	5	7
9	BJ-1	BLB resistance donor	1	2	2	5
10	DV-85	BLB resistance donor	1	2	0	3
11	Camor (AC 17366)	BLB resistance donor	1	2	1	4
12	Java-14	BLB resistance donor	1	1	1	3
13	AC-19-1-1	BLB resistance donor	1	2	1	4
14	Kalimekri	BLB resistance donor	1	2	1	4
15	Cemposelak	BLB differential	1	2	1	4
16	Hashikalmi	BLB resistance donor	1	3	2	6
17	Tadukan	BLB susceptible line	1	2	9	12
18	PR-113	BLB resistant variety, Ajaya derivative	1	2	3	6
19	T(N)-1	BLB susceptible variety	1	2	6	9
20	Jaya	BLB susceptible variety	1	-	2	3
21	Rasi	BLB susceptible variety	1	2	2	5
22	BPT-5204	BLB susceptible variety	1	2	6	9
23	IR-54	BLB resistant variety	1	2	3	6
24	IR-64	BLB resistant variety	1	3	6	10
25	IR-22	BLB resistant variety	1	3	3	7
26	Swarna	-	1	4	1	6
27	IR-20	BLB resistant variety	1	3	3	7
28	Kuntalan	BLB resistance donor	1	2	3	6
29	Nigeria-5	BLB resistance donor	1	2	2	5
30	Zenith-BB	BLB resistant line	2	4	1	7
31	Chugoku-45	BLB resistant line	1	-	6	7
32	IET-8320	BLB resistant pyramid	2	2	1	5
33	IET-8585 (Ajaya)	BLB resistant pyramid	1	4	1	6
34	RP-2151-222-4	BLB resistant pyramid	1	2	4	7
35	IET-14444	BLB resistant pyramid, Ajaya derivative	1	2	4	7
36	CRR1-1	BLB resistant pyramid from CRR1	1	1	4	6
37	CRR1-3	BLB resistant pyramid from CRR1	1	1	2	3

(GATA)_n loci: tools for intraspecific variability in rice

Allelic diversity studies in rice [6,21,29,48,49] have revealed 2–25 alleles per SSLP locus as compared to 2–4 alleles per RFLP locus in cultivated indica and japonica germplasm [27]. In our studies too, up to 19 alleles were detected in *Oryza sativa* cultivars at the three (GATA)_n-containing loci. Even in the closely related Indian elite rice cultivars, 3–10 alleles were obtained, indicating that microsatellite-based PCR markers are very useful in studying variability among genotypes having a narrow genetic base. The main reason for this high degree of polymorphism at the intraspecific level is due to the high

mutation rate at microsatellite loci, which results in additions and deletions of repeat motifs.

In humans, SSRs with fewer than 10 repeat motifs tend to show little, if any, polymorphism [13,52]. However, our studies on rice have revealed that SSRs with less number of repeat motifs also can be extremely informative in plants. In *Beta*, *Arabis*, *Lycopersicon* and soybean also polymorphism is observed, despite the small number of repeats per SSR locus [53–56], suggesting that microsatellite loci with small number of repeats can

be potentially informative in plants and should not be ignored.

Cross species amplification at (GATA)_n loci

Wu and Tanksley [6] and Panaud *et al* [48] used STMS markers to amplify DNA of wild *Oryza* species, indicating their utility in analyzing interspecific crosses. We have observed that STMS markers resulted in a higher level of variation as revealed by a greater number of alleles, when used to amplify genomic DNA of wild rice accessions as compared to *O. sativa* genotypes. When wild germplasm of rice was analyzed, 7–22 alleles were detected at two loci viz. OS1A6 and OS1H10, while the locus OS2E7 resulted in a ladder-like pattern indicating an extremely high level of polymorphism. Provan *et al* [57] have observed that amplification of wild rice accessions with primers flanking mononucleotide repeats in rice chloroplast genome, which are analogous to nuclear SSRs, resulted in higher diversity values as compared to *O. sativa* accessions.

There are several examples of cross species amplification of SSR loci in animal studies, where transferability of these loci in divergent taxa has been reported. In plants also, although reports on cross-species amplification are fewer, similar observation have been made in different species of *Cicer* [58] and *Glycine* [56]. Using soybean SSR flanking primers, Peakall *et al* [56] have amplified DNA from several legume genera such as *Albizia*, *Vigna*, *Trifolium*, etc. Weising and Gardner [59] observed chloroplast DNA variation in members of Solanaceae and most other angiosperm species when they used primers flanking microsatellite motifs in tobacco. The *O. sativa* primers obtained in this study showed amplification not only in different *Oryza* species but also in representative cereal genera viz. maize, wheat, barley and oat, indicating that they may be located in the conserved regions of the genome. In an earlier study, we have successfully amplified DNA from different *Oryza* species as well as wheat, maize, barley, oat, pearl millet and rye at two microsatellite-containing loci, OS1E6 and RM122 [36]. However, with the exception of our studies, very few primer sequences from rice have been shown to be conserved enough to amplify in related genera, although some amplify DNA of wheat, oat, barley, maize, sorghum, millet, sugarcane and bamboo reliably [27,46]. Roder *et al* [60] have used microsatellites from wheat to amplify barley and rye DNA. Recently, Ishii and McCouch [61] observed microsynteny at microsatellite loci in the chloroplast genomes of rice and maize. Using primers flanking these loci they amplified wild and cultivated accessions of rice and eight other members of the family Gramineae viz. maize, teosinte, wheat, oat, barley, pearl millet, sorghum and sugarcane. All these studies

indicate that microsatellite loci could be useful for comparative genome mapping in cereals.

Microsatellite-containing loci with greater number of repeats are more likely to be polymorphic [16] and the high level of polymorphism at such loci can be attributed to two molecular mechanisms, viz. replication slippage or unequal crossing over. Levinson and Gutman [62] have proposed that if replication slippage is an important mechanism, a longer repeat would tend to show more variation, since the chance of replication errors is higher for a longer stretch of repeated sequence. Our studies on wild genotypes revealed a positive correlation between the average number of repeats and the amount of variation. The OS2E7 locus with 13 imperfect repeats resulted in a large number of alleles, which could not be scored. On the other hand, the OS1H10 locus with seven perfect repeats resulted in 22 alleles while seven alleles were obtained at the OS1A6 locus with seven imperfect repeats, indicating that replication slippage must be the molecular mechanism responsible for generating new alleles at these SSR loci [62–64]. Such a phenomenon was also observed in the case of (TAA)_n-containing loci in chickpea [58].

Utility of microsatellites in coding and non-coding regions

Two of the loci reported here have shown significant homology to short stretches of the genes encoding for *Oryza longistaminata* receptor kinase-like protein and *Oryza sativa* pathogenesis related-protein. Further, these homologous regions lie within the OS1H10 and OS2E7 loci that are amplified in PCR, 4bp prior to the (GATA)_n repeat in case of OS2E7 and 5 bp after the (GATA)_n repeat in OS1H10. Also, *Xa21*, which is one of the genes responsible for imparting resistance to bacterial blight in rice, encodes a putative receptor kinase [65,66] and (CGG)_n repeats have been found in the exonic region of the receptor kinase-like protein, *Xa21* [29]. Considering these factors, we conclude that the reason for the bacterial blight resistance lines grouping together in one cluster, distinct from the other clusters containing the remaining genotypes, is the short but significant homology of these (GATA)_n loci to the *Oryza longistaminata* receptor kinase-like protein.

We have also observed that the bacterial blight resistant lines clustered into different groups based on the resistance genes present in them. We had earlier detected a similar phenomenon when we used 11 different STMS markers close to the bacterial blight resistance genes, to analyze the same set of 37 rice genotypes [67]. Using OS2E7 flanking primers all the near-isogenic lines except IRBB-8 had only one band of 143 bp. OS1H10 flanking primers on the other hand resulted in three common bands of 240 bp, 257 bp and 331 bp in five of the near-

isogenic lines, IRBB-5, IRBB-8, IRBB-10, IRBB-13 and IRBB-21, along with several other bands and a common band of 165 bp in IRBB-4 and IRBB-7. The presence of such bands in the near-isogenic lines, having different bacterial blight resistance genes probably suggests that these microsatellite loci may be associated with disease resistance genes.

Tautz *et al* [68] have shown that there is a frequent occurrence of nearly all possible SSR motif arrays in both coding and non-coding regions of eukaryotic genomes. Microsatellite repeats have been found in several cloned and completely sequenced genes of rice and many of these microsatellite loci have been mapped [6,29]. Dinucleotide repeats of RM120 are present in the 3' untranslated region (UTR) of the rice *Adh-2* gene for alcohol dehydrogenase [69], while those of RM203 are found upstream of the *phy-18* gene for phytochrome in rice [70]. Genetic mapping of these two microsatellite loci has resulted in simultaneous mapping of these two genes onto the rice genetic map [6], indicating that flanking microsatellite polymorphisms can be used to map known genes even when there may not be determinable polymorphism using coding sequences as probes. The screening of a cDNA library has revealed the presence of seven out of twelve surveyed SSRs in the transcribed regions in rice [42]. SSRs have also been found linked to several genes of agronomic importance in rice. Expansion of (CT)_n repeats at the microsatellite locus, RM190, located at the 5' UTR of the *waxy* gene [71] was implicated in phenotypic variation for starch quality in commercial US rice cultivars [26]. Two microsatellite markers were found to be closely linked to the giant embryo gene *ge^s* [24], responsible for controlling enlarged embryo size and high protein, oil and vitamin content in rice, while three markers have shown linkage to the *xa-5* gene, responsible for imparting resistance to bacterial blight in rice [25]. On chromosome 1, two STMS markers were found to be located in the vicinity of a QTL for rice yield [22].

The availability of greater number of SSCP markers will help to saturate the existing RFLP, SSCP and AFLP maps of rice, thus increasing the power of genome analysis and to fine map different genes and QTLs of agronomic and economic importance in rice. Furthermore, the ability of few characterized microsatellite loci to amplify in different cereal genera could be useful for synteny mapping when they are mapped onto orthologous regions of different cereal genomes. The clustering of bacterial blight resistant genotypes using microsatellite markers in our study postulates diagnostic potential of such markers which can be exploited in marker-assisted breeding programs. However, in order to unequivocally diagnose re-

sistance, a close linkage of these microsatellite loci to R genes will have to be established via mapping.

Conclusions

(GATA)_n-containing loci are powerful tools for detecting simple sequence length polymorphisms in the rice germplasm analyzed. These loci are not only useful in discriminating closely related germplasm with a narrow genetic base but can also be used to in cross species and cross-genus amplification as these loci are found to be conserved in different cereals. In addition the polymorphism obtained at these microsatellite loci could cluster the bacterial blight resistant and susceptible lines into different groups according to the resistance genes present in them.

Materials and methods

Library screening and selection of clones containing (GATA)_n repeats

A genomic library, constructed using *EcoRI* digest of *Oryza sativa* var. Basmati-370 in Charon 40 phage vector (PFU 4.8×10^{-8}), was screened with the oligonucleotide (GATA)₄, by plaque hybridization [72]. The tetranucleotide (GATA)₄ was end-labeled using γ ³²P ATP and T4 polynucleotide kinase (U.S. Biochemical Corp., U.K.), according to Sambrook *et al* [72] and hybridization was carried out at (Tm-5)°C. To reduce the number of false positives, three rounds of screening were performed, wherein the clones that hybridized with (GATA)₄ probe in the first round were transferred onto new filters and re-hybridized.

Subcloning and sequencing of the positive clones

Phage DNAs isolated from the selected putative positive clones were digested with *TaqI* and *MspI* and the DNA fragments obtained were ligated into the *AclI* site of pUC18 [73] or the *ClaI* site of pBluescript cloning vectors. These recombinant plasmids were used to transform *Escherichia coli* host strain XL1-Blue competent cells and white colonies were transferred into microtitre plates containing Luria Bertani medium. Colony hybridization was performed using end-labeled (GATA)₄ probe [72] in order to select the recombinant clones containing (GATA)_n repeat.

Plasmid DNA, prepared using alkaline lysis method [72] and purified by polyethylene glycol method [72], was sequenced by Sanger's dideoxy chain termination method using Sequenase version 2.0 DNA sequencing kit (U.S. Biochemical Corp., U.K.).

Homology search and design of primers

(GATA)₅ was used in BLAST instead of (GATA)₄, used for hybridization, so that longer GATA stretches containing a minimum of 20 nucleotides, could be identified.

PCR primers, flanking the (GATA)_n repeat motif, were designed and were synthesized by Research Genetics Inc (USA) and the primer synthesis facility at NCL (India).

Plant material

The PCR primers synthesized above were used to screen wild and cultivated rice genotypes, along with four representative cereal genotypes. 26 wilds, 6 cultivated rice varieties and related genera (Table 2) were obtained from the International Rice Research Institute, Manila, Philippines, while the remaining two wilds and all the other cultivated rice genotypes (Table 2), along with 47 Indian elite rice varieties (Table 3) and 37 bacterial blight resistant and susceptible lines (Table 4) were obtained from Directorate of Rice Research, Hyderabad, India. Seeds of maize were procured from Indian Agricultural Research Institute, New Delhi, while seeds of wheat were obtained from Agharkar Research Institute, Pune, India and those of barley and oat were made available by U.S.D.A. A.R.S., National Small Grains Collection, Aberdeen, U.S.A.

DNA was extracted from leaf tissue by hexa-decyl-trimethyl-ammonium bromide (CTAB) method [74].

Polymerase chain reaction using primers flanking the microsatellite motif

Each reaction was performed in 10 µl volume [6] containing 200 µM each of dCTP, dGTP and dTTP, 25 µM of dATP, 1 µCi of ³²P dATP, 20 ng of each primer, 50 ng template DNA, 0.24 U of *Taq* DNA polymerase (Boehringer Mannheim, Germany), in 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl and 0.5 mM spermidine.

DNA amplifications were carried out in a MJ Research thermal cycler, wherein an initial denaturation of 5 min was performed at 94°C followed by 40 thermal cycles of denaturation at 94°C for 1 min, primer annealing at specific annealing temperature (depending on the T_m of the primer pair) for 2 min and extension at 72°C for 1.5 min with a final extension at 72°C for 5 min.

After amplification, half volume of stop solution (98% deionized formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol) was added to individual reactions. Samples were denatured at 70–80°C for 2–3 min and then electrophoresed on 6% denaturing polyacrylamide gels containing 7 M urea in 0.5 X TBE (Tris-borate-EDTA) buffer (pH 8.0), at 1500 V and visualized by autoradiography.

ΦX174 DNA, digested with *HaeIII* (Bangalore Genei, India) was dephosphorylated using calf intestinal phosphatase (CIP) (Boehringer Mannheim, Germany) and end-labeled using ³²P ATP according to Sambrook *et al* [72] and loaded as molecular weight standard.

Data analysis

Band sizes of each of the alleles were calculated in base pairs, using the program, SeqAid II (tm) version 3.5. [75], wherein the distance of migration of the bands in the marker lane was considered as a standard, against which the sizes of the fragments of interest in the remaining lanes were calculated.

The fragments amplified using these primer pairs were scored as present (1) or absent (0) and the data entered into a binary matrix was analyzed using Winboot program [76]. Coefficients of similarity were calculated using Dice coefficient and cluster analysis was performed by agglomerative technique using the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method. The relationships between the different rice lines studied were portrayed graphically in the form of a dendrogram.

Acknowledgements

APD thanks the Council of Scientific and Industrial Research (CSIR), New Delhi for the award of Senior Research Fellowship. International Rice Research Institute, Manila, Philippines is thanked for providing DNA of 26 wilds, six cultivated rice varieties and three related genera. Directorate of Rice Research (DRR), Hyderabad is acknowledged for providing seed material of all the remaining rice genotypes used in our study. Indian Agricultural Research Institute (IARI), New Delhi and Agharkar Research Institute are acknowledged for providing seed material of maize and wheat, respectively. The authors thank USDA ARS, National Small Grains Collection, Aberdeen, U.S.A. for providing seeds of barley and oat. This work was supported by grant from the Rockefeller Foundation, U.S.A., #Grant No RF 95001, #327. The experiments described in this manuscript comply with the current laws of the country in which they were performed.

References

- Litt M, Luty JA: **A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle action gene.** *Am J Hum Genet* 1989, **44**:397
- Weber JL, May PE: **Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction.** *Am J Hum Genet* 1989, **44**:388-396
- Tautz D: **Hypervariability of simple sequence as a general source for polymorphic DNA markers.** *Nucleic Acids Res* 1989, **17**:6463-6471
- Epplen JT: **On simple repeat GAT/CA sequences in animal genomes: a critical reappraisal.** *J Hered* 1988, **79**:409-417
- Goodfellow PN: **Variation is now the theme.** *Nature* 1992, **359**:777-778
- Wu KS, Tanksley SD: **Abundance, polymorphism and genetic mapping of microsatellites in rice.** *Mol Gen Genet* 1993, **24**:225-235
- Bell CJ, Ecker JR: **Assignment of 30 microsatellite loci to the linkage map of Arabidopsis.** *Genomics* 1994, **19**:137-144
- Tautz D, Renz M: **Simple sequences are ubiquitous repetitive components of eukaryotic genomes.** *Nucleic Acids Res* 1984, **12**:4127-4138
- Field D, Willis C: **Long, polymorphic microsatellites in simple organisms.** *Proc R Soc Lond* 1996, **263**:209-215
- Schlotterer C, Vogel C, Tautz D: **Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations.** *Genetics* 1997, **146**:309-320
- Beckmann JS, Soller M: **Towards a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellites.** *Bio / technology* 1990, **8**:930-932
- Love JM, Knight AM, McAleer MA, Todd JA: **Towards construction of a high resolution map of the mouse genome using PCR-analyzed microsatellites.** *Nucleic Acids Res* 1990, **18**:4123-4130

13. Weber JL: **Human DNA polymorphisms and methods of analysis.** *Curr Opin Biotechnol* 1990, **1**:166-171
14. Morgante M, Olivieri AM: **PCR-amplified microsatellite markers in plant genetics.** *Plant J* 1993, **3**:175-182
15. Thomas MR, Scott NS: **Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs).** *Theor Appl Genet* 1993, **86**:985-990
16. Weber JL: **Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms.** *Genomics* 1990, **7**:524-530
17. Rafalski JA, Tingey SV: **Genetic diagnostics in plant breeding: RAPDs, microsatellites, and machines.** *Trends Genet* 1993, **9**:275-280
18. Gupta PK, Balyan IS, Sharma PC, Ramesh B: **Microsatellite in plants-a new class of molecular markers.** *Curr Sci* 1996, **70**:45-54
19. Jarne P, Lagoda PJJ: **Microsatellites, from molecules to populations and back.** *Trends Ecol Evol* 1996, **11**:424-429
20. Olufowote JO, Xu Y, Chen X, Park WD, Beachell HM, Dilday RH, Goto M, McCouch SR: **Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers.** *Genome* 1997, **40**:370-378
21. Yang GP, Saghai Maroof MA, Xu CG, Zhang Q, Biyashev RM: **Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice.** *Mol Gen Genet* 1994, **245**:187-194
22. Xiao J, Li J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Yuan L: **A wild species contains genes that may significantly increase the yield of rice.** *Nature* 1996, **384**:223-224
23. Xiao J, Li J, Yuan L, McCouch SR: **Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers.** *Theor Appl Genet* 1996, **92**:637-643
24. Koh HJ, Heu MH, McCouch SR: **Molecular mapping of the gene controlling the super-giant embryo character in rice (*Oryza sativa* L.).** *Theor Appl Genet* 1993, **93**:257-261
25. Blair MW, McCouch SR: **Microsatellite sequence tagged site markers diagnostic for the rice bacterial leaf blight resistance gene *xa5*.** *Theor Appl Genet* 1997, **95**:174-184
26. Ayres NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA, Park WD: **Microsatellites and a single nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm.** *Theor Appl Genet* 1997, **94**:773-781
27. McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M: **Microsatellite marker development, mapping and applications in rice genetics and breeding.** *Plant Mol Biol* 1997, **35**:89-99
28. Deshpande VG, Ranjekar PK: **Repetitive DNA in three Gramineae species with low DNA content.** *Hoppe-Seyler Z Physiol Chem* 1980, **361**:1223-1233
29. Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S: **Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.).** *Theor Appl Genet* 2000, **100**:713-722
30. Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR: **Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.).** *Theor Appl Genet* 2000, **100**:697-712
31. Ramakrishna W, Lagu MD, Gupta VS, Ranjekar PK: **DNA fingerprinting in rice using oligonucleotide probes specific for simple repetitive DNA sequences.** *Theor Appl Genet* 1994, **88**:402-406
32. Ramakrishna W, Chowdari KV, Lagu MD, Gupta VS, Ranjekar PK: **DNA fingerprinting to detect genetic variation in rice using hypervariable DNA sequences.** *Theor Appl Genet* 1995, **90**:1000-1006
33. Gupta VS, Ramakrishna W, Rawat SR, Ranjekar PK: **(CAC)₅ detects DNA fingerprints and sequences homologous to gene transcripts in rice.** *Biochem Genet* 1994, **32**:1-8
34. Joshi SP, Gupta VS, Aggarwal RK, Ranjekar PK, Brar DS: **Genetic diversity and phylogenetic relationship as revealed by inter-simple sequence repeat (ISSR) polymorphisms in the genus *Oryza*.** *Theor Appl Genet* 2000, **100**:1311-1320
35. Ramakrishna W, Davierwala AP, Gupta VS, Ranjekar PK: **Expansion of (GA)_n dinucleotide at a microsatellite locus associated with domestication in rice.** *Biochem Genet* 1998, **36**:323-327
36. Davierwala AP, Ramakrishna W, Ranjekar PK, Gupta VS: **Sequence variations at a complex microsatellite locus in rice and its conservation in cereals.** *Theor Appl Genet* 2000, **101**:1291-1298
37. Chowdari KV, Davierwala AP, Gupta VS, Ranjekar PK, Govila OP: **Genotype identification and assessment of genetic relationships in pearl millet [*Pennisetum glaucum* (L.) R. Br.] using microsatellites and RAPDs.** *Theor Appl Genet* 1998, **97**:154-162
38. Chowdari KV, Venkatachalam SR, Davierwala AP, Gupta VS, Ranjekar PK, OP Govila: **Hybrid performance and genetic distance as revealed by the (GATA)₄ microsatellite and RAPD markers in pearl millet.** *Theor Appl Genet* 1998, **97**:163-169
39. Sastry JG, Ramakrishna W, Sivaramakrishnan S, Thakur RP, Gupta VS, PK Ranjekar: **DNA fingerprinting detects genetic variability in the pearl millet downy mildew pathogen (*Sclerospora graminicola*).** *Theor Appl Genet* 1995, **91**:856-861
40. Parasnis AS, Ramakrishna W, Chowdari KV, Gupta VS, Ranjekar PK: **Microsatellite (GATA)_n reveals sex-specific differences in Papaya.** *Theor Appl Genet* 1999, **99**:1047-1052
41. Rychlik W, Rhoads RE: **A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA.** *Nucleic Acids Res* 1989, **17**:8543-8551
42. Panaud O, Chen X, McCouch SR: **Frequency of microsatellite sequences in rice (*Oryza sativa* L.).** *Genome* 1995, **38**:1170-1176
43. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410
44. Davierwala AP, Chowdari KV, Kumar Shiv, Reddy APK, Ranjekar PK, Gupta VS: **Use of three different marker systems to estimate genetic diversity of Indian elite rice varieties.** *Genetica* 2000, **108**:269-284
45. Zhao X, Kochert G: **Characterization and genetic mapping of a short, highly repeated, interspersed DNA sequence from rice (*Oryza sativa* L.).** *Mol Gen Genet* 1992, **231**:353-359
46. Zhao X, Kochert G: **Phylogenetic distribution and genetic mapping of a (GGC)_n microsatellite from rice (*Oryza sativa* L.).** *Plant Mol Biol* 1993, **21**:607-614
47. Akagi H, Yokozeki Y, Inagaki A, Fujimura T: **Microsatellite DNA markers for rice chromosomes.** *Theor Appl Genet* 1996, **93**:1071-1077
48. Panaud O, Chen X, McCouch SR: **Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.).** *Mol Gen Genet* 1996, **252**:597-607
49. Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR: **Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.).** *Theor Appl Genet* 1997, **95**:553-567
50. Wang Z, Weber JL, Zhong G, Tanksley SD: **Survey of plant short tandem DNA repeats.** *Theor Appl Genet* 1994, **88**:1-6
51. Epplen JT: **The methodology of multilocus DNA fingerprinting using radioactive or non-radioactive oligonucleotide probes specific for simple repeat motifs. In advances in electrophoresis.** Vol. 5 Edited by A. Chrambach, M.J. Dunn and B.J. Radola. VCH Publishers Inc., New York 1992:62-112
52. Garza JC, Slatkin M, Freimer NB: **Microsatellite allele frequencies in humans and chimpanzees with implications for constraints on allele size.** *Mol Biol Evol* 1995, **12**:594-603
53. Morchen M, Cuguen J, Michaelis G, Hanni C, Saumitou-Laprade P: **Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L.** *Theor Appl Genet* 1996, **92**:326-333
54. van Treuren R, Kuittinen H, Karkkainen K, Baena-Gonzalez K, Savolainen O: **Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*.** *Mol Biol Evol* 1997, **14**:220-229
55. Smulders MJM, Bredemeijer G, Ruskortekaas W, Arens P, Vosman B: **Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species.** *Theor Appl Genet* 1997, **84**:264-272
56. Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A: **Cross species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for the transferability of SSRs in plants.** *Mol Biol Evol* 1998, **15**:1275-1287
57. Provan J, Corbett G, McNicol JW, Powell W: **Chloroplast DNA variability in wild and cultivated rice (*Oryza* spp.) revealed by**

- polymorphic chloroplast simple sequence repeats. *Genome* 1997, **40**:104-110**
58. SM Udupa, Robertson LD, Weigand F, Baum M, Kahl G: **Allelic variation at (TAA)_n microsatellite loci in a world collection of chickpea (*Cicer arietinum* L.) germplasm.** *Mol Gen Genet* 1999, **261**:354-363
 59. Weising K, Gardner RC: **A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genome of dicotyledonous angiosperms.** *Genome* 1999, **42**:9-19
 60. Roder M, Plaschke J, Koenig SU, Boerner A, Sorrells ME, Tanksley SD, MW Ganai: **Abundance, variability and chromosomal location of microsatellites in wheat.** *Mol Gen Genet* 1995, **246**:327-333
 61. Ishii T, McCouch SR: **Microsatellites and microsynteny in the chloroplast genomes of *Oryza* and eight other Gramineae species members.** *Theor Appl Genet* 2000, **100**:1257-1266
 62. Levinson G, Gutman GA: **Slipped strand mispairing: a major mechanism for DNA sequence evolution.** *Mol Biol Evol* 1987, **4**:203-221
 63. Wolff RK, Plaek KR, Jeffreys AJ, White R: **Unequal crossing over between homologous chromosomes is not the major mechanism involved in generation of new alleles at VNTR loci.** *Genomics* 1991, **5**:382-384
 64. Innan H, Terauchi R, Miyashita T: **Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*.** *Genetics* 1997, **146**:1441-1452
 65. Song WY, Wang GP, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P: **A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*.** *Science* 1995, **270**:1804-1806
 66. Ronald PC: **The molecular basis of disease resistance in rice** *Plant Mol Biol* 1997, **35**:179-186
 67. Davierwala AP, Reddy APK, Lagu MD, Ranjekar PK, Gupta VS: **Marker Assisted Selection Of Bacterial Blight Resistance Genes In Rice.** *Biochemical Genetics* 2001, **39**:261-278
 68. Tautz D, Trick M, Dover GA: **Cryptic simplicity in DNA is a major source of genetic variation.** *Nature* 1986, **322**:652-656
 69. Xie Y, Wu R: **Rice alcohol dehydrogenase genes: anaerobic induction, organ specific expression and characterization of cDNA clones.** *Plant Mol Biol* 1989, **13**:53-68
 70. Kay SA, Keith B, Shinozaki K, Chye ML, Chua NH: **The sequence of the rice phytochrome gene.** *Nucleic Acids Res* 1989, **17**:2865-2866
 71. Bligh HFJ, Till RI, Jones CA: **A microsatellite sequence closely linked to the *Waxy* gene of *Oryza sativa*.** *Euphytica* 1995, **86**:83-85
 72. Sambrook J, Fritsch EF, Maniatis T: **Molecular cloning: A Laboratory Manual.** Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2nd ed, 1989
 73. Yanisch-Perron C, Vieira J, Messing J: **Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18, and pUC19 vectors.** *Gene* 1985, **33**:103-119
 74. Rogers SO, Bendich AJ: **Extraction of DNA from plant tissues.** In: Gelvin SB, Schilperoort RA (eds) *Plant Molecular Biology Manual A6.1.* Kluwer Academic Publishers, Dordrecht, 1988
 75. Schaffer HE, Sederoff RR: **Improved estimation of DNA fragment lengths from agarose gels.** *Anal Biochem* 1981, **115**:113-122
 76. Yap IV, Nelson RJ: **Winboot: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms.** *IRRI discussion paper series no 14, Manila, Philippines*, 1996

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright



BioMedcentral.com

Submit your manuscript here:

<http://www.biomedcentral.com/manuscript/>

editorial@biomedcentral.com