

4

Biology of waterborne pathogenic mycobacteria

G. Cangelosi, J. Clark-Curtiss, M. Behr, T. Bull and T. Stinear

4.1 INTRODUCTION

4.1.1 Taxonomy and terminology

The waterborne mycobacteria are members of a large and very significant family of human pathogens. *Mycobacterium* is the single genus in the family Mycobacteriaceae, order Actinomycetales. Over 70 *Mycobacterium* species have been defined, at least 30 of which cause disease in humans or animals. *Mycobacterium* species fall into two groups: the slow growers and the rapid growers. This descriptive division predates modern genotype-based taxonomic methods; however, it has been shown to be consistent with genotypic taxonomy (Rogall *et al.* 1990). Most pathogenic mycobacteria are slow growers, the most notable exception being the rapid-growing *M. fortuitum* complex. Mycobacteria

© 2004 World Health Organization. *Pathogenic Mycobacteria in Water: A Guide to Public Health Consequences, Monitoring and Management*. Edited by S. Pedley, J. Bartram, G. Rees, A. Dufour and J. Cotruvo. ISBN: 1 84339 059 0. Published by IWA Publishing, London, UK.

are easily identifiable by the mycolic acids in their cell envelopes. Rich in other lipids as well, mycobacterial cell envelopes can account for up to 40% of total cell weight. These structures are responsible for the ability of the bacteria to resist decolorization by weak acids after staining, hence the diagnostic term “acid-fast bacilli”.

The *M. tuberculosis* complex, a cluster of closely related species that includes the causative agent of TB, ranks with HIV and malaria parasites as among the most significant microbial pathogens of humans. The genus also includes *M. leprae*, the causative agent of Hansen disease (leprosy). Neither *M. tuberculosis* nor *M. leprae* are encountered in water. A variety of other *Mycobacterium* species, some of which are very common in water, occasionally infect humans. When isolated from patients’ samples, these organisms are sometimes referred to as “atypical” or “NTM”; however, the term “environmental mycobacteria” is more informative.

From the standpoint of human health, the most significant of the environmental mycobacteria are the MAC and *M. ulcerans*. MAC is considered to include MAA, MAP, *M. avium* subspecies *silvaticum* and *M. intracellulare*. These divisions are well defined based on phenotypic as well as genotypic (rRNA) criteria (Boddinghaus *et al.* 1990; Frothingham & Wilson 1994). Recently, a proposal was made to confine the designation MAA to a cluster of closely related strains that are commonly isolated from birds and rarely from humans (avian strains) and assign a new name, *M. avium* subspecies *hominissuis*, to the more diverse group of strains that are more commonly isolated from mammalian (including human) environments (Mijs *et al.* 2002). This proposal was based upon genotypic and phenotypic analysis of a large number of strains. However, at the time of this writing it is still a proposal, not a universally accepted standard. Therefore, the standard terminology is used here.

Many human and environmental isolates of MAC do not fall neatly within any of the designations listed above. Such isolates can exhibit genotypic features of two or more different MAC species, and are usually given designations such as “MAC_x”. The existence of such intermediate types may help explain the common view of MAC as a continuum of more or less indistinguishable species and subspecies. However, as emphasized elsewhere in this book, the pathogens we know as MAA, MAP and *M. intracellulare* are distinct in many significant ways, including environmental niches, host preferences and clinical manifestations. Recognition of these distinctions is important to the understanding of the biology of MAC.

M. ulcerans is the causative agent of Buruli ulcer, a debilitating disease characterized by large necrotic skin ulcers that is currently widespread throughout West and Central Africa (refer to Chapter 8). *M. ulcerans* is closely related to the fish pathogen *M. marinum*. Analysis of their full 16SrRNA

sequences show > 99.8% identity and comparisons of other gene sequences show a similar high level of sequence conservation (Stinear *et al.* 2000b). Yet despite this high level of genetic relatedness *M. ulcerans* and *M. marinum* are phenotypically distinct. Notably, *M. ulcerans* is slow growing with a doubling time of 36 hours *in vitro*, whereas *M. marinum* is a robust environmental organism that doubles every 6 hours *in vitro* and is readily isolated from diverse aquatic environments. It is extremely difficult to isolate *M. ulcerans* from the environment, although the pathogen may occupy niche environments such as the salivary glands of particular aquatic insects (Masollier *et al.* 2002). *M. marinum* causes a tuberculoid disease in fish and other poikilotherms and a relatively minor skin infection in humans. As with other mycobacterial pathogens, *M. marinum* can replicate within the host macrophage and provokes the formation of a granuloma by the host. In contrast, *M. ulcerans* is not known to cause disease in fish and in humans it produces large necrotic skin lesions caused by massive necrosis of subcutaneous fat. Histopathology shows a marked absence of a host inflammatory immune response and massive numbers of bacilli are found extracellularly. This unusual pathology has been linked to the presence of a macrolide toxin produced by *M. ulcerans* called mycolactone (George *et al.* 1999). *M. marinum* does not produce mycolactone.

In addition to the use of rRNA-based systematics, the species and subspecies of environmental mycobacteria have also been classified by a number of phenotypic characteristics including Runyon Group, pigment production, serotype and host range. However, genotypic classifications based upon rRNA structural gene and spacer sequences, species-specific insertion elements and other genetic markers described in later sections, are considered by most to be more reliable markers. They are becoming more practical with the increasingly widespread availability of commercial DNA probes (Accuprobe, Genprobe, Inc., San Diego, California), PCR and DNA sequencing capabilities.

4.1.2 Evolution and diversity of MAC

Epidemiological evidence indicates that humans are infected by MAC that is living in the environment, not via person-to-person transmission. Although viable MAC cells can occasionally be cultured from the stools of AIDS patients, there is little evidence to suggest that infective populations of the bacteria commonly make their way back to the environment. Therefore, the populations of MAA and *M. intracellulare* cells to which humans are exposed are not likely to have human environments in their recent evolutionary histories. Their pathogenic capabilities were selected by environmental challenges encountered outside of human hosts. Unfortunately, very little is known about the lives of MAC cells in the environment.

It is possible that extracellular environments select for traits in MAC that are coincidentally advantageous in human infections. For example, resistance to environmental toxins might also confer protection from the killing effects of antibiotics and host immune cells. Alternatively, the bacteria may routinely encounter and defend themselves against the phagocytic cells of animal hosts. MAP is well known to infect livestock while MAA and *M. intracellulare* infect a variety of mammals and birds in serovar-specific fashion. Certain MAA strains are commonly excreted in the faeces of birds after which they can persist in the soil for extended periods of time. Although there is little evidence for direct transmission from animals to humans a few serovars and strain types can be recovered from both, consistent with the possibility that virulence mechanisms are maintained in animal reservoirs (Mijs *et al.* 2002; Inderlied *et al.* 1993).

An alternative model is that environmental mycobacteria commonly reside within environmental amoebae, where they encounter an environment similar to that found in human phagocytic cells (Cirillo *et al.* 1997; Steinert *et al.* 1998; Miltner & Bermudez 2000). This is supported by laboratory studies which have shown that MAA cells grown within *Acanthamoeba* are more virulent than extracellularly grown MAA. However, it is not known how often this happens in the environment. Nonhuman reservoirs must be explored if we are to understand how environmental mycobacteria infect humans. Such studies may also lead to improved nonhuman disease models, and refined methods for detecting harmful MAC populations in drinking-water.

Compared to closely related groups such as the *M. tuberculosis* complex, MAC is diverse phenotypically as well as phylogenetically. This diversity is evident between species (e.g. MAA and *M. intracellulare*), between individual isolates within a species and even within a single clinical isolate. Most clinical isolates of MAC form multiple colony morphotypes that vary with regard to infectivity, susceptibility to antibiotics and ability to survive in various environments. MAC also exhibits considerable genetic polymorphism, mediated in part by mobile insertion elements that are abundant in their genomes (Arbeit *et al.* 1993; von Reyn *et al.* 1995; Eckstein *et al.* 2000; Matsiota-Bernard *et al.* 2000; Laurent *et al.* 2002).

The heterogeneity of MAC has important implications with regard to epidemiology and risk assessment. It is possible that some strains found in the environment are especially infectious to humans, while others may be relatively harmless. This would compromise the predictive value of environmental monitoring efforts that assume uniform levels of infectivity. Therefore, a significant priority for MAC research is the identification of genetic or phenotypic markers that can distinguish the most infective strains from those

that are relatively harmless to humans. Such markers could help refine epidemiological analysis and lead to more accurate methods of risk assessment.

4.2 MYCOBACTERIAL GENOMES

Our understanding of the biology of *M. tuberculosis* and *M. leprae* improved dramatically with the determination of their genomic nucleotide sequences. Further rapid progress is being made with studies on genome-wide patterns of gene expression and gene function. Our understanding of MAC and similar environmental pathogens is on the cusp of similar progress, with the impending completion of the genomic nucleotide sequences of MAA, MAP, *M. ulcerans*, and *M. marinum*.

4.2.1 The MAA genome

The genome sequence of MAA strain 104, a clinical isolate originating from an AIDS patient in California, is nearing completion by The Institute for Genomic Research (<http://www.tigr.org>). The sequence has a high G+C content (68.5%) and is about 5.48 mb long. It has been annotated by M. Behr and co-workers at McGill University, Montreal, Canada (Semret *et al.* submitted). Approximately 4480 coding sequences were identified that are likely to be genes. In contrast, the genome of *M. tuberculosis* is about 4.4 mb long and has 3959 likely genes (Cole *et al.* 1998; <http://genolist.pasteur.fr/TubercuList/>). Approximately 385 MAA genes have no counterpart in *M. tuberculosis*. Presumably, some of these MAA-specific genes confer the ability to live and grow in the environment, which *M. tuberculosis* lacks. Others code for unique cell surface properties that are discussed below.

When using genome sequence information on MAC it is important to bear in mind that the organism's heterogeneity is such that there is probably no truly representative strain. The choice of MAA strain 104 for genome sequencing was made in the mid-1990s based on several criteria. As an AIDS isolate it was representative of the majority of MAC disease cases at the time. It is virulent in nonhuman disease models and intrinsically multi-drug resistant. It is a stable "red transparent" morphotypic clone, which is somewhat unusual in a virulent clinical isolate (Mukherjee *et al.* 2001). Its relative stability would have simplified the process of sequence determination and assembly at the possible cost of using a strain that is less typical than an isolate that exhibits normal morphotypic switching.

As with *M. tuberculosis* the genome of MAA contains a large number of mobile genetic elements termed IS. These elements are usually small (1 to 2 kb), with little or no genetic information other than that required for transposition

(hopping) from site to site within the genome. The variable patterns of IS insertion can be used to generate strain-specific “bar codes” for molecular epidemiological analysis. These codes are relatively stable in *M. tuberculosis*, but in MAA they exhibit considerable diversity even within a single clinical isolate (Laurent *et al.* 2002). IS are thought to be important drivers of genetic diversification in mycobacteria. When an IS hops into a gene that gene is usually disrupted and its function is lost. Moreover, genetic recombination between neighbouring IS can result in large-scale deletions and rearrangements of genomic DNA. This phenomenon has been shown to play an important role in the generation of spontaneous “rough” colony type variants of MAA (Eckstein *et al.* 2000).

The positions of insertion elements IS1245 and IS999 (Laurent *et al.* 2002) within the MAA strain 104 genome are shown in Figure 4.1. IS999 insertion sites in three other strains are also shown. Regions of the genome with high IS concentrations are likely to contain DNA sequences that are not required for life in at least some environments. An example of such a region is the ssGPL gene cluster which is deleted in rough colony type variants. These IS-rich regions are potential hot spots for loss or rearrangement over the course of MAA evolution.

A microarray of the MAA genome has been prepared by M. Behr and co-workers (Semret *et al.*, submitted manuscript). A microarray is a wafer dotted with small nucleotide probes, each specific for an individual MAC gene. Hybridization of a DNA or RNA sample to the array, followed by sophisticated bioinformatic analysis, reveals the presence or absence of genes or gene products in the sample. This tool will allow us to rapidly assess and quantify the genomic diversity found among large numbers of clinical and environmental isolates. Comparison of gene deletion patterns by “chromosome alignment” may lead to the identification of genomic signatures characteristic of virulent strains, a useful outcome for risk assessment. Microarray analysis of conditional gene expression (RNA samples) during intracellular growth might also reveal virulence genes.

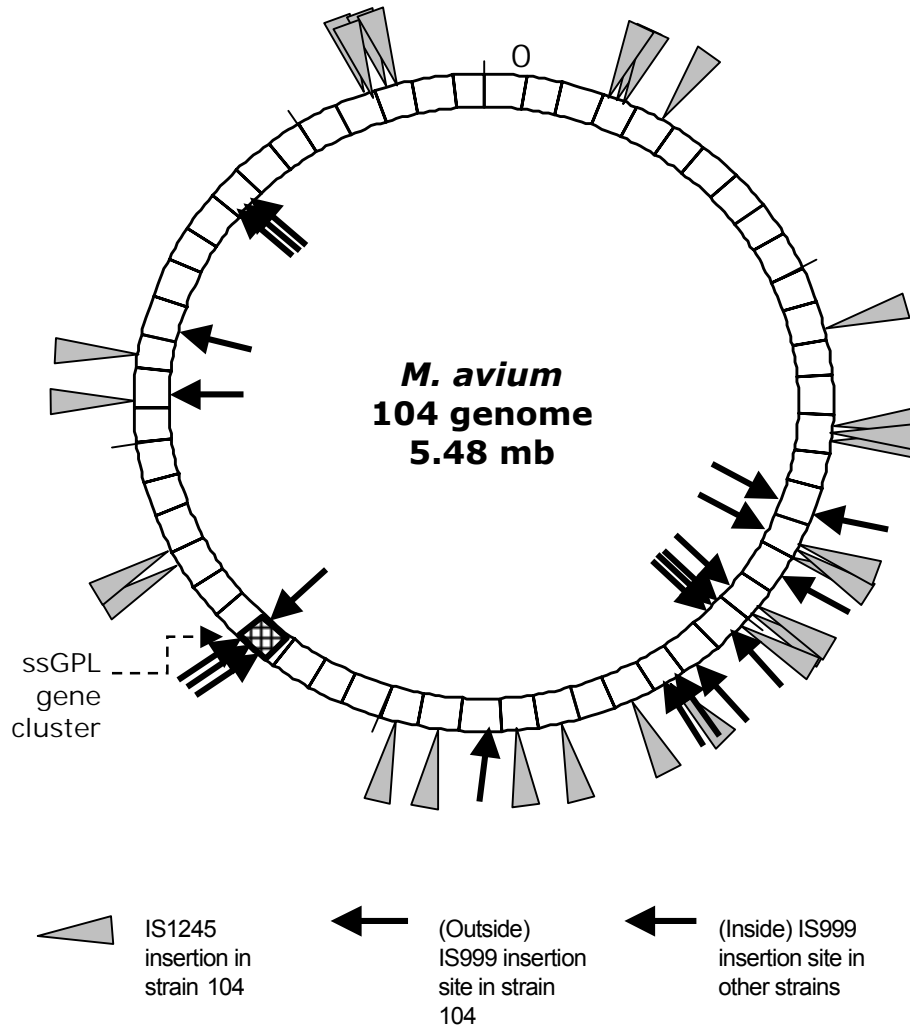


Figure 4.1 IS1245 and IS999 insertion sites in MAA strains 104, 102, HMC02 and HMC10. Insertion sites of IS1245 and IS999 in strain 104 were mapped by computer analysis of the preliminary 5.48 mb genome sequence generated by the Institute for Genomic Research (<http://www.tigr.org>). Positions of these sites are indicated on the outside of the circle. Insertion sites of IS999 in strains 102, HMC02 and HMC10 were mapped by inverted PCR and sequencing as described (Laurent *et al.* 2002). This strategy yielded positions of the subset of IS999 insertions that could be amplified in this fashion. This depended on the random occurrence of specific restriction endonuclease sites in the genomic regions bordering the insertions. Therefore, the data represent a random sample of IS999 insertion sites in these strains.

4.2.2 The MAP genome

A collaboration to sequence the genome of the virulent bovine MAP isolate K-10 is nearing completion (<http://www.cbc.umn.edu/ResearchProjects/AGAC/Mptb/Mptbhome.html>). The genome, at 4.8 mb, is 0.7 mb shorter than the MAA strain 104 genome but remarkably similar in sequence and organization (98% overall homology). The G+C% content at 69% is also comparable to MAA.

A comparative alignment of preliminary drafts of the MAP and MAA genomes has provided some insights into the biological distinctions between the two organisms. MAA has about 700 kb of sequence not found in MAP, while MAP has about 140 kb of sequence, containing 60-70 probable genes, not found in MAA. The MAP specific regions are often located at the same genomic positions as MAA specific regions. These contain all previously identified MAP specific genes (Poupart *et al.* 1993; Ellingson *et al.* 1998; Tizard *et al.* 1998; Bannantine *et al.* 2002; Nielsen & Ahrens 2002) and are organized within six large regions ranging from 10 to 60 kb, and several minor regions of 1 to 5 kb (Bull, unpublished). Many of the MAP specific regions contain, or are flanked by, remnants of bacteriophage (viral) sequences, suggesting that they may have been acquired from other bacterial species via viral cross-infection. MAP appears to have evolved from an MAA-like progenitor by a combination of gene deletion and acquisition (Brennan & Nikaido 1995; Tizard *et al.* 1998). This model is supported by more recent microarray analysis of multiple MAA and MAP strains (Semret *et al.*, submitted).

Functional genomic analysis of MAP specific regions will prove useful to understanding the unique biology of MAP. Computer analysis of these regions has already identified several genes with high degrees of homology to genes involved in pathogenicity by other bacteria. A major 60 kb MAP specific region encodes several gene sets (operons) including membrane transport systems and a novel putative mycobactin-like peptide synthesis pathway that may help to sequester important bacterial growth factors such as Fe-III and cobalt (Stratmann *et al.* in press). Systems such as these that scavenge trace nutrients are often important for the survival of bacteria in intracellular environments. A cluster of MAP specific genes homologous to *mce* genes is present. This is in addition to the common *mce* genes found in many other mycobacterial species which may be involved with specific host entry and survival.

Other MAP specific regions carry putative catalase, peroxidase and nitrile hydrolase genes associated with protection from free-radical mediated intracellular killing. One of these regions (previously designated Locus 6) is deleted in some attenuated (non-virulent) MAP vaccine strains (Hermon-Taylor & Bull 2002). Additional MAP specific genes that may play roles in virulence are described in section 4.4.3. The possibility that MAP carries virulence genes not found in MAA is consistent with the distinctive ecological niche that MAP occupies, namely that of a chronic persistent pathogen of animals.

The MAP genome also contains multiple insertion elements including IS1311 (Marsh *et al.* 1999), ISMav2 (Strommenger *et al.* 2001) and IS900 (Green *et al.* 1989). IS900 is unique to MAP (Hermon-Taylor *et al.* 2000; Bull *et al.* 2003a) and is involved in most of the known diversifications between MAP strains. There is a high degree of clonality (genetic homogeneity) between different MAP strains which may be due to the unusually slow growth rate of MAP (22-26 hrs compared with MAA 10-12 hrs) (Harris & Barletta 2001) and its relatively recent dissemination throughout the world (Moreira *et al.* 1999; Pavlik *et al.*, 1999). Nonetheless, over 30 MAP types with variations in the genomic insertion positions of IS900 are known (Bull *et al.* 2000).

4.2.3 Genomes of other environmental mycobacteria

In addition to MAA and MAP there are two other genome projects that are nearing completion: *M. ulcerans* at the Pasteur Institute (<http://genopole.pasteur.fr/Mulc/BuruList.html>) and *M. marinum* at the Sanger Centre (http://www.sanger.ac.uk/Projects/M_marinum). The *M. marinum* genome has been estimated to be about 6.5 mb long, consistent with extensive metabolic and environmental versatility. Relative to *M. ulcerans*, *M. marinum* is a genetically heterogeneous species (Stinear *et al.* 2000b; Ucko *et al.* 2002). The *M. ulcerans* genome is estimated to be smaller, at around 5.8 mb, but a more accurate size assessment will come with further sequence assembly. The *M. ulcerans* genome is extraordinarily rich in IS with two different elements (IS2404 and IS2606) accounting for 10% of the total genome. The biological consequences of the presence of these high copy number IS are as yet unknown but it seems possible that they may contribute to the relatively fastidious nature of *M. ulcerans*.

Recently a very large plasmid in *M. ulcerans* was found to encode all the genes necessary for the synthesis of mycolactone (Stinear *et al.* submitted). *M. ulcerans* may be a clonal derivative of *M. marinum* that arose by the acquisition of a plasmid from another microorganism combined with the loss of genetic information on the chromosome. The role that mycolactone plays in the natural ecology of *M. ulcerans* remains to be determined, but it seems unlikely that its primary role is pathogenesis. This molecule may confer a fitness advantage for the survival of *M. ulcerans* in a niche environment such as the salivary glands of aquatic insects. The development of *M. ulcerans* from an *M. marinum*-like ancestor may prove to be a good illustration of how significant human pathogens evolve from environmental forebears.

Currently there is no effort underway to sequence the genome of *M. intracellulare*. Given the increasing recognition of *M. intracellulare* as a species with distinct (from MAA) pathobiological characteristics, we believe that this pathogen should be the next focus of genomic analysis.

4.3 BACTERIAL PHYSIOLOGY

4.3.1 The cell envelope and its role in virulence and antimicrobial resistance

Perhaps the most distinctive feature of mycobacterial physiology is the cell envelope. The ability of pathogenic mycobacteria to survive in hostile environments such as the phagocytic vesicles of immune cells may be related in part to cell wall impermeability. The same is probably true of the pathogen's intrinsic resistance to diverse antimicrobial agents (Inderlied *et al.* 1993; Jarlier *et al.* 1994; Portillo-Gomez *et al.* 1995; Heifets 1996; Cangelosi *et al.* 2001).

The MAC cell wall is a complex array of hydrocarbon chains perforated by porins through which nutrients and other compounds pass into and out of the cell. The electron-dense peptidoglycan layer is surrounded by a hydrophobic arabinogalactan-peptidoglycan-mycolic acid layer characteristic of all mycobacteria. This layer is surrounded by a second electron-dense layer made up, in part, of ssGPL found only in MAC (Belisle *et al.* 1991, 1993; Inderlied *et al.* 1993; Wayne *et al.* 1993; Belisle & Brennan 1994). The ssGPL consists of a core nonspecific GPL modified by serovar-specific oligosaccharide side chains. The core non-specific GPL which is common to many environmental mycobacteria has a tetrapeptide structure linked to a 6-deoxy-L-talose, which in MAC is further modified with variable oligosaccharide structures to form the ssGPL.

The synthesis of ssGPL is one the distinguishing characteristics of MAC relative to *M. tuberculosis*. Genes coding for ssGPL synthesis are clustered in a region spanning approximately 50 kb of the MAC genome (Eckstein *et al.* 2000; Laurent *et al.* 2003). The ssGPL is one of the major antigens in host environments, and is responsible for the "watery" (hydrophilic) appearance of MAC colonies *in vitro*. The pathobiological role of ssGPL has not been determined. However, these molecules have been reported to accumulate in the phagosome during intracellular growth contributing to the formation of a capsule around the bacteria (Tereletsky & Barrow 1983; Rulong *et al.* 1991). Stable rough colony type variants that lack GPL frequently arise during extracellular growth, whereas strains cultivated from patient samples are invariably smooth.

4.3.2 Morphotypic switches

Most MAC isolates form multiple colony types (morphotypes) that vary with regard to infectivity, susceptibility to antibiotics and ability to grow in various environments. The transition from the smooth colony type to rough is irreversible. Rough mutants fall into two categories: those that lack all traces of GPL and those that produce a lipopeptide core of GPL that is not glycosylated. Both categories result from spontaneous deletions within the ssGPL cluster (Belisle *et al.* 1991, 1993). These

deletions are mediated by homologous recombination between neighbouring insertion elements (Eckstein *et al.* 2000).

Additional morphotypic switches in MAC are less well understood. The transparent-to-opaque switch is reversible at frequencies ranging from 10^{-4} to 10^{-5} per generation (Woodley & David 1976; Inderlied *et al.* 1993; Prinzis *et al.* 1994). Transparent variants are more virulent and more drug resistant than their opaque counterparts. Opaque variants grow more quickly on laboratory media but are rarely isolated from environmental or clinical samples. That may be an artefact of decontamination protocols applied to clinical and environmental samples, including sodium hydroxide and CPC, which are toxic to opaque variants. However, MAC colonies grown directly from the blood of AIDS patients without decontamination are transparent suggesting that this morphotype is the true form in which the pathogens exist *in vivo* (Meylan *et al.* 1990).

A separate switch, termed red-white, becomes visible when clinical isolates are grown on agar media containing the lipoprotein stain CR (Cangelosi *et al.* 1999, 2001; Mukherjee *et al.* 2001; Laurent *et al.* 2003). The red to white and opaque to transparent switches are independent of each other, such that red opaque, red transparent, white opaque and white transparent forms can be isolated from most strains. Relative to red variants, white variants are more resistant to multiple antibiotics *in vitro*, more common in patient samples and more virulent in disease models. When stable red variants are inoculated into disease models (mice and human macrophages) they switch en masse to the white morphotype. This occurs over the course of one or two generation times consistent with a switch at the level of gene expression rather than selection of a pre-existing white subpopulation. However, red transparent variants are occasionally recovered from patient samples, and the genome sequence strain MAA 104 is a stable red variant.

A gene, *crs*, has been identified that is required for the synthesis of the major CR binding site, possibly a cell wall glycolipid (Laurent *et al.* 2003). We have proposed that white variants, but not red variants, express surface components that mask underlying CR binding sites. These components, or components that are co-expressed in white variants, may also be responsible for the reduced antibiotic susceptibility associated with the white morphotype (Cangelosi *et al.* 2001). Consistent with this model, mutational loss of *crs* results in reduced CR binding and a white morphotype, but does not increase multi-drug resistance. Recently, a polyketide synthase (lipid biosynthetic) gene, *pks12*, and a second surface-associated gene of unknown function, Maa2520, were found by mutational analysis to be required for the multi-drug resistant white morphotype (Cangelosi *et al.*, submitted manuscript). However, mutational analysis ruled out a role for ssGPL in multi-drug resistance (Laurent *et al.*, 2003).

Like other MAC strains, MAP also forms smooth and rough colony type variants on agar media. Colonies of MAP grow more slowly than those of MAA and *M. intracellulare*, and only on media supplemented with the siderophore mycobactin.

An intriguing physiological feature of MAP is its ability to persist intracellularly in a cell wall deficient form, similar to a spheroplast (Chiodini *et al.* 1986; Wall *et al.* 1993). This form is non-acid fast by the ZN method. It is presumed to result from an induced cessation or radical alteration of cell wall mycolic acid production, and it may be part of a mechanism for avoiding host immune surveillance. It is not observed in other MAC. The cell wall deficient form has been found intracellularly in macrophages associated with the gut wall of up to 92% of patients with CD and 26% of controls (Sechi *et al.* 2001; Naser *et al.* 2002; Bull *et al.* 2003). Reversion to the conventional acid-fast vegetative phenotype after prolonged culture (up to two years) is possible (Chiodini *et al.* 1984; Schwartze *et al.* 2000) showing that this morphotypic switch is not permanent. Cell wall deficient forms have also been shown to persist intracellularly for more than a year when “fed” to amoebae. It is not certain, however, if MAP cells divide whilst in this chronic persistent state.

4.3.3 Metabolism and catabolism

The mycobacteria are aerobic organisms. Many environmental species such as *M. fortuitum* are fast-growing with *in vitro* doubling times in the range of 2 hours, whereas MAC is slow growing with doubling times exceeding 15 hours. Consistent with its slow growth, MAC has only a single copy of the *rrn* (ribosomal RNA) operon and a correspondingly low ribosome copy number.

MAC can utilize a variety of carbon and energy sources. Growth is stimulated by glycerol and fatty acids. The availability of fatty acids and the ability of the pathogens to utilize them are thought to be important factors in the colonization of host tissues. As with *M. tuberculosis*, optimal growth of MAC *in vitro* is observed on relatively elaborate media such as the Middlebrook series, supplemented with glycerol, oleic acid and low purity (96%) BSA. The function of the BSA supplement is not known, but it is thought to detoxify harmful by-products of fatty acid metabolism. In contrast to some pathogenic mycobacteria, MAC grows well on minimal media such as Sauton's, which consists of nothing more than glycerol, citric acid, L-asparagine as a nitrogen source and trace salts of potassium, magnesium, iron and zinc. Thus the organisms have comprehensive biosynthetic capabilities, consistent with their ability to colonize dilute environments.

There are some interesting exceptions to the rule of nutritional independence in MAC. In contrast to virulent “white” variants, spontaneously occurring non-virulent “red” variants of MAA require the presence of BSA for growth *in vitro* (Millones & Cangelosi, unpublished results). This observation might help explain the relatively poor survival of this morphotype in certain environments, including host cells and tissues (Mukherjee *et al.* 2001). In contrast to MAA, MAP strains require growth media supplemented with the siderophore mycobactin, a distinction that has been useful for bacteriological identification.

4.4 BIOLOGY OF MAC IN HOST ENVIRONMENTS

4.4.1 Entry and survival in host cells

MAA is thought to enter the human body through the bronchial or intestinal mucosa. The bacteria are ingested by host macrophages primarily (but probably not exclusively) via complement receptors and complement component C3-mediated phagocytosis (Swartz *et al.* 1988; Bermudez *et al.* 1990; Schlesinger *et al.* 1990, 1991). Complement-mediated phagocytosis does not trigger an oxidative burst within the macrophage thus facilitating mycobacterial survival after ingestion. Once inside the non-activated macrophage, MAA prevents fusion between the phagosome in which it resides and the lysosome thereby escaping the normal lytic functions of the macrophage (Frehel *et al.* 1986, 1991; Crowle *et al.* 1991; de Chastellier *et al.* 1993; Sturgill-Koszycki *et al.* 1994). MAA grows and multiplies within the phagosomal compartment until the macrophage lyses. If the released bacilli are phagocytosed by other non-activated macrophages, the cycle of growth, multiplication, release and phagocytosis continues unabated until the infected host succumbs or mounts an immune response. However, if MAA is ingested by activated macrophages, phagosome-lysosome fusion is not impeded and the bacilli are killed. Induction of a cell-mediated immune response is necessary for controlling the infection and in severely immunocompromised individuals (e.g. AIDS patients) this type of response does not function effectively and the infection is not controlled.

4.4.2 MAA genes involved in intracellular life

Understanding the metabolic activities of MAA during growth in macrophages should provide clues as to how the bacilli survive in this hostile environment. One approach has been to analyse global patterns of gene expression during intracellular growth. Plum and Clark-Curtiss (1994) were the first to use a cDNA subtractive approach to identify MAA genes that were up-regulated for expression or uniquely expressed in bacteria growing within primary human macrophages. A gene designated macrophage-induced gene (Mig) was identified. The Mig protein is an acyl-Coenzyme A synthetase that uses saturated medium-chain fatty acids, unsaturated long-chain fatty acids and some aromatic carbon acids as substrates (Morsczek *et al.* 2001). Mig protein may be involved in the metabolism of fatty acids for synthesis of mycobacterial lipids during intracellular growth.

More recently, Hou *et al.* (2002) used the SCOTS technique to identify 46 genes that are up-regulated or uniquely expressed by MAA during growth in human macrophages. This analysis provided some insights into *M. avium* metabolism during intracellular growth. For example, both the tricarboxylic acid cycle and the glyoxalate shunt appeared to function during intracellular growth. The tricarboxylic acid cycle is the central metabolic pathway responsible for generation of CO₂, adenosine

triphosphate, reduced nucleotides and precursors of several amino acids. The glyoxalate shunt functions to prevent loss of carbon molecules by bypassing the steps in which CO₂ is generated. In other bacteria the kinds of carbon sources available affect the operation of these cycles with the glyoxalate shunt becoming operative when fatty acids are used as carbon sources. The glyoxalate shunt is not used exclusively but operates simultaneously with the tricarboxylic acid cycle (Cronan & LaPorte 1996).

Expression of genes encoding several enzymes involved in biosynthetic pathways for amino acids and mycolic acids also appeared to be up-regulated during intracellular growth of MAA (Hou *et al.* 2002). In addition, genes encoding enzymes involved in mycobactin biosynthesis were up-regulated (Hou *et al.* 2002). Mycobactins are siderophores produced by mycobacteria to enable them to obtain iron, an essential nutrient for all organisms, but one which is not usually readily available to intracellular organisms. Another up-regulated gene was a homologue of the *M. tuberculosis narK3* gene (Hou *et al.* 2002). This gene codes for a nitrite extrusion protein. Excess nitrite is toxic to some mycobacteria.

Several genes that code for proteins involved in regulation of gene expression were also up-regulated for expression during growth in macrophages. Finally, a number of genes encoding homologues to *M. tuberculosis* proteins that may be important in mycobacterial pathogenesis were also observed to be up-regulated in MAA during growth in macrophages. These included genes belonging to two of the *mce* operons and two genes encoding PPE proteins (Hou *et al.* 2002). Certain *mce* proteins have been implicated in entry and intracellular survival (Arruda *et al.* 1993; Graham & Clark-Curtiss 1999), while PPE proteins have been postulated to participate in antigenic variation (Cole *et al.* 1998).

Additional global gene expression analyses have targeted proteins. Honer zu Bentrup *et al.* (1999) and Sturgill-Koszycki *et al.* (1994) employed two-dimensional gel electrophoresis to identify a protein, isocitrate lyase that is up-regulated for expression in MAA grown in mouse macrophages. Additional experiments are necessary to definitively prove that up-regulated proteins are essential for intracellular life. Moreover, there probably are additional proteins important to intracellular life which remain to be identified. Nevertheless, valuable information regarding MAA metabolism during intracellular growth has been obtained. Identification of additional up-regulated genes will further enhance our understanding of MAA physiology in this environmental niche.

Do all MAA strains express the same genes after phagocytosis by human macrophages? The MAA strain used for the SCOTS analyses was a serotype 4 strain isolated from an HIV-infected human. Do MAA strains that are able to infect humans express genes that are not expressed in strains that are not pathogenic in humans? Answers to these questions may provide better ways to assess the risks of infection by MAC-contaminated water supplies.

4.4.3 MAP genes involved in intracellular life

Accumulating evidence supports a role for unique cell wall structures in intracellular survival and virulence of MAP. A gene designated *desA1*, which codes for a desaturase associated with mycobacterial cell wall synthesis (Jackson *et al.* 1997), is intact in 90% of human isolates but is truncated by the insertion of an *IS900* element in all animal and environmental isolates (Bull *et al.* 2000). Some animal MAP isolates, when exposed to long-term intracellular persistence in amoebae, can promote the auto-excision of *IS900* from inside the *desA1* gene thereby restoring its function (Bull, unpublished). The activation of genes such as *desA1* in intracellular environments may therefore be important factors in MAP pathogenesis.

Differential expression analysis using a subgenome microarray on intracellular and extracellular cultured MAP, has highlighted a set of genes, the GS cassette, which make and modify fucose (Bull *et al.* 2000a) and are up-regulated on cell entry. The cassette is found on a 9 kb MAP specific region bounded by a repeat sequence. It has a relatively low G+C% content, suggesting that was acquired horizontally from another bacterial species. The GS cassette is found in *M. avium* subspecies *silvaticum*; however, in that organism the putative fucose acetylation gene is truncated by the insertion of *IS1612*, an element not found in MAP (Bull *et al.* 2000b). Acetylated fucose as a terminal sugar of cell wall ssGPL is the defining moiety of serotype 2 in MAA. Homologues of the GS cassette are also present in *Mycobacterium bovis* but are deleted in BCG vaccine strains (Behr *et al.* 1999). Work is in progress to determine the significance of these observations relative to MAP cell wall synthesis and virulence.

4.5 KEY RESEARCH ISSUES

This chapter presented a broad overview of what is known, and not known, about the biology of MAC and similar environmental mycobacteria in host and external environments. The following issues were singled out as priorities for future research.

4.5.1 Diversity

This issue is especially important with regard to MAA and *M. intracellulare*. These species are versatile and heterogeneous and it is possible that some strains found in the environment are infectious while others may be relatively harmless. Risk assessment may be ineffective if it is based on the assumption that all environmental strains are uniformly infectious. Strain-to-strain diversity can and should be quantified using available disease models combined with molecular epidemiological and genomic tools such as microarrays. The results could significantly improve the way we approach MAC epidemiology and risk assessment.

4.5.2 Molecular markers of virulence

If environmental MAC isolates are heterogeneous it may be possible to identify markers of pathogenic strains. If readily detectable such markers could help refine epidemiological analysis and lead to more accurate methods of risk assessment. Public health efforts to control MAC exposure will thereby benefit from molecular analysis of MAC pathogenicity.

4.5.3 Taxonomy

The pathogens we know as MAA, MAP and *M. intracellulare* are distinct in many significant ways, including environmental niches, host preferences and clinical manifestations. Recognition of these distinctions is essential to the understanding of the biology of MAC.

4.5.4 Evolutionary context

Populations of environmental mycobacteria to which humans are exposed do not appear to have human environments in their recent evolutionary histories. Their pathogenic capabilities were selected by environmental challenges encountered outside of human hosts. The ecology of these organisms in their natural habitats must be better understood if we hope to fully understand their pathogenic capabilities.

4.5.5 Genomic analysis

The impending completion of the genome sequences of several environmental mycobacteria, combined with tools for global analysis of gene expression such as microarrays and SCOTS, will greatly improve our understanding of these organisms and our ability to address the biological issues listed above. Genomic analysis should be expanded to include *M. intracellulare* and additional environmental mycobacteria. Comparative genomic analysis of mycobacteria will continue to bolster our understanding of how environmental microorganisms become human pathogens.

Acknowledgements

Data presented in this chapter were generated with funding from grant AI25767 from The National Institutes of Health and grant G8E10521 from the U.S. Environmental Protection Agency. Preliminary MAA genome sequence data was obtained from The Institute for Genomic Research website at <http://www.tigr.org>. Sequencing of the MAA genome was accomplished with support from the National Institutes of Health.