

# Molecular bacteriology: a diagnostic tool for the millennium

Tyrone L Pitt, Nicholas A Saunders

From a practical viewpoint, clinical bacteriology can be subdivided into three main activities. First and foremost is detection, isolation, and identification of pathogens. This is usually achieved by culture of specimens on specific media in defined conditions, accompanied by microscopy and tests of biochemical activity of an isolate. Second is the determination of antibiotic susceptibility, and occasionally, strain markers of clinical relevance—for example, serotypes, toxins, and so on. The third activity is epidemiological identification or comparison of isolates of the same species for infection control or population analysis. This has, until recently, been the preserve of the reference laboratory but devolution of increasingly complex techniques to diagnostic settings is gathering pace. In this article we review some areas in which molecular techniques are being applied for laboratory diagnosis and we examine developments in two basic strategies for test development and the technology required for their delivery to the laboratory bench.

## Detection, isolation, and identification of pathogens

It can be argued that for acute bacterial infections, the clinical laboratory provides solely confirmation or denial of a clinical diagnosis, because in many cases treatment of the patient will already have begun. Most bacteriological reports are generated at least 48 hours after the receipt of the specimen. This delay largely reflects the staining and microscopy of the specimen and the growth of the organism in pure culture. The limitations of microscopy and culture include poor sensitivity of staining methods; uncultivable, slow growing, or fastidious bacteria; and the potential hazard to staff of handling highly dangerous pathogens. Other conventional methods such as antibody or antigen detection may suffer from false positive or false negative reactions, cross reactions, background titres, and non-specificity. Despite these drawbacks, conventional methods continue to serve the diagnosis of infectious diseases well, but there is considerable room for novel molecular approaches.

Nucleic acid techniques for detection and identification of bacteria fall in three basic groups: (1) target amplification by polymerase chain reaction (PCR), transcription mediated amplification, nucleic acid sequence based amplification (NASBA), and so on; (2) probe amplification using ligase chain reaction or Q-beta replicase; and (3) signal amplification, as in branched DNA assay.<sup>1</sup> Numerous commercially available systems using these procedures for bacterial detection and identification

have been proposed, and some are already widely used in diagnostic laboratories. By using gene targets which are either universal (rRNA and housekeeping genes) or species specific (unique 16S and 23S rRNA sequences or other loci), it is possible to rapidly detect the presence of an increasing number of species either by PCR or by hybridisation with specific probes. Greisen *et al* list PCR primers for the 16S rRNA gene and three series of oligonucleotide probes for the detection of the PCR product.<sup>2</sup> The first series comprised a universal bacterial probe and broad range probes for Gram positive bacteria, *Bacteroides-Flavobacterium*, and Gram negative species. The second series included major species isolated from CSF in bacterial meningitis, and the third series was designed for the detection of organisms most usually associated with contamination of CSF specimens. Depending on the specimen and technology used it is possible to detect the presence and even quantitate a given species within minutes, particularly with the use of real time PCR as discussed below.

As with conventional methods there are many limitations to nucleic acid amplification technology. The most frequent are contamination of the sample by postamplification products, and the presence of substances that interfere with the PCR reaction. Other factors that may affect the sensitivity of a particular method are poor technique and contamination owing to lack of separation of work areas, primer design, stringency of assay, type of specimen, and the ubiquity of the organism sought. The sensitivity of many PCR methods is often much higher than phenotypic methods and many claim to be able to detect as few as 2 to 10 bacteria per ml of specimen, a level far below the threshold of conventional culture methods. PCR methods may also detect dead organisms or persistent bacterial products following resolution of infection and so give false positive results. For frequently occurring species, highly subspecific assays are called for in order to rule out natural or contaminating bacterial flora.

## Detection of antibiotic resistance genes

The increasing resistance of bacteria to antimicrobial compounds and the spread of multi-antibiotic resistant strains of primary pathogens such as *Mycobacterium tuberculosis* and *Salmonella typhi*, and the opportunist species *Staphylococcus aureus* and enterococci continue to generate public concern. Surveillance programmes, education of prescribers, and restriction of certain antibiotics have all been advocated as a means of controlling the rise in antibiotic resistance. However, these strategies need to be informed, at least in part, by an under-

Laboratory of Hospital Infection, Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT, UK  
T L Pitt

Hepatitis and Retrovirus Laboratory, Central Public Health Laboratory  
N A Saunders

Correspondence to:  
Dr Pitt  
email tpitt@phls.nhs.uk

standing of the mechanisms of resistance used by different bacteria.

The genetic basis of several antibiotic resistance mechanisms is known, and nucleic acid assays for them have been described.<sup>3</sup> Examples include, among others, the detection of rifampicin resistance in *M tuberculosis*, where mutations in a region of the *rpoB* gene, which encodes the  $\beta$  subunit of RNA polymerase, lead to the alteration of the rifampicin binding site,<sup>4</sup> the *mecA* gene which determines methicillin resistance in *S aureus*,<sup>5</sup> vancomycin resistance genes in enterococci,<sup>6</sup> and genes controlling enzymes that modify aminoglycosides.<sup>7</sup> Commercially available single tube assays using hybridisation capture techniques for the detection of resistance genes are becoming more mainstream, and a clear goal will be single tests to detect the widest range of resistance mechanisms.

A recently described innovation may have the potential to detect, subtype, and identify antimicrobial resistance of microbes in microscope slide smears. This technique uses peptide nucleic acid (PNA) molecules (pseudopeptides) with DNA binding capacity. The PNA molecules have a polyamide backbone instead of the sugar phosphate of DNA and RNA, and nucleotide bases attached to this backbone are able to hybridise by specific base pairing with complementary DNA or RNA sequences. Using fluorescein labelled PNA probes for concomitant detection of tuberculous and non-tuberculous species of mycobacteria, the method showed good specificity and sensitivity for the differentiation between the species in smears of cultures.<sup>8</sup> PCR amplification of target genes is also possible on conventionally stained microscope slides and this brings nearer the prospect of very rapid and sensitive test methods which are limited only by the specificity of the oligonucleotide primer for the target (McLauchlin J, personal communication)

### Strain characterisation and epidemiological typing

Traditionally, the characterisation of bacteria beyond the subspecific level has been the domain of the reference laboratory. However, this is changing and most teaching hospital microbiology departments and even some district hospitals have invested in equipment and materials to carry out studies of this nature. The last decade has witnessed an explosion of techniques for strain identification that vary in their discriminatory potential and reproducibility. These factors determine whether the DNA profiles generated can be interpreted as definitive of a specific strain population or a convenient fingerprint to facilitate the comparison of isolates within a defined epidemiological incident. Fingerprinting techniques such as random amplified polymorphic DNA (RAPD) and pulsed field gel electrophoresis (PFGE) of total DNA macrorestriction digests are widely used, but they produce complex banding patterns that may not be reproducible outside the individual laboratory. Strict standardisation of protocols has been advocated along with exchange of digitised electronic pat-

tern data to reduce variability and increase definition of strain types, but to date no fingerprinting system has been internationally adopted as a reference method. On the other hand DNA probe methods, in which a probe is hybridised to complementary target sequences immobilised on a membrane, are generally more reproducible and allow definition of strains or more broad subspecific groups, depending on the conservation of the probe target. Indeed, highly specific probes which recognise relatively unique domains in the bacterial genome lend themselves to the establishment of a binary system based simply on the reactivity or not of the probe with the strain under test. These data are amenable to automation and integrated computer analysis. A recent study on *S aureus* strains found binary typing to have greater discrimination than PFGE and RAPD, and high reproducibility.<sup>9</sup> This approach will no doubt be adopted more widely for different pathogens in the future.

The increasing accessibility of rapid and relatively inexpensive large scale DNA sequencing will clearly have a significant impact on strain characterisation and bacterial population studies. The technique of multilocus sequence typing (MLST)<sup>10</sup> is suited more to the investigation of bacterial phylogeny and evolution of population lineages than for typing of strains in hospital outbreaks. MLST indexes variation in a limited number of housekeeping genes and identifies alleles directly from the nucleotide sequences of internal fragments of these genes. In the next decade the total genome sequences of most pathogenic bacteria will be widely available. A technique such as fluorescence amplified fragment length polymorphism, which objectively samples the whole genome, will provide not only well defined and reproducible strain type markers but also allow a comparison to be made between the predicted fragments (from sequence) and those observed with an isolate.<sup>11</sup> Sequences associated with specific functions or pathogenicity are also likely to be revealed. Sequence data are very portable and permit the establishment of global databases for individual species over the internet that can be added to by any investigator using a standard protocol. Such approaches will undoubtedly remain research tools for the foreseeable future, but the data generated by these studies will probably affect the design and development of diagnostics directed at a particular strain population associated with virulence in the host.

### Future developments

The inexorable rise in the development and use of methods for molecular bacteriology is based upon the central position of the chromosome as the blueprint for life. The fact that this blueprint is encoded by simple chemicals polymerised in a particular order is the icing on the cake and has led to the current favourable situation. Now, at the close of the century, nucleic acids can be synthesised, sequenced, and manipulated by a variety of simple, reliable methods. Furthermore, the development of amplification methods such as PCR allows the specific

detection of minuscule quantities of nucleic acid. In our view the basic techniques for nucleic acid chemistry that are in place at the end of this millennium have laid the basis for a golden age of rapidity, sensitivity, and specificity in diagnostic bacteriology. Furthermore, it is becoming clear that developments of nucleic acid based tests will allow their application beyond the laboratory directly in clinical practice. This near patient testing will enhance evidence based primary care.<sup>12</sup> A disadvantage of this change in practice may be that fewer primary specimens are submitted to diagnostic laboratories; however, it may also be anticipated that the quality of material submitted for reference tests will be enhanced. In the following sections we discuss two areas of technical innovation in biotechnology which we expect to reach maturity early in the new millennium.

### Probe arrays

Nucleic acid probes have obvious advantages for the characterisation of bacterial nucleic acids. The probe is mixed with the test DNA under conditions that allow hybridisation only when a specific sequence is present. If the probe is of an appropriate length, even a single mismatch with the test DNA will effectively abolish hybridisation. Early attempts to use probes in diagnostic and reference work were hampered by the relatively low sensitivity of the systems used for detection of the target DNA. However, the potential for the use of probe systems has been greatly increased by the advent of nucleic acid amplification systems such as PCR and NASBA that provide large quantities of the probe target sequence.

Matched probes with different target sequences that hybridise with high specificity under a single set of conditions are the basis of the new probe array methods. In many cases low density arrays, typically consisting of less than 100 probes, are sufficient for the application and these can be produced by attaching premade oligonucleotides to solid supports. High density arrays may consist of many thousands of different probes and are constructed by in situ oligonucleotide synthesis on a glass support, either by photolithography or by other proprietary methods. Hybridisation of the prelabelled target nucleic acids to the bound probe is achieved in a similar way, irrespective of the array format. Bound nucleic acid is detected directly if it is labelled with a fluorescent or radioactive ligand. Other ligands used to label the target nucleic acid are detected indirectly using enzyme conjugates.

Low density probe arrays have applications in bacterial identification,<sup>13</sup> strain typing,<sup>14</sup> and drug resistance monitoring.<sup>15</sup> A good example is the system used for typing *M tuberculosis* known as spoligotyping<sup>14</sup> (for spacer oligotyping) which relies on variation between strains at the direct repeat (DR) locus. This locus consists of identical tandem repeat sequences interspersed with spacers that vary in length and base sequence. In spoligotyping, a series of oligonucleotides is covalently coupled to strips of a nylon filter. Spacers present within *M tuberculosis* DNA extracted from test sam-

ples are amplified and labelled by PCR and then hybridised to the membrane. A strain specific pattern of negative and positive hybridisation signals is produced, depending upon the specific complement of spacer sequences present.

High density probe arrays can be used for rapid sequencing of amplified DNA fragments<sup>16</sup> and they have potentially many applications. In addition, since it is already possible to manufacture arrays consisting of about half a million probes, it should be possible to combine different applications—such as species identification, strain typing, and detection of drug resistance markers—on a single chip. The chip method relies on prior knowledge of the consensus sequence of the gene and any common variants. On the chip, sets of four probes occupying adjacent positions are identical, matching part of the consensus sequence, except at a central base which is different in each probe. The next four probes on the chip array match the consensus sequence one base down from the first set, again with a centrally located variable base. Subsequent sets of four probes carry on along the consensus sequence in this way. The target sequence only hybridises to the perfectly matched probes and the sequence can therefore be determined with reasonable precision. The initial cost of fabrication of these arrays is high, but long production runs bring down the unit costs considerably. Applications of gene probes in bacteriology have been limited. However, a study reported by Gingeras *et al* shows how gene chips might be applied in future to provide a range of data from a single test.<sup>17</sup> In this work, probes for *Mycobacterium* species identification based on the *rrn* genes and for determination of susceptibility to rifampicin in strains of the *M tuberculosis* complex based on the *rpoB* gene were combined on a single array.

### Real time PCR

Real time PCR is a powerful technique combining sample amplification and analysis in a single reaction tube. Conventional PCR is designed to result in the amplification of a specific DNA fragment and subsequent analysis is required to determine whether a product with the correct sequence has been made. Analysis can be confined to simple gel electrophoresis to determine that a product of the correct size is present post-PCR; however, this evidence is considered insufficient in the diagnostic setting. Sequence based analysis of the product is preferred and this can be achieved by direct sequencing or more conveniently with probes such as in PCR enzyme linked oligosorbent assays (PCR-ELOSA).<sup>13</sup> Real time PCR machines, such as the LightCycler<sup>18</sup> or Perkin-Elmer 7700, speed up the process greatly by combining amplification with a means for detection of the specific product by fluorescence so that both steps occur conveniently in a single reaction tube.

Detection methods for real time PCR are generally dependent upon liquid phase oligonucleotide probe hybridisation which results in changes in the levels of fluorescence as the tar-

get amplicon accumulates during cycling. All of the probe systems rely on the same physical principle known as fluorescent resonant energy transfer (FRET). An example is the Sybr Green I/Cy5 system<sup>19</sup> (also known as the Bi-probe system). The primary fluor Sybr Green I acts as the photon donor when it is bound to double stranded DNA and Cy5 acts as the acceptor. Cy5 emits photons within a wavelength band monitored by the real time PCR machine. The number of photons produced with FRET is increased when DNA bound Sybr Green I and Cy5 are brought into close proximity as occurs when probe binds to its target sequence within the amplicon (fig 1).

Real time PCR has significant advantages over conventional PCR in terms of the simplicity and number of manual procedures needed and consequently the considerable cost savings that are possible in both labour and materials are likely to outweigh the high initial cost of the instruments. A clear advantage inherent in real time PCR is that the potential for contamination during analysis is effectively eliminated since the tubes are not opened post-PCR.

The use of real time instruments greatly simplifies the quantitation of initial target DNA levels (fig 2). Fluorescence is monitored during each temperature cycle so that product accumulation can be detected in its exponential phase. At this point the quantity of amplicon is still proportional to the number of copies of template. Thus the initial bacterial loads within a set of samples can be compared by examination of the curves describing product accumulation. Absolute values may be calculated if suitable standards are included in each run. Quantitative data from PCR are likely to prove useful in determining the significance of positive findings, especially for specimens taken from non-sterile sites or where specimen contamination is suspected.

Mutations in DNA sequences can also be detected by real time PCR, as the stability of the probe-target duplex can be measured. Differences between the base sequences of target genes at the probe binding site produce large differences in stability. Thus known mutations within genes resulting in changes in antibiotic susceptibility or other key phenotypes can be detected accurately and quickly using a single reaction mixture. The ability to detect subtle differences between the DNA sequences is also of potential value for strain typing.

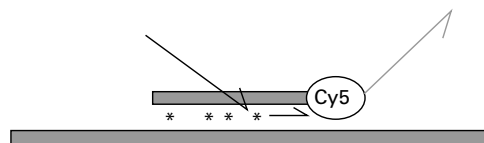


Figure 1 The Bi-probe system. When the Cy5 labelled probe (upper short bar) hybridises to the target strand (lower bar) Sybr Green I (represented by the symbol \*) is bound to the complex. Photons from the real time PCR machine's light source (black arrow) excite the Sybr Green I which then emits photons (short arrow) of the right wavelength to excite the Cy5. The excited Cy5 emits light (light grey arrow) at a third wavelength that is detected by the instrument.

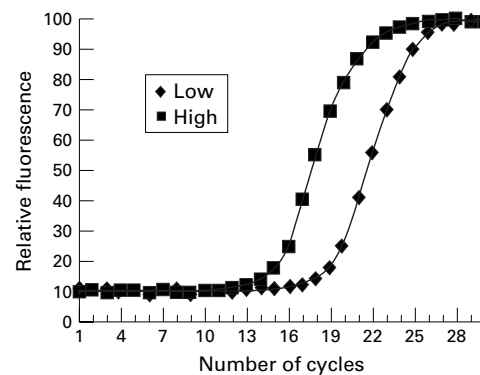


Figure 2 Quantification by real time polymerase chain reaction (PCR). The graph represents the results expected when samples with large and small numbers of template molecules are subjected to PCR with measurement of product formation by fluorescence at each cycle. For each sample no deflection from baseline is seen during the early cycles, while the level of product formed is below the detection threshold. During these early cycles it is assumed an exponential increase in copies of the target occurs. After 13 cycles (high copy number) or 16 cycles (low copy number), there is an exponential increase in fluorescence for several cycles. Following the cycles of exponential increase, a linear path is followed and finally a plateau is reached at a similar fluorescence value for each sample. Quantitation is possible because there is a linear relation between the starting template copy number and the cycle at which a significant increase in fluorescence occurs during the exponential amplification phase.

Real time PCR is currently finding many applications in the bacteriology laboratory and it is likely that it will be used more widely in future. In bacteriology the initial impact will be improvements to diagnostic PCR and quantification of targets. However, this technically simple innovation will also prove useful in the rapid detection of antibiotic resistance and virulence markers, for strain identification, and even for epidemiological typing.

- Whelen CA, Persing DH. The role of nucleic acid amplification and detection in the clinical laboratory. *Annu Rev Microbiol* 1996;50:349-73.
- Greisen K, Loeffelholz M, Purhoit A, et al. PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. *J Clin Microbiol* 1994;32:335-51.
- Courvalin P. Genotypic approach to the study of bacterial resistance to antibiotics. *Antimicrob Agents Chemother* 1991;35:1019-23.
- Hirano K, Abe C, Takahashi M. Mutations in the rpoB gene of rifampin resistant *Mycobacterium tuberculosis* strains isolated mostly in Asian countries and their rapid detection by line probe assay. *J Clin Microbiol* 1999;37:2663-6.
- Chambers HF. Methicillin resistance in staphylococci: molecular and biochemical basis and clinical implications. *Clin Microbiol Rev* 1997;10:781-91.
- Patel R, Uhl JR, Kohner P, et al. Multiplex PCR detection of vanA, vanB, vanC-2 and vanC-2/3 genes in enterococci. *J Clin Microbiol* 1997;35:703-7.
- Swenson JM, Ferraro MJ, Sahn DF, et al. Multilaboratory evaluation of screening methods for detection of high-level aminoglycoside resistance in enterococci. *J Clin Microbiol* 1995;33:3008-18.
- Stender H, Lund K, Petersen KH, et al. Fluorescence in situ hybridization assay using peptide nucleic acid probes for differentiation between tuberculous and nontuberculous mycobacterium species in smears of mycobacterium cultures. *J Clin Microbiol* 1999;37:2760-5.
- van Leeuwen W, Verbrugh H, van der Velden J, et al. Validation of binary typing for *Staphylococcus aureus* strains. *J Clin Microbiol* 1999;37:664-74.
- Maiden MCJ, Bygraves JA, Feil E, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* 1998;95:3140-5.
- Arnold C, Metherell L, Willshaw G, et al. Predictive fluorescent amplified-fragment polymorphism analysis of *Escherichia coli*: high resolution typing method with phylogenetic significance. *J Clin Microbiol* 1999;37:1274-9.
- Borriello SP. Near patient microbiological tests. *BMJ* 1999;319:298-301.
- Patel S, Yates M, Saunders NA. PCR-enzyme-linked immunosorbent assay and partial rRNA gene sequencing: a

- rational approach to identifying mycobacteria. *J Clin Microbiol* 1997;35:2375–80.
- 14 Kamerbeek J, Schouls L, Kolk A, *et al.* Simultaneous detection and strains differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;35:907–14.
  - 15 de Beenhouwer H, Lhiang Z, Jannes G, *et al.* Rapid detection of rifampicin resistance in sputum and biopsy specimens from tuberculosis patients by PCR and line probe assay. *Tuber Lung Dis* 1995;76:425–30.
  - 16 Kozal MJ, Shah N, Shen N, *et al.* Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nature Med* 1996;2:753–8.
  - 17 Gingeras TR, Ghandour G, Wang E, *et al.* Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res* 1998;8:435–8.
  - 18 Wittwer CT, Ririe KM, Andrew RV, *et al.* The LightCycler™: A microvolume multisample fluorimeter with rapid temperature control. *BioTechniques* 1997;22:176–81.
  - 19 Cane PA, Cook P, Ratcliffe D, *et al.* Use of real-time PCR and fluorimetry to detect lamivudine resistance-associated mutations in hepatitis B virus. *Antimicrob Agents Chemother* 1999;43:1600–8.