9

Mycobacterial Genomes

David C. Alexander and Jun Liu

Summary

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains a major cause of death around the world. Diseases caused by nontuberculous mycobacteria are increasingly associated with immunocompromised individuals. The availability of whole-genome sequences of mycobacterial species in the past several years has revolutionized TB research. This chapter provides an overview of the biology of mycobacteria and the diseases that they cause, with emphasis on how recent advances in genomics have improved our knowledge of the lifestyle and phylogeny of these organisms.

Key Words: Mycobacterial genomes; *Mycobacterium tuberculosis*; Mycobacteria.

"If one judges of the importance of a disease according to its distribution, and according to the degree in which it menaces health and induces death prematurely, tuberculosis assumes the first rank in human pathology...it has been known as far back as the memory of man extends, and has unceasingly decimated the race for hundreds and thousands of years," (1)

Prof. Dr. Georg Cornet, 1904 (19th century Bacteriologist)

1. Introduction

One hundred years later, mycobacterial diseases retain their first rank as menaces to human and animal health. Despite global initiatives and five decades of chemotherapeutics, tuberculosis (TB) caused by *Mycobacterium tuberculosis* remains a common bacterial disease. An estimated 2 billion people are infected with *M. tuberculosis* and 2 million succumb to TB each year *(2)*. Leprosy, caused by *Mycobacterium leprae*, inflicts disfigurement and untold human suffering. Effective treatments for leprosy are available, but attempts at eradication have failed, and more than 600,000 new cases are reported each year *(3)*. Buruli ulcer, a deadly necrotizing skin disease caused by *Mycobacterium ulcerans*, is often described as an emerging infection, but in endemic regions it is more common than leprosy. Diseases caused by atypical, nontuberculous mycobacteria, such as members of the *Mycobacterium avium* complex (MAC), used to be rare. However, the epidemic of HIV infection and AIDS has been accompanied by a surge in these opportunistic mycobacterial infections, which are often difficult to treat. Animals also suffer. Vigilant farming practices have reduced the incidence of *Mycobacterium bovis* infection, but outbreaks of bovine TB still occur. Johne's disease, a fatal inflammatory bowel disease of livestock, caused by the *Mycobacterium avium* subspecies *paratuberculosis*, remains endemic in domestic herds *(4)*.

Historically, mycobacterial research has been hampered by the fastidious nutritional requirements, extraordinarily slow growth rates of these organisms and, especially with

From: *Bacterial Genomes and Infectious Diseases*
Edited by: V. L. Chan, P. M. Sherman, and B. Bourke © Humana Press Inc., Totowa, NJ

M. tuberculosis, the high risk of contagion. During the past 15 yr, the development of effective molecular biology tools has rendered these bacteria more amenable to genetic and biochemical studies. The availability of whole-genome sequences, starting with *M. tuberculosis* strain H37Rv in 1998 *(5)*, has revolutionized TB research. The genome sequences of *M. bovis (6)*, *M. leprae (7)*, and *M. tuberculosis* strain CDC1551 *(8)* have now been published. Additional sequencing projects are underway for the genomes of *M. ulcerans*, *Mycobacterium marinum*, *Mycobacterium microti*, *Mycobacterium smegmatis*, two *M. avium* subspecies, several strains of *M. tuberculosis* and the vaccine strain, Bacille Calmette-Guérin (BCG).

This chapter provides an overview of mycobacteria and the diseases that they cause, with an emphasis on how recent advances in genomics have enriched our understanding of both the biology and phylogeny of these organisms. The focus of the first section is TB and *M. tuberculosis*. As illustrated in the second section, the role of other mycobacteria should not be underestimated. The final section focuses on specific gene families and virulence factors that distinguish mycobacteria from other prokaryotes.

2. *Mycobacterium tuberculosis* and the Genus *Mycobacterium*

The genus *Mycobacterium* comprises more than 70 species *(9)*. A few, notably *M. tuberculosis*, *M. leprae*, and *M. ulcerans*, cause significant morbidity and mortality. Others, including *M. kansasii*, *M. fortuitum*, *M. abscessus*, *M. xenopi*, *M. chelonae*, and the *M. avium* complex, are responsible for occasionally lethal opportunistic infections *(10–12)*. However, the vast majority are harmless environmental organisms, common in water and soil *(13)*. Under the microscope, mycobacteria are small, rod-shaped bacteria. They are Gram-positive organisms, but are best distinguished by their characteristic acid-fast staining. This acid-fastness is a property of the mycobacterial cell wall, an unusual, lipid-rich structure that forms a hydrophobic, low permeability barrier and provides innate protection against many antimicrobial agents. Traditionally, mycobacteria have been classified according to growth rate and pigmentation (e.g., the Runyon Groups), and further subdivided on the basis of biochemical reactions (e.g., niacin production, nitrate reduction, drug resistance), serotypes, bacteriophage susceptibility, and cell wall lipid profiles. However, these have been superseded by molecular methods, especially DNA sequencing and polymerase chain reaction-based tests, which are rapid and require little starting material, both of which are important considerations when dealing with slow-growing and hazardous organisms. Sequencing of 16S ribosomal DNA (rDNA) and the 16S–23S rDNA internal transcribed spacer (ITS) region has been used to establish a phylogeny of *Mycobacterium* species. The latter has revealed that the ITS of fast-growing mycobacteria is longer and structurally distinct from the ITS of slow-growers, and supports the traditional distinction based on growth rate *(14)*. Sequencing allows discrimination between isolates that are phenotypically indistinguishable *(15)*, and has uncovered phylogenetic differences (i.e., sequevars) within individual species *(14)*.

3. *Mycobacterium tuberculosis* and the Global Impact of Tuberculosis

M. tuberculosis was first described in 1882 by the eminent microbiologist Robert Koch *(16)*, but it has been with us since antiquity. Known as tuberculosis, consumption, phthisis, and the white plague, evidence of *M. tuberculosis* disease has been found in

ancient manuscripts, sculptures, and wall paintings. In recent years, *M. tuberculosis* DNA has been extracted from mummies *(17,18)*. At the end of the 19th century, TB was the leading cause of death in the Western world, killing one in seven. Advances of the 20th century, including antibiotics, succeeded in almost eliminating the disease from Europe and the Americas. Despite a recent resurgence in Eastern Europe, the annual incidence of TB in most of the developed world remains low, with less than 50 cases per 100,000 people *(19)*. Globally, the situation is much worse. With more than 250 cases per 100,000 people, the incidence of TB is highest in Africa, but because of its larger population, the total number of cases is greater in Asia. The World Health Organization estimates that *M. tuberculosis* is responsible for 2 million deaths and 8.8 million new infections each year, 80% of which occur in developing countries *(20)*. One important impediment to global efforts to control TB is the ongoing HIV epidemic *(21,22)*. There is an overlap in the geographical distribution of these infectious diseases and, in the year 2000, an estimated 11% of new adult TB cases were also infected with HIV. At greatest risk for both *M. tuberculosis* and HIV infections are people in their prime working and reproductive years, between 14 and 49 yr of age. Widespread illness in this age group has profound social and economical effects. Without a workforce to maintain and fund the local medical, educational, and business infrastructure, the health of an entire community suffers.

The enormous cost of *M. tuberculosis* is matched by the complex lifestyle of this facultative intracellular pathogen *(23,24)*. It is transmitted between people, most often by an aerosol route. The cough of a tuberculous individual generates tiny nasal droplets, no larger than 5 µm in diameter, which contain live bacteria. New infections occur when these droplets are inhaled and penetrate to the alveoli of the respiratory tract, where *M. tuberculosis* is ingested by alveolar macrophages. Although macrophages typically destroy invading microbes, *M. tuberculosis* has the ability to subvert these phagocytic cells *(25,26)*. During the first 2 wk of infection, the bacteria slowly but continuously replicate inside of the macrophages. Two to eight weeks postinfection, cell-mediated immunity develops. At this stage, activated T-lymphocytes and noninfected macrophages act to control the growth of *M. tuberculosis*. In most individuals (~90%), the infection stops here. The immune response generates a granuloma around the *M. tuberculosis* that prevents the bacteria from spreading. With the infection contained, active disease does not develop. However, in some infected individuals, especially children under 5 yr and immunocompromised adults, the primary *M. tuberculosis* infection cannot be contained. The bacilli continue to replicate, host tissue is destroyed, and active TB develops. Although most often associated with the lungs (pulmonary TB), *M. tuberculosis* can attack anywhere in the body including the bones (Pott's disease), brain (TB meningitis), lymph nodes (scrofula), and intestinal tract.

M. tuberculosis is a tenacious pathogen. Even when the primary infection is contained, the bacteria within the granuloma can survive for decades, persisting in a special dormant state *(27,28)*. When the immune system is compromised, by such factors as malnutrition, HIV infection, diabetes, renal disease, chemotherapy, or extensive corticosteroid therapy, reactivation of the disease can occur *(29)*. The protective granuloma disintegrates, and the long dormant *M. tuberculosis* revives and spreads unchecked.

Most cases of TB will respond to antibiotics. Standard regimens involve daily treatment with four drugs: isoniazid, rifampin, pyrazinamide, and ethambutol, for 2 mo,

followed by 4 mo of isoniazid and rifampin. Failure to comply can result in a relapse and the emergence of multi-drug resistant strains, which no longer respond to these agents *(30–32)*. Effective treatment of multi-drug resistant strains requires quarantine of the patient and up to 24 mo of drug therapy. Atypical nontuberculous mycobacterial organisms, responsible for opportunistic infections in AIDS patients, tend to exhibit natural antibiotic resistance. Combined with the immunocompromised state of the host, these infections are, therefore, extraordinarily difficult to eradicate.

4. The Genome of *Mycobacterium tuberculosis*

The complete genome sequence of *M. tuberculosis* H37Rv (a virulent strain isolated in 1905 and then propagated in vitro) was published in 1998 *(5)*. The circular genome comprises of 4,411,532 bp and has a mean guanine and cytosine content of 65.6%. The original annotation identified 3974 genes encoding 3924 proteins and 50 stable RNA. Initially overlooked, an additional 82 protein encoding genes have since been added. Genes have been identified via (1) sequence homology to known proteins in other microorganisms, (2) experiments using two-dimensional electrophoresis and mass spectrometry, and (3) bioinformatic techniques that examine *M. tuberculosis* codon usage. The approx 4000 genes are classified into 11 broad functional groups. Of these, 52% are assigned with precise or putative functions, with the remaining 48% being conserved hypotheticals or functionally unknown genes.

The publication of the H37Rv genome was followed by completion of the *M. tuberculosis* strain CDC1551 genome sequence (isolated in 1995 and responsible for a outbreak in the United States) *(8,33)*, and the partial genome sequencing of *M. tuberculosis* strain 210 (a representative of the W/Beijing strains, responsible for the majority of cases in Asia and the former Soviet Union, as well as outbreaks in the United States) *(34,35)*. Analysis of these sequences reveals single nucleotide polymorphisms (SNPs, including both synonymous and nonsynonymous substitutions), large sequence polymorphisms (LSPs; genetic deletions/insertions), plus variations in the numbers and types of mobile elements (e.g., transposons and prophages) among *M. tuberculosis* isolates. For example, comparison of the whole genome sequences of H37Rv (4.41 Mbp) and CDC1551 (4.40 Mbp) revealed approx 1100 SNPs *(8)*. Approximately 65% of the SNPs are nonsynonymous substitutions, which is unusual because it is generally thought that many nonsynonymous substitutions are lost during purifying selection, as demonstrated in other bacteria, such as *Escherichia coli* and *Salmonella enterica*. LSP analysis indicates that *M. tuberculosis* exhibits less genetic diversity than other bacteria. Only 74 LSPs longer than 10 bp were identified between H37Rv and CDC1551 *(8)*. In a larger study of 100 clinical isolates, a total of 68 distinct deletion events, ranging in size from 105 bp to approx 12 kb, were identified *(36)*. Together, these LSPs represent approx 186 kb (~4.2%) of the H37Rv genome and affect 224 (~5.5%) genes, including genes in all major functional categories. However, among individual isolates no more than 41 kb or 50 genes were deleted. In contrast, differences between the K12 and O157 strains of *E. coli* affect more than 1300 genes *(37)*. Even so, these studies indicate that a degree of polymorphism does exist between different *M. tuberculosis* strains, which is consistent with the phenotypical diversity observed among clinical isolates *(38)*. Numerous methods, most commonly restriction fragment length profiles, spoligotypes, IS*6110* profiles, and SNP analysis are used to characterize clinical isolates *(39–41)*. These differences

provide insight into the epidemiology of outbreaks, infectivity, and virulence of individual strains. Notably, half of LSPs between H37Rv and CDC1551 involve genes encoding Pro–Pro–Glu (PPE) and Pro–Glu (PE)-polymorphic GC-rich repetitive sequence (PGRS) family proteins *(8)*, which are considered antigens important for human immunity (*see* **Subheading 11.2.**). Genetic studies also provide insight into the evolution of *M. tuberculosis*. Current evidence indicates that the W/Beijing strains, such as 210, are more ancestral than CDC1551 and H37Rv, and contain genes that are no longer present in the more recently derived strains *(38,42,43)*.

5. The *Mycobacterium tuberculosis* Complex

The *M. tuberculosis* complex refers to a number of genetically related human and animal pathogens that share 99.9% similarity at the nucleotide level and are indistinguishable by 16S rDNA sequencing. These include *M. tuberculosis*, *M. africanum*, *M. microti* (voles), *M. caprae* (goats), *M. bovis*, as well as the BCG vaccine strains and a variety of isolates from unusual sources, such as *M. pinnipedii* (from seal lions and fur seals), and the dassie bacillus (from *Procavia capensis*, the hyrax, or dassie). The animal strains are responsible for zoonotic transmission of TB to humans, especially via ingestion of infected meat or milk. Indeed, it was long believed that TB was an animal disease that managed to jump the species barrier. However, genetic interrogation of *M. tuberculosis* complex isolates, together with genome sequencing of *M. microti*, *M. bovis* AF2122/97 (a virulent strain from Great Britain) *(6)*, and BCG Pasteur (a vaccine strain maintained in Paris since 1923), reveals that the opposite is true. *M. tuberculosis* originated with humans and was transmitted to animals *(44)*. The current phylogeny indicates that the *M. tuberculosis* complex evolved from a human strain via successive, unidirectional deletion events (*see* Fig. 1).

M. africanum refers to human TB isolates from parts of Africa which, on the basis of biochemical tests, were considered distinct from both *M. tuberculosis* and *M. bovis*. However, this classification scheme has proven unreliable as biochemical markers do not correlate with genetic data, which identify at least three groups of *M. africanum* *(42,45)*. Some so-called *M. africanum* strains are genetically indistinguishable from *M. tuberculosis*. A second group contains a single genomic deletion, or region of difference (RD), called RD9, which affects seven genes. In the third group of *M. africanum*, RD9 is deleted along with the LSPs, RD7, RD8, and RD10. These four deletions are conserved across all animal strains. An additional five deletions are found in both *M. caprae* and *M. bovis*. Taken together, these RDs provide a scheme for the reductive evolution of the *M. tuberculosis* complex from human to cattle. Additional deletions define branches within the *M. tuberculosis* complex. For example, a small deletion, called TbD1, is a marker of "modern" TB *(42)*. Multiple genomic deletions distinguish the dassie bacillus from other species *(46)*. Additional variations are unique to specific host–pathogen pairs. The biological roles of individual RDs have yet to be established. Some may be hot spots for genetic rearrangement, but most are believed to influence disease transmission and progression. Consistent with both of these ideas, several overlapping, but nonidentical deletions have occurred in discrete isolates of the *M. tuberculosis* complex. For example, the RD1 deletion, which has occurred independently in *M. microti*, the dassie bacillus, and BCG impairs the secretion of key immunodominant antigens and impacts bacterial virulence. In contrast, the RD5 deletion, which affects five genes in

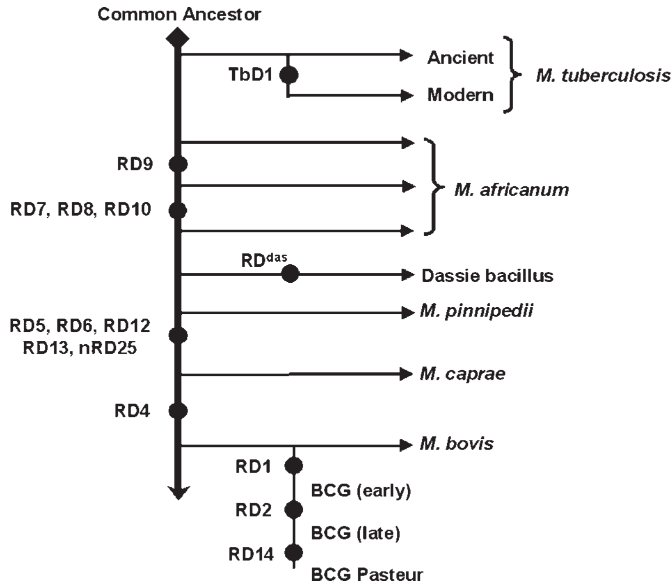

Fig. 1. Phylogenetic tree of the *Mycobacterium tuberculosis* complex. This phylogeny was generated using a variety of molecular markers, but only major genomic deletion events (filled circles) are indicated. Events on the thick vertical axis have accumulated over time and affect successive members of the complex. Events on the thin horizontal axes are only found in specific members of the complex. (Adapted from refs. *42,44–46,54,55,58*.)

M. microti, seven genes in the dassie bacillus, and eight genes in BCG Pasteur, is associated with transposition of the mobile genetic element, insertion sequence (IS)*6110 (44)*.

The role of SNPs in the evolution of the *M. tuberculosis* complex is less clear. The genome of *M. bovis* is 99.95% identical to that of *M. tuberculosis*. Fewer than 2500 SNPs have been identified and at least one-third are synonymous changes that do not alter amino acid sequence *(6)*. However, frameshifts and other variations do occur. A point mutation in *ald* (alanine dehydrogenase) prevents *M. bovis* and BCG strains from catabolizing the amino acid alanine *(47)*. *M. bovis*, although not BCG, also has a mutation in *pykA* (pyruvate kinase), such that glycolytic intermediates are not converted to pyruvate or used in the TCA cycle. Other SNPs alter the antigenic repertoire of *M. bovis*, and are likely to impact host–pathogen interactions *(6)*. Curiously, some genes that are defective in *M. bovis* are also psuedogenes in *M. leprae*.

5.1. Bacille Calmette-Guérin

Live attenuated vaccines have reduced the incidence, and even eliminated, many important bacterial and viral diseases. The original BCG was derived between 1908 and 1921 by 230 in vitro passages of an *M. bovis* strain. This attenuated BCG was found to be nonpathogenic in guinea pigs, yet sufficiently immunogenic to protect against a challenge with virulent *M. tuberculosis (48,49)*. In the preantibiotic era, the promise of an effective vaccine against TB made BCG popular and, starting in 1923, stocks were distributed around the world. BCG vaccination of newborns remains common in many countries with an estimated 100 million doses of BCG administered each year. BCG is safe and probably protects children against TB meningitis. However, in randomized

controlled trials the efficacy of BCG against adult pulmonary disease has ranged from 0 to 80%, and remains a matter of debate *(50,51)*. Efficacy may be diminished by prior exposure to environmental mycobacteria, which can alter the immune response to the vaccine *(52,53)*. However, BCG itself is also likely to blame. Between 1923 and the 1960s (when lyophilized stocks were finally established), BCG strains continued to be propagated in vitro. Different laboratories employed different culture media and passaged strains according to different schedules. BCG continued to adapt to laboratory conditions, plus there was selection for strains that produced few side effects, yet still promoted tuberculin conversion (wrongly considered an indicator of immunogenicity). The net result has been the creation of numerous, genetically heterogeneous BCG substrains. Although the original BCG of 1921 has been lost, genetic analysis of BCG substrains indicates that the initial attenuation event was the deletion of RD1 *(54,55)*. Indeed, experimental deletion of RD1 impairs the virulence of *M. tuberculosis* H37Rv in a mouse model of TB *(56,57)*. As with the *M. tuberculosis* complex, evolution of BCG substrains has involved multiple unidirectional deletion events. For example, the RD2 region is deleted from all substrains obtained from the Pasteur Institute after 1931, whereas nRD18 is only deleted in strains derived after 1933 *(54,58)*.

The need for an effective vaccine against *M. tuberculosis* remains. Although some research is directed at the generation of attenuated *M. bovis* or *M. microti* strains, DNA vaccines, and protein preparations, much of research continues to focus on improving the BCG strain *(59–61)*. Despite its shortcomings, BCG has an excellent safety record and, as a live persistent vaccine, it exhibits long lasting and complex immunostimulatory properties. Work with recombinant BCG, modified to produce immunogenic mycobacterial antigens, also shows promise against TB *(62,63)*. Recombinant BCG is also being used to fight other infections *(64,65)*. For example, BCG expressing an antigen from *Schistosoma mansoni* appears to protect mice from this helminthic infection *(66)*. In addition, BCG exhibits antitumor properties and is used as effective treatment for bladder cancer *(67)*.

6. *Mycobacterium leprae*

On February 28, 1873, G. H. Armauer Hansen identified bacilli in nodules removed from a patient with leprosy *(68)*. Previous epidemiological evidence suggested that leprosy was transmissible, but it was Hansen's finding that established leprosy as an infectious bacterial disease. Leprosy, often called Hansen's disease, affects millions and is endemic in India, Vietnam, and the Philippines, with approx 630,000 new infections occurring each year *(3)*. Despite its global importance, leprosy remains difficult to study. Animal models do not accurately emulate human disease, and the mechanisms of transmission are poorly defined. *M. leprae* is a slow growing obligate intracellular bacterium that has never been cultured in vitro. The complete genome sequence of *M. leprae* strain TN revealed that this organism has undergone massive gene decay; that is, an extreme form of the reductive evolution seen in the *M. tuberculosis* complex *(7)*. The genome is 3.3 Mbp, which is 1.1 Mbp smaller than that of *M. tuberculosis*. Of the estimated 3720 *M. leprae* open reading frames, more than 1100 are pseudogenes that no longer encode proteins. Even so, biosynthetic pathways for most molecules (e.g., amino acids, nucleotides) and cell wall components (e.g., peptidoglycan, arabinogalactan, mycolic acids, lipids) remain intact *(7,69)*. Conversely, many catabolic pathways,

transcriptional regulators, transport proteins, polyketide synthesis systems, detoxification enzymes, and DNA repair processes are impaired. These defects likely account for the fastidious growth requirements of *M. leprae*. Many putative virulence factors, such as the *mce* genes, PE/PPE genes, and all PE-PRGS genes are also absent or defective. As in some strains of BCG, the *mma3* gene, required for synthesis of methoxymycolates, is defective *(70)*. Similar to *M. avium* subsp. *paratuberculosis*, the mycobactin siderophore biosynthesis (*mbt*) operon is nonfunctional, suggesting that iron metabolism is impaired. At the same time, *M. leprae* has an extra NRAMP-like metal transporter, which may compensate for the *mbt* defect. Other genes not found in *M. tuberculosis*, and possibly beneficial to intracellular survival of *M. leprae*, include a uridine phosphorylase, a eukaryote-like adenylate cyclase, and a putative sugar transport system.

The chronology of gene decay in *M. leprae* has been a matter of some debate. Did the loss of regulatory genes precipitate or follow the loss of metabolic functions? The isolation of an ancestral *M. leprae*, with a more intact genome, could answer this question, but no such organism has been identified. Epidemiological studies of *M. leprae* are hampered by the lack of molecular typing methods. One contribution of genome sequence analysis has been the identification of polymorphic regions suitable for molecular typing *(71,72)*. Such characterization is integral to an improved understanding of disease transmission. Although several natural reservoirs have been suggested, including armadillos, insects, soil, and water, the sources for human infection remain a mystery.

7. The *Mycobacterium avium* Complex

MAC includes a variety of genetically related species with diverse pathogenic potential *(10)*. *M. avium* subsp. *avium* (*Maa*) is common in the environment. It causes avian tuberculosis and sporadic infections of wild mammals (e.g., deer), as well as opportunistic infections in immunocompromised humans. *M. avium* subsp. *silvaticum* (*Mas*), the so-called wood pigeon bacillus, is primarily a bird pathogen. *M. avium* subsp. *paratuberculosis* (*Map*) causes Johne's disease and, although the hypothesis remains controversial, has been implicated as a cause of Crohn's Disease, a chronic inflammatory bowel disease in humans *(73–75)*. MAC organisms exhibit greater heterogeneity than members of the *M. tuberculosis* complex. Multiple sequevars have been revealed by rDNA analysis and unidirectional deletion events cannot account for relationships between all isolates *(9,76)*. Different branches appear to have acquired new genetic material via horizontal transfer *(77)*. Genome comparison of *Maa* strain 104 (a human pulmonary isolate) and *Map* strain K10 (a Johne's disease isolate from a cow) has emphasized variations both within MAC, and between MAC and the *M. tuberculosis* complex.

Maa strain 104 (5.4 Mbp) and *Map* strain K10 (4.5 Mbp) have larger genomes than *M. tuberculosis* and encode several hundred more genes. Orthologs to many *M. tuberculosis* genes exist, but there are some notable differences. For example, *Maa* encodes one-third as many PE/PPE genes as *M. tuberculosis*. Although the function of these repetitive proteins is unknown, they are thought to contribute to the antigenic diversity of mycobacteria. The RD1 region, considered important for virulence, is missing from MAC. At the same time, *Maa* possesses more transcriptional regulatory genes and dedicates a larger portion of its genome to lipid metabolism. Extra genes, especially those involved in transcriptional regulation or associated with cell wall functions, likely help *Maa* adapt to volatile environmental conditions. In general, the genomic differ-

ences between MAC and the *M. tuberculosis* complex seem to reflect their disparate lifestyles. *M. tuberculosis* is an obligate pathogen with a relatively stable intracellular niche and little exposure to other bacteria. In contrast, MAC are environmental organisms, which must contend with changing environmental conditions and have greater opportunity for horizontal gene transfer via interactions with disparate bacteria and phages. *Map* is capable of infecting cattle, but is also capable of survival for months in a barnyard or field *(78)*. *Maa* infects birds, thrives in drinking water and hot tubs, and is an endosymbiont of free-living protists *(79)*. Indeed, it has been suggested that the ability of *Maa* to colonize phagocytic amoebae and protozoans prefigures that hallmark of mycobacterial infections: the invasion and subversion of macrophages *(80)*.

Maa is considered to be the ancestral form of MAC, but the enormous diversity between *Maa* isolates has precluded a robust analysis of phylogeny. Isolates from diverse sources (e.g., a deer in New Zealand and a bird from the Netherlands) may be indistinguishable, yet samples from similar sources (immunocompromised people living in the same city) can differ at multiple sites and exhibit discrepancies in genome size of greater than 250 kb *(81)*.

Traditionally, ISs have been used to define species. Isolates containing IS*900* are called *Map (82)*. IS*900* negative isolates are called *Maa* when positive for IS*1245*, and *Mas* if positive for IS*901*. However, inconsistencies exist, and some nonMAC organisms may contain these mobile elements. New, genome-based approaches likely will reveal the evolutionary history of MAC *(81)*. Such methods have already contributed to the phylogeny of *Map (83)*. Long suspected on the basis of culture characteristics, it is now clear that *Map* can be subdivided into bovine (cattle) and ovine (sheep) branches *(84,85)*. However, it remains unclear if these genetic differences are related to idiosyncrasies of the host–pathogen interaction, or if they represent some geographical bias. Identification of markers restricted to individual MAC subspecies is key to the development of sensitive and specific diagnostic tests *(86)*. Reliable tests are not yet available for MAC diseases, but would be valuable for the efficient detection and treatment of conditions, such as Johne's disease *(4)*.

8. *Mycobacterium ulcerans* and *Mycobacterium marinum*

Buruli ulcer is a devastating skin disease. The ulcers are difficult to treat and can consume as much as 70% of the skin surface before causing death *(87)*. First described in the scientific literature in 1948, after an outbreak in an Australian resort town *(88)*, historical evidence suggests the disease has long been endemic in Africa. It takes its name from the Buruli region of Uganda *(89)*. The causative agent, *M. ulcerans*, is among the slowest growing mycobacteria, with an in vitro generation time of more than 30 h. It is associated with wetlands. Contaminated water and water-borne insect larvae are implicated in the infection cycle. The related organism, *M. marinum*, is a lethal pathogen of fish, amphibians, and reptiles *(90)*. In humans, *M. marinum* is responsible for skin diseases, such as swimming-pool granuloma and fish-fancier's finger. As the names suggest, this opportunistic infection also results from contact with contaminated water. However, *M. marinum* infections are rarely lethal.

The genome of *M. ulcerans* has yet to be completely sequenced, but its associated plasmid, pMUM1, has been deciphered *(91)*. This 174-kb plasmid contains genes for the biosynthesis of mycolactone, a ketolide with immunosuppressive properties.

Mycolactone is thought to be the virulence factor responsible for the Buruli ulcer, and the key difference between *M. ulcerans* and *M. marinum*. A thorough comparison of these organisms with *M. tuberculosis* has yet to be completed. The genomes of *M. marinum* and *M. ulcerans* are in the 6 Mbp range, and likely encode numerous genes required for their aquatic lifestyles. Even so, rDNA sequencing suggests that *M. marinum* and *M. ulcerans* are closely related to the *M. tuberculosis* complex. Large regions of synteny are present, and virulence-associated genes, including the RD1 region, are known to be intact *(92)*.

M. marinum has found favor as an experimental model of mycobacterial infections because it grows more quickly and is safer to work with than *M. tuberculosis (93–96)*. More importantly, *M. marinum* can be used to study genuine host–pathogen interactions. Mice are widely used to study tuberculosis, but *M. tuberculosis* is not a mouse pathogen and the mouse model, although useful, does not accurately reflect human disease. In contrast, *M. marinum* naturally infects both fish and frogs. The combination of *M. marinum* and zebrafish may prove to be a particularly useful model. Numerous genetic tools are available for their study and, like humans, zebrafish exhibit both innate and adaptive immune responses.

9. *Mycobacterium smegmatis*

Mycobacterium smegmatis was once believed to cause syphilis. It is now recognized as a harmless saprophyte, common in soil. Fast growing, it has served as a model for the study of mycobacteria *(97–99)*. Sequencing of the *M. smegmatis* mc^{2}155 laboratory strain reveals that, like the other environmental mycobacteria, it has a large genome. It also shares many of the physiological characteristics particular to the mycobacteria. However, *M. smegmatis* is quite different from the pathogenic mycobacteria and is unable to survive in macrophages *(100)*. Its genome shows little synteny with *M. tuberculosis* and many putative virulence genes are absent.

10. Mycobacteriophages

The first mycobacteriophage was described in 1947 *(101)*. Since then, several hundred have been isolated and every mycobacterial genome sequence contains at least one prophage. Phages are likely a key mediator of diversity in the MAC complex. It is not known if prophage gene products contribute to mycobacterial virulence, but considering the importance of prophage-derived toxins in other actinomycetes (e.g., the diphtheria and tetanus toxins of *Corynebacterium* spp.) their role in disease pathogenesis would not be surprising.

Mycobacteriophages have been widely employed as diagnostic tools. Phage-based strain typing has now been superseded by other methods, but luminescent reporter phages and phage replication assays still are used for rapid detection of mycobacteria in clinical and environmental samples, and to determine antibiotic resistance *(102–106)*. Phage-based systems also are used for genetic manipulation of mycobacteria, including allelic replacement and transposon-delivery *(107–111)*. Although little is known about individual phages, sequencing projects, such as those conducted at the Pittsburgh Bacteriophage Institute have revealed great diversity in both genome size and gene content *(112–117)*.

11. Mycobacterial Gene Families

Analysis of the genome sequences of different *Mycobacterium* species indicates that genome size may vary and individual genes may lack orthologs, but most functions are conserved. These include basic metabolic activities common to all prokaryotes, such as DNA replication, transcription, cell division, and small molecule biosynthesis, along with some mycobacterial-specific functions. Perhaps most importantly, several large gene families that were either previously unknown or poorly understood have been identified through the genome projects. Analysis of *Mycobacterium* genomes indicates that there is a core set of approx 200 highly conserved genes encoding mycobacterial-specific functions *(118)*. Half of these are associated with cell wall biosynthesis. Others fall into a "conserved hypothetical" category for which roles remain to be determined. Several are PE/PPE genes, and a few are classified as regulatory genes or virulence determinants, such as the *mce* genes. Not included in this core set are antigens of the 6 kDa early secretory antigenic target (ESAT)-6 and 10 kDa culture filtrate protein (CFP)-10 family, as individual loci are not perfectly conserved, and similar proteins are found in other actinomycetes. Likewise, the resuscitation promoting factors have homologs in other organisms.

11.1. Mycobacterial Lipid Metabolism

Genome sequencing revealed that a large number of *M. tuberculosis* genes (~250) encode enzymes involved in lipid metabolism. In contrast, *E. coli*, which has a similar genome size as *M. tuberculosis*, contains approx 50 enzymes involved in lipid metabolism. As mentioned earlier, the lipid-rich cell wall is a defining characteristic of mycobacteria (Fig. 2) *(119)*. The excess of lipid metabolic enzymes in *M. tuberculosis* correlate with the unusual chemical composition of the structure. The wall consists of three covalently linked polymers: peptidoglycan, arabinogalactan, and mycolic acid (Fig. 2) together with a variety of complex lipids, including lipoglycans (e.g., lipoarabinomannan, lipomannan, and the related phosphatidylinositol mannosides), glycopeptidolipids, sulfolipids, trehalose-containing glycolipids, phthiocerol dimycocerosates, phenolic glycolipids, and triacylglycerols *(120,121)*. The proportion of these lipids varies from species to species and is also affected by changing growth conditions *(122)*.

The lipid domain of mycobacterial cell wall forms an asymmetric bilayer. The outer leaflet of this bilayer contains various surface glycolipids, whereas the inner leaflet is composed exclusively of mycolic acids. The mycolic acid layer displays exceptionally low fluidity and low permeability. It is this barrier that is responsible for the natural resistance of mycobacteria to many antimicrobial agents, including antibiotics and host immune factors *(121,123,124)*. Important roles also have been suggested for the surface glycolipids, especially the multiple methyl-branched fatty acids: sulfolipids, phthiocerol dimycocerosates, phenolic glycolipids, diacylated trehaloses, and polyacylated trehaloses *(125)*. Genes for phthiocerol and phenolphthiocerol dimycocerosate synthesis are present on a *M. tuberculosis* pathogenicity island (PAI). The latter lipid is necessary for the growth of *M. tuberculosis* in the lungs of infected mice *(126)*. Another component of interest is the lipoglycan, lipoarabinomannan. The *M. tuberculosis* version of this lipid exerts immunomodulatory effects, including the downregulation of cell-mediated immunity *(127)*. Curiously, lipoarabinomannan from *M. chelonae* has

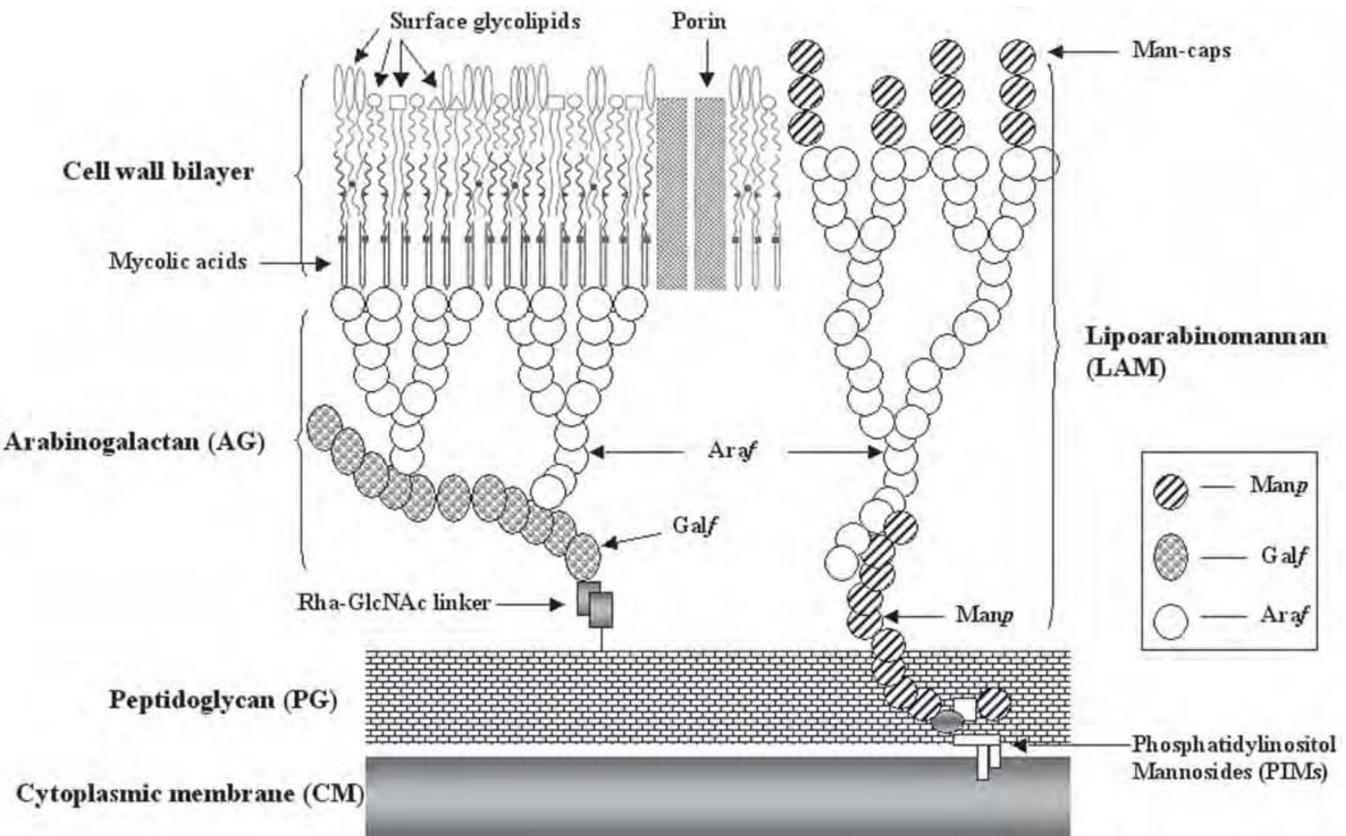

Fig. 2. Cell envelope of *Mycobacterium*. The cell wall core consists of three covalently linked polymers: peptidoglycan, arabinogalactan, and mycolic acids. The cell wall lipid domain forms an asymmetric lipid bilayer with mycolic acids constituting exclusively the inner leaflet and the extractable glycolipids occupying the outer leaflet of the bilayer. Lipoarabinomannan is thought to be associated with the cell wall through a phosphatidylinositol anchor to the cytoplasmic membrane.

no such ability. The difference appears to stem from species-specific variations in the structure of lipoarabinomannan *(127)*.

Homology searches to genes of known function in other bacteria have assisted in the identification of mycobacterial genes involved in cell wall lipid synthesis. Speculations on the role of individual genes can be tested by mutagenesis of the target gene and biochemical analysis of purified gene products. Such insight into the genetics and enzymology of cell wall biosynthesis and assembly makes it possible to identify cellular targets for the development of new drugs. Several front-line antitubercular drugs, such as isoniazid, ethambutol, and pyrazinamide, are now known to target lipid biosynthesis and cell wall assembly.

11.2. The PE and PPE Gene Families

The novel PE and PPE gene families are also highly abundant, accounting for 9% of all genes in *M. tuberculosis*. The proteins are distinguished by the eponymous PE or PPE motifs at N-terminal residues 8 and 9, or 8 to 10, respectively. Each family is further subdivided on the basis of characteristic C-terminal motifs. For example, all PE proteins have a conserved N-terminal domain (~110 amino acids). One subgroup includes short proteins with no C-terminal region. Members of the second group, PE-PGRS, have large C-terminal domains that contain multiple (sometimes hundreds) of tandem repeats of the glycine-rich motif, Gly–Gly–X (often Gly–Gly–Ala). The remaining PE proteins have substantial (100–400 amino acids) C-terminal domains, but the repeated Gly–Gly–X motif is absent. Analysis suggests that some proteins in this third group share an approx 225 amino acid motif in which the primary sequence is degenerate, but the secondary structure is conserved *(128)*.

The role of the PE/PPE genes has been the source of much speculation. One prominent idea is that they contribute to antigenic variation among strains of *M. tuberculosis*, and influence the host immune response. Comparative analysis of the PE-PGRS proteins of *M. tuberculosis* strains H37Rv and CDC1551 revealed variations resulting from frame-shift mutations, as well as in-frame insertions and deletions. Usually, the PE domains were unaffected and just the PGRS domains differed between strains. Size variations also were seen in clinical samples of *M. tuberculosis* by Western blot using PE-PGRS-specific antibody. In addition, some PE-PGRS proteins are surface exposed *(129)*. Others are antigenic and recognized by sera obtained from TB patients and those vaccinated with BCG *(130,131)*. However, their enzymatic functions, if any, are unknown and the importance of the conserved and variable regions is a mystery.

Additional evidence for the role of the PE/PPE genes in pathogenesis stems from the recent identification of three PAIs conserved among *M. tuberculosis* H37Rv, CDC1551, and *M. bovis (132)*. Prominent in PAI2 and PAI3 are PPE and PE-PGRS family genes. In vivo studies also indicate that these genes are important. For instance, when *M. marinum* resides in host granulomas or macrophages, two PE-PGRS genes are preferentially expressed *(95)*. Furthermore, disruption of PE/PPE family genes can lead to growth attenuation in the mouse model of TB *(133)*.

11.3. ESAT-6/CFP-10 Antigens

Exponentially growing *M. tuberculosis* secretes numerous proteins into the surrounding media. Two of these, ESAT-6 and CFP-10, are immunodominant antigens *(134,135)*.

However, neither protein is produced by BCG. The genes encoding ESAT-6 (*esx*A/Rv3875) and CFP-10 (*esx*B/Rv3874) belong to the RD1 region that is absent from all BCG vaccine strains. Independent deletion events also resulted in the loss of these genes from *M. microti* and the dassie bacillus *(46)*. ESAT-6 and CFP-10 are important for pathogenicity. Experimental deletion of RD1 from *M. tuberculosis* results in a diminution of virulence, whereas restoration of this region to *M. microti* enhances its virulence *(56,57, 136)*. The two proteins form a heterodimeric complex *(137)*. CFP-10 most likely acts as a chaperone. ESAT-6 has cytolytic properties and contributes to the cell-to-cell spread of mycobacteria within the host *(92,136)*. The antigens are coexpressed with genes encoding a novel type of secretion system. Work with both *M. tuberculosis* and *M. marinum* indicates that, even if the *esx*A and *esx*B genes are intact, disruption of this secretory apparatus blocks export of ESAT-6 and CFP-10, such that cytolysis does not occur *(92)*. Interaction of CFP-10, but not ESAT-6, with components of the secretion system has been demonstrated *(138)*. This entire gene cluster is conserved in *M. marinum*, *M. smegmatis*, and probably *M. kansasii* and *M. szulgai*. A homologous region is present in *M. leprae*, but contains several pseudogenes In contrast, the cluster is completely absent from *M. avium (137,139)*.

Ten paralogous gene pairs (*esx*C/*esx*D to *esx*V/*esx*W) are annotated in the *M. tuberculosis* H37Rv genome. Although expected to form heterodimeric complexes similar to ESAT-6/CFP-10, their roles have yet to be established. Several of these pairs are associated with their own versions of the novel secretion system. These clusters are variably present across mycobacterial genomes, but at least one, *esx*G/*esx*H and its associated secretion apparatus, is conserved between *M. tuberculosis*, *M. avium*, *M. smegmatis*, and *M. leprae (139)*. Related systems are present in other actinomycetes, including *Corynebacterium* and *Streptomyces*, but their importance is currently unknown *(139)*.

ESAT-6 and CFP-10 are being used in the development of new diagnostics and new vaccines. Skin testing with PPD (the purified protein derivative of *M. tuberculosis*) is a sensitive method for the diagnosis of TB. Unfortunately, it lacks specificity because false-positive reactions commonly occur among BCG vaccinated individuals. In contrast, ESAT-6 and CFP-10 are not produced by BCG. The two immunodominant antigens are recognized by sera from the majority of TB patients, but not sera from people vaccinated with BCG. As such, ESAT-6 and CFP-10 can effectively discriminate between vaccinated and infected individuals *(140)*.

Vaccine development has employed two approaches toward the antigens. The first involves deletion of *esx*A and *esx*B from *M. bovis* and *M. tuberculosis* to produce a live attenuated vaccine *(141,142)*. The opposite strategy has been to use ESAT-6 and CFP-10 for vaccination. They have been used individually, in protein cocktails, and expressed in either recombinant BCG or alternate carrier strains, such as attenuated *Salmonella (63,143–146)*. Both approaches show promise, but a deletion strain would permit continued use of the antigens as diagnostic tools.

11.4. Mammalian Cell Entry

The mammalian cell entry (MCE) proteins are putative virulence factors present in diverse mycobacteria *(147)*. Their name stems from the finding that a cloned fragment of the *M. tuberculosis* H37Rv *mce1* region permits nonpathogenic *E. coli* cells to invade

cultured HeLa cells *(148)*. The canonical *mce* region is a polycistronic operon of eight genes following the order *yrbAB mceABCDEF*. However, among the four *mce* operons in *M. tuberculosis* and nine in *Maa*, some variations exist. The *M. tuberculosis mce2* operon contains two extra genes, one upstream of *yrb2A* and a second between *mce2B* and *mce2C (149)*. Consistent with a putative role in invasion, these proteins are predicted to be cell wall-associated and/or secreted. *In silico* modeling of Mce1A suggests its structure resembles that of Colicin N, a pore-forming bacterial toxin *(150)*. This attractive, albeit hypothetical, model predicts that binding of Mce1A to its cognate receptor is followed by a conformational change of the protein and perforation of the target cell membrane. The putative receptor-binding surface of the *in silico* structure corresponds to a known MceA1 epitope, recognized by a monoclonal antibody. This epitope is not conserved in MceA2, MceA3, or MceA4, implying that the proteins bind to different receptors and possibly target different cells. In agreement with this model, a truncated form of recombinant MceA1 protein, which includes the putative receptor-binding domain, promotes the uptake of latex beads by HeLa cells, whereas recombinant Mce2A protein does not *(151)*. Although the roles of individual *mce* operons have not been determined, gene expression profiling indicates that they are differentially transcribed. In synthetic media, *mce1* is strongly expressed by exponentially growing *M. tuberculosis*, but only weakly by stationary phase cells. The inverse pattern is found with *mce4*; strong expression in stationary phase but none during exponential growth. The *mce* genes are also transcribed in vivo. *M. tuberculosis* from the spleens of experimentally infected Guinea pigs were found to express *mce4*. Bacilli from the lungs of infected rabbits express *mce1*, *mce3*, and *mce4 (149)*. Appropriate spatial and temporal expression of the *mce* operons may be important events in the infectious process. Primary mouse macrophages infected with a *mce1*-defective *M. tuberculosis* strain have an altered cytokine profile and may be unable to stimulate T-cell immunity. Mice infected with the same mutant *mce1* strain develop histologically aberrant lung granulomas and are unable to control bacterial replication *(152)*. However, the impact of *mce* defects likely varies among different mycobacterial species and in different hosts. Homologs of the *mce3* operon have been deleted from both *Map* and *M. bovis*, yet these organisms remain effective pathogens in cattle.

11.5. Resuscitation Promoting Factors

The resuscitation promoting factor (Rpf) was first described in another actinomycete, *Micrococcus luteus (153)*. Following long-term culture, these cells lose viability and are unable to form colonies when plated onto fresh solid medium. However, the senescent cells are not dead, just dormant and nonculturable. Viability and colony formation by *Micrococcus* are restored by the addition of Rpf, a small protein secreted by actively growing cells. As mentioned earlier, dormancy is a feature of the bacteria persisting in TB granulomas. Moreover, *M. tuberculosis*, BCG, and *M. smegmatis* all respond to exogenously added Rpf *(154,155)*. Genome sequencing has revealed five *rpf* genes in *M. tuberculosis* and a similar number in MAC. *M. smegmatis* also produces Rpf.

The *M. tuberculosis rpf* genes have been systematically deleted *(156)*. Individually, the genes are neither essential for growth in vitro nor for in vivo infection of mice. This suggests that they have redundant or overlapping functions, although strains with mul-

tiple *rpf* deletions have yet to be tested. Analysis indicates that all *M. tuberculosis rpf* genes are transcribed during logarithmic growth. Expression patterns vary at intermediate time points, but all transcripts are also detected in 4-mo-old cultures. During the acute phase of infection, *rpf* expression is also present in *M. tuberculosis* from mouse lungs *(156)*. Additional work on the immunological impact of Rpfs is required, but they have attracted some attention as vaccine candidates *(157)*.

12. Concluding Remarks

Mycobacterial research is experiencing a renaissance. The wealth of genome sequence data, new molecular tools, plus bioinformatic, proteomic, structural, and functional genomic approaches hold substantial promise for understanding the biology of these unusual microbes, elucidating the molecular mechanisms of pathogenesis, and developing new chemotherapeutic agents and effective vaccines. These approaches ultimately should lead to the effective control of mycobacterial disease and end the scourge of TB.

References

1. Cornet, G. (1904) *Tuberculosis and Acute General Miliary Tuberculosis*. W.B. Sauders and Co., Philadelphia.
2. World Health Organization. Stop TB Annual Report. (2001) World Health Organization, Geneva, Switzerland.
3. World Health Organization (2002) Leprosy. Global situation. *Wkly. Epidemiol. Rec.* **77,** 1–8.
4. Cocito, C., Gilot, P., Coene, M., de Kesel, M., Poupart, P., and Vannuffel, P. (1994) Paratuberculosis. *Clin. Microbiol. Rev.* **7,** 328–345.
5. Cole, S. T., Brosch, R., Parkhill, J., et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393,** 537–544.
6. Garnier, T., Eiglmeier, K., Camus, J. C., et al. (2003) The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. USA* **100,** 7877–7882.
7. Cole, S. T., Eiglmeier, K., Parkhill, J., et al. (2001) Massive gene decay in the leprosy bacillus. *Nature* **409,** 1007–1011.
8. Fleischmann, R. D., Alland, D., Eisen, J. A., et al. (2002) Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* **184,** 5479–5490.
9. Harmsen, D., Dostal, S., Roth, A., et al. (2003) RIDOM: comprehensive and public sequence database for identification of *Mycobacterium* species. *BMC. Infect. Dis.* **3,** 26.
10. Primm, T. P., Lucero, C. A., and Falkinham, J. O. (2004) Health impacts of environmental mycobacteria. *Clin. Microbiol. Rev.* **17,** 98–106.
11. Wayne, L. G. and Sramek, H. A. (1992) Agents of newly recognized or infrequently encountered mycobacterial diseases. *Clin. Microbiol. Rev.* **5,** 1–25.
12. Wolinsky, E. (1992) Mycobacterial diseases other than tuberculosis. *Clin. Infect. Dis.* **15,** 1–10.
13. Tortoli, E. (2003) Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. *Clin. Microbiol. Rev.* **16,** 319–354.
14. Roth, A., Fischer, M., Hamid, M. E., Michalke, S., Ludwig, W., and Mauch, H. (1998) Differentiation of phylogenetically related slowly growing mycobacteria based on 16S-23S rRNA gene internal transcribed spacer sequences. *J. Clin. Microbiol.* **36,** 139–147.
15. Cloud, J. L., Neal, H., Rosenberry, R., et al. (2002) Identification of *Mycobacterium* spp. by using a commercial 16S ribosomal DNA sequencing kit and additional sequencing libraries. *J. Clin. Microbiol.* **40,** 400–406.

16. Koch, R. (1882) Die Aetiologie der Tuberkulose. *Berliner Klinischen Wochenschrift.* **15,** 221–230.
17. Salo, W. L., Aufderheide, A. C., Buikstra, J., and Holcomb, T. A. (1994) Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc. Natl. Acad. Sci. USA* **91,** 2091–2094.
18. Zink, A. R., Sola, C., Reischl, U., et al. (2003) Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *J. Clin. Microbiol.* **41,** 359–367.
19. Dye, C., Scheele, S., Dolin, P., Pathania, V., and Raviglione, M. C. (1999) Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *JAMA* **282,** 677–686.
20. World Health Organization. Global Tuberculosis Control: Surveillance, Planning, Financing. WHO Report 2004. (2004) World Health Organization, Geneva, Switzerland.
21. Gazzard, B. (2001) Tuberculosis, HIV and the developing world. *Clin. Med.* **1,** 62–68.
22. Porter, J. D. (1996) Mycobacteriosis and HIV infection: the new public health challenge. *J. Antimicrob. Chemother.* **37,** 113–120.
23. Cosma, C. L., Sherman, D. R., and Ramakrishnan, L. (2003) The secret lives of the pathogenic mycobacteria. *Annu. Rev. Microbiol.* **57,** 641–676.
24. Smith, I. (2003) *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* **16,** 463–496.
25. Amer, A. O. and Swanson, M. S. (2002) A phagosome of one's own: a microbial guide to life in the macrophage. *Curr. Opin. Microbiol.* **5,** 56–61.
26. Deretic, V. and Fratti, R. A. (1999) *Mycobacterium tuberculosis* phagosome. *Mol. Microbiol.* **31,** 1603–1609.
27. Stewart, G. R., Robertson, B. D., and Young, D. B. (2003) Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* **1,** 97–105.
28. Wayne, L. G. (1994) Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur. J. Clin. Microbiol. Infect. Dis.* **13,** 908–914.
29. Lawn, S. D., Butera, S. T., and Shinnick, T. M. (2002) Tuberculosis unleashed: the impact of human immunodeficiency virus infection on the host granulomatous response to *Mycobacterium tuberculosis*. *Microbes. Infect.* **4,** 635–646.
30. Espinal, M. A. (2003) The global situation of MDR-TB. *Tuberculosis* **83,** 44–51.
31. Mukherjee, J. S., Rich, M. L., Socci, A. R., et al. (2004) Programmes and principles in treatment of multidrug-resistant tuberculosis. *Lancet* **363,** 474–481.
32. Nachega, J. B. and Chaisson, R. E. (2003) Tuberculosis drug resistance: a global threat. *Clin. Infect. Dis.* **36,** S24–S30.
33. Valway, S. E., Sanchez, M. P., Shinnick, T. F., et al. (1998) An outbreak involving extensive transmission of a virulent strain of *Mycobacterium tuberculosis*. *N. Engl. J. Med.* **338,** 633–639.
34. Glynn, J. R., Whiteley, J., Bifani, P. J., Kremer, K., and van Soolingen, D. (2002) Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg. Infect. Dis.* **8,** 843–849.
35. van Soolingen, D., Qian, L., de Haas, P. E., et al. (1995) Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. *J. Clin. Microbiol.* **33,** 3234–3238.
36. Tsolaki, A. G., Hirsh, A. E., DeRiemer, K., et al. (2004) Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc. Natl. Acad. Sci. USA* **101,** 4865–4870.
37. Perna, N. T., Plunkett, G., Burland, V., et al. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409,** 529–533.

38. Sreevatsan, S., Pan, X., Stockbauer, K. E., et al. (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. USA* **94,** 9869–9874.
39. Barnes, P. F. and Cave, M. D. (2003) Molecular epidemiology of tuberculosis. *N. Engl. J. Med.* **349,** 1149–1156.
40. van Soolingen, D. (2001) Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J. Intern. Med.* **249,** 1–26.
41. Mostrom, P., Gordon, M., Sola, C., Ridell, M., and Rastogi, N. (2002) Methods used in the molecular epidemiology of tuberculosis. *Clin. Microbiol. Infect.* **8,** 694–704.
42. Brosch, R., Gordon, S. V., Marmiesse, M., et al. (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. USA* **99,** 3684–3689.
43. Gutacker, M. M., Smoot, J. C., Migliaccio, C. A., et al. (2002) Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains. *Genetics* **162,** 1533–1543.
44. Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A., and Behr, M. A. (2002) Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. *J. Infect. Dis.* **186,** 74–80.
45. Mostowy, S., Onipede, A., Gagneux, S., et al. (2004) Genomic analysis distinguishes *Mycobacterium africanum*. *J. Clin. Microbiol.* **42,** 3594–3599.
46. Mostowy, S., Cousins, D., and Behr, M. A. (2004) Genomic interrogation of the dassie bacillus reveals it as a unique RD1 mutant within the *Mycobacterium tuberculosis* complex. *J. Bacteriol.* **186,** 104–109.
47. Chen, J. M., Alexander, D. C., Behr, M. A., and Liu, J. (2003) *Mycobacterium bovis* BCG vaccines exhibit defects in alanine and serine catabolism. *Infect. Immun.* **71,** 708–716.
48. Behr, M. A. and Small, P. M. (1999) A historical and molecular phylogeny of BCG strains. *Vaccine* **17,** 915–922.
49. Crispen, R. (1989) History of BCG and its substrains. *Prog. Clin. Biol. Res.* **310,** 35–50
50. Behr, M. A. and Small, P. M. (1997) Has BCG attenuated to impotence? *Nature* **389,** 133–134.
51. Brewer, T. F. and Colditz, G. A. (1995) Relationship between bacille Calmette-Guerin (BCG) strains and the efficacy of BCG vaccine in the prevention of tuberculosis. *Clin. Infect. Dis.* **20,** 126–135.
52. Brandt, L., Feino, C. J., Weinreich, O. A., et al. (2002) Failure of the *Mycobacterium bovis* BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis. *Infect. Immun.* **70,** 672–678.
53. Buddle, B. M., Wards, B. J., Aldwell, F. E., Collins, D. M., and de Lisle, G. W. (2002) Influence of sensitisation to environmental mycobacteria on subsequent vaccination against bovine tuberculosis. *Vaccine* **20,** 1126–1133.
54. Behr, M. A., Wilson, M. A., Gill, W. P., et al. (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284,** 1520–1523.
55. Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C., and Stover, C. K. (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J. Bacteriol.* **178,** 1274–1282.
56. Lewis, K. N., Liao, R., Guinn, K. M., et al. (2003) Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. *J. Infect. Dis.* **187,** 117–123.
57. Pym, A. S., Brodin, P., Brosch, R., Huerre, M., and Cole, S. T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol. Microbiol.* **46,** 709–717.

58. Mostowy, S., Tsolaki, A. G., Small, P. M., and Behr, M. A. (2003) The in vitro evolution of BCG vaccines. *Vaccine* **21,** 4270–4274.
59. Doherty, T. M. and Andersen, P. (2002) Tuberculosis vaccine development. *Curr. Opin. Pulm. Med.* **8,** 183–187.
60. Kumar, H., Malhotra, D., Goswami, S., and Bamezai, R. N. (2003) How far have we reached in tuberculosis vaccine development? *Crit. Rev. Microbiol.* **29,** 297–312.
61. Young, D. B. and Stewart, G. R. (2002) Tuberculosis vaccines. *Br. Med. Bull.* **62,** 73–86.
62. Horwitz, M. A. and Harth, G. (2003) A new vaccine against tuberculosis affords greater survival after challenge than the current vaccine in the guinea pig model of pulmonary tuberculosis. *Infect. Immun.* **71,** 1672–1679.
63. Pym, A. S., Brodin, P., Majlessi, L., et al. (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat. Med.* **9,** 533–539.
64. Stover, C. K., Bansal, G. P., Hanson, M. S., et al. (1993) Protective immunity elicited by recombinant bacille Calmette-Guerin (BCG) expressing outer surface protein A (OspA) lipoprotein: a candidate Lyme disease vaccine. *J. Exp. Med.* **178,** 197–209.
65. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., et al. (1991) New use of BCG for recombinant vaccines. *Nature* **351,** 456–460.
66. Varaldo, P. B., Leite, L. C., Dias, W. O., et al. (2004) Recombinant *Mycobacterium bovis* BCG expressing the Sm14 antigen of *Schistosoma mansoni* protects mice from cercarial challenge. *Infect. Immun.* **72,** 3336–3343.
67. Shelley, M. D., Court, J. B., Kynaston, H., et al. (2004) Intravesical Bacillus Calmette-Guerin in Ta and T1 Bladder Cancer. In: *The Cochrane Library* **4,** John Wiley and Sons, Ltd, Chichester, UK.
68. Hansen, G. H. A. (1875) On the etiology of leprosy. *Br. J. Foreign Med. Chir. Rev.* **55,** 459–489.
69. Brennan, P. J. and Vissa, V. D. (2001) Genomic evidence for the retention of the essential mycobacterial cell wall in the otherwise defective *Mycobacterium leprae*. *Lepr. Rev.* **72,** 415–428.
70. Behr, M. A., Schroeder, B. G., Brinkman, J. N., Slayden, R. A., and Barry, C. E. (2000) A point mutation in the *mma*3 gene is responsible for impaired methoxymycolic acid production in *Mycobacterium bovis* BCG strains obtained after 1927. *J. Bacteriol.* **182,** 3394–3399.
71. Groathouse, N. A., Rivoire, B., Kim, H., et al. (2004) Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. *J. Clin. Microbiol.* **42,** 1666–1672.
72. Shin, Y. C., Lee, H., Walsh, G. P., Kim, J. D., and Cho, S. N. (2000) Variable numbers of TTC repeats in *Mycobacterium leprae* DNA from leprosy patients and use in strain differentiation. *J. Clin. Microbiol.* **38,** 4535–4538.
73. Chacon, O., Bermudez, L. E., and Barletta, R. G. (2004) Johne’s disease, inflammatory bowel disease, and *Mycobacterium paratuberculosis*. *Annu. Rev. Microbiol.* **58,** 329–363.
74. Greenstein, R. J. (2003) Is Crohn’s disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne’s disease. *Lancet Infect. Dis.* **3,** 507–514.
75. Hermon-Taylor, J. and Bull, T. (2002) Crohn’s disease caused by *Mycobacterium avium* subspecies *paratuberculosis*: a public health tragedy whose resolution is long overdue. *J. Med. Microbiol.* **51,** 3–6.
76. Novi, C., Rindi, L., Lari, N., and Garzelli, C. (2000) Molecular typing of *Mycobacterium avium* isolates by sequencing of the 16S-23S rDNA internal transcribed spacer and comparison with *IS1245*-based fingerprinting. *J. Med. Microbiol.* **49,** 1091–1095.
77. Krzywinska, E., Krzywinski, J., and Schorey, J. S. (2004) Naturally occurring horizontal gene transfer and homologous recombination in *Mycobacterium*. *Microbiology* **150,** 1707–1712.

78. Whittington, R. J., Marshall, D. J., Nicholls, P. J., Marsh, I. B., and Reddacliff, L. A. (2004) Survival and dormancy of *Mycobacterium avium* subsp. *paratuberculosis* in the environment. *Appl. Environ. Microbiol.* **70,** 2989–3004.
79. Falkinham, J. O., Norton, C. D., and LeChevallier, M. W. (2001) Factors influencing numbers of *Mycobacterium avium*, *Mycobacterium intracellulare*, and other Mycobacteria in drinking water distribution systems. *Appl. Environ. Microbiol.* **67,** 1225–1231.
80. Skriwan, C., Fajardo, M., Hagele, S., et al. (2002) Various bacterial pathogens and symbionts infect the amoeba *Dictyostelium discoideum*. *Int. J. Med. Microbiol.* **291,** 615–624.
81. Semret, M., Zhai, G., Mostowy, S., et al. (2004) Extensive genomic polymorphism within *Mycobacterium avium*. *J. Bacteriol.* **186,** 6332–6334.
82. Bull, T. J., Hermon-Taylor, J., Pavlik, I., El-Zaatari, F., and Tizard, M. (2000) Characterization of IS*900* loci in *Mycobacterium avium* subsp. *paratuberculosis* and development of multiplex PCR typing. *Microbiology* **146,** 2185–2197.
83. Amonsin, A., Li, L. L., Zhang, Q., et al. (2004) Multilocus short sequence repeat sequencing approach for differentiating among *Mycobacterium avium* subsp. *paratuberculosis* strains. *J. Clin. Microbiol.* **42,** 1694–1702.
84. Dohmann, K., Strommenger, B., Stevenson, K., et al. (2003) Characterization of genetic differences between *Mycobacterium avium* subsp. *paratuberculosis* type I and type II isolates. *J. Clin. Microbiol.* **41,** 5215–5223.
85. Motiwala, A. S., Strother, M., Amonsin, A., et al. (2003) Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: evidence for limited strain diversity, strain sharing, and identification of unique targets for diagnosis. *J. Clin. Microbiol.* **41,** 2015–2026.
86. Bannantine, J. P., Hansen, J. K., Paustian, M. L., et al. (2004) Expression and immunogenicity of proteins encoded by sequences specific to *Mycobacterium avium* subsp. *paratuberculosis*. *J. Clin. Microbiol.* **42,** 106–114.
87. van der Werf, T. S., Stinear, T., Stienstra, Y., van der Graaf, W. T., and Small, P. L. (2003) Mycolactones and *Mycobacterium ulcerans* disease. *Lancet* **362,** 1062–1064.
88. MacCallum, P., Tolhurst, J. C., Buckle, G., and Sissons, H. A. (1948) A new mycobacterial infection in man. *J. Pathol. Bacteriol.* **60,** 93–122.
89. Clancey, J. K. (1964) Mycobacterial skin ulcers in Uganda: description of a new mycobacterium (Mycobacterium Buruli). *J. Pathol. Bacteriol.* **88,** 175–187.
90. Decostere, A., Hermans, K., and Haesebrouck, F. (2004) Piscine mycobacteriosis: a literature review covering the agent and the disease it causes in fish and humans. *Vet. Microbiol.* **99,** 159–166.
91. Stinear, T. P., Mve-Obiang, A., Small, P. L., et al. (2004) Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc. Natl. Acad. Sci. USA* **101,** 1345–1349.
92. Gao, L. Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J. N., and Brown, E. J. (2004) A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol. Microbiol.* **53,** 1677–1693.
93. Chan, K., Knaak, T., Satkamp, L., Humbert, O., Falkow, S., and Ramakrishnan, L. (2002) Complex pattern of *Mycobacterium marinum* gene expression during long-term granulomatous infection. *Proc. Natl. Acad. Sci. USA* **99,** 3920–3925.
94. Ramakrishnan, L. and Falkow, S. (1994) *Mycobacterium marinum* persists in cultured mammalian cells in a temperature-restricted fashion. *Infect. Immun.* **62,** 3222–3229.
95. Ramakrishnan, L., Federspiel, N. A., and Falkow, S. (2000) Granuloma-specific expression of *Mycobacterium* virulence proteins from the glycine-rich PE-PGRS family. *Science* **288,** 1436–1439.

96. Ramakrishnan, L., Valdivia, R. H., McKerrow, J. H., and Falkow, S. (1997) *Mycobacterium marinum* causes both long-term subclinical infection and acute disease in the leopard frog (*Rana pipiens*). *Infect. Immun.* **65,** 767–773.
97. Andrew, P. W. and Roberts, I. S. (1993) Construction of a bioluminescent mycobacterium and its use for assay of antimycobacterial agents. *J. Clin. Microbiol.* **31,** 2251–2254.
98. Mayuri, Bagchi, G., Das, T. K., and Tyagi, J. S. (2002) Molecular analysis of the dormancy response in *Mycobacterium smegmatis*: expression analysis of genes encoding the DevR-DevS two-component system, Rv3134c and chaperone alpha-crystallin homologues. *FEMS Microbiol. Lett.* **211,** 231–237.
99. Triccas, J. A., Parish, T., Britton, W. J., and Gicquel, B. (1998) An inducible expression system permitting the efficient purification of a recombinant antigen from *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* **167,** 151–156.
100. Wei, J., Dahl, J. L., Moulder, J. W., et al. (2000) Identification of a *Mycobacterium tuberculosis* gene that enhances mycobacterial survival in macrophages. *J. Bacteriol.* **182,** 377–384.
101. Gardner, G. M. and Weiser, R. S. (1947) A bacteriophage for *Mycobacterium smegmatis. Proc. Soc. Exp. Biol. Med.* **66,** 205–206.
102. Bardarov, S. J., Dou, H., Eisenach, K., et al. (2003) Detection and drug-susceptibility testing of *M. tuberculosis* from sputum samples using luciferase reporter phage: comparison with the Mycobacteria Growth Indicator Tube (MGIT) system. *Diagn. Microbiol. Infect. Dis.* **45,** 53–61.
103. Carriere, C., Riska, P. F., Zimhony, O., et al. (1997) Conditionally replicating luciferase reporter phages: improved sensitivity for rapid detection and assessment of drug susceptibility of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **35,** 3232–3239.
104. Hazbon, M. H., Guarin, N., Ferro, B. E., et al. (2003) Photographic and luminometric detection of luciferase reporter phages for drug susceptibility testing of clinical *Mycobacterium tuberculosis* isolates. *J. Clin. Microbiol.* **41,** 4865–4869.
105. Riska, P. F. and Jacobs, W. R. Jr. (1998) The use of luciferase-reporter phage for antibiotic-susceptibility testing of mycobacteria. *Methods Mol. Biol.* **101,** 431–455.
106. Riska, P. F., Su, Y., Bardarov, S., et al. (1999) Rapid film-based determination of antibiotic susceptibilities of *Mycobacterium tuberculosis* strains by using a luciferase reporter phage and the Bronx Box. *J. Clin. Microbiol.* **37,** 1144–1149.
107. Bardarov, S., Bardarov, J. S. J., Pavelka, J. M. J., et al. (2002) Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology* **148,** 3007–3017.
108. Bardarov, S., Kriakov, J., Carriere, C., et al. (1997) Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **94,** 10,961–10,966.
109. Jacobs, W. R. Jr., Snapper, S. B., Tuckman, M., and Bloom, B. R. (1989) Mycobacteriophage vector systems. *Rev. Infect. Dis.* **11,** S404–S410.
110. Pearson, R. E., Jurgensen, S., Sarkis, G. J., Hatfull, G. F., and Jacobs, W. R. J. (1996) Construction of D29 shuttle phasmids and luciferase reporter phages for detection of mycobacteria. *Gene* **183,** 129–136.
111. Jacobs, W. R. J., Tuckman, M., and Bloom, B. R. (1987) Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature* **327,** 532–535.
112. Pedulla, M. L., Ford, M. E., Houtz, J. M., et al. (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113,** 171–182.
113. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. USA* **96,** 2192–2197.

114. Hatfull, G. F. and Sarkis, G. J. (1993) DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Mol. Microbiol.* **7,** 395–405.
115. Ford, M. E., Sarkis, G. J., Belanger, A. E., Hendrix, R. W., and Hatfull, G. F. (1998) Genome structure of mycobacteriophage D29: implications for phage evolution. *J. Mol. Biol.* **279,** 143–164.
116. Ford, M. E., Stenstrom, C., Hendrix, R. W., and Hatfull, G. F. (1998) Mycobacteriophage TM4: genome structure and gene expression. *Tuber. Lung Dis.* **79,** 63–73.
117. Mediavilla, J., Jain, S., Kriakov, J., et al. (2000) Genome organization and characterization of mycobacteriophage Bxb1. *Mol. Microbiol.* **38,** 955–970.
118. Marmiesse, M., Brodin, P., Buchrieser, C., et al. (2004) Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex. *Microbiology* **150,** 483–496.
119. Brennan, P. J. (2003) Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis* **83,** 91–97.
120. Brennan, P. J. and Nikaido, H. (1995) The envelope of mycobacteria. *Ann. Rev. Biochem.* **64,** 29–63.
121. Liu, J. and Nikaido, H. (1999) A mutant in *Mycobacterium smegmatis* defective in the biosynthesis of mycolic acids accumulates meromycolates. *Proc. Natl. Acad. Sci. USA* **96,** 4011–4016.
122. Kolattukudy, P. E., Fernandes, N. D., Azad, A. K., Fitzmaurice, A. M., and Sirakova, T. D. (1997) Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. *Mol. Microbiol.* **24,** 263–270.
123. Liu, J., Barry, C. E., Besra, G. S., and Nikaido, H. (1996) Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *J. Biol. Chem.* **271,** 29,545–29,551.
124. Liu, J., Rosenberg, E. Y., and Nikaido, H. (1995) Fluidity of the lipid domain of cell wall from *Mycobacterium chelonae*. *Proc. Natl. Acad. Sci. USA* **92,** 11,254–11,258.
125. Minnikin, D. E., Kremer, L., Dover, L. G., and Besra, G. S. (2002) The methyl-branched fortifications of *Mycobacterium tuberculosis. Chem. Biol.* **9,** 545–553.
126. Cox, J. S., Chen, B., McNeil, M., and Jacobs, W. R. J. (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402,** 79–83.
127. Nigou, J., Gilleron, M., and Puzo, G. (2003) Lipoarabinomannans: from structure to biosynthesis. *Biochimie* **85,** 153–166.
128. Adindla, S. and Guruprasad, L. (2003) Sequence analysis corresponding to the PPE and PE proteins in *Mycobacterium tuberculosis* and other genomes. *J. Biosci.* **28,** 169–179.
129. Sampson, S. L., Lukey, P., Warren, R. M., van Helden, P. D., Richardson, M., and Everett, M. J. (2001) Expression, characterization and subcellular localization of the *Mycobacterium tuberculosis* PPE gene Rv1917c. *Tuberculosis* **81,** 305–317.
130. Choudhary, R. K., Mukhopadhyay, S., Chakhaiyar, P., et al. (2003) PPE antigen Rv2430c of *Mycobacterium tuberculosis* induces a strong B-cell response. *Infect. Immun.* **71,** 6338–6343.
131. Okkels, L. M., Brock, I., Follmann, F., et al. (2003) PPE protein (Rv3873) from DNA segment RD1 of *Mycobacterium tuberculosis*: strong recognition of both specific T-cell epitopes and epitopes conserved within the PPE family. *Infect. Immun.* **71,** 6116–6123.
132. Zubrzycki, I. Z. (2004) Analysis of the products of genes encompassed by the theoretically predicted pathogenicity islands of *Mycobacterium tuberculosis* and *Mycobacterium bovis*. *Proteins* **54,** 563–568.
133. Sassetti, C. M. and Rubin, E. J. (2003) Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. USA* **100,** 12,989–12,994.

134. Andersen, P., Andersen, A. B., Sorensen, A. L., and Nagai, S. (1995) Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice. *J. Immunol.* **154,** 3359–3372.
135. Berthet, F. X., Rasmussen, P. B., Rosenkrands, I., Andersen, P., and Gicquel, B. (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology* **144,** 3195–3203.
136. Hsu, T., Hingley-Wilson, S. M., Chen, B., et al. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc. Natl. Acad. Sci. USA* **100,** 12,420–12,425.
137. Sorensen, A. L., Nagai, S., Houen, G., Andersen, P., and Andersen, A. B. (1995) Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. *Infect. Immun.* **63,** 1710–1717.
138. Okkels, L. M. and Andersen, P. (2004) Protein-protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis*. *J. Bacteriol.* **186,** 2487–2491.
139. Gey, V. P. N., Gamieldien, J., Hide, W., Brown, G. D., Siezen, R. J., and Beyers, A. D. (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biol.* **2,** 1–18.
140. van Pinxteren, L. A., Ravn, P., Agger, E. M., Pollock, J., and Andersen, P. (2000) Diagnosis of tuberculosis based on the two specific antigens ESAT-6 and CFP10. *Clin. Diagn. Lab. Immunol.* **7,** 155–160.
141. Collins, D. M., Kawakami, R. P., Wards, B. J., Campbell, S., and de Lisle, G. W. (2003) Vaccine and skin testing properties of two avirulent *Mycobacterium bovis* mutants with and without an additional *esat-6* mutation. *Tuberculosis* **83,** 361–366.
142. Wards, B. J., de Lisle, G. W., and Collins, D. M. (2000) An *esat*6 knockout mutant of *Mycobacterium bovis* produced by homologous recombination will contribute to the development of a live tuberculosis vaccine. *Tuber. Lung Dis.* **80,** 185–189.
143. Brandt, L., Elhay, M., Rosenkrands, I., Lindblad, E. B., and Andersen, P. (2000) ESAT-6 subunit vaccination against *Mycobacterium tuberculosis*. *Infect. Immun.* **68,** 791–795.
144. Mollenkopf, H. J., Groine-Triebkorn, D., Andersen, P., Hess, J., and Kaufmann, S. H. (2001) Protective efficacy against tuberculosis of ESAT-6 secreted by a live *Salmonella typhimurium* vaccine carrier strain and expressed by naked DNA. *Vaccine* **19,** 4028–4035.
145. Mustafa, A. S. and Al-Attiyah, R. (2003) Tuberculosis: looking beyond BCG vaccines. *J. Postgrad. Med.* **49,** 134–140.
146. Olsen, A. W., Hansen, P. R., Holm, A., and Andersen, P. (2000) Efficient protection against *Mycobacterium tuberculosis* by vaccination with a single subdominant epitope from the ESAT-6 antigen. *Eur. J. Immunol.* **30,** 1724–1732.
147. Haile, Y., Caugant, D. A., Bjune, G., and Wiker, H. G. (2002) *Mycobacterium tuberculosis* mammalian cell entry operon (*mce*) homologs in *Mycobacterium* other than tuberculosis (MOTT). *FEMS Immunol. Med. Microbiol.* **33,** 125–132.
148. Arruda, S., Bomfim, G., Knights, R., Huima-Byron, T., and Riley, L. W. (1993) Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. *Science* **261,** 1454–1457.
149. Kumar, A., Bose, M., and Brahmachari, V. (2003) Analysis of expression profile of mammalian cell entry (*mce*) operons of *Mycobacterium tuberculosis*. *Infect. Immun.* **71,** 6083–6087.
150. Das, A. K., Mitra, D., Harboe, M., et al. (2003) Predicted molecular structure of the mammalian cell entry protein Mce1A of *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* **302,** 442–447.
151. Chitale, S., Ehrt, S., Kawamura, I., et al. (2001) Recombinant *Mycobacterium tuberculosis* protein associated with mammalian cell entry. *Cell Microbiol.* **3,** 247–254.

152. Shimono, N., Morici, L., Casali, N., et al. (2003) Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the *mce1* operon. *Proc. Natl. Acad. Sci. USA* **100,** 15,918–15,923.
153. Mukamolova, G. V., Kaprelyants, A. S., Young, D. I., Young, M., and Kell, D. B. (1998) A bacterial cytokine. *Proc. Natl. Acad. Sci. USA* **95,** 8916–8921.
154. Mukamolova, G. V., Turapov, O. A., Young, D. I., Kaprelyants, A. S., Kell, D. B., and Young, M. (2002) A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **46,** 623–635.
155. Shleeva, M., Mukamolova, G. V., Young, M., Williams, H. D., and Kaprelyants, A. S. (2004) Formation of 'non-culturable' cells of *Mycobacterium smegmatis* in stationary phase in response to growth under suboptimal conditions and their Rpf-mediated resuscitation. *Microbiology* **150,** 1687–1697.
156. Tufariello, J. M., Jacobs, W. R. Jr., and Chan, J. (2004) Individual *Mycobacterium tuberculosis r*esuscitation-promoting factor homologues are dispensable for growth in vitro and in vivo. *Infect. Immun.* **72,** 515–526.
157. Yeremeev, V. V., Kondratieva, T. K., Rubakova, E. I., et al. (2003) Proteins of the Rpf family: immune cell reactivity and vaccination efficacy against tuberculosis in mice. *Infect. Immun.* **71,** 4789–4794.