Lepr Rev (1993) **64,** 289–294

Editorial

THE MOLECULAR BIOLOGY OF *MYCOBACTERIUM LEPRAE*

Introduction

Although *Mycobacterium leprae* was one of the first organisms to be established as causing human disease,[1] less is known about it than virtually any other bacterium of medical importance. This is because it has not yet been grown in bacteriological culture, and hence many of the approaches adopted for studying the basic biology and biochemistry of other bacteria are not applicable. To some extent this has been circumvented by the use of experimental animals. However, the limitations of working *in vivo*, and, in particular, the uniquely long generation time of *M. leprae*, even when growing optimally in the tissues of experimental animals, have proved major obstacles.

The recent application of a molecular biological approach to studying *M. leprae* is proving crucial in our basic understanding of the biology of the organism. Studies on its relationship to other bacteria, including mycobacteria, are now possible, while it is also possible to obtain large amounts of compounds, such as protein antigens or enzymes, which can be used to study interactions with the host at the molecular level. It should prove possible in the near future to use the basic information on the biology of *M. leprae* to devise strategies for circumventing the need to culture *M. leprae* for many clinical microbiological applications, such as drug-sensitivity testing, while further advances in understanding the basic biology of the organism should lead to a greater understanding of the host–pathogen relationship.

The molecular biological approach to studying *M. leprae*

The availability of large numbers of *M. leprae* from infected armadillo tissue provides access to *M. leprae* nucleic acids. The preparation of libraries of *M. leprae* DNA fragments in cosmid or bacteriophage *E. coli* vectors can serve as an inexhaustible source of *M. leprae* DNA.[2,3] By studying the nucleic acids and using molecular biological techniques much is being learned about the basic biology of *M. leprae*. In future, it seems likely that this knowledge will be put to practical use, and techniques which circumvent the need to culture the organism will be devised.

0305-7518/93/064289+06 $01·00 © Lepra

THE TAXONOMIC POSITION OF *M. LEPRAE* BASED ON NUCLEIC ACID ANALYSIS

Mycobacteria belong to the high guanosine plus cytosine (G+C) subdivision of Gram positive bacteria, which also includes such actinomycetes as *Streptomyces*. Most mycobacteria have between 60 and 67 mole per cent G+C in their DNA, compared to 72% for streptomycetes, and 49% for *E. coli*. Interestingly *M. leprae* has a lower G+C% than other mycobacteria (56%; Clark-Curtiss *et al.*,[2] Imaeda *et al.*[3]), raising questions about its taxonomic position.

Using a molecular biology approach it is possible to address this. The 16S ribosomal RNA (rRNA) gene is convenient for studying phylogenetic relationships. rRNA is present in all free-living organisms; some stretches of the nucleotide sequence are variable, whereas others are conserved by functional constraints. The conserved sequences permit the alignment of the variable sequences; the degree of variation within these sequences is a reflection of the taxonomic and phylogenetic relationship between organisms. When applied to the mycobacteria, it is clear that the division between slow-growing and fast-growing species is a true phylogenetic division, that is, the division is clearly recognizable in terms of the 16S rRNA sequence,[4] and that *M. leprae* is correctly positioned in the slow-growing mycobacteria group.[5–7]

CLONING AND EXPRESSION OF *M. LEPRAE* GENES

One of the major areas of impact of molecular biology on the study of *M. leprae* is cloning of genes and high level expression to provide large amounts of protein. This has been widely applied for the production of recombinant proteins of immunological interest. This is a particularly attractive approach to use with an organism that cannot be grown in culture.

The approach was initiated by Young *et al.*,[8] who cloned small fragments of randomly sheared *M. leprae* DNA into the *E. coli* phage λgt11. In this vector, the cloned DNA is inserted into the coding region of the *E. coli* β-galactosidase gene, and the resulting protein is a hybrid of β-galactosidase and the *M. leprae* DNA-encoded protein. Expression of the *M. leprae* protein can then be detected immunologically, for example using monoclonal antibodies which have been raised against *M. leprae*. Once the *M. leprae* fragment has been identified, it is then possible to obtain the entire gene, to sequence it, and to manipulate high-level expression of the protein so that it can be purified and used for immunological studies.

This has now been carried out for a number of *M. leprae* proteins by a variety of different groups. In the initial experiments, mouse monoclonal antibodies were used to screen for expression of *M. leprae* antigens.[8] Characterization of the proteins recognized by the monoclonal antibodies revealed that many of them belonged to a well-characterized group of highly conserved proteins—heat shock proteins.[9,10] Several of these have now been expressed at high level and are being used to investigate their interaction with the immune response.

Alternative approaches for screening *M. leprae* expression libraries for immunologically important proteins have involved using sera from patients[11–13] or T cells.[14] These studies have identified proteins other than those selected by monoclonal antibodies.

Although the studies outlined above have demonstrated that *E. coli* is an excellent host for cloning and expression of mycobacterial proteins, for many applications using

mycobacteria themselves as cloning hosts is necessary. Many studies are now in progress to develop mycobacterial vectors so that experiments involving gene transfer between mycobacteria can be carried out.[15–20]

MOLECULAR METHODS FOR RAPID DETECTION AND IDENTIFICATION OF *M. LEPRAE*

There has been a great deal of interest in recent years in using molecular techniques to detect and identify microorganisms. Such techniques would be particularly attractive where the microbe in question could not be cultured, as with *M. leprae*. The polymerase chain reaction (PCR) is a technique used for amplifying small quantities of DNA (or RNA) to the point where they are readily visualized. The technique exploits two basic principles of biology; first that the affinity of single-stranded DNA for its complementary sequence is strong and specific, and secondly that it is possible to use one strand of DNA as a template to synthesize its complementary strand. In essence, the PCR involves the cyclical synthesis of DNA copies from a single template strand; the resulting increase in the number of copies of the template DNA is exponential, allowing, in principle, a single copy to become amplified so that it is readily detectable on a gel.

The reaction involves synthesizing copies of a region of DNA ('the target sequence') from oligonucleotide primers which bind to opposite strands of DNA and flank the target sequence. Each cycle of the PCR involves separating the strands of target DNA by heat denaturation, binding of the oligonucleotide primers to their complementary sequences within the target DNA ('primer annealing') and synthesis of new DNA between the primer sequences ('primer extension'). Each newly-synthesized DNA strand then becomes a template for the next cycle, so that after 20 or 30 cycles the original sequence has been amplified by approximately 10-million fold. The various stages of each cycle (denaturation, annealing and primer extension) occur optimally at different temperatures; by using a DNA polymerase which is not destroyed at high temperatures, the entire reaction can be performed in a single tube by simply changing the temperature on a cyclical basis, a process which can be achieved automatically using a microprocessor-controlled thermocycler.

PCR has been used to study DNA from fossils, to study the relatedness of different species of plants, animals and bacteria and in the forensic laboratory to identify criminals. The most obvious use in the clinical microbiology laboratory is in diagnosis and species identification, where its combination of exquisite sensitivity (minute quantities of DNA can be amplified to detectable levels) and specificity (the degree of specificity depends on the selection of primers; however, the starting material can be a complex mixture of host tissue and contaminating organisms) make the technique potentially extremely powerful.

A number of approaches have been described in which PCR has been used to detect and identify *M. leprae* in infected tissue. The main differences are at the point of DNA extraction and in the choice of target sequences. The latter consideration embodies the specificity of the technique—a target sequence with species-specific regions on which the oligonucleotide probes are based will provide a species-specific detection system. Some of the target sequences which have been used for the detection of *M. leprae* are shown in Table 1. PCR represents a powerful, sensitive and specific procedure which has been shown to work in experimental systems for detecting *M. leprae*. However, its role in the clinical diagnostic or clinical research laboratory is unclear. The current keen interest in the application of PCR technology for the detection of *M. leprae* should be pursued with critical objectivity.

Table 1. PCR *M. leprae* target sequences

Target sequence	Reference
Antigen encoding sequences	
18 kD	Williams *et al.*[21]
36 kD	Hartskeerl *et al.*[22]
65 kD	Woods & Cole[23]
Non-antigen encoding sequences	
Repetitive sequence	Woods & Cole[23]
Ribosomal RNA sequences	Cox *et al.*[6]

THE *M. LEPRAE* GENOME MAP

As discussed above, libraries of *M. leprae* DNA can serve as an inexhaustible source of material. However, a great deal of work is still required to obtain the piece of DNA of interest. The establishment of an 'ordered' library and a genetic map of the *M. leprae* genome[24] provides an important resource which should prove to be of great benefit to scientists working on the molecular biology and genetics of *M. leprae*.

Essentially, an ordered library involves 'fingerprinting' the cloned DNA fragments such that their relative position on the *M. leprae* chromosome can be determined. The position of known genes can then be located by hybridization with previously identified probes, and a comprehensive picture of the organization of the genome obtained. Currently an ordered library and map has been constructed, such that virtually complete coverage of the chromosome has been obtained and the position of 72 genetic loci determined.[24]

The value of such a library and map is that it provides an important resource for scientists interested in the molecular biology of *M. leprae*. For example, it has been possible to identify and study the gene encoding the target enzyme of rifampicin, the β subunit of RNA polymerase, making it possible to devise methods to rapidly diagnose rifampicin resistance (see below). This could be achieved with the minimum of effort because the locus had been identified using probes from *E. coli*.

THE MOLECULAR BASIS OF RIFAMPICIN RESISTANCE, AND RAPID DETECTION OF RESISTANT ISOLATES

Detection of rifampicin resistance currently takes between 3 weeks and 2 months for *M. tuberculosis* and between 6 and 12 months for *M. leprae*. Thus treatment is given empirically, without prior knowledge of the sensitivity status. Recently the molecular basis of rifampicin resistance of *M. leprae*[25] and *M. tuberculosis*[26] has been determined, and techniques for diagnosing resistance within hours, rather than weeks or months, are being developed.[26]

Resistance to rifampicin involves alteration of the RNA polymerase; because most *E. coli* rifampicin resistance mutations occur within a short stretch of DNA in the gene that encodes the RNA polymerase subunit β (*rpoB*), this part of the gene was concentrated on to define mutations in *M. tuberculosis* and *M. leprae*. For *M. tuberculosis* 98% of resistance mutations could be identified within a region of 23 amino acids,[26] while

all 9 isolates of rifampicin resistant *M. leprae* had mutations in the same site.[25] Telenti *et al.*[26] have used this information to devise a rapid test for rifampicin-resistant *M. tuberculosis*. Using a technique known as PCR–SSCP (Polymerase Chain Reaction–Single strand conformation polymorphism), they are able to distinguish rifampicin resistant isolates of *M. tuberculosis* within hours, rather than weeks. A similar approach could be used for the rapid detection of rifampicin resistant *M. leprae*.

Conclusions

Although the underlying problems of studying *M. leprae* in the laboratory remain, i.e. it still cannot be cultured *in vitro*, and takes many months to grow in laboratory animals, the availability of large amounts of organisms from the armadillo have enabled studies on the biochemistry, immunochemistry and molecular biology of *M. leprae* to move forward. I have not attempted, in this brief review, to document all the facts known about the organism but have tried to emphasize the sort of approaches which are now available for studying the basic biology of the organism. By using the examples of PCR and rifampicin resistance, I have tried to illustrate how these basic studies might provide the basis for developing techniques which are useful in the clinical laboratory and circumvent the need to culture the organism.

There is a great deal of (justified) optimism about the success of MDT and the potential for control and elimination of leprosy. However, our experience with tuberculosis emphasizes that this optimism should be tempered with caution; in spite of the fact that effective chemotherapy against *M. tuberculosis* has been available for several decades, TB is still a major public health problem in developing countries, and the spectre of increased incidence and multiple drug resistance in industrially developed countries is now emerging.[27] Although prospects for eradicating leprosy are now brighter than at any other time, *M. leprae*, like *M. tuberculosis*, may yet have some surprises in store.

Acknowledgment

This review is based on a presentation to the Royal Society of Tropical Medicine and Hygiene held to mark the 30th anniversary of the Ridley–Jopling classification. The full paper will be published in the *Transactions of the Royal Society of Tropical Medicine and Hygiene*, and I wish to thank the Society for permitting publication of this modified version.

National Institute for Medical Research M. J. COLSTON
The Ridgeway
Mill Hill
London NW7 1AA

References

1 Hansen GA. Undersogelser angaaende spedalskhedens aarsager. *Norsk Magazin. f. Laegevid*, 1874; **4:** 1–88.
2 Clark-Curtiss JE, Jacobs WR, Docherty MA, Ritchie LR, Curtiss R, III. Molecular analysis of DNA and construction of genomic libraries of *Mycobacterium leprae*. *J Bacteriol*, 1985; **161:** 1093–1102.

[3] Imaeda T, Kirchheimer WF, Barksdale L. DNA isolated from *Mycobacterium leprae*: genome size, base ratio, and homology with other related bacteria as determined by optical DNA-DNA reassociation. *J Bacteriol*, 1982; **150:** 414–17.
[4] Stahl DA, Urbance JW. The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. *J Bacteriol*, 1990; **172:** 116–24.
[5] Smida J, Kazda J, Stackebrandt E. Molecular-genetic evidence for the relationship of *Mycobacterium leprae* to slow-growing pathogenic mycobacteria. *Int J Lepr*, 1988; **56:** 449–54.
[6] Cox RA, Kempsell K, Fairclough L, Colston MJ. The 16S ribosomal RNA of *Mycobacterium leprae* contains a unique sequence which can be used for identification by the polymerase chain reaction. *J Medical Microbiology*, 1991; **35:** 284–90.
[7] Kempsell KE, Yuan-en Ji, Estrada-G, ICE, Colston MJ, Cox RA. The nucleotide sequence of the promoter, 16S rRNA and spacer region of the ribosomal RNA operon of *Mycobacterium tuberculosis* and comparison with *Mycobacterium leprae* precursor RNA. *J General Microbiology*, 1992; **138:** 1717–27.
[8] Young RA, Mehra V, Sweetser D, Buchanan T, Clark-Curtiss J. Davis RW, Bloom BR. Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*. *Nature*, 1985; **316:** 450–2.
[9] Young DB, Lathigra R, Hendrix R, Sweetser D, Young RA. Stress proteins are major immune targets in leprosy and tuberculosis. *Proc Natl Acad Sci USA*, 1988; **85:** 4267–70.
[10] Young RA, Elliot TJ. Stress proteins, infection and immune surveillance. *Cell*, 1989; **59:** 5–8.
[11] Laal S, Sharma YD, Prasad HK, Murtaza A, Singh S, Tangri S, Misra RS, Nath I. Recombinant fusion protein identified by lepromatous sera mimics native *Mycobacterium leprae* in T cell responses across the leprosy spectrum. *Proc Natl Acad Sci USA*, 1991; **88:** 1054–58.
[12] Sela S, Thole JER, Ottenhoff HM, Clark-Curtiss JE. Identification of *Mycobacterium leprae* antigens from a cosmid library: Characterisation of a 15 kilodalton antigen that is recognised by both humoral and cellular immune system in leprosy patients. *Infect Immun*, 1991; **59:** 4117–24.
[13] Cherayil BJ, Young RA. A 28-kDa protein from *Mycobacterium leprae* is a target of the human antibody response in lepromatous leprosy. *J Immunol*, 1988; **141:** 4370–5.
[14] Mustafa AS, Gill HK, Nerland A, Britton WJ, Mehra V, Bloom BR, Young RA, Godal T. Human T-cell clones recognise a major *M. leprae* protein antigen expressed in *E. coli*. *Nature*, 1986; **319:** 63–5.
[15] Jacobs WR, Tuckman M, Blooom BR. Introduction of foreign DNA into mycobacteria using a shuttle plasmid. *Nature*, 1987; **327:** 532–5.
[16] Lee MH, Pascopella L, Jacobs WR, Hatfull GF. Site-specific integration of mycobacteriphage L5: Integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guerin. *Proc Natl Acad Sci USA*, 1991; **88:** 3111–15.
[17] Ranes MG, Rauzier J, Lagranderie M, Gheorghiu M, Gicquel B. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: Construction of a "mini" mycobacterium–*Escherichia coli* shuttle vector. *J Bacteriol*, 1990; **172:** 2793–7.
[18] Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, Snapper SB, Barletta RG, Jacobs WR, Bloom BR. New use of BCG for recombinant vaccines. *Nature*, 1991; **351:** 456–60.
[19] Snapper SB, Lugosi L, Jekkel A, Melton RE, Kieser T, Bloom BR, Jacobs WR. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. *Proc Nat Acad Sci USA*, 1988; **85:** 6987–91.
[20] Aldovini A, Young RA. Humoral and cell-mediated immune responses to live recombinant BCG-HIV vaccines. *Nature*, 1991; **351:** 479–82.
[21] Williams DL, Gillis TP, Booth PJ, Looker D, Watson JD. The use of a specific DNA probe and polymerase chain reaction for the detection of *Mycobacterium leprae*. *J Infect Dis*, 1990; **162:** 193–200.
[22] Hartskeerl RA, de Wit MYL, Klatser PR. Polymerase chain reaction for the detection of *Mycobacterium leprae*. *J Gen Microbiol*, 1989; **135:** 2357–65.
[23] Woods SA, Cole ST. A rapid method for the detection of potentially viable *Mycobacterium leprae* in human biopsies: a novel application of PCR. *FEMS Microbiol. Lett*, 1989; **65:** 305–10.
[24] Eiglmeier K, Honore N, Woods SA, Caudron B, Cole ST. Use of an ordered cosmid library to deduce the genomic organisation of *Mycobacterium leprae*. *Molecular Microbiology*, 1993; **7:** 197–206.
[25] Honore N, Cole ST. The molecular basis of rifampicin-resistance in *Mycobacterium leprae*. *Antimicrobial Agents & Chemotherapy*, 1993; **37:** 414–18.
[26] Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, Matter L, Schopfer K, Bodmer T. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *The Lancet*, 1993; **341:** 647–50.
[27] Culliton B. Drug-resistant TB may bring epidemic. *Nature*, 1992; **356:** 473.