Diagnosis, epidemiology and pathogenesis of bacterial infections in the molecular era

S Sethi, T F Murphy, K L Klingman

Introduction
The rapid development of techniques in molecular biology over the past 20 years has ushered in a new era in microbiology and infectious diseases. The development of these powerful new methods has already had a profound impact in the clinical microbiology laboratory and in research on the molecular epidemiology and pathogenesis of infectious diseases. The breadth of these advances will continue to expand at an even greater rate as new tools are applied to the study of old problems as well as to the study of the ever increasing number of emerging infectious diseases.

The purpose of this review is to present an overview of some of the applications of molecular techniques to the study of human bacterial pathogens. Rather than dealing with details of methodology, this review will focus on the impact of new technology on clinical applications and research in the diagnosis, epidemiology and pathogenesis of infectious diseases. While the main focus will be human bacterial pathogens, it should be noted that the study of all microbes has been affected by application of new approaches. The selected examples which are discussed below do not represent an exhaustive list but illustrate important areas in which the use of molecular tools has been most productive.

Diagnosis of bacterial infections
For the past 100 years, the laboratory diagnosis of infectious diseases has relied on direct methods such as bacterial culture. These well described techniques can be used in most clinical laboratories. More recently, immunoassays such as immunofluorescence, immunoblot and enzyme linked immunosorbent assay have become useful adjuncts in the clinical laboratory. These techniques, usually used in the setting of large tertiary health care centres, facilitate the aetiological diagnosis of infectious diseases in many situations.

ADVANTAGES OF MOLECULAR BASED TECHNIQUES
There are two significant limitations to the above mentioned methods. Firstly, culture techniques are limited because some pathogens require a lengthy period of time to be identified, exhibit similar characteristics to non-pathogenic species, thereby necessitating specialised testing, or have fastidious growth requirements making detection by culture difficult. Secondly, the newer immunoassays are generally labour intensive or require expensive equipment, or both. These two factors account for difficulty in applying these diagnostic methods in smaller hospitals, clinics, less developed countries, or to the field. Time is a crucial variable in any diagnostic test, irrespective of the centre at which it is performed. Timely determination of the specific aetiological agent of an infectious process is critical. Rapid, accurate diagnosis of aetiology will improve the quality of health care and reduce health care costs in several ways: (1) earlier institution of appropriate therapy will reduce infection related morbidity and mortality; (2) early diagnosis in the field may prevent the rapid dissemination of an epidemic; (3) narrow spectrum antimicrobial agents can be used earlier, decreasing expenditure and the risk of adverse drug events; and (4) invasive diagnostic procedures can be deferred or avoided.

Molecular biology based techniques offer the potential for decreasing test time and broadening the scope of potential pathogens for which a clinical laboratory can test. The decision when to use a molecular technique such as the polymerase chain reaction (PCR), restriction fragment length polymorphism (RFLP) analysis or nucleic acid probe technology depends on the balance between the cost, difficulty, sensitivity, and specificity of the standard techniques compared with newer molecular techniques.

Molecular methods which will have broad applicability in the diagnosis of infectious diseases include the use of nucleic acid probes and nucleic acid amplification with PCR.

Nucleic acid probes
Nucleic acid hybridisation techniques were the first application of molecular biology to be used in the clinical microbiology laboratory. Nucleic acid probe test systems, such as GenProbe, are in widespread use for the diagnosis of sexually transmitted diseases from clinical samples, as culture confirmation assays for mycobacterial diseases, and for the detection of Legionella...
from clinical specimens. Probe based test systems for assaying various clinical samples have been described. These include studies of: sputum and bronchoalveolar lavage for the presence of Mycobacterium tuberculosis, Mycoplasma, and other respiratory pathogens; urine for the presence of urinary tract bacterial pathogens; stool for the presence of enteric pathogens; and blood and blood products for bacteria and other blood borne pathogens such as Babesia and malaria. Some of these, such as the mycobacterial probes for use in sputum samples, will become standard. This technology also has great potential for use in outbreak situations where laboratories must be set up rapidly in less developed areas as minimal equipment is required for many probe based systems. These test systems are designed to be used either directly on clinical samples for diagnosis or to confirm culture results. The test usually involves a DNA or RNA oligomer (usually 10–30 bases long) labelled with either a radioisotope, an enzyme, or a chemiluminescent substrate. The probes are hybridised either in solution or on a nylon membrane after DNA or RNA has been extracted from the clinical specimen (for example, sputum, cerebrospinal fluid or leucocytes) or culture material. The probe must be specific for the organism being tested and not bind to human DNA or RNA or other contaminating substances (such as proteins) in the sample. The benefits of these test systems are that they can be more sensitive and specific than culture detection methods, and faster than culture based methods. Although individual assays may be more expensive than traditional culture based methods, the fast turnaround time and the higher sensitivity translate into cost savings in the laboratory and in patient care. Invasive testing procedures can be avoided, patients can be treated earlier in the course of their illness and they can be treated with specific rather than empirical drug regimens.

In situ hybridisation is essentially a form of DNA or RNA probe analysis carried out on tissue histopathological specimens rather than on nylon membranes or in solution. In most clinical situations this involves the application of a labelled oligomer to formalin fixed, paraffin wax embedded tissues or cytological specimens. These techniques are used clinically to detect a variety of viral, bacterial and fungal pathogens in tissue samples. These methods have been most useful in the detection of infection by human papilloma virus (HPV) and in advancing our understanding of the epidemiology and pathogenesis of HPV, for which there is no culture technique yet available. In situ hybridisation techniques have also been important in linking Epstein–Barr virus (EBV) with various lymphoproliferative disorders and nasopharyngeal carcinoma. Because these techniques are more sensitive than routine histopathological stains, they are increasingly important in the diagnosis of infections in immunosuppressed subjects. In situ hybridisation with the addition of PCR technology will further enhance the sensitivity of the detection of pathogens in fixed tissue samples.

**Nucleic acid amplification**

For several clinical infections, the currently available diagnostic tests have limitations and the change to a molecular based technique will be rapid once sufficiently reliable tests become available. DNA and RNA amplification techniques use molecular methods to accomplish what older culture methods do; they amplify a process to a point where it can be detected. Culture media in Petri plates biologically amplifies one bacterium until it forms a colony which can be detected visually on the plate. PCR uses enzymatic duplication and detection methods to amplify one molecule of DNA. Nucleic acid amplification systems do not require an organism to be viable, one reason for its greater sensitivity over culture systems. A major drawback of PCR based systems is amplification of contaminating, irrelevant DNA.

PCR is most likely to supplant current methods in situations in which the diagnosis relies on cultures of microbes with fastidious growth requirements (for example, Borrelia pertussis, Borrelia burgdorferi, Pneumocystis carinii), when culturing is dangerous (Bacillus anthracis, HIV, Lassa fever), the organism grows slowly (Mycobacteria), or the culture method is not sensitive enough to detect the pathogen in the clinical sample (herpes simplex virus in spinal fluid). PCR technology has great potential for screening a single clinical sample, such as spinal fluid, a swab from a genital ulcer, or sputum for multiple potential pathogens. For example, a sputum sample or bronchoalveolar lavage fluid from a patient with pneumonia could be screened for various bacteria, fungi and viruses which are most likely to be the aetiological agent of pulmonary infection. Quantitative PCR can also be used to follow a therapeutic response to treatment of infections such as malaria, tuberculosis, HIV, and hepatitis C. PCR has become the method of choice for the laboratory diagnosis of HIV infection in neonates and infants.

In some clinical settings the type of organism present may not be known. Ribosomal RNA probes have proved to be useful in the detection of bacterial DNA. Oligonucleotide probes based on shared 16S rRNA sequences among all bacterial species (eubacterial primers) have been developed. These are amplified by PCR in a sample. The nucleotide sequence of the amplified DNA can be used to determine the identity of the agent in the sample. This technique has been successfully used to discover new pathogens, such as those causing Whipple's disease, cat-scratch fever, and human Ehrlichiosis. As our understanding of genetics improves, one can envisage an algorithm of PCR and a battery of primers based on these types of shared sequences.

Another area in which PCR techniques will be used diagnostically is in the determination of resistance to available anti-infective drug therapies. The application of these techniques will greatly enhance our ability to treat patients effectively early in the course of their illness. A major impediment to the current widespread use of molecular biological techniques to deter-
mine drug sensitivity is a lack of understanding of the genetic basis of drug resistance.3

Although PCR in its various forms has great promise, it probably will not totally replace older techniques in the clinical microbiology laboratory. The major drawbacks are cost, the labour intensive, technical nature of the work and lack, in some instances, of clear benefit to switch from the older method to a newer PCR method. The lack of clear benefit is often related to the observation that improved analytical sensitivity of the PCR based test may not correlate with improved clinical sensitivity.1,7 It may not be clinically relevant if a small number of live organisms or large numbers of non-viable organisms are present in a sample. As more PCR methods are used in the clinical laboratory and automation improves, the cost factor will decline in importance.1

**Epidemiology**

Molecular epidemiology in relation to the study of human disease is the use of molecular techniques to understand factors that influence the risk of developing a particular disease.13 Molecular epidemiology of infectious pathogens mainly deals with the differentiation of microbial strains and species especially in tracing the spread of a microbial species in an outbreak.14 Wider application of molecular epidemiology can permit an understanding of the pathogenesis and mechanisms of transmission of pathogens.15 Insights into phylogenetic evolution of bacteria and species relatedness can also be gained by molecular epidemiology.16

**CHARACTERISTICS OF TYPING SYSTEMS**

For an epidemiological technique to be useful in the study of infectious pathogens, it should have certain characteristics; these are listed in table 1 and discussed below.17 The superiority of molecular or genotypic techniques over traditional phenotypic techniques (for example, biotyping, phage typing, morphology) with regard to these characteristics has been clearly established.

**Table 1 Characteristics of typing assays**

<table>
<thead>
<tr>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminatory ability</td>
</tr>
<tr>
<td>Typability</td>
</tr>
<tr>
<td>Reproducibility</td>
</tr>
<tr>
<td>Sensitivity</td>
</tr>
<tr>
<td>Ease of application</td>
</tr>
</tbody>
</table>

Discriminatory ability is the ability of an assay to discriminate between species and, more importantly, between clonally related but distinct strains of a species. This is the most vital characteristic of an epidemiological typing technique. Genotypic techniques have greater discriminatory ability than phenotypic techniques.

Typability is the proportion of strains which can be typed by a particular typing method. Phenotypic methods depend on growth of the pathogen in vitro and expression of a particular characteristic. Thus, fastidious pathogens or pathogens lacking the relevant phenotypic characteristics are non-typable by the phenotypic methods. Genotypic methods vary with respect to their typing ability. Plasmid DNA based methods often have low typability as some strains may lack plasmids. Chromosomal DNA methods have very high rates of typability because of the universal presence of chromosomal DNA and restriction sites.

Reproducibility is the ability of an assay to give the same result for a particular sample when performed many times. Before a technique can be applied for general use, reproducibility in several laboratories needs to be addressed. Phenotypic characteristics of pathogens are notoriously variable in vitro and with passage. The genotype is stable and therefore, molecular methods have a distinct advantage with regard to reproducibility.

Sensitivity of a molecular method is more of an issue when the technique is used for diagnosis. PCR based techniques can detect very low concentrations of pathogens in clinical samples. If species specific primers are used for PCR, diagnosis and typing could be accomplished in a single step. This is especially useful for fastidious pathogens.

Ease of application is the factor governing whether a typing method will remain confined to research laboratories or become widely used in clinical laboratories. PCR based methods are extremely promising as they are easily applied and can be used to process a large number of specimens simultaneously. Problems of contamination by extraneous DNA is a limitation of these methods.

**CLASSIFICATION OF MOLECULAR METHODS**

Various typing methods have been developed for and applied to epidemiological studies of microbial pathogens. A simple classification of methods is presented in table 2.

**Table 2 Classification of molecular methods used in bacterial epidemiology**

<table>
<thead>
<tr>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Plasmid analysis</td>
</tr>
<tr>
<td>B. Restriction enzyme digestion (RFLP) plasmid DNA chromosomal DNA</td>
</tr>
<tr>
<td>C. DNA amplification (PCR) of specific genes with repetitive DNA element primers (Rep-PCR) with arbitrary primers (randomly amplified polymorphic DNA or RAPD) with ribosomal RNA primers (PCR ribotyping)</td>
</tr>
<tr>
<td>D. Combination of B and C RFLP analysis of PCR products</td>
</tr>
</tbody>
</table>

A DNA fingerprint is developed by each of these methods. This DNA fingerprint differs among strains because of polymorphism—that is, length differences. Agarose gel electrophoresis permits size fractionation of the DNA and detection of this polymorphism. The agarose gel can be stained and the banding pattern of DNA studied directly.18 Alternatively, the DNA can be transferred to a nylon membrane and hybridised with DNA probes of specific sequences.19,20

In the next section, two of these methods as they have been applied to mucosal Gram negative bacterial pathogens will be discussed.
RFLP ANALYSIS OF GENOMIC DNA

RFLP analysis (also known as restriction endonuclease analysis) is one of the earliest molecular methods to be applied to the epidemiological study of microbial pathogens. The organism is cultivated and chromosomal DNA is isolated and digested with restriction enzymes. The resulting mixture of DNA fragments is separated by agarose gel electrophoresis to produce the DNA fingerprint.

In the early applications of this technique the restriction enzymes used were frequent DNA cutters whose recognition site is a 4–6 base pair DNA sequence. Therefore, the pattern produced has a large number of bands, making interpretation difficult and limiting discriminatory ability. A variation of this technique is RFLP analysis by pulsed field gel electrophoresis (PFGE), also known as contour clamped homogeneous electric field (CHEF) electrophoresis. This method facilitates the separation of large fragments of DNA. Chromosomal DNA is nitrified and digested within an agarose matrix. In this manner the chromosomal DNA is protected from the shearing forces inherent in conventional methods of DNA purification. Restriction enzymes which have infrequent 8 base pair recognition sites are used to produce large DNA fragments. The banding pattern produced on electrophoresis, therefore, is of a small number of large molecular weight bands which is easier to interpret and has a greater discriminatory ability.

*Branhamella catarrhalis* is a Gram negative mucosal pathogen which has become the third most common cause of otitis media in children and acute exacerbations of chronic bronchitis in adults.11 There was no well established typing system available for this pathogen. On protein electrophoresis, the outer membrane proteins of different strains of this bacteria are remarkably similar. A high percentage of β-lactamase producing strains make antibiotic resistance patterns insensitive for strain analysis. Plasmids are present but not in great enough numbers to allow plasmid analysis. The epidemiology of this bacterium therefore had not been well established.

Nosocomial outbreaks of *B catarrhalis* had been studied successfully with conventional RFLP analysis showing the epidemiological relatedness of the strains isolated and patient to patient transmission.22 However, as discussed above, the large number of bands seen preclude accurate analysis of a large number of strains isolated in an endemic setting. We had such a collection of *B catarrhalis* strains obtained from patients with bronchiectasis over several years. PFGE distinguished the 37 strains accurately into 13 classes and established the dynamics of colonisation by *B catarrhalis* in this patient population.18

REPETITIVE DNA SEQUENCE PCR TYPING

Amplification in vitro of specific portions of DNA by PCR has found wide applications in molecular epidemiology of pathogenic organisms. Initial applications of PCR used primers based on species specific DNA sequences followed by sequencing, RFLP or hybridisation of the PCR products. These techniques were laborious and the primers could not be used with multiple bacterial species. Recent developments, such as Rep-PCR,19 RAPD20 and PCR ribotyping,24 use primers based on DNA sequences which are present in many species of bacteria, while still retaining their ability to distinguish the various strains within a species. These sequences are repetitive—that is, present in multiple copies per genome but not constant in their position and frequency among various strains. Thus, the PCR amplifies genomic DNA present between the repetitive sequences; the size and number of these PCR products vary among strains. On agarose gel electrophoresis, the PCR products produce a DNA fingerprint. Repetitive extragenic palindromic (REP) elements and enterobacterial repetitive intergenic consensus (ERIC) sequences are examples of repetitive sequences which have been used as primers for typing.

Non-PCR methods include *Haemophilus influenzae* (NTHI) is a common mucosal pathogen in otitis media and chronic bronchitis.21 Several typing systems utilising phenotypic characteristics—for example, biotyping and outer membrane protein patterns, have been developed. These systems have drawbacks such as lack of reproducibility, phenotypic variation, and poor discriminatory ability. An efficient, reproducible typing system would be invaluable for the understanding of the dynamics of colonisation and pathogenesis of chronic infections by NTHI in chronic respiratory diseases such as chronic obstructive pulmonary disease (COPD) and cystic fibrosis.

Van Belkum et al20 used primers based on ERIC sequences in a repetitive DNA sequence PCR assay to study 40 strains of *NTHI* from patients with cystic fibrosis and COPD. They reliably and reproducibly distinguished strains with the DNA fingerprints obtained. When compared with RFLP analysis, the PCR fingerprints were equally discriminatory and much simpler to perform and interpret. If these results are substantiated with a larger number of strains from varied epidemiological settings, an adequate typing system for NTHI could be established.

The research potential of molecular epidemiology is well established. Genotypic methods, though superior to phenotypic methods, have limitations. Each molecular method examines the microbial genome in a different manner. Thus, when different methods are applied to a population of strains, varied results can be obtained.18 The gold standard among these methods has yet to be defined. Interpretation of patterns obtained with genotypic techniques lacks standardisation. What are the limits of similarity in a banding pattern which defines clonally related strains versus different clones? These limits are somewhat arbitrary and vary among investigators. In general, the RFLP methods are cumbersome, need expensive equipment and are difficult with large numbers of strains. The PCR based methods do not have these disadvantages but are prone to contamination and need strict quality controls. Of the two
groups of methods, the PCR based methods are more promising for application in clinical microbiology laboratories. Cost effectiveness of these methods in a clinical setting needs to be addressed.

Pathogenesis
The application of molecular approaches to the study of bacterial infections has advanced the understanding of mechanisms of pathogenesis over the past decade. The ability to characterise the molecular interaction between host and pathogen is critical for understanding pathogenesis so that better methods of prevention and treatment can be designed.

**ADHERENCE OF BACTERIAL PATHOGENS TO HOST CELLS**
The initial step in the pathogenesis of bacterial infections is often colonisation of a mucosal surface. In order for colonisation to occur molecules on the bacterial cell surface (adhesins) must bind to molecules on the epithelial cell surface (receptors). This adhesin–receptor interaction partly accounts for the host and organ specificity of bacterial infection. For example, bacterial species which are exclusively human pathogens (for example, *Neisseria gonorrhoeae, H influenzae, B catarrhalis*) may derive their host specificity from the interaction of adhesin molecules with receptors which are unique to human epithelial cells. Adhesin–receptor interactions probably partly account for tissue specificity as well. Respiratory tract pathogens express adhesins which bind receptors on epithelial cells of the respiratory tract whereas genitourinary tract pathogens express adhesins which bind receptors on urogenital epithelial cells. Elucidating the mechanisms and specificity of these molecular interactions is important for understanding the pathogenesis of bacterial infection.

Molecular studies of adhesins of NTHI have helped to elucidate the molecular mechanisms of adherence to human respiratory epithelial cells. The role of a high molecular weight outer membrane protein (HMW-1) in adherence to human cells was investigated. A kanamycin resistance cassette was placed into the HMW-1 gene which was in a plasmid vector. A strain of NTHI was transformed with the linearised plasmid and a kanamycin resistant transformant which did not express the HMW-1 was constructed. This isogenic mutant showed noticeably decreased adherence to Chang epithelial cells, a human cell line, compared with the parent strain, which expressed HMW-1. This series of experiments established that HMW-1 is an important adhesin for NTHI. The relation between HMW-1 and other adhesins has been explored and the role of the adhesins in binding to other cell types has been studied. The capability to construct isogenic mutants is a powerful method for elucidating the molecular mechanisms of adherence of pathogens to human cells.

The differential adherence of bacteria to specific host cells plays a role in the pathogenesis of infection. *N gonorrhoeae* expresses a family of variant outer membrane proteins (Opa) which mediate adherence to human cells. The bacterial chromosome contains 11 different copies of the gene. Each allele can be turned on and off independently. Expression of some opa alleles enables the gonococcus to bind and invade epithelial cells. When the organism expresses other opa alleles, the specificity for epithelial cells is lost; instead, the bacterium interacts with polymorphonuclear neutrophils. This host cell specificity which is intrinsic to Opa proteins expressed by the gonococcus confers important virulence properties on the organism. Depending on the host tissue and the stage of infection, the bacterium expresses the Opa protein with appropriate host cell specificity. Understanding these mechanisms is critical for comprehending the pathogenesis of infection so that new measures to treat and prevent bacterial infections can be developed.

**MOLECULAR STRUCTURE OF BACTERIAL OUTER MEMBRANE PROTEINS**
An understanding of the molecular architecture of Gram negative outer membrane proteins has had an important influence on approaches to the study of pathogenesis of infections caused by these organisms. A variety of molecular approaches has been applied for studying the antigenic structure of outer membrane proteins. These include the development of monoclonal antibodies, cloning outer membrane protein genes, sequence determination and analysis, recombinant expression of selected regions of protein molecules, mutagenesis studies, and others.

An understanding of molecular structure of surface molecules is also essential. For example, investigators have been studying the role of NTHI in lower respiratory tract infection in adults with chronic bronchitis for decades. A serological approach has been used to evaluate a potential role of NTHI in these infections. The development of an antibody response to the putative pathogen has been accepted as evidence that the organism causes the disease. Several studies from different centres have used this approach. Some studies showed no antibody response, some studies showed higher titres of antibodies in patients compared with controls but no correlation with clinical exacerbations, and some studies showed an antibody response to NTHI following infection (reviewed in ). These apparently conflicting results defied explanation until more was known about the antigenic structure of the outer membrane of this bacterium.

The molecular architecture of the P2 porin protein of NTHI accounts for the apparently conflicting results of serological studies in patients with chronic bronchitis. P2 is the most abundant protein in the outer membrane and contains an immunodominant, strain specific epitope. Preliminary studies indicate that the predominant immune response to NTHI is directed against this strain specific immunodominant epitope. The host is thereby protected from infection by the same strain but
remains susceptible to infection by other strains of the species. Therefore, serological studies which used laboratory strains of NTHI rather than the patient’s homologous strain failed to detect the predominant immune response. Furthermore, an understanding of the antigenic structure of the major outer membrane proteins and the host immune response to specific epitopes on these outer membrane proteins allows one to explain why adults with chronic bronchitis experience recurrent infections by the same bacterial species in spite of having an intact immune system.

Over the past decade much has been learned about the outer membrane structure of Gram negative bacteria. This work has led to the identification of a common theme in the molecular structure of porin molecules. These protein molecules consist of a β pleated sheet with 16 strands that traverse and remain largely buried within the outer membrane. Interpersed within the relatively conserved backbone of the protein are eight peptide loops that are potentially exposed on the bacterial surface. These surface accessible loops show antigenic diversity among strains of a species. This structure has important implications for the understanding of the pathogenesis of recurrent infections. As the surface exposed loops of the major porin protein show antigenic heterogeneity among strains, an immune response to this part of the molecule will be relatively strain specific.

These two examples (molecular mechanisms of adherence and molecular structure of outer membrane proteins) are just two of numerous observations in which molecular approaches have enabled observations regarding pathogenesis of bacterial infections to be made. Studies of the molecular mechanisms of bacterial pathogenesis promises to be a productive area of investigation. The elucidation of these mechanisms will lead directly to the development of new therapeutic and prophylactic modalities for a variety of bacterial infections.

Summary
The rapid progress in the development of powerful techniques in molecular biology over the past 20 years has had an enormous impact on the disciplines of infectious diseases and microbiology. Our understanding of human bacterial pathogens has undergone significant advances, particularly in diagnosis, epidemiology and pathogenesis. Nucleic acid probes and amplification by PCR have forever changed the laboratory diagnosis of many bacterial infections by providing specific, accurate and rapid means to identify DNA from pathogens in clinical samples. The application of molecular techniques has led to the development of typing systems based on the genetic structure of bacterial pathogens. This has provided investigators with highly effective tools for studying epidemiology, transmission, evolutionary relations, and pathogenesis of bacterial pathogens. The ability to isolate and manipulate bacterial genes has resulted in a virtual explosion of new information on molecular mechanisms of bacterial pathogenesis. These important advances in diagnosis, epidemiology and pathogenesis will lead directly to better means of treatment and prevention of a variety of bacterial infections in the next decade.

Bacterial infection in the molecular era


