MINI-REVIEW

Applications of random amplified polymorphic DNA (RAPD) in molecular ecology

H. HADRYS, M. BALICK* and B. SCHIERWATER†
Yale University, Department of Biology, 165 Prospect Street, New Haven, CT 06511 and †New York Botanical Garden, Bronx, NY 10458, USA

Abstract

Molecular genetic markers have been developed into powerful tools to analyse genetic relationships and genetic diversity. As an extension to the variety of existing techniques using polymorphic DNA markers, the Random Amplified Polymorphic DNA (RAPD) technique may be used in molecular ecology to determine taxonomic identity, assess kinship relationships, analyse mixed genome samples, and create specific probes. Main advantages of the RAPD technology include (i) suitability for work on anonymous genomes, (ii) applicability to problems where only limited quantities of DNA are available, (iii) efficiency and low expense.

Keywords: DNA fingerprinting, DNA probes, kinship analysis, paternity determination, RAPD, taxonomic identifications

Received 3 February 1992

Introduction

Within the past few years molecular genetic approaches have become of increasing importance to studies in behavioural ecology and population biology. For instance, DNA fingerprinting technologies have revolutionized approaches to our understanding of animal social systems by permitting analyses of kinship relationships (e.g. Burke & Bruford 1987; Burke et al. 1991a; Jones, Lessels & Krebs 1991; Gyllensten, Jakobsson & Temrin 1991; Packer et al. 1991; Pemberton, Bancroft & Amos 1991; Schlötterer, Amos & Tautz 1991; Smith et al. 1991).

Despite constant progress in methodology, application of DNA markers to many problems in behavioural and population ecology has been limited by technical considerations (Lewin 1989; Kirby 1990; Burke et al. 1991a; Pemberton et al. 1991). The most frequently used DNA markers include RFLPs (restriction fragment length polymorphisms) visualized by Southern blot hybridization to different types of single-locus or multilocus probes (Burke et al. 1991b) and PCR amplified simple sequence microsatellite loci (Tautz 1989). Potential applications are frequently thwarted by the requirement for significant quantities of DNA in the case of RFLP analysis or by lack of relevant DNA sequence information in the case of conventional PCR-based techniques. Recent criticisms are that DNA fingerprinting requires special molecular training, is labour-intensive, and is relatively expensive (Weatherhead & Montgomery 1991).

It has been argued that DNA fingerprinting is so essential to behavioural ecology and population biology that it would be highly unfortunate to have its application limited to a few specialized laboratories rather than to the broad community working in these fields (Weatherhead & Montgomery 1991). Major progress in technology development is expected in two directions: (1) increase in analytical power per unit effort, and (2) simplification in technology, and ultimately reduction in expense. Use of random amplified polymorphic DNA (RAPD) markers, detected by PCR amplification of small inverted repeats scattered throughout the genome, adds a new technology of DNA fingerprinting to the molecular analysis of relatedness between genotypes. The introduction of RAPD fingerprinting is a substantial contribution toward the second direction.
RAPD technology is only some 2 years old, and consequently published applications are limited. However, the rapidly growing interest in using RAPD technology justifies an early review on the current literature and the potentials of the method. In this paper we shall review the principle and original applications of the RAPD technology, discuss its applications in molecular ecology, and point out the particular advantages as well as limitations of RAPD markers.

**Principle of RAPD analyses**

The PCR-based RAPD technique (Williams et al. 1990) is an attractive complement to conventional DNA fingerprinting in ecology. RAPD analysis is conceptually simple. Nanogram amounts of total genomic DNA are subjected to PCR using short synthetic oligonucleotides of random sequence. The amplification protocol differs from the standard PCR conditions (Erlich 1989) in that only a single random oligonucleotide primer is employed and no prior knowledge of the genome subjected to analysis is required. When the primer is short (e.g. 10-mer), there is a high probability that the genome contains several priming sites close to one another that are in an inverted orientation. The technique essentially scans a genome for these small inverted repeats and amplifies intervening DNA segments of variable length. The profile of amplification products depends on the template–primer combination and is reproducible for any given combination (see below). The amplification products are resolved on agarose gels and polymorphisms serve as dominant genetic markers, which are inherited in a Mendelian fashion (Williams et al. 1990; Carlson et al. 1991; Martin, Williams & Tankesley 1991; Welsh, Peterson & McClelland 1991). Amplification of non-nuclear RAPD markers is negligible because of the relatively small non-nuclear genome sizes.

Two modifications of detecting RAPD markers have been described as DNA Amplification Fingerprinting (DAF) and Arbitrarily Primed Polymerase Chain Reaction (AP-PCR). DAF uses short random primers of 5–8 bp and visualizes the relatively greater number of amplification products by polyacrylamide gel electrophoresis and silver staining (Caetano-Anolles, Bassam & Gresshoff 1991). AP-PCR uses slightly longer primers (such as universal M13) and amplification products are radioactively labelled and also resolved by polyacrylamide gel electrophoresis (Welsh & McClelland 1990; Welsh et al. 1991b).

Standard RAPD analysis is performed according to the original methods (Williams et al. 1990) using short oligonucleotide primers of random sequence which are commercially available (Operon Technologies, Inc., Alameda, Calif.). Only high-molecular-weight, i.e. non-degraded, DNA should be subjected to RAPD analyses. Amplification products can be resolved by gel electrophoresis on 1.4% agarose gels.

**Applications**

Here we illustrate several potential applications of RAPD fingerprinting in molecular ecology, including determination of taxonomic identities, detection of interspecific gene flow, assessment of kinship relationships, analysis of mixed genome samples, and production of specific probes.

**Determination of taxonomic identity**

By employing different oligonucleotide primers, molecular characters can be generated that are diagnostic at different taxonomic levels. For any given primer, RAPD amplification products can be broadly classified into two groups: variable (polymorphic) or constant (non-polymorphic). These definitions are relative for a given operational taxonomic unit (OTU). For instance, consider a RAPD analysis of several individuals within a species, and several species within a given genus (Fig. 1). Constant fragments diagnostic for a genus may be identified, as well as fragments which are polymorphic between species within the genus. Both types of product can be exploited for establishing systematic relationships. In this example, constant fragments operationally identify members of a certain genus exclusively if the fragment is a unique polymorphism in a comparison of genera (genus-specific character in Fig. 1). Note, the determination of taxonomic relatedness is only valid between taxa for which the diagnostic RAPD fingerprint patterns have been established. Similarly, fragments polymorphic at the species level will operationally identify members of a given species if the fragment is constant among all members of that species (species-specific character in Fig. 1). An example of such a marker is provided in RAPD fingerprints from two dragonfly species in Fig. 2. Fragments polymorphic among individuals may also be utilized to determine clonal identity, as is frequently required for asexually reproducing organisms. Clone-specific markers have been identified in hydroids (Schierwater, unpubl.), clonal “individual”-specific markers in fungal mycelia (Smith, Bruhn & Anderson 1992), cultivar-specific markers in broccoli and cauliflower (Hu & Quiros 1991), strain-specific markers in mice (Welsh et al. 1991a), species-specific markers in irises (Arnold, Buckner & Robinson 1991) and tomato (Klein-Lankhorst et al. 1991) and genus- and family-specific markers in palms (M. Balick & S. Dellaporta, in preparation). Thus RAPD products can be generated that
serve as diagnostic molecular characters at different taxonomic levels.

As illustrated in Fig. 1, the specificity of any single RAPD marker may range from the level of the individual to higher taxonomic levels. However, taxon identification by diagnostic RAPD markers can only be done by comparison within a set of genotypes of known RAPD amplification profile for a given primer. The specificity of RAPD markers has to be determined empirically for each genome within a set of genomes under investigation by screening several primer-template combinations (analogous to the screening of restriction enzyme–probe combinations in Southern-blot-based fingerprinting). Homology of diagnostic markers, especially between members of higher OTUs, may be established by Southern analysis to exclude the possibility of co-migration of fragments of the same size from non-homologous loci (see below).

**Analyses of interspecific gene flow and hybrid speciation**

A straightforward conclusion of the outlined potential of the RAPD method to identify diagnostic markers for different OTUs is that RAPD can be applied to analyse fusion of genotypes at different taxonomic levels. At the level of the individual, RAPD markers may be applied to parentage analysis (see below); at the population or species level RAPD may be used to detect hybrid populations or species. The detection of genotype hybrids relies upon the identification of diagnostic RAPD markers for the parental genotypes under investigation.
Arnold et al. (1991) have demonstrated interspecific gene flow between two Louisiana iris species, *Iris fulva* and *I. hexagona*, by analyses of species-diagnostic RAPD markers. Using two 10-bp and one 16-bp primers they reported four species-specific markers from different populations of *I. fulva*. These markers were missing in *I. hexagona*, but were present at intermediate frequencies in experimental F1 hybrids (*I. fulva x I. hexagona*) and at variable frequencies in a natural contemporary hybrid population. Conclusions from RAPD analyses on the occurrence of interspecific gene flow between two iris populations were consistent with results from RFLP analyses of the chloroplast genome. Furthermore, variable frequencies of species-diagnostic markers were also found in the putative hybrid species *I. nelsonii*. Here, RAPD markers may be useful in investigating the role of hybridization in the origin of *I. nelsonii*.

Using the AP-PCR modification of RAPD, Welsh et al. (1991a) identified F1 hybrids from different inbred maize lines. Other groups have begun to use RAPD for analyses of hybridization events where allozyme studies have not proven to be sensitive enough for hybrid genotypes, for example, in hybridization along vertical zonations in natural populations of *Daphnia* (B. Streit, personal communication) or in plant breeding programmes (Crowhurst et al. 1991; Hu & Quiros 1991; Martin et al. 1991; Quiros et al. 1991).

**Determination of paternity and kinship relationships**

By employing fragments that are polymorphic among individuals, RAPD analysis may be used to assess paternity and kinship relationships in large offspring samples. A common problem in behavioural ecology is to determine the actual father from a number of potential fathers. The earliest application of RAPD fingerprinting to paternity analyses resolved the question of paternity in an unknown mating system of the dragonfly *Anax parthenope* (Hadrys 1991; H. Hadrys et al., in preparation). *A. parthenope* males guard ovipositing females over extended periods of time and thereby regularly take the chance of suffering severe injury or even death by attacks from conspecific males trying to split the tandem pairs. The presence of spermatodesmids, i.e. sperm bundles encased in a slowly degrading matrix, has suggested that the male may guard a female in order to assure a subsequent mating success rather than immediate fertilization success. RAPD analyses of tandem males, tandem females and the offspring clutches identified the tandem male as the father of the immediately oviposited offspring clutches. Figure 3 shows RAPD fingerprints of the guarding male, the guarded female, the offspring, and several unrelated males. The RAPD markers polymorphic among potential fathers show the guarding male to be the actual father. For statistical analyses the number
of polymorphic markers can be increased by increasing the number of diagnostic primers, and conventional band-sharing coefficients can be applied (Lynch 1990, 1991; Burke et al. 1991a; Keane et al. 1991; H. Hadrys et al., in preparation). Band-sharing coefficients between unrelated individuals are highly dependent on the primer-template combination used (see Fig. 2). Although we found background band-sharing coefficients of RAPD markers generally higher than those known from multilocus probe fingerprinting, this does not represent a serious problem. Because linkage between different arbitrary priming sequences is extremely unlikely, the number of independent polymorphic markers analysed can be rapidly increased by pooling markers revealed by several primers. The amplification of monomorphic RAPD markers may be kept to a minimum by choosing the right primer-template combination, and any monomorphic markers may be removed from the analysis to decrease background band sharing.

In principle RAPD markers can formally be treated as Mendelian alleles, and for parentage analysis analytical approaches may be developed which are based on knowledge of allelic frequencies, e.g. as used in statistical analyses of single-locus fingerprint profiles. In practice, however, allelic frequencies of scorable dominant RAPD markers in diploid organisms might be difficult to estimate and markers revealed by the same primer may be linked (cf. Williams et al. 1990; Carlson et al. 1991).

The segregation and linkage of RAPD markers has been demonstrated in several genetic studies. The expected nuclear transmission of RAPD markers to F1 offspring has been reported for hybrids of maize inbred lines (Welsh et al. 1991a) as also has the segregation of RAPD polymorphisms in experimental crosses of conifers (Carlson et al. 1991) and the linkage of RAPD markers to resistance genes in lettuce (Michelmore, Paran & Kesseli 1991; Paran, Kesseli & Michelmore). RAPD markers have also been used in demonstrating the introgression of two parental genomes in an iris hybrid species (Arnold et al. 1991).

Conventional RFLP fingerprinting techniques are ill-suited for the analysis of paternity and estimation of reproductive success in species with large offspring clutches, because of the need to determine paternity for each individual offspring. RAPD fingerprinting provides a ready alternative for such cases. Synthetic offspring may be produced by mixing equal amounts of the DNA of the mother and the potential father (Fig. 3). The amplification products from the synthetic offspring should ideally contain the full complement of bands (H. Hadrys et al., in preparation; cf. Carlson et al. 1991; see also below) that appear in any single offspring of these parents. However, certain combinations of alleles may lead to amplification artefacts (e.g. heteroduplex formation, see below), and certain markers may only get amplified from an offspring but not from either of its parents. The occurrence of non-parental bands in offspring from known primate pedigrees has raised concerns in parentage determinations when single individuals are analysed (Riedy, Hamilton & Aquadro 1992). The occurrence of non-parental bands in offspring from known primate pedigrees has raised concerns in parentage determinations when single individuals are analysed (Riedy, Hamilton & Aquadro 1992). In contrast, the synthetic offspring is a complete representation of both parental genomes and will match the profile of a large sample of offspring, since in both the 'synthetic' and the real offspring clutch, allele combinations that may cause amplification artefacts are represented in equal amounts. The degree of mismatch between synthetic offspring and actual offspring clutches is indicative of mixed paternity, which can be measured by quantitative determination of mixed genome samples.

Analysing mixed genome samples

The RAPD technique may be used to generate quantitative estimates of the relative proportions of different genomes in mixed DNA samples. In many polygamous mating systems, especially in insects, sperm competition and mixed paternity may occur. Here, confirmation of the presence of more than one paternal genotype would be highly desirable. Using DNA from a dragonfly, Orthetrum coerulescens, we reconstituted mixed-genome samples experimentally by varying the relative proportion of DNA from two male individuals over two orders of magnitude. This DNA was amplified and the relative amounts of individual male-diagnostic amplification products quantified by densitometry. The relative intensity of diagnostic bands from two different individuals varied predictably with the relative DNA concentrations of each genome in the reaction, and the DNA concentration of a given genome could be estimated from band intensities as long as the relative amount of this genome was at least 20% (Fig. 4; see also Michelmore et al. 1991). Note that this approach is based on well amplified polymorphic bands, requires precise knowledge of the diagnostic markers for each genome being scanned, and requires prior calibration experiments.

Generating novel specific probes

Any diagnostic RAPD marker can be eluted from the gel, reamplified, radiolabelled with 32P and serve as an inexhaustible supply of probe in Southern analyses (Fig. 5). Such probes may be used to exclude the possibility of
co-migration of fragments of different sequence but similar size. Other applications include generation of probes for taxonomic analysis or the quantitative estimation of the presence of a certain genome in a mixed sample by Southern analysis (cf. Michelmore et al. 1991; Jeffreys et al. 1991). The specificity of the probes can be further improved by eluting diagnostic RAPD bands from the gel, reamplifying, subcloning and sequencing the fragments, and eventually selecting a partial consensus sequence as a probe. Williams et al. (1990) report that six of 11 RAPD markers tested as probes in RFLP analyses were useful hybridization probes, because all of them hybridized to single-copy DNA. The other five, however, were not useful, because they hybridized to middle- or highly repetitive DNA in the soybean genome. RAPD probes have also been used to detect RFLPs in tomato species (Martin et al. 1991; Klein-Lankhorst et al. 1991).

**Difficulties and limitations of RAPD fingerprinting**

The following technical considerations and potential difficulties merit attention:

**The size of the primer**

Primer size will determine the degree of specificity in genome scanning. It may be expected that primers of short length will amplify an unreasonably large number of sequences and that larger primers will amplify too few sequences to be routinely informative. Beyond a certain primer size (c.15-mer) increasing primer length may also increase non-specific primer annealing, consequently increasing the probability of random non-reproducible amplification patterns. All studies using standard RAPD
conditions (fragment separation on agarose gels) have found 10-bp primers to be an efficacious size. A G+C content of the primer similar to the G+C content of the analysed genome will maximize the frequency of binding sites and hence amplification products.

Sensitivity to reaction conditions

Being PCR-based, the principal limitations of RAPD fingerprinting arise from its sensitivity to reaction conditions, and slight changes in the conditions may affect the reproducibility of amplification products (Williams et al. 1990; Arnold et al. 1991; Carlson et al. 1991; Klein-Lankhorst et al. 1991). The technique is sensitive to (a) shape of the temperature profile, (b) type of polymerase used and (c) Mg²⁺ concentration. The amplification profile is sensitive to Taq or DNA concentration. The shape of the temperature profile is a property of the thermal cycler and must be standardized. Only strictly standardized reaction conditions will guarantee reproducible amplification products. Furthermore, we found that the optimal concentration of template DNA per reaction may vary substantially from typical conditions (25 ng per reaction) depending on the primer–template combination used (cf. Carlson et al. 1991).

The possibility of co-migration

An assumption of the use of the RAPD technique is that amplified fragments are unique, i.e. that the procedure does not amplify two distinct fragments which co-migrate on gels because of similar size. Co-migration in the RAPD technique is easily detected by eluting individual PCR products from gels and reprobing the products via Southern analysis (Fig. 5). Alternatively, polyacrylamide gel electrophoresis may be used to increase the resolution of band separation.

Non-reproducible amplification products

As with other genetic markers, some RAPD fragments may be ambiguous and not easy to score (Williams et al. 1990). These unclear and non-reproducible fragments, which may derive from non-specific priming or from heteroduplex formation between related amplification products (or other secondary structure artefacts, which can prevent normal amplification patterns) are not useful as genetic markers. However our own work, as well as the work of several others (e.g. Williams et al. 1990; Arnold et al. 1991; Hu & Quiros 1991; Klein-Lankhorst et al. 1991) have all shown that if the RAPD amplification is repeated two or more times, the majority of markers is clearly reproducible and scoreable. As in many cases of using PCR, sometimes amplification products are found

Fig. 5 Use of amplification products as probes. (a) RAPD fingerprint gel (primer B14 5’TCCGCTCTGG) of a family of Orthetrum coerulescens (lanes 1–9). Lane 5 is the synthetic offspring of the parents of the family, and lane 7 are larvae from the last oviposited egg clutch that lack a diagnostic marker (arrow). (b) A sample (5 μl) of the diagnostic marker was eluted from the gel and reamplified in a 100-μl reaction volume containing 50 μCi of ³²PdATP under standard RAPD PCR conditions (Williams et al. 1990). The radioactive amplification product was used as a hybridization probe to a Southern blot of the original RAPD fingerprint. The autoradiogram shows the diagnostic marker to be homologous in different lanes and also identifies a smaller RAPD fragment (arrow) containing homologous sequence to the probe. Note that the probe hybridizes substantially more strongly to the synthetic offspring (lane 5), which contains the pooled RAPD markers from both parents. This result was reproducible with synthetic offspring from different parents; its basis is not yet understood and it suggests that care should be taken in using probes with synthetic offspring samples. Also note, the probe used here did not hybridize to any RAPD markers (revealed with the same primer) from the unrelated dragonfly Anax parthenope.
even in the absence of template DNA in the reaction (Innis et al. 1990; Klein-Lankhorst et al. 1991). However, in all reported cases those 'ghost' bands disappear if the template DNA under investigation is added to the reaction.

Other considerations

The most common DNA fingerprint technologies differ substantially in (i) complexity of technological procedures, (ii) required amount of DNA, (iii) sequence information needed for a genome being scanned, (iv) analytical power of assigning genotype relatedness, (v) expense in terms of labour and money, (vi) broadness of applications. In this context, RAPD fingerprinting seems likely to have wide potential for applications in molecular ecology, and requires the least in technology, labour and expenses. The cost of producing one individual DNA fingerprint by Southern hybridization can be very substantial (Weatherhead & Montgomerye 1991); in contrast, the average expense for one RAPD fingerprint can be as low as US$2.00. On the other hand, RAPD markers are the least informative of all known DNA markers and they are dominant (heterozygosity is normally not detectable). Consequently, the analytical power of RAPD markers is not competitive with analyses using sequence information or single locus probe fingerprint technologies. However, RAPDs are detected more easily than RFLPs and can be competitive with RFLPs even in analyses of genomes with high levels of heterozygosity (Williams et al. 1990; Carlson et al. 1991; Hu & Quiros 1991).

Potential Future Applications

We note here several additional applications currently under development.

1. **Sex determination.** In many ecological (as well as agricultural and legal) applications it would be convenient to have available markers that were sex-specific. We expect that little difficulty will be encountered in developing RAPD markers with this characteristic.

2. **Generation of specific PCR primers for anonymous genomes.** A major limitation in the application of PCR to ecological problems is the absence of sequence information for the vast majority of organisms. We suggest that this difficulty may be overcome for many applications by using a RAPD-based strategy for developing 'designed' PCR primers. Specifically, RAPD primers with an embedded restriction site may be used to detect fragments showing the desired properties (e.g. detecting a particular taxon). These fragments may then be cloned and sequences used to develop specific 'designed' PCR primers for diagnostic markers.

3. **Quantitative analysis of mixed biosamples.** Analogous to the analysis of mixed paternity samples in dragonflies, analysis of field samples of different species or other OTUs (e.g. plankton sampling) may be performed.

4. **Phylogeny.** RAPD markers may prove to be useful characters for cladistic analysis.

Acknowledgments

Many substantial contributions to the manuscript derive from the ideas of Dr Stephen L. Dellaporta and Dr Leo W. Buss. We thank Matt Dick, Robert Dorritt and Neil Blackstone for critical comments. Our work is supported by DAAD, DFG, NATO, NIH, NSF and ONR.

References


This paper is a result of collaborative work between the groups of Leo Buss and Stephen Dellaporta (Yale, Biology) to develop and apply new molecular genetic tools to developmental genetics and population ecology. Heike Hadrys is studying the evolution of mating systems in odonates by combining field work with molecular genetic analyses at the Zoological Institute Braunschweig and at Yale University. Michael Ballick collaborates with Stephen Dellaporta on conservation biology and systematics of palms. Bernd Schierwater has performed 3 years of postdoctoral research with Leo Buss on molecular developmental genetics and evolutionary ecology of hydroids.