
DNA polymorphisms amplified by arbitrary primers are useful as genetic markers

John G.K.Williams*, Anne R.Kubelik, Kenneth J.Livak, J.Antoni Rafalski¹ and Scott V.Tingey¹
Central Research and Development Department and ¹Agricultural Products Department, E.I. du Pont de Nemours & Co., Inc., Experimental Station, Wilmington, DE 19880, USA

Received August 23, 1990; Revised and Accepted October 8, 1990

ABSTRACT

Molecular genetic maps are commonly constructed by analyzing the segregation of restriction fragment length polymorphisms (RFLPs) among the progeny of a sexual cross. Here we describe a new DNA polymorphism assay based on the amplification of random DNA segments with single primers of arbitrary nucleotide sequence. These polymorphisms, simply detected as DNA segments which amplify from one parent but not the other, are inherited in a Mendelian fashion and can be used to construct genetic maps in a variety of species. We suggest that these polymorphisms be called RAPD markers, after Random Amplified Polymorphic DNA.

INTRODUCTION

Genetic maps comprising closely-spaced DNA markers are useful for genome analysis. DNA markers that are shown to be genetically linked to a trait of interest can be used for gene cloning, medical diagnostics, and for trait introgression in plant and animal breeding programs (1, 2). In many organisms, however, saturated genetic maps are not available. The DNA markers most commonly used are restriction fragment length polymorphisms (RFLP, 3). Anonymous low copy number genomic clones are frequently used to visualize polymorphisms. Detection of RFLPs by Southern blot hybridizations are laborious and incompatible with the high analytical throughput required for many applications (4). Other polymorphism assays (5) that are based on the polymerase chain reaction (PCR), require target DNA sequence information for the design of amplification primers. The time and cost of obtaining this sequence information is prohibitive for many large scale genetic mapping applications. Here we describe a simple process, distinct from the PCR process, which is based on the amplification of genomic DNA with single primers of arbitrary nucleotide sequence. These primers detect polymorphisms in the absence of specific nucleotide sequence information, and the polymorphisms function as genetic markers, and can be used to construct genetic maps.

MATERIALS AND METHODS

Primer synthesis

Oligodeoxynucleotide primers were synthesized by standard phosphoramidate chemistry on a DuPont Coder 300 DNA synthesizer. After removal of protecting groups in 30% ammonium hydroxide at 55° for 5 h, the samples were dried under vacuum, dissolved in 200 µl of water, and purified by gel filtration on Sephadex G25 (NAP-5 disposable columns, Pharmacia).

Sources of genomic DNA

Human DNA samples from anonymous individuals were obtained from Drs. John Gilbert and Allen Roses of Duke Medical Center, Duke University. Soybean DNA was isolated (6) from the inbred cultivars *Glycine max* variety Bonus, and *Glycine soja* accession PI 81762, and from 88 F₂ individuals segregating from a cross of these two parents obtained from Dr. Theodore Hymowitz, University of Illinois. Corn DNA was isolated (7) from the *Zea mays* lines CM37 and T232 obtained from Dr. Ben Burr, Brookhaven National Laboratory. *Neurospora crassa* DNA samples were isolated from the strains Oak Ridge FGSC 4488 and Mauriceville FGSC 2225 (8), obtained from Dr. R.L. Metzberg, University of Wisconsin. Bacterial DNA samples were obtained from Dr. John Webster, DuPont Co.

Amplification conditions

Amplification reactions were performed in volumes of 25 µl containing 10mM Tris-Cl, pH 8.3, 50 mM KCl, 2 mM MgCl₂, 0.001% gelatin, 100 µM each of dATP, dCTP, dGTP and TTP (Pharmacia), 0.2 µM primer, 25 ng of genomic DNA, and 0.5 unit of *Taq* DNA polymerase (Perkin Elmer Cetus). Amplification was performed in a Perkin Elmer Cetus DNA Thermal Cycler programmed for 45 cycles of 1 min at 94°, 1 min at 36°, 2 min at 72°, using the fastest available transitions between each temperature. Amplification products were analyzed by electrophoresis in 1.4% agarose gels and detected by staining with ethidium bromide.

Annealing temperatures above 40° in the thermal cycling profile prevented amplification by many of the 10 base

* To whom correspondence should be addressed

oligonucleotides tested (data not shown). With some combinations of primer and genomic DNA template a non-discrete size range of amplification products, appearing as a 'smear' as visualized on a gel, could be converted to discretely sized bands by reducing the concentration of either the polymerase or the genomic DNA.

Genetic analysis of amplified DNA polymorphisms.

Amplified polymorphic DNA fragments and RFLPs were mapped in the context of a 436 marker RFLP map of soybean (manuscript in preparation), a 110 marker RFLP map of corn (9), or an 80 marker RFLP map of *N. crassa* (10), by scoring marker segregation in the respective populations used to create these maps. Multipoint maps and LOD scores were calculated using the Mapmaker program (11).

Use of amplified segments as RFLP probes

Polymorphic DNA segments AP3.1 and AP11a.1, amplified (as described above) from the soybeans Bonus and PI 81762 respectively, were resolved by electrophoresis in a 1.4% agarose gel and excised from the gel. A 5 μ l slice of gel containing DNA was added to a 100 μ l reaction mixture and was amplified under the conditions described above, using primers AP3 and AP11a, respectively. The re-amplified DNA samples were labeled with 32 P (BRL Random Primers DNA Labeling System, Life Technologies, Inc.) and used as hybridization probes (12) to detect RFLPs with the restriction enzymes *Pst* I for AP3.1 and *Bcl* II for AP11a.1. These RFLPs were mapped in soybean as described above.

RESULTS

Polymorphism and sequence specificity.

Figure 1A shows the results of an experiment in which single primers were used to amplify segments of genomic DNA from humans, corn, soybean, and *N. crassa*. Primers were designed

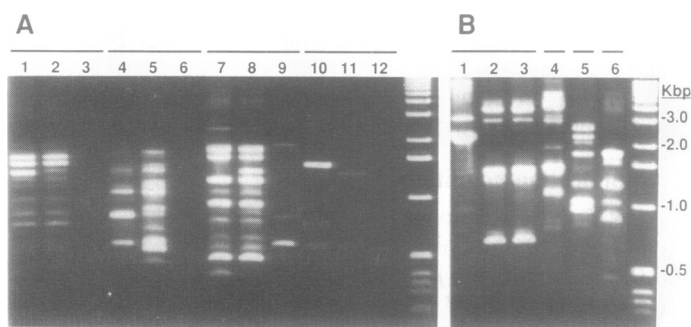


Figure 1a. Amplification of eukaryotic DNA. DNA was amplified from a variety of species using primers of arbitrary nucleotide sequence (Materials and Methods). Amplification products were resolved by electrophoresis in a 1.4% agarose gel which was stained with ethidium bromide and photographed. Molecular weight markers (kilobase pairs, kbp) are as indicated. Lane 1 and 2, human DNA Hu2 and Hu3, respectively, amplified with primer 5'-ACGGTACACT. Lanes 4 and 5, corn CM37 and T232, respectively, amplified with GCAAGTAGCT. Lanes 7 and 8, soybean *G. max* and *G. soja*, respectively, amplified with CGGCCCTGT. Lanes 10 and 11, *N. crassa* Oakridge and Mauriceville, respectively, amplified with CACATGCTTC. Genomic DNA was omitted in control reactions (lanes 3, 6, 9, and 12) to determine whether any of the bands seen with genomic DNA are actually primer artefacts. **Figure 1b.** Amplification of prokaryotic DNA. Lanes 1-3, *Escherichia coli* (strains 037, 641, 642, respectively). Lane 4, *Listeria monocytogenes* (strain 681). Lane 5, *Staphylococcus aureus* (strain 684). Lane 6, *Salmonella typhimurium* (strain 706). All genomic DNA samples were amplified with primer 5'-TCACGATGCA.

in the absence of any nucleotide sequence information for the species tested. The nucleotide sequence of each primer was chosen within the constraints that the primer was 9 or 10 nucleotides in length, between 50 and 80% G+C in composition, and contained no palindromic sequences (4). Several DNA segments were amplified in each sample. While most of these segments were common to both individuals of a given species, some segments were amplified from one individual but not the other. For example, a 1.4-kb DNA segment was amplified by primer AP9 from one human sample but not from the other (Figure 1A, lanes 1 and 2, respectively). At least one such polymorphism was apparent in each of the species examined. To confirm that the observed bands were amplified genomic DNA, and not primer artefacts (13), genomic DNA was omitted from control reactions for each primer. No amplification products were seen for any primer except for primer AP12h (Fig 1, lane 9); this artefact is not significant, however, since it is not produced when genomic DNA is included in the reaction mixture (Fig 1, lanes 7-8). These results show that single primers of arbitrary sequence can be used to amplify genomic DNA segments, and that polymorphisms can be detected between the amplification products of different individuals. Several samples of bacterial DNA were also assayed to determine whether these short primers could be used to amplify DNA segments from small genomes. The results shown in Figure 1B indicate that genomes as small as *E. coli* (4×10^3 kbp) will support amplification, and that bacteria can be distinguished according to the banding patterns of their DNA on an agarose gel.

A set of eleven related oligonucleotide 10-mers was synthesized to determine the contribution of each nucleotide to the specificity of the amplification reaction. Each primer differed from the oligonucleotide 10-mer, 5'-TGGTCACTGA, by substitution of a single nucleotide at a successive position in the sequence. The G+C content of all primers was maintained at 50%, and each primer was used to amplify DNA from two different species of soybean, *Glycine max* and *Glycine soja*. Following amplification, DNA samples were analyzed by agarose gel electrophoresis (Figure 2). Several DNA segments were amplified in each sample, and polymorphisms were apparent for many of the primers (e.g., a 0.65-kb band is present in Figure 2 lane 21, but absent in lane 22). In this experiment, most of the nucleotide substitutions in the primer caused a complete change in the pattern of amplified DNA as compared to the original primer, and in

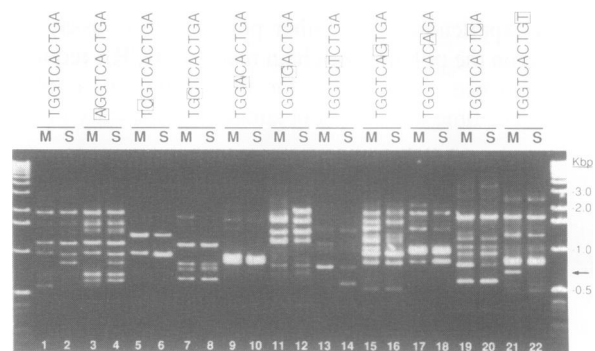


Figure 2. The effect of nucleotide substitutions in the primer on amplification. Genomic DNA from soybean *G. max* ('M') and *G. soja* ('S') was amplified with the indicated primers. Squares indicate nucleotide substitutions relative to the primer 5'-TGGTCACTGA (lanes 1-2). The arrow points to a polymorphic 0.65-kb band (see Results).

many cases revealed new polymorphisms. For example, the pattern obtained with the primer 5'-TGGTCACTGA (Figure 1, lanes 1 and 2) differs from that carrying the G-to-C substitution 5'-TGCTCACTGA (Figure 2, lanes 7 and 8). Less dramatic differences in patterns are seen for substitutions at the 5'-most position (Figure 2, lanes 1-4). These results show that an oligonucleotide 10-mer can act as a primer in the DNA amplification reaction, that polymorphisms can be detected among the amplification products, and that nucleotide changes in the primer (and by inference, the template) determine whether a given DNA segment will be amplified.

Genetic segregation

To learn whether these amplification polymorphisms are useful as genetic markers, and to assess whether the assay is reproducible, 11 polymorphisms generated with various primers were mapped in soybean, using 66 segregating F2 individuals. Each polymorphism was scored as a dominant marker and correlated with the segregation data for 430 soybean RFLP markers, derived from the same 66 individuals (manuscript in preparation). Segregation of the AP11a.1 polymorphism is shown for 16 of these F2 individuals in Fig 3A. Analysis of the data (11) indicates that AP11a.1 maps to linkage group 5 at the position shown in Figure 4. The probability that the marker is at the position indicated is $10^{16.8}$ times greater than the probability that AP11a.1 is unlinked, indicating the certainty of the map assignment. This also shows that the assay is robust, permitting reliable scoring of a polymorphism in a segregating population. The map positions and probabilities for 10 other markers are indicated in Figure 4. Primers AP4c and AP10b each revealed two different and unlinked polymorphisms (markers AP4c.1 and AP4c.2, and AP10b.1 and AP10b.2, respectively), demonstrating that single primers can be used to amplify DNA from dispersed polymorphic loci. The RAPD markers mapped (Figure 4) increased the saturation of the soybean map by filling in some

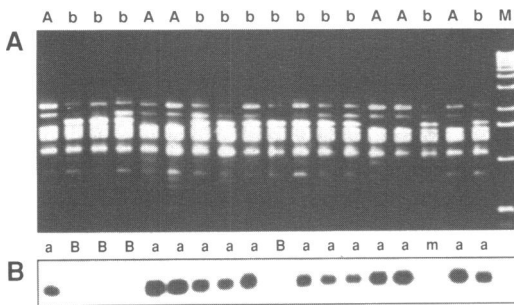


Figure 3. panel A. Segregation of an amplification polymorphism in soybean. Genomic DNA samples from 66 F2 progeny were amplified with primer AP11a (5'-ACCTCGAGCACTGCT). Amplification products from the parents *G. max*, and *G. soja* (lanes 1 and 2, respectively), and sixteen F2 individuals are shown. Arrow 'a' points to a segregating polymorphic band that is clear to score, while arrow 'b' points to a band that appears to be polymorphic in the parents, but cannot be confidently scored among the progeny (see Discussion). Panel B. Segregation of a RFLP detected by a polymorphic amplified DNA probe. The AP11a.1 polymorphism amplified in *G. soja* (Figure 3a, lane 2 band 'a') was used as a hybridization probe to detect a *Bcl II* RFLP. Panel B shows the hybridization of this probe to a Southern blot of *Bcl II* digested genomic DNA from the same individuals shown in panel A. Segregation scores for the amplification polymorphism and the RFLP are shown above panel A and B, respectively. Scores are interpreted in the following way: 'A' = genotype of parent A (lane 1), 'B' = genotype of parent B (lane 2), 'a' = either A or a heterozygote, 'b' = either B or a heterozygote, 'm' = missing data. 'M' identifies the molecular weight markers.

gaps (for example markers AP4 and AP12h1 in LG27), and by extending the map in the telomeric direction (markers AP5a2 in LG21b, AP4c3 in LG 3a, see Figure 4).

Many different random sequence primers were used to assess the quality and frequency of polymorphisms in corn, soybean, and *N. crassa* (data not shown). This was accomplished by determining what percentage of primers could be used to detect polymorphisms that could be mapped with confidence (ie. LOD scores in support of linkage greater than 4.0). The frequencies of polymorphism detection was 1 per primer for corn (number of tested primers, $p = 34$), 0.5 per primer for soybean ($p = 45$), and 2.5 per primer for *N. crassa* ($p = 88$).

Comparison to RFLPs

RFLP analysis was used to confirm the map positions of RAPD markers. Several polymorphic amplified DNA segments that had been previously mapped were excised from an agarose gel, labeled with ^{32}P , and used as hybridization probes to detect RFLPs. Two RFLPs were found using the amplified DNA segments AP3.1 and AP11a.1. Co-segregation of the RFLP

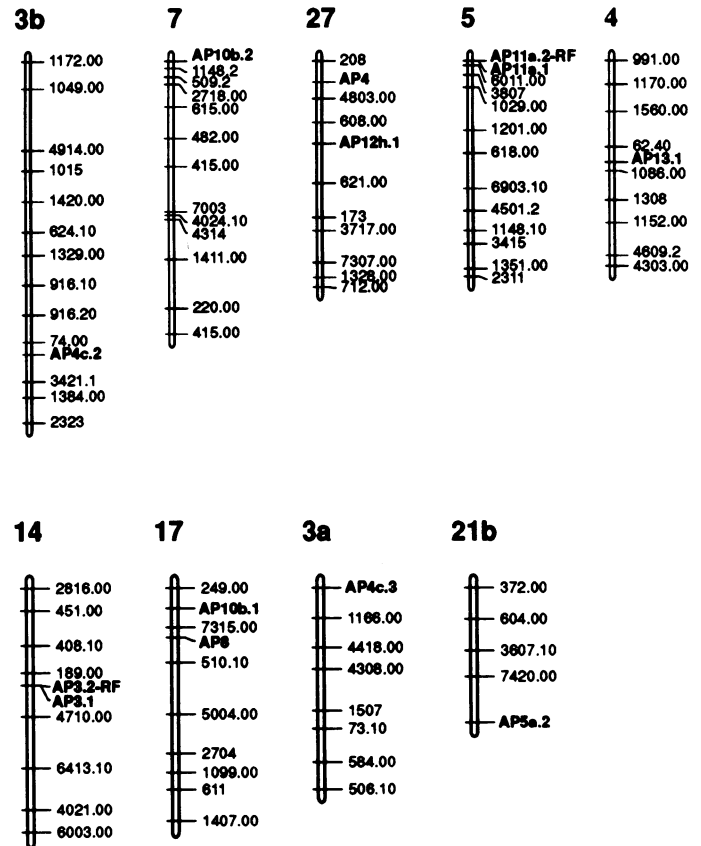


Figure 4. Genetic map of polymorphic loci detected by amplification. Amplification polymorphisms (bold type) and RFLPs identified using probes derived from amplified polymorphic DNA segments (bold type, suffix '-RF') were mapped in the soybean genome on the indicated linkage groups relative to classical RFLP markers (plain type) as described in Results. For each marker, the probability of the indicated map position is as follows: AP3.1 is $10^{10.1}$; AP3.2-RF, $10^{12.0}$; AP4, $10^{6.8}$; AP4c.2, $10^{11.1}$; AP4c.3, $10^{6.3}$; AP5a.2, $10^{3.7}$; AP8, $10^{15.4}$; AP10b.1, $10^{20.4}$; AP10b.2, $10^{5.4}$; AP11a.1, $10^{16.8}$; AP11a.2-RF, $10^{16.8}$; AP12h, $10^{4.3}$; AP13, $10^{12.8}$. Primer sequences are AP3 (5'-TCGTAGCCAA), AP4 (T-CACGATGCA), AP4c (TCTCGATGCA), AP5a (CTGTGTGCTAC), AP8 (TGGTCACTGA), AP10b (GGAAGTAGTG), AP11a (ACCTCGAGCACTGTCT), AP12h (CGGCCCTGT), AP13 (ATTGCGTCCA).

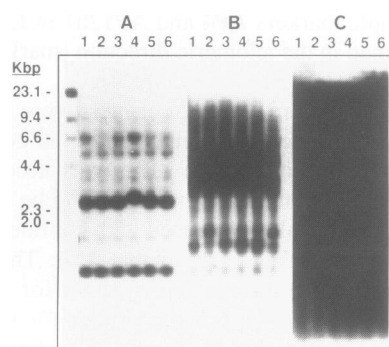


Figure 5. Complexity of genomic DNA sequences amplified with arbitrary primers. Polymorphic amplified DNA fragments were used as hybridization probes on Southern blots of *Eco* RI digested genomic DNA from the soybeans Bonus (lane 1), PI81762 (lane 2), PI416937 (lane 3), N85-2176 (lane 4), PI153293 (lane 5), and PI230970 (lane 6). Panel A shows hybridization of a polymorphic locus, amplified by primer AP8j (TGGTCACTGT), to 'single copy' DNA. Panel B shows hybridization of a polymorphic locus, amplified by primer AP4 (TCA-CGATGCA), to 'middle repetitive' DNA. Panel C shows hybridization of a polymorphic locus, amplified by primer AP13 (ATTGCGTCCA), to 'highly repetitive' DNA.

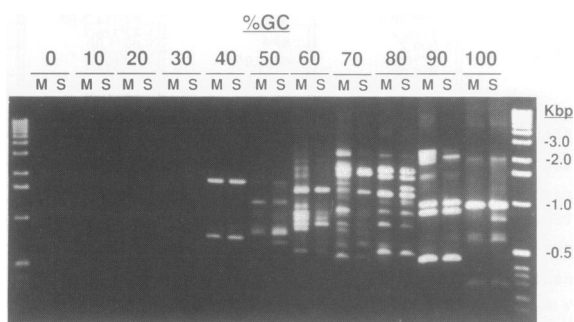


Figure 6. Effect of primer G+C content on amplification. Primers of different G+C content were used to amplify genomic DNA from the soybeans *G. max* ('M') and *G. soja* ('S'). The G+C content and nucleotide sequence of each primer is: 0% (5'-TAATTATTAT), 10 (TAATTATTGT), 20 (TAATTACTGT), 30 (TAATCACTGT), 40 (TAGTCACTGT), 50 (TGGTCACTGT), 60 (CGGTCACTGT), 70 (CGGCCACTGT), 80 (CGGCCCTGT), 90 (CGGCCCGGT), 100 (CGGCCCGGC).

AP11a.2-RF and its cognate amplification polymorphism AP11a.1 is apparent from examination of 16 F₂ individuals, as shown in Figures 3A and 3B. Segregation data for all 66 F₂ individuals confirmed that the RFLP loci AP11a.2-RF and AP3.1-RF map to the same positions as their cognate amplified probes (Figure 4). In performing this experiment, it was noted that several of the amplified polymorphic DNA segments were not suitable as RFLP probes because of hybridization to repetitive DNA. Of 11 amplified probes tested, 6 hybridized to single copy DNA (Figure 5A), 3 hybridized to middle-repetitive DNA (Figure 5B), and 2 hybridized to highly-repetitive DNA in the soybean genome (Figure 5C). This indicates that amplification polymorphisms can provide DNA markers in genomic regions which are not accessible to RFLP analysis due to the presence of repetitive DNA sequences.

Primer base composition and length

To determine the constraints on base composition of the primer, a set of related oligonucleotide 10-mers was synthesized that

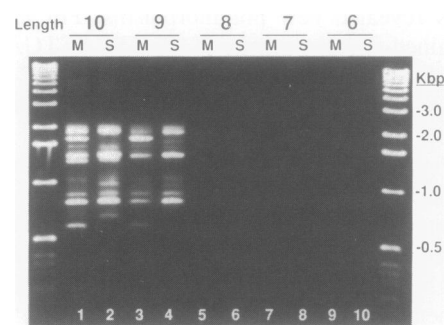


Figure 7. Effect of primer length on amplification. Primers of decreasing length were used to amplify genomic DNA from the soybeans *G. max* ('M') and *G. soja* ('S'). The length and nucleotide sequence of each primer is: 10 (ATTGCGTCCA), 9 (TTGCGTCCA), 8 (TGCGTCCA), 7 (GCGTCCA), 6 (CGCCCA).

ranged from 0% to 100% in G+C content. Each primer was related to the primers of the next-lower and next-higher G+C content by single base changes. Each primer was used to amplify soybean DNA. The results indicated that G+C content in an oligonucleotide 10-mer should be 40% or greater to generate detectable levels of amplification products (Figure 6). This experiment also supported the conclusion that single base changes can cause a complete change in the set of amplified DNA segments. To determine the minimum primer length, a set of primers from 10 to 6 bases in length was designed by deleting successive nucleotides from the 5'-end of the oligonucleotide 10-mer 5'-ATTGCGTCCA. All primers contained 5 G+C's. The results suggested that the minimum useful primer length is an oligonucleotide of nine bases (Figure 7). While these restrictions on primer base composition and length were determined at an annealing temperature of 36° in the thermal cycle profile, the same restrictions were found to apply at an annealing temperature of 15° (data not shown).

DISCUSSION

This study demonstrates that short primers of arbitrary nucleotide sequence may be used to reproducibly amplify segments of genomic DNA from a wide variety of species. Polymorphisms among the amplification products are detected frequently, are useful as genetic markers, and can be detected through examination of an ethidium bromide-stained agarose gel. Since the process described here uses primers of arbitrary nucleotide sequence to access random segments of genomic DNA to reveal polymorphisms, we propose to call these markers RAPD markers (pronounced 'rapid', for Random Amplified Polymorphic DNA) to distinguish them from ASPs (Amplified Sequence Polymorphisms) described previously for DNA segments of known sequence (5).

We believe that the RAPD assay may in some instances detect single base changes in genomic DNA. Most single nucleotide changes in a primer sequence caused a complete change in the pattern of amplified DNA segments (Figures 1 and 6). Effectively, by making a change in the primer, these experiments introduced single mismatches in the primer-genomic DNA duplex at both of the sites that define a DNA segment, and detectable amplification was prevented. By inference, a single base change in the genome may also prevent amplification by introducing a mismatch at just one end of a DNA segment. It should be

emphasized that these results do not imply that all amplifications are the result of perfect pairing between the primer and the DNA template. The number of DNA segments amplified from bacterial samples with much smaller genomes (see Figure 1b) can only be explained on the basis of mismatch between the primer and the DNA template.

Other sources of polymorphisms may include deletions of a priming site, insertions that render priming sites too distant to support amplification, or insertions that change the size of a DNA segment without preventing its amplification. As with any other genetic marker, some polymorphisms are clear and easy to score (e.g., AP11a.1, arrow A in Figure 3A), others appear ambiguous and are not useful as genetic markers (see arrow B in Figure 3A). Ambiguous polymorphisms may result from poor discrimination by a primer between alternative priming sites of slightly different nucleotide sequences.

Nearly all RAPD markers are dominant, as DNA segments of the same length are amplified from one individual but not from another. It is not possible to distinguish whether a DNA segment is amplified from a locus that is heterozygous (1 copy) or homozygous (2 copies) with a dominant RAPD marker. Co-dominant RAPD markers, observed as different-sized DNA segments amplified from the same locus, are detected only rarely. For example, we have recently constructed a genetic map of *N. crassa* comprising 88 RAPD markers; 84 of the markers were dominant and only 4 were co-dominant (Kubelik *et al.*, in preparation). Dominant markers are acceptable for genetic mapping using inbred homozygous parents. (e.g., Figure 3). In many cases where it is important to distinguish heterozygotes from homozygotes, tightly linked RAPD markers, each diagnostic for a different parental genotype, could be used in pairs to assess the genotype of the chromosomal region. The confidence with which heterozygotes could be identified would depend on how tightly linked the paired markers are. In general, the use of paired dominant markers to detect heterozygotes will require twice as many markers as would be needed using co-dominant markers. Polymorphisms within the amplified DNA fragments could also be discovered through restriction enzyme digestion or nucleotide sequence analysis to extend the utility of previously mapped or nonpolymorphic loci.

RAPD markers are well suited for genetic mapping, for plant and animal breeding applications, and for DNA fingerprinting, with particular utility for studies of population genetics. RAPD markers can also provide an efficient assay for polymorphisms, which should allow rapid identification and isolation of chromosome-specific DNA fragments. Hybrid cell lines or genetic stocks carrying deletions or additions of large chromosomal segments could be screened relative to appropriate controls, to identify the region of the genome carrying the deletion or addition. Like most molecular markers, the information content of an individual RAPD marker is very low. It is only when many of these anonymous markers are used to define a genome that they begin to have utility. High density genetic maps comprised of molecular markers have led to the identification of several previously unidentified loci of biological importance (2). RAPD markers hold promise for the automation of the genome mapping, extending the power of genetic analysis to organisms which lack an ample number of phenotypic markers to completely describe their genome. Genetic mapping using RAPD markers has several advantages over other methods: (i) a universal set of primers can be used for genomic analysis in a wide variety of species (ii) no preliminary work, such as isolation of cloned DNA probes,

preparation of filters for hybridizations, or nucleotide sequencing, is required (iii) each RAPD marker is the equivalent of a Sequence Tagged Site (14), which can greatly simplify information transfer in collaborative research programs. Perhaps the most significant advantage to this method is that the determination of genotype can be automated. Genetic maps consisting of RAPD markers can be obtained more efficiently, and with greater marker density, than by RFLP or targeted PCR-based methods.

J. Welsh and M. McClelland, in the companion paper in this volume, have independently shown, using similar methodology, that primers of arbitrary sequence can be used for genomic fingerprinting.

ACKNOWLEDGEMENTS

We would like to thank K. Butler and J. Lynch for technical assistance, and G. Trainor, Andrew Paterson, Barbara Mazur and Howard Hershey for helpful discussions.

REFERENCES

- Landegren, U., Kaiser, R., Caskey, C.T. and Hood, L. (1988)*Science* **242**, 229–237 .
- Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989)*Biotechnology* **7**, 257–264.
- Botstein, D., White, R.L., Skolnick, M.H. and Davis, R.W. (1980)*Am. J. Hum. Genet.* **32**, 314–331.
- Beckmann, J.S. (1988)*Biotechnology* **6**, 1061–1064 .
- Skolnick, M.H. and Wallace, R.B. (1988)*Genomics* **2**, 273–279.
- Murray, M.G. and Thompson, W.F. (1980)*Nucleic Acids Res.*, **8**, 4321–4325.
- Dellaporta, S.L., Wood, J. and Hicks, J.B., in *Molecular Biology of Plants* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1985), pp. 36–37.
- Metzenberg, R.L., Stevens, J.N., Selker, E.U. and Morzycka-Wroblewska, E. (1984)*Neurospora Newsl.* **31**, 35–39 .
- Burr, B., Burr, F.A., Thompson, K.H., Albertson, M.C. and Stuber, C.W. (1988)*Genetics* **118**, 519–526 .
- Metzenber, R.L. and Grotelueschen, J., in *Genetic Maps* S. J. O'Brien, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), pp. 3.22–3.29.
- Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E. and Newburg, L. (1987)*Genomics* **1**, 174–181 .
- Maniatis, T., Fritsch, E.F. and Sambrook, J., *Molecular Cloning. A Laboratory Manual*. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982).
- Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J., *PCR Protocols* (Academic Press, San Diego, 1990).
- Olson, M., Hood, L., Cantor, C. and Botstein, D. (1989)*Science* **245**, 1434–1435 .