Theor Appl Genet (1991) 82: 489–498

THEORETICAL AND APPLIED GENETICS

Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA

M. N. V. Williams, N. Pande, S. Nair, M. Mohan and J. Bennett *

International Center for Genetic Engineering and Biotechnology, NII Campus, Shaheed Jeet Singh Marg, New Delhi-110067, India

Received December 10, 1990; Accepted March 7, 1991
Communicated by G. S. Khush

Summary. Thirty mapped Indica rice genomic (RG) clones were partially sequenced from each end. From such sequence data, pairs of oligonucleotides were synthesized to act as primers for polymerase chain reaction (PCR) amplification of the corresponding loci in crude total DNA preparations. The PCR products from DNA of Indica varieties were of the sizes expected from the sizes of the corresponding RG clones. However, size polymorphisms were seen between PCR products from Indica and Japonica varieties, and among wild *Oryza* species. Restriction fragment length polymorphism (RFLP) was observed between PCR products of Indica varieties simply by electrophoretic analysis of restricted products, without the need for Southern hybridization or radiolabelling. The RFLPs noted between varieties ARC6650 and Phalguna were inherited in recombinant inbred lines derived from a cross between them. The RFLPs were detectable in PCR products amplified from DNA extracted by a simple procedure from single seedlings or leaves, and revealed genetic heterogeneity in cultivated lines. An approach is described that is relevant to the acceleration of classical plant breeding through molecular techniques.

Key words: RFLP – PCR – Rice – Inheritance – Nonradioactive

Introduction

Linkage maps of plant genomes have traditionally been established using biochemical, morphological, and behavioral markers. To these have been added isozyme and DNA markers. The last are particularly powerful, because they have increased the speed and convenience of mapping and promise the production of saturation maps, where the density of markers is so high that any gene of interest is highly likely to cosegregate with flanking DNA markers (Bernatzky and Tanksley 1986). The key technique in the construction of such maps and in the mapping of genes relative to the DNA markers is restriction fragment length polymorphism (RFLP) analysis (Botstein et al. 1980; Burr et al. 1988). This has been used to map both major genes and quantitative trait loci (Paterson et al. 1988; Martin et al. 1989; Tanksley et al. 1989). RFLP analysis detects some of the naturally occurring DNA sequence variations between plants and uses them as genetic markers, which can be followed during progeny analysis in the same way as conventional markers. Southern hybridization of total genomic DNA digested by restriction enzymes (restricted DNA) with a labelled cloned probe reports on DNA sequence variation in and around the locus corresponding to the clone. A shift in the size of a restriction fragment hybridizing to the probe is indicative of a sequence change such as inversion, insertion or deletion, loss or creation of a restriction enzyme site in the region of interest, or in vivo chemical modification of the DNA (Botstein et al. 1980; Chomet et al. 1987).

In spite of its power, RFLP analysis by Southern hybridization poses problems for routine applications in plant breeding. It requires large quantities of high-quality DNA for the detection of single-copy loci, necessitates many manipulations, and detects only a fraction of the existing sequence variability. The advent of the polymerase chain reaction (PCR) provides an additional approach to RFLP analysis which will help to solve these problems. Using short oligonucleotide primers and thermostable DNA polymerase, PCR allows the amplification of specific regions of DNA from complex sequence

* To whom correspondence should be addressed

mixtures such as genomic DNA (Saiki et al. 1988). Very small samples of DNA can be amplified (including DNA from crude or degraded tissue extracts or individual cells) (Paabo et al. 1988; Li et al. 1990) and the procedure is largely automated. The amplified segments can be compared by RFLP analysis directly on stained agarose gels without recourse to Southern hybridization, and if restriction fails to detect a polymorphism between the PCR products amplified from different genotypes, the products can themselves be directly sequenced to reveal all sequence variations between them (Shyamala and Ames 1989). However, unlike Southern hybridization, PCR requires at least a small amount of sequence data on the loci of interest to permit synthesis of appropriate primers (approximately 20 bases long). Here we report partial sequence data for 30 of the RG series of DNA markers mapped onto the rice genome by McCouch et al. (1988). We demonstrate the application of such data to the PCR amplification of the corresponding loci from DNA extracted from *Oryza* species, as well as the ability of PCR to detect novel RFLPs and follow their inheritance, even in single seedlings.

Materials and methods

Plant material

The Directorate of Rice Research (DRR), Hyderabad, India, supplied seeds of two Indica varieties of rice (elite variety Phalguna and ARC6650, a landrace from Assam, India) (Prasada Rao and Kalode 1987), together with seeds of 47 recombinant inbred lines (F_5–F_6) derived from a cross between these parental lines. Other Indica varieties were from DRR, and variety Leah was obtained from M.C. Rush, Louisana State University, USA. The International Rice Research Institute, Los Banos, The Philippines, supplied seeds of the Japonica variety Nipponbare and of the species *O. longistaminata* (accession no. 101378) and *O. officinalis* (accession no. 101150), *O. punctata* (accession no. 101409) and *O. rufipogon* (accession no. 103307).

Plasmids

The RG series of plasmids described by McCouch et al. (1988) were supplied by S. D. Tanksley, Cornell University, Ithaca/NY). These pUC9-based plasmids (previously reported to be in pUC8) were grown in *Escherichia coli* DH5α and extracted by the alkaline lysis procedure (Maniatis et al. 1982).

Restriction endonucleases

Enzymes were obtained from New England Biolabs (USA) or Boehringer Mannheim (Germany) and were used under conditions recommended by the manufacturers.

Rice DNA

Total DNA was extracted from 14-day-old rice seedlings by the method of Walbot (1988) with RNase treatment and phenol extraction, except for single seedlings or leaves.

Plasmid sequencing

The inserts of selected RG plasmids were partially sequenced on one strand by the dideoxy chain termination method of Sanger et al. (1977) using double-stranded plasmid DNA, Mn^{++} buffer modification (Tabor and Richardson 1989) of the U.S. Biochemicals (USA) Sequenase kit, and α-^{35}S dATP. The pUC homologous forward (5′CGCCAAGGTTTTCCCAGTCACGAC-3′) and reverse (5′-AGCGGATAACAATTTCACACAGGA-3′) sequencing primers were synthesized by W. Crockett, Brookhaven National Laboratory, Upton/NY.

PCR primers

Primers for PCR amplification of segments of rice genomic DNA corresponding to RG clones were synthesized on an Applied Biosystems 380 B DNA synthesizer, deprotected with ammonia, and precipitated with ethanol. The primers consisted of 20-bases sequences (20-mers) selected from the terminal sequencing data of RG clones to have a GC content of about 50%, with fewer than four contiguous identical bases.

PCR amplification of rice genomic DNA

PCR reactions contained 50 m*M* KCl, 10 m*M* TRIS-HCl (pH 8.2), 1.5 m*M* $MgCl_2$, 0.001% gelatin, 200 m*M* each of 4 dNTPs, 500 n*M* of each primer, 100–1,000 ng of genomic DNA, and 5 units of *Taq* polymerase (Perkin Elmer/Cetus, USA) in 50–150 µl. Amplification consisted of 30–38 cycles of 1 min at 94 °C (denaturation), 1 min at 55 °C (annealing), and 3–9 min at 72 °C (elongation). Reaction mixes (3–12 µl) or the products of digestion with restriction endonucleases were loaded on 1 or 2% agarose gels containing 2 µg/ml ethidium bromide and electrophoresed at 60–100 V, before visualization under long-wavelength ultraviolet light and photography through a red filter. Marker DNA was 1-kb Ladder (Bethesda Research Laboratories, USA).

Southern hybridization

Rice DNA (5 or 10 µg) was digested with restriction endonucleases and electrophoresed overnight at 25 V on 0.8% agarose gels containing TEB buffer (Maniatis et al. 1982). DNA was transferred to GeneScreen Plus nylon membranes (DuPont NEN, USA) by a simplified Southern hybridization protocol (Schuler and Zielinski 1989), with slight modifications of solutions as follows: 15 min in 0.2 *N* HCl; 2 × 15 min in 1.5 *M* NaCl, 0.5 *M* NaOH; 2 × 10 min in 1 *M* ammonium acetate. All these steps were carried out with gentle shaking. Membranes were subsequently baked in vacuo at 80 °C for 2 h and hybridized (Bernatzky 1988) with RG plasmids that had been ^{32}P-labelled ($>1 \times 10^8$ dpm/µg) using the BRL Nick-translation kit. After hybridization, the membranes were washed under stringent conditions (twice in 2 × SSC, 0.1% SDS at room temperature for 5 min each; twice in 0.5 × SSC, 0.1% SDS at 65 °C for 15 min each; and twice in 2 × SSC at room temperature for 5 min) and autoradiographed. Each blot was used for four or five hybridizations, between which the probe was removed according to GeneScreen protocol [20 min boiling in 10 m*M* TRIS-HCl (pH 8.0), 1 m*M* EDTA, 1% SDS].

Results

Amplification of rice DNA

McCouch et al. (1988) used a large number of single- and low-copy clones derived from IR36 genomic DNA, to prepare a genetic map of the 12 haploid chromosomes of rice. We determined partial sequences for the ends of 30 of these clones. Table 1 lists 40 bases of sequence, suitable

Table 1. DNA sequences of ends of RG clones. Sequences used subsequently for PCR primers are underlined

FORWARD	REVERSE
CHROMOSOME 1	
RG345	
CAGTCCGTATTTGAACTTCATGACGGTTCTGGCATTGTCG	CAGAAGCCATTCTACATGCTTACGGCTGCTCTCCAGAGTG
RG381	
CCTCGTACGTGCTCAGCCGTAGTACTGGCCGGTGCAGGAG	ACACCTATTATATATGCACGGTGTCCATCTATCTTTAGTC
RG236	
TATTCCGCAAACCTGTTGGATGGGAAGTCCCTTGGATCCC	TCTGAGTAAACACACAGAACATTTCAACTGGTCAAATCCC
RG140	
GTAGCTGGTTCAAATCCTCAGGTATTTTGTTAGCTCATAT	GGCCACACCACCATATGCCCATGCACATCCATCCATCGTC
CHROMOSOME 2	
RG157	
CGCTCTATCTCATTTTACAAGCTATTTTGTATACAATATC	GCCTGCTTCAGCAAATTTATGAACTTCTTTGTGAGAAGCC
RG120	
TTCAGTCAATTTATTGACAGGTTTGCTGTCTTCTTTTAAG	GGTCAAAGGTGAATGAATGAATCAATCAATCGATAGGAAA
RG144	
TGGAGGAAAATATCATACTGAAATAATGCAGAGTTGCATC	AACGCGTCGCTGGATTGCCGTACCTTTGTGGATTACGGGG
CHROMOSOME 3	
RG64	
ATGATGCCCAAAAGCTTAGATAGGCAACATCTCAATTAGA	AATACATTCTGCCAATAGAAGTTACTTGCACTAACAAGAA
RG123	
GATATGATGTTGCCATGTAGACCTTGTTTGACTGCTGGGT	GGTTCCTGCCAATTCCTCTTCCTTCTGATTCCATGACCTG
RG172	
ACCGCGTGGGTGCTCTTCTTGACGCCGTTCGAGACCATGA	ACATCTACACACCTAACATATAGACGGCATACAGTGAGTC
CHROMOSOME 4	
RG100	
GAGAAGAGATCGCAAGTAACCTCCTCACTAAGGAGCACGC	CTAGAAAGCGATCGAACCACAGCCACAGCATGCACGGCTG
RG227	
CTGCAAACGAACGAAAGCAAAGTTACAAAGACATGCGCAG	CATACTCCTACGCATGGTGCATATGATCTTGGTCGCTCCA
CHROMOSOME 5	
RG207	
CTGCAGTCTATCTATGATTGCTACGACGAAGATAGCTGGX	AGCAGGTTCATAAAGGCTCAAGTACAGTACTCCTGTGTAA
RG13	
ATTACTCCAGTACTCCATATAGCACTACACCCATACTGTA	TCTAGAGTAGAGGGTGTATGCTCAAAGAAGATAAAAGGCC
RG182	
GATTTAAATTGCTATCTTTGCATGAGACCAACTGGAGTGG	AATAGCTAGGCTTTGCACTGAAACTTTTACTCTTTCTTCT
CHROMOSOME 6	
RG341	
TCTCAACTCAGCTGGATTTCCTTATGGGTCAGTAACACTT	GTCCTCAGCTAGACTCTGACCAACTTGAGATGCTTTGAAT
RG235	
CTTTGCATTGCTGCTCACTTTTATTACTATGCACATGCAT	AGTACATGCATACGCATGCTGATCTATTCTAATTAATCTA
RG190	
TCTTCATCTTCTGCATGTTCAGACAGCTTGTTTCCATCAT	AAGGACTACAAGAATCAGTCTGCGATCCGCAGAATGTTTA
CHROMOSOME 7	
RG351	
ATGCAGAAGTAGTTGTATTTGTCTTGGATCAGGTCTATTA	GTAGTATTTACTCTTGTGTTAGTAGGAGTACTAGTATTTA
RG173	
ATGAGAACTCTGAAACCGCAGCGACAGTGACTGACCGTGG	GCCGCCCTGCATCCCGACCACCGCGCCGGACGTGATCCAC
CHROMOSOME 8	
RG365	
TCAGTTGCTCCTAAAGCTCACCCAGTCAGTTCGCATTATC	CCAGCAAATATACTTGAAATGCAAAATTAAGAACATTCAA
CHROMOSOME 9	
RG386	
CTGCAGCAGAGTACTGTGCAACTGAXCXCTGTTCAGATCA	ATCTGATTGTCCCCATGTATGCTGGCTAATTCTATCAGCC
RG136	
CCATGCATGCATGAACCAATGCAAATGCTTTTGATAATAT	TCAGATTGTAATGTACTCCAACTAATCCATGTTATAGAAA
CHROMOSOME 10	
RG241	
CATACTTTTCTGACCTAGTGCAGTAGTAAGACCACCTCAG	GAACAATTAACTGCTCACAGCAATATGTAACTAGTTCTTT
RG257	
GATTGAGCAGGTACACTACTGAGGTTCTGGGGGTAAAGGT	CCATATGAATGGAGGGAAAAATAATAGGAGAAATTCAAGC
CHROMOSOME 11	
RG303	
TATCACCTCTATCAAAGAAGTGTGCTGAGGTGATAAAATC	ACAACTAAGAGCTAGTACTTTATATCGCAAGTTCCATGAC
RG118	
AGATTACAGTTTTAAGGTAGTTAGCAGCTGCCGTCACATG	TTTCCAGCAGATCATCAGGGATGCCAGTGTTCACCAGCTG
CHROMOSOME 12	
RG143	
CAACCGAATCACCTAAGAATAGACTGGCACAATGAATGCG	CTCAGCAACAGTCACTGCAATTATGCAAACCATACAGAGT
RG214	
CTAAAACGCCTATATCAGCGGCTTTCTAGGGTTCCACTCG	TGTATTCATTTGATGAGAAATCACGCTCAATCCAGGTGGA
RG329	
GCCAAAGCTCTCCAGTACTTAGACATGCTTCTGGACAAAT	AACTGAGAGCTACTAAAATCTCTCACACCCAAGGCCATGA

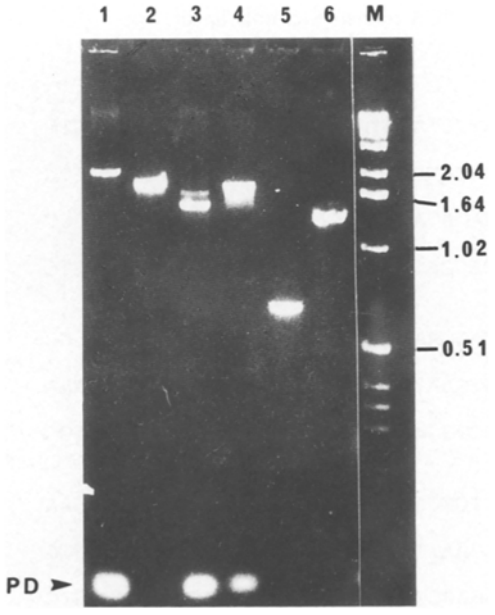

Fig. 1. Ethidium bromide staining of PCR products for six single-copy loci amplified from total DNA of rice variety ARC6650. Lane 1: *RG13*; lane 2: *RG118*; lane 3: *RG214*; lane 4: *RG303*; lane 5: *RG351*; lane 6; *RG386*; lane M: marker DNA. Primer-dimer is indicated (PD)

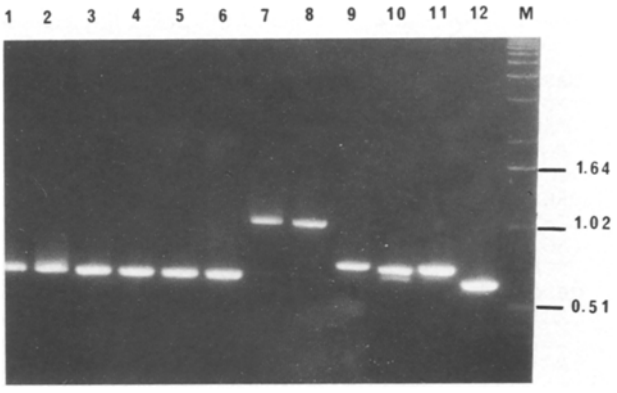

Fig. 2. Ethidium bromide staining of PCR products for locus *RG351* amplified from total DNA of Indica and Japonica rice varieties, as well as wild species. Lane 1: ARC6650; lane 2: Phalguna; lane 3: Basmati 370; lane 4: IR50; lane 5: Swarnadhan; lane 6: Velluthacheera; lane 7: Leah; lane 8: Nipponbare; lane 9: *O. longistaminata*; lane 10: *O. officinalis*; lane 11: *O. punctata*; lane 12: *O. rufipogon*; lane M: marker DNA

for preparation of PCR primers, for the two ends of each clone. All 12 haploid chromosomes of rice are represented.

For several of the sequenced clones, pairs of oligonucleotides (20-mers) were synthesized and used as PCR primers. The sequences of some of the primers used are underlined in Table 1. Figure 1 shows PCR amplification products for six pairs of primers hybridizing to genomic DNA extracted from the Assam landrace ARC6650. For each pair of primers, a single major PCR product was detectable by agarose gel electrophoresis and staining with ethidium bromide. The size of each product was close to that of the original cloned insert derived from IR36. The fact that amplification was successful at high stringency of primer:template annealing indicates adequate sequence conservation between IR36 and ARC6650 with respect to regions hybridizing to the primers. The fact that the products were of the expected sizes indicates that no large-scale inversion, addition, or deletion has occurred within these six segments since divergence of the lineages leading to IR36 and ARC6650. The low-molecular-weight primer-dimer artifact (PD) is seen with some pairs of primers but not with all, a result consistent with its occurrence being dependent on primer sequence.

The size of PCR-amplified product was approximately the same from a number of Indica rice varieties at all the loci examined. The fragment size for some loci such as *RG351,* however, varied when amplified from DNA of Japonica variety Nipponbare and from wild *Oryza* species (Fig. 2). Amplification of a single major band using primers based on IR36 sequence indicates that there is a high degree of conservation over short stretches of DNA in the genus *Oryza*. Differences in the size of the amplified fragments presumably represent insertion or deletion events between the conserved regions. Such differences are expected to become more frequent as genetic distance increases. A large number of desirable agronomic traits such as pest and stress resistance are present in wild *Oryza* species, but crosses between wild and cultivated varieties are difficult and progeny may have to be recovered by embryo rescue techniques. PCR may thus provide an easy means to determine the specific origin of chromosome segments in semisterile progeny embryos from wide crosses.

Detection of RFLPs without Southern hybridization

We are performing RFLP analysis on progeny of a cross between ARC6650 and Phalguna. The objectives of this work are, firstly, to map several insect resistance genes onto the genetic map of the rice genome prepared by McCouch et al. (1988) and, secondly, to use these map data to accelerate breeding for enhanced insect resistance. By identifying RFLP loci linked to resistance genes, pyramiding of multiple genes for resistance to the same insect becomes possible, and it is believed that such multiple resistances are much less likely to break down under field conditions. It is clear from Fig. 2 that PCR can detect size polymorphisms at certain loci, especially when DNAs from distantly related plants are amplified. In more closely related varieties, such as the Indica lines ARC6650 and Phalguna, none of the amplified loci directly show size polymorphism. We have consequently

Table 2. Average frequency of cleavage sites for restriction endonucleases in rice loci amplified by PCR

Restriction endonuclease	Recognition sequence	Average no. of sites per kb	No. of loci examined
*Mae*II	ACGT	1.13	7
*Alu*I	AGCT	4.18	7
*Nla*III	CATG	2.71	7
*Mae*I	CTAG	1.92	7
*Sau*3AI	GATC	1.81	7
*Rsa*I	GTAC	2.94	7
*Taq*I	TCGA	1.02	7
*Hpa*II	CCGG	1.13	7
*Bst*UI	CGCG	0.34	7
*Hha*I	GCGC	1.13	7
*Hae*III	GGCC	1.13	7
*Mse*I	TTAA	2.36	6
*Hin*fI	GANTC	1.81	7
*Scr*FI	CCNGG	0.71	3
*Dde*I	CTNAG	1.22	4
*Fnu*4HI	GCNGC	1.63	4
*Sau*96I	GGNCC	0.57	6
*Mae*III	GTNAC	1.74	6
*Mnl*I	CCTC	1.11	5

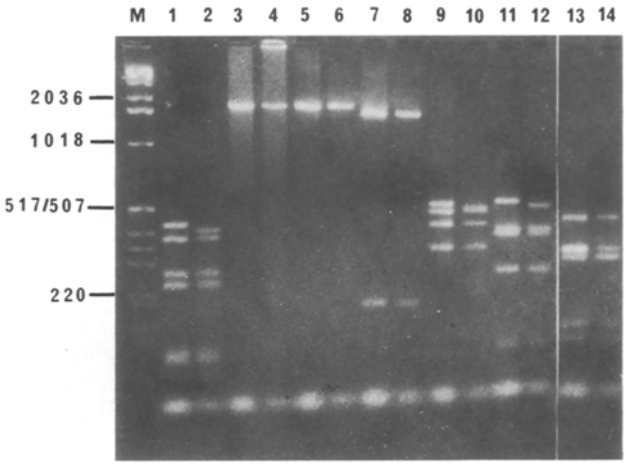

Fig. 3. Ethidium bromide staining of restriction fragments generated from PCR products for locus *RG303* amplified from total DNA of ARC6650 (lanes 1, 3, 5, 7, 9, 11, 13) and Phalguna (lanes 2, 4, 6, 8, 10, 12, 14). Lane M: DNA marker. Lanes 1–14: PCR products digested with *Alu*I (1, 2), *Bst*UI (3, 4), *Hpa*II (5,6), *Hha*I (7, 8), *Mae*I (9, 10), *Mae*III (11, 12), or *Rsa*I (13, 14)

digested the amplification products with several four-nucleotide recognizing restriction endonucleases (4-cutters) to detect base changes or small insertions and deletions. In a random DNA sequence, 4-cutter sites should occur every 256 bp (4^4), and each enzyme should therefore identify 16 bp ($1{,}000/256 \times 4$) of sequence per kilobase. We have examined digestion at 19 different restriction sites which, barring overlap, theoretically represent up to 30% of all bases in any segment of random DNA. A significant proportion of single base pair changes could thus be examined by restriction enzyme digestion. The number of actual sites found may be lower than anticipated, because genomic DNA sequences are not random and fragments smaller than about 50 bp were not resolved on our gels; consequently, sites that fall closer together than this are not recorded. We found that there is significant variation in the number of sites for different enzymes in the mapped single-copy loci (Table 2): sites for certain enzymes such as *Alu*I and *Rsa*I are more common than others and thus may be more useful in preliminary surveys for polymorphism. For example, Fig. 3 shows the patterns of digestion when the *RG303* segment is amplified from ARC6650 and Phalguna and then cut with 4-cutters. Two of these seven enzymes failed to cut the PCR products, two other enzymes cut but the patterns of fragments were indistinguishable, while three enzymes (*Alu*I, *Mae*I, and *Mae*III) yielded patterns where one fragment was clearly different in size between the two varieties. The results suggest that the Phalguna segment amplified by the RG303 primers is actually about 20 bp shorter than the corresponding segment of ARC6650, a difference too small to be detected in the undigested PCR products (about 1,800 bp).

An RFLP that is probably due to base substitution is clearly visible when the enzyme *Rsa*I is used to cut the PCR products (~1,900 bp) of the *RG118* locus after amplification from ARC6650 and Phalguna (Fig. 4). Three *Rsa*I fragments (~660, ~550, and ~300 bp) are identical between the two products, but a 260-bp fragment of Phalguna (lane 2) contains an additional *Rsa*I site in ARC6650 (lane 1), yielding 200- and 60-bp fragments (60-bp fragment not shown in Fig. 4). Mixture of the two amplified DNAs followed by digestion results in the appearance of the 260-, 200-, and 60-bp bands (lane 3), suggesting that the 260-bp band is not a partial digestion product caused by an inhibitor in the ARC6650 DNA.

Analysis of recombinant inbred lines

The mapping of insect resistance genes of ARC6650 and Phalguna onto the genetic map of the rice genome depends upon analysis of the segregation of resistance phenotypes and DNA markers in an appropriate mapping population. This population may consist of F_2 individuals derived from a cross between the two parents, or of a set of recombinant inbred (RI) lines derived from F_2 individuals. The latter is intrinsically the more powerful option, because the mapping population is then essentially immortal and the RFLP mapping data base can be progressively supplemented with additional DNA and phenotypic markers. The F_2-based mapping population cannot be used once those individual plants have died;

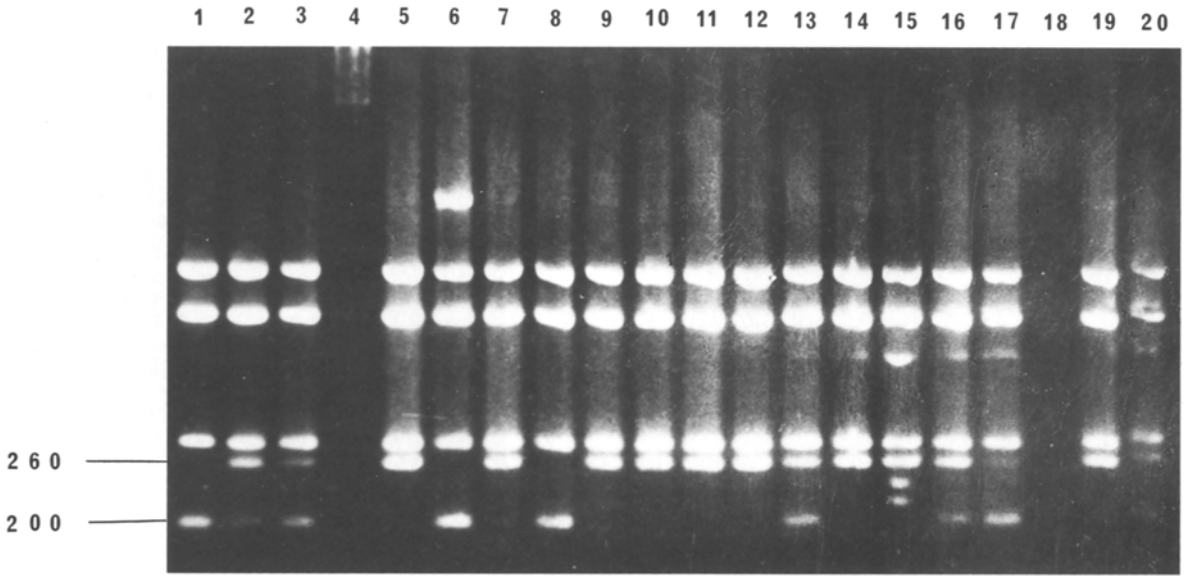

Fig. 4. Ethidium bromide staining of *Rsa*I fragments of PCR products for locus *RG118* amplified from DNA of ARC6650, Phalguna, and 16 recombinant inbred lines derived from a cross between them. Lane 1: ARC6650; lane 2: Phalguna; lane 3: predigestion mix of lanes 1 and 2; lane 4: DNA markers; lanes 5–20: recombinant inbred lines Ryt 74 (5), Ryt 11468 (6), Ryt 9686 (7), Ryt 8653 (8), Ryt 346 (9), Ryt 75 (10), Ryt 318 (11), Ryt 9701 (12), Ryt 65 (13), Ryt 18 (14), Ryt 11516 (15), Ryt 11469 (16), Ryt 12074 (17), Ryt 24 (18), Ryt 365 (19), and Ryt 12197 (20). Some of the bands, in particular the 200-bp band of lane 13, are faint

during their lifetime they can be used to map only a few phenotypes and then only if individual plants are divided into multiple tillers. Accordingly, we have acquired 47 recombinant inbred lines of the ARC6650 × Phalguna cross and are comparing them for RFLPs, both by Southern hybridization and PCR.

Figure 4 shows the inheritance of the additional *Rsa*I site of locus *RG118* among 16 recombinant inbred lines. DNA extracted from about ten seedlings of each line was pooled. DNA from 1 line (lane 18) failed to amplify adequately after 30 cycles, while DNA from the other 15 lines amplified to approximately equal extents. Allowing for two cases of incomplete digestion (lanes 6 and 15), ten lines have inherited an allele from Phalguna (118P), while three lines correspond to the ARC6650 allele (118A). This bias in favor of the Phalguna phenotype will be discussed below. PCR amplification of the *RG303* locus, followed by digestion with *Alu*I, also shows segregation of a small insertion polymorphism (six Phalguna to seven ARC6650) (results not shown). *RG118* and *RG303* both map to chromosome 11, but are greater than 50 map units apart (McCouch et al. 1988), and our PCR results indicate that they segregate independently in the RI lines (Table 3).

Table 3. Inheritance of mapped loci in recombinant inbred lines as determined by PCR. A = ARC6650 allele, P = Phalguna allele, H = heterozygous

Variety	*RG214*	*RG303*	*RG118*
Ryt 74	P	A	P
Ryt 11468	P	P	A
Ryt 346	P	A	P
Ryt 8653	P	P	A
Ryt 9686	P	P	P
Ryt 75	P	A	P
Ryt 318	P	A	P
Ryt 9701	P	P	P
Ryt 65	P	A	H
Ryt 18	P	H	P
Ryt 11516	P	H	P
Ryt 11469	P	A	H
Ryt 12074	P	P	H
Ryt 24	P	P	–
Ryt 365	P	A	P
Ryt 12197	P	H	P

Analysis of individual seedlings

Of particular interest in Fig. 4 are the lines Ryt 65 and Ryt 11469 (lanes 13 and 16), which appear to contain an unequal mixture of the two parental alleles, with the Phalguna allele predominant. This phenomenon was found to be due to the pooling of DNA from several plants of these inbred lines. Figure 5 shows the analysis of DNA from 17 individual seedlings of line Ryt 65. Among these seedlings, 2 were homozygous 118A, 14 were homozygous 118P, and 1 was heterozygous. Thus, the pattern seen in lane 13 of Fig. 4 results from a mixture of seed from a line that was clearly still segregating at locus *RG118* after five to six generations of inbreeding.

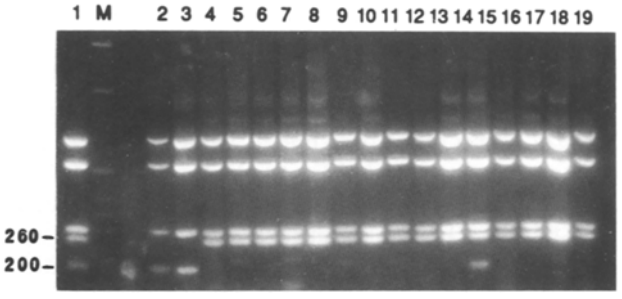

Fig. 5. Ethidium bromide staining of *Rsa*I fragments of PCR products for locus *RG118* amplified from total DNA of single seedlings of Ryt 65. Lane 1: batch DNA prep; lane M: marker DNA; lanes 2–19: single seedlings

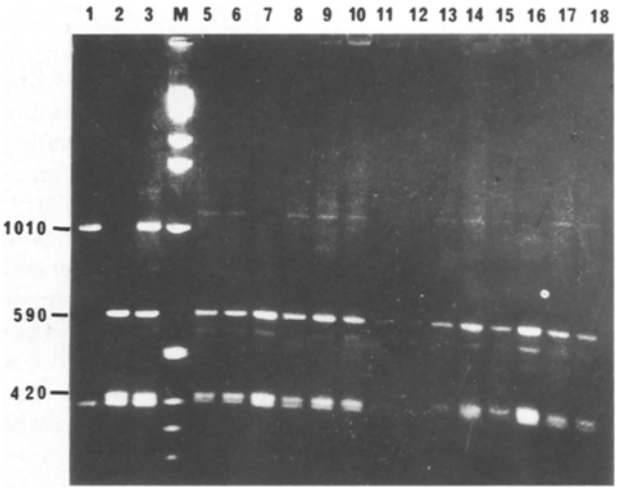

Fig. 6. Ethidium bromide staining of *Mse*I fragments of PCR products for locus *RG214* amplified from total DNA of ARC6650, Phalguna, and 14 recombinant inbred lines derived from a cross between then. Lane designation as in Fig. 4

Even in a selfing plant such as rice, residual heterozygosity is not unexpected at this stage. Single seedling analysis of Phalguna indicated that this parental line also carried a mixed population of alleles at this locus, whereas mixed seed has not been detected in ARC6650 at any of the loci examined. Mixed seed populations in cultivated lines of rice have been previously observed (Miezan and Ghesquire 1986; McCouch et al. 1988). Wang and Tanksley (1989) reported that 26% of Indica, Japonica, and Javanica cultivated varieties contain detectable polymorphism. Different recombinant inbred lines show heterozygosity at different loci (Table 3), consistent with segregation of residual heterozygosity, rather than cross-pollination or postharvest intermixture of seed. Given the sensitivity of DNA-based genetic analysis, it is likely that many inbred varieties and recombinant inbred lines will prove to be populations of genotypes and should be sampled as such.

Bias in recombinant inbred lines

The ideal mapping population for RFLP analysis would be an unbiased set of recombinant inbred lines derived from a cross between two very different parents. This would maximize the number of phenotypically important genes that could be mapped with that particular population (Burr et al. 1988). There should be at least 40 lines in the population and each line should contain the allele of only one parent at as many loci as possible, partly to ensure that lines are true breeding and partly to simplify cosegregation analysis. Even using single seedling descent, at least five seasons of inbreeding are required to reduce the frequency of heterozygosity below about 3%. Consequently, students of RFLP analysis may not initially breed their own recombinant inbred lines; rather, they may use existing sets of inbred lines made available from past breeding programs. This is the approach that we have adopted in our study of ARC6650 and Phalguna. Unfortunately, most breeding programms have strong biases built into them, such as selection for dwarf habit, high yield, early flowering, and so on. This was undoubtedly the case for the progeny of the ARC6650 × Phalguna cross, where the aim was to introgress into high-yielding Phalguna only a few agronomic traits from ARC6650, such as brown plant-hopper resistance (E.A. Siddiq, M.V.S. Sastry, and U. Pradasa Rao, personal communication). For this reason, DNA markers that are tightly linked to loci determining important Phalguna traits will tend to be overrepresented in the population of recombinant inbred lines. Figure 6 shows an example of such a DNA marker: locus *RG 214*. The PCR products for this locus from ARC6650 and Phalguna were about 1,500 bp. Cleavage with *Mse*I yielded two fragments for the ARC6650 allele (~1,010 and 400 bp) and three for the Phalguna allele (~590, 420, and 400 bp); that is, the Phalguna allele contained an additional *Mse*I site in the 1,010-bp fragment. Out of the 14 recombinant inbred lines examined, 12 amplified adequately and showed the Phalguna allele; the 2 that failed to amplify adequately later proved also to have the Phalguna allele. Thus, the *RG214* locus is very tightly linked to a locus of importance in determining a key aspect of the Phalguna phenotype.

Analysis of loci *RG214* and *RG118* by Southern hybridization confirms the patterns of inheritance detected by PCR. However, the polymorphisms detected by the two strategies do not represent the same physical DNA changes (i.e., the enzymes that detect polymorphism do not recognize the same sequences) (data not shown). Fine structure RFLP analysis of PCR-amplified DNA therefore detects a different subset of genetic changes.

Discussion

We have demonstrated the use of PCR to detect RFLPs and to follow their inheritance in rice without resorting to Southern hybridization. PCR primers were synthesized from sequence data that we obtained for the terminal regions of single-copy RG clones of McCouch et al. (1988). These clones had been prepared from genomic DNA of IR36 (an Indica rice), were of known size, and had been mapped onto the 12 haploid chromosomes of rice (McCouch et al. 1988). In most cases PCR amplification from total genomic DNA of Indica rices yielded a single, specific product of the expected size. RFLPs were observed between these PCR products after cleavage with restriction endonucleases and analysis of restriction fragments, by agarose gel electrophoresis and staining with ethidium bromide. Some of these RFLPs were due to size polymorphisms that were too small (about 20 bp) to be seen in undigested PCR products (800–2,200 bp) and could represent footprints of mobile elements (Dorin and Starlinger 1986). Other RFLPs were due to the presence/absence of particular restriction enzyme sites and likely result from point mutations. However, when PCR products of DNAs from Indica, Japonica, and wild rice species were compared, size polymorphisms were detected without the need for digestion with restriction endonucleases.

RFLP analysis is usually performed by Southern hybridization. The advent of highly sensitive, non-radiochemical methods for labelling DNA probes will increase the ease with which Southern hybridization is applied to RFLP analysis. However, PCR-based protocols for RFLP such as described here dispense with the labor-intensive Southern hybridization procedure entirely. There are several contexts in which PCR might be an attractive alternative to Southern hybridization. The first concerns loci that are apparently nonpolymorphic by Southern hybridization. An examination of genetic maps of genomes based on RFLP analysis of random DNA probes reveals many regions that are poorly covered with markers (McCouch et al. 1988). This could be due merely to chance or it could be due to an unexpectedly high degree of relatedness in these regions between the parental lines used to generate the mapping populations. In either case, it is essential that the few markers that have been mapped to these regions are polymorphic for the parental lines used in RFLP analysis. Standard Southern hybridization using hexanucleotide-recognizing restriction enzymes can discriminate between different DNAs at a given locus, if at least one restriction enzyme can be found to cut the DNAs differently in or adjacent to that region. Such analysis is thought to detect primarily insertions and deletions, since polymorphism is more commonly found in large restriction fragments (McCouch et al. 1988; Miller and Tanksley 1990). We routinely use a set of six hexanucleotide-recognizing enzymes to cut the DNAs of ARC6650 and Phalguna, and we find an RFLP with at least one of these enzymes for about 66% of 56 loci. The remaining loci are nonpolymorphic according to this approach and as such are useless for RFLP mapping. The usual approach to the rescue of such loci is to use additional restriction endonucleases, but this can be very wasteful of genomic DNA and is unlikely to detect small DNA changes. Restriction enzyme digestion of PCR products detects a different subset of DNA changes and might thus rescue loci that appear to be monomorphic by Southerns. Even where an RFLP is not detectable, fine structure comparison of DNA from a particular locus is possible from PCR products. Single-base-pair mismatches may be detected by denaturing gradient gel electrophoresis (Meyers et al. 1987) or chemical cleavage (Cotton et al. 1988). Ultimately, direct sequencing of amplified products followed by high stringency oligonucleotide hybridization (Ikuta et al. 1987; Tiercy et al. 1988) may allow exploitation of any base pair change. This additional effort would be justifiable for markers in regions haboring genes of interest, since the closer flanking DNA markers are to such genes, the lower the frequency of recombination between them and the lower the frequency of false negative and false positive results. We hope that the publication here of sequence data for 30 such loci covering all 12 chromosomes of rice will lead others to use this approach.

RFLP patterns detected by Southern hybridization can be affected by modification of DNA such as methylation, since it is known that many restriction enzymes are sensitive to methylation (Nelson and McClelland 1987) and it is likely that other DNA modifications can affect their activity. An important feature of RFLPs detected in PCR products is that they are not subject to artifacts such as those caused by differences in methylation of the DNA. This may be especially important in analysis of DNA from callus cultures and plants regenerated from culture, which show abnormal and often augmented patterns of methylation (Brown et al. 1990).

PCR-based RFLP analysis will be applicable in plant breeding programs once several important phenotypic traits have been mapped relative to cloned DNA markers. For each trait, it would be necessary to synthesize only two pairs of PCR primers, corresponding to the ends of the clones of the two flanking markers. This approach would be especially useful in the case of stress resistance genes. The inheritance of such genes is now monitored largely by phenotypic analysis, and only one trait can be introgressed at a time, unless different tillers are subjected to different stresses, or unless the gene is linked closely with a morphological marker. However, when mapped genes are monitored by detection of flanking DNA markers, many traits can be monitored simultaneously by exploiting the ability of different pairs of

PCR primers to amplify specific loci from the DNA obtainable from a single leaf of individual plants. We have shown here the inheritance of loci *RG118* and *RG303* in individual seedlings of a recombinant inbred line of the ARC6650 × Phalguna cross. It should be noted that PCR not only requires much less DNA than that required for Southern hybridization but also tolerates a much simpler DNA extraction protocol. The ribonuclease treatment and phenol extraction used to extract DNA for Southern hybridization can be dispensed with in the case of PCR amplification. DNA extracted from single leaves of seedlings can be used to determine the genotype of many recombinants early in the growing season, and those plants with a particular set of attributes can be quickly selected for further crossing or phenotypic analysis. Finally, unlike Southern hybridization, the PCR technique is already largely automated and need not use radioactivity. These features may make PCR an attractive alternative in the application of RFLP analysis to plant breeding programs outside specialized laboratories. The major disadvantage of PCR is the current expense of the technology. This is likely to diminish rapidly as sequence data, automated water baths, and primers become more readily available and mapping data on agronomically important genes are accumulated. Exchange of primers between laboratories will also greatly facilitate the use of this approach.

Little is known about the molecular origins of RFLPs. In principle, point mutations, inversions, deletions, or insertions could be responsible (McCouch et al. 1988; Miller and Tanksley 1990). When RFLPs are detected by Southern hybridization, only a minority of DNA sequence alterations responsible for these RFLPs will lie within the regions hybridizing with the probes. The majority will lie at unknown sites outside the hybridizing regions. A complicated cloning strategy would be required to identify the site of the alteration. In contrast, mutations responsible for RFLPs detected by PCR lie within the amplified segments themselves and may be characterized fully by sequencing of the PCR products or fragments thereof. We have preliminary evidence that transposable elements or their footprints have been responsible for some of the polymorphisms detected in the comparison of different rice varieties.

In summary, PCR provides a fast and simple method of following inheritance of DNA markers, using only small quantities of biological material and without using radioactivity. In addition, the technique may allow rescue of apparently nonpolymorphic mapped loci, and may simplify the characterization of mutations responsible for polymorphism.

Acknowledgements. We thank E.A. Siddiq, M.V.S. Sastry, U. Prasada Rao (Directorate of Rice Research, Hyderabad, India), R. N. Misra, A. Krishnamurty (Central Rice Research Institute, Cuttack, India), T.T. Chang (International Rice Research Institute, the Philippines), and M.C. Rush (Louisiana State University, USA) for seed and helpful discussions. We are grateful to S. Tanksley for RG clones, A. Bhardwaj (ICGEB) for synthesis of primers, R. Jefferson (FAO) for comments on the manuscript, and the Rockefeller Foundation for support.

Note added in proof:

The sequences in table 1, together with additional sequence data from RG loci, are available in the EMBL Sequence Data Library, accession numbers X57860 to X57923.

References

Bernatzky R (1988) Restriction fragment length polymorphism. In: Gelvin SB, Schilperoort RA (eds) Plant molecular biology manual. Kluwer, Dordrecht, pp C2:1–18

Bernatzky R, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Botstein D, White RL, Skolnick M, Davis R (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Brown PTH, Kyozuka J, Sukekiyo Y, Kimura Y, Shimamoto K, Lorz H (1990) Molecular changes in protoplast-derived rice plants. Mol Gen Genet 223:324–328

Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW (1988) Gene mapping with recombinant inbred lines in maize. Genetics 118:519–526

Chomet PS, Wessler S, Dellaporta SL (1987) Inactivation of the maize transposable element *Activator* is associated with DNA methylation. EMBO J 6:295–302

Cotton RGH, Rodriguez NR, Campbell RD (1988) Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc Natl Acad Sci (USA) 85:4397–4401

Doring H-P, Starlinger P (1986) Molecular genetics of transposable elements in plants. Annu Rev Genet 20:175–200

Ikuta S, Takagi K, Wallace RB, Itakura K (1987) Dissociation kinetics of 19 base-paired oligonucleotide-DNA duplexes containing different single mismatched base pairs. Nucleic Acids Res 15:797–811

Li H, Cui X, Arnheim N (1990) Direct electrophoretic detection of the allelic state of single DNA molecules in human sperm by using the polymerase chain reaction. Proc Natl Acad Sci USA 87:4580–4584

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY

Martin B, Nienhius J, King G, Schaefer A (1989) Restriction fragment length polymorphisms associated with water use efficiency in tomato. Science 243:1725–1728

McCouch S, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

Meyers RM, Maniatis T, Lerman LS (1987) Detection and localization of single-base-pair changes by denaturing gradient gel electrophoriesis. Methods Enzymol 155:501–527

Miezan K, Ghesquiere A (1986) Genetic structure of African traditional rice cultivars. In: Rice genetics. Proc Int Rice Genet Symp. International Rice Research Institute, Manila, pp 932

Miller JC, Tanksley SD (1990) Effect of different restriction enzymes, probe source, and probe length on detecting restriction fragment length polymorphism in tomato. Theor Appl Genet 80:385–389

Nelson M, McClelland M (1987) The effect of site-specific methylation on restriction-modification enzymes. Nucleic Acids Res 15:219–230

Paabo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year-old brain. Nucleic Acids Res 16:9775–9787

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantiative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Prasada Rao U, Kalode MB (1987) Genetics and linkage between genes for purple coleoptile and leaf with brown planthopper resistance in rice. Indian J Genet 47:333–336

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Ehrlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Schuler MA, Zielinski E (1989) Methods in plant molecular biology. Academic Press, New York, pp 171

Shyamala V, Ames GF-L (1989) Amplification of bacterial genomic DNA by the polymerase chain reaction and direct sequencing after asymmetric amplification: application to the study of periplasmic permeases. J Bacteriol 171:1602–1608

Tabor S, Richardson CC (1989) Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. Proc Natl Acad Sci USA 86:4076–4080

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. BioTechnol 7:257–264

Tiercy J-M, Gorski J, Jeannet M, Mach B (1988) Identification and distribution of three serologically undected alleles of *HLA-DR* by oligonucleotide-DNA typing analysis. Proc Natl Acad Sci USA 85:198–202

Walbot V (1988) Preparation of DNA from single rice seedlings. Rice Genet Newslett 5:149–151

Wang ZY, Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. Genome 32:1113–1118