The Thai Journal of Veterinary Medicine

Volume 29
Issue 3 *September, 1999*

Article 1

9-1-1999

RANDOM AMPLIFIED POLYMORHIC DNA (RAPD) TECHNOLOGY AS USEFUL TOOL FOR BACTERIAL TYPING

Niwat Chansiripornchai

Piyart Subhachalat

Review Article

RANDOM AMPLIFIED POLYMORPHIC DNA (RAPD) TECHNOLOGY AS A USEFUL TOOL FOR BACTERIAL TYPING

Niwat Chansiripornchai* Piyarat Subhachalat**

Abstract

Niwat Chansiripornchai* Piyarat Subhachalat**

RANDOM AMPLIFIED POLYMORPHIC DNA (RAPD) TECHNOLOGY AS A USEFUL TOOL FOR BACTERIAL TYPING

Traditional and molecular genetic typing techniques have been used for discriminating the clonal relationships of bacterial strains. Traditional typing techniques based on phenotypic characteristics such as serotyping are being increasingly replaced by the use of DNA-based methods. The polymerase chain reaction (PCR) has led to typing techniques based on DNA amplification. Randomly amplified polymorphic DNA (RAPD) typing (Arbitrarily Primed- Polymerase Chain Reaction, APPCR) is one of the techniques which is being used increasingly to type micro-organisms, especially during clinical outbreaks. The theoretical aspects and the critical parameters that affect typing and results of RAPD typing are discussed in the text.

Key words : Random Amplified Polymorphic DNA (RAPD), Critical parameter affecting typing, Molecular epidemiology

* Department of Veterinary Medicine, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand.

** Department of Veterinary Pharmacology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand

บทคัดย่อ

นิวัตร จันทร์ศิริพรชัย ปิยะรัตน์ ศุภชลัสถ์

อาร์ เอ พี ดี เทคนิกที่เป็นประโยชน์ ในการจำแนกสายพันธุ์แบคทีเรีย

วิธีการจำแนกความสัมพันธ์ของสายพันธุ์แบคทีเรียที่ใช้มีทั้งวิธีการดั้งเดิมและวิธีการทางพันธุอณูชีววิทยา วิธีการดั้งเดิมอาศัยลักษณะที่แสดงออกของแบคทีเรีย เช่น ซีโรไทป์ ซึ่งวิธีการนี้กำลังถูกแทนที่มากขึ้นด้วยวิธีการที่อาศัยการจำแนกจาก ดี เอ็น เอ ปฏิกิริยาลูกโซ่โพลี่เมอเรสนำไปสู่การจัดจำแนกซึ่งอาศัยการเพิ่มจำนวนของดี เอ็น เอ วิธีอาร์ เอ พี ดี หรือ เอ พี พี ซี อาร์ เป็นหนึ่งในบรรดาวิธีการที่กำลังได้รับความนิยมเพื่อการจำแนกจุลชีพ โดยเฉพาะอย่างยิ่งขณะเกิดการระบาดของโรค ทฤษฎีและตัวแปรสำคัญที่มีผลต่อการจัดจำแนก และวิธีการแก้ปัญหาของเทคนิก อาร์ เอ พี ดี ได้ถูกอธิบายในตัวเรื่อง

คำสำคัญ : เทคนิกอาร์ เอ พี ดี ตัวแปรที่มีผลต่อเทคนิก ระบาดวิทยาระดับโมเลกุล

Introduction

The phenotypic or genotypic characteristics of pathogenic and non-pathogenic bacteria, including biochemical, cell surface, toxin/enzyme production properties and pathogenicity have all been used to develop identification and typing techniques such as staining, serotyping, biotyping, phage typing, ELISA, and culturing (drug resistance and animal experiments).

Generally these tools have been applied to:

1. Investigate outbreaks and to collect epidemiological data for research purposes and disease control.
2. Examine a series of isolates from a single patient with the aim of determining whether
   (i) the patient is a pathogenic carrier
   (ii) infection is recurring and/or
   (iii) the patient is not recuring.
3. Establish possible links between the clinical (syndrome) manifestations and the pathogen isolated from the patient, the outbreak or even the environment and to establish the mechanism of pathogenicity of the organism involved in the infection.
4. To give us more information in respect to our understanding the epidemiology of infection.

Even when, by studies of host-microbe interactions, we have learned how this infectious disease process occurs in the isolated individual, still much has yet to be unravelled as to how this infectious disease is transmitted in and between different populations, and by what means transmission can be prevented.

Veterinary epidemiology not only studies the relationships between different pathogens, their hosts, and the environment, but it also concerned about public health. It provides information and the means for understanding and controlling the spread of disease.

The rapid development of molecular genetics during the last few decades has resulted in new epidemiological techniques, such as "nucleic acid fingerprinting". These powerful new technologies have rapidly changed many traditional methods of microbiological epidemiology. As for bacteriology, these new

technologies have shifted the focus of attention from serological biochemistry and physiochemistry to their genomes.

In the past few decades various *in vitro* nucleic acid amplification procedures been developed, such as the polymerase chain reaction (PCR). Now they are widely used for diagnostic studies and, increasingly, for molecular genetic epidemiological studies as well (Erlich *et al.*,1991). A number of molecular typing methods can be used to classify bacteria based on groups of genetically related isolates that are presumed to derive from a common ancestor and that arise through evolutionary subdivision, or ecologically, as divergent epidemiologically unrelated groups of isolates, from independent sources of infection (Struelens, 1998).

These new molecular epidemiology techniques for typing bacterial pathogens can now be used to answer a number of questions. In particular, what is the extent and mode of transmission of different clones, and what is their prevalence over time and their epidemic or endemic geographical spread?

Traditional typing techniques

For a long time, traditional or phenotypic typing techniques have been used to identify and examine a small number of phenotypic characters, which are often affected by alterations in the expression of the organism's gene (Wachsmuth, 1985; Mekalanos, 1992). Traditional methods involve isolating the organism in pure culture and performing predetermined biochemical or immunologic tests to identify it. These tests are based on the organism's biochemical or structural characteristics, such as their bacterial proteins or enzymes (Dryden *et al.*, 1992; Gaston *et al.*, 1988; Mulligan *et al.*, 1988; Selander *et al.*, 1986; Tabaqchali *et al.*, 1987). In many cases, the tests are time consuming and difficult to interpret, and the techniques used are often not sensitive. Such traditional techniques, like serotyping, are therefore being increasingly replaced. The "certain criteria" (Arbiet, 1999) of all typing techniques is based on their ability to (1) type a strain, or strains, and gather reliable results which can be used to draw specific conclusions, (2) reproduce data and information that conform to what was obtained, (3) discriminate between strains, or groups of strains, such as phenotypic or genetic clusters having common properties, and (4) distinguish between epidemiologically unrelated strains.

Molecular genetic typing techniques

Today, molecular genetic fingerprinting of nucleic acid such as Restriction Enzyme Analysis (REA), standard or pulse field gel electrophoresis (PFGE) and ribotyping, where chromosomal fragments are hybridized with DNA probes based on ribosomal DNA sequences are becoming increasedly used for epidemiological studies. (Archer *et al.*, 1984; Bingen *et al.*, 1994; Branchini *et al.*, 1993; Fry *et al.*, 1991, John *et al.*, 1989; Marco *et al.*, 1993; Peterson *et al.*, 1993; Shlaes *et al.*, 1986). The polymerase chain reaction (PCR) technology has influenced development of many techniques with several areas of application, including the discrimination of bacterial strains. Current PCR typing can be divided into five main groups. Each is described below.

1. PCR-RFLP

PCR-RFLP is dependent on amplifying a defined DNA fragment. The amplification product is digested by one or more restriction enzymes to generate restriction fragment length polymorphisms (RFLPs). These polymorphisms allow discrimination of different strains (Gaydos *et al.*, 1992; Navarrow *et al.*, 1992; Rodrigo *et al.*, 1992).

2. PCR-ribotyping

This technique is a variation on standard ribotyping. DNA fingerprints are generated by amplification rather than by restriction en-

zyme digestion (Gaydos *et al.*, 1992; Navarrow *et al.*, 1992; Rodrigo *et al.*, 1992).

3. AP-PCR

The Arbitrarily-Primed PCR (AP-PCR) technique, which is also known as the Random Amplification of Polymorphic DNA (RAPD) technique, uses short primers to produce a genomic fingerprint, allowing the analysis of organisms in which genomic sequences are largely unknown (Welsh and McClelland, 1990; Williams *et al.*, 1990). One or more primers of variable length are arbitrarily selected and allowed to anneal to the DNA template at low stringency. The PCR products are then run electrophoretically to obtain DNA fingerprints, which differ according to the relationship between the strains under investigation.

4. PCR-AFLP

PCR-AFLP is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. The technique involves three steps; (i) restriction cleavage of the DNA and ligation of oligonucleotide adapters, (ii) selective amplification of sets of restriction fragments and (iii) gel analysis of the amplified fragments. PCR amplification of restriction fragments is achieved by using the adapter and restriction site sequence as target sites for primer annealing. Selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites (Vos *et al.*, 1995).

5. REP-PCR

This method is based on the amplification of Repetitive Extragenic Palindromic (REP) elements found in all bacterial genomes, and relies on primers derived from short extragenic sequences which can identify many bacteria and fungi (Woods *et al.*, 1992; van Belkum *et al.*,1992). If two sequences are located close enough to each other, within a few thousand bases, depending on the polymerase used, they can act as priming sites to amplify the area of DNA between them. This area is known as an interrepeat fragment (Arbeit, 1999). The number and the size of these repeat sequences are variable from strain to strain, and therefore lead to a number of different-sized interrepeat fragments that can be resolved electrophoretically to yield distinct DNA fingerprints (Versalovic, 1991). A method that is similar in concept is known as ERIC-PCR (enterobacterial repetitive intergenic consensus sequence PCR) and is based on a family-conserved sequence found in the Enterobacteriaceae (Hulton *et al.*, 1991; Versalovic *et al.*, 1991). All of these methods produce fingerprints at relatively high annealing temperatures (similar to those employed in standard PCR reaction), although this does not preclude the use of these primers at lower temperatures in a classic RAPD reaction.

Arbitrarily Primed Polymerase Chain Reaction (AP-PCR) or Random Amplified Polymorphic DNA (RAPD) typing

RAPD typing is based on the fact that short primers will hybridize at random sites to initiate DNA amplification (Kostman *et al.*, 1992; Kostman *et al.*, 1995; Welsh and Mc Clelland, 1990). As with REP-PCR, the proximity, number and location of these priming sites vary between strains, and the method produces DNA fingerprints that differ when electrophoresed and visualized. The attractions of the method are evident. No prior knowledge of the template DNA sequence is required, and the technique can be theoretically applied to any organism. The amplification reaction in RAPD typing is characterized by low stringency, which is normally achieved by using low annealing temperatures (Kostman *et al.*, 1992; Kostman *et al.*, 1995; Welsh and Mc Clelland, 1990).

Theoretical aspects

Caetano-Anollés *et al.* (1992b) proposed a reaction mechanism that explains how a single primer amplifies template DNA. Amplification products, synthesized in the first round of the cycle with two primers, become the preferred templates for subsequent amplification cycles. However, amplification products derived from single primers have palindromic ends. These can form hairpin loops that may interfere with primer annealing in subsequent amplification cycles. The extent of hairpin loop formation and stability will vary between each amplification product and therefore, only some of the first-round products will be amplified efficiently in subsequent cycles. These concerns are largely, if not totally, eliminated in amplification reactions using two primers. In late amplification cycles, the production of the amplified product is linear rather than exponential, because decreasing efficiency of template denaturation favours the formation of product-product duplexes rather than primer- template interactions. Because different amplification products, that act as templates, are amplified independently, those present at low concentrations will be preferentially amplified during later amplification cycles. The specificity of the amplification process is thus dependent on primer-template interactions. Changes in the concentration of template DNA or primer will affect the PCR products that are amplified and therefore, result in different DNA fingerprints.

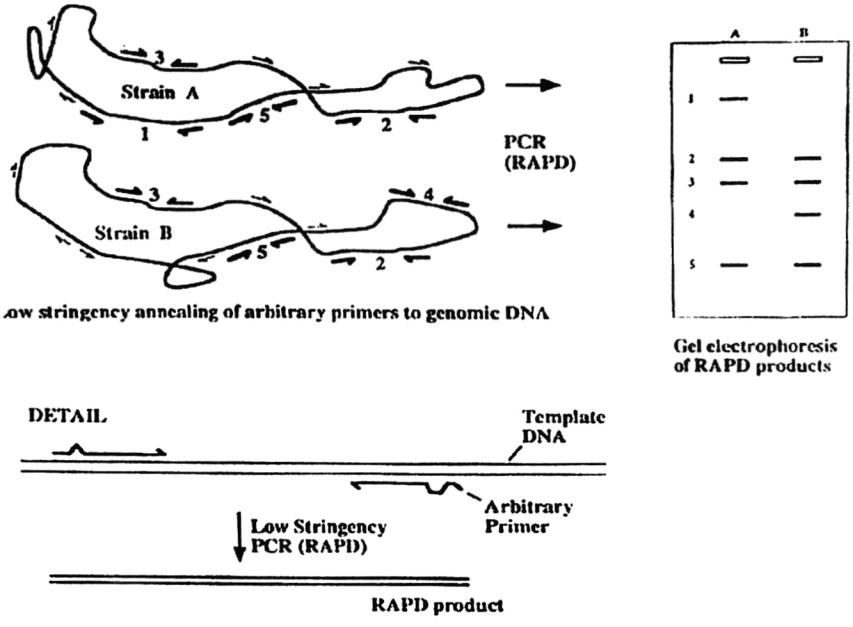

Figure 1. The Random Amplified Polymorphic DNA (RAPD) technique. At the top left, pairs of half arrows in bold, indicate primer annealing to pairs of sites that result in RAPD products; half arrows in gray, indicate primer annealing to individual sites that are not near enough to other equivalent sequences to result in RAPD products. The annealing of primers to pairs of incom pletely match sites, which is postulated to be responsible for more RAPD bands from prokaryotic genomes, is diagrammed in the "DETAIL" section (lower left). (adapted from Berg *et al.*, 1994)

Critical parameters affecting RAPD typing

In principle, each physical and chemical component of RAPD-PCR can be modified to produce a potential increase in yield, specificity, or sensitivity. This section will now present an overview of these factors, which are not independent of each other, and explain how they affect the success of RAPD-PCR.

Reaction conditions

A defining feature in RAPD typing is the relatively low annealing temperature of 25-40 °C required for adequate amplification (Bassam *et al.*, 1992; Young *et al.*, 1994; MacPherson *et al.*, 1993; Saulnier *et al.*, 1993). Nevertheless, short primers have been shown to produce products at 55 °C, although no DNA amplification has been observed at 65-80 °C (Caetano-Anoll's *et al.*, 1992a; Rychlik *et al.*, 1990). Higher temperatures, presumably, prevent primer annealing and subsequent extension. The use of 'hot start' techniques does not seem to improve product yield or alter RAPD profiles (Erlich *et al.*, 1991; Mullis, 1991). Buffer components, such as magnesium ion (Mg^{2+}) and dimethyl sulfoxide (DMSO), influence RAPD reactions in much the way they influence amplification in the standard PCR-based technique (Bassam *et al.*, 1992). The number of cycles does not seem to influence RAPD fingerprints. Power (1996) suggests that there are identical fingerprints with the same primer after 35-50 cycles, and that the relative band intensities do not vary within this range, although less than 30 cycles results in the absence of some bands. This variability is presumably a result of different products being amplified at different times during the amplification cycles, as discussed previously.

Primer considerations

One attraction of RAPD typing is that primers can be designed arbitrarily without prior knowledge of the template sequence required. Primer selection is probably the single most influential parameter in successful RAPD typing (van Belkum *et al.*, 1993). The three main considerations are primer sequence, length and concentration.

Primer sequence. At the low annealing temperatures used in RAPD typing. Fingerprints are obtained depending on the number and the frequency of annealing sites on the template. This is difficult to predict at low temperatures, so a number of primers and/or primer combinations must be screened, to obtain the most suitable for the particular species under investigation. This statement implies that screening should be done with panels of strains of known epidemiology. Nevertheless, some authors suggest that total GC content of primers is the most critical factor in estimating the efficiency of amplification, although there is no clear explanation for the poor amplification observed with some primers (Kwok *et al.*, 1990; Williams *et al.*, 1990). The GC content should be at least 60% and contain no hairpin structure.

Primer length. Optimum primer length depends on template DNA. Theoretical consideration of primer-template interactions suggests that longer primers should possess greater discriminatory ability. The size of the primer is typically 7-15 bases in length. Primers with different arbitrary sequences and different sizes will give different banding patterns for the same DNA. However, MacPherson *et al.* (1993) reported that, for three machines tested, a variation in the length of the primer (20 or 10 nucleotides) had no effect on the reproducibility of the assays.

One expects longer primers to yield more forms when mismatches occur in the centre of the primer sequence. Annealing with longer primers is therefore responsible for extra bands, with longer primers and potentially greater discriminatory ability. Although this is true in many cases, it is not always so in practice, which presumably reflects the influence of template DNA on RAPD interactions. In-

deed, it has been proposed that primers greater than eight nucleotides in length do not increase the information content significantly and therefore longer primers may be unnecessary. In practice, van Belkum *et al.* (1993), reported RAPD typing was more successful when using primers longer than 10 bases. Conversely, short 10 base primers were more discriminatory than longer primers, when typing clinical strains of *Candida albicans* (Power, 1996).

Primer concentration. Khandka (1997) found that increasing the concentration of a primer resulted in the amplification of low molecular weight DNA fragments, while lowering the concentration, resulted in fragments of high molecular weight. However, MacPherson *et al.* (1993) suggested that a DNA concentration of 1 ng gave similar banding patterns, when using the same thermal cycler. Low concentrations of primer (0.05 μM) did not produce any detectable DNA fragments. Increased primer concentrations, of 0.25 μM or higher, generated intensely stained DNA fragments. Although concentrations above 0.5 μM did bring about the synthesis of increasing amounts of very short DNA fragments, they did not improve the clarity of the banding patterns.

In this study, when a concentration of 25 picomoles of a specific primer, per 25 μl of the reaction mixture was used, good informative RAPD products were obtained. This accords with other studies using 20-50 picomoles of a specific primer per 25μl of the reaction mixture. A primer amount in the range of 20-50 picomoles gave more informative RAPD products than when using a lower (5 picomoles) or a higher amount (100 picomoles) (Berg, 1994). At low annealing temperatures, some degree of mismatch can be expected. Annealing at the 3' ends is essential for polymerization (Caetano-Anoll's, 1993; Bassam *et al.*, 1992). Therefore, annealing yields no initiation of product. Short and long primers are expected to anneal nearly perfectly, with the only mismatch at the 5' end.

RAPD fingerprints can vary according to primer concentration. The dynamics of the reaction tend to shift towards shorter products, the higher the ratio of primer to template DNA. van Belkum *et al.* (1993) suggested that, with increasing concentrations, primers will anneal to less specific target sequences, causing more and smaller fragments to be generated.

Template DNA

The quality of the template influences the outcome of the PCR. For instance, impure templates may contain polymerase inhibitors that decrease the efficiency of the reaction. Also, a large amount of RNA in a DNA template can chelate Mg^{2+} and reduce the yield of the PCR.

One of the advantages of RAPD analysis is that it can be performed on the genomic DNA of virtually any organism. However, as with any DNA amplification procedure, concentration of template DNA can influence the number of products, and therefore, RAPD analysis can yield different fingerprints for a given isolate. Interestingly, excess template DNA can result in a dampening of the amplification process, resulting in less bands (Power, 1996). Presumably, this phenomenon is a result of relative shortages of primer, resulting in little or no priming at sites where annealing is less than optimum. Many publications advocate the use of crude template DNA preparations produced from the boiled organism extract, in order to speed up the time needed for RAPD analysis.

Although this undoubtedly works in standard DNA amplification protocols, there are serious shortcomings when applied to RAPD typing of clinical isolates. Power (1996) showed fingerprints generated using purified *C. albicans* DNA, both the strains grouped by patient and the profiles, were consistent and reproducible. On the other hand, when the fingerprints were generated using unpurified DNA, obtained by simply boiling cultures of

the isolates, each fingerprint was unique, but the patterns were not reproducible. There are several explanations for these observations. To begin with, it is difficult to estimate the amount of DNA accurately if sufficient amounts of impurities are present and boiling may generate variable quantities of chromosomal and extrachromosomal DNA. Secondly, inhibitors may be present that inhibit the efficiency of DNA polymerase and thirdly, DNA may shear, resulting in different templates.

The problems are numerous and the solutions difficult and therefore, crude preparations of DNA are not recommended for RAPD typing. The use of pure, high quality DNA, is critical for reproducible RAPD products and patterns.

A further consideration in all molecular typing methods is whether typing should be done on chromosomal DNA or total DNA (which includes plasmid DNA). Two isolates should be regarded as being of the same strain if and only if, they are considered to share not only the same chromosome, but also the same plasmid content. In other words, a difference in chromosome or a difference in plasmid content, reflects a difference in strain. The possibilities for isolates to belong to different stains is further increased and complicated by the presence of transposable elements, which can become part of the genome after transfer from a plasmid.

Some typing schemes are based wholly (plasmid profiles) or in part (antibiograms) on plasmid content for their typing ability. Although these have proved useful in many epidemiological studies (David *et al.*,1991; Fantinatti *et al.*, 1994; Mayer, 1988), the readiness of gain or loss of certain plasmids means that these typing systems are often unstable. For this reason, RAPD typing should, for most purposes, be based strictly on chromosomal DNA alone. Total exclusion of plasmid DNA can never be guaranteed, particularly with large plasmids, but it can be minimized by meticulous techniques. Dilution of DNA extracts prior to amplification ensures that RAPD profiles are minimally influenced by plasmid DNA.

DNA polymerases

Commercially available thermostable DNA polymerases differ, even though they all synthesize DNA at elevated temperatures. Each enzyme is supplied with an optimized reaction buffer which differs in detail from that of its competitors. Additionally, the enzymes themselves differ in their ability to be processed, which can affect the relative quantities of the different product lengths that are synthesized. This is particularly true of the recent long-range thermostable DNA polymerases. These factors will inevitably lead to slight variations between different enzymes, although not necessarily lead to different RAPD fingerprints when using the same starting material. Some enzymes possess proofreading ability, which corrects mismatches in primer-template interactions at the 3'end and allows polymerization to proceed, which effectively increases the number of priming sites on the template.

To ensure that any variation is not a result of the inherent properties of the enzymes that are used, comparisons of RAPD fingerprints generated using different thermostable DNA polymerases require validation of those enzymes. According to Khandka (1997), an increase in the ratio of *Taq* DNA polymerase to DNA in the reaction mixture, increased the number of amplified RAPD products.

Our tests used RAPD beads containing Ampli*Taq* DNA polymerase and the Stoffel fragment. At 97.5 °C, the full-length *Taq* Pol I has an activity half-life of 9 min., and the Stoffel fragment, 21 min. *Taq* Pol I contains a polymerization-dependent 5' to 3' exonuclease activity, whereas the Stoffel fragment does not process that activity (Lawyer, *et al.*, 1993), given that the 5' to 3' exonuclease domain is deleted. Bassam *et al.* (1992) reported that the Stoffel fragment was more tolerant of reaction

conditions, more efficient in the amplification of short products and able to produce more informative fingerprints than the native Ampli *Taq*. Nonetheless, the two different thermostable polymerases produced a more complex RAPD fingerprinting pattern than did either alone. Sobral and Honeycutt (1993) reported a similar finding.

$MgCl_2$ concentration on RAPD profiles

Mg^{2+} forms soluble complexes with dNTPs to produce the actual substrate that the polymerase recognizes. The concentration of free Mg^{2+} depends on the concentrations of compounds that bind the ion, including dNTP, free pyrophosphate, and EDTA. Mg^{2+} influences enzyme activity and increases the T_m of double-stranded DNA. Excess Mg^{2+} in the reaction can increase non-specific primer binding and increase the non-specific background of the reaction. Most standard PCR studies using primers that perfectly match particular template sequences recommend lower $MgCl_2$ concentrations of 1.5 mM than those that are optimal for RAPDs. The first report on the RAPD method used 2 mM $MgCl_2$ for genetic analysis of higher plant genomes (Williams *et al.*, 1990; 1993) and 4 mM for bacterial genomes (Welsh and McCleland, 1990).

The authors used RAPD beads that contained 3 mM $MgCl_2$, which resulted in reproducible RAPD profiles of good resolution representing an array of RAPD products. The concentration of $MgCl_2$ that was used and the type of RAPD patterns obtained, accord with other studies. Berg (1994) reported that no reproducible results were obtained when a lower concentration of 2 mM $MgCl_2$ was used.

Thermal cyclers

A thermal cycler must, at a minimum, be able to do the following: accurately and reproducibly maintain the three PCR incubation temperatures, change from one temperature to another over a definable time, arrive at the selected temperatures without significantly over- or undershooting, and cycle between the temperatures repeatedly and reproducibly. The reaction temperatures actually achieved during amplification cycles will vary according to the make and model of the machine, the mode of temperature control (e.g., plate, tube or preprogrammed) and the vessel that is used (tube or microtitre tray). The thickness of the vessel walls will also affect the rate of heat transfer to the reaction mixture. Therefore, one should preferably use thin-walled reaction tubes that are designed for PCR and that fit precisely into the wells of the particular brand of thermal cycler.

The variation between brands of thermal cyclers has caused some researchers to raise doubts as to the reproducibility of RAPD typing (Caetano-Anollés *et al.*, 1992b; Mac Pherson *et al.*, 1993). This is a serious issue and must be given due consideration. There is evidence that variability arises not only because of variations between the thermal cyclers but also because of the primers that are used (Caetano-Anollés *et al.*, 1992b; MacPherson *et al.*, 1993). Some primers produce the same RAPD fingerprints with the same template DNA regardless of the thermal cycler. Power (1996) suggests that three different primers, chosen for RAPD typing of *E. faecium* isolates, consistently yielded the same respective profiles for each isolate on three different thermal cyclers.

A second consideration is that of application. Model-to-model variability between thermal cyclers may prevent the use of RAPD as a definitive typing technique but does not prevent its use as a comparative typing technique. Since each laboratory tends to use the same equipment and reagents for typing, such variability is not a problem during the investigation of local outbreaks.

Visualization of staining

Visualization is an important step but

often overlooked. The majority of RAPD fingerprints are visualized using ethidium bromide (EtBr) staining and exposure to ultraviolet light. Destaining of the gel after EtBr staining often results in RAPD fingerprints that are different from those seen on the stained gels. Similarly, the length of destaining will also affect the fingerprints. This is probably due to the disappearance of faint bands that are swamped by background fluorescence in stained gels or reduced beyond the point of detection in excessively destained gels. It is not a significant problem when comparing fingerprints on the same gel (though important information may be missed) but it must be taken into account if inter-gel comparisons are to be made. Power (1996) suggests that the fingerprints obtained from the same reaction mixture after no destaining and destaining for 15 minutes in sterile distilled water are obviously different. Evidently, staining procedures should be standardized for proper comparisons.

Interpretation of results even after standardized destaining can be a source of error. There are effectively two methods: visual examination or use of computer software. Both methods have their disadvantages. Two workers may disagree on the subjective visual interpretation of a fingerprinta problem that is exacerbated by faint bands. Computer software offers a degree of impartiality, but someone, again introducing an element of subjectivity, must set the threshold settings for exclusion or inclusion of bands. Nevertheless, inter-gel comparisons are made easier with software, particularly if large databases are used.

Apart from the presence or absence of bands, another problem in interpretation is how many band differences are needed to truly differentiate strains. This problem is inherent in any typing method. Goering (1994) suggests that a one band difference is likely to arise as a result of one genetic event, such as a point mutation or genetic rearrangement, and that such strains can be considered very similar. Two to three band differences can arise by two genetic events when the organism is not similar but may be related. Four or more band differences are definite evidence of a completely unrelated strain. The significance of each band will naturally vary according to the total number of bands generated in a fingerprint. The greater the number of bands, the less significance each individual band assumes. In our experiment, we interpreted one or more band differences as indicative of a different strain.

A degree of irreproducibility in RAPD typing has been attributed to the excess of primer present in the reaction mixture, immediately prior to thermal cycling. Non-specific primer-template annealing is sometimes a problem in DNA amplification reactions and results in non-specific amplification products (van Belkum *et al.*, 1993). If these products occur in RAPD typing, there are methods that can remove or at least reduce non-specific annealing, prior to thermal cycling. These include hot-start techniques, where a component of the reaction mix is withheld until the DNA has been denatured by heating to 90-95°C (Erlich *et al.*, 1991; Mullis, 1991; Persing, 1993) or antibody methods, that block the action of DNA polymerase until the antibody is denatured by the first heating step (Persing, 1993). Such precaution should aid the production of reproducible RAPD fingerprints.

Computing genetic similarity coefficients from RAPD data

The RAPD technique has been used for many types of genetic analyses, including genome mapping, genotype fingerprinting, phylogeny reconstruction, and measuring genetic similarities (van Belkum, 1994; Versalovic *et al.*, 1993). However, it suffers particularly from one limitation, the PCR that is used to produce informative amplification RAPD products, sometimes produces artefactual bands as well. Attempts to optimize PCR protocols designed to completely eliminate

these artefactual bands are often costly and time consuming.

Other attempts at handling RAPD artefacts, such as deleting inconsistent or faint bands, or using only those bands that are reproducible, introduces false negatives into the data. At the same time, simply ignoring artefacts and using all bands, introduces false positives. When RAPD data is used to compute coefficients of genetic similarity, such artefacts can cause significant bias in the estimation of similarity.

Artefactual bands are of two types, namely those caused by variation in experimental conditions and those caused by characteristics of the DNA to be amplified. A procedure is described that allows for correction of the bias caused by the first type of artefact, providing that replicate DNA samples have been extracted, amplified and scored. The resulting data is used to obtain an estimate of the proportion of false positive and negative bands. These values are then used to correct the bias in the computed similarity coefficients. The maximum percent bias, computed from the estimated proportions of false negatives in the RAPD data set, is proposed as a criterion for determining whether bias correction of the similarity coefficient is or is not required. For closely related organisms, two of the three coefficients most widely used with RAPD data, namely the simple matching coefficient and Nei and Li's coefficient, always exhibited a lower percent bias than the other coefficient, known as Jaccard's coefficient. Nei and Li's coefficient, which displays a lower percent bias than the simple matching coefficient, needs to be computed for all samples, not just the new ones. Jaccard's and simple matching coefficients, however, do not need to be recomputed. Only Nei and Li's coefficient has a direct biological meaning, in that it is an estimate of the expected proportion of amplified fragments shared by two samples, because they were inherited from a common ancestor. On this basis, Nei and Li's coefficient is recommended for routine computation of genetic similarities using RAPD data, particularly if PCR artefacts are present (Lamboy, 1994a; b).

The RAPD patterns of individual strains are scored and interpreted in terms of the presence or absence of RAPD electrophoretic bands. The similarity index F between the samples within each respective bacteria species is calculated using the formula (Nei and Li, 1979):

$$F_{xy} = 2n_{xy}/(n_x + n_y)$$

where n_{xy} is the number of RAPD bands shared by two samples, x and y; and nx and ny are the number of scored electrophoresis RAPD bands in those two samples, respectively. The genetic distance d was calculated using Hillis and Moritz (1990) formula

$$d = 1\text{-}F$$

The resulting matrix of pair-wise genetic distances is then used to construct a phenogram.

Comparing typability from RAPD data

An assessment of which typing method is the most efficient must be based on a number of factors: typability, reproducibility and discrimination. Of these characteristics, typability and reproducibility are relatively easy to quantify and are often expressed as simple percentages. The typability of a method is the percentage of distinct bacterial strains that can be assigned a positive typing marker. Reproducibility is the percentage of strains that give the same result on repeated testing. The discriminatory power of a typing method is its ability to distinguish between unrelated strains, with D as its numerical index. In this investigation, we used Simpson's index of diversity written as (Hunter and Gaston, 1988):

$$D = 1 - \frac{1}{N(N\text{-}1)} \sum_{j=1}^{s} n_j(n_j\text{-}1)$$

for comparing the discriminatory powers of typing methods and processes. N is the total number of strains; s is the total number of types

described; and n_j is the number of strains belonging of the j^{th} type.

We believe that this index D, of the discriminatory power for a typing method or process, greatly aids comparisons between typing systems, and that it should be used to assess each and every method and system of typing. An index value that is greater than 0.90 seems to be necessary in order for one to have confidence in the typing results. In our investigation, primers to be used were screened in terms of their indices of discrimination in order to identify useful ones for RAPD analysis.

Concluding remarks

RAPD typing is a technique that is being reported in an increasing number of publications in peer-reviewed journals. Its specific application to clinically significant micro-organisms has been reviewed previously (van Belkum, 1994), though not in technical detail. The main advantage of the technique over more traditional phenotypic methods and some of the more recent molecular methods, is that it is rapid, relatively inexpensive, technically feasible for most laboratories and theoretically applicable to any organism.

Doubts have been raised as to its reproducibility and its use as a definitive typing technique. This study confirmed that it is reproducible when the technique is standardized, which involves standard methods of DNA preparation, consistent volumes and concentrations of reagents, consistent use of the same thermostable DNA polymerase, the same thermal cycler and standard procedures for visualization of fingerprints. If RAPD typing is to become a definitive typing technique, automated systems for DNA preparation and reproducible generation and interpretation of RAPD fingerprints need to be developed. A small step in the right direction is the recent development of RAPD pellets. The pellets only require the addition of water, DNA and primer and their use significantly reduces the number of pipetting steps.

RAPD typing is a comparative typing technique that has a significant impact on the control of diseases outbreaks and the study of epidemiology. It remains imperative that those who perceive it as a quick and easy method, should not misuse the power of the technique and neglect all required precautions and controls. The technique needs therefore, to be further developed and then assessed extensively for its potential as a definitive typing method, which includes establishing criteria for acceptable typability, reproducibility and discrimination.

Much more work is needed before RAPD typing can take its place alongside typing techniques that have been long recongnized and established. Nevertheless, the research for this study suggests that the RAPD method, due to its simplicity, rapidity and low cost, may in the future contribute to better epidem iological control, by means of molecular genetic epidemiological typing of many infectious diseases. More studies of the utility of RAPD typing will aid in the development of more efficient and effective disease control measures and also in the development of new efficient and affordable vaccines.

References

Arbeit, R. D. 1999. Laboratory procedures for epidemiologic analysis. In: Manual of Clinical Microbiology 6th. Murray, P. R., Barn, E. J., Pfaller, M. A., Tenover, F. C. and Yolher, R. H. (eds.). Washington DC, ASM press.

Archer, G. L., Karchmer, A. W., Vishniavsky, N. and Johnston, J. L. 1984. Plas-mid-pattern analysis for the differentiation of infecting from noninfecting *Staphylococcus epidermidis*. J. Infect. Dis. 149: 913-920.

Bassam, B. J., Caetano-Anollés, G. and Gresshoff, P.M.1992. DNA amplifi-

cation on fingerprinting of bacteria. Appl. Microbiol. Biotechnol. 38: 70-76.

Berg, D. E., Akopyants, N. S., and Kersulyte, D. 1994. Fingerprinting Microbial Genomes Using the RAPD or AP-PCR method. Method in Mol. and Cell Biology. 5: 13-24.

Bingen, E. H., Denamur,E. and Elion, J. 1994. Use of ribotyping epidemiological surveillance of nosocomial outbreaks. Clin. Microbiol. Rev. 7: 311-327.

Branchini, M. L., Morthland, V. H., Tresoldi, A. T., Von Nowakonsky, A., Dias, M. B. and Pfaller, M. A. 1993. Application of genomic DNA subtyping by pulsed field gel electrophoresis and restriction enzyme analysis of plasmid. Diagn Microbiol. Infect. Dis. 17: 275-281.

Caetano-Anollés, G., 1993. Amplifying DNA with arbitrary oligonucleotide primers. PCR Methods Appl. 3 : 85-94.

Caetano-Anollés, G., Bassam, B. J. and Gresshoff, P. M. 1992a. DNA fingerprint-ing: MAAPing out a RAPD redefinition? Biotechnology. 10: 937.

Caetano-Anollés, G., Bassam, B. J. and Gresshoff, P. M. 1992b. Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. Mol. Gen. Genet. 235: 157-165.

David, B. P., Purushothaman, V. and Venkatesan, R. A. 1991. Comparison of plasmid profile analysis, antibiogram testing, resistotyping and biotyping in the identification of *Escherichia coli* isolates from poultry. Vet. Rec.129: 94-97.

Dryden, M. S., Talsania, H. G., Martin, S., Cunningham, M., Richardson, J. F., Cookson, B., Marples, R.R., and Phillips, I. 1992. Evaluation of methods for typing coagulase-negative staphylococci. J. Med. Microbiol. 37: 109-117.

Erlich, H. A., Gelfand, D. and Sninsky, J. J. 1991. Recent advances in the polymerase chain reaction. Science 252 : 1643-1651.

Fantinatti, F., Silveira, W. D. and Castro, A. F. 1994. Characteristics associated with pathogenicity of avian septicaemic *Escherichia coli* strains. Vet. Microbiol. 41:75-86.

Fry, N. K., Warwick, S., Saunders, N. A. and Embley, T. M. 1991. The use of 16S ribosomal RNA analyses to investigate the phylogeny of the family Legione llaceae. J. Gen. Microbiol. 137:1215-1222.

Gaston, M. A., Duff, P. S., Naidoo, J., Ellis, K., Roberts, J. I., Richardson, J. F., Marples, R. R. and Cooke, E. M. 1988. Evaluation of electrophoretic methods for typing methicillin-resistant *Staphylococcus aureus*. J. Med. Microbiol. 26:189-197.

Gaydos, C. A., Bobo, L., Welsh, L., Hook, E. W. 3rd, Viscidi, R. and Quinn, T. C. 1992. Gene typing of *Chlamydia trachomatis* by polymerase chain reaction and restriction endonuclease digestion. Sex. Transm. Dis. 19: 303-308.

Goering, R. V. 1994. Guidelines for evaluting pulsed-field restriction fragment patterns in the epidemiological analysis of nosocomial infection. Third International Meeting on Bacterial Epidemiological Markers. Cambridge, UK; Abstract F19.

Hillis, D. M. and Moritz, C. 1990. Molecular Systematics. Sunderland, Mass: Sinauer Associates.

Hulton, C. S., Higgins, C. F. and Sharp, P. M. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli, Salmonella*

typhimurium and other enterobacteria. Mol. Microbiol. 5: 825-834.

Hunter, P.R. and Gaston, M. A. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. 26: 2465-2466.

John, J. F.,Jr. 1989. Molecular analysis of nosocomial epidemics. Infect Dis. Clin. North Am. 3: 683-700.

Khandka, D. K., Tuna, M., Tal, M., Nejidat, A. and Golan-Goldhirsh, A. 1997. Variability in the pattern of random amplified polymorphic DNA. Electrophoresis 18(15):2852-2856.

Kostman, J. R., Alden, M. B., Mair, M., Edlind, T. D., LiPuma, J.J. and Stull, T.L. 1995. A universal approach to bacterial molecular epidemiology by polymerase chain reaction ribotyping. J. Infect. Dis. 171:204-208.

Kostman, J. R., Edlind, T. D., LiPuma, J. J. and Stull, T.L. 1992. Molecular epidemiology of *Pseudomonas cepacia* determined by polymerase chain reaction ribotyping. J. Clin. Microbiol. 30: 2084-2087.

Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Snin-sky, J. J. 1990. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. 18: 999-1005.

Lamboy, W. F. 1994a. Computing genetic similarity coefficients from RAPD data: correcting for the effects of PCR artifacts caused by variation in experimental conditions. PCR Methods Appl. 4(1):38-43.

Lamboy, W. F. 1994b. Computing genetic similarity coefficients from RAPD data: the effects of PCR artifacts. PCR Methods Appl. 4: 31-37.

Lawyer, F. C., Stoffel, S., Saiki, R. K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D. H. 1993. High-level expression, purification and enzymatic tic characterization of full-length *Thermus aquaticus* DNA polymerase and a truncated form deficient in 5' to 3' exonuclease activity. PCR Methods Appl. 2: 275-287.

MacPherson, J. M., Eckstein, P. E., Scoles, G. J. and Gajadhar, A. A. 1993. Variability of the random amplified polymorphic DNA assay among thermal cyclers and effects of primer and DNA concentration. Mol. Cell Probes. 7: 293-299.

Marco, F. and Jimenez de Anta, M. T. 1993. Typing methods: plasmid analysis. Advantages and disadvantages. Enferm. Infect. Microbiol. Clin. 11: 97-101.

Mayer, L. W. 1988. Use of plasmid profiles in epidemiologic surveillance of disease outbreaks and in tracing the transmission of antibiotic resistance. Clin. Microbiol. Rev. 1: 228-243.

Mekalanos, J. J. 1992. Environmental signals controlling expression of virulence determinants in bacteria. J. Bacteriol. 174: 1-7.

Mulligan, M. E., Kwok, R.Y., Citron, D. M., John, J. F. Jr. and Smith, P. B. 1988. Immunoblots, antimicrobial resistance, and bacteriophage typing of oxacillin resistant *Staphylococcus aureus*. J. Clin. Microbiol. 26: 2395-2401.

Mullis, K. B. 1991. The polymerase chain reaction in an anemic mode: how to avoid cold oligodeoxyribonuclear fusion. PCR Methods Appl. 1: 1-4.

Navarrow, E., Simonet, P., Normand, P. and Bardin, R. 1992. Characterization of natural populations of Nitrobactor spp. using PCR/RFLP analysis of the

ribosomal intergenic spacer. Arch. Microbiol. 157: 107-115.

Nei, M. and Li, W. H. 1979. Mathematical model of studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. 10: 5269-5273.

Persing, D.H. 1993. Target selection and optimization of amplification reactions. In: Diagnostic Molecular Microbiology:Principles and Applications Persing, D. H., Smith, T. F., Tenover, F. C. and White, T. J. (eds). Washington, DC: American Society for Microbiology. p. 51-87.

Peterson, L. R., Petzel, R. A., Clabots, C. R., Fasching, C. E. and Gerding, D. N. 1993. Medical technologists using molecular epidemiology as part of the infection control team. Diagn. Microbiol. Infect. Dis. 16: 303-311.

Power, E. G. M. 1996. RAPD typing in microbiology-a technical review. J. Hosp. Infect. 34: 247-263.

Rodrigo, G., Kallenius G, Hoffman E and Svenson S. B. 1992. Diagnosis of mycobacterial infections by PCR and restriction enzyme digestion. Letters Appl. Microbiol. 15: 41-44,.

Rychlik, W., Spencer, W. J. and Rhoads, R. E. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. Nucleic Acids Res. 18: 6409-6412.

Saulnier, P., Bourneix, C., Prevost, G. and Andremont, A. 1993. Random amplified polymorphic DNA assay is less discriminant than pulsed-field gel electrophoresis for typing strains of methicillin-resistant *Staphylococcus aureus*. J.Clin. Microbiol. 31: 982-985.

Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N, and Whittam, T. S. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. 51: 873-884.

Shlaes, D. M. and Currie-McCumber, C. A. 1986. Plasmid analysis in molecular epidemiology: a summary and future directions. Rev. Infect. Dis.8: 738-746.

Sobral, B. W. S. and Honeycutt, R. J. 1993. High throughout genetic mapping of polyploids using PCR generated markers. Theor. Appl. Gennet. 86: 105-112.

Struelens, M. J. 1998. Molecular epidemiologic typing systems of bacterial pathogens: current issues and perspectives. Mem. Inst. Oswaldo. Cruz. 93: 581-585.

Tabaqchali, S., Silman, R., Holland, D. 1987. Automation in clinical microbiology: a new approach to identifying microorganisms by automated pattern matching of proteins labelled with 35S-methionine. J. Clin. Pathol. 40: 1070-1087.

van Belkum, A. 1994. DNA fingerprinting of medically important micro-organisms by use of PCR. Clin. Microbiol. Rev. 7: 174-184.

van Belkum, A., Bax, R., Peerbooms, P., Goessens, W.H., van Leeuwen, N. and Quint, W. G. 1993. Comparison of phage typing and DNA fingerprinting by polymerase chain reaction for discrimination of methicillin-resistant *Staphylococcus aureus strains*. J. Clin. Microbiol. 31: 798-803.

Versalovic, J., Thearith, K. and Lupski, J.R. 1991. Distribution of repetitive DNA sequences in eubacteria and applications to fingerprinting of bacterial genomes. Nucleic Acids Res. 19: 6823-6831.

Versalovic, J., Woods, C.R.Jr., Georghiou, P.R., Hamil, R.J. and Lupski, J. R.1993. DNA-based identification and epidemiological typing of bacterial pathogens. Arch. Pathol. Lab. Med. 117:

1088-1098.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP : A new technique for DNA fingerprinting. Nucleic Acids Res. 23 : 4407-4414.

Wachsmuth, K. 1985. Genotypic approaches to the diagnosis of bacterial infections : plasmid analyses and gene probes. Infect. Control 6:100-109.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535.

Williams, J. G., Reiter, R. S., Young, R. M. and Scolnik, P. A. 1993. Genetic mapping of mutations using phenotypic pools and mapped RAPD markers. Nucleic Acids Res. 21: 2697-2702.

Woods, C. R. Jr., Versalovic, J., Koeuth, T. and Lupski, J. R. 1992. Analysis of relationships among isolates of *Citrobacter diversus* by using DNA fingerprints generated by repetitive sequence-based primers in the polymerase chain reaction. J. Clin. Microbiol. 30; 2921-2929.

Young, K. A., Power, E. G. M., Dryden, M. S. and Phillips, I. 1994. RAPD typing of clinical isolates of *Staphylococcus haemolyticus*. Lett. Appl. Microbiol. 18:86-89.