Transposon Mutagenesis of *Mycobacterium avium* subsp. *paratuberculosis* to Investigate Potential Pathogenicity Islands

Jalal-ud-din Moolji
Supervisor: Dr. Marcel Behr
Department of Microbiology and Immunology
McGill University, Montréal, Québec, Canada
April 2010

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

ABSTRACT

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of Johne's disease, a highly prevalent chronic intestinal disease of cattle. It is also the putative cause of Crohn's disease, a chronic inflammatory bowel disease of humans. The MAP genome contains six segments of DNA called large sequence polymorphisms (LSP^P) that are not present in its closest evolutionary relatives and were probably acquired through horizontal gene transfer. Together, they comprise 125kb, or 2.5% of the MAP genome, and contain 96 open reading frames. A detailed analysis of MAP evolution led us to hypothesize that the LSP^P are pathogenicity islands, encoding genes important for MAP survival or replication in the host. To test this hypothesis, we generated a 5,000-member transposon-mutant library in MAP K-10 and developed a PCR screening method to identify potential mutants of LSP^P genes. We succeeded in isolating a mutant of MAP3776c, which encodes a putative zinc siderophore. It is part of a putative five-gene zinc uptake operon that occupies almost the entirety of the insertion sequence, LSP^P15. The MAP3776c mutant does not appear to have *in vitro* growth defects, and it is able to colonize the livers and spleens of C57Bl/6 mice within one week of intraperitoneal infection. Upcoming data from an ongoing, long-term experiment will determine whether the mutant has altered ability to persist in mice. The mutant of MAP3776c and the transposon-mutant library are useful tools for research on MAP genomics and pathogenicity that might ultimately contribute to improvements in vaccines and immunodiagnostics for Johne's disease.

RÉSUMÉ

Titre de la thèse : La Mutagénèse de Transposon de *Mycobacterium avium* sous-espèce *paratuberculosis* pour Investiguer des Zones de Pathogenicité Potentielles

Mycobacterium avium sous-espèce *paratuberculosis* (MAP) est l'agent causal de la maladie de Johne, une maladie chronique intestinale des bétails très répandue. C'est aussi l'agent causal putatif de la maladie de Crohn, une maladie inflammatoire de l'intestin chronique chez l'humain. Le génome de MAP contient six segments d'ADN appelés des grands polymorphismes de séquences (LSP^P) qui ne sont pas présents chez les cousins évolutionnaires de cette bactérie et qui furent probablement obtenus par transfert génique horizontal. Ensembles, ils constituent 125kb soit 2,5% du génome de MAP, et contiennent 96 cadres de lecture ouverts. Une analyse approfondie de l'évolution de MAP nous a mené à supposer que les LSP^P sont des zones de pathogenicité qui codent des gènes importants pour la survie et la réplication de MAP dans l'hôte. Afin de tester cette hypothèse, nous avons produit une banque de mutants de transposon de MAP de 5 000 membres, et nous avons mis au point des conditions de réaction en chaîne par polymérase (PCR) pour identifier des mutants potentiels des gènes LSP^P. Nous avons réussi à isoler un mutant de MAP3776c, qui encode un sidérophore de zinc. Ce dernier fait partie d'un opéron de cinq gènes pour un système de captation zinc qui prend presque la totalité de la séquence d'insertion LSP^P15. Le mutant ne semble pas avoir de défauts de croissance *in vitro*, et il est capable de coloniser les foies et les rates des souris C57Bl/6 en une semaine à la suite d'une infection intrapéritonéalle. Les données d'expériences en cours à long terme détermineront si le mutant a une capacité modifiée à persister dans les

souris. Le mutant de MAP3776c et la banque de mutants de transposon sont des outils avantageux pour la recherche sur la génomique et la pathogenicité de MAP qui pourraient éventuellement aider à améliorer les vaccins et les diagnostiques immunologiques pour la maladie de Johne.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Marcel Behr, for the opportunity to work on a challenging project in his laboratory. Thank you for your leadership, optimism, and advice – for celebrating my successes and guiding me through difficulties. You have been a role model to me and many generations of students who have passed through your care. I greatly admire your unique combination of talents, perspectives and principles, which will influence me for the rest of my life.

I would also like to thank a former summer student, Dana McIntyre, for her invaluable help in optimizing the PCR protocol. I thank her for her dedication, professionalism and sunny disposition. This project could not have been completed without her help.

I am grateful to Fiona McIntosh, who somehow manages to keep everything running, and with unfailing happiness. I greatly benefited from her positive energy and support. I thank her for many hours of help in making boiled lysates and glycerol stocks from the mutant library, and in grinding mouse organs. I appreciated her organization and the practical training she gave me. She was a key to my laboratory achievements.

I acknowledge Christine Turenne who managed to teach me a surprising number of practical skills in very little time. She radiated kindness and was an unselfish support to anyone who needed her help. Her personality continued to influence the Behr lab long after she left.

The Microbiology and Immunology Honour's Program gave me invaluable laboratory, presentation and scientific writing experience. For this, I thank Dr. James Coulton for enthusiastically organizing the program and Dr. Bernard Turcotte for training me in his lab. I also thank Dr. Coulton for teaching a highly informative graduate course, "Protein-Protein Interactions". Thank you to Dr. Hervé Le Moual for teaching a course of antimicrobial peptides, which was stimulating and a pleasure to attend.

I thank the past and present members of Rs-1 who have supported me during my Masters program. Ashley Fallow was a wonderful moral support and her exemplary Master's thesis greatly helped me to write my own. Thank you to Vlad Kamenski for help with the mouse work. Hafid Soaulhine provided patient advice on cell culture. Dr. Erwin Schurr and his group provided excellent feedback during lab meetings. I acknowledge (in no particular order): Pilar Domenech, Paul Bruère, Rob Kozak, Victoria Pichler, Daniel Pletzer, Alex Dufort, Jean-Daniel Lallande, Speranza Marsala, Gaëlle Kolly, Sam Donato, François Coulombe and Iris Leitner for their help and friendship.

Eve-Annie Pilon edited the French version of the Abstract.
Eve-Annie Pilon a édité la version française du résumé.

For my parents and sister

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADC	Albumin dextrose complex
ATCC	American Type Cell Culture
ATG16L1	Autophagy related 16-like 1
BCG	Bacille de Calmette et Guérin
bp	Base-pairs
cDNA	Complementary DNA
CFU	Colony forming units
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
ELISA	Enzyme-linked immunosorbent assay
ESAT-6	Early secretory antigenic target – 6kDa
GTPase	Guanosine triphosphatase
HPC	Hexadecylpyridinium chloride
hsp60	Heat-shock protein 60
IL	Interleukin
IRGM1	Immunity-related GTPase family M, member 1
IS*900* (or *1110*)	Insertion sequence 900 (or 1110)
Kan.	Kanamycin
kb	Kilobase-pairs
LAMP-1 (or -2)	Lysosomal-associated membrane protein 1 (or 2)
LB media	Luria-Bertani media
LEE	Locus of enterocyte effacement
LSP	Large sequence polymorphism
LSP^P	Insertional large sequence polymorphism
MAC	*Mycobacterium avium* complex
MAP	*Mycobacterium avium* subsp. *Paratuberculosis*
MAP-kinase	Mitogen-activated protein kinase
***mce* operon**	Mammalian cell entry operon
M-cell	Microfold cell
MLSA	Multi-locus sequence analysis
MP buffer	Mycobacteriophage buffer
mRNA	Messenger ribonucleic acid
MTC	*Mycobacterium tuberculosis* complex
NOD2	Nucleotide-binding oligomerization domain containing 2
OADC	Oleic acid albumin dextrose complex
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFU	Plaque-forming units
PPDj	Johnin purified protein deriviative
PPE	Proline-proline-glutamate protein
RCF	Relative centrifugal force
RPM	Revolutions per minute
SCID	Severe combined immunodeficiency
SCOTS	Selective capture of transcribed sequences
siRNA	Small inhibitory ribonucleic acid

SLC11A1	Solute carrier family 11, member 1
T-cell	Thymocyte-derived cell
TE buffer	Tris-ethylenediaminetetraacetate (Tris-EDTA) buffer
TGFβ	Transforming growth factor β
TLR	Toll-like receptor
TNFα	Tumour necrosis factor α
tRNA	Transfer ribonucleic acid

CHAPTER 1: REVIEW OF THE LITERATURE

1.1 JOHNE'S DISEASE

1.1.1 Prevalence and Economic Cost

Johne's disease, also known as paratuberculosis, is a chronic intestinal disease of ruminants of importance to the dairy industry. It is caused by an intracellular pathogen of macrophages, *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Although MAP can infect many animals such as sheep, goats and deer, this thesis will address the best-studied and most economically important host, cattle.

It is difficult to estimate the prevalence of Johne's, as there are no good diagnostic tests, and diseased cattle are almost always culled before symptoms would indicate Johne's specifically (1). None of the current diagnostics have high sensitivity, so a combination of different assays is performed in the hope that an infected animal will test positive with at least one. Moreover, a diagnosis is typically sought for a herd, rather than for an individual cow. With these limitations in mind, estimates of MAP infection have been performed for dairy herds in most of Canada, ranging from 16.7% of in the Maritime provinces to 40% in Alberta (2). In the USA, about 21.6% of dairy herds are MAP-positive, with regional variations (3). MAP is present in every country with a significant dairy industry (4).

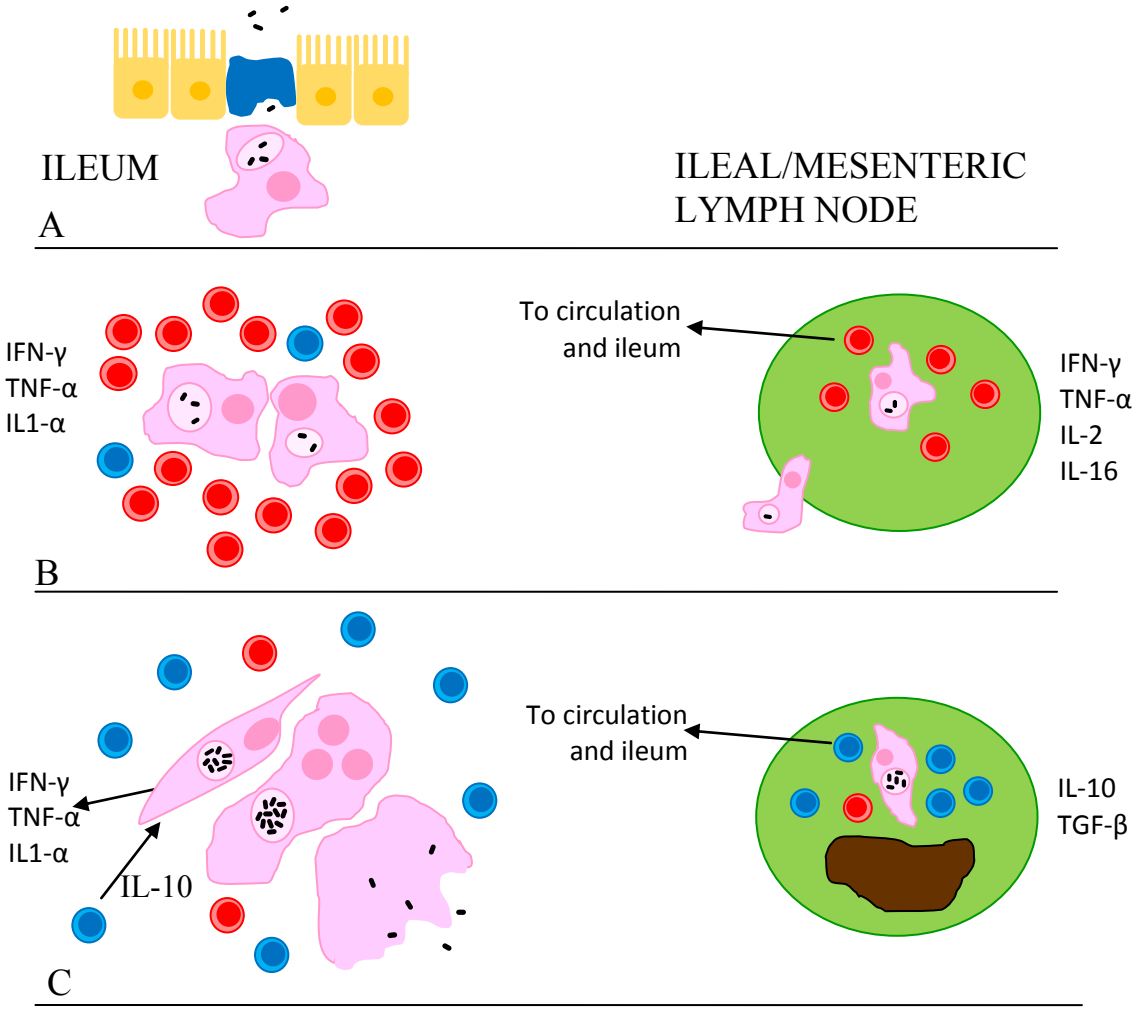
The most thorough study to date of the economic impact of Johne's was done in 1999 by Ott *et al*. (1). They tested 2,542 dairy herds in twenty American states for presence of MAP in individual cows. In their most specific and

conservative model, an individual cow was considered infected if it tested positive in at least two serum enzyme-linked immunosorbent assays (ELISA), or one serum ELISA in a herd where at least 5% of cows showed clinical signs of Johne's. Then, they tracked the milk production, the number of cows born, and the number of cows culled, died or sent for slaughter in all the herds. Based on their calculations, the average infected cow produced $97 less than a healthy cow, mostly due to reduced milk production and early culling. To contextualize this loss, the average profit margin on a single cow was $242. Based on this and their other models, the calculated loss to the entire industry in the United States was $200 to $250 million per year.

1.1.2 Symptoms and Immunology

MAP is transmitted by the faecal-oral route to calves by manure-contaminated teats (5). Minor routes of infection are *in utero*, through the colostrum, and through contaminated water and feed (6). Most cattle are infected shortly after birth, as young calves are more susceptible to infection than older calves and adults (5). Estimates for the minimal infectious dose vary from fifty to 10^3 colony forming units (7, 8). A heavy-shedding adult cow will expel 10^6-10^8 colony forming units per gram of faeces, meaning that a young calf could be infected by consuming only milligrams of contaminated faeces (9).

Once ingested, the bacteria pass through the digestive tract to the Peyer's patches of the lower ileum (Figure 1) (10). These are part of the gut-associated lymphoid tissue, which normally allows the development of healthy gut flora



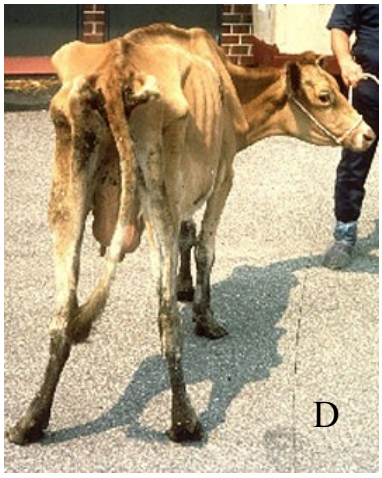


Figure 1: *Host-pathogen interactions in Johne's disease in cattle.* A) MAP is taken up by M-cells in the ileal Peyer's patches, and then by sub-epithelial macrophages. The bacteria cause phagosome maturation arrest. B) The macrophages nucleate a tuberculous granuloma, characterized by T_H1 cells (red) and cytokines. Infected macrophages also migrate to the ileal and mesenteric lymph nodes, where they stimulate a T_H1 response and T-cell proliferation. T-cells from the lymph node enter the circulation and migrate to the granuloma. C) For unknown reasons, many T-cells of the granuloma and lymph node become Tr1 or T_H3 regulatory cells, which produce IL-10 (blue). This counteracts the T_H1 cytokines produced by the macrophages, and changes the granuloma's structure. The macrophages in the granuloma convert to epithelioid cells and multinucleated giant cells. There is rapid bacterial proliferation. Heavily infected cells lyse, releasing bacteria into the gut lumen. The lymph nodes develop necrotic lesions. D) Advanced Johne's disease in a Guernsey cow (photo courtesy of the Johne's Information Center, http:/johnes.org).

while defending against enteric pathogens. Peyer's patches are organized collections of immune cells capped by a layer of specialized microfold cells (M-cells) at the epithelial surface. Beneath the M-cells is a layer of macrophages and dendritic cells, bordered by a cluster of lymphocytes. The function of M-cells is to passively transmit material from the gut lumen to the underlying antigen presenting cells (transcytosis), allowing immune surveillance of the luminal contents. The ingested MAP is engulfed by the M-cells, which transfer viable bacteria to the underlying antigen presenting cells. Then, the MAP is phagocytised by the macrophages, which is its natural host cell. The role of dendritic cells in Johne's pathogenesis has not been determined yet.

MAP survives inside the macrophage phagosome by blocking phagosome-lysosome fusion. The bacterial phagosome is positive for the early endosomal marker, transferrin receptor, but has low amounts of the late endosomal markers LAMP-1 and LAMP-2 (11, 12). Phagosomes containing live MAP acidify only slightly, to pH 6.6, whereas phagosomes containing dead MAP or non-pathogenic *Mycobacteria* acidify to pH 5-5.8 (12). There is preliminary evidence to suggest that MAP inhibits phagosome maturation by phosphorylating MAP-kinase (13) and inhibiting the Rab GTPases (14), though the relevant bacterial factors responsible for these outcomes are undiscovered. About 70% of internalized MAP survive after 24h in J774 macrophage cell culture, compared to only about 1.5% of the environmental bacterium *Mycobacterium smegmatis* (12). Bacterial numbers do not decline after this point, and MAP persists for the long-term (at least 15 days in cell culture). The infected macrophages spread to the ileal and

mesenteric lymph nodes, from which MAP can be cultured within one hour of ingestion (15).

Initial MAP infection is asymptomatic and leads to a long incubation time, called "silent infection" (stage I of Johne's disease) (6). This can last from two to ten years, with an average of about five years. The cattle appear healthy, and appetite is normal. There is no reliable diagnostic test to identify MAP infection at this stage, and no bacteria can be cultured from the faeces. In experimental infections, very minor histopathology and inflammation are sometimes seen in the ileum and adjacent lymph nodes by microscopy, but bacteria are not visible.

Like silent infection, subclinical infection (stage II) has no symptoms and appetite remains normal (6). At this time, tuberculoid granulomas appear in the lower ileum, especially at the ileocecal valve (16). These are tight clusters of immune cells nucleated by macrophages and surrounded by lymphocytes, plasma cells and sometimes eosinophils. Again, bacteria cannot be seen in these lesions. T_H1 cytokines predominate in the granulomas, including IL-6, IL-2, IL-1α, TNFα and especially interferon-γ, which is supported by IL-12 and IL-18 (16, 17). It is theorized that the T_H1 response, particularly interferon-γ, activates infected macrophages to facilitate killing of intracellular bacteria. At the same time, TNFα helps to create compact granulomas, which prevent the dissemination of infected macrophages, thereby limiting the infection to discrete lesions. The cytokine profile of infected lymph nodes is slightly different, containing IL-2, which encourages lymphocyte proliferation, and IL-16, a chemoattractant for helper T-cells (18). The lymph nodes allow the development of antigen-specific, T_H1-

biased helper T-cells, which then circulate systemically. This is the basis for the traditional skin test for Johne's disease, in which MAP antigens are injected subcutaneously and a type-I hypersensitivity reaction is observed. At this stage, small numbers of MAP may be shed in the faeces.

Stage III, or clinical disease, is characterized by diarrhoea and gradual weight loss (6). The appetite remains normal, but milk production decreases and MAP can usually be cultured from the faeces. Symptoms are correlated with the supplanting of tuberculoid granulomas by lepromatous granulomas (16). These are loose aggregates of leukocytes containing large numbers of epithelioid cells, which are highly activated macrophages, and sometimes giant cells, which are syncytial, heavily infected macrophages. There are fewer leukocytes in the granulomas; however, there is extensive inflammation of reactive leukocytes through all the layers of the gut wall and the gut mucosa thickens. MAP can now be seen in large numbers in the ileum and in lymph nodes throughout the body by microscopy. The events responsible for the transition from subclinical infection to clinical disease are unknown.

New immunological models argue that clinical disease occurs when helper T-cells switch from reactive T_H1 cells to immunosuppressive T_H3 and/or Tr1 cells. T_H1 cytokines continue to be produced by macrophages of the granulomas, notably interferon-γ (18, 19). However, the neighbouring $CD4^+$ T-cells produce large amounts of the powerful immunosuppressive cytokine IL-10, which blocks interferon-γ-mediated macrophage activation. As a result, the macrophages are prevented from activating and destroying intracellular bacteria. It has been

hypothesized that antigen-specific T_H1 cells to convert to $CD4^+$ $CD25^+$ regulatory T-cells (Tr1) due to persistent exposure to MAP antigens in the gut, although convincing evidence that these cells express other markers (e.g. Foxp3) is lacking. Large amounts of IL-10 and TGFβ are also present in the lymph nodes which, in combination, might convert activated T-cells to immunosuppressive T_H3 regulatory cells (20, 21). These T_H3 cells could then migrate to the ileum to prevent macrophage activation and formation of functional granulomas through IL-10. Interestingly, interferon-γ is greatly downregulated in the lymph nodes compared to the granulomas, perhaps allowing enhanced bacterial replication within infected macrophages there. There is also an increase in activated γδ T-cells in the peripheral blood, which might contribute to IL-10-mediated suppression (22, 23). The common denominator of the various models is that IL-10 converts T_H1 helper T-cells to a regulatory/suppressor phenotype, preventing macrophage activation and bacterial clearance. New macrophages continue to be recruited to the granuloma by the T_H1 cytokine IL-1α, providing fresh host cells for the replicating bacteria (17). This, in turn leads to uncontrolled bacterial replication, faecal shedding and pathology. Importantly, most of these observations are correlative, meaning that it is difficult to know whether these immunologic events drive the change in clinical stage, or instead are the result of the clinical progression.

Stage III gradually leads to advanced clinical disease (stage IV) (6). There is profuse diarrhoea, hypoproteinemia and severe weight loss. If the animal is not culled, it dies from dehydration or severe emaciation (Figure 1D).

1.1.3 Diagnosis of MAP Infection

Diagnosis of MAP infection has five main purposes: 1) the certification of herds or animals as MAP-free 2) estimating the prevalence of infection 3) informing ways to reduce transmission 4) allowing improvements in farm practices and 5) improving animal welfare (24). Clinical signs are not specific enough for Johne's to be the basis of diagnosis (1). Diseases such as intestinal parasitosis, malnutrition, salmonellosis and winter dysentery are difficult to distinguish from early clinical Johne's. Cows that appear unhealthy are often culled at this stage, frustrating estimates of MAP prevalence based on clinical signs. The ideal diagnostic test would be rapid, cheap, sensitive, highly specific for MAP, and applicable to individual animals rather than whole herds. None of the current tests reproducibly detect subclinical infection, which is badly needed to design effective control strategies.

The current gold standard in MAP diagnosis is faecal culture from an individual infected animal. Methods vary among researchers, but Richard Whittington has written an extensive description of faecal culture methods in a recently published textbook on paratuberculosis (25). Faecal culture involves four steps: 1) decontamination of fast-growing bacteria and fungi from the sample 2) prolonged incubation on media containing chemicals to suppress the growth of contaminants 3) Recognition of MAP colonies and 4) identification of MAP by phenotype or genotype. After the faecal samples are collected, they are suspended in water and either centrifuged or passed through a 3µm filter to remove particulates, which often contain fast-growing enteric bacteria. They are

then treated with a chemical decontaminant like hexadecylpyridinium chloride (HPC), which kills most bacteria, but not MAP. The samples are centrifuged to concentrate the bacteria. The samples are added to specialized media that contain egg yolk, such as Lowenstein-Jensen agar or Herrold's egg yolk agar, supplemented with mycobactin J. Egg yolk is not required if the sample is washed once with water, suggesting that it acts by protecting MAP from inhibition by the chemical decontaminant. Mycobactin J is an iron-binding siderophore necessary for the initial isolation of MAP. The media normally contains a mixture of antibiotics such as PANTA Plus, which includes polymyxin B, amphotericin B, nalidixic acid, trimethoprim, azlocillin and polyoxyethylene stearate to kill contaminants. 7H10 and 7H11 (used in this project) also contain pigments such as malachite green to prevent the growth of non-*Mycobacteria*. The media is then incubated for several months to allow the growth of MAP before being declared culture-negative. If colonies are found, they are first identified by colony morphology (Figure 11). PCR for IS*900*, a transposon that occurs 15-20 times in the MAP genome, is performed on the colonies. This transposon is fairly specific for MAP, although closely related elements are sometimes observed in environmental *Mycobacteria*. An animal that sheds MAP in its faeces is considered infected. However, it is suspected that this method gives frequent false negatives because of the harshness of the decontamination, and current media might not be perfectly optimized for MAP cultivation.

Several methodological changes can be made to address the problems of faecal culture on individual cows. One strategy is to test multiple cows, and label

entire herds MAP infected or MAP-free instead of individual animals (9, 25). This improves the apparent sensitivity of culture, since only a few individual cows need to be culture-positive for the entire herd to be labelled positive. This also enables pooling samples, which is a cost-saving measure. Another strategy is cultivating faecal samples from barns, pastures and water. This technique detects 76% of infected herds from which MAP can be grown from the faeces of individual cows (26). Environmental samples are easy to collect and do not require handling of the cows. This might be a good way to rule-in the presence of MAP on a farm.

A traditional diagnostic is the Johnin purified protein derivative (PPDj) skin test. Briefly, PPDj is prepared by growing MAP in liquid culture, autoclaving it, and centrifuging the sample to precipitate the solid particles. The liquid supernatant is called PPDj (27). In cattle, this is injected under the skin of the caudal skin fold. A few days later, the injection site is examined for the presence of inflammation, indicating type-I (cell-mediated) hypersensitivity. If the swelling is less than 2mm in diameter, the cow is considered non-infected (28). If the swelling is greater than or equal to 4mm in diameter, the cow is considered infected. Intermediate values are inconclusive. The specificity of this test is about 94%, although data about the sensitivity are unavailable. The advantage of this test is that it is quick, cheap and easy. However, it requires handling the animals twice, involves a somewhat subjective reading, and can produce inconclusive results.

The interferon-γ assay also uses cell-mediated immunity, but aims to reduce subjective readings. Blood is drawn from an animal and the peripheral blood leukocytes are extracted (28). These are then incubated in a 96-well plate, and stimulated with PPDj. An ELISA-type procedure is then performed, in which an enzyme-conjugated antibody for interferon-γ is incubated with the cells. The optical density of the plate is read. Unfortunately, the reactive leukocytes (which are antigen-specific T-cells) often are cross-reactive for *M. bovis* antigens. *M. bovis* is the causative agent of bovine tuberculosis and sometimes causes gastrointestinal disease in humans, making it an important public health concern. Therefore, controls ruling out anti-*M. bovis* immunity should be performed. The results of the interferon-γ assay are 50-85% sensitive and about 95% specific in experimentally-infected cattle (28, 29). This test has the advantage of having to handle the animal only once.

There are commercially available antibody ELISA kits for the detection of anti-MAP antibodies. These have the advantage of being applicable to milk, which is already removed from the animals regularly. Unfortunately, few antibodies are produced before clinical infection. This means that although specificity is usually greater than 95%, the sensitivity of antibody ELISA is probably 5-30% in subclinical infection (30). This ELISA is paired with ELISA for anti-*M. bovis* antibodies to rule out tuberculosis. The use of both ELISA partially accounts for the assay's high specificity but low sensitivity. To explain, anti-MAP antibodies often cross-react with *M. bovis* antigens, so specificity is

low. In uncommon cases of no cross-reactivity, the assay is very specific. Once clinical infection begins, the most convincing diagnostic is faecal culture.

PCR of faecal samples for IS*900* has similar sensitivity to faecal culture, but gives relatively frequent false positives and false negatives (31). False positives occur due to cross contamination of DNA in the laboratory. False negatives occur due to the presence of inhibitors in the template. Because of these technical difficulties, PCR has poor reproducibility both within and between laboratories. This problem is worsened by the fact that there are no standard protocols, so each group has its own method. Improvements are needed in PCR diagnostics before they become viable alternatives to other tests. An effective PCR method would have the advantages of being rapid and cheap.

1.1.4 Vaccination

The first paratuberculosis vaccine was developed in France in 1926 by Vallée and Rijnard (32). It consisted of live MAP suspended in an adjuvant of olive oil, paraffin and pumice powder, delivered subcutaneously. Since then, the most common vaccine has been a live MAP-in-oil vaccine delivered subcutaneously. The alternative killed vaccine appears to be safer, but less effective. The main attributes of the live vaccine are listed in Table 1.

A comprehensive study of the effects and costs/benefits of vaccination was performed in Dutch dairy cows by van Schaik *et al*. (33). They showed that vaccination actually reduced milk production and quality in all cows except those already showing clinical disease (in which vaccination had a relatively small effect). The reduction in milk production and quality suggests that the live

Table 1: Characteristics of the current MAP vaccine compared to a hypothetical, ideal MAP vaccine. Reproduced from Huygen *et al.*, 2010 (34).

Vaccine characteristic	Whole-bacilli-in-oil vaccine	Ideal MAP vaccine
Prevents establishment of infection	No	Yes
Prevents clinical disease in uninfected animals	Yes	Yes
Prevents clinical disease in already infected animals	Yes, in early stages of infection	Yes
Injection site lesions	Yes	No
Injury to humans from accidental inoculation	Yes	No
Causes false-positive responses for immune based tests for bovine TB	Yes	No
Causes false-positive responses for immune-based tests for MAP	Yes	No

vaccine has important negative health effects. Furthermore, the number of cows shedding MAP asymptomatically was the same in the vaccinated and unvaccinated groups, suggesting that vaccination did not affect infection rate. Encouragingly, vaccination reduced the number of subclinically infected, non-shedding cows by 90%. Subclinically infected cows produced less milk and were culled at younger ages. It was economically advantageous to vaccinate, even if this reduced milk production, to prevent cows from having even lower productivity due to subclinical disease. As a result, the $15 vaccination increased the productivity of the average cow by $142. In herds where MAP infection is rare, it is unclear if vaccination would still be economical. The results of this study correspond to observations in Iceland, where MAP was introduced in 1933

by twenty imported German sheep (35). Initially, mortality due to Johne's disease was 8-10% of Icelandic sheep herds, but widespread vaccination reduced mortality to almost zero without eradicating MAP infection.

Huygen *et al.* proposed two main reasons for the inability of the current vaccine to eradicate MAP infection totally: the vaccine might not effectively present T_H1 antigens to the immune system, and the vaccine strains still might have immunosuppressive characteristics (34). Defining immunodominant T_H1 antigens is an active area of research. The current strategy involves inoculating animals with secreted or surface proteins, and determining if they provoke a T_H1 or T_H2 response. The authors then propose that the T_H1 or mixed T_H1/ T_H2 antigens be overexpressed in vaccine strains to enhance immune recognition of MAP. In this way, superoxide dismutase (SOD), and the antigen 85 complex (MAP0126c, 1609c, 3531c) were proposed as good candidates for overexpression in vaccine strains or as components of subunit vaccines (36, 37).

Like wild-type MAP, it is likely that the vaccine strain evades the immune system. The vaccine strain 316F is just as virulent as the reference wild-type strain (K10) in mice (34). It has no large genomic deletions, like those that attenuated the human tuberculosis vaccine, *M. bovis* BCG. The vaccine does not cause Johne's in ruminants probably because it is delivered subcutaneously rather than orally. A priority for future research would be to identify and eliminate MAP virulence factors to engineer a safer and more efficacious vaccine.

A third explanation can probably be added to the above list: vaccine-induced immunity might target antigens that are present in established infection,

but not in early infection. This would account for the observation that vaccination does not change infection rate, but limits disease progression. Therefore, it would be useful to identify antigens that are expressed in early infection, such as initial macrophage colonization in the ileum. These could then be over-expressed in the vaccine strains or be made into a mixed subunit vaccine. Antigens that are expressed once infection has been established, but do not contribute to immunity can be eliminated from vaccine strains, but included in diagnostic tests. In this way, vaccination would interfere less with immunodiagnostic assays for MAP.

Diagnostics and vaccines based on MAP-specific antigens might have additional advantages. Diagnostics based on MAP-specific antigens would be specific for MAP rather than other *Mycobacteria*, especially *M. bovis*. Furthermore, a mixed subunit vaccine of MAP-specific antigens would not interfere with diagnostics for bovine tuberculosis. This refinement of vaccines and diagnostic tests would require information about the expression and antigenicity of MAP-specific proteins. Knowing the functions of MAP-specific proteins might be useful in making vaccines against proteins that are important for bacterial survival, and determining how to deliver the vaccines. In this Master's project, a knock-out of a MAP-specific secreted protein was generated, which might be useful in new vaccines or diagnostics.

1.1.5 The Mouse Model

Experiments in large ruminants are costly, and there are few immunological tools for these animals. In contrast, mice are cheap, they come in

many well-characterized breeds, and there are many immunological tools that permit mechanistic dissection of *in vivo* phenomena. Therefore, mice are often used in initial experiments before moving on to large animals. The host susceptibility genes for paratuberculosis, *NOD2* and *TLR2*, were first identified in mice, and then in cows (38-40). Work on γδ T-cells in mice preceded investigations in cattle (19, 23). There are currently several vaccines in the pipeline that have been validated in mice and are moving on to cattle trials (41). Unfortunately, several key features of Johne's are not reproduced in mice, including diarrhoea and death. Faecal shedding of MAP and weight loss are only observed in Beige/SCID mice (42). Nonetheless, mice are currently an important tool considering the barriers to experimental ruminant infection.

Balb/c and C57Bl/6 mice are immunocompetent, but susceptible to MAP infection. Balb/c mice develop granulomas similar to those found in Johne's, including epithelioid cells, macrophages, T-cells and acid-fast bacteria (43). These granulomas increase in size over time. There is also bacterial replication in the spleen. Unlike Johne's, however, granulomas are seen in many organs including the liver, spleen, pancreas and even the uterus and heart. In contrast, C57Bl/6 can sustain MAP infection in the liver, spleen and mesenteric lymph nodes, but show few granulomas in the intestine (44). This might reproduce subclinical infection in cows, in which there is no intestinal pathology, but the lymph nodes are a potential reservoir of bacteria.

C3H/He mice are resistant to MAP infection. They clear most of the bacteria, and develop few granulomas in the liver and spleen (44). These

granulomas decrease in size and number over time. This mouse breed could be used to determine host resistance mechanisms to MAP.

Beige/SCID mice, which lack leukocytes, can be infected with MAP. They develop intestinal ionic imbalance and villi enlargement (42). However, there is limited profusion of lymphocytes into the gut wall, and granulomas are not possible in the absence of T-cells. Interestingly, there is faecal shedding of bacteria. This breed could be a model for clinical or advanced clinical disease, in which adaptive immunity fails to control bacterial replication.

The most efficient route of delivery is intaperitoneal infection, which is the current international standard for mouse infection (41). The more natural, oral route sometimes fails to cause infection even in Beige/SCID mice, and requires large amounts of bacteria (10^8-10^{10}) (42, 45). It is possible that progression from oral infection to clinical disease requires a long incubation time, like in ruminants. In this case, the incubation time might exceed the mouse lifespan.

Given the current state of knowledge, it is unclear which mouse model best represents Johne's, and at which stage of disease. This study used C57Bl/6 mice because they are highly resistant to infection. This should maximize the selective pressure on the infecting bacteria, revealing subtle differences among bacterial strains. Although C3H/He are also resistant, there is no detectable intestinal involvement. Balb/c or other breeds might be useful in future studies on intestinal histopathology.

1.1.6 MAP as the Putative Cause of Crohn's Disease

Crohn's disease is a chronic inflammatory bowel disease of humans, mostly affecting the terminal ileum or the colon. It is a relapsing and remitting disease involving intestinal pain and malabsorption, leading to failure to thrive in children and weight loss in adults (46). The ileum manifests transmural inflammation, with granulomas seen in many but not all cases. In this aspect, the pathology resembles but is not identical to that of Johne's disease. Severe cases can involve fistulas (perforations in the gut wall) and intestinal obstruction. The similarities between Crohn's and Johne's have led to the controversial hypothesis that MAP infection causes Crohn's. The long-term objective of much MAP research is to determine the link (if any) between MAP and Crohn's. This Master's project aims to make a small contribution towards this goal.

Like in Johne's, the granulomas in Crohn's disease consist of epithelioid cells and multinucleated giant cells surrounded by lymphocytes. Traditionally, the cytokine profile of the granulomas has been described as T_H1-biased, as elevated levels of IL-12, interferon-γ, and TNFα are detectable *in situ* (47). There is also increased IL-1β, IL-5, IL-6 and IL-18. This cytokine profile is reminiscent of Johne's. New research suggests that T_H17 cells are involved since IL-23, can be found within lesions, and a mutation in the IL-23-receptor gene protects against Crohn's (47, 48). The T_H17 response has not been studied in Johne's specifically; however, there is data about T_H17 in other mycobacterial infections. MΦ-1 macrophages (activated by GM-CSF) produce IL-23 immediately after mycobacterial infection (49). Mice lacking the *Il-17* gene have

a weaker antigen-specific T_H1 response to *M. bovis* lung-infection than wild-type mice (50). This data points to a role for T_H17-cells in supporting the T_H1 response to mycobacterial infection and provides further support, albeit indirect, for the hypothesis that MAP is involved in the aetiology of Crohn's.

The genetic risk factors for Crohn's suggest a defect in innate immunity leading to impaired clearance of intracellular bacteria. Polymporphisms in *NOD2*, *IRGM1* and *ATGL16L1* predispose specifically to Crohn's disease and not ulcerative colitis, another inflammatory bowel disease (51-53). NOD2 is an intracellular pathogen recognition receptor that recognizes a component of bacterial peptidoglycan, *N*-acetyl muramyl dipeptide. Persons homozygous for a frameshift mutation in NOD2 have an odds ratio of 17.1 for developing disease (54). Last year, it was shown that NOD2 is especially sensitive to the mycobacterial form of this molecule, *N*-glycolyl muramyl dipeptide (55). Once stimulated, NOD2 contributes to TNFα production and macrophage activation. IRGM1 is a GTPase involved in autophagy, which is the cellular process by which cytoplasm, damaged organelles and intracellular bacteria are collected in vacuoles and destroyed. IRGM1 and autophagy are involved in the killing of intracellular *Mycobacteria* by macrophages. Knocking-down *IRGM1* using siRNA made macrophages more permissive to *M. bovis*, resulting in a 50% increase in intracellular *M. bovis*, 4h after infection (56). ATG16L1 is another protein related to autophagy. Recently, it was shown that NOD2 recruits ATG16L1 to the part of the cell membrane during phagocytosis of bacteria (57) and that patients with permissive mutations of NOD2 have impaired autophagy

(58). These studies implicate NOD2 in autophagy, suggesting that these different mutations affect distinct checkpoints along a bacterial recognition pathway, rather than serving different host processes. The fact that NOD2, IRGM1 and ATG16L1 are involved in killing intracellular *Mycobacteria* is consistent with MAP being linked to Crohn's.

However, other observations cast doubt on the hypothesis that MAP is involved in Crohn's. Although antibiotic therapy may result in a significant improvement in symptoms, this quickly becomes ineffective and does not cure the disease (59, 60). Furthermore, current treatments for Crohn's involve immunosuppression using IL-10, anti-TNFα antibodies, or anti-IL-23 antibodies (46, 61). These therapies might be expected to aggravate a MAP infection, in which case they might be expected to worsen rather than alleviate symptoms. These observations are important, but the MAP-in-Crohn's hypothesis is best addressed using microbiological techniques directly to detect (or fail to detect) MAP in Crohn's patients (62). Models based on immunological or pathological observations must be made to fit with unambiguous microbiological detection.

Considerable effort has gone into trying to searching for MAP in diseased Crohn's tissues directly. The most convincing method, culturing MAP from biopsy samples, has received less attention than PCR, and in general, has been very unsuccessful. This has led some to conclude that MAP is almost never present in Crohn's (63). Others have attempted to find MAP in Crohn's biopsies with light microscopy. This is complicated by the fact that MAP is very small (1.2μm × 0.5μm) and can have a "cocco-bacillus" shape in that is difficult to

identify with confidence. Despite these limitations, organisms consistent with MAP can often be found in Crohn's tissues when examined under oil immersion (1000× magnification) (64). So far, the most successful method has been PCR of infected tissues. A recent meta-analysis showed that MAP is selectively present in the tissues of Crohn's patients compared to controls with ulcerative colitis, with an odds ratio of 7.0 (65). This is a striking statistic, considering the odds ratio of fair/poor health in obese Americans (body mass index $\geq$ 40) is only 4.19 (66). However, data from different groups are inconsistent and "in house" PCR often lacks external validation and blinded controls (62). Therefore, more reliable and reproducible assays are needed to detect MAP in diseased tissues. One possibility is to find MAP-specific antigens that can be detected either directly in patient samples or indirectly by screening for serum antibodies. The MAP mutant created in this Master's project may ultimately be useful in developing such an assay.

1.2 THE EVOLUTION OF *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

1.2.1 The Genus Mycobacterium

Only a few *Mycobacteria* are pathogenic, the majority are harmless environmental species living in the water or soil. The distinguishing feature of the genus is the cell wall, which is technically Gram-positive, but is highly unique to *Mycobacterium* (Figure 2) (67). The cell wall consists of the plasma membrane, surrounded by a thin layer of peptidoglycan. This is covalently linked to long-chain lipids, which make the envelope thick and very hydrophobic or

"waxy". This property causes *Mycobacteria* to resist crystal violet staining, so they appear weakly Gram-positive. Identification is traditionally made using the acid-fast stain, which relies on the waxy cell wall. Many unique lipids and lipoproteins intercalate into the cell wall. The mycobacterial envelope stores carbon and energy, and allows the bacterium to resist chemical and physical stresses. In pathogenic *Mycobacteria*, the lipids are also important virulence factors. Many of them have been identified in *Mycobacterium tuberculosis*, including: phenolphthiocerol dimycoserate (PDIM), phenolic glycolipid (PG), trehalose dimycocoserate (TDM), lipoarabinomannan (LAM) and sulfolipid (SL) (67). *Escherichia coli* has only 50 fatty acid metabolism genes, compared to 185 in *M. tuberculosis* and 250 in MAP (68, 69).

The genus can be subdivided into the rapid-growing and the slow-growing *Mycobacteria* (Figure 3A) (70). The rapid-growers are nearly all environmental organisms. Their ability to metabolize many lipids makes them useful for various purposes, such as the bioremediation of oil spills. The slow-growers include numerous important pathogens, including the *Mycobacterium tuberculosis* complex (MTC; Figure 3B). These cause tuberculosis in various animals, including cattle (*M. bovis*), goats (*M. caprae*) and humans (*M. tuberculosis*). The tuberculosis vaccine is an attenuated form of *M. bovis*, called *M. bovis* BCG (Bacille de Calmette et Guérin) (71). The doubling time of these organisms is 18 to 24 hours depending on the culture conditions, compared to 20 minutes for *E. coli*.

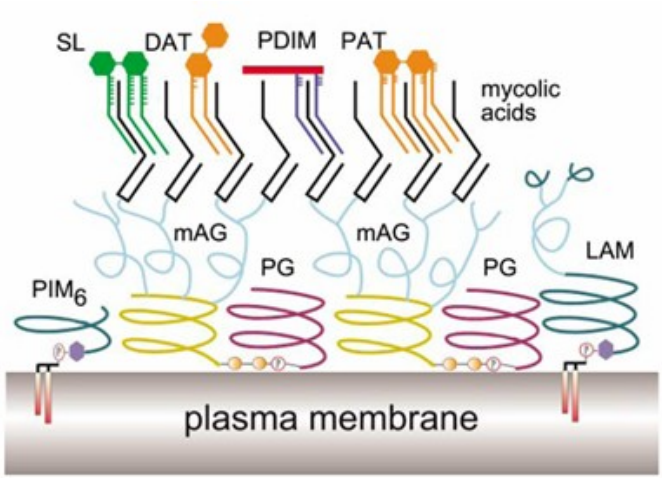


Figure 2: The cell wall of Mycobacterium tuberculosis. PG = peptidoglycan, PIM6 = phosphatidylinositol pentamannoside, mAG = mycolyl arabinogalactan, SL = sulfolipid, DAT = diacyl trehalose, PDIM = phthiocerol dimycocoserate, PAT = pentaacyl trehalose. Reproduced with the publisher's permission (72).

M. leprae is the cause of Hansen's disease (leprosy), which only affects humans. The *M. leprae* genome illustrates an important theme for pathogenic *Mycobacteria*, reductive genomics. Its genome is 27% pseudogenes, and 23.5% non-coding DNA, which might contain "gene remnants mutated beyond all recognition" (73). This is compatible with the notion that pathogenic *Mycobacteria* have lost most of their adaptations to life in the soil and water as they evolved into pathogens. The host environment is highly consistent, so the bacteria need not maintain metabolic flexibility, retaining only core metabolic genes and virulence genes.

Genetically, *M. intracellulare* is the closest relative of *M. avium*, and is part of the *Mycobacterium avium* complex (MAC) (74). It occasionally causes disseminated bacteraemia, or a tuberculosis-like disease in hospitals and the immunocompromised. It can be distinguished from *M. avium* by certain biochemical properties and its 16S rRNA sequence.

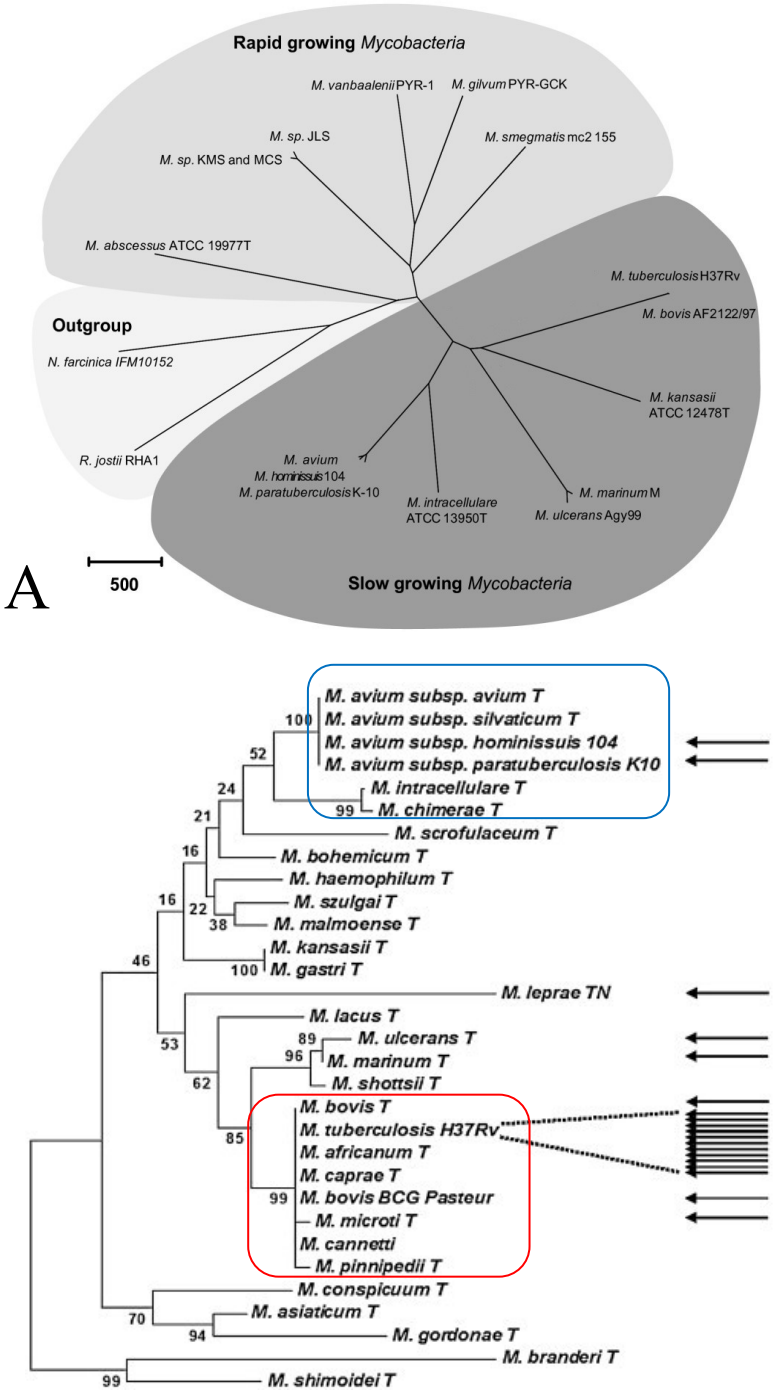


Figure 3: Phylogenies of the genus <u>*Mycobacterium*</u>*.* A) Unrooted phylogeny highlighting the classification into rapid growing and slow growing *Mycobacteria*. The scale is the number of amino acid differences among all predicted proteins (70). B) Rooted phylogenetic tree of slow growing *Mycobacteria*. Arrows indicates genome projects completed or in progress. The *M. avium* complex (MAC) is highlighted in blue and the *M.* tuberculosis complex (MTC) is highlighted in red. The scale represents relative genomic similarity, 100 being the highest (very high genomic similarity, but not 100% identity) (75).

1.2.2 The Evolution of Mycobacterium avium

Mycobacterium avium consists of four subspecies, which were originally distinguished by the environments from which they are normally isolated. *M. avium* subsp. *hominissuis* is considered a free-living, environmental species and can be isolated from the water, the soil and even biofilms on showerheads (74, 76). It also causes an opportunistic bacteraemia in AIDS patients and occasionally infects swine. It is capable of parasitizing amoebae, though the importance of this feature to the subspecies' survival is unknown. In contrast, *M. avium* subsp. *avium* and *M avium* subsp. *sylvaticum* cause avian tuberculosis. The doubling times of these organisms are about the same as *M. tuberculosis*. The final subspecies is the subject of this thesis, *M. avium* subsp. *paratuberculosis* (MAP), whose doubling time is longer, estimated at one to two days depending on culture conditions. All three pathogenic *M. avium* are intracellular pathogens, and do not naturally replicate outside the macrophages of their hosts. Before the genetic era, these organisms were classified as separate species of *Mycobacterium*. However, their 16S rRNA sequences were found to be identical, causing them to be reclassified as subspecies of *M. avium* to reflect their evolutionary closeness (74).

Our group was interested in determining the phylogenetic relationship among the *M. avium* subspecies to understand better how they evolved to occupy such different ecological niches. To do this, we performed multi-locus sequence analysis (MLSA), which involves sequencing conserved housekeeping genes from different bacterial isolates, and then aligning the sequences to obtain a

phylogenetic tree (77). We compared ten housekeeping genes among 55 isolates to obtain the phylogeny given as Figure 4A. As shown, *M. avium* subsp. *hominissuis* is a heterogeneous subspecies where genetic study presents evidence for recombination. This is consistent with its definition as an environmental organism, which must obtain nutrients in numerous habitats and must withstand a variety of physical and chemical stresses. Our phylogeny probably underestimates the genetic diversity of *M. avium* subsp. *hominissuis* because the available strains are biased towards clinical (as opposed to environmental) isolates. In comparison, the other *M. avium* have very little strain diversity within each subspecies. This supports the observation that they are intracellular pathogens occupying highly consistent and controlled environments. Furthermore, the fact that they are intracellular pathogens limits their interaction with other bacteria and phages, making recombination and horizontal gene transfer unlikely.

Interestingly, it appears that the bovine and avian pathogens are more closely related to the environmental subspecies than to each other. This suggests that the pathogenic *M. avium* are descended from two different isolates of *M. avium* subsp. *hominissuis* (Figure 4B). In other words, *M. avium* subsp. *hominissuis* began to cause transmissible disease on two different occasions in two different host species. These two isolates then evolved further in response to their particular host environments to become the current pathogenic subspecies of *M. avium*. The bird pathogen evolved into *M. avium* subsp. *avium* and *sylvaticum* while the ruminant pathogen evolved into the sheep and cow lineages of MAP.

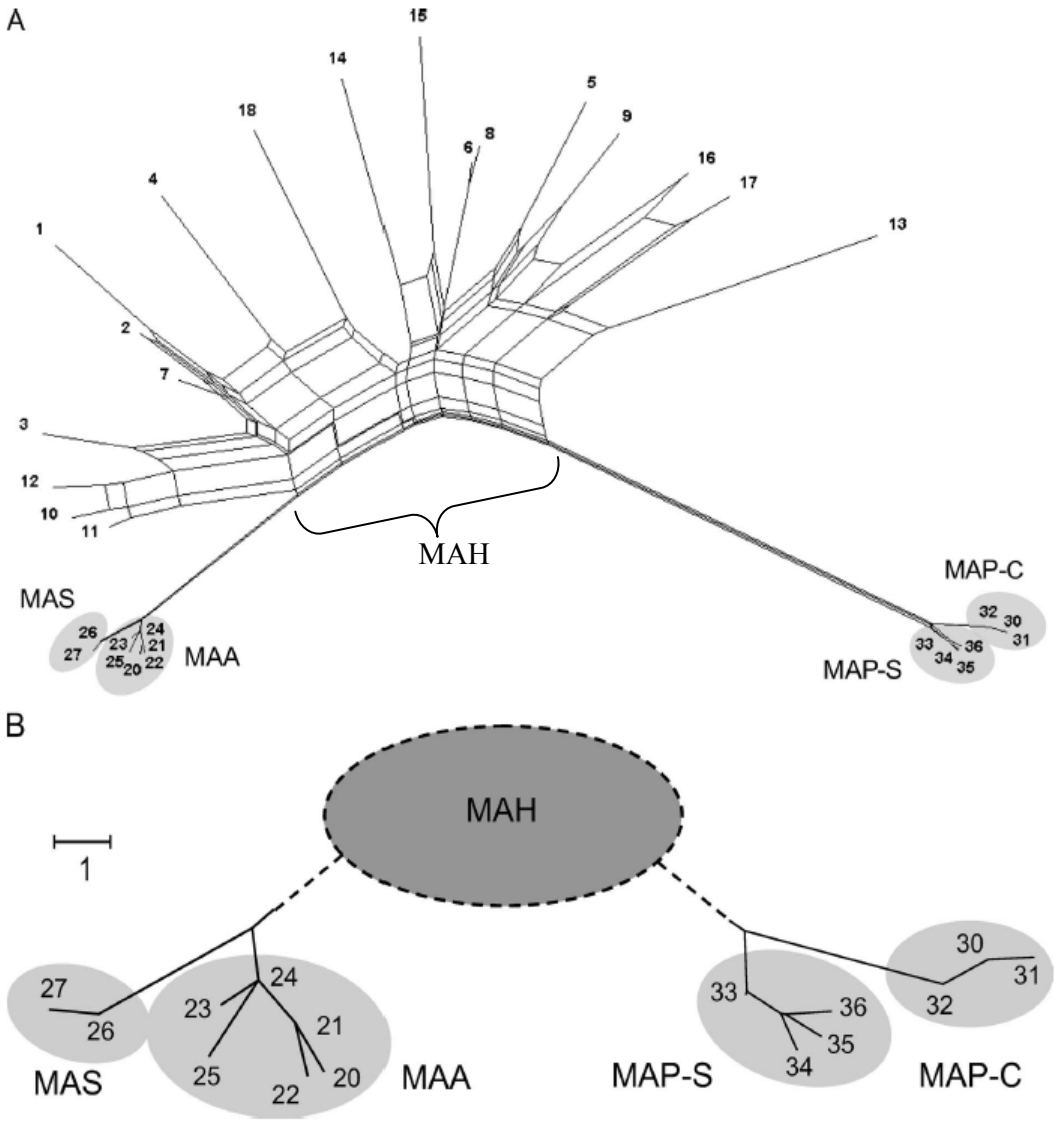


Figure 4: Phylogeny of Mycobacterium avium made with multi-locus sequence analysis (MLSA). A) Phylogeny of *M. avium* showing all sequenced isolates and recombination patterns. Ten housekeeping genes from 55 *M. avium* isolates were sequenced, and the sequences aligned. These included 24 isolates of *M. avium* subsp. *hominissuis*, 8 isolates of *M. avium* subsp. *avium*, 2 isolates of *M. avium* subsp. *sylvaticum* and 21 isolates of *M. avium* subsp. *paratuberculosis.* The web-like pattern within *M. avium* subsp. *hominissuis* indicates possible recombination events between the strains. B) Schematic diagram showing pathogenic *M. avium* as decendants of *M. avium* subsp. *hominissuis*. The scale shows one nucleotide of difference among all ten sequenced genes. MAS = *M. avium* subsp. *sylvaticum*; MAA = *M. avium* subsp. *avium*; MAH = *M. avium* subsp. *hominissuis*; MAP-S = sheep strains of *M. avium* subsp. *paratuberculosis*; MAP-C = cattle strains of *M. avium* subsp. *paratuberculosis* (77).

The alternative possibility was that one pathogenic clone of *M. avium* subsp. *hominissuis* evolved to parasitize a single type of animal. This pathogen then crossed the species barrier to infect a different host, resulting in two pathogenic subspecies. In this case, one would expect the pathogenic *M. avium* to be next to each other on the phylogenetic tree to reflect the sequence of evolutionary events. Instead, the pathogenic subspecies are separated by *M. avium* subsp. *hominissuis*.

The MLSA phylogeny also shows that the genetic diversity of *M. avium* subsp. *hominissuis* is about as large as the separation between *M. avium* subsp. *hominissuis* and MAP. From this, we hypothesized that some *M. avium* subsp. *hominissuis* strains already have most of the adaptations needed to parasitize animals, and need few genetic modifications to become professional pathogens. This hypothesis is consistent with the fact that *M. avium* subsp. *hominissuis* is able to parasitize amoebae and causes opportunistic infections in AIDS patients and pigs. We decided to investigate what these genetic changes might be.

1.2.3 Large Sequence Polymorphisms

MAP contains large genomic insertions and deletions that are not present in other *Mycobacteria*, called large sequence polymorphisms (LSP). These sequences are sometimes used to distinguish MAP isolates from other *Mycobacteria* using PCR (78). However, it was unknown which LSP were variably present in *M. avium* subsp. *hominissuis*, which our previous work showed is the nearest relative of MAP. Furthermore, the distribution of the LSP among different MAP strains was unknown. We hypothesized that the difference

in virulence between *M. avium* subsp. *hominissuis* and MAP was partly due to the presence of virulence genes in the insertion sequences (LSP^P). Therefore, we decided to identify LSP^P that were present in a large panel of MAP strains, but absent in a large panel of *M. avium* subsp. *hominissuis* (79). This was done by amplifying LSP^P using PCR and sequencing the PCR products (to test for their presence) and doing Southern blots (to confirm their absence). We demonstrated that nearly all MAP strains possess six conserved LSP^P comprising about 125kb, or about 2.5% of the entire genome, including 96 open reading frames (Figure 5). These are called LSP^P4, 11, 12, 14, 15 and 16.

LSP^P4 spans 15.3kb and contains a putative operon (MAP0851-0855) and a putative sequence of phage genes (MAP0853-0866). The only genes with predicted bacterial functions encode a metallophosphoesterase (MAP0859) and a truncated cytochrome c oxidase (MAP0861). LSP^P11 is 13kb and contains several phage-derived genes interspersed with hypothetical genes of no convincing similarity to annotated genes in the Kyoto Encyclopaedia of Genes and Genomes (KEGG). It is possible that the entire insertion sequence consists of the remnants of a phage that infected an ancestral MAP strain. The functions encoded by these two insertion sequences are not easy to predict based on bioinformatics alone.

LSP^P12 is 19.4kb long and contains two divergently transcribed operons. The operon from MAP2180-2188 is probably a biosynthetic operon, but its substrate is unknown. The adjoining operon from MAP2189-2194 resembles a mammalian cell entry (*mce*) operon, which were originally identified in

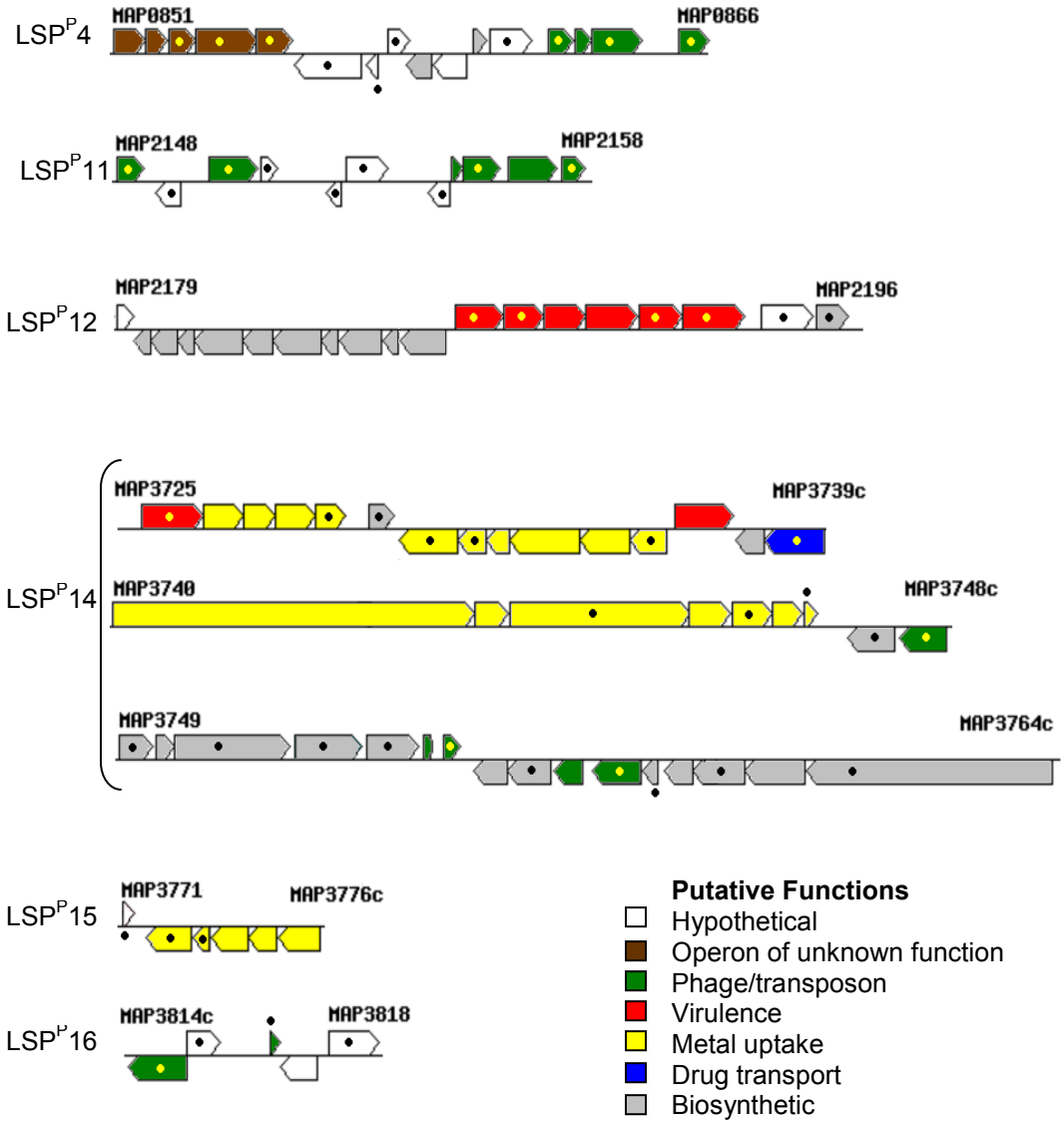


Figure 5: Conserved insertional large sequence polymporphisms (LSP^P) of Mycobacterium avium subsp. paratuberculosis. Putative functions were assigned by aligning the sequence of each gene to all sequences present in the Kyoto Encyclopedia of Genes and Genomes. Priority was given to the functions of predicted operons rather than individual genes. Genes upregulated in Zhu *et al.* are labeled with a black or yellow spot (the spot's colour is irrelevant) (79, 80).

M. tuberculosis (81). These operons are defined by their third gene, encoding *mceA*. Transformation of avirulent *Escherichia coli* with *mceA* alone allows it to invade mammalian cells. This gene is usually preceded by *yrbEA* and *yreEB*, which encode putative integral membrane permeases. *MceA* is followed by *mceB* to *mceF,* which encode proteins with N-terminal, hydrophobic signal sequences. The function of each gene is unknown and the operons are not required for *in vitro* growth. However, there is an emerging suggestion that they code for lipid transport systems, as the *mce4* operon was shown to be important for cholesterol uptake by *M. tuberculosis* in mouse alveolar macrophages (82). The cholesterol was used as a bacterial carbon and energy source. Gioffré *et al*. determined that at least three of the four *mce* operons of *M. tuberculosis* are important for virulence (83). *M. tuberculosis* mutants lacking *mce1*, *mce2* or *mce3* were severely defective in colonizing the lungs and causing pathology in mice. Mice infected with the wild-type strain died within 9 weeks of infection, but mice infected with any of the *mce* mutants never died of disease. The predicted *mce* operon of LSP^P12 encodes *mceA* to *mceF* but lacks *yrbEA* and *yreEB*. The proximity of the two LSP^P12 operons suggests that they might share a common transcriptional regulation system. This implies that the predicted biosynthetic operon has a role in virulence related to the *mce* operon, for example if the product of the former operon was secreted via the latter. MAP2196 encodes a dihyropicolinate reductase, but this is truncated in the sequence for the active site, so it is probably non-functional.

LSP^{P}14 is by far the largest insertion sequence, spanning 65.1kb. It consists of at least five predicted operons (Figure 5). The operon from MAP3726-3729 probably encodes a siderophore uptake system, while MAP3731-3736 probably encodes an inorganic metal uptake system. MAP3740-3746 encode a siderophore biosynthesis operon. MAP3758 encodes an AraC-type transcriptional regulator, which might regulate one of the LSP^{P}12 operons. MAP3725 and MAP3727 probably encode PPE (proline-proline-glutamate) genes. These are widely distributed among pathogenic *Mycobacteria*, and are defined by a 180-residue conserved N-terminal domain and a highly variable C-terminal domain. The function of these proteins is unknown, although they appear to be antigenic and expressed at the bacterial cell surface. Li *et al.* reported that a knock-out of a single PPE gene in *M. avium* subsp. *avium* had impaired growth in macrophage cell culture (84). Mice infected with the mutant had 3-logs fewer bacteria in the liver and 2-logs fewer bacteria in the spleen 3 weeks post infection compared to the wild type and complemented strains. It has been suggested that the numerous and diverse PPE proteins are part of an immune evasion strategy through antigenic variation. That is, the bacteria might secrete the PPE as decoys for the adaptive immune system. MAP3739c seems to be related to an ATP-binding cassette drug transporter found in numerous Actinobacteria. MAP3730 and MAP3738c both encode putative methyltransferases. MAP3748c encodes an integrase from a transposon found frequently in *M. avium*, IS*1110*. The remaining genes have uncharacterized biosynthetic functions or are cryptic genes from transposable elements or phages.

LSP^P15 is only 4.8kb away from LSP^P14, and is 5.4kb long. It encodes a siderophore-based metal import system, probably controlled by a Fur-like transcription factor (MAP3773c). MAP3771 is a fragment of a ribosomal gene, which are sometimes associated with horizontal gene transfer.

LSP^P16 is 6.7kb long and appears to be a transposon that is present in different genomic locations in members of *Mycobacterium*. Therefore, the sequence itself is not unique to MAP, but its location within the genome is. Some of the genes can be assigned transposon-related functions based on sequence homology, but others cannot.

The sequences of the LSP^P suggest that they are genomic islands acquired through horizontal gene transfer. Horizontal gene transfer is the integration of genetic material from one bacterial species into the genome of another species (85). In *Mycobacteria*, this mainly occurs through transduction, which involves phage-infected bacteria. When the phage is replicating, it can package segments of host DNA into the viral particle. The phage particle containing mixed phage/bacterial DNA then infects a new bacterium, sometimes of a different species than the original host. The foreign genetic material from the phage and original host then integrate into the new bacterial chromosome. As a result, the new host acquires genes from another bacterium, which might confer a survival advantage. Phage attachment sites are often identical to tRNA gene sequences, so phages preferentially integrate in tRNA genes. Genomic islands are often unstable because they are derived from phages. Additionally, their G/C content can differ from the rest of the genome because they originate in another species.

Pathogenicity islands are a special kind of genomic island that induces a "quantum leap" in a bacteria's virulence (86). The term is conventionally applied to genomic islands of Gram negative pathogens, such as the locus of enterocyte effacement (LEE) of enteropathogenic *Escherichia coli*. This locus contains seven of the bacteria's most important virulence factors, with functions ranging from invasion of non-polarized cells to interference of host signalling pathways. Without LEE, the bacterium is avirulent. Hacker *et al.* developed criteria to identify pathogenicity islands that can be applied to a wide variety of Gram negative and Gram positive bacteria (87). The features of the LSP^P are compared to this definition in Table 2. The similarity between pathogenicity islands and LSP^P14 led one group to dub LSP^P14 a pathogenicity island (88).

Table 2: Comparison of criteria defining pathogenicity islands to features of LSP^P

Definition of a pathogenicity island (87)	Features of the LSP^P
Carriage of virulence genes	Possibly, but no experimental evidence
Present in pathogenic strains, absent in non-pathogenic strains	Present in MAP, but not environmental *M. avium* subsp. *hominissuis*
Different G/C content relative to rest of genome	G/C Content of the entire genome: 69%
	LSP^P4: 60%
	LSP^P11: 61%
	LSP^P12: 66%
	LSP^P14: 64%
	LSP^P15: 64%
	LSP^P16: 62%
Large chromosomal regions, often greater than 30kb	LSP^P4: 15.3kb
	LSP^P11: 13kb
	LSP^P12: 19.4kb
	LSP^P14: 65.1kb
	LSP^P15: 3.4kb
	LSP^P16: 4.1kb
Contain "compact genetic units", i.e. functional cassettes and operons with little intergenic space	Many putative operons, see Figure 5
Association with tRNA genes and/or insertion sequence elements at their boundaries	LSP^P4 is at a tRNA gene
	LSP^P4 and 11 are flanked by direct repeats
Presence of (often cryptic) 'mobility' genes (insertion sequence elements, integrases, transposases, origins of plasmid replication)	Phage or transposable element genes present in all LSP^P except LSP^P15

1.2.4 Metal Uptake

Our bioinformatic analysis of the LSP^P revealed three putative metal acquisition operons. In LSP^P14, MAP3726-3729 has high homology to a siderophore-based iron uptake system in a soil Actinomycetes, *Frankia*. MAP3740-3746 is probably a biosynthetic operon for an iron siderophore. It is tempting to speculate that this siderophore interacts with the transport system encoded by MAP3726-3729. Between these two operons lies another operon, MAP3736c-3731c. MAP3731c is homologous to an ATP-binding protein in a common cobalt/nickel transport system. However, the other genes in the operon cannot be identified with bioinformatics, and little is known about cobalt/nickel transport in *Mycobacteria*. In LSP^P15, MAP3773c has high homology to *zur* (formerly known as *furB*), a zinc-responsive transcription factor in *M. tuberculosis*. Genes within and just outside of LSP^P15 are homologous to members of the *M. tuberculosis* Zur regulon (Table 3), suggesting that they are involved in zinc transport (89). Interestingly, the Zur regulon is scattered throughout the *M. tuberculosis* genome, but many of its MAP homologs are located in or around LSP^P15.

Host genetic studies of Johne's reveal that divalent cations are important to MAP survival *in vivo*. Polymorphisms in *slc11a1* (formerly *nramp1*) predispose cattle to Johne's (90). This gene encodes a divalent cation transporter present on the phagosomal membrane, and might be involved in removing metals from the infected phagosome. If this pump is non-functional, MAP probably has better access to the essential metals it needs to survive and multiply. However, in

Table 3: Homology between the Zur regulon of M. tuberculosis and genes in and around LSP^{P}15. Cobalamin is an organic compound found in vitamin B_{12} that coordinates cobalt, which is sometimes imported with zinc. ESAT-6 is a secreted protein of *M. tuberculosis* involved in the lysis of host cells. The presence of ESAT-6-like proteins and PPE in the zinc regulon suggest that there is a link between zinc acquisition and virulence.

Member of the *M. tuberculosis* Zur Regulon (89)	Homologous MAP genes in or near LSP^{P}15	Annotation
Rv0106	MAP3770 MAP3772	Cobalamin biosynthesis proteins
Rv0282	MAP3778	Uncharacterized ATPases
Rv2059	MAP3776c	Zinc/manganese transport system substrate-binding proteins
Rv0280	MAP3765	PPE
Rv2359	MAP3773c	Zur transcription factor
Rv3017 Rv3019	MAP3784	ESAT-6-like proteins

the presence of functional SLC11A1, the metal acquisition systems of LSP^{P}14 and 15 might help the bacteria to survive.

Mycobacterial zinc uptake is mediated by a variety of active transporters, though which of these is used during zinc starvation is unknown. Most of these transporters are not specific for zinc, importing manganese, iron or cobalt/nickel as well. Research about transcriptional circuits involving zinc is ongoing, which might lead to information about how the zinc is obtained in the hostile intracellular environment (89, 91).

MAP iron uptake is especially intriguing. MAP cannot be cultured from infected animals or faeces without adding an iron-binding siderophore called mycobactin J to the media (25). MAP bears a mycobactin J biosynthetic operon

(*mbtA-H*), but part of *mbtA* has been deleted, and there are nonsynonymous point mutations in *mbtE* and *mbtF* (79). Disruption of the same operon in *M. tuberculosis* severely reduced bacterial growth in macrophages within five days of infection and caused slow growth in iron-deficient media (92). Presumably, MAP acquired a more adaptive *in vivo* iron uptake system than mycobactin J, making the mycobactin J biosynthetic operon redundant. Studies comparing organic iron sources and their metabolism in MAP have not been done. However, mycobactin J can be replaced by abundant (1%) ferric ammonium citrate in broth culture (93). Growth in this medium slower than with mycobactin J. This suggests that MAP has highly specific iron uptake systems. Such a system might be encoded on one of the iron uptake operons of LSP^P14.

The observation that *in-vitro* but not *in-vivo* culture of MAP requires mycobactin J has not been explained. MAP does not have access to abundant free iron *in vivo*, and it is likely that it uses siderophores.

Experience with *M. tuberculosis* suggests that a MAP iron uptake system would be a promising target for antibiotics. Siderocalin, which binds to the iron siderophore carboxymycobactin, inhibits the growth of *M. tuberculosis* in macrophage cell culture (94). Conversely, increasing systemic iron concentrations increases bacterial replication in tuberculosis caused by *M. tuberculosis* or *M. avium* subsp. *avium* (95, 96). One of the earliest anti-tuberculosis drugs, *p*-aminosalicylic acid (PAS), is believed to target mycobactin synthesis (97). Therefore, if MAP has iron import systems to replace mycobactin,

they might be good targets for antibiotics too. This invites research on the metal uptake operons, especially the iron operons of LSP^P14.

1.2.5 The LSP^P and Transcriptomics

There have never been studies about the expression or function of LSP^P genes specifically. However, some expression data is available from transcriptome experiments. These suggest that the LSP^P are important in macrophage infection, acidic conditions, chemical stress and during excretion in the faeces.

Zhu *et al*. performed a technique called selective capture of transcribed sequences (SCOTS) to determine the MAP transcriptome during prolonged macrophage incubation (80). Briefly, SCOTS involves converting bacterial mRNA into cDNA. Then, a procedure involving selective hybridization and PCR is performed to reduce the number of very common cDNA, thereby enriching for rare species of cDNA. The cDNA are then cloned into *E. coli*, generating a library. The library is sequenced, giving a profile of the mRNA present in the bacteria. Zhu *et al.* compared the transcriptome for growth in media to the transcriptome for macrophage infection. Overall, the intra-phagosomal transcriptome of MAP resembled those of *M. avium* subsp. *avium* and *M. tuberculosis*. It appeared that the phagosome was DNA and cell wall damaging, iron deficient and fatty acid rich. Upregulated genes were involved in a number of metabolic and energy generating pathways, virulence, resistance to oxidative stress, and cell wall biosynthesis. Among the genes that were upregulated compared to growth in media were many LSP^P genes (Figure 5). Almost every

gene in LSP^P4 and LSP^P16 was upregulated. This is consistent with their predicted functions as a prophage and a transposon, respectively. These genetic elements are known to become active during periods of stress (85). Almost all of LSP^P11 was upregulated, though the functions of these genes could not be predicted through bioinformatics. In LSP^P14, a PPE gene (MAP3725) and all the putative metal transport operons were upregulated. The putative zinc-uptake system of LSP^P15 was expressed in macrophages. Interestingly, mRNA was detected for several, but not all genes in the putative operons. This is probably because the SCOTS technique is not totally sensitive, and many expressed mRNA were probably absent in the final library. These transcriptome results show that the LSP^P are expressed, and are probably important during prolonged macrophage infection.

Wu *et al.* determined the MAP transcriptome in various chemical stresses and in cow faeces using microarray technology (98). Unfortunately, most of the LSP^P genes were omitted from the microarray. Some of the included LSP^P genes were shown to be differentially regulated in cow faeces or in acidic media (Table 4). Interestingly, treating the bacteria with the decontamination agent alone changed the expression of many LSP^P genes, indicating that they might have a role in the response to chemical stress. These data suggest that the LSP^P serve many functions within the cell, despite being recent additions to the genome.

Table 4: Microarray data showing fold-change in RNA encoded by LSP^P genes from MAP in different conditions. Where no number is shown, no change in RNA-level was observed. HPC is hexadecylpyridinium chloride, a chemical used to decontaminate the faecal samples.

Gene Name	pH 5.5	1% HPC	Cow Faeces	Annotation
MAP0853	-693.68			Hypothetical
MAP0858	-126.86			Hypothetical
MAP2151		79.41	-734.10	Hypothetical
MAP2158			757.17	Hypothetical
MAP3733c		-248.79	292.04	Part of metal transport operon
MAP3734c			90.01	Part of metal transport operon. ABC membrane transporter
MAP3812c		-452.05	841.50	Peptidoglycan modification
MAP3818			-1170.52	Cytochrome P450

CHAPTER 2: RATIONALE AND OBJECTIVES

The current literature and findings from our lab have led us to hypothesize that the LSP^P are pathogenicity islands in *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Therefore, we hypothesize that certain genes of these LSP^P enhance bacterial survival in infected animals. To test this latter, more specific hypothesis we had two main objectives:

1) To generate knock-outs of LSP^P genes in the reference strain, *M. avium* subsp. *paratuberculosis* K-10.

2) To determine the ability of the LSP^P knock-outs to infect and persist in mice compared to the parental strain. A phenotype in mice would be a good basis for subsequent validation in natural hosts, such as cows.

Genetically manipulating MAP is a particular challenge because of its slow doubling time and resistance to current molecular biology techniques. Our

group's experience was that targeted gene disruption might not work, even after a year of attempts. To avoid this problem, we decided to generate a library of transposon mutants. Each member of the library would be a mutant in which a transposon had integrated into the genome at a random A/T site. The library could then be screened using PCR to identify mutants in which the transposon had integrated inside an LSP^P gene. After validating the mutant, it could be subject to *in vitro* and *in vivo* study, to determine the importance of the gene for growth in the laboratory and persistence in the host.

CHAPTER 3: MATERIALS AND METHODS

All reagents were supplied by Sigma-Aldrich (St. Louis, MO) unless otherwise indicated. All centrifugations were performed at 18,300 RCF (max speed) for 15 minutes unless otherwise indicated.

3.1 BACTERIAL STRAINS AND CULTURE CONDITIONS

All manipulations of *Mycobacterium avium* subsp. *paratuberculosis* K10 and *Mycobacterium smegmatis* mc^2 were performed using biosafety level 2 precautions. The strains were obtained from American Type Cell Culture (ATCC; Manassas, VA).

The broth media for *Mycobateria* was Middlebrook 7H9 (Becton, Dickinson and Co., Sparks MO) supplemented with 0.2% glycerol, 0.05% Tween-80, 1μg/ml mycobactin J (Allied Monitor, Fayette MO) and 10% albumin dextrose complex (ADC: 5% bovine serum albumin fraction V, 2% D-glucose

and 0.81% NaCl; Becton, Dickinson and Co.). Cultures were kept in 1-litre polystyrene storage bottles (Corning, Corning, NY) and incubated at 37°C, rolling at 0.2 RPM. Solid media for the transposon library was Middlebrook 7H11 (Beckton, Dickinson and Co.) supplemented with 1µg/ml mycobactin J and 10% oleic acid albumin dextrose complex (OADC: 5% bovine serum albumin fraction V, 2% D-glucose and 0.81% NaCl; Becton, Dickinson and Co.). *M. smegmatis* and the ΔMAP3776 mutant were subcultured on 7H10 agar (Becton, Dickson and Co.) with 5% glycerol and, like the 7H11, with 1µg/ml mycobactin J and 10% OADC. The mouse experiments also used 7H10 plates, with 2 vials/l of PANTA Plus for the isolating of MAP from livers (per vial: 10,000U polymyxin B, 1,000µg amphotericin B, 4,000µg nalidixic acid, 1,000µg trimethoprim, 1,000µg alzlocillin; Becton, Dickinson and Co.). PANTA was used for the liver only because it is more easily contaminated with gut flora than the spleen during mouse dissection. Media was supplemented with 25µg/ml kanamycin monosulfate when appropriate. All plates were incubated at 37°C. Photographs of colonies were taken with a Coldpix 4500 camera (Nikon, Tokyo).

The *Escherichia coli* used was a *pir*$^{+}$ strain, grown on Luria-Bertani (LB) media. LB broth contained 5% yeast extract, 10% NaCl and 10% Bacto-tryptone. LB plates were made by adding 15g/l of agar to the broth recipe. 50µg/ml of kanamycin monosulfate was added when appropriate. *E. coli* was incubated at 37°C, and broth cultures were incubated in a shaking at 150 RPM.

Measurement of optical densities was done at 600nm wavelength using the following spectrometer: Biochrom Novaspec Plus (General Electric, Piscataway,

NJ). Broth cultures with optical densities between 0.3 and 0.8 were considered in log-phase.

3.2 Mutant Library Generation

A stock of the ΦMycoMarT7 transposon was obtained from David Sherman, Seattle Biomedical Research Institute. This 2kb DNA sequence is capable of replicating and producing infective phages (sometimes called phasmids) in *Mycobacteria* at 30°C, but not at 37°C. It needed to be propagated in *Mycobacteria* to preserve the DNA and to produce phages. This was done according to the protocol of Murry *et al.* (99), which is briefly described below. Since the phage is inactivated by Tween, none of the broth media used in preparing the library contained Tween.

A log-phase culture of *M. smegmatis* mc^2 was washed twice in 7H9. It was then combined with a sample of the transposon in top agar (0.6% agar and 2mM $CaCl_2$) at 42°C. This mixture was poured onto an LB plate, and incubated at 30°C for 48h. At this time, a lawn of bacteria was visible on the plate containing phage-induced plaques. A single plaque was excised from the media, and crushed into mycobacteriophage buffer (MP buffer; 50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM $MgSO_4$, 2mM $CaCl_2$). Thus, the buffer contained a single "clone" of phage. *M. smegmatis* was used for simplicity and speed, since it has a fast doubling time (2-3h) compared to MAP (1-2 days).

Then, the phage-containing buffer was combined with another culture of *M. smegmatis* in top agar, and poured onto LB. The phage was allowed to

replicate until the plate had a "lacy" appearance because of high phage replication. The plate was flooded with MP buffer and rocked at 4°C overnight, and the buffer collected from the surface. This was then passed through a 0.2µm filter to remove any bacteria from the phage stock.

The number of infective phages per millilitre was determined by titering the phage stock with *M. smegmatis*. 250µl of stationary-phase *M. smegmatis* was added to warm top agar, poured onto an LB plate, and left to dry. Then, serial dilutions of the phage stock were made. 10µl of each dilution was spotted onto the surface of the top agar. The plate was incubated at 30°C for 2 days, and the number of plaques in each spot was counted. This was used to calculate the number of plaque forming units (PFU) per millilitre.

The MAP transposon library was created by concentrating 100ml of MAP broth at OD 0.8 into 5ml. This was combined with about 10^{11} PFU of phage and incubated for 4h at 37°C. 250µl aliquots were then plated onto 7H11 containing 25µg/ml kanamycin and grown for about 6 weeks until colonies appeared. The colonies were then picked and added to 2ml centrifuge tubes containing 1.5ml of 7H9 with Tween and kanamycin. These were stored in 9×9 freezer boxes, and incubated at 37°C.

3.3 High-Quality DNA Extraction from MAP

10ml of broth culture (OD 0.5-1) was centrifuged at 3,400 RCF for 15 min. The supernatant was discarded and the pellet resuspended in 0.5ml of TE-buffer, and transferred to a 2ml screw-cap tube. The bacterial suspension was

then incubated at 80°C for 20 minutes, then at 4°C for 5 minutes. 50µl of 10mg/ml aqueous hen's egg lysozyme was added, and incubated at 37°C for two nights. Then, 70µl of 10% sodium dodecyl sulfate (SDS) and 10µl of of 10mg/ml aqueous proteinase K were added, the tube was vortexed gently, and incubated at 65°C for 20 minutes. Following this, 100µl of 5M NaCl and 100µl CTAB/NaCl (0.1g/ml *N*-acetyl-*N,N,N*-trimethyl ammonium bromide and 41mg/ml NaCl) were added, the mixture was vortexed until milky, and incubated at 65°C for 10 minutes. 750µl of 24:1 chloroform/isoamyl alcohol was added, and the mixture was vortexed, then centrifuged at 18,300 RCF for 5 minutes. The upper phase was then transferred to a 1.5ml centrifuge tube. 10µl of 3M sodium acetate and 550µl of isopropanol were added, and the sample was inverted 6-8 times. The sample was incubated overnight at -20°C to precipitate out the DNA. Then, the sample was centrifuged, the supernatant removed, and the pellet resuspended in 9-30µl of water, depending on the subsequent experiment to be performed.

3.4 Transposon-Plasmid Preparation

The transposon contained a kanamycin resistance gene and an *E. coli* origin of replication (oriR6K). Therefore, it could be circularized to form a stable plasmid for *E. coli*. This feature was used to determine the location of the transposons within MAP mutants. Genomic DNA from a mutant was digested with restriction enzymes and circularized. This was then electroporated into *E. coli*, yielding colonies bearing plasmids of the transposon with flanking MAP

DNA. These plasmids were sent for sequencing to determine the flanking sequences of the transposon. The protocol is described below:

4µl of genomic DNA from a transposon mutant (≥ 200ng/µl) was digested overnight with either BamHI or SacII according to the manufacturer's instructions (New England Biolabs, Ipswich, MA). Then, a phenol/chloroform cleanup was performed to isolate the digested DNA. The volume of the sample was brought to 100µl with TE, and 100µl of phenol/chloroform/isoamyl alcohol (25:24:1) was added. The sample was inverted vigorously to mix the two phases, and then centrifuged. The upper phase was placed in a new 1.5ml centrifuge tube, and 100µl of chloroform/isoamyl alcohol (24:1) was added. The tube was inverted vigorously, and centrifuged. The upper phase was removed into a fresh tube, and 250µl of 100% ethanol and 10µl of 3M sodium acetate were added. The sample was incubated at -20°C for at least 20 minutes, and the centrifuged. The supernatant was removed, the pellet was dried, and resuspended in 9µl of water.

The digested DNA was then ligated overnight at 16°C with T4 ligase according to the manufacturer's instructions (New England Biolabs). Another phenol chloroform cleanup was performed to remove the salt and ligase, leaving 9µl of aqueous DNA.

3µl of the ligated DNA was then electroporated into 50µl of electrocompetent *E. coli*. This was done using a 2.0mm cuvette at 2.5kV, 25µF and 200Ω in a GenePulser XCell (BioRad, Hercules, CA). Immediately after electroporation, 450µl of LB broth was added to the cuvette, and it was incubated

at 37°C for 1h. Then, 200µl of the bacteria was plated on LB with kanamycin, and the plates were incubated at 37°C overnight.

The next day, individual colonies were inoculated into 5ml of LB broth with kanamycin, and incubated shaking at 37°C overnight.

The plasmids were subsequently extracted from the *E. coli* using a Qiagen minprep kit (Valencia, CA) according to the manufacturer's instructions. The plasmids were brought to the McGill University and Génome Québec Innovation Centre for sequencing using the primers TnMar_R and TnMar_OriR, which bind inside the transposon and face outwards in opposite directions (Table 6).

3.5 Library Screening

3.5.1 Generation of Template DNA

Using the transposon plasmid protocol, a yellow transposon mutant was identified as having a transposon in MAP3077, which encodes a putative transcription factor. DNA from this mutant was used to optimize the PCR protocol, and as a positive control in the polymerase chain reaction (PCR) protocols. The yellow colour was of no particular significance to the aims of the study. However, the colour allowed easy identification of the mutant in culture.

Boiled lysates were made from the library to generate template DNA for PCR. To make the boiled lysates, 200µl of each mutant from a row of mutants were pooled in a 2ml screw-cap tube (Figure 6). The same was done for every column of mutants (18 in total). The tubes were centrifuged, and the supernatant discarded. Then, the pellets were washed with 1ml of water each, and centrifuged

again. The pellets were resuspended in 200µl of water and incubated at 95°C for 20 minutes. The tubes were then cooled at 4°C for 5 minutes, and centrifuged. The supernatants (which were the row or column boiled lysates) were transferred to new tubes. "Box boiled lysates" were made by adding 100µl of each tube and column boiled lysate to a 2ml screw cap tube. Thus, the box boiled lysate contained DNA from all the mutants in an 81-member box of mutants. 1.81µl of boiled lysate from the ΔMAP3077 strain was introduced into each box boiled lysate as a positive control for PCR screening.

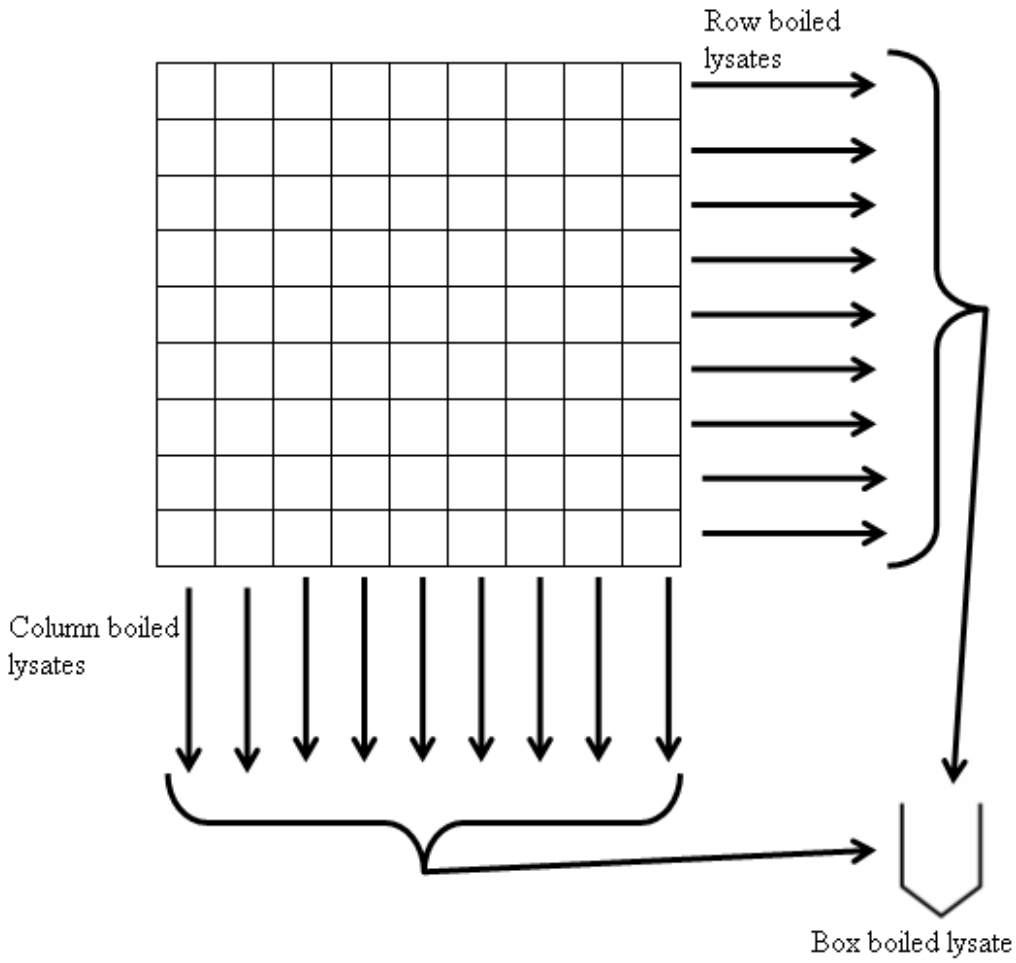


Figure 6: Pooling of mutants from the transposon library to generate boiled lysates for screening. The grid represents a 9×9-slot freezer box used to store the library. Culture from each row and column of mutants was pooled and made into boiled lysate. 100µl of each of these 18 boiled lysates was combined to make the box boiled lysate. Initial screening was performed on the box boiled lysate with the screening PCR protocol. If this round of screening produced a "hit" for one LSP^P gene, a second round of screening using the row and column boiled lysates was done. This would reveal the coordinates of the potential LSP^P gene mutant in the box.

3.5.2 Screening PCR Optimization

A PCR protocol was needed that could identify transposon mutants in the LSP^P from a box of boiled lysates. The original PCR protocol in our lab failed to detect DNA from the ΔMAP3077 mutant in the box boiled lysates. We optimized this protocol by altering the proportions of the standard ingredients, and by trying numerous additives. Primers to detect the positive control were designed to yield amplicons of about 200bp or about 1kb (Table 6, primers for screening optimization). The PCR was considered adequate when both the short and the long amplicons were generated, indicating the protocol had good sensitivity. The original and optimized screening PCR recipes are shown in Table 5.

The PCR program used was 3min at 94°C, followed by 35 cycles of: 94°C for 45s, 55°C for 45s, and 72°C for 1min. The final step was 72°C for 10min. The thermal cycler used was a DNA Engine Tetrad 2 (BioRad, Hercules, CA). The PCR products were then run on a 1% agarose gel containing 1μg/ml ethidium bromide at 100V for about 30min. The products were visualized using a Chemi Genius gel documentation apparatus (Syngene, Frederick, MD).

Table 5: Components of the original and the optimized screening PCR.

Reagent	Original Protocol	Screening Protocol
Boiled lysate	5µl	5µl
Taq buffer (Invitrogen, Carlsbad, CA)	1×	0.8×
Acetamide	5%	5%
dNTP (Invitrogen)	200µM	200µM
$MgCl_2$ (Invitrogen)	2.5mM	1.25mM
Each primer	500nM	500nM
Taq polymerase (Invitrogen)	0.75U	0.75U
Final volume	25µl	25µl

3.5.3 Execution of PCR Screening

Having optimized the PCR protocol, the library was screened for mutants of the LSP^P genes. This was done by performing 96 individual PCR for every box boiled lysate to screen for mutants in every LSP^P gene. For each PCR, one primer was designed to bind about 150bp upstream of the LSP^P gene of interest (Table 6). The second primer was the same for all screening PCR and bound inside the transposon at flanking inverted repeats (MycoMar in Table 6). The inverted repeats allowed the use of a single transposon primer, as the orientation of the transposon in the genome could be ignored. As a positive control, a PCR to detect the ΔMAP3077 DNA was performed for every box boiled lysate screen. A negative control containing all PCR components except for an LSP^P primer was also done. PCR screening was done in 96-well plates and 8-well strip tubes (UltiDent, St. Laurent, QC). Primers were designed using the online tool Primer 3 using the default settings, notably a 60°C melting temperature. Primer 3 is accessible at http://frodo.wi.mit.edu/primer3/.

The screening protocol was used to detect LSP^P mutants in box boiled lysates. If there were PCR products greater than or equal to the minimum size (Table 6), then verification PCR were done on the row and column lysates used to make the box boiled lysate for the candidate gene. If the row and column PCR indicated that a particular mutant might be a LSP^P mutant, then a boiled lysate was made of this mutant specifically, and the stringent "original" PCR was done using the relevant primers to confirm the screening results. Simultaneously, this mutant was subcultured in a larger volume of 7H9 broth. Once the mutant had grown enough, high-quality DNA was extracted and the stringent PCR was performed. PCR products were sequenced to ensure they were indeed the predicted amplicons. As an additional verification, the transposon plasmid protocol was performed on the mutant. The resultant plasmid was also sent for sequencing. The PCR product sequence and the transposon plasmid sequence had to match for a positive identification of an LSP^P mutant to be made. Sequencing was done with the primers TnMar_R and TnMar_OriR (Table 6).

Table 6: *Primers used in this study.* A minimum sized product would occur if the transposon landed just after the first base pair of the start codon.

Primer Name	Sequence (5´ to 3´)	Minimum product size (bp)
MycoMar	cgggacttatcagccaacct	–
MAP3077_1045	ttcacccctactcgttcacc	1045 (screening optimization)
MAP3077_928	ccagccgccgccggcagggat	928 (screening optimization)
MAP3077_209	cggaccgaactcgagaagttg	209 (screening optimization)
MAP3077_262	gcggtttggaagtcgatg	262 (screening optimization)
SigA_L	gtctgggacgaggacgaat	931 (screening optimization)
SigA_R	ttgacatcgtcttggactcg	
TnMar_R	cgcttcctcgtgctttacggtatcg	(Sequencing)
MnMar_OriR	cccgaaaagtgccacctaaattgtaagcg	(Sequencing)
MAP0851_prom_L	ccgaaacgtcacacatcaac	135
MAP0852_prom_L	gttgactttgatcccgaacg	143
MAP0853_prom_L	gtggtgaccgcgtagaagac	146
MAP0854_prom_L	aagcgaactgcttgagaagg	161
MAP0855_prom_L	aagaaattgcggtcgatgag	177
MAP0856c_prom_R	ctcattccaccaccttaccc	130
MAP0857c_prom_R	gagatcgtcgacggataagg	156
MAP0858_prom_L	gctcgaagatgtggaaggtg	159
MAP0859c_prom_R	tggaccgatccttactggac	100
MAP0860c_prom_R	atgacctccgctgaactgac	179
MAP0861_prom_L	cggccaacagatttagtcct	112
MAP0862_prom_L	ctgcttattgggctggtgac	136
MAP0863_prom_L	tgttgatcgctaagcctgtg	197
MAP0864_prom_L	ccgtattggactggctgatt	114
MAP0865_prom_L	agaaggagctgccgaagtatc	175
MAP0866_prom_L	gtactgcgccgagtctttg	166
MAP2148_prom_L	tgtggtaggcgttcacgtag	173
MAP2149c_prom_R	gctcaatggctcacaagga	124
MAP2150_prom_L	ggggcgtggtgtaaatagtg	106
MAP2151_prom_L	accctggccctattacatcc	167
MAP2152c_prom_R	cgacaatggcaacttttcct	176
MAP2153_prom_L	gactcacgcgaagcagtgta	153
MAP2154c_prom_R	cagccatactcgctgtcgta	119
MAP2155_prom_L	cctaagcgccatctcaacc	159
MAP2156_prom_L	gtgacggggtgactacgg	134
MAP2157_prom_L	gtcgggtatggctttcatgt	155
MAP2158_prom_L	tgagaatctccttcggtggt	152
MAP2179_prom_L	gcagggatgtcagcaacttt	104
MAP2180c_prom_R	gcgctcggtgtaccactact	137
MAP2181c_prom_R	ctgaatctgcaagccaatcc	194
MAP2182c_prom_R	gcagggttcacacatctgc	171
MAP2183c_prom_R	cggcgctgtatctggtct	80
MAP2184c_prom_R	tcatgtgcgaaacggactac	176
MAP2185c_prom_R	aacgtcaacaccaccgtctt	135
MAP2186c_prom_R	cacaagacaagaaagccgaag	138
MAP2187c_prom_R	aggagctgcacgaagtgg	131
MAP2188c_prom_R	cccgatccatagtcatccaa	162
MAP2189_prom_L	cggttggcttggatgactat	154
MAP2190_prom_L	gaattaccgctcgcctacaa	174
MAP2191_prom_L	tacggcagcttcgtcaacta	105
MAP2192_prom_L	cgaacggcttctactacaacg	158
MAP2193_prom_L	gcatcaacggtgactcgac	139
MAP2194_prom_L	tgatccctggttttcgaact	169

Table 6 (Cont'd)

Primer Name	Sequence (5´ to 3´)	Minimum product size (bp)
MAP2195_prom_L	tcggctttgagaggtagacg	128
MAP2196_prom_L	gggaggaagtcgaaatgacc	103
MAP3725_prom_L	gtcccgttgattaccacgat	96
MAP3726_prom_L	tactgggacgatgcgagaat	140
MAP3727_prom_L	agatcccgctaggcatcc	77
MAP3728_prom_L	gccctcagcgagctgtat	108
MAP3729_prom_L	cgaacgcttctcggatctaa	115
MAP3730_prom_L	agttggtacggggtcagctt	160
MAP3731c_prom_R	ttatagcgtcgagccttcgt	179
MAP3732c_prom_R	cgaagagatggaacggttgt	133
MAP3733c_prom_R	gccgatcagatactggtgct	171
MAP3734c_prom_R	gtattggtaatcgcccatcg	189
MAP3735c_prom_R	tccgggattgtcaagaagac	140
MAP3736c_prom_R	cagaggacgacaggactgct	160
MAP3737_prom_L	gcccatatcacgtcgaatct	136
MAP3738c_prom_R	atcttcggaattgcaggatg	108
MAP3739c_prom_R	tgtcgagctcatcattttcg	167
MAP3740_prom_L	atcctagctcatgcgtcacc	89
MAP3741_prom_L	gaccaaggtggtcgaacaga	154
MAP3742_prom_L	atttcagcttccgcgactac	198
MAP3743_prom_L	cgtggctactgtcttggtga	108
MAP3744_prom_L	cgttctggacagtggatgac	138
MAP3745_prom_L	aggggttacgcaacgattc	127
MAP3746_prom_L	gctggcgaccatttctacat	158
MAP3747c_prom_R	acccacgcttattgaaaacg	95
MAP3748c_prom_R	gtgaggctgcatgtgtcttg	175
MAP3749_prom_L	tccttaaagcagcgatcaca	188
MAP3750_prom_L	cactgaagctcggcatgtaa	108
MAP3751_prom_L	cctcagaaagtcgccaatgt	195
MAP3752_prom_L	gggtcggcaaattgtcttt	164
MAP3753_prom_L	gtctggacgtgtagggtgct	138
MAP3754_prom_L	ctttcggagttcgacctgac	149
MAP3755_prom_L	tcaaccacaagaagctccaa	125
MAP3756c_prom_R	gggctacacgacaccgtatt	186
MAP3757c_prom_R	gcgtgagatgtccgttttct	95
MAP3758c_prom_R	tcctctacgccgcagtttac	119
MAP3759c_prom_R	ggggcgtggtgtaaatagtg	107
MAP3760c_prom_R	ccgaaaggttacccagtcct	152
MAP3761c_prom_R	gcgaagtagccaaaaggatg	96
MAP3762c_prom_R	cggtggagtctcacctcaac	145
MAP3763c_prom_R	gctgggcaatctcgaactac	203
MAP3764c_prom_R	cgcactcaacatatcggatg	115
MAP3771_prom_L	gttttcgacgacgatccatt	111
MAP3772c_prom_R	tcgacatcgaagaacacacc	170
MAP3773c_prom_R	cctcgacagcgtcatctaca	131
MAP3774c_prom_R	agtcatcgccaaagcagtg	151
MAP3775c_prom_R	gaaccgacgctcgtagtgat	158
MAP3776c_prom_R	gtgcccgtcgtcaagataac	104
MAP3814c_prom_R	gtcgacttggatctgcacct	184
MAP3815_prom_L	attccgcaagcagaagtcc	161
MAP3816_prom_L	aacatgcacagcaggaactg	171
MAP3817c_prom_R	gcatcggccttatttcatgt	121
MAP3818_prom_L	ttgggtcagcaccacaatta	127

3.6 Ziehl-Neelsen Staining

Ziehl-Neelsen (acid fast) staining was done by applying broth cultures of MAP to glass slides. These were left to air dry, and then heat fixed with a Bunsen burner. The slides were flooded with carbol-fuschin for 5 minutes, and then washed with acid alcohol. Acid alcohol is an aqueous solution of 75% v/v methanol with 0.5% v/v concentrated HCl and 5g/l NaCl. The slides were then rinsed with water, and flooded with water for 5 minutes. Methylene blue was added for 1 minute as a counterstain. Slides were viewed with an Olympus BH2 microscope (Center Valley, PA) and photomicrographs were taken with a Coldpix 4500 camera with a Coldpix MDC lens (Nikon, Tokyo).

3.7 Mouse Infection

A competition experiment was performed between wild-type and ΔMAP3776 MAP in mice. The strains were grown to an OD of 0.5, and washed three times in 10ml PBS, and resuspended in 10ml of PBS. The OD was taken after the washes, and the number of bacteria estimated using the ratio OD 1 = 2×10^8 colony-forming units (CFU). Theoretically, the inoculum was a 55/45 mixture of mutant and wild-type MAP, with a total of 2×10^6 CFU in a 200µl of phosphate-buffered saline (PBS). We aimed to have an excess of the mutant in order to test whether the mutant became a minority species with increasing time *in vivo*. Serial dilutions of the inoculum were plated on 7H10 with and without kanamycin to determine the exact number of mutant and wild-type CFU in the inoculum.

Twenty-five 4 to 6-week old C57Bl/6 mice were infected with 200µl inocula intra-peritoneally. Five mice were sacrificed one week after infection to generate an early time point, as proof of productive infection. At this time, the livers and spleens homogenized using small tissue grinders (Kendall, Ridgefield, CT) and serial dilutions of the homogenates were made, and plated on 7H10. The liver plates all contained PANTA Plus, and half of them also contained kanamycin. The spleen plates contained no PANTA Plus, but half of them also had kanamycin. Four weeks after infection, the number of colonies on the spleen plates was counted. The liver plates were counted five weeks post infection, because the presence of PANTA Plus on the plates resulted in a slight delay in colony appearance. The colonies that grew on kanamycin were considered mutants, while the colonies that grew without kanamycin were a combination of wild-type and mutant bacteria. Further mouse sacrifices are planned for 6-weeks post-infection, 3-months post-infection, 6-months post-infection and 1-year post-infection.

CHAPTER 4: RESULTS

4.1 GENERATION OF A 5,000-MEMBER TRANSPOSON-MUTANT LIBRARY

Transposon mutants of *Mycobacterium avium* subsp. *paratuberculosis* K-10 were generated using the ΦMycoMarT7 transposon. 5,000 mutants were picked and stored in sixty-two 9×9 freezer boxes.

4.2 OPTIMIZATION OF SCREENING PCR

Using the transposon-plasmid method, a yellow MAP mutant was identified as having a transposon in the gene for a transcription factor, MAP3077. This mutant was used to develop a PCR protocol to screen the mutant library. After testing multiple reaction conditions and additives, we found that reducing the concentration of Taq buffer by 20% and the concentration of $MgCl_2$ by 50% permitted the detection of the ΔMAP3077 DNA consistently (Figure 7). Even the primer 1045bp away from the MAP3077 transposon allowed detection of the mutant. Unfortunately, there was non-specific banding in the negative controls. This indicated that lowering the salt concentration of the PCR also reduced the specificity of primer binding. Therefore, the screening PCR sacrificed specificity for sensitivity. Nonetheless, the presence of non-specific bands was deemed acceptable, since false positives could be eliminated in later stages of screening using more stringent PCR and the transposon-plasmid method.

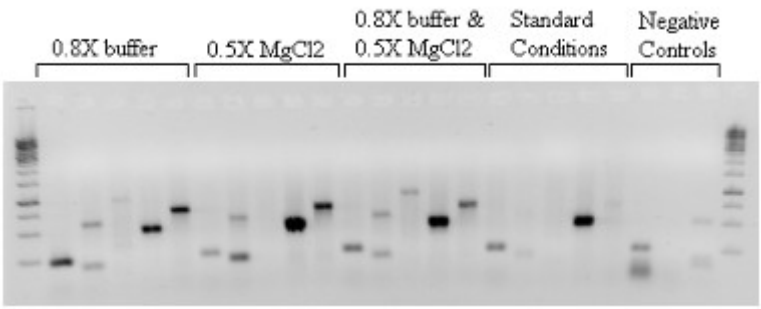
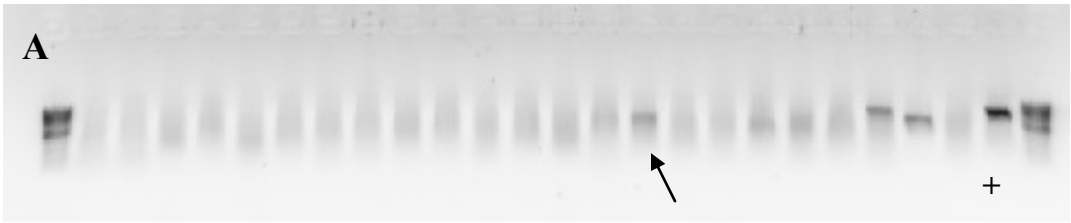


Figure 7: The optimized screening PCR compared to the standard PCR. Boiled lysate from the ΔMAP3077 strain was doped into a box boiled lysate (1% v/v). PCR were then performed to detect the ΔMAP3077 mutant, using different primer sets. The reaction conditions are listed above the gel. Using the standard PCR mix, the ΔMAP3077 strain could only be detected using one primer set. However, reducing the concentration of Taq buffer and $MgCl_2$ allowed the mutant to be clearly detected with all primers tested. The negative control contained box boiled lysate but no ΔMAP3077 DNA, and the reaction conditions were reduced buffer and reduced $MgCl_2$. Some non-specific bands were detected in the negative controls.

4.3 Detection of a ΔMAP3776 Mutant

Having optimized a sensitive PCR protocol, boiled lysates for the entire library were screened for LSP^P mutants. In box 38, the MAP3775c primer produced an amplicon, suggesting that a transposon mutant of MAP3775c or a nearby gene was in box 38 (Figure 8). Screening the row and column boiled lysates for box 38 indicated that this mutant would be in row "a" and column "1" of box 38. The mutant identified by PCR was transferred to a larger flask of 7H9 broth and grown. High-quality DNA extraction was performed, and the PCR for MAP3775c was repeated using a more stringent protocol. This PCR product was then sequenced by Génome Québec. Additionally, the transposon-plasmid technique was applied to the DNA. This produced plasmids bearing the transposon with the flanking regions of MAP DNA in *E. coli*. These plasmids



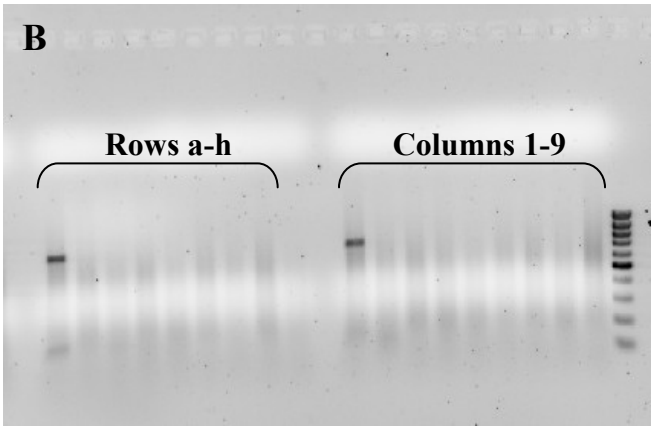


Figure 8: Detection of a transposon mutant near MAP3775c by PCR. A) Subset of PCR done on the box boiled lysate for box 38. The arrow indicates a PCR product for the primer binding upstream of MAP3775c. The plus-sign indicates the positive control for the ΔMAP3077 mutant. The other two bands were later found to be false positives. B) PCR for pools of DNA within box 38 using the MAP3775c primer. A PCR product was obtained in row "a", column "1", indicating that a particular mutant was a potential mutant of MAP3775c or nearby gene.

were extracted and sequenced by Génome Québec. The sequences of the PCR product and the transposon-plasmid indicated that the mutant in box 38, row "a", column "1" had a transposon just after nucleotide 4219501 in the MAP genome, inside MAP3776c, in $LSP^{P}15$ (Figure 9). MAP3776c encodes a putative zinc siderophore, and is part of a putative zinc uptake operon (MAP3776c-3772c). The transposon is inside the predicted dimerization domain, in a region that shows high homology to periplasmic binding proteins for metal ions in numerous bacteria.

START →
4220100
atggcgacggtgggcgcatgttgccgctcgggcgggccagccggggtcccgcggaaccgggccggcgcgcaccgcgacgcggaccgcgtttccggaa
agtctcgcgcattagatatgaaaacggttatcgtttgcggtgtgggtgggttgctctcgcggtggcacggcgcccctatcgctactgtcctggtgctgacagtcgt
aacgagttgctcgtcaagcccgacccaaacagccgaggggtcccgtaatgccgcctcgccatcagccattggcagaacggccgcgccgccgtgtcccac
cgccccactcgctgtcgtggtcagcgtggatcagtggggcgacattgtctccgagctgggcggcgcctgcgcgaacgtcaagaccgtgctcgccagctcgtc
ggtcgacccgcacgattacgagccgtcaccggccgacgccgctgactttatgaacgcgaaacttatcgtggtcaacggtgcagggtacgactcgtgggcat
ccaagctggccggcagctccgcttccggtgccccccttggtgagcgccgcggcagtcacgacaacaccggatggagccaaccctcacctgtgg
4219501
ACAGGTTGGCTGATAAGTC*CCCGGTCTCTAGACCCCTATAGTGAGTCGT*ATTACAAACTGGAACAACACTC
AACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGC
TGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGA
AATGTGCGCGGAACCAGATCCCCACGATGCGTCCGGCGTAGAGGATCTGAAGATCAGCAGTTCAACCTGT
TGATAGTACGTACTAAGCTCTCATGTTTCACGTACTAAGCTCTCATGTTTAACGTACTAAGCTCTCATGTTTA
ACGAACTAAACCCTCATGGCTAACGTACTAAGCTCTCATGGCTAACGTACTAAGCTCTCATGTTTCACGTAC
TAAGCTCTCATGTTTGAACAATAAAATTAATATAAATCAGCAACTTAAATAGCCTCTAAGGTTTTAAGTTTTAT
AAGAAAAAAAAGAATATATAAGGCTTTTAAAGCTTTTAAGGTTTAACGGTTGTGGACAACAAGCCAGGGATG
TAACGCACTGAGAAGCCCTTAGAGCCTCTCAAAGCAATTTTCAGTGACACAGGAACACTTAACGGCTGACA
GAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTC
TTAGACGTCAGGTGGCACTTTTCGGGGTTATAAGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGT
CGACCTCGAGGGGGGGCCCGGTACCGAGGACGCGTCGAATTAATTCCGCTAGCTTCACGCTGCCGCAAG
CACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGA
CCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGCAGGTAG
CTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCC
AGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGG
ATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAG
ATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGAC
AATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACC
GACCTGTCCGGTGCCCTGAATGAACTCCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGC
GTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGC
CGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCG
GCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCA
CGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCA
GCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGAT
GCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTG
TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGG
CTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT
GACGAGTTCTTCTGAGCGGGACTCTGGGGTACGCG*TAATACGACTCACTATAGGG*TCTAGAGACCGGGGA
CTTATCAGCCAACCTGT
4219502
tacctgccctccgcggtgaccgccgttgctgatgcggtaacgcaagaattgagcagaatggaaccgcctgccgccggctacttcagccagcggagagcgc
aattcaccagtgcgacaaggctatacgtgaacctgatcgccaagatcaaggccgaggcggcgggtaaatcctatggggccacagagacggtctttgacta
ccaggcgcaagcggcgggcctggtcaacaagacccctgccgggtaccggcgggcatcggccaacgaatcggaaccctcacccggagacgttgacgcg
ttcctcaccgcccttgccggccggcatatcgacctcttgatctacaacacgcagaccgagggctcaatccccgaggaaatccgatcagccgcggagcagtc
aagcgtgcccgtcgtcaagataaccgagaccgtaccgcccggggagacgtcgtttgaagactggcagtacgggcagctcgtccaactggccaaggcgct
ccatgtcgctgtctga
4218991
END

Figure 9: The sequence of MAP3776c disrupted by the ΦMycoMarT7 transposon. The entirety of MAP3776c containing the transposon is shown. Lower-case letters indicate nucleotides from MAP3776c, while the upper-case letters indicate nucleotides from the transposon. Numbers indicate the standard nucleotide number in MAP. Hilighting shows the features of the transposon: blue = inverted Himar repeats; gray = ori6K, yellow = neomycin phosphotransferase. Red, italicized letters show outward-facing T7 promoter sites. The Genebank accession numbers for MAP3776c and the transposon are AE016958.1 and AF411123, respectively.

4.4 THE ΔMAP3776C MUTANT RESEMBLES WILD-TYPE MAP *IN VITRO*

The ΔMAP3776c strain was tested for *in vitro* defects to determine if the disrupted gene was important for basic bacterial survival. The parental and mutant strains had indistinguishable bacterial morphology by acid fast Ziehl-Neelsen staining (Figure 10) and colony morphology on 7H10 plates (Figure 11). The mutant grew at the same rate as the wild-type in normal 7H9 broth (Figure 12).

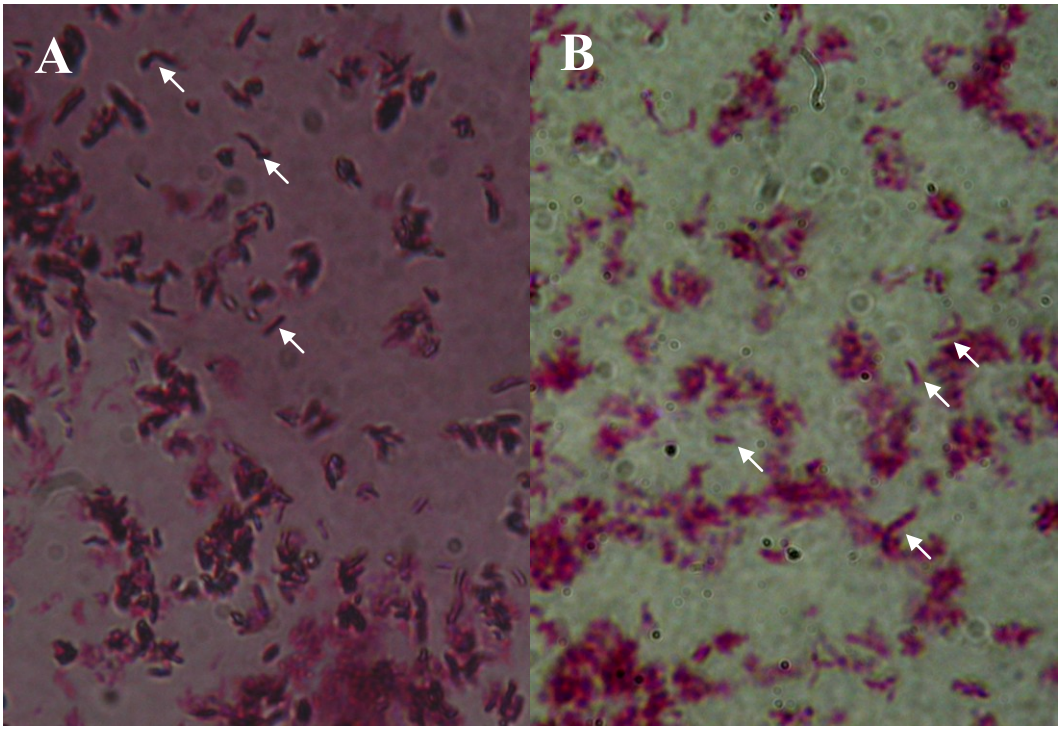


Figure 10: The ΔMAP3776 mutant is indistinguishable from wild-type MAP by Ziehl-Neelsen staining. Ziehl-Neelsen staining was done on liquid cultures of the wild-type and mutant strains, and viewed under oil immersion (1000×). A) Wild-type MAP B) ΔMAP3776 bacteria. Typical bacilli are indicated with arrows. The difference in colour is an effect of the staining procedure and has no effect on bacterial morphology.

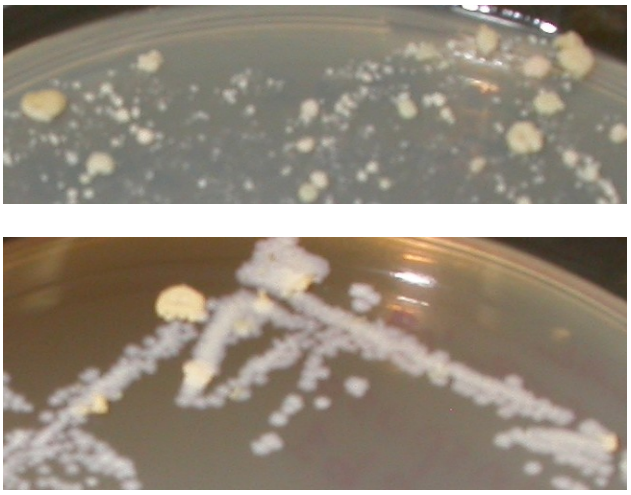


Figure 11: ΔMAP3776c colonies are indistinguishable from wild-type colonies on 7H10 agar. Above, ΔMAP3776c colonies; below, wild-type colonies. MAP colonies are irregular in form, beige/cream in colour, with irregular margins and uneven elevation. After several months, they grow into brittle 3-dimensional clusters. Colony size varies from punctuate to several centimetres in diameter.

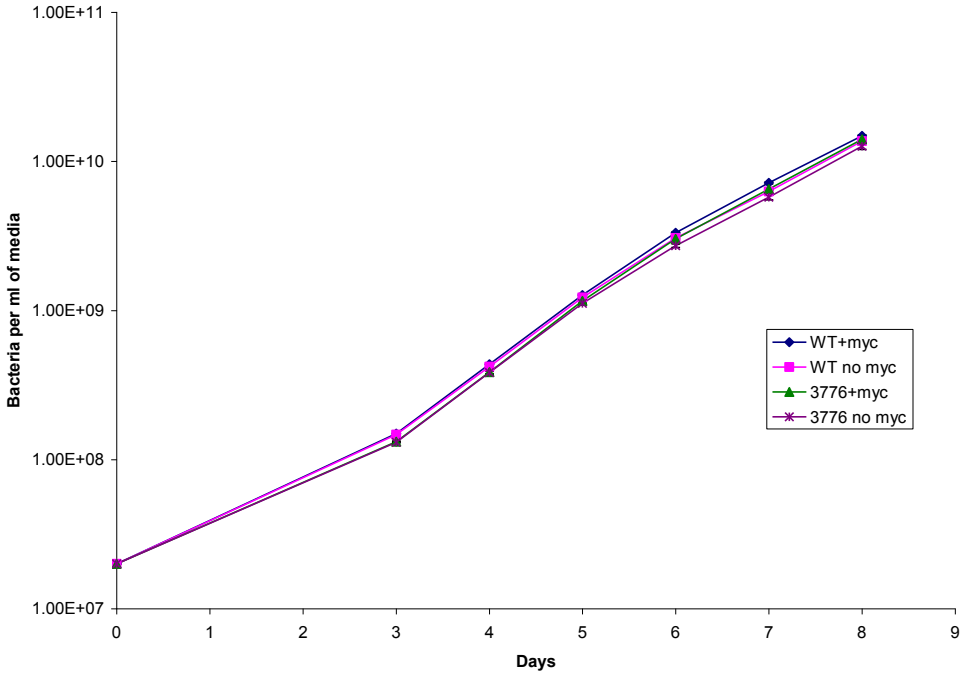


Figure 12: The ΔMAP3776c strain grows at the same rate as wild-type MAP in vitro. The wild-type and ΔMAP3776c were incubated in 7H9 with and without the iron siderophore, mycobactin J. The bacteria were kept in log-phase by continually diluting them, and optical density at 600nm was measured over eight days. The reported "bacteria per ml" takes into account the dilutions. The wild-type and mutant strains grew at the same rate in all media tested. This indicates that MAP3776c was not required for *in vitro* growth. Moreover, MAP3776c did not appear to contribute to iron import through or independently of mycobactin J *in vitro*.

4.5 Preliminary Mouse Data

Twenty-five C57Bl/6 mice were infected intraperitoneally with a mixture of wild-type and ΔMAP3776c MAP. The colony counts for the inoculum were 8.39×10^5 (SD±1.94×10^5) total bacteria, including 5.04×10^5 (SD±4.2×10^4) kanamycin resistant colonies. This is roughly 60% mutant and 40% wild-type. At 7 days post-infection, the mice were sacrificed, and the livers and spleens plated on 7H10 with and without kanamycin. The colony counts are shown in Tables 7 and 8. The statistical variability of the data is too great to draw meaningful conclusions at this time. Five mice will be sacrificed at each of the following time points: 6 weeks, 3 months, 6 months and one year post-infection.

Table 7: MAP colonies in mouse spleens 7-days post infection ($\times10^2$). Spleen 2 was an unsuccessful infection, as almost no bacteria could be grown from the spleen. SD is standard deviation.

	Spleen 1		Spleen 3		Spleen 4		Spleen 5		Mean	SD
+ Kan.	27	37	79	80	44	30	68	97	58 (42%)	27
No Kan.	96	98	192	172	77	93	131	157	139 (100%)	42

Table 8: MAP colonies in mouse livers 7-days post infection ($\times10^2$). Liver 2 was an unsuccessful infection, as almost no bacteria could be grown from the liver. SD is standard deviation.

	Liver 1		Liver 3		Liver 4		Liver 5		Mean	SD
+ Kan.	76	100	49	48	52	22	60	111	65 (75%)	29
No Kan.	211	154	41	31	29	48	99	83	87 (100%)	65

CHAPTER 5: DISCUSSION

The current literature and research from our own group led us to hypothesize that the insertional large sequence polymorphisms (LSP^P) are pathogenicity islands in *Mycobacterium avium* subsp. *paratuberculosis* (MAP). This hypothesis entails that some genes within the LSP^P are virulence genes (Table 2). To test this hypothesis, we generated mutants of LSP^P genes, starting by generating a library of 5,000 transposon mutants using the φMycoMarT7 transposon. A PCR screening process was developed using a validated positive control, and was applied to the mutant library. In this way, we found a mutant of MAP3776c, encoding a putative zinc siderophore in LSP^P15. As a preliminary experiment about the importance of MAP3776c in virulence, we infected C57Bl/6 mice with a combination of wild-type and ΔMAP3776c (kanamycin resistant) bacteria. If MAP3776c is important for MAP pathogenesis, then the wild-type bacteria would probably out-compete the defective mutant bacteria. This would result in a reduction of the number of kanamycin resistant bacteria compared to wild-type bacteria cultivated from the mouse livers and spleens over time. This is not clearly observed in the colony counts one week post-infection (Tables 7 and 8). The statistical variability of the data is too great to draw conclusions about the comparative ability of the mutant to colonize mice relative to the wild-type. However, it is clear that the ΔMAP3776c strain is able to colonize the liver and spleen, and persist for at least one week, indicating that MAP3776c is dispensable for early mouse infection. It is possible that data from mice sacrificed at 6 weeks,

3 months, 6 months and/or one year post-infection will reveal a defect in the mutant compared to the wild-type.

If the mouse experiments show the ΔMAP3776c strain to be attenuated, the next step would be to complement the knock-out to try and rescue the phenotype. MAP3776c is the first gene in an operon that ends with MAP3772c (Figure 5). This means that polar effects are possible, in which inactivation of the first gene of the operon prevents proper transcription of the downstream genes. The transposon contains outward facing T7 promoters, which some researchers believe limit polar effects by strongly and constitutively driving transcription of downstream genes (100). However, T7-based expression systems are not used in *Mycobacteria*, implying that the promoter is relatively weak or inactive in the genus. Quantitative real time PCR should be performed to detect transcription of LSP^P15 genes to determine if the transposon causes polar effects. Depending on the results of this experiment, complementation should be attempted for either MAP3776c alone (1.1kb) or the entire operon (4.4kb). This can be achieved using the pMV261 or pMV361 expression vectors, which use the strong, constitutive, mycobacterial *hsp60* promoter. If complementation restores virulence in mice, MAP3776c should be considered a virulence factor, since it would be dispensable for normal *in vitro* growth (Figures 10 and 11), but would enhance bacterial survival or replication in an animal. By extension, LSP^P15 should be thought of as a pathogenicity island, since it would meet the definition proposed by Hacker *et al*. (Table 2).

Scientifically, LSP^P15 is an ideal insertion element to disrupt in MAP because it is only 5.4kb long and encodes a single operon. This means that the operon has limited genetic "baggage" of irrelevant DNA that arrived in the genome at the same time as more important, adaptive genes. If our hypothesis and bioinformatic predictions are correct, then the operon assists in obataining zinc from the macrophage phagosome. By disrupting the siderophore gene of the operon, we have probably eliminated the operon's main *raison d'être*. Because LSP^P15 contains only one operon, the function of the entire insertion sequence was likely impeded. Therefore, the ΔMAP3776c strain will allow conclusions to me made about an entire LSP^P, which directly addresses the hypothesis about the LSP^P being pathogenicity islands.

The small size of LSP^P15 means that it can be cloned into a vector like pMV261/361 and expressed in *M. avium* subsp. *hominissuis*. Ectopic LSP^P15 might allow this normally avirulent bacterium to surivive better in macrophage cell culture. If *M. avium* subsp. *hominissuis* can be made more virulent by adding LSP^P15, this would be good evidence that the genomic island is in fact a pathogenicity island. Optimization of macrophage experiments began at the same time as mouse experiments and is ongoing.

The fact that MAP3776c encodes a putative secreted protein makes it a particularly interesting candidate for further study. Experiments can be done to detect the protein in culture supernatant or in infected animals. The protein might be immunogenic, and a good candidate for subunit vaccine development. Alternatively, host antibodies to the MAP3776c-encoded protein might be useful

in immunodiagnostics for Johne's disease. If MAP3776c is a secreted protein, it is an attractive candidate for biochemical and immunologic studies.

If MAP3776c is not immunogenic but important for long-term persistence, then the deletion strain would be a good starting point for a live, attenuated vaccine. Preliminary data from this project show that the mutant is able to persist for at least a week in mice, meaning it should be able to prime the immune system. If MAP3776c is important for long-term persistence, then the mutant should die out on its own from zinc starvation after having conferred immunity. However, there is a danger that the vaccine strain would regain virulence through secondary mutations that compensate for the loss of MAP3776c. *M. bovis* BCG is a safe, live vaccine strain because it is missing multiple virulence factors due to large genomic deletions (71). To attenuate MAP to the level of *M. bovis* BCG, several genes apart from MAP3776c would have to be disrupted. This could be done by repeating the transposon mutagenesis protocol described here on the ΔMAP3776c strain using a transposon with a hygromycin resistance gene. A multiple deletion strain of MAP K-10 would be a good basis for a live, attenuated vaccine.

It is formally possible that the wild-type and ΔMAP3776c strains of MAP are equally capable of parasitizing C57Bl/6 mice. Based on this data alone, it would be premature to conclude that MAP3776c has no role in MAP virulence. If MAP3776c encodes a siderophore, there could be complementation *in trans*, in which the wild-type bacteria secrete enough siderophore for all the MAP in the mouse. This would obscure any growth defect of the mutant. Complementation

in trans can be resolved by infecting some mice with the wild-type strain, and others with the mutant strain, and comparing the change in bacterial numbers over time between the two cohorts. Another problem is the mouse model is not a perfect representation of Johne's disease, as discussed in the introduction. To address this issue without resorting to cattle infections, preliminary studies can be conducted in macrophage cell culture, such as the murine cell line, Raw264.7. Focusing on the most important cell type might reveal a phenotype that is not apparent when live mice are infected. Another option is bovine macrophages, which are available from specialized laboratories, but not the American Type Cell Culture (ATCC) repository. In the absence of experiments on livestock, bovine macrophages provide a good way of obtaining data from the relevant host species. Further experiments in mice and macrophage cell culture should be done regardless of the outcome of the mouse experiment in progress.

An important methodological finding of this project was that for DNA extractions, MAP had to be incubated in lysozyme for 48 hours rather than the 2-24 hours needed for *M. tuberculosis*. The extra incubation improved DNA yield from 0-0.2µg/µl to 0.8-1µg/µl. This is of interest to the MAP community, for which DNA extraction is a constant source of frustration. This observation has also led to the hypothesis that MAP is more resistant to lysozyme than other *Mycobateria*, which might be important for its survival in macrophages. Preliminary results suggest that the MAP cell wall is inherently more resistant to lysozymal degradation than other *Mycobacteria*, despite the fact that its slow replication time should increase susceptibility to lysozyme (data not shown).

Interestingly, Koo *et al.* showed that lysozyme was important for resistance to mycobacterial infection in *Drosophila* (101), and Redacliff *et al.* suggested that polymorphisms in the lysozyme gene conferred susceptibility to Johne's in sheep (102). Together, these data clearly invite more research about MAP and lysozyme.

Theoretically, many mutants of LSP^P genes could have been found in the 5,000-member library. Since the LSP^P occupy 2.5% of the MAP genome, the library should have contained nearly 125 LSP^P gene mutants (if we ignore the small intergenic regions). There is no reason to believe the transposon integrated less frequently in the LSP^P because they were from a "foreign" source. We did not detect more mutants of interest probably because of problems in the PCR screening method. The quality and concentration of boiled lysate DNA is highly variable, meaning that the PCR protocol was not optimized for every box. As a result, the positive control often failed. An additional problem was that the low-salt PCR mix meant that there were many non-specific bands, sometimes over a dozen per screen. There was not enough boiled lysate DNA to investigate all the potential "hits", meaning that some promising leads had to be abandoned. Investigating non-specific PCR products also reduced the efficiency of screening. Sometimes, non-specific amplification produced DNA smears for every included primer set, preventing mutants in those boxes from being identified. Further PCR optimization could be performed to improve the yield of LSP^P mutants obtained from screening. Alternatively, several hundred mutants could be transferred to larger flasks of media and grown. This would permit high-grade DNA extractions

from the mutant strains, and subsequently the transposon-plasmid method of determining the transposon insertion sites. Statistically, only 182 mutants need to be successfully screened in this way to have a 99% chance of finding a single LSP^P mutant. Despite their technical difficulty, the protocols needed for this screening procedure are very effective and reproducible in our lab. This method would have the side-benefit of identifying hundreds of MAP mutants that might be useful in other projects about the genomics and pathogenicity of MAP. Having yielded a useful MAP mutant, the mutant library is still a useful tool to obtain more bacterial deletion strains.

As mentioned in the introduction, little is known about zinc transport in *Mycobacteria*. Therefore, the mutant created in the project might be a good tool to investigate the importance of zinc to mycobacterial survival, and the effects of zinc deprivation on the biochemical and transcriptional profile of the bacterium. This is a promising area of future study.

If MAP causes Crohn's disease, then antibiotics targeting MAP zinc uptake might be therapeutically beneficial. In the very long term, and pending more information about the cause of Crohn's, the ΔMAP3776c mutant might be useful in the development of such drugs. Therefore, this research might have links to a human health concern.

CHAPTER 6: CONCLUSIONS AND SUMMARY

In this project, a library of 5,000 transposon mutants of *Mycobacterium avium* subsp. *paratuberculosis* K-10 was generated. A PCR screening method to detect mutants of the insertional large sequence polymorphisms (LSP^P) was developed. Using this method, a mutant was obtained in MAP3776c in LSP^P15, encoding a putative zinc siderophore. A competition experiment between the wild-type and the ΔMAP3776c strains in C57Bl/6 mice was initiated. Preliminary results from one week post infection show that the ΔMAP3776c strain is able to colonize livers and spleens. Results to come might suggest a role for MAP3776c and, by extension, LSP^P15 in virulence. In this case, LSP^P15 should be regarded as a pathogenicity island of *M. avium* subsp. *paratuberculosis*. The ΔMAP3776c deletion strain and the transposon library are useful tools for further research. In the long run, they might contribute to improvements in vaccines and diagnostics for Johne's disease.

REFERENCES

1. **Ott SL, Wells SJ, Wagner BA. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. *Prev Vet Med* 40: 179-92**
2. **Tiwari A, VanLeeuwen JA, McKenna SL, Keefe GP, Barkema HW. 2006. Johne's disease in Canada Part I: clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. *Can Vet J* 47: 874-82**
3. **Wells SJ, Wagner BA. 2000. Herd-level risk factors for infection with Mycobacterium paratuberculosis in US dairies and association between familiarity of the herd manager with the disease or prior diagnosis of the disease in that herd and use of preventive measures. *Journal of the American Veterinary Medical Association* 216: 1450-7**
4. **Barkema HW, Hesselink JW, McKenna SL, Benedictus G, Groenendaal H. 2010. Global prevalence and economics of infection with *Mycobacterium avium* subsp. *paratuberculosis* in ruminants. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 10-21. Wallingford, U.K.: CABI**
5. **Sweeney RW. 1996. Transmission of paratuberculosis. *Vet Clin North Am Food Anim Pract* 12: 305-12**
6. **Fecteau M, Whitlock RH. 2010. Paratuberculosis in cattle. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 144-56. Wallingford, UK: CABI**
7. **Gilmour NJ, Nisbet DI, Brotherston JG. 1965. Experimental oral infection of calves with Mycobacterium johnei. *J Comp Pathol* 75: 281-6**
8. **Chiodini RJ. 1996. Immunology: resistance to paratuberculosis. *Vet Clin North Am Food Anim Pract* 12: 313-43**
9. **Whittington RJ, Fell S, Walker D, McAllister S, Marsh I, Sergeant E, Taragel CA, Marshall DJ, Links IJ. 2000. Use of pooled fecal culture for sensitive and economic detection of mycobacterium avium subsp. paratuberculosis infection in flocks of sheep. *J Clin Microbiol* 38: 2550-6**
10. **Momotani E, Whipple DL, Thiermann AB, Cheville NF. 1988. Role of M cells and macrophages in the entrance of Mycobacterium paratuberculosis into domes of ileal Peyer's patches in calves. *Vet Pathol* 25: 131-7**
11. **Hostetter J, Steadham E, Haynes J, Bailey T, Cheville N. 2003. Phagosomal maturation and intracellular survival of Mycobacterium avium subspecies paratuberculosis in J774 cells. *Comp Immunol Microbiol Infect Dis* 26: 269-83**
12. **Kuehnel MP, Goethe R, Habermann A, Mueller E, Rohde M, Griffiths G, Valentin-Weigand P. 2001. Characterization of the intracellular survival of Mycobacterium avium ssp. paratuberculosis:**

phagosomal pH and fusogenicity in J774 macrophages compared with other mycobacteria. *Cell Microbiol* 3: 551-66

13. **Souza CD, Evanson OA, Weiss DJ. 2006. Mitogen activated protein kinase p38 pathway is an important component of the anti-inflammatory response in Mycobacterium avium subsp. paratuberculosis-infected bovine monocytes. *Microb Pathog* 41: 59-66**
14. **Tooker BC, Burton JL, Coussens PM. 2002. Survival tactics of M. paratuberculosis in bovine macrophage cells. *Vet Immunol Immunopathol* 87: 429-37**
15. **Wu CW, Livesey M, Schmoller SK, Manning EJ, Steinberg H, Davis WC, Hamilton MJ, Talaat AM. 2007. Invasion and persistence of Mycobacterium avium subsp. paratuberculosis during early stages of Johne's disease in calves. *Infect Immun* 75: 2110-9**
16. **Tanaka S, Sato M, Onitsuka T, Kamata H, Yokomizo Y. 2005. Inflammatory cytokine gene expression in different types of granulomatous lesions during asymptomatic stages of bovine paratuberculosis. *Vet Pathol* 42: 579-88**
17. **Coussens PM. 2004. Model for immune responses to Mycobacterium avium subspecies paratuberculosis in cattle. *Infect Immun* 72: 3089-96**
18. **Coussens PM, Verman N, Coussens MA, Elftman MD, McNulty AM. 2004. Cytokine gene expression in peripheral blood mononuclear cells and tissues of cattle infected with Mycobacterium avium subsp. paratuberculosis: evidence for an inherent proinflammatory gene expression pattern. *Infect Immun* 72: 1409-22**
19. **de Almeida DE, Colvin CJ, Coussens PM. 2008. Antigen-specific regulatory T cells in bovine paratuberculosis. *Vet Immunol Immunopathol* 125: 234-45**
20. **Khalifeh MS, Stabel JR. 2004. Upregulation of transforming growth factor-beta and interleukin-10 in cows with clinical Johne's disease. *Vet Immunol Immunopathol* 99: 39-46**
21. **Khalifeh MS, Stabel JR. 2004. Effects of gamma interferon, interleukin-10, and transforming growth factor beta on the survival of Mycobacterium avium subsp. paratuberculosis in monocyte-derived macrophages from naturally infected cattle. *Infect Immun* 72: 1974-82**
22. **Koo HC, Park YH, Hamilton MJ, Barrington GM, Davies CJ, Kim JB, Dahl JL, Waters WR, Davis WC. 2004. Analysis of the immune response to Mycobacterium avium subsp. paratuberculosis in experimentally infected calves. *Infect Immun* 72: 6870-83**
23. **Koets A, Rutten V, Hoek A, van Mil F, Muller K, Bakker D, Gruys E, van Eden W. 2002. Progressive bovine paratuberculosis is associated with local loss of CD4(+) T cells, increased frequency of gamma delta T cells, and related changes in T-cell function. *Infect Immun* 70: 3856-64**
24. **Nielsen SS. 2010. Immune-based diagnosis of paratuberculosis. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 284-93. Wallingford, U.K.: CABI**

25. **Whittington R. 2010. Cultivation of *Mycobacterium avium* subsp. *paratuberculosis*. In *Paratuberculosis: organism, disease, control*, ed. MA Behr, DM Collins, pp. 244-66. Wallingford, U.K.: CABI**
26. **Lombard JE, Wagner BA, Smith RL, McCluskey BJ, Harris BN, Payeur JB, Garry FB, Salman MD. 2006. Evaluation of environmental sampling and culture to determine Mycobacterium avium subspecies paratuberculosis distribution and herd infection status on US dairy operations. *J Dairy Sci* 89: 4163-71**
27. **Bannantine JP, Puastian ML, Kapur V, Eda S. 2010. Proteome and antigens of *Mycobacterium avium* subsp. *paratuberculosis*. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 94-108. Wallingford, U.K.: CABI**
28. **Kalis CH, Collins MT, Hesselink JW, Barkema HW. 2003. Specificity of two tests for the early diagnosis of bovine paratuberculosis based on cell-mediated immunity: the Johnin skin test and the gamma interferon assay. *Vet Microbiol* 97: 73-86**
29. **Huda A, Jungersen G, Lind P. 2004. Longitudinal study of interferon-gamma, serum antibody and milk antibody responses in cattle infected with Mycobacterium avium subsp. paratuberculosis. *Vet Microbiol* 104: 43-53**
30. **Nielsen SS, Toft N. 2008. Ante mortem diagnosis of paratuberculosis: a review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. *Vet Microbiol* 129: 217-35**
31. **Bölske G, Herthnek D. 2010. Diagnosis of paratuberculosis by PCR. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 267-83. Wallingford, U.K.: CABI**
32. **de Lisle GW. 2010. Ruminant aspects of paratuberculosis vaccination. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 344-52. Wallingford, U.K.: CABI**
33. **van Schaik G, Kalis CH, Benedictus G, Dijkhuizen AA, Huirne RB. 1996. Cost-benefit analysis of vaccination against paratuberculosis in dairy cattle. *Vet Rec* 139: 624-7**
34. **Huygen K, Bull T, Collins DM. 2010. Development of new paratuberculosis vaccines. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 353-68. Wallingford, U.K.: CABI**
35. **Fridriksdottir V, Gunnarsson E, Sigurdarson S, Gudmundsdottir KB. 2000. Paratuberculosis in Iceland: epidemiology and control measures, past and present. *Vet Microbiol* 77: 263-7**
36. **Mullerad J, Hovav AH, Fishman Y, Barletta RG, Bercovier H. 2002. Antigenicity of Mycobacterium paratuberculosis superoxide dismutase in mice. *FEMS Immunol Med Microbiol* 34: 81-8**
37. **Mullerad J, Michal I, Fishman Y, Hovav AH, Barletta RG, Bercovier H. 2002. The immunogenicity of Mycobacterium paratuberculosis 85B antigen. *Med Microbiol Immunol* 190: 179-87**

38. **Ferwerda G, Kullberg BJ, de Jong DJ, Girardin SE, Langenberg DM, van Crevel R, Ottenhoff TH, Van der Meer JW, Netea MG. 2007. Mycobacterium paratuberculosis is recognized by Toll-like receptors and NOD2. *J Leukoc Biol* 82: 1011-8**
39. **Pinedo PJ, Buergelt CD, Donovan GA, Melendez P, Morel L, Wu R, Langaee TY, Rae DO. 2009. Association between CARD15/NOD2 gene polymorphisms and paratuberculosis infection in cattle. *Vet Microbiol* 134: 346-52**
40. **Koets A, Santema W, Mertens H, Oostenrijk D, Keestra M, Overdijk M, Labouriau R, Franken P, Frijters A, Nielen M, Rutten V. 2010. Susceptibility to paratuberculosis infection in cattle is associated with single nucleotide polymorphisms in Toll-like receptor 2 which modulate immune responses against Mycobacterium avium subspecies paratuberculosis. *Prev Vet Med* 93: 305-15**
41. **Hines ME, 2nd, Stabel JR, Sweeney RW, Griffin F, Talaat AM, Bakker D, Benedictus G, Davis WC, de Lisle GW, Gardner IA, Juste RA, Kapur V, Koets A, McNair J, Pruitt G, Whitlock RH. 2007. Experimental challenge models for Johne's disease: a review and proposed international guidelines. *Vet Microbiol* 122: 197-222**
42. **Mutwiri GK, Kosecka U, Benjamin M, Rosendal S, Perdue M, Butler DG. 2001. Mycobacterium avium subspecies paratuberculosis triggers intestinal pathophysiologic changes in beige/scid mice. *Comp Med* 51: 538-44**
43. **Tanaka S, Sato M, Taniguchi T, Yokomizo Y. 1994. Histopathological and morphometrical comparison of granulomatous lesions in BALB/c and C3H/HeJ mice inoculated with Mycobacterium paratuberculosis. *J Comp Pathol* 110: 381-8**
44. **Veazey RS, Horohov DW, Krahenbuhl JL, Taylor HW, Oliver JL, 3rd, Snider TG, 3rd. 1995. Comparison of the resistance of C57BL/6 and C3H/He mice to infection with Mycobacterium paratuberculosis. *Vet Microbiol* 47: 79-87**
45. **Mutwiri GK, Butler DG, Rosendal S, Yager J. 1992. Experimental infection of severe combined immunodeficient beige mice with Mycobacterium paratuberculosis of bovine origin. *Infect Immun* 60: 4074-9**
46. **Baumgart DC, Sandborn WJ. 2007. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet* 369: 1641-57**
47. **Bouma G, Strober W. 2003. The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol* 3: 521-33**
48. **Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH. 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314: 1461-3**

49. Verreck FA, de Boer T, Langenberg DM, Hoeve MA, Kramer M, Vaisberg E, Kastelein R, Kolk A, de Waal-Malefyt R, Ottenhoff TH. 2004. Human IL-23-producing type 1 macrophages promote but IL-10-producing type 2 macrophages subvert immunity to (myco)bacteria. *Proc Natl Acad Sci U S A* 101: 4560-5

50. Umemura M, Yahagi A, Hamada S, Begum MD, Watanabe H, Kawakami K, Suda T, Sudo K, Nakae S, Iwakura Y, Matsuzaki G. 2007. IL-17-mediated regulation of innate and acquired immune response against pulmonary Mycobacterium bovis bacille Calmette-Guerin infection. *J Immunol* 178: 3786-96

51. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G. 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411: 599-603

52. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega FM, Briggs J, Gunther S, Prescott NJ, Onnie CM, Hasler R, Sipos B, Folsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S. 2007. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39: 207-11

53. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sanderson J, Jewell DP, Satsangi J, Mansfield JC, Cardon L, Mathew CG. 2007. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 39: 830-2

54. Cho JH. 2001. The Nod2 gene in Crohn's disease: implications for future research into the genetics and immunology of Crohn's disease. *Inflamm Bowel Dis* 7: 271-5

55. Coulombe F, Divangahi M, Veyrier F, de Leseleuc L, Gleason JL, Yang Y, Kelliher MA, Pandey AK, Sassetti CM, Reed MB, Behr MA. 2009. Increased NOD2-mediated recognition of N-glycolyl muramyl dipeptide. *J Exp Med* 206: 1709-16

56. Singh SB, Davis AS, Taylor GA, Deretic V. 2006. Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science* 313: 1438-41

57. Travassos LH, Carneiro LA, Ramjeet M, Hussey S, Kim YG, Magalhaes JG, Yuan L, Soares F, Chea E, Le Bourhis L, Boneca IG, Allaoui A, Jones NL, Nunez G, Girardin SE, Philpott DJ. 2010. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* 11: 55-62

58. Cooney R, Baker J, Brain O, Danis B, Pichulik T, Allan P, Ferguson DJ, Campbell BJ, Jewell D, Simmons A. 2010. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* 16: 90-7
59. Selby W, Pavli P, Crotty B, Florin T, Radford-Smith G, Gibson P, Mitchell B, Connell W, Read R, Merrett M, Ee H, Hetzel D. 2007. Two-year combination antibiotic therapy with clarithromycin, rifabutin, and clofazimine for Crohn's disease. *Gastroenterology* 132: 2313-9
60. Behr MA, Hanley J. 2008. Antimycobacterial therapy for Crohn's disease: a reanalysis. *Lancet Infect Dis* 8: 344
61. Elliott M, Benson J, Blank M, Brodmerkel C, Baker D, Sharples KR, Szapary P. 2009. Ustekinumab: lessons learned from targeting interleukin-12/23p40 in immune-mediated diseases. *Ann N Y Acad Sci* 1182: 97-110
62. Behr MA. 2010. Paratuberculosis in Crohn's disease. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 40-9. Wallingford, U.K.: CABI
63. Shanahan F, O'Mahony J. 2005. The mycobacteria story in Crohn's disease. *Am J Gastroenterol* 100: 1537-8
64. Jeyanathan M, Boutros-Tadros O, Radhi J, Semret M, Bitton A, Behr MA. 2007. Visualization of Mycobacterium avium in Crohn's tissue by oil-immersion microscopy. *Microbes Infect* 9: 1567-73
65. Feller M, Huwiler K, Stephan R, Altpeter E, Shang A, Furrer H, Pfyffer GE, Jemmi T, Baumgartner A, Egger M. 2007. Mycobacterium avium subspecies paratuberculosis and Crohn's disease: a systematic review and meta-analysis. *Lancet Infect Dis* 7: 607-13
66. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, Marks JS. 2003. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 289: 76-9
67. Kremer L, Besra GS. 2005. A Waxy Tale, by *Mycobacterium tuberculosis*. In *Tuberculosis and the Tubercle Bacillus.*, ed. ST Cole, DK Eisenach, DN McMurray, J W.R.J., pp. 287-305. Washington, D.C.: ASM Press
68. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 393: 537-44
69. Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, Banerji N, Kanjilal S, Kapur V. 2005. The complete genome sequence of

Mycobacterium avium subspecies paratuberculosis. *Proc Natl Acad Sci U S A* 102: 12344-9

70. Veyrier F, Pletzer D, Turenne C, Behr MA. 2009. Phylogenetic detection of horizontal gene transfer during the step-wise genesis of Mycobacterium tuberculosis. *BMC Evol Biol* 9: 196

71. Mostowy S, Tsolaki AG, Small PM, Behr MA. 2003. The in vitro evolution of BCG vaccines. *Vaccine* 21: 4270-4

72. Minnikin DE, Kremer L, Dover LG, Besra GS. 2002. The methyl-branched fortifications of Mycobacterium tuberculosis. *Chem Biol* 9: 545-53

73. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409: 1007-11

74. Turenne CY, Alexander DC. 2010. *Mycobacterium avium* complex. In *Paratuberculosis: organism, disease, control.*, ed. MA Behr, DM Collins, pp. 60-72. Wallingford, U.K.: CABI

75. Behr MA. 2008. Mycobacterium du jour: what's on tomorrow's menu? *Microbes Infect* 10: 968-72

76. Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. 2009. Opportunistic pathogens enriched in showerhead biofilms. *Proc Natl Acad Sci U S A* 106: 16393-9

77. Turenne CY, Collins DM, Alexander DC, Behr MA. 2008. Mycobacterium avium subsp. paratuberculosis and M. avium subsp. avium are independently evolved pathogenic clones of a much broader group of M. avium organisms. *J Bacteriol* 190: 2479-87

78. Semret M, Alexander DC, Turenne CY, de Haas P, Overduin P, van Soolingen D, Cousins D, Behr MA. 2005. Genomic polymorphisms for Mycobacterium avium subsp. paratuberculosis diagnostics. *J Clin Microbiol* 43: 3704-12

79. Alexander DC, Turenne CY, Behr MA. 2009. Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. *J Bacteriol* 191: 1018-25

80. Zhu X, Tu ZJ, Coussens PM, Kapur V, Janagama H, Naser S, Sreevatsan S. 2008. Transcriptional analysis of diverse strains Mycobacterium avium subspecies paratuberculosis in primary bovine monocyte derived macrophages. *Microbes Infect* 10: 1274-82

81. Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW. 1993. Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. *Science* 261: 1454-7

82. Pandey AK, Sassetti CM. 2008. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A* 105: 4376-80
83. Gioffre A, Infante E, Aguilar D, Santangelo MP, Klepp L, Amadio A, Meikle V, Etchechoury I, Romano MI, Cataldi A, Hernandez RP, Bigi F. 2005. Mutation in mce operons attenuates Mycobacterium tuberculosis virulence. *Microbes Infect* 7: 325-34
84. Li Y, Miltner E, Wu M, Petrofsky M, Bermudez LE. 2005. A Mycobacterium avium PPE gene is associated with the ability of the bacterium to grow in macrophages and virulence in mice. *Cell Microbiol* 7: 539-48
85. Frost LS, Leplae R, Summers AO, Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 3: 722-32
86. Groisman EA, Ochman H. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* 87: 791-4
87. Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* 23: 1089-97
88. Stratmann J, Strommenger B, Goethe R, Dohmann K, Gerlach GF, Stevenson K, Li LL, Zhang Q, Kapur V, Bull TJ. 2004. A 38-kilobase pathogenicity island specific for Mycobacterium avium subsp. paratuberculosis encodes cell surface proteins expressed in the host. *Infect Immun* 72: 1265-74
89. Maciag A, Dainese E, Rodriguez GM, Milano A, Provvedi R, Pasca MR, Smith I, Palu G, Riccardi G, Manganelli R. 2007. Global analysis of the Mycobacterium tuberculosis Zur (FurB) regulon. *J Bacteriol* 189: 730-40
90. Roupie V, Rosseels V, Piersoel V, Zinniel DK, Barletta RG, Huygen K. 2008. Genetic resistance of mice to Mycobacterium paratuberculosis is influenced by Slc11a1 at the early but not at the late stage of infection. *Infect Immun* 76: 2099-105
91. Riccardi G, Milano A, Pasca MR, Nies DH. 2008. Genomic analysis of zinc homeostasis in Mycobacterium tuberculosis. *FEMS Microbiol Lett* 287: 1-7
92. De Voss JJ, Rutter K, Schroeder BG, Su H, Zhu Y, Barry CE, 3rd. 2000. The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages. *Proc Natl Acad Sci U S A* 97: 1252-7
93. Merkal RS, Curran BJ. 1974. Growth and metabolic characteristics of Mycobacterium paratuberculosis. *Appl Microbiol* 28: 276-9
94. Johnson EE, Srikanth CV, Sandgren A, Harrington L, Trebicka E, Wang L, Borregaard N, Murray M, Cherayil BJ. 2010. Siderocalin inhibits the intracellular replication of Mycobacterium tuberculosis in macrophages. *FEMS Immunol Med Microbiol* 58: 138-45

95. Gomes-Pereira S, Rodrigues PN, Appelberg R, Gomes MS. 2008. Increased susceptibility to Mycobacterium avium in hemochromatosis protein HFE-deficient mice. *Infect Immun* 76: 4713-9

96. Boelaert JR, Vandecasteele SJ, Appelberg R, Gordeuk VR. 2007. The effect of the host's iron status on tuberculosis. *J Infect Dis* 195: 1745-53

97. Brown KA, Ratledge C. 1975. The effect of p-aminosalicyclic acid on iron transport and assimilation in mycobacteria. *Biochim Biophys Acta* 385: 207-20

98. Wu CW, Schmoller SK, Shin SJ, Talaat AM. 2007. Defining the stressome of Mycobacterium avium subsp. paratuberculosis in vitro and in naturally infected cows. *J Bacteriol* 189: 7877-86

99. Murry JP, Sassetti CM, Lane JM, Xie Z, Rubin EJ. 2008. Transposon site hybridization in Mycobacterium tuberculosis. *Methods Mol Biol* 416: 45-59

100. Sassetti CM, Boyd DH, Rubin EJ. 2001. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* 98: 12712-7

101. Koo IC, Ohol YM, Wu P, Morisaki JH, Cox JS, Brown EJ. 2008. Role for lysosomal enzyme beta-hexosaminidase in the control of mycobacteria infection. *Proc Natl Acad Sci U S A* 105: 710-5

102. Reddacliff LA, Beh K, McGregor H, Whittington RJ. 2005. A preliminary study of possible genetic influences on the susceptibility of sheep to Johne's disease. *Aust Vet J* 83: 435-41