lacking complementary strands and thus free to hybridize with target genomic DNA.

In conclusion, these results further open the door to the creation of numerous minisatellite probes to survey for VNTR loci. Such probes can be used not only for restriction fragment length polymorphism (RFLP) analyses of the many biological applications outlined in the introduction, but they can also be employed to screen cloned libraries with the ultimate goal of determining the sequences flanking VNTR loci. The latter information provides perhaps the most accurate and rapid means of utilizing the variation at such loci via the PCR analysis of alternate alleles.

## Acknowledgments

I thank Rebecca German, William Hamilton, Roger Ruff, and Earl Whitson. This research was funded by National Science Foundation Grant BSR 90-96317.

# [21] Use of Randomly Amplified Polymorphic DNA Markers in Comparative Genome Studies

By BRUNELLA MARTIRE BOWDITCH, DARRILYN G. ALBRIGHT, JOHN G. K. WILLIAMS, and MICHAEL J. BRAUN

## Introduction

Detection of genetic variation is essential to a wide range of comparative genetic research endeavors. These include studies as diverse as gene mapping, individual identification, parentage determination, population genetics, and molecular phylogenetics. The speed and accuracy of comparative genetic research often depend on the methodology for detecting variation.

The advent of allozyme electrophoresis and restriction fragment length polymorphism (RFLP) analysis of DNA made it possible to detect many polymorphisms in most organisms at the protein or DNA level.[1] These genetic polymorphisms are crucial in examining the many different fields noted above. Specifically, polymorphisms can aid in the determination of relatedness of groups of taxa, in the analysis of parentage in domestic and wild animal species, in the identification of individuals for captive breeding programs in endangered species, in the comparison of wild and cultivated plants, and in the estimation of levels of inbreeding or outbreeding in

[1] C. Chang and E. M. Meyerowitz, *Curr. Opin. Genet. Dev.* **1**, 112 (1991).



Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.

populations. They have also been essential, for example, both in linkage mapping quantitative trait loci in cultivated plants and in mapping the human genome.[2,3]

Although modern techniques for studying genetic polymorphisms have provided may new vistas of research, they do have some limitations. For instance, RFLP analysis requires fairly large amounts of genomic DNA (often unavailable from rare, ancient, and/or field-collected materials) and cloned probes that may be specific to an organism or group. Targeted PCR (polymerase chain reaction) assays, although requiring much less genomic DNA, depend on DNA sequence knowledge of the organism or gene under study. Williams *et al.*[4] have described a technique for detecting DNA polymorphisms which they call RAPD (random amplified polymorphic DNA). The technique requires only small amounts of DNA, and no prior knowledge of the genome in question is necessary. It is based on random amplification of DNA fragments, via PCR, using short primers of arbitrary sequence. The method detects abundant polymorphism in most organisms. We have used this method in our own systematic studies of genetic variation in natural populations of plants and animals. It clearly has advantages in certain situations, and promises to become a basic tool for genetic studies of closely related organisms.[5]

## Outline of Method

As stated above, the RAPD technique makes use of PCR technology, allowing geometric amplification of DNA templates. With standard PCR, it is first necessary to determine the sequence of the DNA to be analyzed. Then two specific primers (commonly 18-25 bases long), complementary to sequences flanking the target segment, are synthesized to prime the DNA amplification reaction. In contrast, the RAPD method uses a single short primer (e.g., 9 or 10 bases) of arbitrary sequence and a lower annealing temperature than the average PCR reaction (Fig. 1). These two modifications lower the specificity of the reaction so that a number of anonymous but reproducible fragments can be amplified from most complex genomes. Reaction products are analyzed by electrophoresis on agarose gels and ethidium bromide staining, so radioactive isotopes are unnecessary. Mutations that inhibit primer binding or otherwise interfere with amplification can be detected as the absence of the pertinent band in those individuals.

[2] D. Nelson, *Curr. Opin. Genet. Dev.* **1,** 62 (1991).

[3] S. J. O'Brien, *Curr. Opin. Genet Dev.* **1,** 105 (1991).

[4] J. G. K. Williams, A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey, *Nucleic Acids Res.* **18,** 6531 (1990).

[5] H. Hadrys, M. Balick, and B. Schierwater, *Mol. Ecol.* **1,** 55 (1992).

FIG. 1. Schematic showing the use of two primers of common length 18–25 bases in standard PCR versus the use of a single primer of common length 9–11 bases in RAPD reactions.

Because the short oligonucleotides are able to prime reproducible amplification from a variety of genomes under the specified conditions, it is possible to develop a panel of "universal" primers that can be used to detect polymorphism directly in practically any organism using only nanogram quantities of genomic DNA.

## Detailed Protocol for Random Amplified Polymorphic DNA

The final concentrations of all components contained in the RAPD reaction are listed below along with the composition of stock solutions. This is followed by a detailed procedure for mixing a RAPD reaction of 25 μl, which is our usual reaction volume.

*Reaction Components*

11 m*M* Tris-HCl (pH 8.3)
50 m*M* KCl
1.9 m*M* $MgCl_2$
0.001% (w/v) Gelatin or 0.1 mg/ml bovine serum albumin (BSA)
0.1 m*M* each of dATP, dCTP, dGTP, and TTP
0.2 μ*M* primer (equals 5 pmol in a 25-μl reaction)
0.02–0.04 units/μl *Taq* DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT)
0.1–4.0 ng/μl of genomic DNA

*Stock Solutions*

10× RAPD buffer: For 1 ml of stock combine:
100 μl 1 *M* Tris-HCl, pH 8.3 (100 m*M* final concentration)
500 μl 1 *M* KCl (500 m*M*)
19 μl 1 *M* $MgCl_2$ (19 m*M*)
100 μl 10 mg/ml BSA (1 mg/ml)
281 μl Distilled water

10× dNTPs: For 1 ml of stock combine:
10 μl 1 *M* Tris, pH 8.3 (10 m*M* final concentration)

10 μl 0.1 *M* dATP (1 m*M*)
10 μl 0.1 *M* dGTP (1 m*M*)
10 μl 0.1 *M* dCTP (1 m*M*)
10 μl 0.1 *M* TTP (1 m*M*)
950 μl Distilled water

*Procedure*

1. Prepare stock solutions of reaction buffer (10× RAPD buffer) and deoxynucleotides (10× dNTPs). Larger quantities of both solutions can be prepared, subdivided into aliquots, and stored frozen at −20°.

2. Dilute primer and template DNAs in TLE (10 m*M* Tris-HCl, 0.1 m*M* EDTA, pH 8.0) such that 1 μl of the primer dilution contains 5–6 pmol of primer and 1 μl of the template dilution contains the desired amount of DNA for the reaction (between 5 and 15 ng of template DNA per 25-μl reaction has worked best in our hands).

3. Prepare a mix of those reagents common to all tubes in the experiment. Prepare a quantity sufficient for the total number of samples and controls to be run, plus one extra aliquot to ensure adequate volume for pipetting. In a single tube combine the following amounts **per sample:** 2.5 μl of 10× dNTPs, 2.5 μl of 10× RAPD buffer, and 0.5–1 unit *Taq* polymerase (volume will depend on supplier's concentration). Add distilled water to final volume of 23 μl per sample and *mix well.* Making a reaction mix greatly reduces the chances of pipetting errors for individual samples.

4. Label the reaction tubes (600-μl Eppendorf tubes) and add 23 μl of reaction mix (Step 3), 1 μl diluted template DNA (5–100 ng), and 1 μl diluted primer DNA (5–6 pmol). The advantage of adding the reaction mix first is that the surface tension will help pull down the 1 μl volumes, which would otherwise be hard to pipette into the tubes. If using only one template DNA with a number of primers or one primer DNA with a number of templates, the single primer or template DNA can be combined with the mixture in Step 3 instead of adding it separately in Step 4.

5. Mix the 25-μl reactions by flicking the capped tubes with a finger, then spin the tubes briefly in a microcentrifuge to remove any air bubbles.

6. Overlay the reactions with mineral oil to avoid evaporation in the thermocycler (20 μl of oil is usually sufficient) and place the tubes in the thermocycler. Turn on the machine 15 min before starting the reactions to warm it up, as suggested by the instruction manual.

7. Record the location of each of the reaction tubes in the thermocycler block. This is helpful in troubleshooting poor reactions to rule out effects of any temperature inconsistencies within the block itself.

8. Program the thermocycler for 40–45 cycles with the following parameters: denature for 1 min at 94°, anneal for 1 min at 35°–36° (55° for 16-mers), and extend for 2 min at 72°. The total reaction time is approximately 5 hr in a Perkin-Elmer Cetus thermocycler using minimum ramping times between incubation temperatures.

9. Place completed reactions at 4° until gel analysis is performed.

The ranges listed for the amount of template DNA, the amount of primer DNA, the amount of *Taq* polymerase, the number of PCR cycles, and the annealing temperature used do not seem to significantly affect the results of the RAPD reaction. Likewise, substituting BSA for gelatin in the RAPD buffer does not affect the results.

## Detailed Protocol for Gel Analysis of Randomly Amplified Polymorphic DNA Reactions

1. Remove 23 μl of the RAPD reaction to a new Eppendorf tube by placing the pipette tip to the bottom of the reaction tube and drawing the reaction up from the bottom. This avoids most of the mineral oil.

2. Add 2.5 μl of 10× gel loading buffer [30% (w/v) Ficoll, 200 m*M* EDTA (pH 8.0), 0.5% (w/v) bromphenol blue, dissolved in water and stored at −20°] to each sample.

3. Load the samples on 1.4–1.5% (w/v) agarose gels. We use 50-ml minigels made with TBE electrophoresis buffer (89 m*M* Tris base, 89 m*M* boric acid, 2 m*M* EDTA, pH 8.3) and add 5 μl of 5 mg/ml ethidium bromide directly to the molten agarose before pouring the gel. Normally, we use 12-well combs (1.5 mm thick) in the minigels, which hold about 20 μl of sample volume total, but 16-well combs can also be used. Because the ethidium bromide-stained RAPD bands appear sharper the smaller the well size used, larger well sizes are not recommended. DNA length standards in the range of 200–5000 base pairs (bp) should also be included on each gel to facilitate determination of band sizes.

4. Electrophorese the gels at 50 V in TBE buffer for approximately 3 hr. We have found that longer runs at lower voltages yield clearer images of band patterns than shorter runs at a higher voltage. It is crucial to obtain the best separation of bands possible since the gel results are documented only by a photograph.

5. Photograph the gel on an ultraviolet transilluminator using an orange UV filter for the camera lens. For good gels, it is useful to use film which gives both a positive and a negative of the gel so that additional prints of that gel can be made at a later time.

## Considerations for Optimizing Randomly Amplified Polymorphic DNA Reactions

### *Primers*

Although the primers used in this technique are random sequences, there are several criteria that must be met to produce results. Williams *et al.*[4] found that the minimum primer length to detect amplification in ethidium bromide-stained agarose gels was 9 bases, and a minimum of 40% GC content was also required to produce amplification (50–80% GC content is generally used). These criteria were determined at an annealing temperature of 36° but also held true at annealing temperatures as low as 15°.

Primer dimer formation can be a potential problem in PCR amplifications. Although we generally avoided the use of self-complementary sequences when designing primers, we found that dimer formation was not a problem at an annealing temperature of 35° even when 6 bases of a 10-bp primer could potentially hybridize.

Primers can be purchased or synthesized using standard phosphoramidite chemistry. Operon Technologies Inc. (Alameda, CA) has available 25 kits of 20 different 10-base primers.

### *Stringency*

The stringency limits for RAPD reactions are fairly straightforward. Annealing temperatures above 40° prevented amplification by many of the 10-base primers.[4] Annealing temperatures of 35° or 36° are routinely used for 9-, 10-, and 11-base primers, while an annealing temperature of 55° is used for 16-base primers.[6,7] D. Fong (personal communication, 1991) has observed that an annealing temperature of 39° for 10-base primers tends to decrease the total number of amplification products for any given primer, but the bands that do appear are highly reproducible.

### *Number of Cycles, DNA Quality, and Concentration*

DNA samples purified by either stringent methods (equilibrium centrifugation in cesium chloride density gradients) or simple methods (organic extraction) have yielded reproducible RAPD patterns. However, in plants (e.g., members of the evening primrose family, Onagraceae) and

[6] J. G. K. Williams, M. K. Hanafey, J. A. Rafalski, and S. V. Tingey, this series, Vol. 218 p. 704 (1993).

[7] M. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **88,** 1398 (1991).

animals (e.g., molluscs) that have high concentrations of polysaccharides, alternative methods of DNA purification were necessary for successful amplification (Bult *et al.*,[8] L. Adamkewicz, personal communication, 1992). General conditions of the PCR protocol are also important to consider. For RAPD reactions, the number of cycles appears to be more critical than the starting concentration of the template DNA. We have used between 0.0002 and 4.0 ng/$\mu$l of template DNA from a single individual titmouse *(Parus bicolor)* in a 45-cycle RAPD reaction. Equivalent band patterns, as viewed on an ethidium bromide-stained agarose gel, were observed for template DNA concentrations between 0.004 and 0.8 ng/$\mu$l. Template DNA concentrations both higher and lower than this range, however, produced only a subset of the original band pattern (Fig. 2A). These data are supported by the findings of Williams *et al.*,[6] that RAPD band patterns were generally preserved until the template DNA concentration fell below 0.1 ng in a 25-$\mu$l reaction (0.004 ng/$\mu$l). However, when testing 0.6 ng/$\mu$l of the same template DNA at 45, 40, 35, 30, 25, and 20 cycles, some bands in the pattern were poorly reproducible at 30 and 25 cycles, and all bands were absent in 20-cycle runs (Fig. 2B).

### *Taq Polymerase Concentration*

We have found that 0.02 units/$\mu$l of *Taq* polymerase is often sufficient for RAPD reactions using 0.2–0.6 ng/$\mu$l of template DNA with 0.24 pmol/$\mu$l of primer DNA, though the optimal quantity of *Taq* polymerase may depend on the supplier used. Excessive amounts of *Taq* polymerase, or in some cases template DNA, can cause the reaction to appear as a smear of amplified DNA rather than as discrete bands.[4]

### *Ultraviolet Light Range for Photography*

Higher contrast photographs of ethidium bromide-stained agarose gels are obtained if the gel is illuminated with short-wavelength UV radiation (254 nm), whereas longer wavelengths (312 nm) yield less satisfactory results. This enhanced contrast may allow some polymorphisms to be scored under short-wavelength UV light that could not reliably be scored under longer wavelengths. Note that wavelengths shorter than 254 nm should not be used to view a gel band that is to be excised for cloning or direct sequencing since such wavelengths cause breaks in the DNA backbone much more frequently than longer wavelengths do.

[8] C. Bult, M. Kallersjo, and Y. Suh, *Plant Mol. Biol. Rep.* **10**(3), 269 (1992).

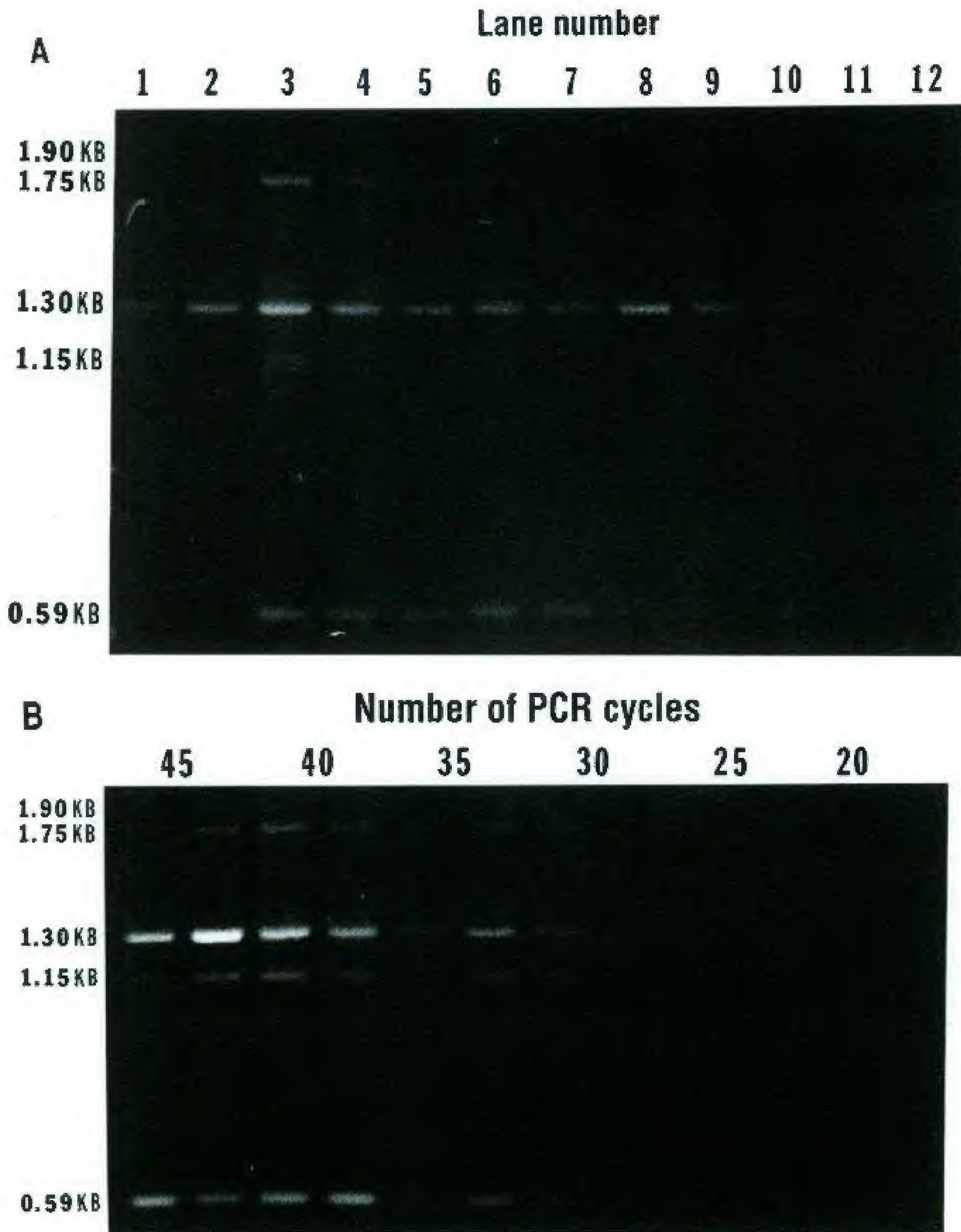


FIG. 2. Effect of DNA concentration and number of PCR cycles on RAPD analysis, shown on ethidium bromide-stained agarose gels, using template DNA from a single individual titmouse *(Parus bicolor)* and the 10-base primer AP5a[4] (5′ CTGTTGCTAC 3′). **(A)** Various concentrations of template DNA amplified through 45 cycles of PCR. Lanes 1–11 contain 100, 50, 20, 10, 5, 2, 1, 0.1, 0.05, 0.01, and 0.005 ng of template DNA, respectively, in 25-$\mu$l reactions. Lane 12 contains no template DNA. **(B)** Constant amount of template DNA (0.6 ng/$\mu$l) amplified with primer AP5a between 20 and 45 cycles. All reagents for the experiment were combined in a single tube, then aliquoted into twelve 25-$\mu$l reactions. Duplicate reactions were performed for each cycle length variation.

*Use of Polyacrylamide Gels*

As compared to ethidium bromide-stained agarose gels, many more amplified bands can be visualized by separating the products of the reaction on a polyacrylamide gel and staining with silver.[9] This gel system also permits detection of bands amplified weakly by primers as short as 5 nucleotides.

## Characteristics of Randomly Amplified DNA Polymorphisms

A typical RAPD amplification will yield from 2 to 10 visible bands, of which most are reproducible, when performed by the methods described herein. For any particular genome, a certain percentage of RAPD primers will not produce satisfactory amplification products. These primer/template combinations either yield no amplification whatsoever or yield very light or hazy bands that are hard to reproduce in replicate reactions. However, the same primers may work perfectly well on the next genome tested, so the failure of a particular primer/template combination is more likely due to lack of readily amplifiable segments in the template DNA rather than a systematic problem with the primer.

A large majority of the strong reaction products fall in the range of 400–2000 bp. This is presumably due to limitations in the resolving power of the agarose gels at lower molecular weights as well as inefficiency of the extension reaction under the described PCR conditions at higher molecular weights. Although optimization of the PCR protocol for higher molecular weight markers is possible, it is probably neither time nor work efficient considering the number of scorable bands obtained in the lower molecular weight range.

Most RAPD markers show dominant/recessive inheritance in diploid organisms, that is, a DNA segment of a certain molecular weight is amplified from some individuals but not from others. In general, it is not possible to distinguish heterozygous individuals from those homozygous for the dominant allele at such loci; both have the "band present" phenotype. Presumably, insertion/deletion events of moderate size (perhaps 50–1500 bp) will result in codominant alleles detectable as amplification products of different size, assuming the mutation has no major effect on the kinetics of amplification. Experience to date, however, indicates that such codominant size polymorphisms are relatively rarely detected in RAPD assays. Williams *et al.*[4] found only four codominant markers among 88 RAPD polymorphisms mapped in the *Neurospora crassa* ge-

[9] G. Caetano-Annoles, B. J. Bassam, and P. M. Gresshoff, *Bio/Technology* 9, 553 (1991).

nome. The frequency may be greater in genomes where short interspersed repetitive elements are common.

When a small number of individuals is analyzed, the statistical inference of allelism may be weak for a particular pair of bands showing a pattern of inheritance consistent with codominance. In this case, homology between the bands can be established by excising them from the gel for sequencing or for use as probes on Southern blots.

## Nature of Randomly Amplified DNA Polymorphisms

Although the RAPD technique can potentially detect all classes of mutations (substitutions, insertions, deletions, inversions), the basis of most polymorphism is presumably substitution. However, substitutions apparently need not occur in the primer binding sites themselves to be detected by the RAPD assay. Braun and Albright[10] describe evidence indicating that the sequence of the primer binding sites is frequently unchanged in RAPD polymorphisms. Primers containing restriction sites were used to detect RAPD polymorphisms on the assumption that such polymorphisms would also be frequently detected as RFLPs when the RAPD bands were used as probes in Southern transfers. Although the predicted restriction fragments were observed, these fragments were rarely polymorphic. The most likely explanation appears to be that the RAPD assay can detect substitutions outside the primer binding sites themselves.

Priming of RAPD assays is carried out at relatively low temperatures that may allow the formation of secondary structure within the single stranded template at or around the primer binding sites. Such secondary structure could inhibit the binding of primer to template sufficiently to prevent amplification. Alternatively, secondary structure might inhibit elongation by *Taq* polymerase. Mutations which stabilize secondary structure may thus be detectable as RAPD polymorphisms even though they do not reside within the primer binding site. Under this hypothesis, the RAPD assay potentially can detect mutations in a much larger region than that covered by the primer binding sites. This hypothesis may help explain the high degree of polymorphism observed in RAPD band patterns for most organisms tested.

## Reproducibility and Experimental Strategies

Because the RAPD assay is often used to identify individuals or populations, reproducibility of individual RAPD reactions must be high. Once

[10] M. J. Braun and D. G. Albright, in preparation (1993).

ideal conditions for a certain reaction are found, it is imperative to keep them constant in order to be able to compare results from different reactions. The same is true for the thermocycler; once a study is begun on a certain instrument it is best to run all samples on the same instrument.

With careful attention to laboratory technique, the reactions are quite reproducible from run to run within the same thermocycler and within different wells of a thermocycler on the same run. The strongest PCR products are the most highly reproducible, with fainter products tending to come and go on multiple runs. For this reason, we suggest that strong products be chosen over faint products when scoring individuals or populations.

An example of band scoring using a particular RAPD primer in marsh wrens *(Cistothorus palustris)* is shown in Fig. 3. Although all bands in the schematic are visible on the original gel photograph, the accuracy with which they might be scored as present or absent is quite variable. Bands d, f, h, and k (Fig. 3) are dominant PCR products that can be easily scored as to their presence or absence. Band e, present in individuals 1 and 4 in Fig. 3, is very close in size to band f but is most likely distinct since individual 5 appears to have both band e and band f. Bands a and c in Fig. 3 vary so much in intensity between individuals that they would be hard to score reliably across a population. The remaining bands (b, g, i, and j, Fig. 3) are quite faint so their reproducibility would be suspect. Thus, only bands d, e, f, h, and k (Fig. 3) are good candidates for scoring within this population.

Because the number of polymorphisms detected is primarily limited by the number of primers tested, it is often possible to identify more polymorphisms than are needed for a particular study. Therefore, the best

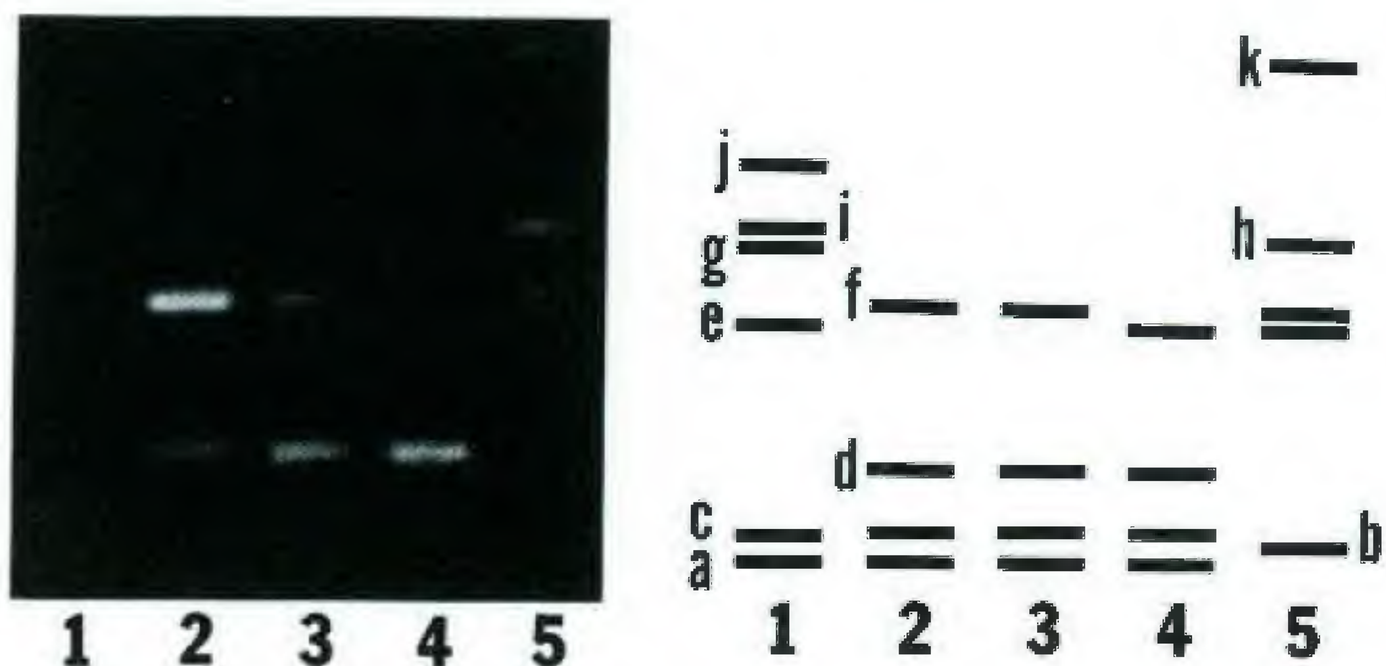


FIG. 3. RAPD analysis on an ethidium bromide-stained agarose gel using 0.6 ng/$\mu$l template DNA from several individual marsh wrens *(Cistothorus palustris)* and the 10-base primer AP4c[4] (5′ TCTCGATGCA 3′). Both the gel photograph and a schematic of visible bands are shown.

strategy to be used with RAPD technology is to screen many primers and select only those that give highly reproducible bands for scoring, rather than trying to optimize every primer/template combination.

By the same rationale, the presence of a band in the pattern of an individual can be more reliably scored than absence of a band. For this reason, we have on several occasions repeated the RAPD assay on all individuals with the band-absent (negative) phenotype for particular polymorphisms, and we recommend this precaution, especially for those investigators becoming familiar with the technique. In our experience, repeating the assay has rarely resulted in a changed score when the positive phenotype consists of a strong PCR product, but replication of experiments serves as a gauge of the accuracy of results and is a useful aid to choosing polymorphisms that are easily reproducible.

A significant source of false-negative scores stems from DNA samples containing impurities that inhibit PCR amplification. Such samples are readily detected by noting the relative intensities of PCR products in an RAPD pattern other than the particular polymorphism being scored (Fig. 4, lane 9). If bands that are monomorphic throughout a population are weak or absent in an individual sample, it is likely that this sample contains some inhibitory contaminant. Further purification of the sample by any of a number of standard protocols normally corrects this problem.

Finally, band patterns seem fairly reproducible between individual Perkin-Elmer Cetus thermocycler units (especially the strong products), but we have had trouble using the precise protocol described here with other brands of thermocyclers. This is most likely due to the fact that different brands of thermocyclers have quite different temperature cycling profiles even when programmed identically. The largest variations occur in ramping times between the various cycle temperatures and in the way the machine determines those temperatures (i.e., temperature probe in the thermocycler block versus temperature probe in a sample itself). The RAPD technique can likely be adapted to any brand of thermocycler by optimizing the programmed reaction times for the machine used.

## Use of Randomly Amplified Polymorphic DNA Technology in Comparative Genome Studies

With RAPD technology, polymorphisms can be detected in closely related organisms such as those that compose a species complex, different populations of a single species, or individuals within a population. For this reason, the RAPD technique is likely to find wide application in gene mapping, in individual and strain identification, and in those issues in ecology and population biology requiring genetic analysis of relatedness or

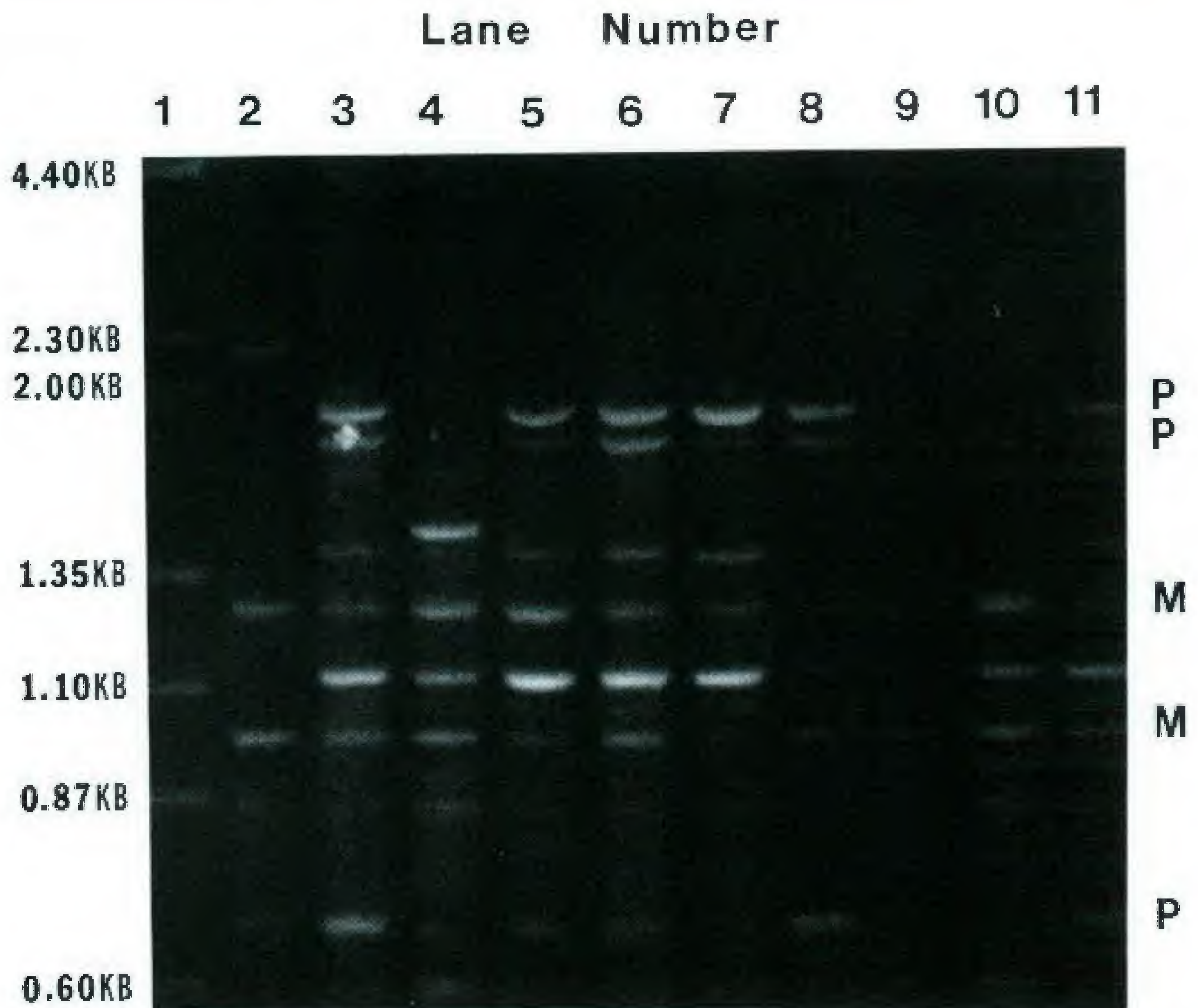


FIG. 4. Typical RAPD analysis on an ethidium bromide-stained agarose gel. **Lane 1** contains DNA length standards. **Lanes 2–11** contain 0.6 ng/$\mu$l each of genomic DNA from different individual marsh wrens *(Cistothorus palustris)* amplified with the 10-base primer AP5a[4] (5′ CTGTTGCTAC 3′). Note that most amplified products lie between 0.6 and 2.3 kb, and there are about 8 to 10 well-defined PCR products per lane. Some products are monomorphic in this population (M), whereas others are polymorphic (P). Amplification of all bands in lane 9 is weak, and testing of this individual should be repeated before polymorphic fragments are scored as present or absent.

identity. We believe that RAPD technology will be very useful, for example, in studying breeding/mating systems for both plant and animal species, such as determining the number of individuals that have originated from outcrossing versus self-crossing in plants. We have successfully applied the RAPD method to several questions involving geographical "species" or populations in both avian and plant systems.

*Avian Systems*

Two different species of birds were analyzed using the RAPD technique. Five individuals each from an eastern and western population of marsh wrens *(Cistothorus palustris)* were compared, and five individuals

each of two subspecies of titmice *(Parus bicolor)* were also compared. In both species, 25 of 31 primers tested were able to amplify genomic DNA. In several cases, primers that failed to produce scorable results in one avian species were successful in the other species, supporting the idea that lack of amplification from a particular primer/genome combination is more often due to chance than nonsuitability of the primer. For the wrens, 17 of the 25 usable primers showed some individual polymorphisms. The 25 primers produced 128 scorable bands, 50 of which were polymorphic. Thus, 55% of all primers tested detected polymorphisms, for an average of 1.6 polymorphic bands per primer tested. Similarly, in titmice, 18 of the 25 usable primers showed some individual polymorphisms. A total of 116 bands was produced, 53 of which were polymorphic. Thus, 58% of primers tested detected an average of 1.7 polymorphic bands per primer tested.

### *Plant Systems*

Ninety primers were screened on a single individual from the weedy *Euphorbia esula* L. complex. Of those tested, 60 primers amplified the DNA. Subsequently, the 60 primers were tested on an individual from each of six morphologically indistinguishable populations within the complex. Forty of the 60 primers used produced fragment patterns which differentiated at least four of the individuals, and 23 of the 40 usable primers gave a unique pattern for every geographical population screened. Of the 231 scorable bands present in the fragment patterns produced by the primers giving population-specific patterns, 82 bands were unique to one of the six populations, and 8 bands were shared by all six populations.

RAPD analysis was also performed on DNA extracted from the monocots *Heliconia latispatha* Bentham and *Phenakospermum guianense* L. C. Rich. Endl. et Miq. Of a total of 35 primers tried, 28 amplified DNA from *Heliconia latispatha,* and 27 amplified DNA from *Phenakospermum guianense.* RAPD technology has been used to examine interspecific gene flow between two species of Louisiana irises, *Iris fulva* Ker-Gawl. and *Iris hexagona* Walt.[7] Six of the seven primers tested detected differences between the two species. Three of the primers were then used to confirm the hybridity of *Iris nelsonii* and its origin from *Iris fulva* and *Iris hexagona.* Finally, Williams *et al.*[4] used RAPD analysis to construct genetic maps in soybeans (*Glycine max* L. Merr. and *Glycine soja* Sieb. et Zucc.), corn (*Zea mays* L. Off.), and the mold *Neurospora crassa.*

### *Other Systems*

Currently, RAPD analysis is being used to study a wide variety of organisms. Successful amplifications have been obtained from mosquito

DNA[11], isopods (D. Fong, personal communication, 1991), molluscs (L. Adamkewicz, personal communication, 1992), humans, and bacteria.[4] It seems likely that the technique can be adapted to virtually any organism.

Welsh and McClelland[12] have used an approach similar to the RAPD technique that they call arbitrarily primed PCR (AP–PCR). In this technique, longer primers (which may have been made for some other purpose) are used as random primers by beginning with two cycles of low stringency PCR followed by further amplification at higher stringencies. The results are repeatable, and have permitted them to distinguish genetically 24 strains within 5 species of *Staphylococcus,* 11 strains of *Streptococcus pyogenes,* and 3 varieties of rice *(Oryza sativa).*

Although RAPDs promise to be useful in parentage analysis and other DNA typing studies, the occasional appearance of nonparental bands in offspring of known parentage has raised concern in some cases.[13] Such anomalous bands may be amplification artifacts caused by heteroduplex formation between alternative parental alleles[5] or between repetitive elements from distinct loci in the genome. Repetition of the analysis may resolve some anomalies. In a study of domestic horses, 5 cases of nonparental bands were observed among 59 sire–dam–offspring triplets analyzed. In all 5 cases, on repetition, either the anomalous band disappeared in the offspring or appeared in one of the parents (E. Bailey, personal communication, 1992). It may be possible to account for most or all nonparental bands by the analysis of "synthetic offspring," namely, an equal mixture of DNAs from putative parents.[5]

## Applications of Randomly Amplified Polymorphic DNA Technology in Systematics

The use of RAPD technology in phylogenetics is limited to extremely closely related organisms. In this instance, the systematic characters are defined as the RAPD fragments of a certain molecular weight, while the character states are the presence or absence of that band. It must be noted that the absence of a band can potentially be the phenotype of many different alleles at a RAPD locus, while the presence of a band demonstrates an amplifiable sequence of a specific length. In other words, the character state "band absent" may in reality encompass many different character states such as inversion, different secondary structure, or any number of point mutations. This implies that the likelihood of band loss through mutation will generally be greater than the likelihood of regaining

[11] R. C. Wilkerson, T. J. Parsons, D. G. Albright, T. A. Klein, and M. J. Braun, *Insect Mol. Biol.* **1** (4), in press (1993).

[12] J. Welsh and M. McClelland, *Nucleic Acids Res.* **18,** 7213 (1990).

[13] M. F. Riedy, W. J. Hamilton, and C. F. Aquadro, *Nucleic Acids Res.* **20,** 918 (1992).

the same band: the transition probabilities between character states are asymmetrical. This situation is analogous to that which arises when restriction site data are used for phylogenetic inference[14–16] and may require similar precautions in analysis.

It is very important to realize that in coding characters we are making a statement of homology. The great advantage of RAPD technology is the ability to obtain DNA polymorphisms without having to sequence or otherwise characterize the genomic DNA of interest. On the other hand, because the amplifications are random, we cannot be sure that the comigrating gel bands we see are homologous in every sample analyzed. The inference of homology is strong when total sequence divergence between taxa is low and many RAPD bands are shared. At higher taxonomic levels, it is likely that only a few shared bands would be generated and their homology would be highly questionable. To avoid gross errors, it is important to limit this type of study to closely related organisms and not infer systematic relationships at higher taxonomic ranks.

[14] V. A. Albert, B. D. Mishler, and M. W. Chase, *in* "Molecular Systematics of Plants" (P. S. Soltis, D. E. Soltis, and J. J. Doyle, eds.), p. 369. Chapman and Hall, NY, 1992.
[15] D. M. Hillis, M. W. Allard, and M. M. Miyamoto, this volume [34].
[16] W. Wheeler, *Syst. Biol.*, submitted for publication.

# [22] Molecular Approaches to Mammalian Retrotransposon Isolation

*By* SANDRA L. MARTIN and HOLLY A. WICHMAN

## Introduction

Interspersed repetitive DNA, which accounts for a significant proportion of the DNA found in mammalian genomes, appears to arise largely from the replicative dispersal of transposable elements. Because of the high copy number of these elements and their mobility within the genome, it is reasonable to propose that their sequences play an important role in the evolution of mammalian genomes and thus of mammals themselves, yet the extent of that role remains to be elucidated. From an evolutionary point of view, it is interesting to ask why mammalian genomes tolerate so much (apparently) excess DNA. One view is that transposable elements persist despite neutral or even deleterious effects on their host because they have evolved selfish mechanisms to maintain themselves as genomic parasites.[1–3] An alternative view is that, although they may evolve as selfish

[1] W. F. Doolittle and C. Sapienza, *Nature (London)* **284,** 601 (1980).
[2] L. E. Orgel and F. H. C. Crick, *Nature (London)* **284,** 604 (1980).
[3] D. Hickey, *Genetics* **101,** 519 (1982).


Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.