# Epidemiology and Pathogenesis of Bovine Mycobacteria

Shannon C Duffy

Laboratory of Dr. Marcel A Behr

Department of Microbiology and Immunology

McGill University, Montreal

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

# ABSTRACT

Zoonoses, or the diseases transmitted from animals to humans, make up ~60% of human diseases and ~75% of emerging pathogens. The threat of zoonotic disease transmitted by bovines is of concern due to their global population and the reliance of many societies on bovines for dairy and meat. There are several mycobacterial species which infect bovines and have zoonotic potential. This includes *Mycobacterium bovis* and *M. orygis,* which are members of the *M. tuberculosis* complex (MTBC) and can cause bovine and zoonotic tuberculosis (TB). The TB vaccine strain is an attenuated strain of *M. bovis* called bacille Calmette-Guérin (BCG) which causes disease on rare occasions in children with primary immunodeficiencies. Finally, *M. avium* subsp. *paratuberculosis* (*MAP*) is known to cause the ruminant gastrointestinal disease paratuberculosis and is hypothesized to be etiologically linked to Crohn's disease.

In this thesis, I explored the epidemiology and pathogenesis of these organisms through two aims. First, I aimed to determine the proportion of MTBC subspecies causing disease in South Asia. I hypothesized that disease would be caused primarily by *M. tuberculosis* with a smaller proportion caused by *M. bovis, M. orygis,* and BCG. Second, I aimed to study the pathogenesis of bovine mycobacteria by developing mouse models of oral infection. I hypothesized that streptomycin pre-treatment would improve mycobacterial organ infection following gavage and would model disease in mice.

In chapter II, we conducted a molecular epidemiological screen of 940 positive TB cultures from a clinic in Vellore, India using assays designed to differentiate between members of the MTBC. We then performed an analysis of all publicly available genomes of MTBC subspecies from South Asia (n=715) to determine whether what was observed in India was also

found in nearby countries. As expected, most cases were caused by *M. tuberculosis.* We found that zoonotic TB occurred, albeit infrequently, and was caused by *M. orygis.* Although some cases of BCG disease were detected, we found no evidence of *M. bovis* in the cultures from India nor in the genomes from South Asia.

In chapter III, we aimed to develop an improved assay for BCG disease that could be implemented in the clinic in Vellore, which previously did not have molecular tests for its detection. We developed a two-step real-time PCR assay which detected BCG using single nucleotide polymorphisms (SNPs). This assay was also designed to differentiate between early and late strains of BCG which have previously been suggested to be linked to different rates of adverse events. We demonstrated that the assay was accurate and robust, and we validated its use on a pilot data set of 19 clinical samples from BCG-suspected patients, 10 of which were identified as BCG.

In chapter IV, we studied the pathogenesis of bovine mycobacteria by creating an oral mouse model which utilizes streptomycin pre-treatment to improve organ infection in the gut. We found that streptomycin pre-treatment significantly increased organ colony forming units (CFUs) following gavage. When this model was applied to *M. bovis* and *M. orygis,* this led to a persistent infection of the intestines and mesenteric lymph nodes (MLNs) as well as pulmonary disease consistent with zoonotic TB. When the model was applied to *MAP*, persistent local infection occurred in the intestines and MLNs up to 24 weeks, but with no evidence of inflammation or disease. This suggested that additional modifications may need to be explored to generate a small animal model of *MAP* enteritis.

Taken together, this thesis encompasses a broad range of fundamental novel information regarding the epidemiology and pathogenesis of several bovine mycobacteria. This thesis has

provided new tools to detect bacteria in the clinic as well as generated important insights concerning the geographic distribution of zoonotic TB pathogens and the pathogenesis of mycobacterial infections in the intestines.

# RESUMÉ

Les zoonoses, ou les maladies transmises des animaux aux humains, représentent environ 60% des maladies humaines et environ 75% des agents pathogènes émergents. La menace de la maladie zoonotique transmises par les bovins est un sujet de préoccupation en raison de leur population mondiale et de la dépendance de nombreuses sociétés à ces animaux pour les produits laitiers et la viande. Il y a plusieurs espèces de mycobactéries qui infectent les bovins et qui présentent un potentiel zoonotique. Cela inclut *Mycobacterium bovis* et *M. orygis,* qui sont membres du complexe *M. tuberculosis* (CMTB) et qui peuvent causer la tuberculose (TB) bovine et zoonotique. La souche vaccinale contre TB est une souche atténuée de *M. bovis* appelée bacille Calmette-Guérin (BCG) qui peut provoquer occasionnellement la maladie chez les enfants présentant des immunodéficiences primaires. Enfin, *M. avium* subsp. *paratuberculosis* (*MAP*) est connue pour causer la paratuberculose, une maladie gastro-intestinale chez les ruminants, et on émet l'hypothèse qu'elle pourrait être liée étiologiquement à la maladie de Crohn.

Dans cette thèse, j'ai exploré l'épidémiologie et la pathogenèse de ces organismes avec deux objectifs. Premièrement, j'ai cherché à déterminer la proportion de sous-espèces du CMTB causant des maladies en Asie du Sud. J'ai émis l'hypothèse que la maladie serait principalement causée par *M. tuberculosis* avec une plus petite proportion causée par *M. bovis, M. orygis,* et BCG. Deuxièmement, j'ai cherché à étudier la pathogenèse des mycobactéries bovines en développant des modèles murins d'infection orale. J'ai émis l'hypothèse qu'un prétraitement à la streptomycine améliorerait l'infection mycobactérienne dans les organes après gavage et permettrait de modéliser les maladies chez les souris.

Dans le chapitre II, nous avons effectué une étude d'épidémiologie moléculaire de 940 cultures tuberculeuses positives provenant d'une clinique à Vellore, en Inde en utilisant des tests conçus pour différencier les membres du CMTB. Nous avons ensuite effectué une analyse de tous les génomes disponibles au public des sous-espèces de CMTB d'Asie de Sud (n=715) afin de déterminer si ce qui a été observé en Inde se retrouvait également dans les pays voisins. Comme prévu, la plupart des cas étaient causés par *M. tuberculosis.* Nous avons constaté que la tuberculose zoonotique se produisait, bien que rarement, et était causée par *M. orygis.* Bien que quelques cas de la maladie de BCG aient été détectés, nous n'avons trouvé aucune preuve de la présence de *M. bovis* dans les cultures de l'Inde ni dans les génomes d'Asie du Sud.

Dans le chapitre III, nous avons cherché à développer un test amélioré pour la maladie de BCG qui pourrait être mis en œuvre dans la clinique de Vellore, qui n'avait auparavant pas de tests moléculaires pour sa détection. Nous avons mis au point un test PCR en temps réel en deux étapes qui détecte la BCG en utilisant des polymorphismes d'un seul nucléotide (PSNs). Ce test était également conçu pour différencier les souches précoces et tardives du BCG dont on a précédemment suggéré qu'elles étaient liées à des taux différents d'évènements indésirables. Nous avons démontré que le test était précis et robuste et nous avons validé son utilisation sur un ensemble de données pilotes de 19 échantillons cliniques de patients suspectés, dont 10 ont été identifié comme ayant la BCG.

Dans le chapitre IV, nous avons cherché à étudier la pathogenèse des mycobactéries bovines en créant un modèle expérimental oral chez les souris qui utilise un prétraitement à la streptomycine pour améliorer l'infection dans les intestins. Nous avons constaté que le prétraitement à la streptomycine a significativement amélioré les unités formant colonies (UFCs) dans les organes après gavage. Lorsque ce modèle a été appliqué à *M. bovis* et *M. orygis,* cela a

mené à une infection persistante des intestins et des ganglions mésentériques (GMs) ainsi qu'à une maladie pulmonaire qui était similaire à la TB zoonotique. Lorsque ce modèle a été appliqué à *MAP,* une infection locale persistante s'est produite dans les intestins et GMs pendant 24 semaines, mais avec aucune preuve d'inflammation ou de maladie. Cela a suggéré qu'il faudrait peut-être explorer d'autres modifications supplémentaires pour générer un modèle d'entérite avec *MAP* chez le petit animal.

Dans l'ensemble, cette thèse englobe une vaste gamme de nouvelles informations fondamentales concernant l'épidémiologie et la pathogenèse de plusieurs mycobactéries bovines. Cette thèse a fourni de nouveaux outils pour détecter les bactéries dans la clinique et a généré des idées importantes concernant la distribution géographique des pathogènes de la TB zoonotique et la pathogenèse des infections mycobactériennes dans les intestins.

# ACKNOWLEDGEMENTS

First and foremost, I have to express my gratitude to the members to the Behr lab who have made this experience so positive. To Marcel Behr, I have been very fortunate to have you as a mentor. You have an uncanny ability to provide guidance and valuable instruction to your students while still promoting independent exploration and growth. I can't begin to accurately convey how much I've learned under your leadership. Thank you. To Sarah Danchuk, I don't know how I would've made it through without you. You've been an advisor and a support system and a best friend all in one. I'll be eternally grateful for your encouragement these last five years. To Ori Solomon, you've brought so much laughter and fun to our lab. The number of times your presence has eased a stressful moment are too many to count. To Jaryd Sullivan, you were always someone I could count on for a piece of sound advice or to lend me a hand whenever I needed help. Thank you for your friendship. Finally, to Fiona McIntosh, you have been foundational to the Behr lab. Without your guidance, none of this would've been possible.

I also need to give a huge thank you to Andréanne Lupien. You are the person I always go to whenever I have a question even remotely science related. Your seemingly boundless knowledge on all subjects has been invaluable. You are an incredible teacher, and I am so thankful for your support. To Vivek Kapur, Joy Michael, Sreenidhi Srinivasan, and Manigandan Venkatesan, thank you for the tremendous opportunities you have allowed me to take part in which are detailed in chapters II and III of this thesis. It has been a dream come true to work with you all on these projects. And to Corinne Maurice and Martin Olivier, I am grateful for all of the help I have received from you over the years as a part of my advisory committee. Your feedback and suggestions helped shape this thesis into one I am so proud of.

Finally, I need to acknowledge my incredible family, without whom I never would've made it here in the first place. To my parents, Peter and Deborah, no words will ever properly express the degree of my gratitude. You are the most loving, kind-hearted, and supportive parents one could wish for. To Brianna, I am incredibly lucky to know such an exceptional person, and even more lucky to have that person be both my sister and my best friend. I am so grateful to have you in my life. To Ned Drew, you make everything more fun. Thank you for joining our family. To Turney McKee, I truly cannot thank you enough for everything you've done. It's hard to imagine how different my PhD might have been without your infallible support. Your extraordinary talent for making me laugh so hard I cry have carried me through this whole experience. And of course, thank you to Watson for being an unending source of love and comfort in my life.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AIDS | Acquired immunodeficiency syndrome |
| AIEC | Adherent-invasive *Escherichia coli* |
| ART | Anti-retroviral therapy |
| ATT | Anti-tubercular therapy |
| BCG | *Mycobacterium bovis* bacille Calmette-Guérin |
| Bp | Base pair |
| BWA-MEM | Burrows-Wheeler aligner maximum exact matches |
| CCAC | Canadian Council on Animal Care |
| CFP | Culture filtrate protein |
| CFU | Colony-forming units |
| CGD | Chronic granulomatous disease |
| ChIP-seq | Chromatin immunoprecipitation followed by high-throughput sequencing |
| CNS | Central nervous system |
| COPD | Chronic obstructive pulmonary disease |
| COVID-19 | Coronavirus disease 2019 |
| CRP | C reactive protein |
| $C_T$ | Cycle threshold |
| DC-SIGN | DC-specific intercellular adhesion molecule-3-grabbing non-integrin |
| DNA | Deoxyribonucleic acid |
| EHEC | Enterohaemorrhagic *Escherichia coli* |
| EPEC | Enteropathogenic *Escherichia coli* |

| | |
|---|---|
| ESAT | Early secreted antigenic target |
| ESR | Erythrocyte sedimentation rate |
| FAO | Food and Agriculture Organization |
| GALT | Gut-associated lymphoid tissue |
| Ge | Genome equivalents |
| H&E | Hematoxylin and eosin |
| HIV | Human immunodeficiency virus |
| HSCT | Hematopoietic stem cell transplant |
| IBD | Inflammatory bowel disease |
| IFN | Interferon |
| IRB | Institutional review board |
| IS | Insertion sequence |
| Lbp/Hlp | Laminin-binding/ histone-like protein |
| LJ | Lowenstein-Jensen |
| LMIC | Low-income and middle-income country |
| LOD | Limit of detection |
| $LSP^P$ | Large sequence polymorphism of paratuberculosis |
| *MAA* | *Mycobacterium avium* subspecies *avium* |
| MABC | *Mycobacterium abscessus* complex |
| MAC | *Mycobacterium avium* complex |
| *MAH* | *Mycobacterium avium* subspecies *hominissuis* |
| ManLAM | Mannose-capped lipoarabinomannan |
| *MAP* | *Mycobacterium avium* subspecies *paratuberculosis* |

| | |
|---|---|
| *MAS* | *Mycobacterium avium* subspecies *silvaticum* |
| MDP | Muramyl dipeptide |
| MGIT | Mycobacteria growth indicator tube |
| MLN | Mesenteric lymph node |
| MPB | Major protein of *Mycobacterium bovis* |
| MPIL | Multiplex PCR of IS*900* integration loci |
| MSMD | Mendelian susceptibility to mycobacterial diseases |
| MTBC | *Mycobacterium tuberculosis* complex |
| NADPH | Nicotinamide adenine dinucleotide phosphate |
| NCBI | National Center for Biotechnology Information |
| NMIBC | Non-muscle-invasive bladder cancer |
| NTM | Nontuberculous mycobacteria |
| PANTA | Polymyxin B, amphotericin B, nalidixic acid, trimethoprim and azlocillin |
| PBMC | Peripheral blood mononuclear cell |
| PCP | Pneumoncystis pneumonia |
| PCR | Polymerase chain reaction |
| PID | Primary immunodeficiency disease |
| PPD | Purified protein derivative |
| PTB | Pulmonary tuberculosis |
| PZA | Pyrazinamide |
| RAxML | Randomized axelerated maximum likelihood |
| RD | Region of difference |
| RI-MUHC | Research Institute of the McGill University Health Centre |

| | |
|---|---|
| RT-PCR | Real-time polymerase chain reaction |
| SCID | Severe combined immunodeficiency disorder |
| SNP | Single nucleotide polymorphism |
| SPAdes | St. Petersburg genome assembler |
| SRA | Sequence reach archive |
| Strep-R | Streptomycin-resistant |
| TB | Tuberculosis |
| Th | T-helper |
| TNF | Tumour necrosis factor |
| UIP | Universal Immunisation Programme |
| USD | United States dollar |
| vSNP | Validate single nucleotide polymorphism tool |
| WGS | Whole-genome sequencing |
| WHO | World Health Organization |
| WOAH | World Organisation for Animal Health |
| WT | Wildtype |
| ZN | Ziehl-Neelson |

# CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

---

The contributions to original knowledge described in this thesis are the following:

1. Designed novel assays to rapidly differentiate between multiple MTBC subspecies including *M. orygis*
    a. Previously published assays could not distinguish *M. orygis* from other MTBC subspecies
    b. Developed both novel conventional 3-primer and 6-primer PCRs and Taqman real-time PCR assays
2. First to conduct a large-scale screen for MTBC subspecies from positive cultures in India using assays to detect *M. orygis*
    a. Determined that zoonotic TB in India is predominately caused by *M. orygis* with no evidence of infection with *M. bovis*
3. First compilation of all publicly available WGS of MTBC subspecies from South Asia which demonstrated the absence of *M. bovis* in the region
4. Designed a novel rapid SNP-based multiplex real-time PCR assay for BCG disease
    a. Distinguished between early and late strains to enable vaccine pharmacovigilance
    b. Applied a rule-in strategy to simplify BCG identification
5. First to publish application of oligo-mediated recombineering to generate targeted SNPs in *MAP, MAH,* and *M. orygis*
    a. Simplified, faster system with high efficiency to manipulate and study these mycobacterial species
6. First to publish outcomes of experimental infection with *M. orygis*

    a. High doses of *M. orygis* given via the oral route led to persistent organ infection and pulmonary disease

7. First assessment of streptomycin pre-treatment in an oral mouse model of mycobacteria
    a. Streptomycin pre-treatment improved enteric infection in both *MAP*- and *M. bovis*-infected mice
    b. Model allowed for generation of zoonotic TB-like disease progression in *M. bovis*- and *M. orygis*-infected mice

# CONTRIBUTIONS OF AUTHORS

This thesis was written by Shannon C Duffy with input and editing from Marcel A Behr. This work was prepared according to the guidelines for a manuscript-based (article-based) thesis provided by McGill University Graduate and Postdoctoral Studies. The following is a description of the contributions of each author to the five chapters of this thesis:

**Chapter 1 – Literature review**

This chapter was written by Shannon C Duffy with guidance and editing from Marcel A Behr aside from section 5.5. Section 5.5 is a published book chapter (Paratuberculosis: Organism, Disease, Control (2nd ed.), pp 29-44, CAB International, Nov 2020) written by both Shannon C Duffy and Marcel A Behr. Marcel A Behr conceptualized the structure of the chapter and Shannon Duffy wrote the original draft. Marcel Behr made significant contributions to the writing and both authors were involved in the editing of the chapter.

**Chapter II - Reconsidering *Mycobacterium bovis* as a proxy for zoonotic tuberculosis: a molecular epidemiological surveillance study**

This chapter is a published manuscript (The Lancet Microbe, Volume 1, Issue 2, E66-E73, June 2020) with the following authors: Shannon C Duffy, Sreenidhi Srinivasan, Megan A Schilling, Tod Stuber, Sarah N Danchuk, Joy S Michael, Manigandan Venkatesan, Nitish Bansal, Sushila Maan, Naresh Jindal, Deepika Chaudhary, Premanshu Dandapat, Robab Katani, Shubhada Chothe, Maroudam Veerasami, Suelee Robbe-Austerman, Nicholas Juleff, Vivek Kapur, and Marcel A Behr.

Their contributions are as follows: Shannon C Duffy designed the molecular assays, prepared the clinical samples, did the PCR assays, assisted with whole-genome sequencing (WGS), analysed the results, and wrote the first draft of the manuscript. Sreenidhi Srinivasan validated the assays, prepared the clinical samples, did the PCR assays, and assisted with WGS. Sarah Danchuk prepared the clinical samples, did the PCR assays, and assisted with WGS. Megan Schilling, Tod Stuber, and Suelee Robbe-Austerman did the bioinformatics analysis and constructed and interpreted the phylogenetic trees. Joy Sarojini Michael contributed to the study design, provided the clinical samples, and supervised the laboratory studies at Christian Medical College in Vellore, India. Manigandan Venkatesan did the Sanger sequencing assays and analysed the associated data. Nitish Bansal, Deepika Chaudhry, Rubab Katani, and Shubadha Chothe did the WGS and Sanger sequencing assays and the associated analyses. Naresh Jindal and Sushila Maan supervised the WGS studies at Lala Lajpat Rai University of Veterinary and Animal Sciences in Hisar, India. Premanshu Danapat and Maroudam Veerasami contributed to the analyses of animal-associated isolates. Nicholas Juleff assisted in the study design and interpretation. Vivek Kapur and Marcel A Behr directed the study design, supervised the project, acquired funding, and contributed to writing the manuscript. All authors reviewed and edited the final version of the manuscript.

## Chapter III - Development of a Multiplex Real-Time PCR Assay for *Mycobacterium bovis* BCG and Validation in a Clinical Laboratory

This chapter is a published manuscript (Microbiology Spectrum, Volume 9, Number 2, e0109821, Sept 2021) with the following authors: Shannon C Duffy, Manigandan

Venkatesan, Shubhada Chothe, Indira Poojary, Valsan Philip Verghese, Vivek Kapur, Marcel A Behr, and Joy Sarojini Michael.

Their contributions are as follows: Shannon C Duffy designed the assays, did all quality assessments of the assays, analyzed the results, and wrote the first draft of the manuscript. Manigandan Venkatesan prepared clinical samples and performed the assays on clinical samples. Shubhada Chothe and Indira Poojary validated the assays. Valsan Philip Verghese assisted in the interpretation of clinical data. Vivek Kapur, Marcel A Behr, and Joy Sarojini Michael directed the study design, supervised the project, acquired funding, and contributed to writing the manuscript. All authors reviewed and edited the final version of the manuscript.

**Chapter IV - Establishment of persistent enteric mycobacterial infection following streptomycin pre-treatment**

This chapter is a published manuscript (Gut Pathogens, Volume 15, Article 46, Oct 2023) with the following authors: Shannon C Duffy, Andréanne Lupien, Youssef Elhaji, Mina Farag, Victoria Marcus, and Marcel A Behr.

Their contributions are as follows: Shannon C Duffy generated engineered bacterial strains, performed all *in vitro* and *in vivo* experiments, analyzed the data, and wrote the first draft of this manuscript. Andréanne Lupien assisted with *in vivo* experiments and study design. Youssef Elhaji performed preliminary experiments and assisted with study design. Mina Farag and Victoria Marcus assessed histopathology and assisted with study design. Marcel A Behr directed study design, supervised the project, acquired funding, and contributed to writing this manuscript. All authors reviewed and edited the final version of this manuscript.

**Chapter V – Discussion**

Shannon C Duffy wrote this chapter with guidance and editing from Marcel A Behr.

# CHAPTER I

# Literature Review

---

## 1. The relationship between animal and human health

The potential for development of human disease from animal sources has long been recognized as a significant threat to public health. Zoonoses, or the diseases transmitted from animals to humans, represent an estimated ~60% of human pathogens[1]. The concept of One Health refers to approaching disease control and surveillance in a collaborative and transdisciplinary manner which recognizes the close interactions between human health, animal health, and the environment[2]. Zoonotic diseases are also the source of ~75% of emerging new pathogens[1]. It has been shown that the emergence of zoonotic diseases is often driven by human actions such as deforestation, climate change, tourism, and trade[2]. As society continues to develop, the threat of emerging diseases is increasing[3]. A One Health approach to controlling and monitoring disease is an essential concept to safeguard public health and establish pandemic preparedness[2].

The impacts of zoonotic diseases have acted as major landmarks in our history. The bubonic plague, caused by the transmission of *Yersinia pestis* from infected fleas, was responsible for the death of ~25 million people in the 14$^{th}$ century[4]. The 1918 influenza pandemic which caused the death of an estimated 50 million people was avian in origin[5,6]. Human immunodeficiency virus (HIV) arose following transmission of virus from African primates and caused a global crisis beginning in the 1980s[7]. This disease has led to the death of ~40.1 million people since its discovery[8]. Ebola outbreaks are also thought to be incited following transmission from an infected animal such as a bat or non-human primate to humans[9]. The largest Ebola epidemic to

date occurring in west Africa from 2014-2016 caused global panic and led to the death of over 11,000 people[10]. Finally, the origin of the recent COVID-19 pandemic which led to worldwide shutdowns and caused the death of ~7 million people to date is still debated[11]. However, a prevailing theory suggests the virus had zoonotic origins similar to the related SARS and MERS outbreaks[12].

The risk of zoonoses presented by bovines (cattle, buffalo, bison) is of particular concern due to the widespread societal reliance on these animals for dairy products and meat[13]. There are an estimated 1.4 billion cattle worldwide[14], emphasizing the importance of understanding the risks of zoonotic disease transmitted by these animals and the application of a One Health approach to their monitoring and prevention. The genus *Mycobacterium* includes several bacterial pathogens which infect bovines and have zoonotic potential. This includes animal-associated members of the *Mycobacterium tuberculosis* complex as well as non-tuberculous mycobacterial species.

## 2. Introduction to mycobacteria

The genus *Mycobacterium* was first named in 1896 by Lehmann and Neumann[15]. Mycobacteria are a part of the Actinomycetales order and the only genus within the *Mycobacteriaceae* family. To date, there are 195 species in the genus, the most famous of which being *Mycobacterium tuberculosis* and *M. leprae.* The genus comprises a diverse collection of species which include environmental, opportunistic, and pathogenic bacteria. Mycobacteria are aerobic, meaning they grow in the presence of oxygen. They are generally rod-shaped (bacilli), but some species such as *M. avium* are shorter and more rounded (coccobacillary). On solid agar, colonies may be smooth or rough and are often buff or yellow in color[16]. The unique features of the mycobacterial cell envelope are what makes it acid fast and allows it to be distinguished from

other bacteria using a Ziehl-Neelson (ZN) stain. Mycobacteria have a cell wall containing peptidoglycan which is surrounded by a layer of arabinogalactan and an outer layer of mycolic acids. This lipid-rich structure is what allows mycobacteria to resist decolorization and appear pink following ZN staining[17].

The genus *Mycobacterium* can be broadly characterized into 3 groups (Figure 1). The first is the *Mycobacterium tuberculosis* complex which is a group of 19 subspecies which have high genetic similarity but differ in their host specificity[18]. The second is the leprotic bacteria *M. leprae* and *M. lepromatosis* which are uncultivatable but are known to cause leprosy, a disease affecting the skin, peripheral nerves, upper respiratory tract, and eyes[19,20]. The remaining mycobacteria fit into the third diverse group called nontuberculous mycobacteria[21].

### *2.1. Mycobacterium tuberculosis complex*

The *Mycobacterium tuberculosis* complex (MTBC) is a group of highly related mycobacterial subspecies known to cause tuberculosis (TB) in their hosts. The subspecies within the MTBC are 99.9% identical at the nucleotide level and share the same 16S sequence[22]. TB is an infectious disease which is typically spread by aerosol and which most often presents as pulmonary disease[23]. However, it can also affect other organs in the body such as the lymph nodes, bones, and gastrointestinal tract in a presentation called extrapulmonary TB. The most common symptoms associated with TB are cough, weight loss, fever, and night sweats[24]. This disease is one of the leading causes of death worldwide with an estimated 10.6 million cases and 1.4 million deaths in 2021[23].

The MTBC includes 9 lineages of human-adapted members known to cause TB in humans[25–27]. Lineages 1-4 and 7 are referred to as *M. tuberculosis sensu stricto*[26]. Lineages 5 and

6 have previously been called *M. africanum* and are so named due to their geographic restriction to west Africa[28]. The remaining lineages *M. tuberculosis* lineage 8 and 9 were first described in 2020 and 2021 respectively and both are suggested to also be restricted to different parts of Africa[25,27]. *M. canettii* is sometimes considered as a part of the MTBC but may also be listed as an outgroup. It is an environmental bacterium rarely isolated from humans within the Horn of Africa. It is thought to share features with the suggested common ancestor of the MTBC[29].

The complex also includes 9 animal-associated members which are known to cause tuberculosis in a variety of animal hosts. Among animal-associated members are *M. pinnipedii* which is known to cause TB in seals and sea lions, *M. microti* which causes TB in voles, *M. mungi* which is a pathogen of mongooses, *M. caprae* which has been isolated from livestock animals such as goats, and *M. suricattae* which infects meerkats. Other members include the chimpanzee bacillus and dassie bacillus which cause TB in chimpanzees and rock hyraxes respectively[18,26]. The 2-remaining animal-associated members of the MTBC *M. bovis* and *M. orygis* have been frequently isolated from bovines, a range of other animals, as well as humans. The vaccine strain against TB, *M. bovis* bacille Calmette-Guérin (BCG) is an attenuated strain of *M. bovis*[30]. These 3 subspecies *M. bovis, M. bovis* BCG, and *M. orygis* will be discussed in detail within this thesis.

### *2.2. Nontuberculous mycobacteria*

Nontuberculous mycobacteria (NTM) refer to a large and diverse group of mycobacterial species. They can be broadly divided into those that are fast-growing (e.g. *M. abscessus* complex, *M. smegmatis*) and those that are slow-growing (e.g. *M. ulcerans, M. avium* complex). Some NTMs are purely environmental and are often found in the soil or in water, some species

can cause opportunistic disease in immunocompromised hosts, whereas others are professional pathogens[21].

The *M. abscessus* complex (MABC) consists of 3 subspecies: *M. abscessus* subspecies *abscessus, M. abscessus* subspecies *bolletii,* and *M. abscessus* subspecies *massilense*[31]. The MABC species can be found in water, dust, and soil. However, they have also been linked to pulmonary, skin, and soft tissue infections[32]. Skin infections are most often found following injury, cosmetic procedures, or transplantation[32–34]. MABC subspecies are intrinsically resistant to a number of antibiotics including aminoglycosides and quinolones and, in some subspecies, macrolides. This makes infections with MABC subspecies very difficult to treat. MABC infection is of particular concern in patients with cystic fibrosis whose most common cause of death is chronic bacterial infections[32].

*M. smegmatis* is a model species for mycobacterial research. It is both fast-growing and non-pathogenic, making it an easy-to-work with and safe alternative to other slow-growing, pathogenic mycobacterial species. It grows in a wide range of temperatures and readily accepts plasmids and genetic manipulation. As a model organism, the genome of *M. smegmatis* $mc^2155$ is also well-characterized. Over 2,800 of its genes have orthologs with >50% amino acid identity in *M. tuberculosis.* Therefore, this species is often used in the laboratory as a tool for training new students, for testing new molecular methods, and for studying genes shared in mycobacterial species that are more difficult to work with[35].

*M. marinum,* as its name suggests, is often associated with aquatic environments. It was first isolated from diseased fish in 1926 and was noted to cause tuberculous lesions in their spleen and liver with occasional involvement of other parts of the body such as the gills[36]. In humans, it can cause granulomatous skin lesions typically in people with an open wound who

take care of fish tanks or are spending a lot of time doing waterborne activities[37]. *M. marinum* can also be used as a model for *M. tuberculosis* because of its genetic similarity and ability to cause disease. The genome of *M. marinum* shares 3,000 orthologs with *M. tuberculosis* with >85% amino acid identity[38] and experimental infection of fish with *M. marinum* models tuberculosis pathogenesis[37].

*M. ulcerans* is the most genetically similar species to *M. marinum,* sharing >98% of the nucleotide sequence[39]. *M. ulcerans* is the causative agent of the Buruli ulcer. This disease causes painless ulcers to develop on the skin which can be treated with surgery. Cases of Buruli ulcer are most often seen in riverine areas with warm and humid climates. It is still not entirely understood how the disease is transmitted to people, but *M. ulcerans* has been recovered from mollusks and fish[40].

*M. haemophilum,* like both *M. marinum* and *M. ulcerans,* is associated with skin lesions and prefers a lower temperature (30-32°C) for growth *in vitro. M. haemophilum* requires the presence of hemin or large amounts of ferric ammonium citrate for laboratory growth due its inability to produce the iron-binding siderophore mycobactin[41,42]. This NTM typically causes cutaneous lesions on the extremities of severely immunocompromised individuals, such as those with HIV or those who recently received an organ transplant. In rare cases, *M. haemophilum* has caused infection in other locations such as the skeletal muscle and bone[41]. Whole-genome sequencing of this bacteria revealed a close genetic relatedness to *M. leprae.* Since *M. leprae* is uncultivatable, *M. haemophilum* has potential to be used as a model to study their shared genes[42].

Finally, the *M. avium* complex (MAC) consists of 12 mycobacterial species with >99.4% sequence identity of the 16S rRNA gene and >97.3% sequence identity of the hsp65 gene[43]. Most of its members are environmental bacteria capable of causing opportunistic infections. It

includes *M. intracellulare* which is often found in soil and has been shown to cause pulmonary disease in patients with pre-existing lung disease such as chronic obstructive pulmonary disease (COPD) or those with acquired immunodeficiency syndrome (AIDS)[44,45]. Another member *M. columbiense* was first described in 2006 after its isolation from HIV-infected patients with severe pneumonia[46]. The complex also includes *M. chimera* which was linked to a series of endocarditis cases following contamination of equipment used in heart surgery[47]. Finally, the 4 subspecies of *M. avium* are also a part of this complex which includes *M. avium* subspecies *paratuberculosis* (*MAP*), a well-known cattle pathogen[48].

Overall, this thesis will consider 4 bovine or bovine-derived mycobacteria with the potential to cause disease in humans: 3 MTBC subspecies (*M. bovis, M. bovis* BCG, *M. orygis*) and 1 NTM (*MAP*).

## 3. Zoonotic tuberculosis

This section discusses zoonotic tuberculosis, which is caused by the transmission of an MTBC subspecies from an animal to a human. One of the most likely routes of zoonotic TB infection is thought to occur through drinking unpasteurized milk from bovine TB-infected animals. Zoonotic TB is defined by the World Health Organization (WHO), the World Organisation for Animal Health (WOAH), and the Food and Agriculture Organization (FAO) as a form of TB caused by *M. bovis*[49]. However, as it will become clear throughout this thesis, other animal-associated members of the MTBC such as *M. orygis* are also capable of causing bovine and zoonotic TB[50].

*3.1 Mycobacterium bovis*

The discovery and characterization of *M. bovis* was a significant milestone in the history of microbiology and had important implications for future research. In 1882, Robert Koch reported his isolation of the tubercule bacillus, the causative agent of TB, from both human and bovine samples[51]. In 1898, Theobald Smith published data demonstrating differences in bacterial characteristics and virulence between the bovine and human tubercule bacillus. He stated his belief that the use of the names bovine tubercule bacillus and human tubercule bacillus may be misleading because it may incorrectly suggest that one bacillus is incapable of infecting the opposite host[52,53]. However, Koch proclaimed in 1901 that the bovine tubercule bacillus was not able to infect humans and therefore diseased cattle did not present a significant threat to human health[52,54]. This controversy was the subject of considerable debate over the next decade. Although Koch's theory was later disproven, this idea had a lasting effect on the views of zoonotic diseases and their ability to transmit between humans and animals for many decades[52]. A Royal Commission proved that the bovine tubercule bacillus could in fact infect and cause TB in humans, however this was not the typical cause of human TB[55]. This Royal Commission was the first of its kind to not only seek the opinions of scientific experts but also fund additional research. This became the template for future committees[52].

The bovine tubercule bacillus was eventually renamed to *M. bovis* in 1970 by Karlson and Lessel[52,56]. In 1997, the reference strain *M. bovis* AF2122/97 was isolated from a cow sick with bovine TB. This strain was fully sequenced and annotated in 2003. Although *M. bovis* shares >99.95% of its genome with *M. tuberculosis*[57], there are several basic laboratory tests which have historically been applied to differentiate the 2 subspecies. In early TB research, the bovine tubercule bacillus was differentiated from the human tubercule bacillus based on their

pathogenicity in different animals. It was found that both bacilli were virulent in guinea pigs, but the human tubercle bacillus had less severe outcomes in rabbits[58]. Later, biochemical tests were applied to differentiate between the 2 subspecies. For example, *M. bovis* strains are microaerophilic, meaning they prefer to grow just beneath the surface of Lebek semisolid agar whereas *M. tuberculosis* will grow at the surface[59]. A test more commonly applied to distinguish *M. bovis* is susceptibility to pyrazinamide, a frontline antibiotic used to treat TB. Most strains of *M. bovis* are intrinsically resistant to pyrazinamide whereas *M. tuberculosis* is susceptible[60]. However, use of any one of these kinds of tests is not a reliable method to definitively identify *M. bovis* as variation does exist among strains[61]. These tests are also time-consuming due to the slow growth rate of *M. tuberculosis* and *M. bovis.* The doubling time is ~24 hours for *M. tuberculosis* and 16-20 hours for *M. bovis*[62,63]. Today, it is much more common to use a molecular approach. Within the MTBC, there are genomic regions named regions of difference (RDs) which vary between subspecies and can be used to differentiate between different branches of the complex[64,65]. Previous studies have shown that RDs can be used to reliably identify *M. bovis* from clinical samples[66–68].

As its name suggests, *M. bovis* has primarily been isolated from bovines and is commonly identified as the cause of bovine TB in Africa and Europe. However, *M. bovis* has also been isolated from a wide variety of other animals including badgers, elk, wild boars, possums, cats, chimpanzees and humans[69].

### *3.2 Mycobacterium orygis*

*M. orygis* is a lesser-known, understudied animal-associated member of the MTBC. This status may in part be explained by the fact that *M. orygis* is relatively new to the scientific

community compared to *M. bovis.* The oryx bacillus was first isolated from an oryx, which is a species of antelope, within a zoo in 1976. Researchers noted that the bacteria were pathogenic to guinea pigs but not to rabbits[70]. In 1994, van Sooligan and colleagues published that they thought the oryx bacillus to be a subgroup of *M. bovis*[71]. However, it wasn't until 2012 that the oryx bacillus was characterized as a phylogenetically distinct member of the MTBC and renamed to *M. orygis* by van Ingen and colleagues[72].

The reference strain *M. orygis* 51145 was originally collected in Quebec and was sequenced and annotated in 2021[73]. *M. orygis* has a unique pattern of RDs which can be used to differentiate it from *M. bovis* and other MTBC subspecies[72]. It also carries specific single nucleotide polymorphisms (SNPs) which are unique. For example, the G698C mutation in *Rv0444c* causes *M. orygis* to produce high levels of the antigenic proteins MPB70 and MPB83. *M. bovis* and some strains of BCG also produce high levels of these proteins, but through independent mutations (C320T and C551T)[74]. The strains of *M. orygis* studied by van Ingen often produced a spoligotyping pattern previously labeled as *M. africanum,* suggesting it may have been mislabeled in previously published studies. Like *M. bovis*, it has a wide host range with isolates being identified in antelopes, cattle, waterbucks, rhesus monkeys, and humans. But unlike *M. bovis, M. orygis* is susceptible to all frontline antibiotics against tuberculosis including pyrazinamide[72].

As more data accumulated identifying *M. orygis* in animals and humans, a trend began to emerge which suggested that *M. orygis* may be endemic to South Asia. When *M. orygis* was characterized in 2012, all isolates collected were derived either from animals in zoos or humans originating from South Asia (one was from Southeast Asia). This was followed by a report of a woman originally from India which transmitted *M. orygis* to a cow in New Zealand[75]. *M. orygis*

was then isolated from a rhinoceros, blue bull, and spotted deer in Nepal[76,77]. In 2017, 8 more isolates were identified in New York, all of whom migrated from India, Pakistan, or Nepal[78]. Finally, *M. orygis* was reported from 18 cattle and 2 rhesus monkeys in Bangladesh by Rahim and colleagues[79]. These findings suggesting the endemicity of *M. orygis* in South Asia are what initiated the project outlined in chapter II of this thesis where we used molecular assays to screen cultures from TB patients for zoonotic TB in India. Our study identified cases of zoonotic TB caused by *M. orygis* without any evidence of disease caused by *M. bovis*[50]. Since its publication in 2020, several additional studies have been conducted which continue to support this theory. *M. orygis* was isolated in several species of ungulates in India and 5 additional cases of human *M. orygis* infections were identified in Norway in patients from South Asia[80–82].

There are 2 circulating theories to explain the distribution of *M. orygis* in South Asia and *M. bovis* in other areas of the world. The first is that there were 2 independent cattle domestication events which resulted in *M. bovis* becoming the pathogen of *Bos taurus* which is the predominant species of cattle found today in Europe and the Americas and *M. orygis* becoming the pathogen of *Bos indicus* which is more commonly found in South Asia[26]. The second theory is that *M. orygis* emerged first and traveled to South Asia with the migration of humans before the establishment of *M. bovis*[79]. Considering that South Asia reports ~4.02 million (~38%) of the world's TB cases and screening for zoonotic TB is not routinely done, the threat posed by *M. orygis* to humans requires further investigation[23,83].

### *3.3 Bovine tuberculosis*

Like human TB, bovine TB is a chronic disease which primarily affects the lungs. The disease can be spread by infected air droplets or by ingesting milk from an infected cow. There

are no published studies to date concerning *M. orygis* pathogenesis or experimental infections, but there are many studies using *M. bovis.* It is understood that *M. bovis* is primarily engulfed in the phagosome of alveolar macrophages. *M. bovis* will block phagosome maturation and use virulence factors such as early secreted antigenic target (ESAT)-6 to lyse the phagosome and exit into the cytosol to replicate. The infected macrophage will ultimately burst and release cytokines and chemokines which recruit additional immune cells which will also be infected with the bacteria. Some infected cells spread to local lymph nodes and present antigen to recruit lymphocytes to the infection site. This leads to the eventual development of caseating granulomas which contain a necrotic core surrounded by macrophages, multinucleated giant cells, and lymphocytes[84]. Diseased animals experience cough and decreasing lung function, loss of appetite, weight loss, and weakness leading to eventual death[85].

The presence of bovine tuberculosis is a threat both to surrounding livestock as well as to public health due to the potential for zoonotic infection. However, with the implementation of national control programs, bovine TB is well-controlled in many countries. These programs primarily rely on a test and cull strategy to maintain healthy herds. However, the presence of wildlife hosts may prevent the complete eradication of the disease. For example, in the United Kingdom and the Republic of Ireland, badgers are well-known to act as a reservoir for *M. bovis* and have thwarted efforts to eliminate the disease within the cattle population[86]. Similarly in New Zealand, the Australian brushtail possum acts as a maintenance host for *M. bovis* and has created numerous hurdles for controlling the disease in livestock[87].

Around the world, bovine tuberculosis remains a considerable problem and a major economic burden. A 2020 meta-analysis estimated the pooled global prevalence to be 13.12%[88]. In India, which has the world's largest cattle population, there are an estimated 21.8 million

animals infected with bovine TB[89]. The factors associated with economic costs of disease vary by country. In countries where bovine TB prevalence is low due to the presence of national control programs, economic costs are primarily related to testing, program maintenance, and trade barriers. In countries which lack these control programs due to economic or social constraints, bovine TB costs are primarily associated with reduced production and animal mortality[90]. In Ethiopia, which does not have a control program in place, the economic costs associated with bovine TB within the dairy industry was estimated to be 132 million USD over 10 years[91]. In contrast, in the United Kingdom, which does have a control program, there are £120 million in costs per year associated with the disease[92].

*3.4 Zoonotic tuberculosis*

The concern that TB could be transmitted from infected bovines into humans has existed since the discovery of the tubercle bacilli. The process of pasteurization was developed in 1864 originally to remove contaminating microorganisms from wine and beer, but it was later applied to decontaminate milk[93]. About 40% of animals slaughtered during this time had clear signs of tuberculosis[52]. Therefore, prior to the widespread adoption of pasteurization, milk presented a significant risk to the development of TB, particularly in young children where it often presented as extrapulmonary disease[55,61]. It was estimated that bovine TB caused ~25% of child TB cases during this time[94]. In 1937, Dr. Stanley Griffith reported that 91% of cervical lymph node TB cases were caused by bovine TB in children under 5 years old[95].

In many countries today, pasteurization is required by law and bovine TB is managed through national control programs, which minimize the risk of zoonotic TB. However, zoonotic TB is still a concern in many places around the world where bovine TB remains prevalent or

where drinking unpasteurized milk is still a common practice. In 2019, the WHO estimated that there were 140,000 TB cases caused by *M. bovis.* The prevalence was estimated to be highest in Africa (68,900 cases) followed by South-East Asia (43,400 cases)[96]. However, it is likely that these numbers underestimate the total number of cases because surveillance for zoonotic TB is not done in many high burden countries and this estimate did not account for the contribution made by other animal associated MTBC members. A 2021 meta-analysis of zoonotic TB cases caused by *M. bovis* reported that prevalence estimates among studies were highly variable (ranging 0.4-76.7%) and that this was dependent on the diagnostic techniques applied[97].

Zoonotic TB in humans can present with many of the same symptoms as infection with *M. tuberculosis* such as cough, fever, weight loss, and night sweats[98]. However, as mentioned, zoonotic TB is more often associated with extrapulmonary disease. In a systematic review of primary infection sites with *M. tuberculosis* vs zoonotic TB (*M. bovis* or *M. caprae*), the median proportion of extrapulmonary disease in zoonotic TB cases was over 2-fold higher than infections with *M. tuberculosis.* The median proportions of affected sites were 30% lymph nodes, 25% genitourinary system, 13% bones and joints, 6% intestines and peritoneum, and 5% nervous system[99]. Although less information about the clinical features of *M. orygi*s is currently known, in a recent study of 8 cases of *M. orygis* in India, all had extrapulmonary disease and 2 had concomitant lung involvement. Of the 8 cases, *M. orygis* was shown to affect the spine (3), genitourinary system (1), lymph node (2), abdomen (1), and spleen/liver (1)[100]. The higher frequency of extrapulmonary disease in zoonotic TB poses several challenges to its surveillance and control. Extrapulmonary disease is variable and can therefore be more difficult and require more time for a diagnosis. It is also much more challenging to acquire samples from extrapulmonary disease sites compared to sputum. Therefore, in many clinics in high burden

countries, culturing of extrapulmonary TB is not routinely done[101]. If zoonotic TB is more often extrapulmonary, it may be missed more frequently due to lack of surveillance.

To treat TB, typically patients are prescribed a 6-month regimen of anti-tubercular therapy (ATT) including rifampicin, isoniazid, pyrazinamide, and ethambutol. Since *M. bovis* is intrinsically resistant to pyrazinamide, patients are recommended to be treated with a 9-month regimen of rifampicin, isoniazid, and ethambutol[102]. In cases of extrapulmonary TB which involve the central nervous system, bones or joints the treatment recommendation is extended to a total of 9-12 months[103]. In a study of 110 patients infected with *M. bovis* treated with these 3 drugs in The Netherlands, the intention-to-treat success rate was 83%[102,104]. This is similar to the global success rate of ATT against *M. tuberculosis* infection which was 86% in 2020[23]. Information regarding treatment of *M. orygis* is limited, but in the aforementioned study of 8 extrapulmonary cases in India, 2 were lost to follow-up, 4 were successfully treated with ATT for 9-18 months, and 2 were successfully treated with ATT for 12-23 months and surgery[100]. Treatment for drug susceptible pulmonary TB infections is known to be a long and difficult process due to the large number of medications prescribed and their various side effects. Treatment of zoonotic TB, although a rarer circumstance, likely poses further challenges due to its proclivity to cause extrapulmonary disease and the intrinsic antibiotic resistance of *M. bovis.*

*3.5 Control strategies*

In 2015, the WHO set the ambitious goal of reducing TB incidence rates by 90% by 2035[105]. As of 2022, incidence rates had only dropped by 10%[23]. To complete this goal, massive efforts need to be made to reduce transmission of the disease. However, just as badgers and brushtail possums prevent elimination of *M. bovis* in cattle populations in the United Kingdom

and New Zealand[86,87], the persistence of TB in bovines may create additional hurdles to achieve these large reductions in TB incidence in humans. This situation is further complicated by recent reports which have isolated *M. tuberculosis sensu stricto* from bovine TB-infected animals in Africa and South Asia[106,107]. There is the possibility these cases occurred by reverse zoonosis where a human infected with TB transmitted disease to the animal. However, *M. tuberculosis* may have also been transmitted animal-to-animal or animal-to-human. If *M. tuberculosis* is circulating among cattle herds and being transmitted back to humans, this could confound our ability to accurately identify cases of zoonotic TB and reduce potential routes of exposure.

Zoonotic TB control requires the application of a One Health approach where information from human, veterinary, and environmental health are simultaneously taken into account to conduct research and design strategies for disease surveillance and management[2,108]. When applying One Health to zoonotic TB, this includes concepts such as improving detection through molecular-based diagnostics (human health), monitoring dairy products for TB contamination (animal health), and managing interactions between livestock and environmental or wildlife reservoirs (environmental health)[109]. However, the concept of One Health is frequently overlooked. In a systematic review of zoonotic TB surveillance programs, only 4/119 countries investigated had implemented programs which considered One Health practices[109]. It is also important that the interconnectedness of human, animal, and environmental health is well understood by animal caretakers and the public. A 2021 study conducted in Ethiopia, which has the largest cattle population in Africa and a heavy burden of bovine TB, revealed that 58.5% of farmer owners/manager screened were not aware of the zoonotic risks posed by bovine TB[110]. A more comprehensive approach must be implemented to better understand and manage zoonotic TB and its contribution to the global TB burden.

## 4. BCG disease

This next section will discuss the vaccine strain used against tuberculosis which was originally derived from *M. bovis* and which can cause disease in certain individuals. The importance of BCG disease detection in these individuals will be explored in chapter III of this thesis.

### *4.1 Mycobacterium bovis bacille Calmette-Guérin*

*M. bovis* BCG is the vaccine strain against tuberculosis originally developed by Albert Calmette and Camille Guérin[30]. It is one of the most widely used vaccines with over 100 million doses given each year[111]. BCG is an attenuated strain of *M. bovis* that was originally isolated from a diseased cow in 1902[112]. It has been suggested by some that the selection of *M. bovis* as the starting point for the development of an attenuated strain originally stemmed from Robert Koch's erroneous statement that the bovine tubercule bacillus was incapable of causing disease in humans[61]. Calmette and Guérin had discovered that growing *M. bovis* in glycerol and potato medium with the addition of ox bile led to a less virulent version of the bacteria. They then continued to passage *M. bovis* in this media roughly 230 times at which point the bacteria no longer caused tuberculosis in infected animals[30,113]. The full mechanism of attenuation is still not entirely understood, but loss of RD1 which includes virulence factors ESAT-6 and culture filtrate protein (CFP)-10 have been shown to reduce virulence[114]. In July of 1921, an infant was given the first vaccination with BCG in Paris. Over the next several years, over 100,000 infants were vaccinated[30].

Studies measuring the efficacy of BCG vaccination have been varied. In the early 1950s, 54,239 participants in the United Kingdom were vaccinated with BCG. The participants were

followed over 20 years and researchers reported a 77% efficacy throughout the study[115]. Conversely, a study conducted around the same time of 64,136 participants in the United States found next to no benefit of BCG[116,117]. The results from these studies impacted the policies enacted for BCG vaccination in the respective countries. In the United States, routine BCG vaccination was not administered, and BCG was instead given to those deemed high risk. However, in the UK and many other countries around the world, BCG vaccination was routinely done[30]. In 2022, a systematic review reported BCG to have a 18% effectiveness, with protection only significant in children under 5 years old[118]. Evidence of BCG protection against miliary and meningeal TB in particular has supported its continued use in many countries today[111,119].

In 1930, 251 newborns were vaccinated with BCG in Lübeck, Germany. Of those, 228 infants later developed signs of TB and 72 (29%) died of the disease. Although it was later determined that the BCG vaccine had been contaminated with *M. tuberculosis* rather than the BCG strain reverting to virulence, the safety of the vaccine was severely criticized following this event[120]. Today, it is understood that BCG is associated with varying degrees of adverse events with low incidence rates. This includes superficial reactions at the injection site, cutaneous lesions distinct from the injection site, and lymphadenitis[121]. A 2020 study from Korea which reviewed the incidence of BCG vaccine adverse events over 5 years determined that lymphadenitis was the most common reaction (18.2 cases per 100,000 doses)[122]. Systemic adverse events are rare but may also occur. Among these are osteitis, osteomyelitis , and finally disseminated BCG disease[121], an outcome which will be discussed in more detail in the next section.

There has been some variability observed in BCG efficacy and safety, and one theory which may explain this phenomenon stems from the fact that there are 14 different BCG strains.

Rather than a single strain being developed and replicated throughout the world, BCG was distributed to labs within different countries who maintained the bacteria using varying media. This resulted in the eventual emergence of different BCG strains with significant heterogeneity among them[30,113]. Some have suggested that certain BCG strains may provide more protection or may be associated with more adverse events than other strains. BCG strains can be divided into what are known as "early" and "late" strains. The early strains are those which were obtained from the Pasteur Institute before 1927 and includes commonly used strains such as BCG Russia and BCG Japan. The late strains are those that were obtained after 1927 and includes strains such as BCG Danish and BCG Pasteur. One of the key differences between early and late strains is that early strains produce high levels of the antigenic proteins MPB64, MPB70, and MPB83, which suggests they could be more immunogenic. Reports comparing efficacy and safety between strains have been variable. A 2020 study comparing BCG Russia, Danish, and Japan found no differences in morbidity or mortality among strains[123]. In the previously aforementioned study in Korea, adverse events were 41.6 cases per 100,000 doses following BCG Danish vaccination versus 25.9 cases per 100,000 doses following BCG Japan vaccination[122].

BCG has also been investigated as a potential therapeutic or preventative measure for aliments unrelated to TB. A meta-analysis of BCG clinical trials found that vaccination reduced all-cause mortality in neonates by 38%[124]. It is thought that BCG leads to off-target immunological effects which can provide non-specific protection, however the exact mechanisms of how this occurs is still uncertain[125]. In 1959, Old and colleagues found that BCG vaccination was associated with increased resistance to tumor transplantation in mice[126,127]. This led to the investigation of BCG as an immunotherapy against acute lymphoblastic leukemia,

melanoma, and eventually non-muscle-invasive bladder cancer (NMIBC)[128–130]. Today, treatment of intermediate to high risk NMIBC tumors with BCG is recommended to reduce progression and recurrence of disease[131]. BCG vaccination was also recently explored as a method to improve COVID-19 outcomes. Unfortunately however, no significant differences were found from the placebo group in a 2023 clinical trial[132]. Finally, since work with BCG does not require a biosafety level 3 laboratory like other MTBC subspecies, BCG is also useful as a safer organism to test new methods and to study MTBC shared genes and mechanisms[133].

*4.2 Clinical features of disseminated BCG disease*

Development of disseminated BCG disease is a rare event, only occurring at an incidence of 0.06-3.4 cases per million[122,134,135]. However, when it does occur it is associated with severe morbidity and a high mortality rate (~60%)[134,136,137]. The disease typically develops within a few months from birth and subsequent vaccination[122,136]. Disseminated BCG disease is suspected in young patients under 2 years of age who are experiencing symptoms such as fever, weight loss, abscess, failure to thrive, and lymphadenopathy. This should be followed by radiologic assessment and a biopsy to be sent for culture and histopathology. A diagnosis of disseminated BCG is confirmed by systemic symptoms which involve 2 or more sites beyond the vaccination site along with microbiological and/or histopathological identification of BCG[136].

Since the disease is systemic, it is associated with a wide variety of clinical symptoms and can affect a variety of organ systems. In a study of 44 cases of disseminated BCG disease occurring over 10 years in Iran, the most common clinical features observed were lymphadenopathy (86%), abscess (84%), abnormal levels of C reactive protein (CRP) (84%), abnormal abdominal sonography (75%), abnormal erythrocyte sedimentation rate (ESR) (70%),

hepatomegaly (68%) splenomegaly (66%), and weight loss (64%)[138]. Less common clinical features that have been observed in cases of BCG disease include osteomyelitis, meningitis, and nephritis[137,139].

*4.3 Risk factors for disease development*

Primary immunodeficiency diseases (PIDs) are a collection of inherited immune disorders which are often associated with recurrent infections, adverse events following vaccines, and certain types of cancers. BCG vaccination is contraindicated in individuals with PIDs because they are much more likely to suffer from BCG disease[140]. In a systematic review of 1,691 individuals with PIDs who were vaccinated, 41.5% were diagnosed with BCG disease[141]. However, these disorders are often undiagnosed prior to BCG vaccination which is routinely done shortly after birth in many countries as per national immunization programs[96]. The PIDs which are more commonly associated with BCG disease are severe combined immunodeficiency disorder (SCID), chronic granulomatous disease (CGD) and Mendelian susceptibility to mycobacterial diseases (MSMD)[140].

SCID patients lack functional T lymphocytes and are sometimes deficient in B lymphocytes and NK cells as well, which impedes their ability to combat infections and respond appropriately to live attenuated vaccines such as BCG. There are a number of genetic mutations which have been associated with the development of SCID, but no studies have yet associated specific polymorphisms with different outcomes of BCG disease[140]. In a study of 349 patients with SCID who were vaccinated with BCG at birth, 118 (34%) developed BCG disease and 46 (13%) died[142].

CGD is caused by a mutation which impairs the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase enzyme complex[143]. This prevents the production of reactive oxygen species which are important for killing of microorganisms by macrophages, monocytes, and neutrophils[140]. In a study of 23 patients with BCG disease and CGD, the 10-year survival rate was 34%. The probability of survival was shown to be strongly correlated with the degree of dissemination, where mortality rate increased with the number of different organ systems affected[143].

MSMD is caused by mutations affecting the secretion of IFN-γ either directly by affecting the IFN-γ receptor or indirectly via genes such as the IL-12 receptor. BCG, like many mycobacterial species, infects cells intracellularly. In response, the host mounts a Th1 immune response mediated by IFN-γ and other cytokines to promote phagosomal killing and production of cytotoxic nitric oxide[144]. As the name suggests, deficiencies in this IFN-γ response are therefore associated with increased susceptibility to disseminated BCG or NTM infections as well as infection with *M. tuberculosis*[140]. The frequency of BCG disease has been shown to be highest in cases of MSMD compared to the other PIDs[141].

In addition to PIDs, HIV status is another major risk factor for the development of BCG disease post-vaccination. HIV is known to infect CD4+ T lymphocytes and lead to deficient immune responses to secondary infections[145]. The chance of transmission of HIV from an infected mother to her infant prior to or during birth (vertical transmission) is about 15-45%. With anti-retroviral therapy (ART), this risk of transmission is greatly reduced[146]. However, vertical transmission continues to frequently occur in many high-burden countries. In 2018, approximately 160,000 infants were infected with HIV, primarily in sub-Saharan Africa where TB is prevalent and BCG vaccination is routinely administered[23,145,147]. In a 2004-2006 study

conducted in South Africa, the incidence of BCG disease was 992 per 100,000 HIV-infected babies[148]. This is approximately 10,000-fold greater than the general population[122,134,135].

*4.4 Treatment and follow-up*

To treat BCG disease, both the mycobacterial infection and the immune status of the patient should be considered. Like its predecessor *M. bovis,* the BCG vaccine strain is resistant to pyrazinamide. There are no official treatment guidelines listed to treat BCG. However, case reports usually describe a minimum treatment with the anti-tuberculosis drugs rifampicin, ethambutol, and isoniazid. Other drugs have also frequently been added to the regimen such as an aminoglycoside (e.g. streptomycin, amikacin) or a fluoroquinolone (e.g. levofloxacin, moxifloxacin)[136,143,149]. There have been several documented cases of resistance to isoniazid, which may necessitate the inclusion of these other drugs in the regimen[143,150]. Antibiotic therapy may also be combined with surgical intervention to drain abscesses or excise infected lymph nodes[150]. Like treatment for TB, BCG treatment times are long, with case reports ranging from 6 to 25 months[136,150].

In a study of 134 cases of BCG disease in China, over 80% of them had an identified PID[151]. Thus, suspicion or diagnosis of disseminated BCG should spur a clinical investigation into the immune status of the patient. Patient outcomes have been shown to improve with early intervention of therapies directed at those with PIDs[140]. For example, in patients with MSMD and CGD, the administration of recombinant IFN-γ can be an effective method to help them mount proper immune responses against an infection. This only works in cases where the IFN-γ receptor is intact[152]. Case studies with comparisons of the effect of IFN-γ treatment versus no treatment for BCG disease survival are limited in size. In one study of 22 CGD patients, 11 of

which received IFN-γ therapy, there was no significant differences in survival between the 2 groups[143]. However, in a second study of 19 patients, 7 of which received IFN-γ therapy, only those who received therapy were cured within one year of treatment. Of the remaining 12 patients, 3 died and the remaining were cured only after IFN-γ therapy was started over a year after the start of antibiotic treatment[153].

Hematopoietic stem cell transplant (HSCT) offers a potentially curative treatment to those with certain PIDs including SCID, CGD, and MSMD[154,155]. If the genetic mutations causing disease are intrinsic to the hematopoietic stem cells, it is possible to cure the disease by replacing these cells with those without the mutation. This process involves the ablation of the patient's bone marrow using chemotherapy followed by donation of hematopoietic stem cells from a matched donor. This process has improved in recent years with new technology to more easily find donors and reduce the chance of graft versus host disease[155]. In 31-year retrospective in Germany, HSCT was found to be curative in most SCID patients with BCG disease. However, the authors noted that BCG disease was likely to worsen following HSCT and thus cases must be carefully monitored and treatment must be paired with intensified antibiotic regimens and other therapies.

## 5. Paratuberculosis and Crohn's disease

The final bovine pathogen to be discussed in this thesis is *MAP,* the cause of paratuberculosis in ruminants and a hypothesized cause of Crohn's disease in humans. This pathogen has particular relevance to Canada which has a high incidence rate for both of these 2 diseases.

*5.1 Mycobacterium avium subspecies paratuberculosis*

A major cattle NTM with potential for zoonosis is *MAP*. *MAP* is an obligate intracellular pathogen known to cause a deadly chronic gastrointestinal disease called paratuberculosis[156]. The disease was first described in 1895 by Johne and Frothingham in a cow that exhibited signs of a wasting disease and had an inflamed intestine infiltrated with abundant acid-fast bacteria. They originally though it died due to intestinal TB, but an infected sample failed to produce disease when injected into guinea pigs[157]. The bacteria was first isolated in 1912 by Twort and Ingram who originally named the bacteria *M. enteritidis chronicae pseudotuberculosae bovis* Johne[156]. The name *MAP* was not given until 1990 and the reference genome *MAP* K10 was sequenced in 2005[158,159]. In addition to the impact of *MAP* on animal health, *MAP* is of zoonotic concern because there is a highly debated theory that it may be the etiological agent of Crohn's disease in humans. The evidence concerning this potential link will be discussed in detail within this chapter.

*MAP* is an intracellular pathogen of intestinal macrophages. It invades the intestine often through microfold cells found in Peyer's patches of the gut-associated lymphoid tissue (GALT). Once in the sub-epithelium, *MAP* will typically invade macrophages. The best characterized interaction for how this occurs is through binding of its ManLAM with the mannose receptor of the macrophage. *MAP* has also been shown to invade dendritic cells through interaction of ManLAM and DC-SIGN. Once intracellular, *MAP* is taken up into the phagosome and blocks phagosome acidification and fusion with the lysosome[160]. Infected macrophages may also travel by the afferent lymphatic vessels into the mesenteric lymph nodes where *MAP* is able to persist[161]. Initially, an appropriate Th1 immune response is mounted against the bacteria and the disease remains subclinical. However, it is thought that the transition to clinical disease is

instigated by a change from a Th1 to a Th2 response[162]. *MAP* infections can be multibacillary where there are large numbers of macrophages infected with many acid-fast bacteria, or they can be paucibacillary where there are few bacteria per macrophage and typically more T cells. Both of these forms can lead to the same gross pathological and clinical outcomes[163].

*MAP* is one of the slowest growing cultivatable mycobacteria with a generation time ranging from 1.3 to 4.4 days[164]. As mentioned, *MAP* is a member of the *M. avium* complex and is 1 of 4 subspecies of *M. avium*. The other 3 members are *M. avium* subspecies *avium* (*MAA*) which causes TB in birds, *M. avium* subspecies *silvaticum* (*MAS*) which is rare and causes TB in wood pigeons, and *M. avium* subspecies *hominissuis* (*MAH*) which is primarily environmental but can cause opportunistic infections in pigs and humans[48]. *MAH* is the closest relative subspecies of *MAP.* When *MAP* emerged from *MAH* it gained 125 kB of genes which are found on genomic islands called large sequence polymorphisms of paratuberculosis ($LSP^P$s). There are 6 in total: $LSP^P$4, $LSP^P$11, $LSP^P$12, $LSP^P$14, $LSP^P$15, and $LSP^P$16[165]. Considering that *MAP* is an obligate pathogen and its predecessor *MAH* is environmental, the genes unique to *MAP* found on $LSP^P$s may carry important information which allow the bacteria to cause disease.

*5.2 Paratuberculosis*

Paratuberculosis is a disease which affects ruminants such as cattle, sheep, goats, and deer. It is sometimes called Johne's disease when referring to disease specifically in cattle. Paratuberculosis is most often spread by the fecal to oral route, but it has also been shown to spread in the milk, colostrum, and *in utero*[166]. It is a slowly progressing condition where clinical symptoms do not typically arise for 2-5 years post-infection[167]. When disease manifests, it is characterized by granulomatous enteritis with a thick corrugated intestinal wall most often in the

terminal ileum paired with enlarged mesenteric lymph nodes and edema[168]. Malabsorption occurs and infected animals will experience symptoms including weight loss, diarrhea, reduced milk production, and eventual death[166,168].

Paratuberculosis is widespread around the world and a major economic burden. Herd-level prevalence of paratuberculosis varies by region and type of animal but estimates range from 10% to 70%. Typically beef herds have lower prevalence of the disease than dairy herds which is likely a result of their shorter life spans[166]. Paratuberculosis has particular relevance in Canada where an estimated 46% of dairy herds are positive for *MAP*[169]. This burden is associated with a cost of ~$28 million (USD) per year due to reduced milk production and premature culling of infected animals[170]. In the United States, where approximately 70.4% of dairy herds are positive for *MAP,* this number is as high as $198 million per year[170,171].

Like bovine TB, paratuberculosis control relies on a test and cull strategy to maintain healthy livestock. Despite the presence of paratuberculosis control programs based on this practice in use in Canada, the disease continues to persist[172]. There are several major factors which contribute to this. The first is that paratuberculosis remains subclinical for years prior to disease manifestation. During this time, the animal does not display any obvious signs of infection, but may still be excreting bacteria into the environment which could infect other animals in the herd[166]. The second related factor is the limitation of the diagnostics for paratuberculosis. There are several tests available to detect *MAP* including fecal culture and PCR or ELISA of the serum or milk. However, these tests are either time-consuming (fecal culture requires months) or have poor sensitivity or specificity, particularly for subclinical animals[166,173]. Additionally, although *MAP* does not replicate outside of a host, it is capable of persisting for up to 55 weeks in environment[174]. There is therefore the risk of re-introduction of *MAP* into

previously negative herds through environmental exposure. Considering these circumstances, it remains very difficult to identify and maintain all animals within a herd as *MAP* negative to eliminate the disease.

*5.3 Crohn's disease*

Crohn's disease and ulcerative colitis make up the 2 main forms of inflammatory bowel disease (IBD). Crohn's disease is a chronic inflammatory condition which can affect the entire gastrointestinal tract. Patients with Crohn's often experience frequent abdominal pain, diarrhea, and perianal lesions. Pathologically, the disease is characterized by non-continuous "skip" lesions of transmural inflammation along the intestine potentially accompanied by creeping fat. More severe complications requiring surgical intervention can also arise such as the formation of strictures which can cause obstructions or fistulae which cause intestinal contents to leak. These symptoms of Crohn's disease are difficult to manage and can have a severe impact on a patient's quality of life[175].

Canada has one of the highest prevalence values of Crohn's disease globally (up to 319 per 100,000 people), second only to Germany (322 per 100,000 people). Notably however the data is limited in countries within Africa and Asia[176]. Crohn's disease has been found to occur more frequently in Ashkenazi Jewish people, those living in cities, and those living in the northern hemisphere. Additionally, risk of Crohn's disease has been shown to increase in those who smoke, those who were exposed to antibiotics early in life, and those who take oral birth control medication[175]. However, the cause of Crohn's disease is unknown. Research suggests that it is a complex disease which arises from a combination of host and environmental factors including genetic predisposition, immune dysregulation, and a microbial trigger[177–179]. *MAP* is

one bacterium suggested to play a role in its etiology, however adherent-invasive *Escherichia coli* (AIEC) and microbial dysbiosis have also been explored[180,181].

To treat Crohn's disease, there are several therapeutic options. Factors including severity of disease and history of the patient are considered to determine which of the available options would be the best fit. Corticosteroids, which have anti-inflammatory and immunosuppressant effects, have been a mainstay in Crohn's disease care and are recommended to treat a flare up of mild to moderate severity. However, they are associated with a variety of side effects and are not recommended for the long-term maintenance of remission[182]. Instead, patients may be prescribed immunosuppressants such as azathioprine, which have been shown to be effective at maintaining remission but not statistically better than placebo at inducing remission. Biologics such as the anti-TNF drugs infliximab (Remicade) and adalimumab (Humira) are recommended for higher risk patients since they are more effective at maintaining remission than azathioprine. Biologics may be also be used for remission induction[175]. Unfortunately, 40-50% of patients who initially respond positively to anti-TNF therapy will stop responding 6-12 months post-treatment. There are several newer biologics which may provide an alternative option such as ustekinumab (Stelara) which targets IL-12 and IL-23 and vedolizumab (ENTIVIO) which blocks lymphocyte migration to the intestine by blocking a lymphocytic integrin[182]. Although biologics have improved the care of Crohn's patients, further advancement is still essential. In a recent analysis, the net remission rates after 56 weeks of biologics ranged only from 6.7% (vedolizumab) to 34.8% (ustekinumab)[183]. There is an obvious need for more effective therapeutics to manage this chronic disease and to improve patients quality of life.

*5.4 Potential link between paratuberculosis and Crohn's disease*

The following section is a book chapter published in 2020 which discusses the hypothesis that *MAP* is etiologically linked to Crohn's disease. This chapter examines the data in support or against this potential relationship published since the previous edition of this book in 2010.

Chapter 3 – Paratuberculosis and Crohn's disease

Paratuberculosis: Organism, Disease, Control (2nd ed.), pp 29-44, CAB International, Nov 2020

Authors: Shannon C. Duffy and Marcel A. Behr

*5.4.1 Introduction*

The potential aetiological relationship between *Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) and Crohn's disease was first proposed over 100 years ago[184]. This hypothesis was stimulated by clinical and pathological similarities observed between paratuberculosis and Crohn's disease. Clinically, both conditions are characterized by intermittent diarrhoea and weight loss; pathologically, the conditions often present with disease within the ileocecal area, and the formation of ulcerations and intramural granulomas[185]. Yet, several differences between the two diseases have also been noted[186].

The cause of Crohn's disease remains unknown, but it is generally considered to involve a genetically determined immune dysregulation to an intestinal microbiological exposure[177–179]. Of note, while a number of *MAP*-centred reviews discuss the possibility of a *MAP*–Crohn's link[187–189], with varying degrees of conviction, most Crohn's aetiology reviews do not mention *MAP*, as this is not presently a leading hypothesis among inflammatory bowel disease (IBD) researchers. The reasons for this general lack of interest in the *MAP* hypothesis are variable but appear to include intellectual fatigue (the hypothesis has not advanced in a century), uncertainty

(inconsistent results from different studies) and the limited effect of anti-mycobacterial drugs (16% absolute efficacy of anti-*MAP* treatment in a clinical trial published by Selby *et al*., 2007)[190]. Each of these will be addressed below, after a review of the data that have accumulated since the previous edition of this book. Our objective is to provide an up-to-date summary of the latest research, which provides evidence for and against the potential relationship between *MAP* and Crohn's disease (Tables 1 and 2). In doing so, we intend to identify what remains unresolved and what further studies could ultimately answer this century-old question, one way or the other.

*5.4.2 Clinical and Pathological Comparison of Paratuberculosis and Crohn's Disease*

Crohn's disease is a chronic systemic inflammatory condition presenting primarily with gastrointestinal pathology. It is a debilitating disease, which typically leads to weight loss, abdominal pain, diarrhoea and obstruction, among others[191]. The diseased tissue in both Crohn's and paratuberculosis is characterized by transmural ulceration, non-caseating granulomas, infiltration of lymphocytes and macrophages, mucosal cobblestoning and creeping mesenteric fat[192]. More recent comparative pathology studies have emphasized that both disease tissues display relatively the same frequency of granulomatous enteritis, lymphoplasmacytic enteritis, oedema and lymphangiesctasia[193].

However, differences have been observed between these two conditions. A comparative pathological study between 29 Crohn's disease, 17 ulcerative colitis and 35 bovine paratuberculosis samples described Crohn's disease lesions as more destructive than paratuberculosis, as defined by the degree of damage to the epithelium, ulcer and fissure formation, among others[194]. Conversely, a recent study found that vasculitis was observed in nearly all paratuberculosis tissues and only a small percentage of Crohn's disease tissues. The

most prominent pathological difference between classical paratuberculosis and Crohn's disease is the ability to readily visualize acid-fast bacteria in the diseased tissue from cattle[193]. The detection of *MAP* in tissues of humans with Crohn's disease has proven inconsistent or unsuccessful, a fact that will be discussed in more detail in a future section.

Together, these observations demonstrate that the rationale implicating *MAP* in Crohn's disease is backed by the many similarities with paratuberculosis. However, it is noted that granulomatous pathology is not pathognomonic for a single aetiology; other granulomatous conditions range from a wood splinter to berylliosis and include bacterial infections (e.g. leprosy) and fungal infections (e.g. histoplasmosis). Furthermore, while differences in pathological presentation between paratuberculosis and Crohn's disease might argue against a common aetiological agent, Crohn's disease is generally accepted to develop due to a combination of genetic, immune and environmental factors. Given the differences in genetics, immunology and environmental exposures between humans and cattle, the final clinical and pathological outcomes of a similar microbial exposure need not be identical.

### *5.4.3 Epidemiological Evidence Investigating Transmissible Infection*

An extensive amount of recent research has supported the presence of *MAP* in several potential routes of human exposure. Both viable *MAP* and *MAP* DNA have been detected in milk[195,196] including in commercially available milk[197,198] revealing the ability of this pathogen to withstand pasteurization. *MAP* has also been detected in meat[199], cheese[200,201], infant formula[202,203], water sources[204–206] and aerosols[207]. From these data, it can be agreed upon that there are numerous opportunities for humans to be exposed to *MAP*. Yet the ability to detect *MAP* in the environment does not provide direct evidence of transmission of bacteria into

humans, nor does it demonstrate that the burden in these sources exceeds the unknown infectious dose for humans.

A 2009 report described three individuals who developed Crohn's disease that lived in close proximity to one another but otherwise had no recorded contact. The author hypothesized that exposure may have occurred due to a shared water system[208]. Another study published by the same author described seven unrelated children who developed Crohn's disease after moving to Forest, Virginia, where the number of cases of IBD was 47 times more than expected. Of these seven individuals, five (71.4%) were positive for *MAP*-specific antibodies (anti-p35 and anti-p36)[209]. *MAP* infection was suggested to have occurred due to contaminated water from nearby dairy farms. Although *MAP* has been reported in drinking water, neither of these studies investigated whether *MAP* was present in the water system of the patients' homes. In contrast to these results, a questionnaire-based assessment of the risk of IBD development and milk consumption in over 400,000 participants found that those consuming more milk were at a significantly reduced risk for development of Crohn's disease[210]. Again however, the presence of *MAP* was not evaluated in this study. One study, which did measure *MAP* presence in both patients and the potential source of exposure, was conducted in a region of Italy where *MAP* is endemic in cattle. The authors found *MAP* DNA in 23/35 (65.7%) biopsies from patients with Crohn's disease and in the circulating tap water[211]. It is important to note that *MAP* was detected in 7/35 (20%) controls in this study. Although the detection of *MAP* was significantly greater in patients with Crohn's disease, these findings reinforce that exposure is not sufficient to cause disease.

An argument raised in regard to the epidemiology of the disease is that there is no evidence of an increased risk to developing Crohn's disease associated with persons in contact

with *MAP*-infected animals such as farmers or veterinarians[212]. Since the previous edition, there have been few published studies that have addressed this argument. A 2010 cross-sectional survey of 1476 cattle farmers and veterinarians found no relationship between Crohn's disease and *MAP* exposure[213], although there were only seven cases of Crohn's disease recorded in total. In contrast, a study conducted on 58 animal attendants with reported gastrointestinal problems cultured more *MAP* compared with controls[214]. These data suggest a role of additional confounding factors such as location or the type of livestock.

Additional research is required in order to gain a clearer picture of the epidemiology surrounding the link between *MAP* and Crohn's disease. But there are several circumstances concerning these types of studies that should be considered. First, research aimed at associating exposure to *MAP* (today) with presentation of disease (later) is impeded by the unknown incubation period; in cattle bacteria typically persist undetected in a subclinical state for 2-5 years before disease manifests[167]. Therefore, studies starting with ascertained *MAP* exposure may fail to link cases, which present years later; conversely, studies starting with Crohn's disease cases are unable to do microbial analysis of the environment to which one was exposed in the preceding years. Second, the idea that contact with infected animals must lead to an increased risk in development of the disease is not found with other bacteria whose aetiological role is accepted. For example, cattle are the known reservoirs for *E. coli* O157:H7, however no increased risk was associated with disease development in those in direct contact with infected cattle[215]. Third, Crohn's disease development is thought to involve a combination of factors including genetic susceptibility and immune dysregulation. If *MAP* only plays a pathogenic role in such susceptible hosts, the vast majority of those exposed would not develop disease. Future

research aimed at rigorously evaluating this argument must be long-term and evaluate both bacteriological exposure and host predisposition.

### *5.4.4 Detection of MAP in Crohn's Disease Patients*

Earlier systematic reviews have concluded that *MAP* detection is significantly enriched in Crohn's disease patients[216–218]. An updated systematic review conducted by Waddell *et al*., 2015 also found a significant association across several methods, but they noted heterogeneity between studies detecting *MAP* by PCR vs hybridization[219]. It is important to note that the primary data compiled in such systematic reviews may be subject to publication bias. As it is more difficult to publish a negative result, investigators in the field need to be reminded that negative data remain critical in the future to updating the collective literature on *MAP* and Crohn's disease.

Isolation in pure culture from diseased tissue is classically the gold standard for demonstrating the presence of a bacterium and serves as the first step in determining if the bacterium is associated with a particular disease (if found to be more common in diseased than in control subjects). However, this method comes with a set of challenges, some inherent to *MAP* microbiology and some inherent to research involving human subjects. *MAP* is among the slowest-growing culturable mycobacteria, and proper decontamination is difficult to achieve in tissue samples containing many other competing bacteria. Yet, other chapters of this book provide details on how to isolate this same slow-growing bacterium from manure, demonstrating technical feasibility. Complicating factors in human studies include the nature of the sample and the status of the patient. In Crohn's patients who are clinically stable, one can only obtain mucosal biopsies or stool samples. If the organisms reside beneath the mucosa (e.g. the

muscularis, the mesenteric lymph nodes), then the chance of detecting them is greatly reduced. In most Crohn's patients who have progressed to the point of requiring a surgical excision, there has been a trial of pre-operative antibiotics to avoid surgery (e.g. metronidazole and ciprofloxacin) and there are usually peri-operative antibiotics given to reduce the risk of a post-operative infection[220]. The effect of these *in vivo* antibiotic treatments on the ability to isolate *MAP* in culture from the sample is unknown.

Overall, recent reports have detected *MAP* by culture in a small percentage of Crohn's disease patients. One study cultured *MAP* from 2/18 (11.1%) Crohn's patients, although none could be cultured from ulcerative colitis or non-IBD controls[221]. Banche *et al*. (2015) cultured *MAP* from 13/76 (17.1%) samples from patients with Crohn's disease compared with 2/44 (4.5%) controls[222]. Another study did not culture *MAP* from Crohn's disease patients, however only five patients were included in the study[223]. In a larger study including 75 individuals recently diagnosed with Crohn's disease, two (2.7%) were *MAP* positive by culture compared with 2/135 (1.5%) controls[224].

Due to the difficulties with *MAP* culture, most studies combine culture with molecular methods such as IS*900* PCR to detect *MAP*. However, the presence of bacterial DNA does not equate with actively growing bacteria, as exemplified by the use of ancient DNA in anthropology. The majority of molecular method-based studies have reported a higher frequency of *MAP* in Crohn's disease patients, but they are also subject to considerable heterogeneity. Indeed, many recent studies have reported a significant association with Crohn's disease[221,222,225,226]; however, PCR detection rates continued to range from 0–100% across studies[227–229]. A 2015 systematic review indicated that a significant amount of heterogeneity could be explained by choice of primers and the sampling frame[219]. Even within the rubric of

PCR-based detection of IS*900*, there are technical variants, such as the nested IS*900* PCR assay. Several recent studies have applied this method, nearly all of which reported a positive association[193,230–234]. However, one study did not detect *MAP* in 75 recently diagnosed Crohn's disease patients using nested IS*900* PCR[224].

Another method used to detect *MAP* in tissue is staining. Staining for acid-fast bacilli and looking under oil-immersion microscopy was described by Jeyanathan *et al*. (2007), who applied methods previously validated on tissue from mice experimentally infected with *MAP* (positive control) and *M. tuberculosis* (negative control)[235,236]. Banche *et al*. (2015) reported non-significant differences in detection where 3/76 (3.95%) Crohn's disease samples were positive by acid-fast staining compared with 0/44 (0%) non-IBD controls[222]. Zarei-Kordshouli *et al*. (2019) detected acid-fast bacteria in 2/30 (7%) Crohn's disease subjects compared with 0/30 (0%) non-IBD controls and 27/30 (90%) caprine paratuberculosis samples[193]. Studies using immunohistochemistry also failed to detect differences between samples from Crohn's disease and controls[226,237]. However, *MAP* infection in humans is often observed to be paucibacillary, which can be more difficult to visualize using acid-fast staining. In addition, some studies have suggested that this lack of visualization is due to *MAP* entering a cell wall-deficient spheroplast form in human tissue, which will not be detected by acid-fast staining[238]. Given that cell wall-deficient forms of *MAP* have not to our knowledge been described in the natural bovine and ovine hosts, this would be a unique situation in clinical microbiology where a zoonotic agent undergoes a morphologic transition during the human spillover infection. An alternative possibility is that *MAP* does undergo a transformation into a cell wall-deficient form in its natural hosts, but this has not been reported to date.

There is some evidence to suggest intra-study variability in results due to changes in

*MAP* status in individuals. In 2009, Kirkwood *et al*. reported significant *MAP* association in biopsies and peripheral blood mononuclear cells (PBMCs) of children with Crohn's disease by IS*900* PCR and culture compared with non-IBD controls[239]. The children participating in this study were then followed up for up to 4 years. *MAP* was found to persist only within a subset of patients and the authors suggested that infection status at disease onset may be an important factor in predicting infection persistence[240]. Overall, the recent data suggest that, although a positive association between *MAP* and Crohn's disease has been consistently reported by systematic reviews, heterogeneity between and within studies continues to be reported. Future studies applying multiple externally validated and standardized methods may allow for more robust cross-study comparison.

### *5.4.5 Comparative Genotyping of Human and Bovine-Derived MAP Isolates*

An argument previously raised against the potential link between *MAP* and Crohn's disease was the perception that the genotypes of human-derived *MAP* isolates are significantly different from those of bovine origin[186]. This argument originated from research using multiplex PCR of IS*900* integration loci (MPIL) fingerprinting, which found that bovine isolates clustered together, whereas human and ovine isolates were more diverse[241]. Many research studies followed reporting similarity between *MAP* cattle and human isolates, using several different comparative genomics methods: DNA microarrays, IS*1311* PCR restriction enzyme analysis, short-sequence repeat genotyping and, most recently, whole genome sequencing[242–246]. The latter, most definitive method, has repeatedly shown that human *MAP* isolates are similar to bovine isolates, in different countries. Bannantine *et al*. (2014) reported the whole genome sequence of *MAP* isolated from the breast milk of a person with Crohn's disease and determined

the isolate was highly similar to those of bovine origin[245]. Another study that sequenced the genome of a human *MAP* isolate 43,545 also found that it was most closely related to bovine type[244,246]. Wynne *et al*. (2011) performed comparative genomics on seven human *MAP* isolates (four from Crohn's disease patients), two bovine isolates and one ovine isolate[244]. Sequencing again revealed that all human-derived isolates were similar to one another and closely related to a bovine isolate.

When considering the presented evidence, it can be concluded that most human *MAP* isolates studied have the genotype of bovine *MAP*. While some variation has been described between human isolates[247], it is noted that single nucleotide polymorphism (SNP) genotyping of *MAP* isolates from 26 dairy herds has revealed heterogeneity, both within and between herds[248]. Unless *MAP* is proposed to have 'jumped' once into humans, and that clone subsequently spread uniquely in humans, there is little reason to anticipate that there should be a human-specific genotype. Additionally, if a *MAP* lineage associated with sheep was isolated from a human sample, this would not invalidate the zoonotic potential of *MAP*; rather it would point to a different potential source.

### *5.4.6 MAP Immune Responses in Crohn's Disease Patients*

Immune responses that might be detectable in Crohn's disease include humoral and cell-mediated responses. In serological studies, Verdier *et al*. (2013) found significantly increased IgG against four *MAP* antigens (glycosyl-transferase-d, Johnin-purified protein derivative (PPD), heparin-binding haemagglutinin, L5P) in gut lavage fluids[249]. Another study also found a positive association in antibody titres against two *MAP* antigens (MAP3865C133-141 and MAP3865125-133)[221]. Lefrançois *et al*. (2011) described a *MAP* laminin-binding/

histone-like protein (Lbp/Hlp) that is bound by serum antibodies of patients with Crohn's disease[250]. In addition, Xia *et al*. (2014) found that anti-PtpA antibodies were significantly increased in Crohn's disease patients compared with healthy controls[251]. However, many of these antigens are also found in other mycobacterial species. Therefore, seropositivity may not indicate reactivity to *MAP* but rather cross-reactivity due to previous exposure with another mycobacteria. These studies are also limited by the fact that detection of *MAP*-reactive antibodies indicates prior exposure to the bacteria but does not specify ongoing infection. In animals with paratuberculosis, detection of *MAP* antibodies is often a lagging indicator, where culture positivity precedes seropositivity[252]. When considered with the above culture data, it does raise the possibility of some Crohn's disease occurring as a post-infectious pathology, given that the prevalence of antibodies to *MAP* is generally much greater than the prevalence of a live *MAP* culture. To our knowledge, only one study has looked at serology and culture in the same matched subjects, so this hypothesis is difficult to explore from the published literature. Zamani *et al*. (2017) found that 13/28 (46%) and 11/28 (39%) of Crohn's disease patients were seropositive for MAP3865c$_{133\text{-}141}$ and MAP3865c$_{125\text{-}133}$ peptides, respectively. *MAP* was cultured from two of these patients[221].

Crohn's disease lesions typically have an abundance of CD4+ T cells that will produce inflammatory cytokines including IL-17 and IFN-γ, which are characteristic of a Th1/Th17 immune response. In contrast to the expanding literature on a humoral immune response to *MAP* in Crohn's disease, there is a relative paucity of evidence for *MAP*-reactive T cells present in patients with Crohn's disease. Sibartie *et al*. (2010) found that PBMCs co-incubated with *MAP* produced significantly more T cells in patients with Crohn's disease compared with ulcerative colitis or controls[253]. Similarly, T-cell lines were shown to proliferate more in response to *MAP*

when isolated from Crohn's disease patients compared with ulcerative colitis. These T cells produced both IL-17 and IFN-γ, supporting the presence of a Th1/Th17 immune response[254]. In intestinal biopsies of Crohn's disease patients, the frequency of T cells reactive to mycobacteria was found to be significantly higher than those reactive against *Escherichia coli*. These T-cell lines were also shown to produce both IL-17 and IFN-γ[255]. In contrast however, Nakase *et al*. (2011) compared the cytokine production of monocytes from Crohn's disease patients in response to *MAP* and *M. avium*[256]. Although significantly more tumour necrosis factor α (TNF-α) was produced by monocytes infected with *MAP,* both IL-12p40 and IL-6 were not produced, arguing against the induction of a Th1/Th17 response.

A key component of the rationale linking *MAP* to Crohn's disease concerns the gene mutations associated with an increased risk of disease. Many of these genes are involved in innate immune recognition of intracellular bacteria, suggesting that those who develop Crohn's disease are at an increased risk of certain types of infection. Loss of function mutations in *NOD2*, an intracellular pathogen recognition receptor, are significantly associated with Crohn's disease[257,258]. The NOD2 receptor will bind to bacterial muramyl dipeptide (MDP). Mycobacteria are known to produce a glycolylated form of MDP, which is a more potent and efficacious activator of NOD2[259,260]. Other genes include *IRGM1*, which encodes a GTPase involved in resistance to intracellular pathogens[261,262] including mycobacteria[263,264], and *ATG16L1*, which is involved in induction of autophagy and the formation of the autophagosome[265,266]. Taken together, one interpretation of the genomic data suggests that Crohn's disease occurs in individuals with a genetically determined susceptibility to infection with an intracellular bacterium that produces glycolyl-MDP, such as *MAP*. However, several studies aimed at determining the relationship between *MAP* infection status and Crohn's disease

susceptibility genes including *NOD2* have not found a positive association between them[229,267]. An alternative interpretation of the data is that the genetic predisposition to Crohn's disease overlaps not with the risk of developing a persistent infection, but rather with the risk of developing an inflammatory reaction to that intracellular pathogen. Research conducted by Fava *et al*. (2016) has shown that the predisposition to leprae reactions, which are inflammatory reactions during or after treatment of leprosy, have a remarkably similar genetic profile to Crohn's disease[268]. Supporting this paradigm, Wagner *et al*. (2013) found that susceptibility gene mutations in *TLR4*, which promotes proinflammatory responses in the gut[269], and *IL-10Ra*, the receptor for an immunoregulatory cytokine[270], were significantly associated with *MAP*-positive Crohn's disease patients[271].

Overall, recent research has provided more evidence of antibodies to *MAP* antigens in Crohn's disease along with more limited evidence of a potential induction of a Th1/Th17 immune response against *MAP* infection in Crohn's disease patients. More data are required concerning host immune responses to *MAP* in order to gain a clearer picture of how *MAP* could contribute to the development of Crohn's disease.

### *5.4.7 The Effect of Immunosuppressive Therapy on MAP Infection*

Anti-inflammatory agents such as anti-TNF-α and corticosteroids are common effective treatments for Crohn's disease. These treatments are also recognized to have immunosuppressive effects. In the case of tuberculosis, it is known that these treatments are associated with risk of disseminated disease[187]. Indeed, Crohn's disease patients treated with TNF-α inhibitors are at an increased risk of tuberculosis when compared with untreated individuals[272]. The lack of reported disseminated *MAP* infection following immunosuppressive therapy has served as a strong

argument against its potential link with Crohn's disease. However, there is evidence to suggest that *MAP* responses to immunosuppressive therapy may not follow the established paradigm of other mycobacteria. Research using calf infection models have suggested that prior depletion of CD4 cells has no effect on *MAP* disease progression[273]; this stands in stark contrast to the expectation from the human literature on AIDS-associated CD4 depletion and *M. avium* disease. An additional observation is that *M. leprae*, the cause of leprosy, has not been reported to cause disseminated disease following anti-TNF-α and steroid treatments[274]; indeed, anti-TNF-α has been used as an adjunctive therapy to manage hyperinflammatory phenotypes of leprosy, albeit while also receiving antimicrobial therapy. In this section we will review the recent research which has aimed to understand anti-TNF-α in *MAP* infection.

Several studies have suggested that Crohn's disease immunosuppressive therapies lead to a measurable decrease in *MAP* infection status. Bach *et al*. (2011) demonstrated that subjects with Crohn's disease had significantly higher titers of antibodies against the *MAP* proteins PtpA and PknG when compared with controls and that these antibodies were decreased in those who were treated with infliximab[275,276]. Similarly, Xia *et al*. (2014) found that anti-PtpA antibodies were significantly decreased in Crohn's disease patients treated with the immunosuppressant azathioprine, but not 5-aminosalycilic acid, steroids or their combination with azathioprine[251]. In a study involving experimental infection of macrophages with *MAP*, macrophages from IBD patients treated with infliximab retained significantly lower *MAP* colony-forming units when compared with those without treatment or healthy controls. The authors suggested that this finding could be a result of an increased induction of *MAP* dormant forms, which may not proliferate well despite immunosuppression[277], an explanation that warrants further investigation.

In order to understand the effect of TNF-α inhibitors, more research is required to delineate the role of TNF-α in host responses to *MAP*. Recent studies have suggested that *MAP* infection modulates TNF-α production to support its survival in the host, potentially explaining how anti-TNF-α treatments may negatively impact *MAP*. The gut culture supernatant TNF-α levels are increased in Crohn's disease patients positive for *MAP* when compared with *MAP*-negative patients[278]. Human macrophages have been shown to secrete increased levels of TNF-α when infected with *MAP*, but not when infected with other mycobacteria such as *M. avium*[256]. Additionally, two SNPs in the TNFα receptor, which were previously associated with undesired outcomes of anti-TNFα therapy, have been associated with increased susceptibility to infection with *MAP*[279]. In opposition to this argument however, Campos *et al*. (2011) found that macrophages from Crohn's disease patients had a defective TNF-α response to *MAP* compared with healthy controls, although this effect was not specific to bacterial challenge with *MAP*[280].

Taken together, recent data suggest that a lack of observed *MAP* dissemination following immunosuppressive treatments neither proves nor disproves a potential role of *MAP* in Crohn's disease. More research is required to elucidate the role of TNF-α in *MAP* pathogenesis and how this may further our understanding of patient responses to anti-TNF-α therapy.

### *5.4.8 The Use of Anti-Mycobacterial Therapy for Treatment of Crohn's Disease*

Another argument against the role of *MAP* in Crohn's disease is that anti-mycobacterial therapies are ineffective at treating patients with Crohn's disease. This argument primarily stems from the results of a double-blind randomized control trial conducted in 2007, where 213 Crohn's disease patients were treated with anti-mycobacterial drugs or placebo[190]. The authors reported an absolute benefit of 16% at week 16, but no evidence of a sustained benefit through

the remainder of the 2-year study. However, an intention-to-treat reanalysis of the data later revealed that the benefit observed at 16 weeks was maintained at weeks 52 and 104[281]. Notwithstanding the duration of the observed benefit, a major constraint of this trial was the lack of evaluation of *MAP* status prior to inclusion in the study. The study evaluated whether patients got better, which is a clinically valid outcome, but could not provide insights into a pathogenic role of *MAP*.

The varying interpretations of these data has left the potential efficacy of antimycobacterial therapy unresolved. A systematic review conducted by Feller *et al*. (2010) evaluated the outcomes of Crohn's disease remission and relapse for patients treated with antibiotics vs placebo[282]. This study found a significant benefit in studies treating patients with clofazimine, an antibiotic known to be active against mycobacteria, but no benefit was observed in studies involving classic anti-tuberculosis drugs. A second systematic review conducted in 2011 evaluated the effect of antibiotic therapy in IBD treatment using only randomized control trials. The authors found a statistically significant effect of antibiotics over placebo in treating both active and quiescent Crohn's disease[283]. It is important to note however, that these systematic reviews included a variety of antibiotics, including those active and those inactive against mycobacteria.

In 2013, RedHill Biopharma began a phase III double-blind placebo-controlled randomized control trial to assess the effect of RHB-104 vs placebo for treatment of moderate to severe active Crohn's disease. An important difference from the 2007 study is that the changes in *MAP* blood status would be evaluated by PCR, although *MAP* positivity was not a requirement for inclusion in the trial (clinicaltrials. gov NCT01951326). RHB-104 is a formulation of rifabutin, clofazimine, and clarithromycin shown to be effective against *MAP* strains isolated

from Crohn's disease patients[284,285]. RedHill Biopharma has announced that the treatment group was superior to placebo in achieving remission at week 26 (37% vs 23%) with statistically greater responses. They also reported those receiving RHB-104 experienced a significant benefit in attaining an earlier remission and in maintaining remission to week 52[286]. Additionally, the company stated that RHB-104 was shown to have a greater calculated maintenance of remission over other standard-of-care therapies for Crohn's disease such as Infliximab[287]. At the time of writing, the RedHill study remains unpublished.

Overall, these results indicate a statistically significant, although clinically modest benefit of antimycobacterial therapy in the treatment of Crohn's disease. More data are required to determine whether this clinical benefit supports a direct involvement of *MAP* in development of Crohn's disease.

### *5.4.9 Conclusions and Directions Forward*

Although there are many recognizable similarities between paratuberculosis and Crohn's disease, several inconsistencies exist. There have been ~200 papers published on the potential aetiological role of *MAP* in the development of Crohn's disease since the first edition of this chapter. Yet, despite this new information, the hypothesized relationship has yet to be proven or invalidated. At the time of the previous edition, there were several outstanding arguments that argued against a link. Although one argument has been largely refuted by recent data – the differences in the genotypes of human and bovine *MAP* isolates – the remaining arguments have yet to be resolved. The intent of this chapter is to emphasize these pending areas of uncertainty and highlight what types of future research studies would be most beneficial to ultimately resolve this debate.

Epidemiological research is limited by the many challenges involved in associating *MAP* exposure to disease outcomes, including its slow growth rate and the multifactorial nature of Crohn's disease. Although clusters of Crohn's disease have been described in the population, how these relate to *MAP* is uncertain. Future studies would benefit from implementing bacterial genotyping in order to establish a transmissible infection from the suggested source of exposure. Additionally, the majority of research investigating a relationship between occupational exposure and disease development have not found an association with direct contact with livestock. However, the notion that occupational risk should be considered a defining point for establishing an aetiological relationship is not consistently found in other types of intestinal infection. Therefore, whether this argument should play a significant role in evaluating this theory going forward should be considered. Studies seeking to detect *MAP* in humans have demonstrated a positive association between *MAP* and Crohn's disease as described by multiple systematic reviews. However, the interpretation of the results across multiple studies is largely impeded by small sample sizes and the use of different, non-standardized detection methods. To rigorously evaluate an association between *MAP* and Crohn's disease, future studies must follow a recognized standard method in order to ease cross-comparison analysis. In addition, the discrepancy observed between acid-fast staining detection of *MAP* in paratuberculosis and Crohn's disease could theoretically be explained by the formation of a cell wall-deficient form in humans, however what role (if any) this form plays in paratuberculosis must be investigated if the theory of a spheroplast infection in humans is to be supported.

Key knowledge gaps exist in our understanding of the immunology of *MAP* infection in humans, which limits our ability to determine its involvement in the development of Crohn's disease. Recent research has broadly established the potential for a *MAP*-induced Th1/Th17

immune response in humans. Future studies should aim to delineate the role of TNF-α as a host defence mechanism or as a response initiated by *MAP* to modulate the host immunity to promote its own infection. This may provide evidence to explain the counterintuitive host responses to Crohn's disease immunosuppressive therapies if *MAP* is implicated in disease aetiology. Additionally, many of the genes positively associated with disease are involved in innate immune responses to intracellular bacterial infection such as with *MAP*. However, no relationship with *MAP*-positive status has been established with a number of susceptibility genes. As research in this field has been primarily focused on searching for *MAP*, we may be overlooking other potential candidates, which involve similar host immune responses.

In addition to the need for further bench-level research, incoming data from the recently concluded RHB-104 phase III trial will be a key piece of information in determining how the elimination of *MAP* infection affects outcomes of Crohn's disease. Both this study and a previous report have shown a 15–20% absolute benefit of antibiotic therapy in Crohn's disease treatment. However, the direct measurement of *MAP* and its relationship with disease severity has not yet been evaluated. If reversal of *MAP* infection status is shown to have a significant impact on the benefit observed from RHB-104, this would provide strong support of a pathogenic role of *MAP* in development of Crohn's disease.

[End of book chapter.]

## 6. Animal models of bovine mycobacteria

To study how the bovine mycobacteria *M. bovis*, *M. orygis,* and *MAP* cause disease and their potential for zoonoses, it is important to have a representative small animal model in the laboratory. These models can provide fundamental value in understanding the role of

mycobacterial genes in disease pathogenesis and translational value in testing potential vaccines and drugs. A murine model of infection has limited space requirements, is accessible, inexpensive, and uses an extremely well-characterized host for which there are also readily available genetically defective variants. At the time of writing, there are no published studies of infection models of these bacteria in mice which model the oral route of exposure, and which mimic natural disease progression. Chapter IV of thesis explores avenues of accomplishing this.

*6.1 Existing models of Mycobacterium bovis and Mycobacterium orygis*

*M. bovis* experimental infections have been conducted in both bovines and mice as well as several other species such as badgers, rabbits, deer, and pigeons[288–291]. Bovine models have been performed by experimental infection where an inoculum of *M. bovis* is administered directly to the animal. Experimental infections are useful for studying bacterial pathogenesis because the inoculum is defined and comparable between animals. This has been done by a variety of infection routes including the intranasal, intratracheal, aerosol, and oral route[292–295]. In a comparison of intratracheally and orally infected calves, the route of infection was shown to direct the location of lesion development: intratracheally infected calves developed lesions primarily in the thorax and lungs whereas orally infected calves developed lesions primarily in the abdomen[295]. Studies of *M. bovis* in cattle have also been performed by natural infection where uninfected calves were housed in the same area as one that was experimentally infected[296]. Natural infection models can be useful for understanding disease dynamics and spread as well as understanding the range of infection outcomes which can occur. This type of experiment was used to show how *M. bovis* elimination in cattle herds is hindered by transmission of *M. bovis* from wildlife reservoir hosts such as badgers[86]. Although bovine infections have the obvious

benefits of being a natural host for *M. bovis*, these experiments are expensive, labor-intensive, and require large, specialized facilities and equipment.

Mouse experimental infections offer a comparably less expensive option to investigate *M. bovis* pathogenesis. This makes them more amenable to larger experiments involving more animals and experimental conditions. In the literature, there are studies assessing infection outcomes in mice that have been infected intranasally, intravenously, orally, and by aerosol[297–300]. Intranasal infection of C3HeB/FeJ mice with *M. bovis* led to histopathological features shared by experimentally infected cattle including formation of classic tuberculous lung granulomas with a caseous necrotic core surrounded by neutrophils, foamy macrophages, and multinucleated giant cells. However, approximately half of the mice had a severe response to infection and died 4-5 weeks post-aerosol following rapid weight loss and development of large necrotic lung lesions with abundant acid-fast bacilli[297]. Additionally, oral infection of house mice with *M. bovis* caused 33% death in under 5 weeks with tuberculous lesions primarily restricted to the lungs. The surviving mice regained lost weight and stabilized until the 60 day experimental endpoint[300].

The published research concerning *M. orygis* has largely focused on its epidemiology and isolation in humans. There is currently no research in the literature which has performed experimental infections with *M. orygis* to study its pathogenesis. In the Behr lab, there are several experiments currently ongoing characterizing infection outcomes of aerosol infection of *M. orygis* in mice. PhD candidate Sarah Danchuk has uncovered that aerosol infection of *M. orygis* leads to severe lung disease and ~50% mortality 4 weeks post-infection in C57BL/6 mice (unpublished data). This is in stark contrast to aerosol infection with *M. tuberculosis* where the median survival time is ~40 weeks[301].

The current mouse models of infection with *M. bovis* and *M. orygis* do not mimic what has been found in cases of zoonotic TB in humans. A large proportion of zoonotic TB is extrapulmonary sometimes with concomitate lung involvement and the most common route of exposure for zoonotic TB is thought to be through ingestion of unpasteurized milk or contaminated milk products. To create a more representative mouse model of zoonotic infection with *M. bovis* or *M. orygis*, bacteria would ideally be administered via the oral route and would lead to the development of a slower-progressing tuberculosis with potentially extrapulmonary organ involvement.

*6.2 Existing models of Mycobacterium avium subspecies paratuberculosis*

Like *M. bovis*, experimental infections of *MAP* have been performed in bovines, mice and other animals including other ruminants (goats, sheep, deer) and rabbits. In bovine infections, there are both long-term and short-term models. Long-term models are used to study disease outcomes, test vaccines and diagnostics, and evaluate therapeutics. Published long-term models have administered *MAP* via the oral, intragastric, or parenteral routes, however the oral route is recommended since this models the natural route of infection. Two consecutive doses of $5x10^8$ CFU over 2 days to young calves has been shown to lead to detectable intestinal colonization 4-12 weeks post-infection. The short-term models take place over a few hours to a couple weeks and are useful for studying host-pathogen interactions. One model involves cannulation of the ilea to directly deliver *MAP* to the site of infection. This is useful for measuring local immune responses to *MAP.* Another short-term model is the ligated intestinal loop model where the ilea and potentially the jejunum are excised and injected with *MAP.* This

experiment takes place in 12 hours or less and therefore has been most useful to study early host immune responses[302].

For the same reasons previously mentioned in regards to zoonotic TB pathogens, small animal experimental models have a number of advantages to study *MAP.* Although mice are the most accessible option, there have been a couple of studies using rabbits which had notable, although inconsistent success[302]. In 2005, Vaughan and colleagues infected pregnant white rabbits with 3 oral doses of $5x10^8$ CFU *MAP*. Once the rabbit juveniles were born, a subset of these was also infected with 3 oral doses of $10^8$ CFU *MAP.* Clinical signs and histopathological lesions characteristic of paratuberculosis were observed in ¼ adult rabbits and 3/16 juveniles between 3-24 months of age[303]. However, other studies using a rabbit model have reported *MAP* colonization but no clinical signs of disease development[304,305].

There are a number of published studies using various mouse strains and infection strategies to try to model *MAP* infection. However, a representative mouse model which uses the natural route of infection (oral) in an immunocompetent mouse to induce intestinal disease has not yet been achieved. The model which is the most reproducible is infection of high doses of *MAP* through the intraperitoneal route[302]. This leads to consistent colonization of the spleen and liver with the formation of epithelioid granulomas[306]. This type of model has been useful to study the role of specific genes on *MAP* viability through infection with transposon libraries or *MAP* knockout strains[307,308]. This model has also been attempted on immunodeficient mice including SCID mice where *MAP* injection led to mild infiltration of inflammatory cells into the intestinal lamina propria[309]. Unfortunately, infection with *MAP* via the oral route may lead to colonization but it does not lead to clinical symptoms or intestinal pathology. Cooney and colleagues infected BALB/c mice with 2 doses of $10^9$ CFU and reported chronic infection in the mesenteric lymph

nodes at 24-weeks post-infection with some minor evidence of inflammation in the liver[310]. Perhaps the model which was closest to modeling *MAP* disease was what was reported by Veazy and colleagues in 1995. They found that oral infection with $10^{11}$ CFU of *MAP* in 3-week old C57BL/6 mice resulted in granulomatous lesions 11 months post-infection, however these were restricted to the mesenteric lymph nodes[311]. Mouse intrinsic resistance to oral infection is not uncommon and has been found with other known enteropathogens such as enteropathogenic *E. coli* (EPEC) and enterohemorrhagic *E. coli* (EHEC)[312]. To create a representative *MAP* model in mice, new strategies will need to be employed.

*6.3 Streptomycin pre-treatment models*

*Salmonella enterica* serovar Typhimurium is a pathogen which causes enterocolitis in both bovines and humans. However, oral infection of this pathogen in mice leads to systemic disease rather than inflammation in the intestines. Barthel and colleagues found that when they pre-treated mice with streptomycin to reduce colonization resistance in the intestines and then followed that with an oral infection of a streptomycin-resistant strain of *S. enterica* serovar Typhimurium, this led to reproducible colonization ($10^6$ to $10^9$ CFU/g) and inflammation within the ceca of C57BL/6 mice. Colonization was also found in the ileum, mesenteric lymph nodes, spleen, and liver although the CFU levels were lower. When they repeated the experiment using a non-pathogenic *Lactobacillus* species, no obvious signs of inflammation were observed, indicating that the infection outcomes were specific to *S. enterica* serovar Typhimurium rather than the result of the antibiotic pre-treatment[313].

A streptomycin pre-treatment model of infection with *M. bovis, M. orygis,* or *MAP* is an intriguing option to model the disease they cause. If streptomycin can reduce competing bacteria

within the mouse, this may allow for an orally administered slow-growing mycobacteria to colonize the intestine more easily and may lead to representative disease development.

**7. Figure**

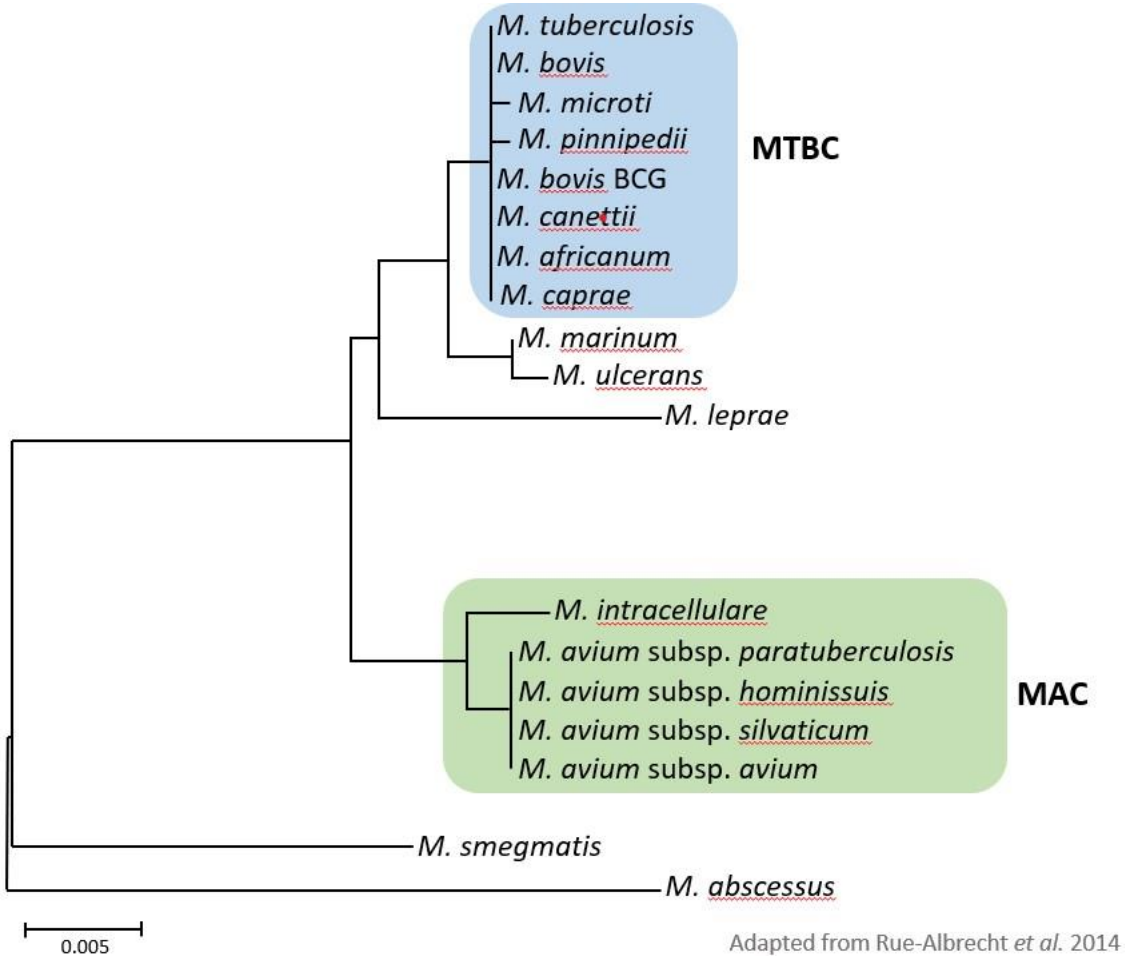



**Figure 1: Phylogenetic tree of the *Mycobacterium* genus**

This phylogenetic tree represents the evolutionary relationships of different species within the genus *Mycobacterium*. This includes members of the MTBC such as *M. tuberculosis* and *M. bovis,* leprotic bacteria including *M. leprae,* and NTMs such as members of the MAC. This tree was adapted from Rue-Albrecht and colleagues[314].

## 8. Tables

| Arguments | Responses from emerging data | References |
|---|---|---|
| 1. Clinical and pathological similarities of paratuberculosis and Crohn's disease | Confirmed | Zarei-Kordshouli *et al*., 2019 |
| 2. Presence in food chain (milk, meat) and water supplies | Confirmed | Beumer *et al*., 2010; Carvalho *et al*., 2012; Rhodes *et al*., 2014; Espeschit *et al*., 2018; Gerrard *et al*., 2018; Lorencova *et al*., 2019 |
| 3. Increased detection of *MAP* in Crohn's disease tissues by culture, PCR, FISH | Confirmed | Feller *et al*., 2007; Abubakar *et al*., 2008; Waddell *et al*., 2015 |
| 4. Positive blood cultures of *MAP* in Crohn's disease patients | Unresolved | Parrish *et al*., 2009; Mendoza *et al*., 2010 |
| 5. Increased serological responses to *MAP* in Crohn's disease patients | Confirmed | Verdier *et al*., 2013; Xia *et al*., 2014; Zamani *et al*., 2017 |
| 6. Therapeutic responses to combination anti-tuberculosis therapy that include macrolide antibiotics | Confirmed | Feller *et al*., 2010; Khan *et al*., 2011 |

MAP, *Mycobacterium avium* subsp. *Paratuberculosis*;PCR, polymerase chain reaction;FISH, fluorescence in situ hybridization;

**Table 1** Pros. Arguments in favour of the potential aetiological role of *Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) in Crohn's disease. Based on Sartor, 2005

| Arguments | Responses from emerging data | References |
|---|---|---|
| 1. Difference in clinical and pathological responses in paratuberculosis and Crohn's diseases | Confirmed | Momotani *et al*., 2012; Zarei-Kordshouli *et al*., 2019 |
| 2. Lack of epidemiological support of transmissible infection | Unresolved | Pierce, 2009; Pierce *et al*., 2011; Pistone *et al*., 2012; Opstelten *et al*., 2016 |
| 3. No evidence of transmission to humans in contact with animals infected with *MAP* | Unresolved | Qual *et al*., 2010; Singh *et al*., 2011a |
| 4. Genotypes of Crohn's disease and bovine *MAP* isolates not similar | Refuted | Paustian *et al*., 2008; Singh *et al*., 2009; Bannantine *et al*., 2014; Wynne *et al*., 2011; Timms *et al*., 2015 |
| 5. Variability in detection of *MAP* by PCR (0–100% in Crohn's disease and ulcerative colitis tissues) and serological testing | Confirmed | Waddell *et al*., 2015 |
| 6. No evidence of mycobacterial cell wall by histochemical staining | Unresolved | Jeyanathan *et al*., 2007; Banche *et al*., 2015; Zarei-Kordshouli *et al*., 2019 |
| 7. No worsening of Crohn's disease with immunosuppressive agents or HIV infection | Unresolved | Allen *et al*., 2011 |
| 8. No documented cell mediated immune responses to *MAP* in patients with Crohn's disease | Refuted | Sibartie *et al*., 2010; Olsen *et al*., 2009, 2013 |
| 9. No therapeutic response to traditional antimycobacterial antibiotics | Unresolved | Selby *et al*., 2007; Behr and Hanley, 2008; RedHill Biopharma, 2018a, b |

*MAP*, *Mycobacterium avium* subsp. *paratuberculosis*;PCR, polymerase chain reaction;

**Table 2** Cons. Arguments against the potential aetiological role of *Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) in Crohn's disease. Based on Sartor, 2005

# AIMS AND HYPOTHESES OF RESEARCH

In this thesis I will be discussing epidemiologic and pathogenetic studies of bovine mycobacteria. *M. bovis, M. orygis,* and *MAP* have a significant impact on animal health and contribute to substantial global economic losses each year. Their zoonotic potential makes them an additional public health concern. *M. bovis* and *M. orygis* have both been shown to cause bovine and zoonotic TB. However, the contribution of zoonotic TB to the global TB burden is still uncertain and the geographic distribution of these 2 pathogens is not entirely understood. The vaccine strain BCG is an attenuated strain of *M. bovis* which can also contribute to human disease in certain groups, however methods of rapid detection of BCG are not always available. Finally, *MAP* is the known cause of paratuberculosis in ruminants and the hypothesis that it may be involved in Crohn's disease has existed but made very little progress for several decades. In order to study the pathogenesis of these mycobacteria, a small animal model of disease would be a valuable tool.

My PhD thesis has 2 aims to address both the epidemiology and pathogenesis of bovine mycobacteria.

**Aim 1:** To understand the proportion of MTBC subspecies (*M. tuberculosis, M. bovis,* BCG, and *M. orygis*) causing disease in the context of South Asia.

**Aim 2:** To study the pathogenesis of bovine mycobacteria by developing representative mouse models of infection.

These aims are aligned with 2 hypotheses:

**Hypothesis 1:** I hypothesize that MTBC disease in South Asia is caused primarily by *M. tuberculosis* with a smaller proportion of cases caused by *M. bovis, M. orygis* and BCG.

**Hypothesis 2:** I hypothesize that an oral streptomycin pre-treatment model of infection with bovine mycobacteria would improve organ infection and model disease progression.

# PREFACE TO CHAPTER II

In chapter I of this thesis, I have stated that the WHO defines zoonotic tuberculosis as tuberculosis caused by *M. bovis.* I have also described the scientific evidence suggesting that *M. orygis* is endemic to South Asia and explained how cases of zoonotic tuberculosis are likely missed due to lack of surveillance and its tendency to present as extrapulmonary disease. In chapter II, I describe the findings from a collaboration with researchers from Christian Medical College Vellore, a large referral hospital in the south of India which regularly collects cultures from cases of extrapulmonary tuberculosis. I developed several methods of rapidly screening positive cultures using RD regions that can differentiate between MTBC members. After screening 940 cultures using these assays, I found zoonotic tuberculosis was infrequent and was caused by *M. orygis.* No cases were caused by *M. bovis,* but several were identified as BCG. We then expanded our search and analyzed all of the publicly available MTBC genomes published from South Asian countries (n=715). More cases of *M. orygis* were identified, but again there was no evidence of zoonotic TB caused by *M. bovis* in South Asia. This led us to conclude that *M. orygis* is likely the primary cause of zoonotic TB in South Asia and therefore the WHO definition of zoonotic TB should be expanded to include other MTBC subspecies.

# CHAPTER II

## Reconsidering *Mycobacterium bovis* as a proxy for zoonotic tuberculosis: a molecular epidemiological surveillance study

---

Published Manuscript- The Lancet Microbe, Volume 1, Issue 2, E66-E73, June 2020

*Shannon C Duffy, Sreenidhi Srinivasan, Megan A Schilling, Tod Stuber, Sarah N Danchuk, Joy S Michael, Manigandan Venkatesan, Nitish Bansal, Sushila Maan, Naresh Jindal, Deepika Chaudhary, Premanshu Dandapat, Robab Katani, Shubhada Chothe, Maroudam Veerasami, Suelee Robbe-Austerman, Nicholas Juleff, Vivek Kapur, Marcel A Behr*

**Department of Microbiology and Immunology** (S C Duffy BSc, S N Danchuk BSc, Prof M A Behr MD)**, McGill International Tuberculosis Centre** (S C Duffy, S N Danchuk, Prof M A Behr)**, and Department of Medicine** (Prof M A Behr)**, McGill University, Montreal, QC, Canada**

**Infectious Diseases and Immunity in Global Health Program, Research Institute of the McGill University Health Centre, Montreal, QC, Canada** (S C Duffy, S N Danchuk, Prof M A Behr)

**Department of Animal Science and the Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA** (S Srinivasan PhD, M A Schilling PhD, R Katani PhD, S Chothe PhD, Prof V Kapur PhD)

**National Veterinary Services Laboratories, Animal and Plant Health Inspection Service, US Department of Agriculture, Ames, IA, USA** (T Stuber BSc, S Robbe-Austerman PhD)

**Department of Clinical Microbiology, Christian Medical College Vellore, Vellore, India** (Prof J S Michael MD, M Venkatesan MSc)

**Department of Veterinary Public Health and Epidemiology** (N Bansal PhD, N Jindal PhD, D Chaudhary PhD)**, and Department of Animal Biotechnology** (Prof S Maan PhD)**, College of Veterinary Sciences, Lala Lajpat Rai University of Veterinary and Animal Sciences, Hisar, India**

**Eastern Regional Station, Indian Veterinary Research Institute, Indian Council of Agricultural Research, Kolkata, India** (P Dandapat PhD)

**Cisgen Biotech Discoveries, Chennai, India** (M Veerasami PhD)

**Bill & Melinda Gates Foundation, Seattle, WA, USA** (N Juleff PhD)

## 1. Summary

**Background** Zoonotic tuberculosis is defined as human infection with *Mycobacterium bovis*. Although globally, India has the largest number of human tuberculosis cases and the largest cattle population, in which bovine tuberculosis is endemic, the burden of zoonotic tuberculosis is unknown. The aim of this study was to obtain estimates of the human prevalence of animal-associated members of the *Mycobacterium tuberculosis* complex (MTBC) at a large referral hospital in India.

**Methods** We did a molecular epidemiological surveillance study of 940 positive mycobacteria growth indicator tube (MGIT) cultures, collected from patients visiting the outpatient department at Christian Medical College (Vellore, India) with suspected tuberculosis between Oct 1, 2018, and March 31, 2019. A PCR-based approach was applied to subspeciate cultures. Isolates identified as MTBC other than *M tuberculosis* or as inconclusive on PCR were subject to whole-

genome sequencing (WGS), and phylogenetically compared with publicly available MTBC sequences from south Asia. Sequences from WGS were deposited in the National Center for Biotechnology Information Sequence Read Archive, accession number SRP226525 (BioProject database number PRJNA575883).

**Findings** The 940 MGIT cultures were from 548 pulmonary and 392 extrapulmonary samples. A conclusive identification was obtained for all 940 isolates; wild-type *M bovis* was not identified. The isolates consisted of *M tuberculosis* (913 [97·1%] isolates), *Mycobacterium orygis* (seven [0·7%]), *M bovis* BCG (five [0·5%]), and nontuberculous mycobacteria (15 [1·6%]). Subspecies were assigned for 25 isolates by WGS, which were analysed against 715 MTBC sequences from south Asia. Among the 715 genomes, no *M bovis* was identified. Four isolates of cattle origin were dispersed among human sequences within *M tuberculosis* lineage 1, and the seven *M orygis* isolates from human MGIT cultures were dispersed among sequences from cattle.

**Interpretation** *M bovis* prevalence in humans is an inadequate proxy of zoonotic tuberculosis. The recovery of *M orygis* from humans highlights the need to use a broadened definition, including MTBC subspecies such as *M orygis*, to investigate zoonotic tuberculosis. The identification of *M tuberculosis* in cattle also reinforces the need for One Health investigations in countries with endemic bovine tuberculosis.

**Funding** Bill & Melinda Gates Foundation, Canadian Institutes for Health Research.

## 2. Introduction

Globally, tuberculosis is the most deadly disease arising from a single infectious agent, leading to 1·4 million deaths each year, primarily in low-income and middle income countries (LMICs).[1] Tuberculosis in cattle (bovine tuberculosis) is also a considerable animal health

problem and endemic in most LMICs, costing an estimated US$3 billion worldwide each year.[2] WHO, the World Organisation for Animal Health, and the Food and Agriculture Organization of the United Nations define zoonotic tuberculosis as human infection with *Mycobacterium bovis,* a member of the *Mycobacterium tuberculosis* complex (MTBC).[1,3,4] As such, detection of *M bovis* is often used as a proxy for measuring the prevalence of zoonotic tuberculosis. On the basis of this definition, WHO estimated that the annual number of human tuberculosis cases due to zoonotic tuberculosis was 143 000 in 2018.[1]

By 2035, WHO is aiming to reduce the incidence of tuberculosis by 90% as a part of its End TB Strategy.[5] India has the largest burden of human tuberculosis globally, with more than 2·6 million cases and 400 000 deaths reported in 2019.[1] The cattle population in India exceeds 300 million, a population larger than any other country.[6] Yet in India, bovine tuberculosis is both uncontrolled and endemic, with an estimated 21·8 million infected cows as of 2017.[7] Some previous studies have estimated the prevalence of zoonotic tuberculosis in India to be around 10%; however, these studies were limited in their ability to differentiate between MTBC members.[8–10] Other evidence in the past decade has suggested that *Mycobacterium orygis,* an MTBC member described in 2012, might be endemic to south Asia;[11–14] however, robust estimates of *M orygis* prevalence in humans and cattle are lacking.

In the present study, we sought to evaluate the definition of zoonotic tuberculosis, and investigate the prevalence of MTBC members in a large referral hospital in southern India, via molecular analyses of 940 positive broth cultures. Our findings are subsequently evaluated within the framework of other sequences collected in south Asia. We sequence 25 isolates by whole-genome sequencing (WGS) and compare these sequences with representative genomes of

the MTBC lineages circulating worldwide, and with 715 publicly available genomes collected from cattle and humans in south Asia.

## 3. Research in context

### *3.1 Evidence before this study*

WHO, the World Organisation for Animal Health, and the Food and Agriculture Organization of the United Nations currently define zoonotic tuberculosis as human infection with *Mycobacterium bovis* of animal origin. However, other members of the *Mycobacterium tuberculosis* complex (MTBC), such as *Mycobacterium orygis*, have been reported as the cause of tuberculosis in humans, especially in patients from India. Consequently, the burden of zoonotic tuberculosis might be underestimated by surveillance studies restricted to *M bovis*. In India, which has the largest number of human tuberculosis cases and the largest cattle population worldwide, and where bovine tuberculosis is endemic, the burden of zoonotic tuberculosis is unknown. On July 17, 2019, we searched PubMed using the terms "zoonosis" AND "tuberculosis" OR "zoonotic tuberculosis" AND "India" without date or language restrictions, which returned 45 results. Only three studies had screened for zoonotic tuberculosis in humans, none of which used methods capable of reliably differentiating between MTBC subspecies. These three studies were also small in scale (maximum of 331 samples), and none subspeciated positive cultures or attempted to identify zoonotic subspecies other than *M bovis.*

### *3.2 Added value of this study*

This study is the first to screen pulmonary and extrapulmonary culture isolates in India with tests to detect different MTBC subspecies capable of causing zoonotic tuberculosis. This

study represents an advance from previous surveillance reports in that screening was done directly on broth cultures with a two-step protocol, involving identification by PCR, and confirmatory testing by whole-genome sequencing. Our findings were further validated by comparisons in the context of other publicly available sequences collected from humans and cattle from south Asia. Our research showed that in a large referral hospital in southern India, *M bovis* infection is uncommon, and that zoonotic tuberculosis might instead be caused by *M orygis* and potentially *M tuberculosis*.

### *3.3 Implications of all the available evidence*

The findings indicate that *M bovis* is an inadequate proxy for the detection of zoonotic tuberculosis infection. In regions where other MTBC subspecies are the primary pathogen among the cattle population, *M bovis* prevalence might substantially underestimate the burden of zoonotic tuberculosis. The operational definition of zoonotic tuberculosis provided by WHO, the World Organisation for Animal Health, and the Food and Agriculture Organization for the United Nations should be broadened to include other MTBC members capable of causing human disease such as *M orygis*. This research highlights the importance of a One Health approach to tuberculosis control in India, and elsewhere, in which medical surveys should be informed by corresponding veterinary data.

## 4. Methods

### *4.1 Study design and isolates*

In this molecular epidemiological surveillance study, we used 940 positive mycobacteria growth indicator tube (MGIT) cultures previously collected at Christian Medical College in

Vellore, India between Oct 1, 2018, and March 31, 2019. This study took place from Jan 5 to June 9, 2019. Institutional review board approval (approval number 11725, dated Dec 19, 2018) and ethical clearance from Christian Medical College were obtained for this study. All patient data were denominalised before inclusion of each isolate in our study. An isolate was defined as a pure pathogen collected from an individual patient. A sample was defined as the patient specimen that was previously cultured to obtain an isolate.

All PCR screening was done at Christian Medical College. WGS was done at Lala Lajpat Rai University of Veterinary and Animal Sciences in Hisar, India. WGS analysis and construction of phylogenetic trees was done at The Pennsylvania State University, State College, PA, USA, in collaboration with the United States Department of Agriculture in Ames, IA, USA. The study design and procedure are outlined in appendix 1 (p 1).

The MGIT cultures screened in this study were obtained from patients who visited the outpatient department at Christian Medical College with signs and symptoms of suspected tuberculosis (from Oct 1, 2018, to March 31, 2019). Pulmonary isolates were defined as those collected from the lung fluid or tissue; extrapulmonary isolates were defined as those collected from tissue other than the lungs. These definitions served to differentiate disseminated disease from disease confined to the lungs and their adjacent structures.

*4.2 PCR assays*

Cultures were screened with PCR assays previously established and validated at McGill University in Montreal, QC, Canada. DNA was prepared for screening by boiling. We planned to screen the first 600 isolates by conventional PCR between Jan 5 and Jan 25, 2019, to assess the proportion of *M tuberculosis,* as opposed to *M bovis, M orygis*, or other MTBC subspecies. Two

deletion-based conventional PCR assays were developed for this screening. A three-primer PCR was designed to detect the presence or absence of region of difference (RD) 9, which is present in *M tuberculosis* and absent from other MTBC subspecies (appendix 1 p 2).[11] A six-primer PCR was developed to detect differences in the deletion size of RD12. The RD12 region is present in *M tuberculosis* and absent from *M bovis* and *M orygis*; however, the *M orygis* deletion is larger than in *M bovis* and the insertion sequence IS6110 is inserted (appendix 1 p 2).[11] All primers were designed with Primer3 software (version 0.4.0; appendix 1 p 3).[15] PCR reaction conditions are described in appendix 1 (p 4). PCR products were separated by gel electrophoresis on a 2% (w/v) agarose gel containing a 1:10 000 dilution of ethidium bromide.

A five-probe multiplex real-time (rt) PCR assay was developed to screen the remaining 340 isolates between May 11 and May 20, 2019. Forty isolates identified as *M tuberculosis* and all isolates identified as MTBC other than *M tuberculosis* by conventional PCR were also assessed by rtPCR to confirm consistent identity across assays. Four of the probes used have been described by Halse and colleagues,[16] which detected the presence of RD1, RD9, RD12, and a conserved region external to RD9. The RD1 probe allows for the differentiation of wild-type *M bovis* from *M bovis* BCG.[11] A fifth probe was designed to target the *M orygis*-specific single nucleotide polymorphism (SNP) in *Rv0444c* (G698C).[17] These probes allowed for differentiation of several members of the MTBC (appendix 1 p 5). The *Rv0444c* probe and primers were designed with Primer Express (version 3.0). All primer and probe sequences, and conditions of the rtPCR reactions are given in appendix 1 (pp 3–4).

*4.3 Sequencing assays*

Isolates identified as MTBC other than *M tuberculosis* or as inconclusive by PCR were processed for WGS. Genomic DNA was extracted as previously described.[18] Libraries were prepared with the Nextera DNA Flex Library Prep Kit (Illumina, San Diego, CA, USA). We assessed the quality of each library with a Fragment Analyzer Automated CE System (Agilent, Santa Clara, CA, USA) using a Next Generation Sequencing Fragment Kit (1–6000 bp; Agilent). Libraries were paired-end sequenced on an Illumina MiSeq System with a MiSeq Reagent Kit (version 3, 600-cycle). Sequences were deposited for download in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) accession number SRP226525 (BioProject database number PRJNA575883). For isolates not amplified by PCR, suggestive of a nontuberculous mycobacteria, *hsp65* sequencing was done by the Sanger method, as previously described.[19]

*4.4 Bioinformatics*

Genome sequences were assessed with the validate SNP (vSNP) tool of the US Department of Agriculture-Veterinary Services. The vSNP pipeline involved a two-step process. Step 1 determined SNP positions called against a best-matched reference (either *M tuberculosis* strain H37Rv or *M bovis* strain AF2122). Step 2 assessed SNPs called between related isolate groups, which shared a common SNP, to output SNP alignments, tables, and phylogenetic trees. For an SNP to be considered in a group there needed to be at least one locus with an allele count of 2, a quality score greater than 150, and map quality greater than 56, as per standard thresholds in vSNP. The output SNP alignment was used to assemble a maximum likelihood phylogenetic

tree with RAxML software (version 8.2; GTRCATI model).[20] Additional details of the pipeline are provided in appendix 1 (pp 6–7).

To situate sequences generated in this study within the context of global MTBC diversity, we used SPAdes (version 3.14.1) to assemble 373 genomes representing recognised lineages associated with humans and animals, derived from the NCBI SRA database, and constructed a phylogenetic tree with kSNP3.0 (appendix 2 p 1).[21] We followed the kSNP3.0 manual instructions using the kmer value calculated with the kchooser tool.

To compare the newly sequenced genomes with sequences from south Asia, we searched the NCBI SRA using the search terms ("*Mycobacterium bovis*" OR "*Mycobacterium tuberculosis*" OR "*Mycobacterium africanum*" OR "*Mycobacterium orygis*" OR "*Mycobacterium canetti*" OR "*Mycobacterium caprae*" OR "*Mycobacterium bovis* BCG" NOT "H37Rv" NOT "H37Ra") from ("India" OR "Bangladesh" OR "Nepal" OR "Sri Lanka" OR "Pakistan"). All sequences were downloaded from the SRA with the fasterq-dump tool from the SRA toolkit (version 2.9.6) and sequences were filtered for quality, according to the criteria in appendix 1 (p 8). Sequences of the required quality were run through vSNP (appendix 2 p 2). Phylogenetic trees were constructed with vSNP to compare the sequences from this study with the genomes from south Asia. SRA reference sequences of the MTBC lineages were also included for comparison (appendix 2 p 3). Phylogenetic trees were rooted to *M tuberculosis* strain H37Rv. As a post-hoc assessment, to determine whether human *M orygis* isolates were dispersed among cattle isolates, and whether cattle *M tuberculosis* isolates were dispersed among human isolates, sequences generated in this study were phylogenetically compared with *M orygis* sequences from south Asia and *M tuberculosis* sequences of cattle origin. We visualised all

phylogenetic trees using the interactive tree of life (version 4).[22] Further details on tree assembly are provided in appendix 1 (pp 6–7).

*4.5 Statistical analysis*

The minimum sample sizes were determined to be 300 pulmonary isolates and 300 extrapulmonary isolates to ensure that a zero numerator would indicate a 95% CI of 1·0% or less. Statistical analysis was done with GraphPad Prism (version 8.1.2). Prevalence ratios of extrapulmonary and pulmonary tuberculosis, as a function of age and mycobacterial species, were calculated in GraphPad and compared with Fisher's exact tests, with a p value of less than 0·05 considered to indicate statistical significance. Prevalence ratio 95% CIs were calculated with the Koopman asymptotic score in GraphPad.

*4.6 Role of the funding source*

The funders of the study had no role in data collection, data analysis, or data interpretation. NJu is the Senior Program Officer at the Bill & Melinda Gates Foundation and was involved in study design and writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## 5. Results

DNA was prepared from 940 positive MGIT cultures. In total, 548 pulmonary samples and 392 extrapulmonary samples were included in our study. Table 1 describes the characteristics of patients from whom isolates were obtained. Pulmonary isolates were from sputum (504 samples), pleural samples (30), lung biopsies (eight), endotracheal aspirate (three),

bronchoalveolar lavage (two), and chest wall abscess (one). The extrapulmonary sample types are listed in appendix 1 (p 9). Extrapulmonary tuberculosis was most prevalent in patients of younger age. Extrapulmonary isolates were obtained from 19 (73·1%) of 26 patients younger than 10 years, and 373 (40·8%) of 914 patients aged 10 years and older (prevalence ratio 1·8 [95% CI 1·3–2·2]; p=0·0018). Meanwhile, pulmonary tuberculosis was most prevalent in patients of older age. Pulmonary isolates were obtained from 88 (80·0%) of 110 patients aged 60 years and older, and 460 (55·4%) of 830 patients younger than 60 years (prevalence ratio 1·4 [1·3–1·6]; p<0·0001). Most patients were from 20 states and territories in India (884 [94%] of 940); all other patients were from Bangladesh (54 [6%]) and Nepal (2 [<1%]; table 1, figure 1A). Within India, most isolates were collected from patients living in the states of Tamil Nadu (341 [39%] of 884 Indian isolates), West Bengal (187 [21%]), and Andhra Pradesh (100 [11%]). Pulmonary isolates comprised the majority of isolates from each location, except for Bangladesh where 35 (64·8%) of 54 isolates were extrapulmonary (figure 1B).

Isolates were assessed by conventional and rtPCR to evaluate the proportions of MTBC subspecies. Following PCR screening, 25 isolates were selected for WGS (appendix 1 p 10). Based on PCR, these isolates were provisionally identified as *M orygis* (seven)*, M bovis* BCG (six), inconclusive (eight), and *M tuberculosis* isolates negative for RD12 (two). The two remaining isolates, identified as *M tuberculosis*, were assigned for WGS in error. Of the eight inconclusive isolates, seven were categorised as such because of delayed amplification of the RD12 probe, and one was selected following poor amplification of the RD1 probe.

The proportions of mycobacterial species identified in our study after all genotyping are described in table 2 and appendix 3. No wild-type *M bovis* was identified. Seven (0·7%) of the 940 isolates were identified as *M orygis*, six of which were extrapulmonary. *M orygis* was

enriched in extrapulmonary isolates compared with pulmonary isolates, with a prevalence ratio of 8·4 (95% CI 1·3–52·9; p=0·023). Six of the seven *M orygis* isolates were from patients from Bihar, Jharkhand, and West Bengal in northeast India (figure 1B, appendix 3). The remaining *M orygis* isolate was from a patient from Karnataka in south India. 913 isolates (97·1%) were identified as *M tuberculosis.* The two isolates that lacked RD12 on PCR were assigned as *M tuberculosis* and not *Mycobacterium canettii* by WGS. The seven isolates categorised as inconclusive because of delayed amplification of the RD12 probe contained an SNP in the RD12 reverse primer sequence, and were assigned as *M tuberculosis*. The other isolate categorised as inconclusive because of poor amplification of RD1 was found to have a 5962 bp deletion spanning Rv3871–Rv3877 including the site of the RD1 probe, and was assigned as *M tuberculosis*. Five isolates (0·5%) were confirmed as *M bovis* BCG by WGS. All of these isolates were identified in patients aged 3 years or younger, consistent with expectations for a population in which BCG vaccine is routinely administered. 15 isolates (1·6%) were identified as non-tuberculous mycobacteria. The most common non-tuberculous species were *Mycobacterium abscessus* and *Mycobacterium intracellulare* (three isolates each; appendix 1 p 11). The identification of one isolate was inconsistent across methods, suggestive of tube mislabelling: the isolate was identified as *M bovis* BCG by rtPCR and *M tuberculosis* by WGS, and was assigned as *M tuberculosis*.

A phylogenetic tree was constructed of the 25 isolates sequenced in this study in the context of the MTBC lineages circulating worldwide (figure 2). The 13 *M tuberculosis* isolates were dispersed across *M tuberculosis* lineage 1 (Mtb L1; nine isolates), Mtb L2 (one), Mtb L3 (one), and Mtb L4 (two). Of the nine Mtb L1 genomes collected in our study, seven were found to cluster together; these seven isolates shared an SNP in the RD12 reverse primer sequence. All

five confirmed *M bovis* BCG isolates identified as the Russian strain. No isolates clustered with wildtype *M bovis* or *M caprae*. The seven isolates that clustered with previously sequenced *M orygis* isolates were separated from each other by 66–282 SNPs (appendix 1 p 12).

A search of the NCBI SRA for MTBC isolates from south Asia generated 1640 genomes. These sequences were filtered before tree assembly, giving 715 high-quality sequences (appendix 2 p 2). The downloaded genomes were phylogenetically compared with our newly sequenced isolates and their respective metadata (figure 3A, appendix 2 p 4). No wild-type *M bovis* was identified in the 715 genomes. As expected, Mtb L1 and Mtb L3 were enriched. Almost all Mtb L1 genomes were from India (two sequences were from Bangladesh and three were from Pakistan); whereas, the Mtb L2, L3, and L4 sequences were from a variety of other countries in south Asia. Sequences of cattle origin from south Asia were represented in Mtb L1 (four isolates) and the *M orygis* lineage (20 isolates). On post-hoc assessment, the four isolates of cattle origin in Mtb L1 were dispersed among human sequences; similarly, the seven human *M orygis* isolates from this study were dispersed among sequences collected from cattle (figure 3B).

## 6. Discussion

Zoonotic tuberculosis is increasingly recognised as a potential threat to the control of tuberculosis.[23] *M bovis* was identified as a cause of bovine tuberculosis in 1898, nearly a century before descriptions of other distinct MTBC subspecies infecting cattle, at which point the scientific community had arrived at a general consensus that zoonotic risks associated with tuberculosis were caused by *M bovis.*[24] *M bovis* as the causal agent continues to be defined as zoonotic tuberculosis by WHO, the World Organisation for Animal Health, and the Food and

Agriculture Organization of the United Nations,[1,3,4] despite mounting evidence that human infections can also be caused by other members of the MTBC, such as *M orygis*.[11–14,25] To evaluate *M bovis* as a proxy for zoonotic tuberculosis and investigate the potential role of other MTBC subspecies, we did a molecular epidemiological screen of 940 clinical tuberculosis isolates from a large referral hospital in southern India. Our search was then broadened to include 715 MTBC sequences from south Asia deposited in the SRA database. Notably, no wild-type *M bovis* was identified in our study. Seven cases of tuberculosis were found to be caused by *M orygis,* and these were more likely to present as extrapulmonary tuberculosis (table 2). These findings suggest that *M bovis* might be an inadequate proxy for zoonotic tuberculosis infection in regions where *M bovis* is not the predominant MTBC member in livestock.

Detection of *M orygis* in this study is concordant with previous reports linking *M orygis* to patients with tuberculosis from south Asia.[11–14] Marcos and colleagues[12] identified eight *M orygis* isolates in individuals originating from India, Pakistan, or Nepal. All seven *M orygis* isolates collected by Lavender and colleagues[13] were from patients from India. Additionally, *M orygis* has been detected in cattle and rhesus monkeys in Bangladesh.[11,14] Collectively these data indicate that members of the MTBC complex other than *M bovis* might be relatively more prevalent in livestock in these countries. In 2018, Brites and colleagues[26] proposed that the distribution of *M orygis* and *M bovis* could be traced back to two independent cattle domestication events, wherein *M orygis* became a pathogen of cattle in south Asia, primarily Bos indicus, while *M bovis* became a pathogen of Bos taurus. Virulence studies comparing pathogenicity between these two subspecies in infected cattle breeds are yet to be done. Conversely, Rahim and colleagues[14] suggested that the global distribution of these subspecies indicates the emergence of *M orygis* before *M bovis* from a common MTBC ancestor, wherein *M*

*orygis* might have been dispersed to south Asia with the migration of humans and become established before the arrival of *M bovis.* However, previous establishment of an MTBC lineage does not negate the potential for introduction of another lineage at a later point in time, as shown by the widespread dispersal of Mtb L2 and L4.[27]

We observed *M tuberculosis* isolates from south Asia to be dispersed across lineages 1–4 with an enrichment of Mtb L1 and L3, concordant with previous reports.[27] MTBC sequences of cattle origin from south Asia were distributed among lineages Mtb L1 and *M orygis*, highlighting the need to understand pathogenicity and transmission dynamics of these pathogens in cattle. *M tuberculosis* transmission from humans to cattle has been reported, which might have important implications for transmission control in countries with a high burden of endemic tuberculosis.[28,29] Collectively, our data support the occurrence of zoonotic tuberculosis in India but suggest an association with *M orygis*, and possibly *M tuberculosis*, rather than *M bovis.*

In this molecular epidemiological study, we screened almost 1000 clinical tuberculosis isolates from positive MGIT cultures. Our approach has several advantages compared with previous studies.[8–10] The use of positive broth cultures ensured that active tuberculosis cases were microbiologically confirmed, unlike PCR-based investigations done on direct clinical samples without culture corroboration. The two-step protocol enabled provisional identification by PCR and assessment of inconclusive results by WGS. In addition, compared with conventional PCR, our five-probe multiplex rtPCR allowed for streamlined differentiation of many MTBC subspecies in a single reaction (appendix 1 p 5), showing promise for easy adoption in routine screening for zoonotic tuberculosis, depending on the needs and resources of the location. Furthermore, the detection of *M bovis* BCG served as a positive control for the PCR assays to identify *M bovis*, were it present. However, deletion of RD1, along with WGS analysis,

confirmed that these isolates were the Russian BCG strain, consistent with the vaccine currently used in India. An important limitation is our single-centre surveillance approach, meaning the isolates collected might be biased to locations within or nearby Tamil Nadu or to the patients who travelled to the study site. Future studies need to be done in other areas of India and other countries in south Asia to obtain a representative dataset. Furthermore, WGS analysis indicated that RD12 might be an inadequate marker for detection of *M tuberculosis*. This region was deleted in two *M tuberculosis* isolates, and seven *M tuberculosis* isolates were first categorised as inconclusive due to an SNP in the RD12 reverse primer sequence.

In conclusion, this study indicates that *M bovis* might be uncommon in India, and thus its detection is potentially an ineffective proxy for the prevalence of zoonotic tuberculosis. The operational definition of zoonotic tuberculosis should be broadened to include other MTBC subspecies capable of causing human disease. The increasing evidence supporting *M orygis* endemicity in south Asia and the identification of *M tuberculosis* in cattle highlight the importance of a One Health approach, involving multisectoral collaboration across the veterinary and clinical sectors, to the control of tuberculosis in India.

## 7. Contributors

SCD designed molecular assays, prepared clinical samples, did PCR assays, assisted with whole-genome sequencing (WGS), analysed results, and wrote the first draft of the manuscript. SS validated assays, prepared clinical samples, did PCR assays, and assisted with WGS. SND prepared clinical samples, did PCR assays, and assisted with WGS. MAS, TS, and SR-A did the bioinformatics analysis, and constructed and interpreted phylogenetic trees. JSM contributed to study design, provided clinical samples, and supervised laboratory studies at Christian Medical

College in Vellore, India. MVen did Sanger sequencing assays and analysed associated data. NB, DC, RK, and SC did WGS and Sanger sequencing assays and associated analyses. NJi and SM supervised WGS studies at Lala Lajpat Rai University of Veterinary and Animal Sciences in Hisar, India. PD and MVee contributed to analyses of animal-associated isolates. NJu assisted in study design and interpretation. VK and MAB directed study design, supervised the project, acquired funding, and contributed to writing the manuscript. All authors reviewed and edited the final version of the manuscript.

## 8. Declaration of interests

NJu is the Senior Program Officer at the Bill & Melinda Gates Foundation. VK reports grants from the Bill & Melinda Gates Foundation, during the conduct of this study. MAB reports grants from the Canadian Institutes for Health Research, during the conduct of this study. All other authors declare no competing interests.

## 9. Acknowledgments

The current study was supported by the Bill & Melinda Gates Foundation (grant OPP1176950 to VK) and the Canadian Institutes for Health Research (grant FDN148362 to MAB). We would like to thank Shalini Elangovan and Deepa Mani for assisting in the preparation of clinical samples. We would also like to thank Fiona McIntosh for providing technical support with assay development. Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published maps and tables.

## 10. References


1. WHO. Global tuberculosis report 2019. Geneva: World Health Organization, 2019.
2. Waters WR, Palmer MV, Buddle BM, Vordermeier HM. Bovine tuberculosis vaccine research: historical perspectives and recent advances. *Vaccine* 2012; **30:** 2611–22.
3. WHO, Food and Agriculture Organization of the UN, World Organisation for Animal Health. Zoonotic tuberculosis. 2017. https://www.who.int/tb/areas-of-work/zoonotic-tb/ZoonoticTBfactsheet2017.pdf (accessed Nov 18, 2019).
4. WHO, Food and Agriculture Organization of the UN, World Organisation for Animal Health. Roadmap for zoonotic tuberculosis. 2017. https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/Tuberculosis/Roadmap_zoonotic_TB.pdf (accessed Nov 18, 2019).
5. WHO. The end TB strategy. Global strategy and targets for tuberculosis prevention, care and control after 2015. May 2014. http://www.who.int/tb/strategy/End_TB_Strategy.pdf?ua=1 (accessed Nov 18, 2019).
6. United States Department of Agriculture. Livestock and poultry: world markets and trade. April 9, 2020. https://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf. (accessed May 22, 2020).
7. Srinivasan S, Easterling L, Rimal B, et al. Prevalence of bovine tuberculosis in India: a systematic review and meta-analysis. *Transbound Emerg Dis* 2018; **65:** 1627–40.
8. Bapat PR, Dodkey RS, Shekhawat SD, et al. Prevalence of zoonotic tuberculosis and associated risk factors in central Indian populations. *J Epidemiol Glob Health* 2017; **7:** 277–83.

9. Prasad HK, Singhal A, Mishra A, et al. Bovine tuberculosis in India: potential basis for zoonosis. *Tuberculosis (Edinb)* 2005; **85:** 421–28.
10. Shah NP, Singhal A, Jain A, et al. Occurrence of overlooked zoonotic tuberculosis: detection of *Mycobacterium bovis* in human cerebrospinal fluid. *J Clin Microbiol* 2006; **44:** 1352–58.
11. van Ingen J, Rahim Z, Mulder A, et al. Characterization of *Mycobacterium orygis* as *M tuberculosis* complex subspecies. *Emerg Infect Dis* 2012; **18:** 653–55.
12. Marcos LA, Spitzer ED, Mahapatra R, et al. *Mycobacterium orygis* lymphadenitis in New York, USA. *Emerg Infect Dis* 2017; **23:** 1749–51.
13. Lavender CJ, Globan M, Kelly H, et al. Epidemiology and control of tuberculosis in Victoria, a low-burden state in south-eastern Australia, 2005–2010. *Int J Tuberc Lung Dis* 2013; **17:** 752–58.
14. Rahim Z, Thapa J, Fukushima Y, et al. Tuberculosis caused by *Mycobacterium orygis* in dairy cattle and captured monkeys in Bangladesh: a new scenario of tuberculosis in south Asia. *Transbound Emerg Dis* 2017; **64:** 1965–69.
15. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res* 2012; **40:** e115.
16. Halse TA, Escuyer VE, Musser KA. Evaluation of a single-tube multiplex real-time PCR for differentiation of members of the *Mycobacterium tuberculosis* complex in clinical specimens. *J Clin Microbiol* 2011; **49:** 2562–67.
17. Saïd-Salim B, Mostowy S, Kristof AS, Behr MA. Mutations in *Mycobacterium tuberculosis Rv0444c*, the gene encoding anti-sigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. *Mol Microbiol* 2006; **62:** 1251–63.

18. van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol* 1991; **29:** 2578–86.
19. Kapur V, Li LL, Hamrick MR, et al. Rapid mycobacterium species assignment and unambiguous identification of mutations associated with antimicrobial resistance in *Mycobacterium tuberculosis* by automated DNA sequencing. *Arch Pathol Lab Med* 1995; **119:** 131–38.
20. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30:** 1312–13.
21. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 2015; **31:** 2877–78.
22. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019; **47:** W256–59.
23. Olea-Popelka F, Muwonge A, Perera A, et al. Zoonotic tuberculosis in human beings caused by *Mycobacterium bovis*—a call for action. *Lancet Infect Dis* 2017; **17:** e21–25.
24. Smith T. A comparative study of bovine tubercle bacilli and of human bacilli from sputum. *J Exp Med* 1898; **3:** 451–511.
25. Lipworth S, Jajou R, de Neeling A, et al. SNP-IT tool for identifying subspecies and associated lineages of *Mycobacterium tuberculosis* complex. *Emerg Infect Dis* 2019; **25:** 482–88.

26. Brites D, Loiseau C, Menardo F, et al. A new phylogenetic framework for the animal-adapted *Mycobacterium tuberculosis* complex. *Front Microbiol* 2018; **9:** 2820.

27. Wiens KE, Woyczynski LP, Ledesma JR, et al. Global variation in bacterial strains that cause tuberculosis disease: a systematic review and meta-analysis. *BMC Med* 2018; **16:** 196.

28. Sweetline Anne N, Ronald BS, Kumar TM, Kannan P, Thangavelu A. Molecular identification of *Mycobacterium tuberculosis* in cattle. *Vet Microbiol* 2017; **198:** 81–87.

29. Ocepek M, Pate M, Zolnir-Dovc M, Poljak M. Transmission of *Mycobacterium tuberculosis* from human to cattle. *J Clin Microbiol* 2005; **43:** 3555–57.

## 11. Figures

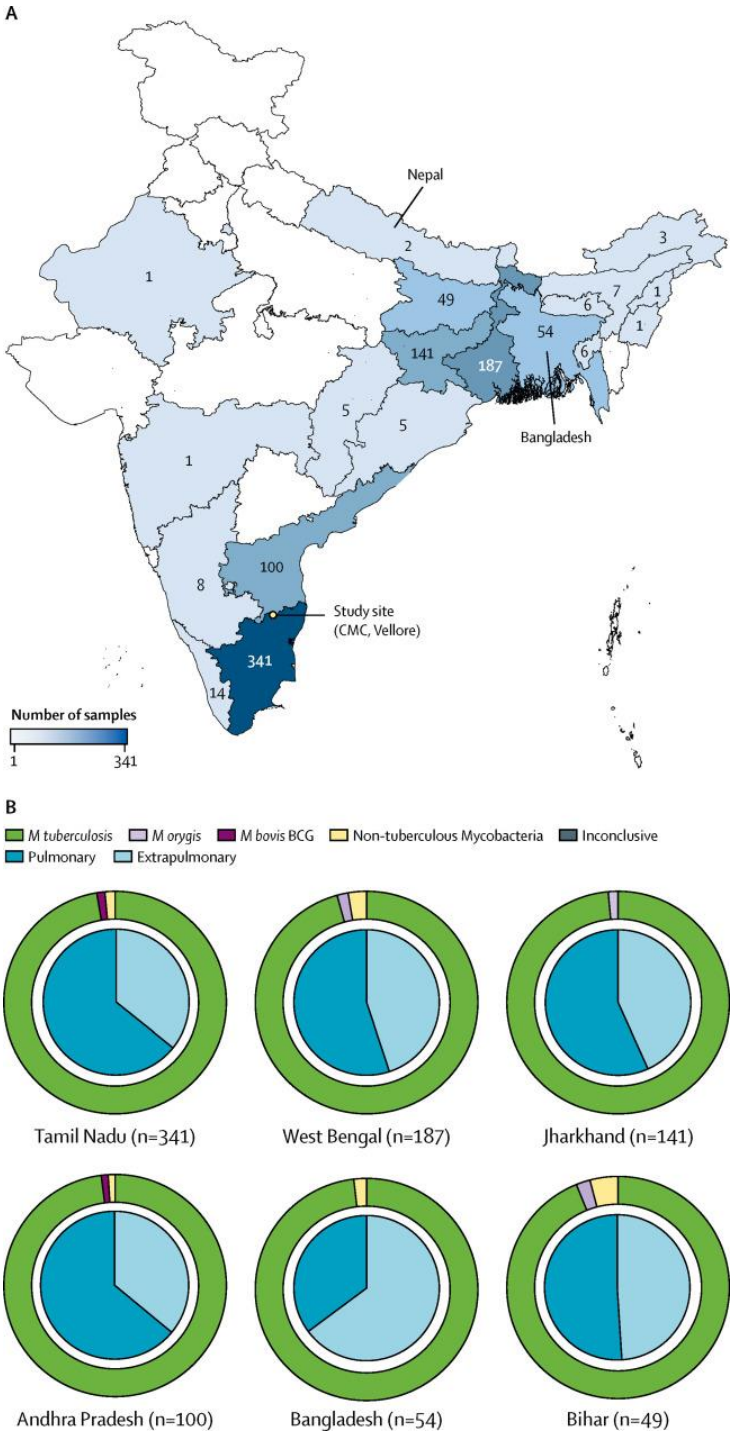



**Figure 1** Distribution and sample types of patient isolates within India and surrounding countries (A) Geographical distribution of the collected isolates. Numbers indicate isolates collected per area. No isolates were screened from locations in white. (B) Sample types of isolates from

locations where 20 or more samples were collected. Inner pie charts show the proportion of pulmonary and extrapulmonary isolates collected from each location. Outer doughnut charts indicate the proportion of mycobacterial subspecies collected from each location. Exact numbers are provided in appendix 3. CMC=Christian Medical College. *M tuberculosis=Mycobacterium tuberculosis. M orygis=Mycobacterium orygis. M bovis=Mycobacterium bovis*.

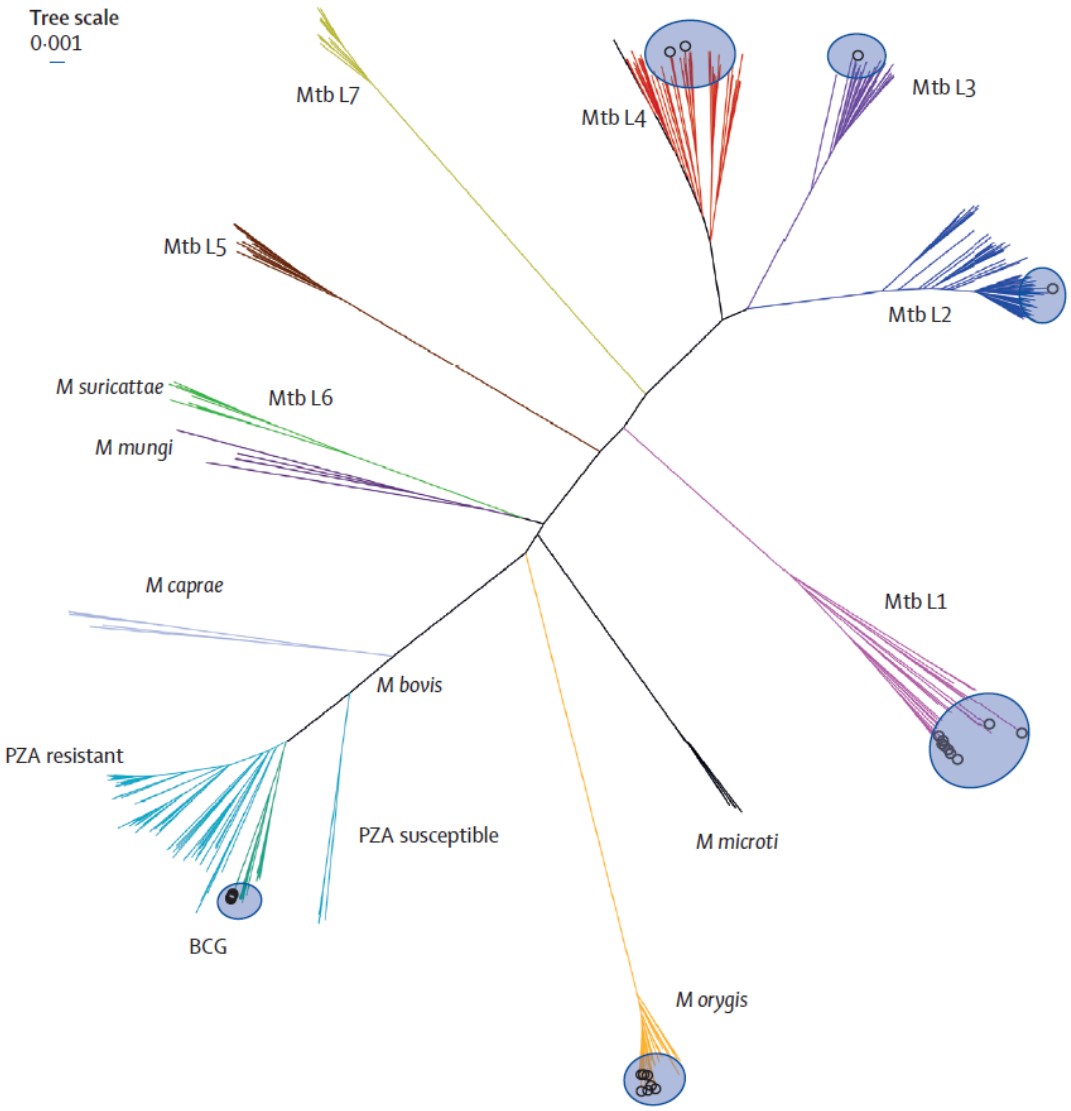



**Figure 2** Phylogenies of newly sequenced isolates in the context of genetic diversity among global MTBC isolates

The unrooted tree shows clustering of 25 isolates sequenced in this study (circles) in the context of representative samples (n=373) of MTBC lineages from around the world (appendix 2 p 1). MTBC=*Mycobacterium tuberculosis* complex. Mtb L1=*Mycobacterium tuberculosis* lineage 1. *M suricattae=Mycobacterium suricattae*. *M mungi=Mycobacterium mungi. M caprae=Mycobacterium caprae. M microti=Mycobacterium microti. M bovis=Mycobacterium bovis*.PZA=pyrazinamide. *M orygis=Mycobacterium orygis*.

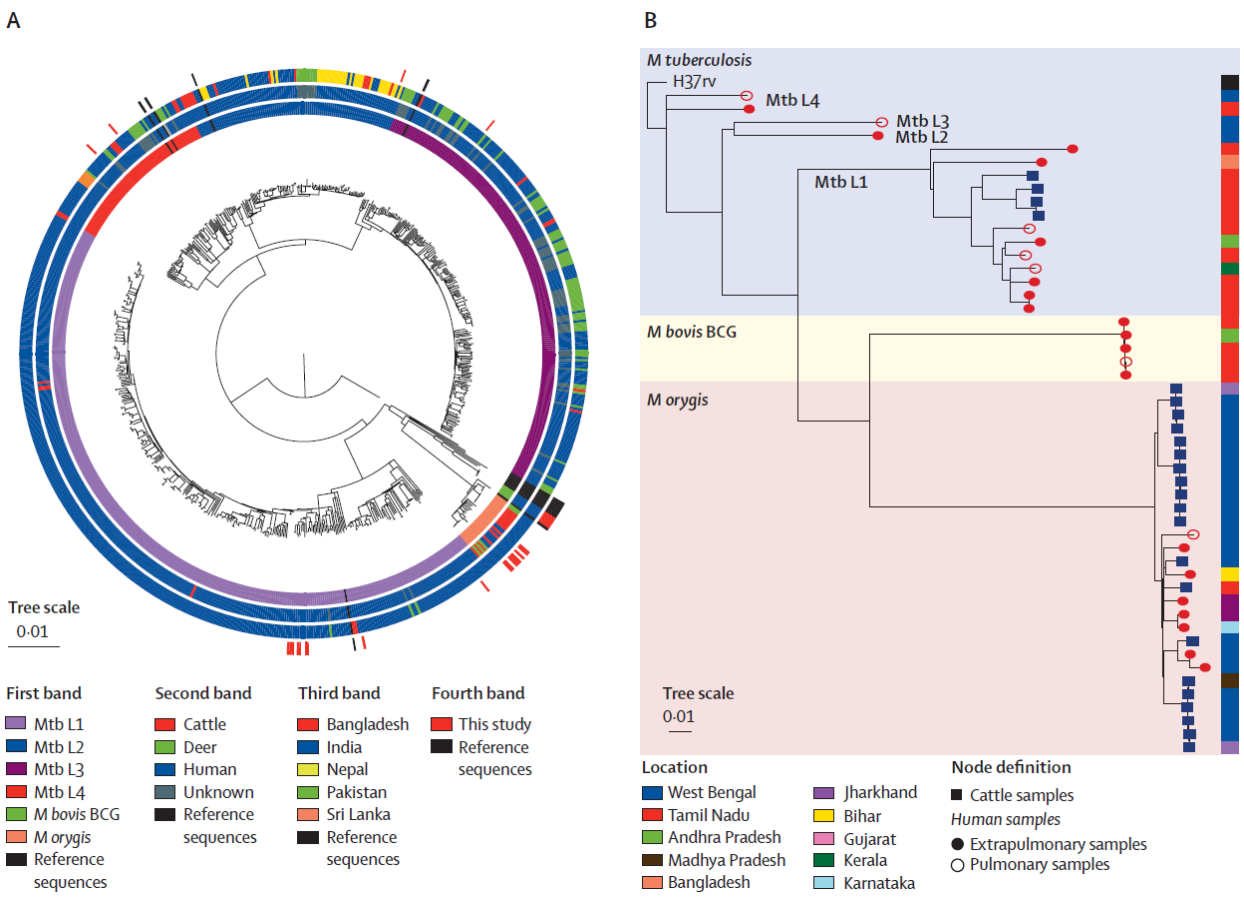



**Figure 3** Phylogenies of newly sequenced isolates in the context of *Mycobacterium tuberculosis* complex isolates in south Asia

(A) The outermost band shows our 25 sequenced isolates in the context of south Asian sequences. Other bands represent 715 MTBC sequences downloaded from the National Center for Biotechnology Information Sequence Read Archive. The sequences represent different lineages (first band from the middle), different host species (second band), and different countries (third band). Cattle isolates included one bison isolate as part of the Bovidae family. (B) Represented in the dendrogram are isolates from this study, all downloaded *M orygis* sequences (all of cattle origin; 20 sequences), and all downloaded Mtb L1 sequences of cattle origin (four). One sample was from Bangladesh and the rest were from Indian states. *M*

*tuberculosis=Mycobacterium tuberculosis*. Mtb L1=*M tuberculosis* lineage 1. *M bovis=Mycobacterium bovis. M orygis=Mycobacterium orygis*.

## 12. Tables

| | Pulmonary sample (n=548 patients) | Extrapulmonary sample (n=392 patients) | Total (n=940 patients) |
|---|---|---|---|
| **Age, years** | | | |
| 0–9 | 7 (1%) | 19 (5%) | 26 (3%) |
| 10–19 | 54 (10%) | 42 (11%) | 96 (10%) |
| 20–29 | 112 (20%) | 101 (26%) | 213 (23%) |
| 30–39 | 100 (18%) | 83 (21%) | 183 (19%) |
| 40–49 | 99 (18%) | 71 (18%) | 170 (18%) |
| 50–59 | 88 (16%) | 54 (14%) | 142 (15%) |
| 60–69 | 64 (12%) | 14 (4%) | 78 (8%) |
| ≥70 | 24 (4%) | 8 (2%) | 32 (3%) |
| **Sex** | | | |
| Female | 185 (34%) | 178 (45%) | 363 (39%) |
| Male | 363 (66%) | 214 (55%) | 577 (61%) |
| **Location** | | | |
| India | 528 (96%) | 356 (91%) | 884 (94%) |
| South | 297 (56%) | 168 (47%) | 465 (53%) |
| East | 209 (40%) | 173 (49%) | 382 (43%) |
| Northeast | 15 (3%) | 13 (4%) | 28 (3%) |
| Central | 5 (1%) | 0 | 5 (1%) |
| North | 2 (<1%) | 1 (<1%) | 3 (<1%) |
| West | 0 | 1 (<1%) | 1 (<1%) |
| Bangladesh | 19 (3%) | 35 (9%) | 54 (6%) |
| Nepal | 1 (<1%) | 1 (<1%) | 2 (<1%) |

**Table 1** Characteristics of patients from whom isolates were obtained and screened for zoonotic tuberculosis

Data are number of patients (%). Percentages do not always equal 100% due to rounding. For the purposes of the present study, regions of India are divided by location as follows: south India includes the Andaman and Nicobar Islands, Andhra Pradesh, Karnataka, Kerala, Lakshadweep,

Puducherry, Tamil Nadu, and Telangana; east India includes West Bengal, Bihar, Jharkhand, and Odisha; northeast India includes Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura; central India includes Chhattisgarh and Madhya Pradesh; north India includes Jammu and Kashmir, Himachal Pradesh, Punjab, Chandigarh, Uttarakhand, Haryana, National Capital Territory of Delhi, Rajasthan, and Uttar Pradesh; and west India includes Dadra and Nagar Haveli, Daman and Diu, Goa, Gujarat, and Maharashtra.

| | Pulmonary isolates (n=548) | Extrapulmonary isolates (n=392) | Total isolates (n=940) |
|---|---|---|---|
| *Mycobacterium tuberculosis* | 535 (97·6%) | 378 (96·4%) | 913 (97·1%)* |
| *Mycobacterium orygis* | 1 (0·2%) | 6 (1·5%) | 7 (0·7%)† |
| *Mycobacterium bovis* | 0 | 0 | 0 |
| *Mycobacterium bovis* BCG | 1 (0·2%) | 4 (1·0%) | 5 (0·5%)† |
| Non-tuberculous mycobacteria | 11 (2·0%) | 4 (1·0%) | 15 (1·6%) |

WGS=whole-genome sequencing. *Including two *M tuberculosis* isolates negative for RD12 on PCR but confirmed by WGS (both extrapulmonary); eight inconclusive isolates on PCR identified as *M tuberculosis* by WGS (six pulmonary, two extrapulmonary); and one pulmonary isolate identified as *M bovis* BCG by PCR and *M tuberculosis* by WGS. †All detected by PCR and confirmed by WGS.

*Table 2:* **Isolate subspecies determined by PCR and WGS**

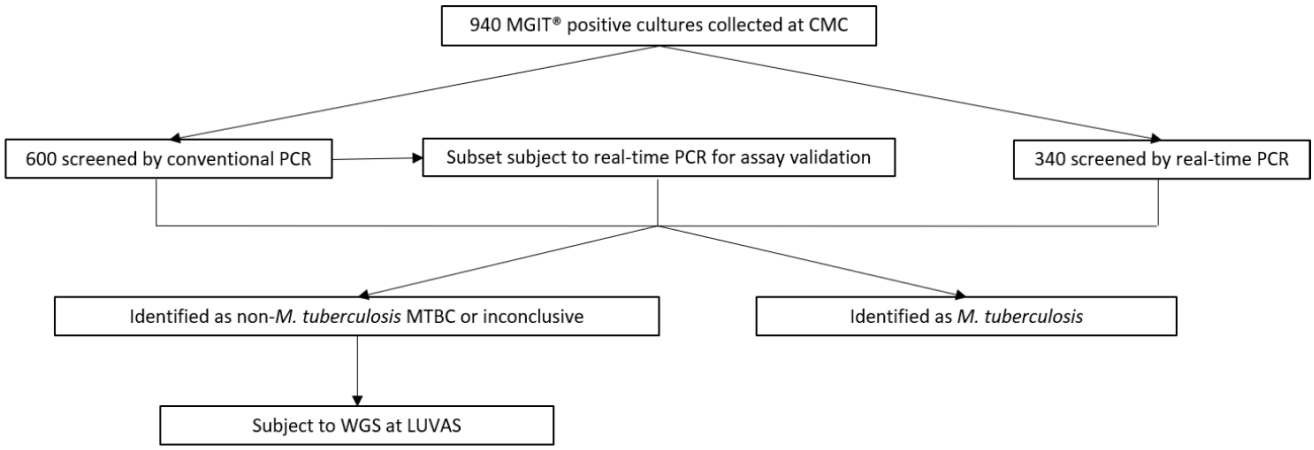
**Table 2** Isolate subspecies determined by PCR and WGS

WGS=whole-genome sequencing.

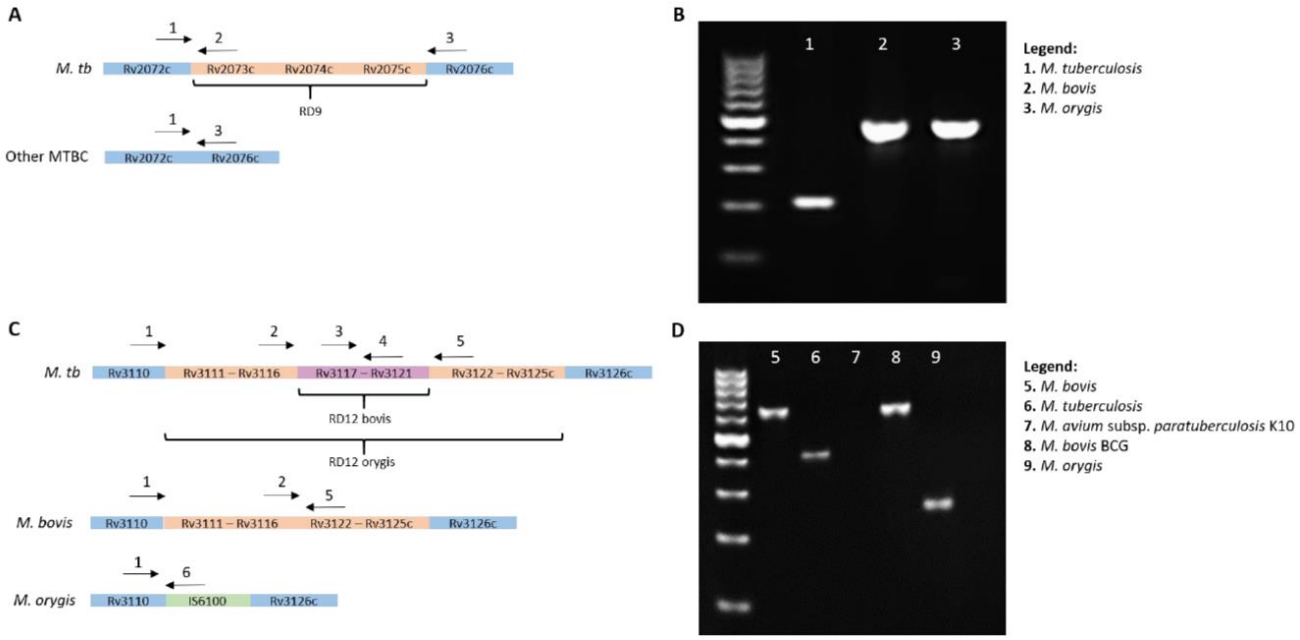
* Including two *M tuberculosis* isolates negative for RD12 on PCR but confirmed by WGS (both extrapulmonary); eight inconclusive isolates on PCR identified as *M tuberculosis* by WGS (six pulmonary, two extrapulmonary); and one pulmonary isolate identified as *M bovis* BCG by PCR and *M tuberculosis* by WGS.

† All detected by PCR and confirmed by WGS.

## 13. Supplementary figures

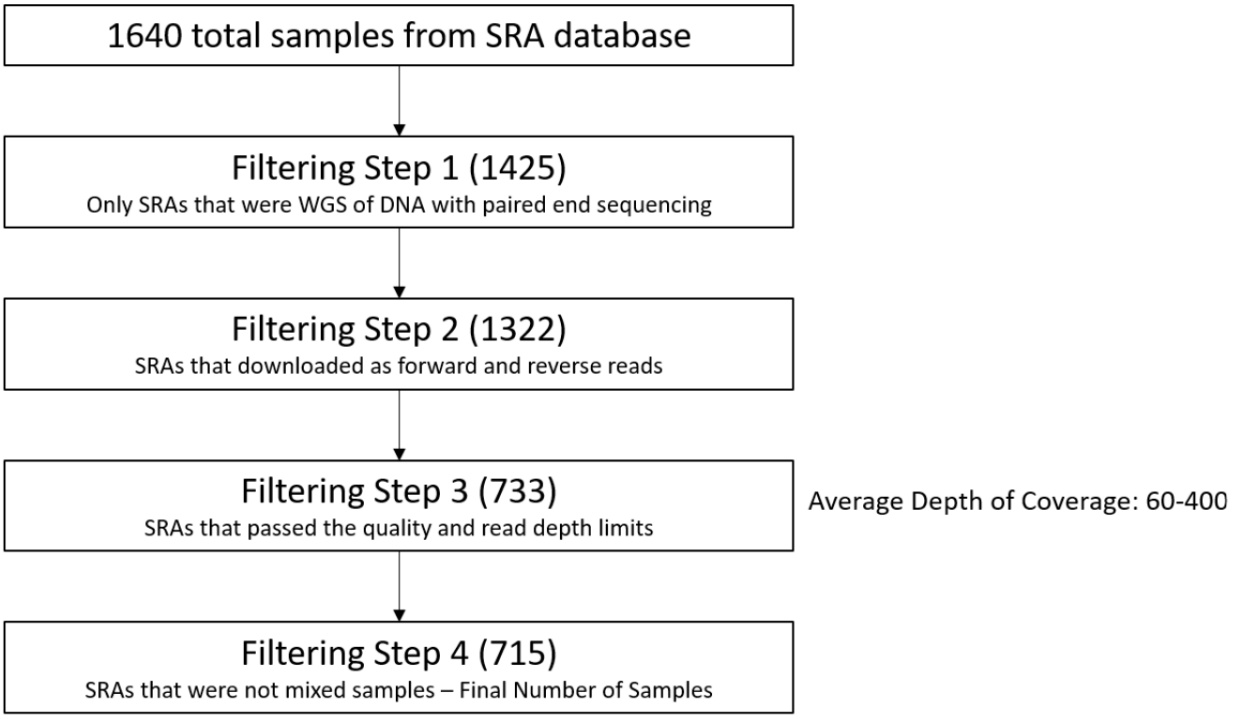



**Appendix 1 (p 1) Supplementary Figure 1** Study design




**Appendix 1 (p 2) Supplementary Figure 2** Conventional PCR to detect differences in deletions of RD9 and RD12

(A) A three-primer PCR reaction was developed to detect the presence of absence of RD9, which is found in *M. tb* but is absent in all other MTBC members. (B) A 209bp band is amplified when *M. tb* is present and a 410bp band is amplified when other MTBC members are present. (C) A six-primer PCR was developed to detect differences in the deletion size of RD12. RD12 is present in *M. tb* and absent in *M. bovis* and *M. orygis.* In *M. orygis,* the RD12 deletion is larger and is replaced with IS6100. (D) A 409bp band is amplified when *M. tb* is present, a 615bp band is amplified when *M. bovis* or BCG is present, and a 264bp band is amplified when *M. orygis* is present.



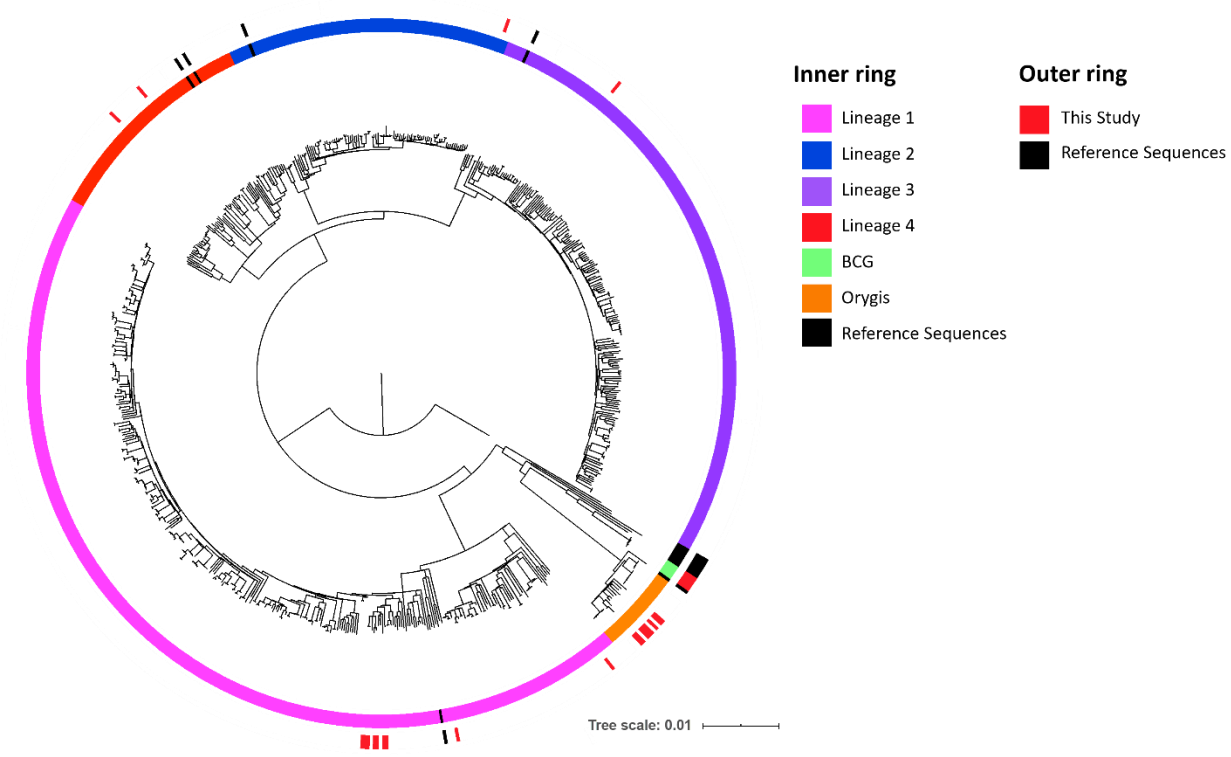

**Appendix 1 (p 8) Supplementary Figure 3** Selection and filtering pipeline of downloaded SRA MTBC genomes from South Asia

A total of 1640 SRAs were downloaded from NCBI with the search terms ("*Mycobacterium bovis*" OR "*Mycobacterium tuberculosis*" OR "*Mycobacterium africanum*" OR "*Mycobacterium orygis*" OR "*Mycobacterium canetti*" OR "*Mycobacterium caprae*" OR "*Mycobacterium bovis* BCG" NOT "H37Rv" NOT "H37Ra") from ("India" OR "Bangladesh" OR "Nepal" OR "Sri Lanka" OR "Pakistan"). Through multiple filtering steps, the total number of SRAs analyzed in the phylogenetic tree in Fig 2B was 715. The lists of SRAs after each filtering step can be found in the Supplemental Tables 4-8.

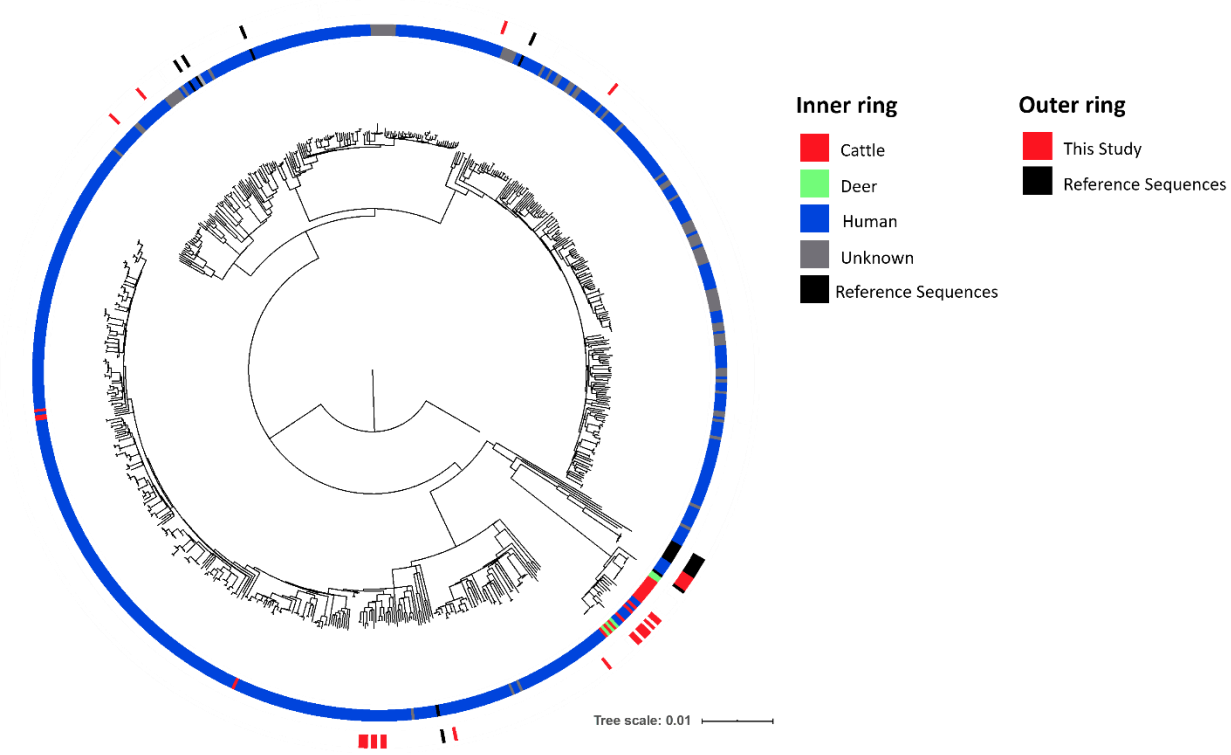
| | P429 | E120 | E36 | E374 | E313 | E65 | E157 |
|---|---|---|---|---|---|---|---|
| P429 | 0 | 85.47 | 88.27 | 89.10 | 89.66 | 88.84 | 89.52 |
| E120 | 282 | 0 | 86.30 | 87.55 | 88.13 | 87.27 | 87.84 |
| E36 | 271 | 285 | 0 | 91.53 | 92.02 | 91.19 | 91.77 |
| E374 | 252 | 259 | 214 | 0 | 93.29 | 92.38 | 93.07 |
| E313 | 239 | 247 | 202 | 170 | 0 | 92.93 | 93.64 |
| E65 | 258 | 265 | 223 | 193 | 177 | 0 | 97.39 |
| E157 | 242 | 253 | 208 | 175 | 161 | 66 | 0 |

**Appendix 1 (p 12) Supplementary Figure 4** SNP distances between 7 *M. orygis* isolates from this study

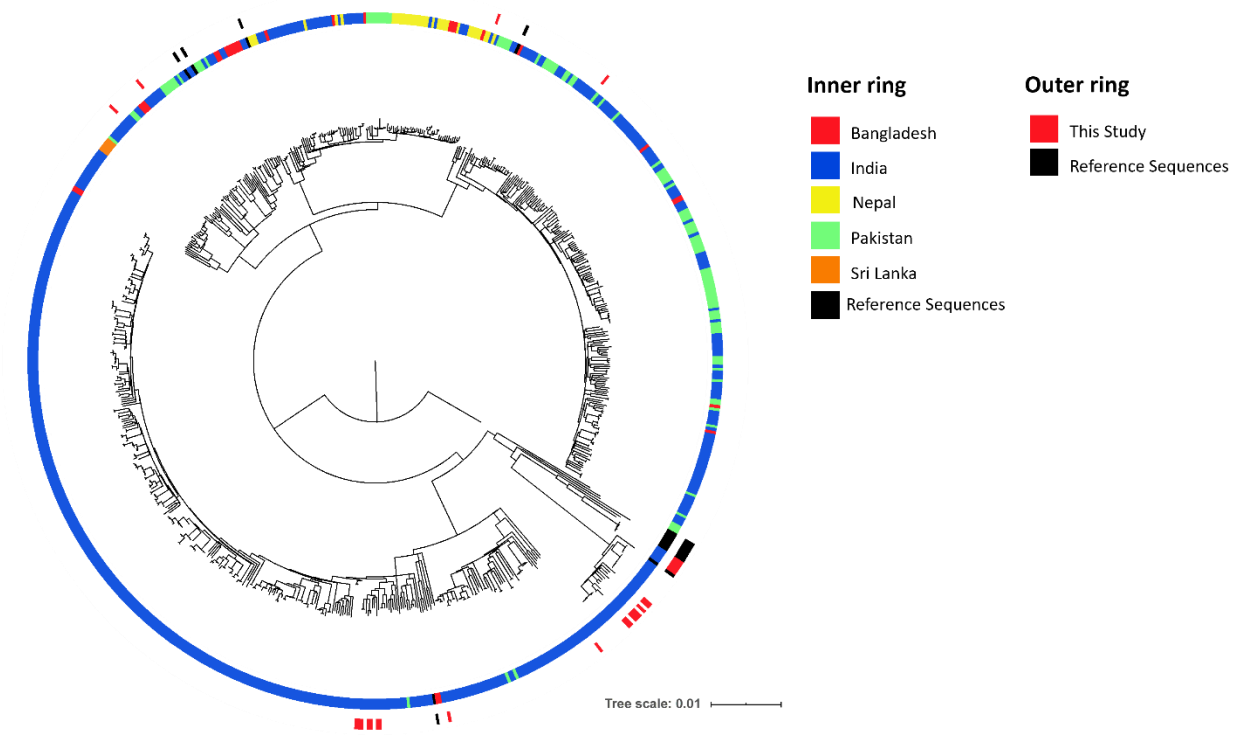
The intensity of the color corresponds to the distance between isolates. The bottom portion of the matrix indicates the number of SNPs between the two isolates. The top portion indicates the percent of total SNPs shared between them.




**Appendix 1 (p 13) Supplementary Figure 5** Maximum likelihood phylogenetic tree of newly sequenced isolates and 715 MTBC genomes collected from South Asia with lineage metadata




**Appendix 1 (p 14) Supplementary Figure 6** Maximum likelihood phylogenetic tree of newly sequenced isolates and 715 MTBC genomes collected from South Asia with host metadata




**Appendix 1 (p 15) Supplementary Figure 7** Maximum likelihood phylogenetic tree of newly sequenced isolates and 715 MTBC genomes collected from South Asia with country metadata

## 14. Supplementary tables

| Assay | Name | Sequence (5' → 3') | Dye | Quencher |
|---|---|---|---|---|
| Conventional PCR | RD9_Forward | CCGATACCATGCAACAACGG | | |
| | RD9_Reverse1 | CGGTCTCTCCGAGCATTC | | |
| | RD9_Reverse2 | GCTCGAGCTAGACCTGCAC | | |
| | RD12Mtb_Forward | GTATTTGCGCCCATATCCTGG | | |
| | RD12Mtb_Reverse | CCTGGCTTCAAGCACCATTC | | |
| | RD12Mbovis_Forward | GGCCATCAACGTCAAGAACCTC | | |
| | RD12Mbovis_Reverse | CGAACTCGTATTTTGTGGCCAC | | |
| | RD12Morygis_Forward | GTGGAAATGGAAGCGTTGACC | | |
| | RD12Morygis_Reverse | GGTACCTCCTCGATGAACCAC | | |
| Real-time PCR | Rv0444c_Probe | CTCGGCTGACCCGA | FAM | MGBNFQ |
| | Rv0444c_Forward | GATGCTGGGCACCATTGTC | | |
| | Rv0444c_Reverse | GCCCACCGGTACCATCTTG | | |
| | RD1 Probe | CACTCTGAGAGGTTGTCA | VIC | MGBNFQ |
| | RD1_Forward | CCCTTTCTCGTGTTTATACGTTTGA | | |
| | RD1_Reverse | GCCATATCGTCCGGAGCTT | | |
| | RD9 Probe | AGGTTTCA+CCTTCGAC+CC | TEX615 | BHQ |
| | RD9_Forward | TGCGGGCGGACAACTC | | |
| | RD9_Reverse | CACTGCGGTCGGCATTG | | |
| | RD12 Probe | TGCGCTGACCCCAC | NED | MGBNFQ |
| | RD12_Forward | CGTTGGAACGCGAAATACG | | |
| | RD12_Reverse | CCAGGATATGGGCGCAAAT | | |
| | extRD9_Probe | G+TT+CTTCAG+CTGGT+CC | CY5 | BHQ |
| | extRD9_Forward | GCCACCACCGACTCATAC | | |
| | extRD9_Reverse | CGAGGAGGTCATCCTGCTCTA | | |

**Appendix 1 (p 3) Supplementary Table 1** Primer and probe sequences for conventional and real-time PCR assays

The real-time PCR RD1, RD9, RD12, and ext-RD9 probes and primers are as described in Halse *et al*. The Rv0444c, RD1, and RD12 probes are Taqman MGB probes. The RD9 and ext-RD9 probes are locked nucleic acid probes. A '+' indicates insertion of a locked nucleic acid base.

| Assay | Master mix | Thermocycling conditions |
|---|---|---|
| Conventional PCR | 6.25 µl of 10X *Taq* buffer (Thermo Scientific) | Initial denaturation 94°C for 3 minutes |
| | 6.25 µl acetamide 50% (wt/vol) | 35 cycles of: |
| | 1.6 mM $MgCl_2$ | - Denaturation at 94°C for 30 seconds |
| | 0.2 mM deoxynucleoside triphosphates (dNTPs) | - Annealing at 55°C for 1 minute |
| | 2.5 U per reaction *Taq* polymerase (Thermo Scientific) | - Elongation at 72°C for 1 minute, |
| | 500 nM of each primer | Final elongation step at 72°C for 10 minutes |
| | 5µl of template DNA | |
| | Sterile water | |
| | 50µl total volume | |
| Real-time PCR | 10 µl TaqMan multiplex master mix (Applied Biosystems) | 95°C for 10 minutes |
| | 450 nM of each primer | 40 cycles of: |
| | 125 nM of each probe | - 95°C for 15s |
| | 1 µl of template DNA | - 60°C for 1 minute |
| | Sterile water | |
| | 20µl total volume | |

**Appendix 1 (p 4) Supplementary Table 2** Master mix preparation and thermocycling conditions for conventional and real-time PCR assays

| | RD1 | RD9 | RD12 | Rv0444c | Ext-RD9 |
|---|---|---|---|---|---|
| ***M. tuberculosis*** | + | + | + | - | + |
| ***M. orygis*** | + | - | - | + | + |
| ***M. bovis*** | + | - | - | - | + |
| ***M. bovis* BCG** | - | - | - | - | + |
| ***M. africanum*** | + | - | + | - | + |
| ***M. microti*** | - | - | + | - | + |
| **NTMs** | - | - | - | - | - |

**Appendix 1 (p 5) Supplementary Table 3** Interpretation of RT-PCR results to determine MTBC sample identity

Appendix 2 containing supplementary Tables 4-7 has been included as an additional Excel file with this thesis.

**Appendix 2 Supplementary Tables 4-7**

| Sample type | Number |
|---|---|
| Abdomen/Peritoneal fluid | 15 |
| Anal fistula | 2 |
| Biopsy | 5 |
| Bone | 43 |
| Bone marrow | 2 |
| Brain abscess | 1 |
| Colon | 5 |
| Cerebrospinal fluid | 27 |
| Fluid | 5 |
| Lymph node | 162 |
| Muscle abscess | 5 |
| Pus | 56 |
| Pericardium | 1 |
| Skin | 1 |
| Synovium | 8 |
| Tissue | 40 |
| Urine | 13 |
| Unspecified | 1 |
| Total | 392 |

**Appendix 1 (p 9) Supplementary Table 8** Number and tissue types of extrapulmonary samples

Extrapulmonary samples were defined as cultures from tissues other than the lungs or lung fluid.

| Isolate number | Identification by real-time PCR | Average coverage | Genome coverage (%) | Phred quality score R1 | Phred quality score R2 | SNP count | Identification by WGS |
|---|---|---|---|---|---|---|---|
| P70 | *M. bovis* BCG | 103.2 | 99.35 | 34.6 | 31.1 | 803 | *M. bovis* BCG Russia |
| P280 | *M. bovis* BCG | 80.4 | 99.15 | 34.5 | 31.1 | 1834 | *M. tuberculosis* lineage 2 |
| E50 | *M. bovis* BCG | 54.4 | 99.49 | 33.7 | 30.3 | 784 | *M. bovis* BCG Russia |
| E110 | *M. bovis* BCG | 85.0 | 99.68 | 33.8 | 29.6 | 797 | *M. bovis* BCG Russia |
| E280 | *M. bovis* BCG | 77.2 | 99.33 | 34.1 | 28.0 | 804 | *M. bovis* BCG Russia |
| E396 | *M. bovis* BCG | 98.7 | 99.38 | 33.9 | 28.8 | 817 | *M. bovis* BCG Russia |
| P326 | Inconclusive | 53.5 | 99.27 | 33.9 | 29.2 | 2322 | *M. tuberculosis* lineage 1 |
| P414 | Inconclusive | 82.2 | 99.42 | 33.6 | 27.8 | 2412 | *M. tuberculosis* lineage 1 |
| P448 | Inconclusive | 78.7 | 99.36 | 34.8 | 30.3 | 987 | *M. tuberculosis* lineage 4 |
| P465 | Inconclusive | 121.2 | 99.46 | 34.1 | 28.7 | 2476 | *M. tuberculosis* lineage 1 |
| E343 | Inconclusive | 277.6 | 99.53 | 35.1 | 32.5 | 2607 | *M. tuberculosis* lineage 1 |
| E369 | Inconclusive | 75.1 | 99.33 | 33.8 | 29.5 | 2429 | *M. tuberculosis* lineage 1 |
| E379 | Inconclusive | 115.2 | 99.51 | 34.2 | 29.3 | 2499 | *M. tuberculosis* lineage 1 |
| E415 | Inconclusive | 153.8 | 99.53 | 33.9 | 28.9 | 2510 | *M. tuberculosis* lineage 1 |
| E277 | *M. tuberculosis* | 138.6 | 99.39 | 34.8 | 30.4 | 1691 | *M. tuberculosis* lineage 3 |
| E428 | *M. tuberculosis* | 105.6 | 99.66 | 34.0 | 29.0 | 2594 | *M. tuberculosis* lineage 1 |
| P429 | *M. orygis* | 12.2 | 97.49 | 34.3 | 30.7 | 2299 | *M. orygis* |
| E36 | *M. orygis* | 72.9 | 98.26 | 33.5 | 28.2 | 2547 | *M. orygis* |
| E65 | *M. orygis* | 34.6 | 97.75 | 33.4 | 29.2 | 2408 | *M. orygis* |
| E120 | *M. orygis* | 11.1 | 97.15 | 33.9 | 31.1 | 2330 | *M. orygis* |
| E157 | *M. orygis* | 64.7 | 98.26 | 34.1 | 28.9 | 2478 | *M. orygis* |
| E313 | *M. orygis* | 58.8 | 98.28 | 34.1 | 31.6 | 2483 | *M. orygis* |
| E374 | *M. orygis* | 139.1 | 98.45 | 34.6 | 31.1 | 2566 | *M. orygis* |
| E186 | *M. tuberculosis* RD12 absent | 90.3 | 98.74 | 34.1 | 29.1 | 968 | *M. tuberculosis* lineage 4 |
| E363 | *M. tuberculosis* RD12 absent | 107.4 | 99.00 | 34.5 | 30.2 | 2474 | *M. tuberculosis* lineage 1 |

**Appendix 1 (p 10) Supplementary Table 9** Selection, library preparation and whole genome sequencing data of 25 selected isolates

| Sample name | PCR ID | Top BLAST match | Seq length | Query coverage | % Identity |
|---|---|---|---|---|---|
| E133 | NTM | *Mycobacterium phocaicum* | 430 | 100.00% | 99.77% |
| E153 | NTM | *Mycobacterium engbaekii* | 440 | 99.00% | 99.32% |
| E193 | NTM | *Mycobacterium abscessus* | 440 | 100.00% | 100.00% |
| P24 | NTM | *Mycobacterium sp. K328YA* | 438 | 94.00% | 100.00% |
| P30 | NTM | *Mycobacterium alvei* | 437 | 100.00% | 98.40% |
| P390 | NTM | *Mycobacterium abscessus* | 432 | 100.00% | 100.00% |
| P427 | NTM | *Mycobacterium intracellulare* | 419 | 100.00% | 100.00% |
| P146 | NTM | *Mycobacterium simiae* | 443 | 99.00% | 99.32% |
| P149 | NTM | *Mycobacterium intracellulare* | 424 | 100.00% | 99.76% |
| E22 | NTM | *Mycobacterium abscessus* | 357 | 100.00% | 99.72% |
| P219 | NTM | *Mycobacterium yongonense* | 359 | 100.00% | 99.44% |
| P281 | NTM | *Mycobacterium fortuitum* | 362 | 100.00% | 100.00% |
| P426 | NTM | *Mycobacterium intracellulare* | 333 | 100.00% | 100.00% |
| P81 | NTM | *Mycobacterium parascrofulaceum* | 381 | 100.00% | 99.48% |

**Appendix 1 (p 11) Supplementary Table 10** Hsp65 sanger sequencing results of non-tuberculous mycobacteria (NTM) isolates

Appendix 3 containing supplementary Tables 11-12 has been included as an additional Excel file with this thesis.

**Appendix 3 Supplementary Table 11-12** Meta data and pivot tables

## 15. Supplementary methods

**Appendix 1 (p 6-7) Supplemental methods**

*Bioinformatics*

Sequences were assessed using the United States Department of Agriculture Animal and Plant Health Inspection Service Veterinary Services pipeline vSNP (https://github.com/USDA-VS/vSNP). The vSNP pipeline involved a two-step process. Step 1 determined SNP positions called within the sequence. Paired FASTQ files were processed using BWA-MEM to align reads to a reference genome *M. tuberculosis* H37Rv (NC_000962.3) for sequences included in this study ([1]). Duplicate reads were tagged and removed using the Mark Duplicates tool from Picard v 2.20.2 (http://broadinstitute.github.io/picard). SNPs were called using FreeBayes v. 1.3.1 ([2]). Unmapped reads shorter than 64 base pairs were removed and low-coverage contigs with an average k-mer coverage of less than 5 were removed. Depth of coverage was calculated using Pysam (https://github.com/pysam-developers/pysam) and positions with zero coverage were added to the VCF file. Step 2 assessed SNPs called between closely related isolate groups to output SNP alignments, tables and phylogenetic trees. For a SNP to be considered in a group there must have been at least one position with an allele count (AC) =2, quality score >150 and map quality > 56. Once determined, SNPs were aligned using the following workflow. If the quality score for a SNP position was greater than 150, the alternate allele was called if AC=2. However, if AC=1, the position was called ambiguous. Deletions were called when the alternate

allele was a gap. If the quality score was between 50 and 150, the allele was marked N. If the quality score was less than 50, then the reference allele was called. Uninformative SNPs were not included. BAM files were used to visualize SNP calls. Unreliable positions due to read alignment error may have been removed from the analysis. The output SNP alignment was used to assemble a maximum likelihood phylogenetic tree using RAxML ([3]).

*Phylogenetic tree assembly*

To compare the 25 newly sequenced genomes in the context of sequences from South Asia, a NCBI Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra) search was performed using the search terms ("*Mycobacterium bovis*" OR "*Mycobacterium tuberculosis*" OR "*Mycobacterium africanum*" OR "*Mycobacterium orygis*" OR "*Mycobacterium canetti*" OR "*Mycobacterium caprae*" OR "*Mycobacterium bovis* BCG" NOT "H37Rv" NOT "H37Ra") from ("India" OR "Bangladesh" OR "Nepal" OR "Sri Lanka" OR "Pakistan"). This search yielded 1640 genomes. These sequences were then filtered prior to tree assembly (Supplementary Figure 3). A total of 215 were excluded because they were not from studies that performed with paired end sequencing. A further 103 were excluded as they did not contain forward and reverse read files. Another 589 sequences whose average depth of coverage was not between 60-400 were excluded. Finally, 18 were excluded after running vSNP step 2 due to the samples being mixed and generating multiple SNPs at all locations in the genome. In total, 715 sequences remained (Supplementary Table 5). All sequences were download from the SRA using the fasterq-dump tool from the sra toolkit v. 2.9.6 (https://ncbi.github.io/sra-tools/) and sequences were run through steps 1 and 2 of vSNP. Phylogenetic trees were constructed using vSNP to compare the 25 sequences from this study with the total 715 available genomes from South Asia

and a subset. Reference sequences were also included for comparison (Supplementary Table 6). Phylogenetic trees were rooted to *M. tuberculosis* H37Rv. To compare the sequences collected in this study in the context of the global MTC, treeSPAdes (http://cab.spbu.ru/software/spades/) was used to assemble reads for kSNP3 ([4]). Genomes assemblies had expected complete genome sizes. The kSNP3 manual instructions were followed using kchooser calculated kmer value. All phylogenetic trees were visualized using the Interactive Tree of Life (iTOL) with their respective metadata (Supplementary Table 7) ([5]).

References:


1. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv e-prints* 2013; arXiv:1303.3997
2. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv e-prints* 2012; arXiv:1207.3907
3. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30**: 1312–3.
4. Gardner SN, Slezak T, Hall, BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genomes. *Bioinformatics* 2015; **31**: 2877-8.
5. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acid Res* 2019; **47:** W256-9.

# PREFACE TO CHAPTER III

---

While completing the research described in chapter II, we were informed that there was currently no method available at Christian Medical College which could rapidly identify cases of BCG disease. We also found that deletion-based assays which use RDs may not be the ideal method to differentiate between MTBC subspecies because some strains had mutations in or around these regions. We determined that a rule-in approach may be more reliable than a rule-out approach-based assay. As described in chapter I, BCG disease is associated with high mortality rates and its treatment requires specific management that is unique from TB. A rapid and accurate diagnosis of BCG is important to improve patient outcomes. In this chapter, I describe a diagnostic assay I developed to identify BCG and its use in a small pilot study of BCG-suspected patients at Christian Medical College. For this diagnostic, I developed a real-time PCR assay which identifies BCG using BCG-specific SNPs and is capable of differentiating between early and late BCG strains in order to determine whether one BCG strain is contributing to more cases of BCG disease than another. Overall, we demonstrated that the assay was an accurate and robust method of rapidly identifying BCG and we identified BCG disease in 10/19 suspected cases at Christian Medical College.

# CHAPTER III

## Development of a Multiplex Real-Time PCR Assay for *Mycobacterium bovis* BCG and Validation in a Clinical Laboratory

---

Published Manuscript - Microbiology Spectrum, Volume 9, Number 2, e0109821, Sept 2021

Shannon C Duffy[1,2], Manigandan Venkatesan[3], Shubhada Chothe[4], Indira Poojary[5], Valsan Philip Verghese[6], Vivek Kapur[4,5], Marcel A Behr[1,2,7], Joy Sarojini Michael[3]

**Affiliations**

[1]Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada.

[2]McGill International TB Centre, Montreal, Quebec, Canada.

[3]Department of Clinical Microbiology, Christian Medical College, Vellore, Vellore, Tamil Nadu, India.

[4]Department of Animal Science, The Pennsylvania State University, University Park, Pennsylvania, USA.

[5]Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA.

[6]Department of Pediatric Infectious Diseases, Christian Medical College, Vellore, Vellore, Tamil, India.

[7]Department of Medicine, McGill University, Montreal, Quebec, Canada.

## 1. Abstract

*Mycobacterium bovis* bacillus Calmette-Guérin (BCG) is a live attenuated vaccine which can result in local or disseminated infection, most commonly in immunocompromised individuals. Differentiation of BCG from other members of the *Mycobacterium tuberculosis* complex (MTBC) is required to diagnose BCG disease, which requires specific management. Current methods for BCG diagnosis are based on mycobacterial culture and conventional PCR; the former is time-consuming and the latter often unavailable. Further, there are reports that certain BCG strains may be associated with a higher rate of adverse events. This study describes the development of a two-step multiplex real-time PCR assay which uses single nucleotide polymorphisms to detect BCG and identify early or late BCG strains. The assay has a limit of detection of 1 pg BCG boiled lysate DNA and was shown to detect BCG in both pure cultures and experimentally infected tissue. Its performance was assessed on 19 suspected BCG clinical isolates at Christian Medical College in Vellore, India, taken from January 2018 to August 2020. Of these 19 isolates, 10 were identified as BCG (6 early and 4 late strains), and 9 were identified as other MTBC members. Taken together, the results demonstrate the ability of this assay to identify and characterize BCG disease from cultures and infected tissue. The capacity to identify BCG may improve patient management, and the ability to discriminate between BCG strains may enable BCG vaccine pharmacovigilance.

**IMPORTANCE** Vaccination against tuberculosis with bacillus Calmette-Guérin (BCG) can lead to adverse events, including a rare but life-threatening complication of disseminated BCG. This complication often occurs in young children with immunodeficiencies and is associated with an ~60% mortality rate. A rapid method of reliably identifying BCG infection is important because BCG requires treatment unique to tuberculosis. BCG is resistant to the first-

line antituberculosis drug pyrazinamide. Additionally, diagnosis of BCG disease would lead to further investigation of a possible underlying immune condition. We have developed a diagnostic assay to identify BCG which improves upon previously published methods and can reliably identify BCG from bacterial culture or directly from infected tissue. This assay can also differentiate between strains of BCG, which have been suggested to be associated with different rates of adverse events. This assay was validated on 19 clinical isolates collected at Christian Medical College in Vellore, India.

**KEYWORDS:** real-time PCR identification, BCG disease, disseminated BCG, *Mycobacterium bovis* BCG, *Mycobacterium tuberculosis*

## 2. Introduction

*Mycobacterium bovis* bacillus Calmette-Guérin (BCG) is a live attenuated vaccine which is widely administered for protection against tuberculosis[1]. BCG vaccination can be associated with local adverse events such as injection site abscesses or lymphadenitis. It may also lead to systemic infection, causing osteomyelitis or disseminated BCG[2]. Disseminated disease is rare and often occurs in young children with primary immunodeficiency diseases (PIDs), such as severe combined immunodeficiency (SCID), chronic granulomatous disease (CGD), and Mendelian susceptibility to mycobacterial disease (MSMD)[3,4]. PIDs are a group of genetic disorders which cause impaired immunity and can result in increased susceptibility to infection and vaccine complications[4]. Children with HIV are also at an increased risk of BCG complications[5]. The estimated incidence of disseminated BCG is 0.06 to 3.4 cases per million, but nearly 1 per 100 in HIV-infected babies[3,5,6]. Although the World Health Organization recommends against giving BCG to patients with documented immunodeficiency conditions, in

India, as per the National Immunisation Programme, BCG is given soon after birth when these conditions have not yet been diagnosed[1,2].

To diagnose BCG disease, laboratory differentiation of mycobacterial isolates is essential, as BCG has similar growth kinetics and morphology to *M. tuberculosis* strains. The treatment of patients with BCG infection will differ from those with tuberculosis since BCG is resistant to one of the first-line antituberculosis drugs, pyrazinamide[7]. In addition, a diagnosis of BCG infection should lead to an immediate investigation into the patient's immune status[4,8]. A BCG diagnosis would also confirm the absence of a contagion, and a public health inquiry into the source of infection would not need to be conducted. Current methods of diagnosis of BCG often rely on the detection of regions of difference (RD) by PCR[5,9-13]. RD-based approaches apply a rule-out strategy, where the detection of RDs is used to eliminate a potential mycobacterial species identity. A rule-in approach using lineage-defining single nucleotide polymorphisms (SNPs) may provide a simplified method of positively identifying mycobacteria. Additionally, conventional gel-based PCR is an open system which requires the running of a gel and is therefore not a pragmatic option in many clinical labs. This method can also only target one region of the genome at a time. Real-time PCR is advantageous because it is a closed system, requires a single step, and is easily multiplexed.

Given the increased data on *M. tuberculosis* SNPs and the availability of real-time PCR reagents that enable multiplexed analyses, we developed a two-step real-time PCR assay to rapidly identify BCG from a bacterial culture or directly from infected tissue. The two steps were designed to (i) identify BCG from other members of the MTBC and (ii) differentiate between early and late strains of BCG, something of interest in India, where two different BCG strains (BCG Russia and BCG Danish) with potentially different rates of adverse events are in use[2, 14-16].

This assay was validated on 19 clinical isolates previously collected at Christian Medical College in Vellore, India, and our results suggest that the use of this assay may better inform patient management and enable BCG vaccine pharmacovigilance.

## 3. Results

*3.1 Specificity, reaction efficiency, and limit of detection.*

The specificity of the assay was tested on DNA from boiled lysate preparations of 26 isolates from 20 different mycobacterial species, including 8 BCG strains. The results for the step 1 and step 2 assays for all DNA correlated 100% as expected (Table S1 in the supplemental material). The expected step 1 amplification curves of several mycobacterial species are shown in Fig. 1. BCG was positive for all 3 probes (IS*1081*, *kdpD*, and *pncA*), *M. bovis* was positive for IS*1081* and *pncA*, *M. tuberculosis* was positive for IS*1081* only, and *M. abscessus* was negative for all 3 probes. The expected step 2 amplification curves of a BCG early and late strain are shown in Fig. 2. BCG Russia, an early strain, was positive for *crp*, and BCG Danish, a late strain, was positive for *mmaA3*. The reaction efficiencies of the probes were calculated by plotting the cycle threshold ($C_T$) values from a serial dilution of BCG Russia boiled lysate DNA for step 1 (Fig. S1 in the supplemental material) and both BCG Russia and BCG Danish boiled lysate DNA for step 2 (Fig. S2 in the supplemental material). The reaction efficiencies for all probes ranged between 95.72 and 103.99%, which is within the range of desired values of 90 to 110% (Table 1). The limit of detection for step 1 was 1 pg of BCG Russia boiled lysate DNA. For step 2, the limit of detection was also 1 pg, whether using BCG Russia or BCG Danish boiled lysate DNA (Table 1). One picogram of BCG Russia DNA or 1 pg of BCG Danish DNA is estimated to equal 212.3 or 210 genome copies, respectively.

*3.2 Assay reproducibility and robustness.*

The inter- and intra-assay reproducibility was evaluated by comparing the $C_T$ values of 3 separate plates with 5 technical replicates per plate. The $C_T$ values remained consistent across all replicates: the $C_T$ values remained under 2 $C_T$ of difference across all replicates using step 1 (Fig. S3A in the supplemental material) and under 3 $C_T$ of difference across all replicates using step 2 (Fig. S3B in the supplemental material). The ability of all probes to produce the expected amplification plot in the presence of excess nonspecific DNA was tested using 1 ng target DNA (either BCG Russia or BCG Danish as indicated) with the addition of 10 ng of DNA from *M. abscessus*. All probes amplified as expected despite the presence of nonspecific DNA (Fig. S4 in the supplemental material).

*3.3 Identification from mice.*

To determine whether the assay could be used to detect BCG directly from tissue, DNA was extracted from the spleen, liver, and lungs of 3 C57BL/6 mice which were intravenously infected with *M. bovis* BCG Russia. In all samples, the step 1 assay correctly identified the presence of BCG DNA (Fig. 3A), and the step 2 assay correctly identified the presence of an early BCG strain (Fig. 3B).

*3.4 Validation with clinical samples.*

The assay was validated on a set of 19 isolates previously collected at Christian Medical College in Vellore, India (Table 2). Samples were taken from pus (6), lymph node (4), gastric aspirate (4), thigh sinus tract (1), bone (1), pleural tissue (1), colonic ulcer (1), and mastoid tissue (1). The samples came from patients in the following 5 states in India: Tamil Nadu (8), Andhra

Pradesh (5), Karnataka (2), West Bengal (2), and Odisha (1). One sample was taken from a patient from Bangladesh. DNA was extracted from a culture on the mycobacteria growth indicator tube (MGIT) (13) or Lowenstein-Jensen (LJ) medium (5) or directly from a pus swab (1) (Table S2 in the supplemental material). Of the 19 isolates included, 10 (52.6%) were confirmed to be BCG. Four (40%) of the BCG isolates were from female patients. The median age of the BCG-infected patients was 6.5 months (range 3 to 12 months). In total, 4 of the BCG isolates came from patients diagnosed with immune conditions: 2 BCG isolates were from patients with SCID and 2 from those with MSMD (Table S2 in the supplemental material). Of the 10 BCG isolates, 6 (60%) were identified as a BCG early strain, and 4 (40%) were identified as a BCG late strain. Sample amplification plots of 2 clinical isolates, 1 early strain and 1 late strain, are provided in Fig. S5 in the supplemental material. The other 9 isolates included in this study were identified as MTBC by the step 1 assay. Two (~22.2%) of the MTBC isolates came from female patient samples. Of the 9 MTBC isolates, 7 (~77.8%) came from patients between 1 and 2 years old (the maximum age of inclusion).

## 4. Discussion

A rapid method of differentiating BCG from other mycobacterial species is essential to providing optimal care for BCG infection, particularly in cases of disseminated BCG, which is associated with a mortality rate of ~60%[6,8,17-20]. A method of determining whether an early or late strain is causing a case of disseminated BCG may also provide an opportunity for pharmacosurveillance, to determine whether certain strains are linked to a higher rate of adverse events[2,14]. Here, we have developed a two-step multiplex real-time PCR assay to rapidly identify BCG isolates from other mycobacterial species and differentiate between BCG early and late

strains. This assay was shown to detect BCG from a culture, directly from the tissue of infected mice, or from a patient pus swab in a case where a culture was unavailable. The ability to identify BCG directly from a clinical sample circumvents time-consuming mycobacterial culture to inform patient treatment. This assay was tested on a small validation set of 19 clinical isolates previously collected at Christian Medical College in Vellore, India, where two different BCG strains are in use: BCG Russia (an early strain) and BCG Danish (a late strain). There was no assay for BCG detection in use at Christian Medical College prior to conducting this study. If clinicians suspected a patient was infected with BCG based on their age, disease presentation, and immune status, they were treated accordingly. Additionally, we are not aware of any real-time PCR assays which detect BCG using SNPs, nor are we aware of any which can discriminate between BCG strains. Of the 19 isolates tested, 10 were identified as BCG, 6 of which were early strains.

The clinical management of a patient diagnosed with BCG disease will differ compared to a patient with *M. tuberculosis*. Unlike *M. tuberculosis*, BCG infection would not trigger the need for contact tracing to identify the source of infection and perhaps other infected persons. Confirmation of BCG infection may also affect which drugs would be included in the patient treatment regimen. Since BCG is derived from *M. bovis*, all strains are naturally resistant to pyrazinamide, a first-line antituberculosis drug[7]. The assay described in this study confirms pyrazinamide resistance using the *pncA* probe. In addition, the *mmaA3* probe used to identify late strains detects a SNP which is associated with a loss of methoxymycolate production and is suggested to cause low levels of isoniazid resistance in late strains[7,21]. A diagnosis of BCG infection should also lead to further investigation of possible underlying immune conditions, which may not have been recognized at the time of disease presentation[4,8]. In a previous study of

349 SCID patients, 118 developed disseminated BCG disease, a 33,000-fold increase over the general population[17]. In this study, of the 10 isolates identified as BCG, 2 came from patients confirmed to have SCID and 2 came from patients confirmed to have MSMD (Table S2 in the supplemental material). Although BCG vaccination is contraindicated in people with these conditions, the Universal Immunisation Programme (UIP) recommends BCG vaccination in India at birth or as early as possible (https://www.nhp.gov.in/universal-immunisation-programme_pg), so these rare immune conditions are usually not recognized until after vaccination.

There are 14 different BCG strains, which have evolved with considerable genetic and phenotypic heterogeneity among them. BCG strains differ from one another by genetic deletions, SNPs, and duplications[15]. It has been suggested that some BCG strains may be associated with a higher rate of adverse events[2,14-16]. This possibility is of particular interest in India, where both an early and a late BCG strain are currently in use. In Finland, BCG osteitis rates were as high as 36.9 per 100,000 while using the early strain BCG Sweden. Upon switching to the late strain BCG Glaxo in 1978, that rate decreased to 6.4 per 100,000[2,14]. Conversely, in phase 1 of a randomized control trial in which BCG Danish was compared with BCG Russia, lymphadenitis was reported in 6 out of 582 patients given BCG Danish and 1 out of 575 patients given BCG Russia[16]. In this study, 6 out of 10 BCG isolates were early strains, presumably BCG Russia. The other 4 were identified as late strains, presumably BCG Danish. The sample size of this study is too small to draw conclusions concerning an association between the BCG strain and adverse events; however, screening is ongoing, and this question may be revisited with the established assay as the total sample size rises.

To conclude, we have developed a rapid method of reliably identifying BCG early and late strains which improves upon previously described assays. This assay may be used to inform patient treatment without the time constraints of mycobacterial culture. The ability to differentiate between early and late strains may also inform BCG vaccine pharmacovigilance in a country in which two vaccine strains are currently in use.

## 5. Materials and methods

*5.1 Assay design.*

A two-step real-time PCR protocol was designed to identify BCG isolates from clinical samples. Step 1 identifies BCG from other MTBC subspecies. It is a three-probe multiplex real-time PCR assay which detects the MTBC insertion element IS*1081*, the BCG-specific SNP *kdpD* c247t, and the pyrazinamide resistance SNP *pncA* c169g[7,22]. Step 2 differentiates BCG isolates as early or late strains. It is a two-probe real-time PCR assay which detects the BCG early strain SNP in *crp* c140t and the BCG late strain SNP in *mmaA3* g293a[7]. The probes and primers were designed using Primer Express v. 3.0 (Applied Biosystems, Foster City, CA, USA). All primer and probe sequences are listed in Table 1. The IS*1081*, *kdpD*, *crp,* and *mmaA3* probes are TaqMan MGB probes (Applied Biosystems), and the *pncA* probe is a PrimeTime BHQ probe (Integrated DNA Technologies, Coralville, IA, USA). All primers were purchased from Invitrogen (Carlsbad, CA, USA).

*5.2 Real-time PCR.*

All real-time PCRs were performed in a 20-µl reaction mixture containing 1 µl DNA at a concentration of 10 ng/µl unless otherwise specified. Step 1 reaction mixtures were prepared

with 10 µl TaqMan multiplex master mix (Applied Biosystems), 100 nM IS*1081* forward and reverse primers, 300 nM forward and reverse *kdpD* primers, 900 nM *pncA* forward and reverse primers, 125 nM IS*1081* probe, 125 nM *kdpD* probe, and 500 nM *pncA* probe. Step 2 real-time PCR mixtures were prepared using 10 µl TaqMan multiplex master mix, 300 nM *crp* and *mmaA3* forward and reverse primers, 250 nM *crp* and *mmaA3* probes, and 1 µl dimethyl sulfoxide (Sigma-Aldrich, St. Louis, MO, USA). Real-time PCRs were performed using a CFX96 touch real-time PCR detection system (Bio-Rad, Hercules, CA, USA) under the following thermocycling conditions: one cycle at 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 63°C for 1 min. A summary of how results are interpreted is found in Fig. S6 in the supplemental material.

*5.3 Specificity.*

The specificity of the two-step assay was evaluated with the following bacterial species: *M. tuberculosis*, *M. bovis*, *M. orygis*, *M. africanum*, *M. caprae* (2), *M. microti*, *M. avium* subsp. *paratuberculosis* (4), *M. avium* subsp. *avium* (3), *M. avium* subsp. *hominissuis*, *M. kansasii*, *M. abscessus*, *M. smegmatis*, BCG Russia, BCG Moreau, BCG Sweden, BCG Japan, BCG Danish, BCG Pasteur, BCG Prague, and BCG Glaxo. If more than one isolate of each mycobacterium was tested, the number of isolates is indicated in parentheses. Cultures of each bacterial species were grown in 7H9 Middlebrook medium (Becton, Dickinson, Franklin Lakes, NJ, USA), ADC (Becton, Dickinson), 0.2% glycerol (Sigma-Aldrich), and 0.05% Tween 80 (Sigma-Aldrich) to approximately an optical density at 600 nm ($OD_{600}$) of 0.5 to 0.8. The medium was supplemented with 1 µg/ml ferric mycobactin J (Allied Monitor, Fayette, MO, USA) for *M. avium* subsp. *paratuberculosis* cultures. One milliliter of culture was added to a

screw cap microcentrifuge tube and centrifuged for 5 min at 13,000 rpm. The supernatant was discarded, and the pellet was resuspended in 200 µl molecular grade water. The tube was boiled at 95°C for 40 min and then centrifuged again for 5 min at 13,000 rpm. The supernatant was transferred to a new tube for real-time PCR use. Boiled lysate DNA was used in this study to demonstrate the assay robustness and evaluate the assay using a DNA preparation that would be a pragmatic and convenient option for many labs. The prepared DNA was stored at −20°C until ready for testing by real-time PCR.

*5.4 Reaction efficiencies, limits of detection, reproducibility, and robustness.*

The reaction efficiencies of all probes were determined by plotting the $C_T$ values obtained from a standard curve of 10-fold dilutions of DNA isolated from BCG Russia or BCG Danish. The reaction efficiency was calculated using the formula $E = (10^{-1/m} - 1) \times 100\%$, where $m$ is the slope of the line. The limit of detection of all probes was determined by the minimum amount of DNA to yield a positive amplification result from the above standard curves. The inter- and intra-assay reproducibility was evaluated by comparing the $C_T$ values of the step 1 and step 2 assays completed on 3 separate plates with 5 technical replicates per plate. The ability of the assay to function in the presence of excess nonspecific DNA was tested by performing real-time PCR with 1 ng of target DNA (BCG Russia or BCG Danish) and 10 ng of *M. abscessus* DNA.

*5.5 Mouse infection.*

The ability of the assay to identify BCG directly from tissue was determined using infected mice. Three C57BL/6 mice (Jackson Laboratories, Bar Harbor, MA, USA) were intravenously infected with $2.3 \times 10^8$ CFU BCG Russia. After 3 weeks, the mice were sacrificed,

and spleen, liver, and lung samples were collected. The organs were homogenized, and infection was confirmed by serial dilution and plating onto Middlebrook 7H10 agar (Becton, Dickinson), OADC enrichment (Becton, Dickinson), 0.5% glycerol (Sigma-Aldrich), and PANTA antibiotic mixture (Becton, Dickinson). DNA was extracted from organ homogenates as previously described[23]. DNA was then quantified and analyzed by the two-step real-time PCR assay.

*5.6 Clinical samples.*

The assay was validated using 19 clinical isolates previously collected at Christian Medical College, Vellore, India, from January 2018 to August 2020. This study obtained approval and ethical clearance from the Institutional Review Board (IRB) of Christian Medical College, Vellore (IRB approval number 11725, dated 19 December 2018). An isolate is defined as a pure microbe cultured from an individual patient. A sample is defined as the patient specimen provided for culture. Isolates were collected from patients presenting with suspected tuberculosis who were 2 years of age or younger or by request of the clinician due to suspected BCG disease. Clinicians suspected BCG involvement with a BCG site abscess or axillary lymphadenitis, or when disseminated MTBC was isolated in an infant with an immunodeficiency disorder such as SCID or MSMD. DNA was extracted from positive mycobacteria growth indicator tube (MGIT) or Lowenstein-Jensen (LJ) cultures as available. In one sample, a culture was not available, and DNA was extracted directly from a pus swab. DNA was prepared by boiling for samples collected before March 2019 or by using the QIAamp DNA minikit (Qiagen, Hilden, Germany) for samples collected after March 2019. All DNA was analyzed by the step 1 real-time PCR assay. If an isolate was identified as BCG in step 1, it was analyzed by step 2.

## 6. Acknowledgements

This research was supported by the Canadian Institutes for Health Research (grant FDN148362 to M.A.B.) and the Bill & Melinda Gates Foundation (OPP1176950 to V.K., M.A.B., and J.S.M.). The funding agency had no role in study design, data collection, analysis, interpretation, or the decision to submit this work for publication.

We thank Fiona McIntosh and Sarah Danchuk for technical support with the assay development.

## 7. References

1. World Health Organization. 2020. *Global tuberculosis report 2020.* World Health Organization, Geneva, Switzerland.
2. World Health Organization. 2012. Information sheet observed rate of vaccine reactions Bacille Calmette-Guérin (BCG) vaccine. World Health Organization, Geneva, Switzerland.
3. Lee PPW. 2015. Disseminated Bacillus Calmette-Guérin and susceptibility to mycobacterial infection—implications on Bacillus Calmette-Guérin vaccinations. *Ann Acad Med Singap* 44:297–301.
4. Norouzi S, Aghamohammadi A, Mamishi S, Rosenzweig SD, Rezaei N. 2012. Bacillus Calmette-Guérin (BCG) complications associated with primary immunodeficiency diseases. *J Infect* 64:543–554. doi: 10.1016/j.jinf.2012.03.012.
5. Hesseling AC, Johnson LF, Jaspan H, Cotton MF, Whitelaw A, Schaaf HS, Fine PEM, Eley BS, Marais BJ, Nuttall J, Beyers N, Godfrey-Faussett P. 2009. Disseminated bacille

Calmette-Guérin disease in HIV-infected South African infants. *Bull World Health Org* 87:505–511. doi: 10.2471/BLT.08.055657.

6. Bernatowska EA, Wolska-Kusnierz B, Pac M, Kurenko-Deptuch M, Zwolska Z, Casanova J-L, Piatosa B, van Dongen J, Roszkowski K, Mikoluc B, Klaudel-Dreszler M, Liberek A. 2007. Disseminated bacillus Calmette-Guérin infection and immunodeficiency. *Emerg Infect Dis* 13:799–801. doi: 10.3201/eid1305.060865.
7. Abdallah AM, Hill-Cawthorne GA, Otto TD, Coll F, Guerra-Assunção JA, Gao G, Naeem R, Ansari H, Malas TB, Adroub SA, Verboom T, Ummels R, Zhang H, Panigrahi AK, McNerney R, Brosch R, Clark TG, Behr MA, Bitter W, Pain A. 2015. Genomic expression catalogue of a global collection of BCG vaccine strains show evidence for highly diverged metabolic and cell-wall adaptations. *Sci Rep* 5:15443. doi: 10.1038/srep15443.
8. Ong RYL, Chan S-WB, Chew SJ, Liew WK, Thoon KC, Chong C-Y, Yung CF, Sng L-H, Tan AM, Bhattacharyya R, Jamuar SS, Lim JY, Li J, Nadua KD, Kam K-Q, Tan NW-H. 2020. Disseminated Bacillus-Calmette-Guérin infections and primary immunodeficiency disorders in Singapore: a single center 15-year retrospective review. *Int J Infect Dis* 97:117–125. doi: 10.1016/j.ijid.2020.05.117.
9. Talbot EA, Williams DL, Frothingham R. 1997. PCR identification of *Mycobacterium bovis* BCG. *J Clin Microbiol* 35:566–569. doi: 10.1128/jcm.35.3.566-569.1997.
10. Pinsky BA, Banaei N. 2008. Multiplex real-time PCR assay for rapid identification of *Mycobacterium tuberculosis* complex members to the species level. *J Clin Microbiol* 46:2241–2246. doi: 10.1128/JCM.00347-08.

11. Monajemzadeh M, Shahsiah R, Zarei A, Alamooti AA, Mahjoub F, Mamishi S, Khotai G, Pazira R, Eram N. 2010. Frequency of Bacille Calmette-Guérin (BCG) and *Mycobacterium tuberculosis* in tissue biopsy specimens of children vaccinated with BCG. *Am J Clin Pathol* 133:102–106. doi: 10.1309/AJCPXLZPRHX9L0YG.
12. Reddington K, O'Grady J, Dorai-Raj S, Niemann S, van Soolingen D, Barry T. 2011. A novel multiplex real-time PCR for the identification of mycobacteria associated with zoonotic tuberculosis. *PLoS One* 6:e23481. doi: 10.1371/journal.pone.0023481.
13. Teo JWP, Cheng JWS, Jureen R, Lin RTP. 2013. Clinical utility of RD1, RD9, and *hsp65* based PCR assay for the identification of BCG in vaccinated children. *BMC Res Notes* 6:434. doi: 10.1186/1756-0500-6-434.
14. Kröger L, Korppi M, Brander E, Kröger H, Wasz-Höckert O, Backman A, Rapola J, Launiala K, Katila M-L. 1995. Osteitis caused by Bacille Calmette-Guérin vaccination: a retrospective analysis of 222 cases. *J Infect Dis* 172:574–576. doi: 10.1093/infdis/172.2.574.
15. Abdallah AM, Behr MA. 2017. Evolution and strain variation in BCG, p 155–169. *In* Gagneaux S (ed), *Strain variation in the Mycobacterium tuberculosis complex: its role in biology, epidemiology, and control.* Springer Nature, Cham, Switzerland.
16. Schaltz-Buchholzer F, Berendsen M, Roth A, Jensen KJ, Bjerregaard-Andersen M, Sørensen MK, Monteiro I, Aaby P, Benn CS. 2020. BCG skin reactions by 2 months of age are associated with better survival in infancy: a prospective observational study from Guinea-Bissau. *BMJ Glob Health* 5:e002993. doi: 10.1136/bmjgh-2020-002993.
17. Marciano BE, Huang C-Y, Joshi G, Rezaei N, Carvalho BC, Allwood Z, Ikinciogullari A, Reda SM, Gennery A, Thon V, Espinosa-Rosales F, Al-Herz W, Porras O, Shcherbina A,

Szaflarska A, Kiliç Ş, Franco JL, Gómez Raccio AC, Roxo P, Jr, Esteves I, Galal N, Grumach AS, Al-Tamemi S, Yildiran A, Orellana JC, Yamada M, Morio T, Liberatore D, Ohtsuka Y, Lau Y-L, Nishikomori R, Torres-Lozano C, Mazzucchelli JTL, Vilela MMS, Tavares FS, Cunha L, Pinto JA, Espinosa-Padilla SE, Hernandez-Nieto L, Elfeky RA, Ariga T, Toshio H, Dogu F, Cipe F, Formankova R, Nuñez-Nuñez ME, Bezrodnik L, Marques JG, Pereira MI, Listello V, et al.. 2014. BCG vaccination in patients with severe combined immunodeficiency: complications, risks, and vaccination policies. *J Allergy Clin Immunol* 133:1134–1141. doi: 10.1016/j.jaci.2014.02.028.

18. Aelami MH, Alborzi A, Pouladfar G, Geramizadeh B, Pourabbas B, Mardaneh J. 2015. Post-vaccination disseminated Bacillus Calmette Guerin infection among children in southern Iran. *Jundishapur J Microbiol* 8:e25663. doi: 10.5812/jjm.25663.
19. Amanati A, Pouladfar G, Kadivar MR, Dashti AS, Jafarpour Z, Haghpanah S, Alborzi A. 2017. A 25-year surveillance of disseminated Bacillus Calmette-Guérin disease treatment in children in southern Iran. *Medicine (Baltimore)* 96:e9035. doi: 10.1097/MD.0000000000009035.
20. Sadeghi-Shabestari M, Rezaei N. 2009. Disseminated bacille Calmette-Guérin in Iranian children with severe combined immunodeficiency. *Int J Infect Dis* 13:e420–e423. doi: 10.1016/j.ijid.2009.02.008.
21. Behr MA, Schroeder BG, Brinkman JN, Slayden RA, Barry CE, III.. 2000. A point mutation in the *mma3* gene is responsible for impaired methoxymycolic acid production in *Mycobacterium bovis* BCG strains obtained after 1927. *J Bacteriol* 182:3394–3399. doi: 10.1128/JB.182.12.3394-3399.2000.

22. Pelayo MCG, Uplekar S, Keniry A, Lopez PM, Garnier T, Garcia JN, Boschiroli L, Zhou X, Parkhill J, Smith N, Hewinson RG, Cole ST, Gordon SV. 2009. A comprehensive survey of single nucleotide polymorphisms (SNPs) across *Mycobacterium bovis* strains and *M. bovis* BCG vaccine strains refines the genealogy and defines a minimal set of SNPs that separate virulent *M. bovis* strains and *M. bovis* BCG strains. *Infect Immun* 77:2230–2238. doi: 10.1128/IAI.01099-08.

23. van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. 1991. Occurrence and stability of insertion sequences in Mycobacterium tuberculosis complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol* 29:2578–2586. doi: 10.1128/jcm.29.11.2578-2586.1991.

# 8. Figures

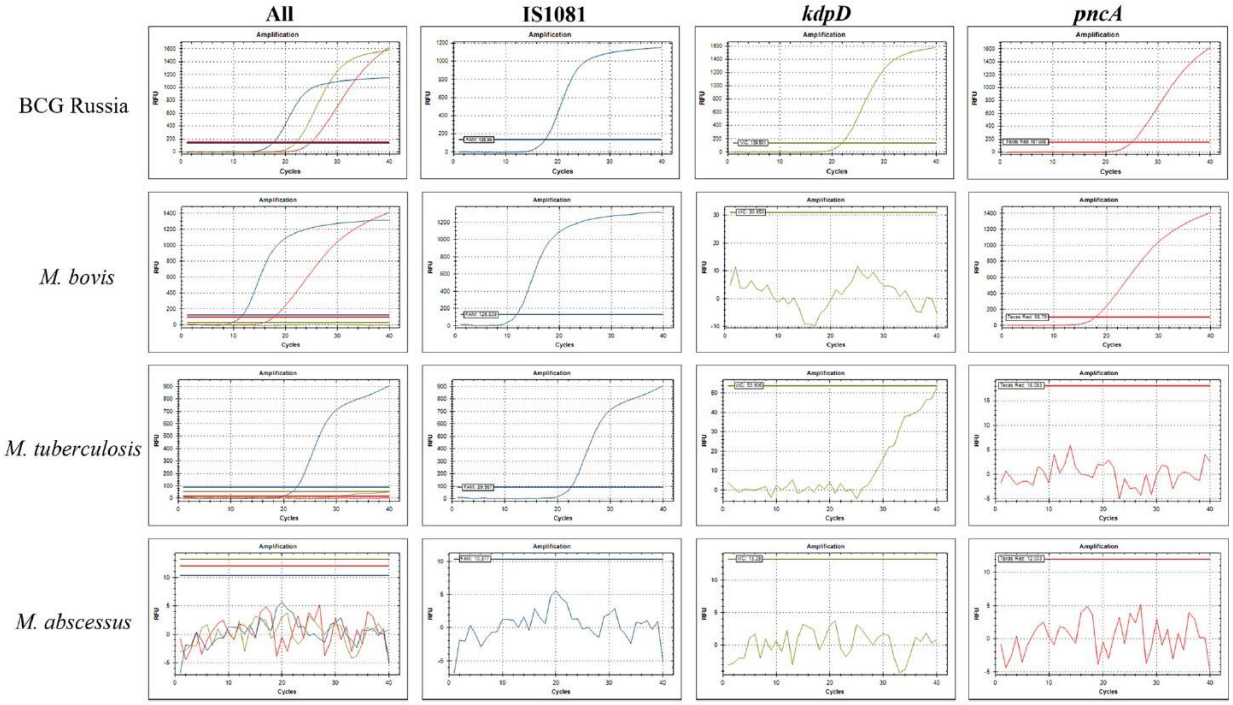



**Figure 1** Step 1 assay specificity

A sample amplification curve of several mycobacteria with different probe amplification patterns using 10 ng of DNA: an MTBC subspecies *M. tuberculosis*, pyrazinamide-resistant *M. bovis*, a BCG strain BCG Russia, and a nontuberculous mycobacterium *M. abscessus*.

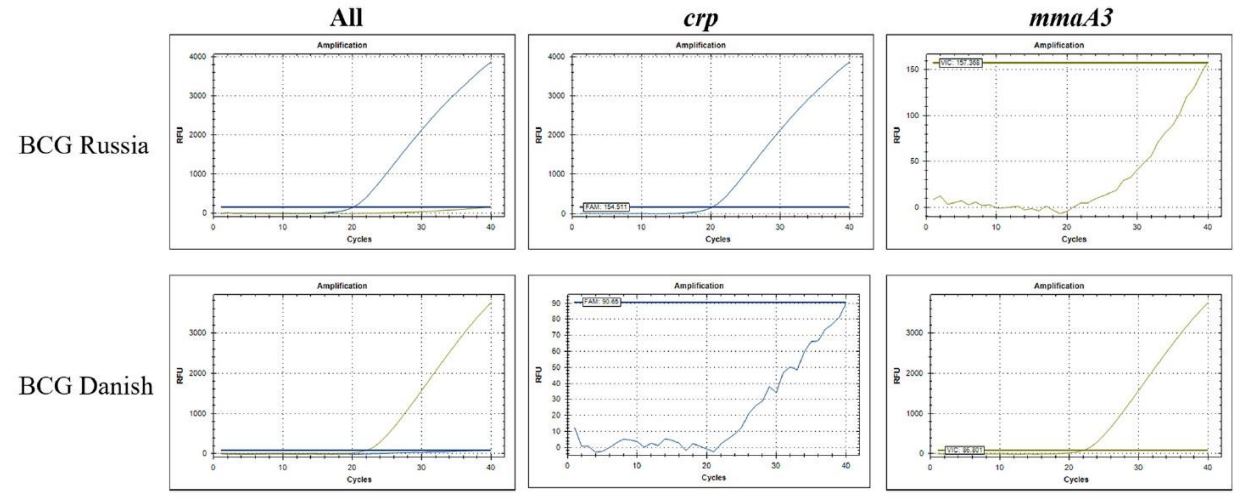



**Figure 2** Step 2 assay specificity

A sample amplification curve for the early strain BCG Russia and the late strain BCG Danish using 10 ng DNA.

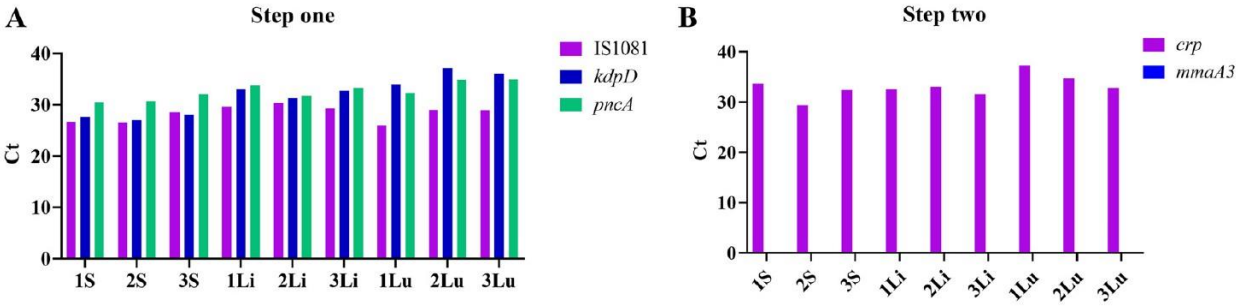



**Figure 3** Real-time PCR results of DNA extracted directly from the tissue of BCG-infected mice

The $C_T$ values of DNA analyzed using the step 1 (A) and step 2 (B) assays. The number indicates the mouse, S indicates DNA extracted from the spleen, Li indicates DNA extracted from the liver, and Lu indicates DNA extracted from the lungs.

## 9. Tables

| Assay | Name | Sequence (5′–3′) | Dye | Quencher | $R^2$ | Reaction efficiency (%) | Limit of detection (pg) |
|---|---|---|---|---|---|---|---|
| Step 1 (identify BCG) | IS1081_Probe | CTGAAGCCGACGCCCTGTGC | FAM | MGBNFQ | 0.9971 | 95.72 | 1 |
| | IS1081_Forward | GGCTGCTCTCGACGTTCATC | | | | | |
| | IS1081_Reverse | CGCTGATTGGACCGCTCAT | | | | | |
| | *kdpD*_Probe | ATGATCTCGCCGCGC | VIC | MGBNFQ | 0.9997 | 103.99 | 1 |
| | *kdpD*_Forward | GCAACAAGACCGCGAAACTG | | | | | |
| | *kdpD*_Reverse | GCAGTACTGCCTCCACATCGA | | | | | |
| | *pncA*_Probe | TGACGACTTCTCCGGCA | TEXAS RED | BHQ | 0.9945 | 101.08 | 1 |
| | *pncA*_Forward | AGCGGCGGACTACCATCAC | | | | | |
| | *pncA*_Reverse | TGTCCAGACTGGGATGGAAGT | | | | | |
| Step 2 (identify strain) | *crp*_Probe | TCGGCTGTACATCAT | FAM | MGBNFQ | 0.9996 | 96.92 | 1 |
| | *crp*_Forward | CCGTGGACACACGGTCTTC | | | | | |
| | *crp*_Reverse | GGCGACCGATCTTGACCTT | | | | | |
| | *mmaA3*_Probe | TCGTCGACTTGACC | VIC | MGBNFQ | 0.9908 | 100.04 | 1 |
| | *mmaA3*_Forward | CGAGCGCTACGACGTCAA | | | | | |
| | *mmaA3*_Reverse | CTGGCAGTAGGCGTGCTGAT | | | | | |

**Table 1** Sequences, reaction efficiencies, and limits of detection of probes and primers

| Isolate no. | Age (mo) | Sex | Location | Sample type | Assay result |
|---|---|---|---|---|---|
| 1 | 3 | M | Karnataka | Abscess pus | BCG late strain |
| 2 | 3 | F | Andhra Pradesh | Gastric aspirate | MTBC |
| 3 | 4 | M | Karnataka | Lymph node pus | BCG late strain |
| 4 | 5 | F | Tamil Nadu | Lymph node | BCG early strain |
| 5 | 5 | F | Tamil Nadu | Lymph node pus | BCG late strain |
| 6 | 6 | M | Tamil Nadu | Lymph node | MTBC |
| 7 | 6 | M | Andhra Pradesh | Lymph node | BCG early strain |
| 8 | 7 | M | Andhra Pradesh | Thigh sinus tract biopsy specimen | BCG early strain |
| 9 | 9 | F | Tamil Nadu | Lymph node | BCG early strain |
| 10 | 9 | F | Tamil Nadu | Lymph node pus | BCG late strain |
| 11 | 10 | M | West Bengal | Talar bone | BCG early strain |
| 12 | 12 | M | Tamil Nadu | BCG site pus | BCG early strain |
| 13 | 12 | F | Odisha | Gastric aspirate | MTBC |
| 14 | 12 | F | Tamil Nadu | Lymph node | MTBC |
| 15 | 18 | M | Andhra Pradesh | Gastric aspirate | MTBC |
| 16 | 19 | M | Andhra Pradesh | Pleural tissue | MTBC |
| 17 | 20 | M | Bangladesh | Colonic ulcer biopsy specimen | MTBC |
| 18 | 22 | M | West Bengal | Gastric aspirate | MTBC |
| 19 | 24 | M | Tamil Nadu | Mastoid tissue | MTBC |

**Table 2** Patient information and assay results of 19 suspected BCG clinical isolates

## 10. Supplementary figures

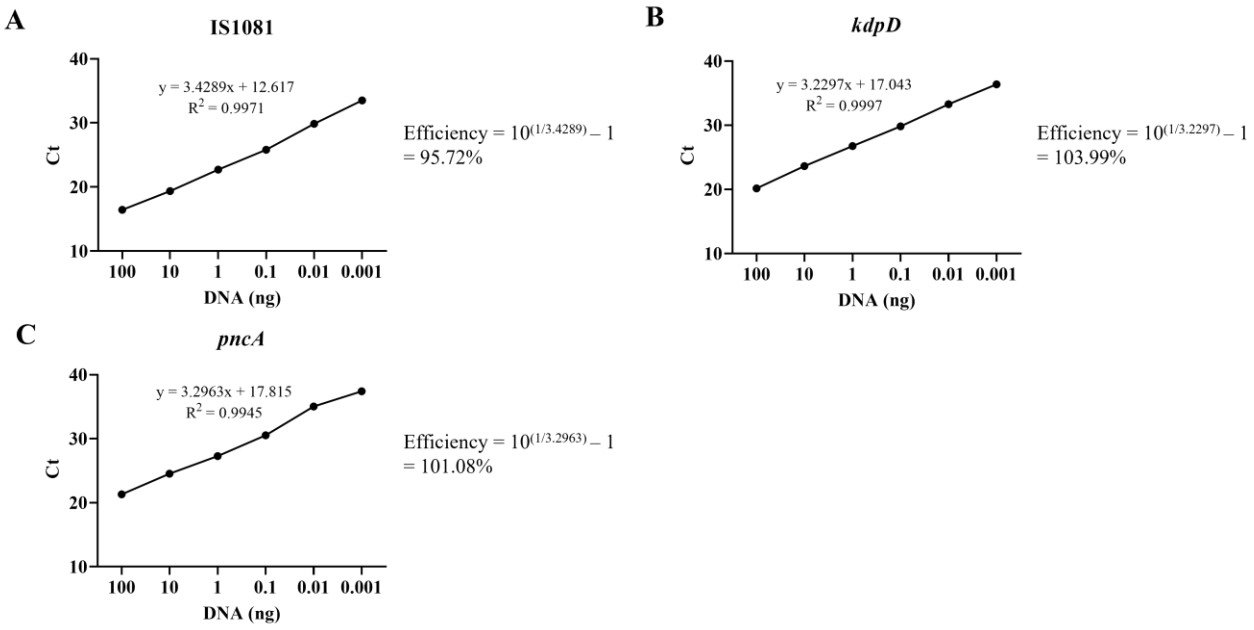



**Figure S1** Reaction efficiency calculations of step one probes

(A-C) A plot of the Ct values of a 10-fold serial dilution of BCG Russia DNA and calculation of the reaction efficiency for the IS1081 probe (A), the *kdpD* probe (B) and the *pncA* probe (C).

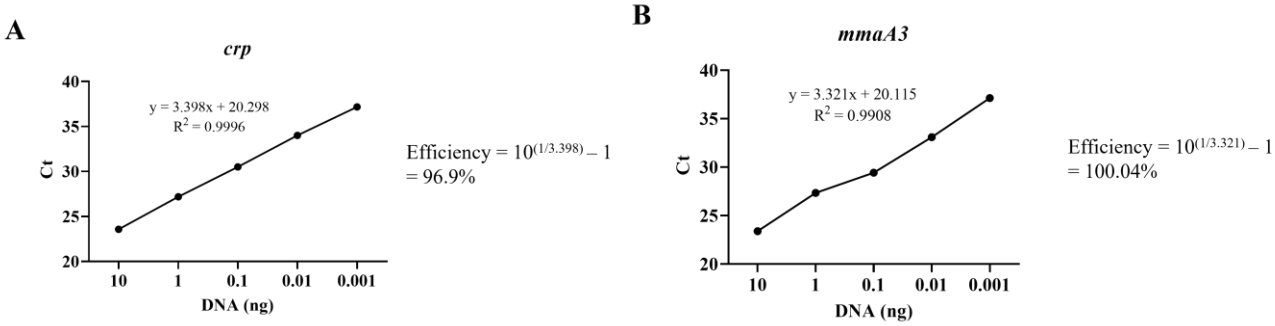



**Figure S2** Reaction efficiency calculations of step two probes

(A) A plot of the Ct values of a 10-fold serial dilution of BCG Russia DNA and calculation of the reaction efficiency of the *crp* probe. (B) A plot of the Ct values of a 10-fold serial dilution of BCG Danish DNA and calculation of the reaction efficiency of the *mmaA3* probe.

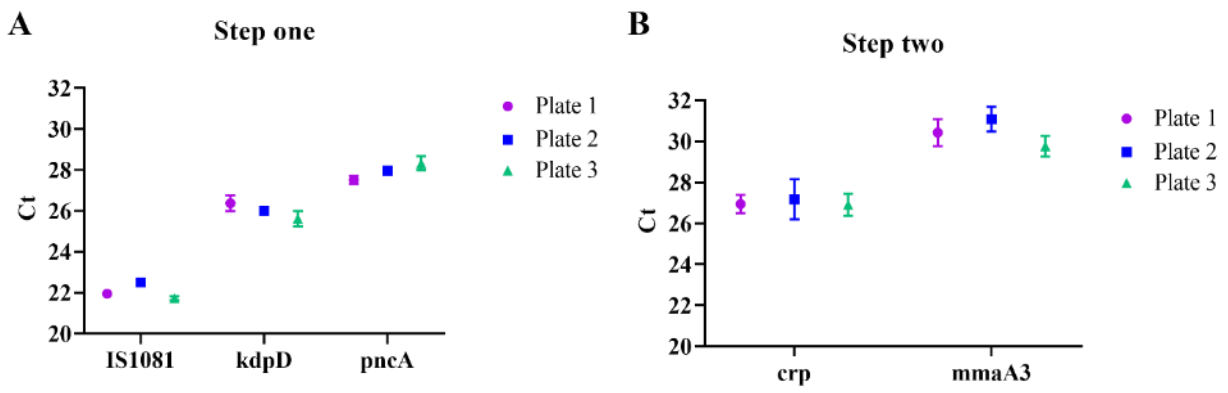



**Figure S3** Inter- and Intra-assay reproducibility

(A, B) The Ct values were compared using 1 ng of BCG Russia DNA with step 1 (A) or step 2 (B) assay. The inter-assay reproducibility was evaluated by comparing the Ct values across five technical replicates. The intra-assay reproducibility was evaluated by comparing the Ct values across three assay replicates.

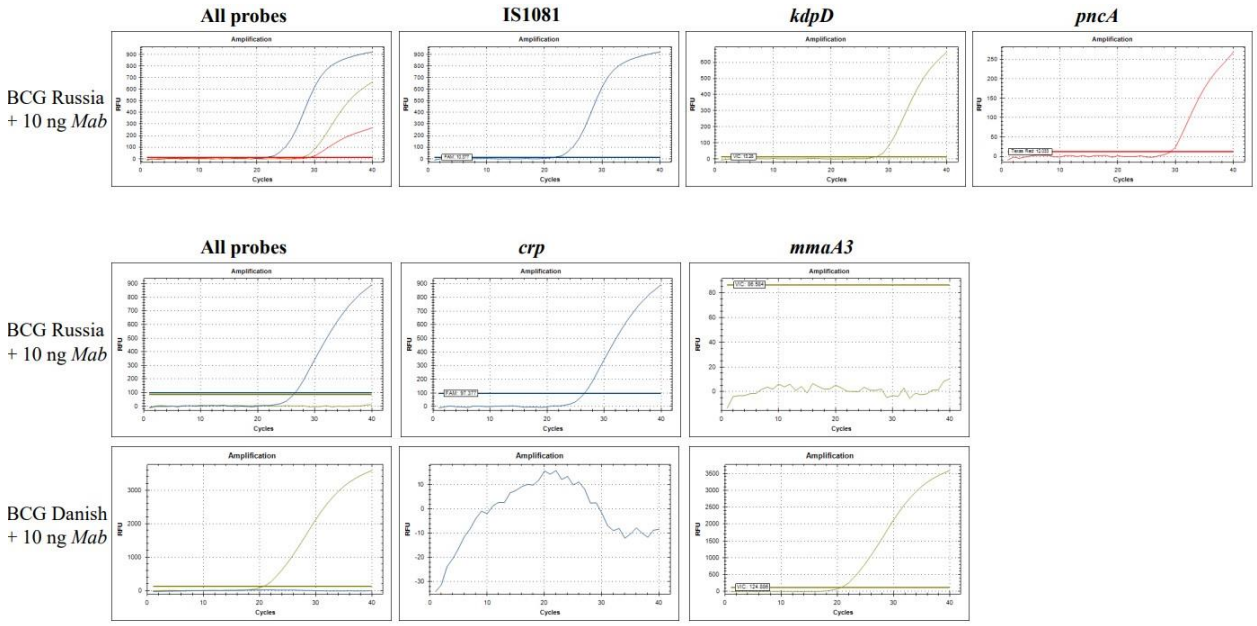



**Figure S4** Assay performance in presence of 10ng of excess non-specific DNA

The ability of the assay to produce the expected amplification plot in the presence of excess non-specific DNA was evaluated. (A) A total of 10 ng of DNA from *M. abscessus* (*Mab*) was added to a step one reaction mixture containing 1 ng of BCG Russia DNA. (B) A total of 10 ng of DNA from *Mab* was added to a step one reaction mixture containing 1 ng of either BCG Russia or BCG Danish DNA.

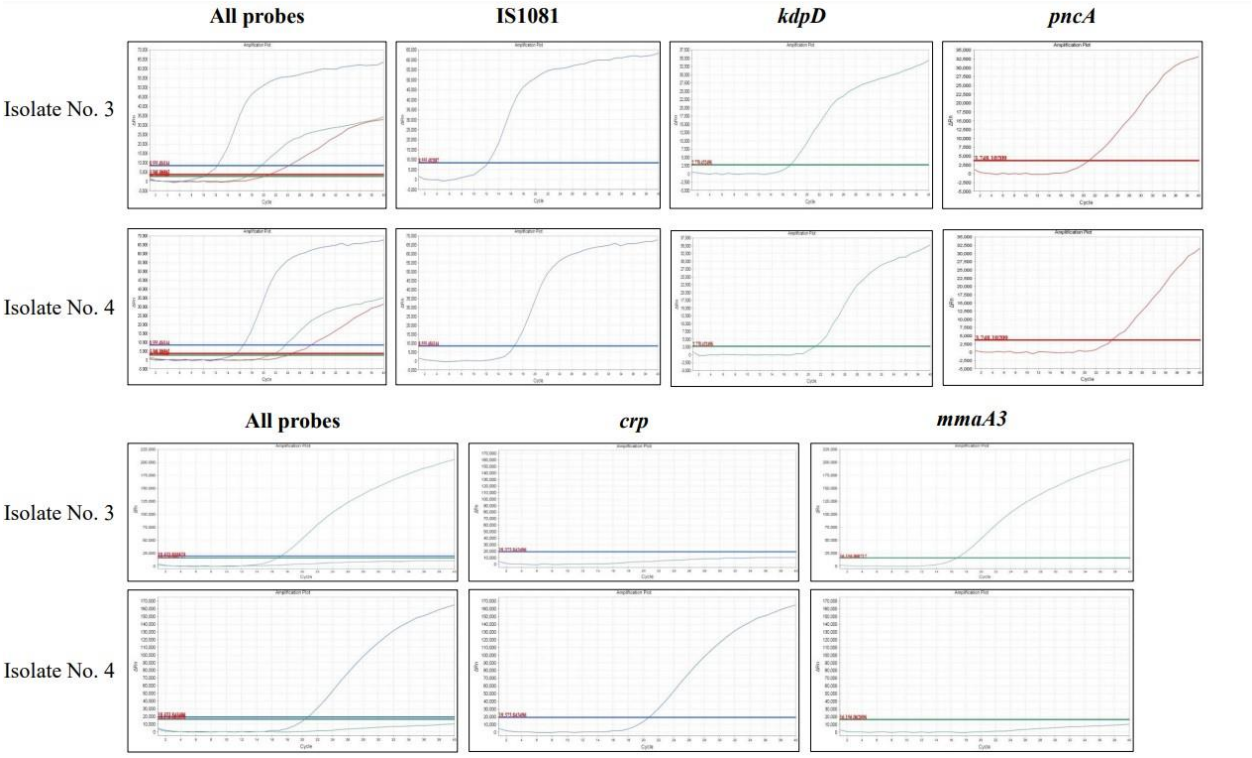



**Figure S5** Clinical isolates amplification plots

(A, B) Example amplification plots of two clinical isolates using the step one (A) and step two (B) assay. Isolate number 3 was identified as a BCG late strain and isolate number 4 was identified as a BCG early strain.

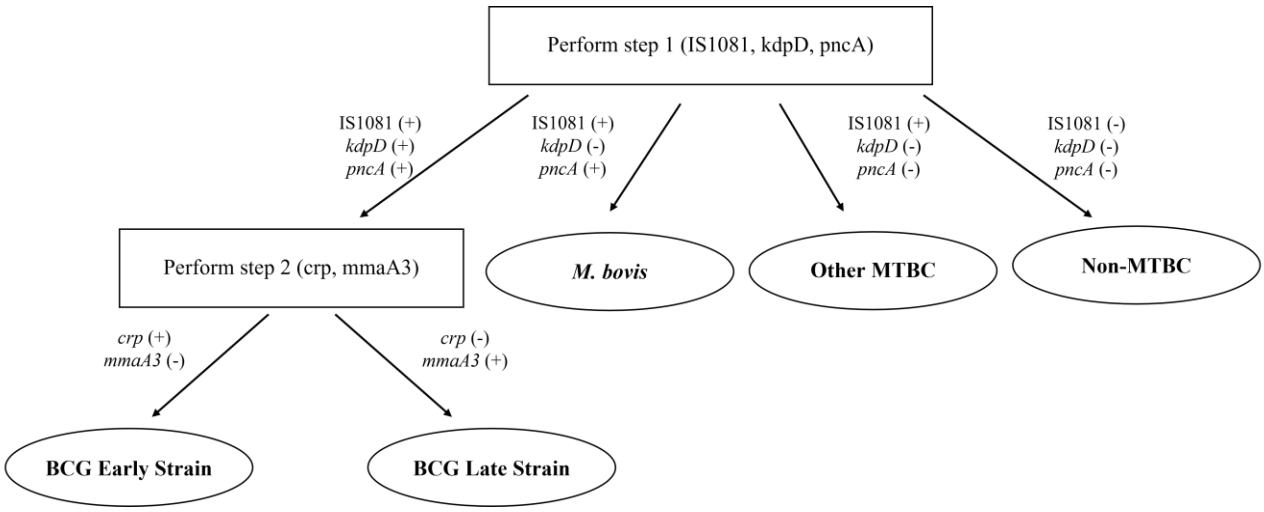



**Figure S6** Assay workflow and interpretation

## 11. Supplementary tables

| Isolate | IS1081 | *kdpD* | *pncA* | *crp* | *mmaA3* |
|---|---|---|---|---|---|
| *M. tuberculosis* H37Rv | + | - | - | | |
| *M. bovis* AF2122/97 | + | - | + | | |
| *M. orygis* 51145 | + | - | - | | |
| *M. africanum* MT18 7400 | + | - | - | | |
| *M. caprae* 6032 | + | - | - | | |
| *M. caprae* 65749 | + | - | - | | |
| *M. microti* | + | - | - | | |
| *M. abscessus* ATCC 19977 | - | - | - | | |
| *MAP* K10$^{+}$ | - | - | - | | |
| *MAP* Holland | - | - | - | | |
| *MAP* 6758 | - | - | - | | |
| *MAP* 8645 | - | - | - | | |
| *MAH* 104 | - | - | - | | |
| *MAA* 15769 | - | - | - | | |
| *MAA* 35713 | - | - | - | | |
| *MAA 35718* | - | - | - | | |
| *M. kansasii* ATCC 12478 | - | - | - | | |
| *M. smegmatis* MC$^{2}$155 | - | - | - | | |
| BCG Russia | + | + | + | + | - |
| BCG Danish | + | + | + | - | + |
| BCG Pasteur | + | + | + | - | + |
| BCG Sweden | + | + | + | + | - |
| BCG Moreau | + | + | + | + | - |
| BCG Glaxo | + | + | + | - | + |
| BCG Japan | + | + | + | + | - |
| BCG Prague | + | + | + | - | + |

**Table S1** Summary of specificity results from 26 isolates

Boxes filled in gray indicate not applicable

| Isolate No. | Age (months) | Sex | Location | Sample type | Clinical disease | Year of diagnosis | Extraction from | Extraction method | Assay result |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | M | Karnataka | Abscess pus | Failure to thrive; BCG abscess | 2020 | LJ | Qiagen kit | BCG Late strain |
| 2 | 3 | F | Andhra Pradesh | Gastric aspirate | TB meningitis and PTB | 2020 | MGIT | Qiagen kit | MTBC |
| 3 | 4 | M | Karnataka | Lymph node pus | Axillary lymphadenitis | 2019 | LJ | Qiagen kit | BCG Late strain |
| 4 | 5 | F | Tamil Nadu | Lymph node | Axillary lymphadenitis | 2018 | MGIT | Heat prep | BCG Early strain |
| 5 | 5 | F | Tamil Nadu | Lymph node pus | Axillary lymphadenitis | 2020 | LJ | Qiagen kit | BCG Late strain |
| 6 | 6 | M | Tamil Nadu | Lymph node | Submandibular lymphadenitis | 2019 | MGIT | Heat prep | MTBC |
| 7 | 6 | M | Andhra Pradesh | Lymph node | MSMD, disseminated BCG infection | 2018 | MGIT | Heat prep | BCG Early strain |
| 8 | 7 | M | Andhra Pradesh | Thigh sinus tract biopsy | Post-immunization abscess with sinus tract on thigh | 2020 | LJ | Qiagen kit | BCG Early strain |
| 9 | 9 | F | Tamil Nadu | Lymph node pus | SCID, PCP, failure to thrive, disseminated BCG infection | 2019 | MGIT | Heat prep | BCG Early strain |
| 10 | 9 | F | Tamil Nadu | Lymph node pus | MSMD, axillary lymphadenitis | 2020 | Pus swab | Qiagen kit | BCG Late strain |
| 11 | 10 | M | West Bengal | Talar bone | Ankle tubercular arthritis | 2019 | MGIT | Qiagen kit | BCG Early strain |
| 12 | 12 | M | Tamil Nadu | BCG site pus | SCID, BCG site ulceration | 2018 | MGIT | Heat prep | BCG Early strain |
| 13 | 12 | F | Odisha | Gastric aspirate | Congenital TB | 2020 | MGIT | Qiagen kit | MTBC |
| 14 | 12 | F | Tamil Nadu | Lymph node | TB cervical lymphadenopathy | 2020 | MGIT | Qiagen kit | MTBC |
| 15 | 18 | M | Andhra Pradesh | Gastric aspirate | PTB | 2019 | MGIT | Heat prep | MTBC |
| 16 | 19 | M | Andra Pradesh | Pleural tissue | TB pleural effusion | 2019 | MGIT | Heat prep | MTBC |
| 17 | 20 | M | Bangladesh | Colonic ulcer biopsy | Chronic active colitis | 2019 | MGIT | Heat prep | MTBC |
| 18 | 22 | M | West Bengal | Gastric aspirate | TB meningitis and PTB | 2019 | MGIT | Qiagen kit | MTBC |
| 19 | 24 | M | Tamil Nadu | Mastoid tissue | Tubercular mastoiditis | 2020 | LJ | Qiagen kit | MTBC |

**Table S2** Clinical and DNA extraction information for 19 suspected BCG clinical isolates

PTB – pulmonary tuberculosis, MSMD – Mendelian susceptibility to mycobacterial disease, SCID – severe combined immunodeficiency disorder, PCP – pneumocystis pneumonia, CNS – central nervous system, MGIT – mycobacterial growth indicator tube, LJ – Lowenstein Jensen medium, MTBC – *Mycobacterium tuberculosis* complex

# PREFACE TO CHAPTER IV

In this next chapter, we aimed to develop a model to study the pathogenesis of bovine mycobacteria within the lab. We sought to apply this model to *MAP,* a pathogen which has particular relevance to Canada due to the high prevalence of paratuberculosis and Crohn's disease. As mentioned in chapter I, there are currently no mouse models of infection with *M. orygis, M. bovis,* or *MAP* which mimic their natural route of infection and progression to disease. An appropriate mouse model would act as a valuable tool to study bacterial pathogenesis and to test novel diagnostics or vaccines. In this chapter, I describe my development of a streptomycin pre-treatment mouse model of oral infection with *MAP* and my assessment of the finalized model in zoonotic TB pathogens *M. bovis* and *M. orygis.* I determined that an orally administered streptomycin pre-treatment model of infection with high doses of *MAP* led to chronic infection of the intestines and mesenteric lymph nodes up to 24-weeks post-gavage. However, there was no evidence of inflammation or disease. When this model was applied to *M. bovis* or *M. orygis,* this again led to persistent infection in the intestines and mesenteric lymph nodes, but pulmonary dissemination and disease also occurred. We concluded that this model mimicked what can be observed in zoonotic TB in humans where an oral infection progressed to lung disease. However, other methods, such as an immune-deficient mouse strain, may need to be explored in order to model *MAP* enteritis in mice.

# CHAPTER IV

## Establishment of persistent enteric mycobacterial infection following streptomycin pre-treatment

---

Published manuscript – Gut Pathogens, Volume 15, Article 46,

doi: https://doi.org/10.1186/s13099-023-00573-w

Shannon C Duffy[a-c], Andréanne Lupien[a-d], Youssef Elhaji[e], Mina Farag[f,g], Victoria Marcus[f,g], Marcel A Behr[a-d*]

**Affiliations**

[a]*Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada*

[b]*McGill International TB Centre, Montreal, Quebec, Canada*

[c]*The Infectious Diseases and Immunity in Global Health Program, Research Institute of the McGill University Health Centre, Montreal, QC, Canada*

[d]*Department of Medicine, McGill University, Montreal, Quebec, Canada*

[e]*Diagnostic Genomic Division, Department of Laboratory Medicine and Pathology, Hamad Medical Corporation, Doha, Qatar*

[f]*Department of Pathology, McGill University, Montreal, Quebec, Canada*

[g]*Department of Laboratory Medicine, Division of Pathology, McGill University Health Center, Montreal, Quebec, Canada*

## 1. Abstract

*Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) is the causative agent of paratuberculosis, a chronic gastrointestinal disease affecting ruminants. This disease remains widespread in part due to the limitations of available diagnostics and vaccines. A representative small animal model of disease could act as a valuable tool for studying its pathogenesis and to develop new methods for paratuberculosis control, but current models are lacking. Streptomycin pre-treatment can reduce colonization resistance and has previously been shown to improve enteric infection in a *Salmonella* model. Here, we investigated whether streptomycin pre-treatment of mice followed by *MAP* gavage could act as a model of paratuberculosis which mimics the natural route of infection and disease development in ruminants. The infection outcomes of *MAP* were compared to *M. avium* subsp. *hominissuis* (*MAH*), an environmental mycobacterium, and *M. bovis* and *M. orygis,* two tuberculous mycobacteria. Streptomycin pre-treatment was shown to consistently improve bacterial infection post-oral inoculation. This model led to chronic *MAP* infection of the intestines and mesenteric lymph nodes (MLNs) up to 24-weeks post-gavage, however there was no evidence of inflammation or disease. These infection outcomes were found to be specific to *MAP.* When the model was applied to a bacterium of lesser virulence *MAH,* the infection was comparatively transient. Mice infected with bacteria of greater virulence, *M. bovis* or *M. orygis,* developed chronic intestinal and MLN infection with pulmonary disease similar to zoonotic TB. Our findings suggest that a streptomycin pre-treatment mouse model could be applied to future studies to improve enteric infection with *MAP* and to investigate other modifications underlying *MAP* enteritis.

**Keywords:** *Mycobacterium avium* subsp. *paratuberculosis,* paratuberculosis, Crohn's disease, mouse models, *Mycobacterium avium* subsp. *hominissuis, Mycobacterium bovis, Mycobacterium orygis*

## 2. Introduction

*Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) is the cause of a chronic gastrointestinal disease called paratuberculosis which affects ruminants such as cattle, sheep, goats, and deer[1]. Post-*MAP* infection, animals enter an initial subclinical period lasting 2-5 years which is followed by clinical disease characterized by diarrhea, wasting, and eventual death[2]. Control of paratuberculosis relies on a test and cull strategy to remove infected animals from the herd, however its efficacy is constrained by several factors such as the limitations of *MAP* diagnostics[3,4]. Paratuberculosis therefore remains widespread with global prevalence estimates ranging from 10-70%[3]. This disease poses a major economic burden in many countries, such as in the United States, where it is associated with a cost of $198 million per year[5]. In addition, *MAP* may also pose a threat to public health if, as hypothesized, it is etiologically linked to Crohn's disease[6].

*MAP* is a nontuberculous mycobacterium which emerged from *M. avium* subsp. *hominissuis* (*MAH*), an environmental species which can cause opportunistic disease in humans and pigs[7,8]. *MAP* is an obligate intracellular pathogen of intestinal macrophages[9]. It will invade the intestines through M cells in Peyer's patches and infect macrophages found within the lamina propria[9]. The infected macrophages may remain local or travel to the mesenteric lymph nodes (MLNs) where *MAP* can also persist[10]. Diseased animals will develop granulomatous enteritis with a thickening of the intestinal wall, primarily in the terminal ileum, accompanied by

inflamed MLNs[11]. *MAP* is shed into the feces allowing other animals to be infected via the fecal to oral route[3]. However, *MAP* may also be acquired from the environment due to the ability of the bacteria to persist for long periods of time in water and soil[12,13].

To study *MAP* pathogenesis and develop new methods for paratuberculosis control, a small animal model could be a valuable tool. Bovine models, while offering the obvious benefit of being a natural *MAP* host, are often an impractical choice due to their cost, space, and personnel training requirements. Mouse models alternatively offer several benefits including reduced cost, increased accessibility, and availability of a variety of defined and well-characterized mouse strains. Although intraperitoneal injection of mice with *MAP* leads to reproducible infection of the spleen and liver, it does not mimic the natural route of infection[14]. An oral infection model would match the natural route; however, gavage of *MAP* can lead to inconsistent infection and does not result in intestinal disease[15,16]. This phenomenon has also been observed with other known enteropathogens such as enteropathogenic and enterohemorrhagic *Escherichia coli*[17,18]. In the case of *Salmonella enteria* serovar Typhimurium, gavage infections in mice lead to systemic dissemination rather than local disease in the intestine. However, Barthel *et al.* found that if the mice were pre-treated with streptomycin to reduce colonization resistance posed by gut microbes prior to gavage with a streptomycin-resistant strain of *S. enterica* serovar Typhimurium, intestinal infection does occur with disease which more closely mimics the enterocolitis observed in humans[19].

Here, we investigated whether an oral streptomycin pre-treatment model of *MAP* infection could act as a representative mouse model of *MAP*-induced intestinal infection and disease. To compare how our *MAP* results compared with other mycobacteria, we applied the same infection strategy to *MAH,* a mycobacterium of lesser virulence*,* which is primarily

environmental but has been shown to cause opportunistic infections in immunocompromised humans and pigs[20]. The model was also applied to 2 mycobacteria of greater virulence (*M. bovis*, *M. orygis*) which are causes of bovine and zoonotic TB and are suggested to infect humans through drinking unpasteurized milk[21,22]. Overall, we determined that following streptomycin pre-treatment, *MAP* resulted in chronic intestinal infection with no evidence of intestinal disease, unlike *MAH* which infected only transiently and *M. bovis/M. orygis* which disseminated to the lungs and caused pulmonary disease.

## 3. Materials and methods

### *3.1 Media*

All mycobacterial strains were grown in Middlebrook 7H9 broth (Difco Laboratories, Detroit, MI) with 0.2% glycerol, 0.1% Tween 80, and 10% albumin-dextrose-catalase (Becton, Dickinson and Co., Sparks, MD) with rotation at 37°C. For colony isolation, the bacteria were plated on Middlebrook 7H10 medium supplemented with 10% oleic acid-albumin-dextrose-catalase (Becton, Dickinson and Co.). To grow *MAP,* 0.1% mycobactin J (Allied Monitor, Fayette MO) was added to liquid and solid media. The antibiotics kanamycin (50µg/mL), streptomycin (50µg/mL), or PANTA antibiotic mixture (Becton, Dickinson and Co.) were included in the media when required.

### *3.2 Generation of streptomycin-resistant strains*

To generate streptomycin-resistant (strep-R) strains of *MAP* K10*, MAH* 104*, M. bovis* Ravanel and *M. orygis* 51145, a K43R mutation was introduced into the *rpsL* gene of each species using oligo-mediated recombineering as previously described with minor

modifications[23]. This mutation is known to confer resistance to streptomycin[24]. In brief, the pNit::ET plasmid (from Kenan Murphy – Addgene plasmid #107692) was first introduced into each strain and selected on kanamycin-containing media. The plasmid was confirmed by PCR of the kanamycin cassette (Supplementary Table 1). To introduce the K43R mutation in *MAP, M. bovis* and *M. orygis,* 30mL cultures containing pNit::ET were grown to log phase in 7H9 media with kanamycin and then diluted to an $OD_{600}$ of 0.1. When cultures reached an $OD_{600}$ of 0.8, 30µL of a 1000X stock of isovaleronitrile was added to stimulate the pNit::ET plasmid. When cultures reached an $OD_{600}$ of 1.0, 3mL of 2M glycine was added to the culture. The next day, the cultures were washed 3 times in 10% glycerol. The cells were resuspended in 1mL 10% glycerol and 200µL of cells were transferred to a 2mm gap electroporation cuvette (Fisherbrand, Waltham, MA) containing 1µg of a 70-mer oligo designed to introduce the K43R mutation (Supplementary Table 1). The cells were electroporated using the following settings: 2.5 kV, 1,000Ω, and 25µF. The cells were then transferred to 3mL of 7H9 and incubated at 37°C. After 5 days (*MAP*) or 2 days (*M. bovis, M. orygis*), the cultures were recovered on 7H10 agar with streptomycin.

To generate the *rpsL* mutation in *MAH,* the same procedure was followed except for the following modifications: the starting culture was 50mL, 50µL of isovaleronitrile was used in stimulation, 7.5mL of 2M glycine was added prior to electroporation, washes were performed with pre-warmed 10% glycerol with 2M sucrose, cells were resuspended in 800µL 7H9 with 2M sucrose prior to recovery, and cells were recovered after 1 day.

Colonies which grew on streptomycin were grown up and screened for the K43R mutation by PCR of the *rpsL* gene followed by Sanger sequencing (Supplementary Table 1). Sequences were visualized using Geneious Prime (version 2022.1.1).

*3.3 Whole genome sequencing and analysis*

Genomic DNA was extracted from strep-R *MAP* and the parental strain using the QIAamp UCP pathogen mini kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. Paired end sequencing libraries were prepared using the S4 reagent kit (Illumina, San Diego, CA) and shotgun sequencing was performed using the NovaSeq 6000 S4 PE150 - 35M reads (Illumina). The sequence was aligned to the *MAP* K10 reference genome (NC_002944.2) using BWA-MEM[25]. The reads were sorted using SAMtools and visualized using Integrative Genomics Viewer (IGV)[26,27].

The strep-R *MAP* sequence was analyzed for deletions and for SNPs outside of the engineered K43R mutation in *rpsL.* Duplicate reads in sorted BAM files were removed using Picard (http://broadinstitute.github.io/picard). Variant calling was done using Freebayes v1.3.6 with mapping quality 60, minimum read coverage 10, and minimum allele frequency of 0.5[28]. Variant calls were annotated using SNPEff v.4.3[29]. Variants identified in strep-R *MAP* were compared with the parental strain to identify unique variants.

*3.4 pH experiment*

The gastric pH of a mouse is between 3.0 (fed) and 4.0 (after fasting)[30]. To quantify the amount of *MAP* potentially lost due to the acidic environment of the stomach, strep-R *MAP* was grown to an $OD_{600}$ of 1.0 (~$2x10^8$ colony-forming units (CFU)/mL) in 7H9 media. The culture was then split and resuspended in PBS with 0.1% Tween-80 at either a pH of 7.0 or 3.0 and incubated at 37°C for 1, 2, or 4 hours. At the indicated timepoint, the cultures were spun down and resuspended in 7H9 then serially diluted and plated on 7H10 agar.

*3.5 Animals*

Mice were housed in a pathogen-free environment at the Research Institute of the McGill University Health Centre (RI-MUHC). All animal experiments were in accordance with the guidelines of the Canadian Council on Animal Care (CCAC) and all protocols were approved by the animal resource division of the RI-MUHC. C57BL/6 and BALB/c mice were purchased from The Jackson Laboratory (Bar Harbor, ME) at 7-weeks of age and quarantined in-house for 1 week prior to infections. All mouse infections were sex and age matched.

*3.6 MAP mouse infection model*

Infection outcomes were compared using a strep-R *MAP* inoculum of $10^8$ or $10^9$ CFU. Bacterial inoculum was prepared by growing strep-R *MAP* to an $OD_{600}$ of 0.5 (~$1x10^8$ CFU/mL) in 7H9 with streptomycin, washing the cells once in PBS with 0.01% Tween-80, and then resuspending the bacteria in either 1/5 or 1/50 the original volume in PBS with 0.01% Tween-80 to generate a stock of $5x10^8$ CFU/mL or $5x10^9$ CFU/mL. Prior to all gavage steps, mice were fasted of food and water for 3 hours. The first day, mice were given 200µL of streptomycin (100mg/mL) by oral gavage. Twenty-four hours later mice were orally inoculated with 200µL of the prepared strep-R *MAP* stock (equivalent to a dose of $10^8$ or $10^9$ CFUs). Mice given the higher dose were given a second $10^9$ CFU strep-R *MAP* infection by oral gavage the following day. When indicated, mice were given 100µL of 3% sodium bicarbonate 30 minutes[31] prior to oral gavage of strep-R *MAP.* At the indicated timepoints, mice were sacrificed, and the small intestine, large intestine, MLNs, spleen, liver, and lungs were taken for quantification of organ CFUs. The small intestine, large intestine, and MLNs were also sent for histopathology assessment.

*3.7 MAH murine infection model*

A $5x10^9$ CFU/mL bacterial stock of strep-R *MAH* was prepared by growing the bacteria to an $OD_{600}$ of 0.5 (~$1x10^8$ CFU/mL). The culture was washed once in PBS with 0.01% Tween-80 and then resuspended in 1/50 of the original volume in PBS with 0.01% Tween-80. Mice were fasted of food and water for 3 hours prior to all oral gavage steps. Mice were given 200µL of streptomycin (100mg/mL) followed by two consecutive doses of 200µL of strep-R *MAH* ($5x10^9$ CFU/mL) at 24-hour intervals. Mice infected with strep-R *MAH* were euthanized 48-hours, 4-, 8-, 12-, and 24-weeks post-infection and the small intestine, large intestine, and MLNs were assessed for CFUs.

*3.8 M. bovis and M. orygis mouse infection model*

All *M. bovis* and *M. orygis* experiments took place in the containment level 3 facilities at the RI-MUHC as a fully virulent *M. bovis* strain (*M. bovis* Ravanel) was used rather than an *M. bovis* BCG vaccine strain. *M. orygis* is a cause of zoonotic TB in people in or migrating from South Asia[21], and a clinical isolate collected in Canada (*M. orygis* 51145) was used[32]. A $5x10^9$ CFU/mL bacterial stock of strep-R *M. bovis*/*M. orygis* was prepared by first growing strep-R *M. bovis*/*M. orygis* in 7H9 to an $OD_{600}$ of 0.5 (~$5x10^7$ CFU/mL). The culture was washed once in PBS with 0.01% Tween-80 and then resuspended in 1/100 of the original volume in PBS with 0.01% Tween-80. Mice were fasted of food and water for 3 hours prior to all oral gavage steps. Mice were given either 200µL of streptomycin (100mg/mL) or no pre-treatment followed by two consecutive doses of strep-R *M. bovis*/*M. orygis* spaced 24-hours apart. Due to the recognized virulence of *M. bovis*/*M. orygis* in mice[33,34], infected mice were weighed and monitored for survival throughout the 24-week experiment. Mice were euthanized at 4- and 24- weeks post-

infection and the small intestine, large intestine, MLNs, and lungs were assessed for CFUs and histopathology.

*3.9 Organ CFU quantification*

At each timepoint, the small and large intestines were excised, separated, and cut open longitudinally. The fecal matter and mucus were removed mechanically using the flat end of curved tweezers, and the organs were washed 3 times in PBS with 0.01% Tween-80 before being placed in 1mL 7H9. The intestines were processed in this way to identify *MAP* CFUs that had invaded into the organ tissue and to avoid CFUs passing through the intestines as a result of the inoculation at the earlier timepoints. The spleens, livers, and lungs were placed directly in 1mL 7H9. The organs were then homogenized with an Omni Tissue Homogenizer TH (Omni International, Kennesaw, GA) for 45 seconds. Finally, the MLN chain was directly pushed through a 70µm sterile cell strainer (Fisher Scientific, Waltham, MA) into 1mL 7H9. The resulting homogenates for all collected organs were serially diluted in 7H9 and plated on 7H10 containing PANTA to quantify CFUs. If no colonies were counted on the lowest dilutions plated, the organ was assigned the limit of detection (LOD).

*3.10 F57 real-time PCR of fecal pellets*

DNA was extracted from fecal pellets of C57BL/6 and BALB/c mice that were uninfected or 12-weeks post-*MAP* infection using the QIAmp PowerFecal Pro DNA kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The *MAP* K10 genome is 4,829,781 base pairs which corresponds to ~5.3fg per copy[158]. A standard curve for qPCR was prepared from $1x10^7$ genome equivalents (ge) (53ng) to 1 ge (5.3fg). A Maxima SYBR

Green/ROX real-time PCR assay (Thermo Fisher Scientific, Waltham, MA) was performed using primers for the single copy gene F57 (Supplementary Table 1) following the manufacturer's instructions.

*3.11 Histopathology*

Small intestines and large intestines from all infected and uninfected control mice, mesentery from *MAP*-infected and uninfected mice, and lungs from *M. bovis*-infected, *M. orygis*-infected, and uninfected mice were sent to the histology core at McGill University. The organs were paraffin-embedded and 4µm sections were cut and stained by hematoxylin and eosin (H&E) staining. Positive control slides of enteritis were generously provided from Dr. Laura Sly's lab who has a model of spontaneous enteritis in $SHIP^{-/-}$ mice[36]. All slides were reviewed by a pathologist at the MUHC. All slides were photographed with a Nikon Eclipse NI microscope.

*3.12 Lipocalin-2 ELISA*

Fecal pellets were collected from infected and uninfected control mice and stored at -80°C until processed. Fecal pellets were weighed and placed in a 1.5mL screwcap tube. PBS with 0.1% Tween-20 was added at a volume of 10ul per 1mg of feces. The tubes were vortexed at maximum speed for 20 minutes and then spun down in a microcentrifuge at 12,000 x g for 10 minutes at 4°C. The supernatant was transferred into a new screwcap tube and stored at -20°C until ready for ELISA of lipocalin-2, a broad marker of intestinal inflammation in mice[37]. The assay was performed using the mouse lipocalin-2/NGAL DuoSet ELISA (Bio-techne, Minneapolis, MO) and DuoSet ELISA ancillary reagent kit (Bio-techne) according to the manufacturer's instructions.

*3.13 Statistics*

Statistical analyses were conducted with GraphPad Prism (version 9.3.1). Grouped data are graphed as individual datapoints with the sample median. Multiple groups comparisons were performed using a two-way ANOVA with Sidak's multiple comparisons test. To compare 2 groups, an unpaired t-test was used. Analysis of pooled organ data was performed using a Mann-Whitney test.

## 4. Results

*4.1 Streptomycin pre-treatment increases MAP bacterial burden following oral infection*

A strep-R *MAP* strain was generated by introducing a K43R mutation into the *rpsL* gene via oligo-mediated recombineering. The SNP introduction was confirmed via Sanger sequencing (Supplementary Fig. 1A) and the strain was shown to grow on 7H10 agar with 50μg/mL streptomycin (Supplementary Fig. 1B). Whole-genome sequencing was performed on the strep-R *MAP* and parent strain to determine whether any off-target mutations occurred in the generation of strep-R *MAP*. No deletions were detected and only 1 additional point mutation was detected in *mtrB,* a gene which makes up a two-component regulatory system. The *mtrAB* system has previously been associated with multidrug resistance in *M. avium,* therefore it is possible that this is a compensatory mutation[38]. To test whether streptomycin pre-treatment would improve infection of *MAP* following oral gavage, C57BL/6 mice were given 20mg of streptomycin or no pre-treatment followed by a dose of $10^8$ CFU strep-R *MAP* 24-hours later. After 48 hours, mice were euthanized and the organ CFUs were compared between the 2 groups (Fig. 1A). The mean strep-R *MAP* CFUs were ~14X greater (*p=0.04) in the large intestine in mice that received streptomycin pre-treatment (Fig. 1B).

*4.2 Increasing gastric pH does not improve MAP infection following oral inoculation*

In order to determine whether *MAP* intestinal load could be further increased following streptomycin pre-treatment, the effect of the acidic gastric pH of the mouse stomach on *MAP* death following gavage was tested. The effect of pH on *MAP* survival was tested both *in vitro* and *in vivo.* Strep-R *MAP* was grown to an $OD_{600}$ of 1.0 (~$2x10^8$ CFU) and then transferred to PBS at a pH of either 7.0 or 3.0. There were significantly fewer CFUs in cultures exposed to an acidic pH after 1- (*p=0.019), 2- (**p= 0.009) and 4-hours (**p=0.009). One hour in PBS at a pH of 3.0 decreased the number of viable strep-R *MAP* by ~0.5-log and 4 hours decreased the number of viable strep-R *MAP* by >1-log (Fig. 1C).

To investigate whether increasing the mouse gastric pH would therefore improve *MAP* infection post-gavage, mice were administered sodium bicarbonate prior to gavage of strep-R *MAP*. Mice were pre-treated with 20mg of streptomycin and then 24-hours later half of the mice were given 3% sodium bicarbonate 30 minutes[31] before all mice were given $1x10^8$ CFU strep-R *MAP* (Fig. 1D). After 48 hours, there were no differences in bacterial burden found between mice that received sodium bicarbonate and those that did not in the large intestine (Fig. 1E), small intestine (Fig. 1F) or MLNs (Fig. 1G). This indicated that although a low pH reduced strep-R *MAP* CFUs *in vitro,* increasing gastric pH *in vivo* did not improve strep-R *MAP* organ infection post-gavage and therefore was not performed in later experiments.

*4.3 Two consecutive oral doses of strep-R MAP led to infection of the mesenteric lymph nodes*

The effect of increasing the dosage of strep-R *MAP* organ infection was next evaluated. C57BL/6 mice were given streptomycin pre-treatment followed either by a single dose of $1x10^8$ CFU strep-R *MAP* or 2 doses of $1x10^9$ CFU given 24-hours apart (Fig. 2A). Organ CFUs were

assessed at 48-hours, 4-, and 8-weeks post-infection. No differences were observed between the doses in infection of the large intestine (Fig. 2B) or small intestine (Fig. 2C). However, only mice that received the 2 high doses of strep-R *MAP* had CFUs in the MLNs (Fig. 2D), suggesting that a greater dose is required for *MAP* migration to the lymph nodes. This was significant at 48-hours (*p=0.013) 4- (**p=0.009) and 8-weeks (***p=0.001) post-infection. At 48-hours post-infection, strep-R *MAP* CFUs in the higher dose group were greatest in the large intestine (~$10^3$-$10^4$ CFU/g) and MLNs (~$10^1$ CFU/MLN chain). Later at 4- and 8-weeks post-infection, strep-R *MAP* CFUs were primarily restricted to the MLNs (~$10^1$-$10^2$ CFU/MLN chain). This persistence of *MAP* within the MLNs is consistent with the known course following natural infection[10]. The higher dose infection model was therefore chosen for the remaining mouse experiments.

*4.4 The streptomycin pre-treatment MAP model leads to chronic infection without disease*

To determine whether the streptomycin pre-treatment *MAP* model led to long-term infection, the inoculation was repeated with timepoints extending to 24-weeks post-infection (Fig. 3A). Strep-R *MAP* was found to persist primarily in the large intestine (Fig. 3B) and MLNs (Fig. 3C) with CFUs sporadically observed in the small intestine (Fig. 3D). At initial uptake, *MAP* infects the large intestine (~$10^3$-$10^4$ CFU/g) and MLNs (~$10^1$ CFU/MLN chain). At 4- and 8-weeks post-infection *MAP* was primarily detected in the MLNs (~$10^1$-$10^2$ CFU/MLN chain). By 12-weeks post-infection *MAP* was detected at greatest abundance in the large intestine (~$10^2$-$10^3$ CFU/g). At 24-weeks post-infection, *MAP* infection remained in some mice but had been cleared in most. *MAP* was not consistently detected in the spleen, liver, or lungs throughout the 24-week experiment (Supplementary Fig. 2). Considering 12-weeks post-gavage was the peak of

*MAP* infection in the large intestine, fecal shedding was assessed via quantitative PCR of the F57 gene between uninfected controls and mice infected 12 weeks prior. No observable differences were detected between these 2 groups at this timepoint and the ge values were consistent with background (Supplementary Fig. 3).

To determine whether chronic infection led to inflammation or observable differences in histopathology, organs were sent for processing and H&E staining. The slides of the large intestine, small intestine, and MLNs were compared with slides from controls at the 12- and 24-week timepoints by a pathologist. These timepoints were chosen since *MAP* disease progression typically occurs after a long subclinical period and the 12-week timepoint was the peak of bacterial load in the large intestine. No changes were observed in the histopathology of infected animal organs compared to uninfected controls (Fig. 3E). To investigate whether more subtle inflammatory changes had occurred, fecal pellets were also processed for quantification of lipocalin-2, a broad marker for inflammation in mice[37] (Fig. 3F). No differences were observed in levels of lipocalin-2 compared to uninfected controls. Together this data suggests that a streptomycin pre-treatment model can result in long-term (up to 24 weeks) organ infection of *MAP* in C57BL/6 mice without signs of inflammation or disease.

### *4.5 The infection outcomes of a streptomycin pre-treatment MAP model are comparable between C57BL/6 and BALB/c mice*

To identify whether infecting a mouse strain more susceptible to mycobacteria[39] would lead to increased bacterial burden or disease induction, the model was compared between C57BL/6 and BALB/c mice (Fig. 4A). Overall, no significant differences were observed between the 2 mouse strains in strep-R *MAP* infection of the large intestine (Fig. 4B), small intestine (Fig.

4C) or MLNs (Fig. 4D). This suggests that *MAP* organ infection is not affected by the differences between the genotypes of the C57BL/6 and BALB/c mice. Based on the comparable results across the different mouse strains, we pooled the results together to explore any time-dependent trends. Analysis of total CFU burden revealed that infection peaked at 12-weeks post-gavage and that the burden was diminished by 24-weeks (*p=0.045) (Fig. 4E). This trend was seen also by analysis of each organ with a significant difference observed in the MLNs (*p=0.049) (Supplementary Fig. 4). Consistent with the findings from C57BL/6 mice, no changes were observed in the histopathology of BALB/c mice compared to unexposed controls (Fig. 4F). The levels of lipocalin-2 were also comparable between the 2 groups, other than at 24-weeks where a minor increase was observed in infected animals (*p=0.042) (Fig. 4G). When streptomycin pre-treatment was compared with no pre-treatment in BALB/c mice, strep-R *MAP* CFUs were shown to increase in the large intestine and MLNs with streptomycin pre-treatment. Slight increases in lipocalin-2 were observed in the streptomycin group compared to no pre-treatment (Supplementary Fig. 5). Overall, this data suggests that infection outcomes are largely unchanged by using BALB/c mice.

*4.6 The streptomycin pre-treatment infection strategy with MAH leads to comparatively transient infection*

To understand how infection outcomes compare when the same infection strategy was applied to an environmental mycobacterium, the infection was repeated using *MAH,* the closest relative to *MAP*[8]. A strep-R strain of *MAH* was generated using oligo-mediated recombineering to introduce a K43R mutation in the *rpsL* gene and was confirmed by Sanger sequencing (Supplementary Fig. 6A). C57BL/6 mice were given 20mg streptomycin followed by 2

consecutive doses of $1x10^9$ CFU strep-R *MAH* (Fig. 5A). Organ CFUs were found in the large and small intestine only 48-hours post-infection, after which strep-R *MAH* was cleared (Fig. 5B, 5C). Bacterial load ranged from ~$10^3$-$10^4$ CFUs/g in the large intestine, which is comparable to *MAP* CFUs at 48-hours (Fig. 3B). In the MLNs, infection was observed in some mice at 48-hours and 4-weeks post-infection, after which it was also cleared (Fig. 5D). Collectively, this data suggests that the persistence phenotype of *MAP* observed following this infection strategy is specific to the bacteria rather than the mode of infection.

*4.7 Application of the streptomycin pre-treatment strategy with M. bovis or M. orygis leads to pulmonary infection and disease*

To investigate whether the streptomycin pre-treatment infection model could lead to disease with more pathogenic species, infection outcomes were compared with *M. bovis* and *M. orygis,* 2 bovine mycobacteria causing tuberculosis. Strep-R *M. bovis/M. orygis* strains were generated by oligo-mediated recombineering and the K43R mutation was confirmed by Sanger sequencing (Supplementary Fig. 6B). For *M. bovis* infections, C57BL/6 mice were given 20mg streptomycin or no pre-treatment followed by 2 consecutive oral doses of $1x10^9$ CFU of strep-R *M. bovis* each spaced 24-hours apart (Fig. 6A). Strep-R *M. bovis* infection consistently occurred in the large intestine (Fig. 6B), small intestine (Fig. 6C), and MLNs (Fig. 6D). Dissemination was also detected in the lungs (Fig. 6E). At 24-weeks post-infection, streptomycin pre-treatment led to significantly increased CFUs in the large intestine (*p=0.041), small intestine (*p=0.048) and lungs (**p=0.0085). Over the course of the experiment, 2 mice in the streptomycin pre-treatment group had to be euthanized at 5-weeks (mouse 1) and 15-weeks (mouse 2) post-infection (Fig. 6F). Assessment of the histopathology of their lungs indicated that mouse 1 likely

died due to lung disease due to the large areas of consolidation and immune cell infiltration observed. The lungs of mouse 2 appeared normal which indicated it may have died due to reasons unrelated to *M. bovis* infection (Fig. 6G). At the experimental endpoint of 24-weeks, sections of the lungs of the remaining mice were also sent for histopathology which indicated that 3 out of 4 mice had some evidence of lung disease including airway consolidation and interstitial lymphocytic infiltration (Fig. 6H). None of the *M. bovis*-infected mice had evidence of disease in their small or large intestine.

To determine whether similar infection outcomes would occur with another cause of bovine and zoonotic TB, the infection was repeated with strep-R *M. orygis* (Supplementary Fig 7A). Similar outcomes were observed in this experiment where gavage led to local infection of the large intestine, small intestine, and MLNs with dissemination to the lungs at 24-weeks post-infection (Supplementary Fig. 7B-E). One mouse died at 13-weeks post-infection without evidence of lung disease observed by H&E staining (Supplementary Fig. 7F, G). The surviving mice that were euthanized at the experimental endpoint had some evidence of lung disease with areas showing reactive peri-bronchial lymphoid aggregates, interstitial lymphocytic infiltration and foamy histiocytes (Supplementary Fig. 7H). None of the *M. orygis*-infected mice had evidence of disease in their small or large intestine.

## 5. Discussion

Paratuberculosis remains a widespread concern for animal health and a significant economic burden. The limitations of available diagnostics and vaccines for paratuberculosis deters elimination of the disease. The development of a mouse model of *MAP* which mimics the natural route of infection and leads to intestinal disease would be a valuable tool for testing new

methods of detecting or preventing paratuberculosis. However, oral infection models of *MAP* are currently lacking. Here, we investigated whether streptomycin pre-treatment, a method previously shown to support *S. enterica* serovar Typhimurium induced enteritis, could improve *MAP* infections and model paratuberculosis in mice.

Streptomycin pre-treatment was found to significantly improve *MAP* infection of the large intestine (Fig. 1B). The benefit of streptomycin pre-treatment to increase organ bacterial load was also found when infecting another mouse strain (BALB/c) or when inoculating another mycobacterium (*M. bovis*) (Supplementary Fig. 5, Fig. 6). To further enhance *MAP* infection, the dosage was increased to 2 consecutive doses of $10^9$ and only mice that received this higher dose were shown to have *MAP* infection in the MLNs (Fig. 2). During the clinical phase of paratuberculosis, inflamed MLNs are one of the frequently observed signs of disease[11], thus this higher dosage led to a more representative *MAP* model. Streptomycin pre-treatment followed by a high dose *MAP* gavage led to a persistent *MAP* infection in the intestines and MLNs which peaked at 12-weeks. Our strategy for assessing intestinal CFUs was designed to detect bacterial burden within the organ tissue. It is possible that additional CFUs may have been detected within the intestines if the mucus was not washed away during processing. Although the improvement in *MAP* burden using this streptomycin pre-treatment was promising, there was no evidence of inflammation or disease in either C57BL/6 or BALB/c mice (Fig. 3, Fig. 4). Furthermore, there was no evidence that *MAP* was being shed through the feces at the time of the maximum infection burden. Future studies attempting to model paratuberculosis in mice may benefit from a streptomycin pre-treatment strategy to enhance *MAP* infection. However, given that disease was not generated at an inoculum that is difficult to exceed, for technical reasons, additional host

manipulations such as an immunodeficient mouse strain are likely required to generate *MAP* enteritis.

In addition to streptomycin pre-treatment, our study investigated the effect of gastric pH on infection outcomes. Although *in vitro* studies showed that a low pH like the environment of the mouse stomach did lead to reduced *MAP* CFUs, when sodium bicarbonate was used to increase the mouse gastric pH prior to gavage with *MAP*, no differences were observed in the organ CFUs compared to mice given no sodium bicarbonate (Fig. 1C-G). It is possible that the effect seen *in vitro* did not translate into a detectable difference *in vivo*, simply because of the amount of time *MAP* spends in mouse stomach. Schwartz *et al.* reported the mouse gastric emptying time to be 74 ± 17 minutes but also reported considerable variability between animals[40]. Given that the tissue burden of *MAP* was much lower than the inoculum we delivered via gavage, our results suggest that there are other pH-independent factors which hinder *MAP* infection post-oral inoculation.

To determine how infection outcomes may vary when using the same infection strategy with mycobacteria of lesser or greater virulence, the infection was compared with *MAH,* a primarily environmental bacteria, and *M. bovis/M. orygis,* bacteria known to cause bovine tuberculosis. Gavage of *MAH* led to a transient infection in the large intestine and MLNs which was cleared after 48-hours or 4-weeks post-infection respectively (Fig. 5). This indicated that the persistence of *MAP* was specific to the bacteria rather than the method of streptomycin pre-treatment. When mice were inoculated with *M. bovis* or *M. orygis,* this led to both infection of the intestines and MLNs, and dissemination to the lungs. After 24 weeks, these mice had clear signs of lung disease which varied in severity. Overall, these infection outcomes were consistent with a model of zoonotic tuberculosis where exposure via the oral route led to pulmonary

disease. This indicates that *MAP* localization and lack of disease development within the streptomycin pre-treatment model was again bacteria-specific rather than a product of the infection strategy.

In conclusion, our data suggests that streptomycin pre-treatment is an effective method for improving infection of *MAP* post-gavage. When compared to *MAH, M. bovis,* and *M. orygis*, *MAP* gavage led to local persistent infection which was not observed with the other bacteria. Future studies aiming to develop an oral infection model of *MAP* may build upon these data to enhance infection and/or to produce an inflammatory response to *MAP* infection. A mouse model of paratuberculosis could act as an important platform for testing novel diagnostics and vaccines to improve management of paratuberculosis.

## 6. Abbreviations

| | |
|---|---|
| CCAC | Canadian Council on Animal Care |
| CFU | Colony-forming units |
| ge | Genome equivalents |
| H&E | Hematoxylin and eosin |
| LOD | Limit of detection |
| *MAH* | *Mycobacterium avium* subsp. *hominissuis* |
| *MAP* | *Mycobacterium avium* subsp. *paratuberculosis* |
| MLN | Mesenteric lymph node |
| RI-MUHC | Research Institute of the McGill University Health Centre |
| Strep-R | Streptomycin-resistant |
| WT | Wildtype |

## 7. Declarations

*7.1 Ethics approval and consent to participate*

All animal experiments were in accordance with the guidelines of the Canadian Council on Animal Care and all protocols were approved by the animal resource division of the RI-MUHC.

*7.2 Consent for publication*

Not applicable.

*7.3 Availability of data and materials*

All data generated or analyzed during this study are included in this published article and its supplementary information.

*7.4 Competing interests*

The authors declare that they have no competing interests.

*7.5 Funding*

This study was supported by a Foundation Grant from the Canadian Institutes for Health Research (FDN–148362 to Marcel A Behr). Shannon C Duffy was supported by a doctoral training award from the Fonds de Recherche Santé Québec.

*7.6 Authors' contributions*

SCD generated engineered bacterial strains, performed all *in vitro* and *in vivo* experiments, analyzed the data, and wrote the first draft of this manuscript. AL assisted with *in vivo* experiments and study design. YE performed preliminary experiments and assisted with study design. MF and VM assessed histopathology and assisted with study design. MAB directed study design, supervised the project, acquired funding, and contributed to writing this manuscript. All authors reviewed and edited the final version of this manuscript.

*7.7 Acknowledgements*

We would like to thank Fiona McIntosh, Jaryd Sullivan, Sarah Dempsey and Sarah Danchuk for assistance with animal experiments and to Ori Solomon for assistance with analysis of the whole genome sequences. We are very thankful to the animal resource division of the RI-MUHC for the monitoring and care of laboratory mice. We would also like to thank Laura Sly and Susan Menzies for providing positive control slides of $SHIP^{-/-}$ mice and the histology core of the Rosalind and Morris Goodman Cancer Institute at McGill University for processing histopathology slides. Thank you to Corinne Maurice and Eve Beauchemin for sharing protocols for performing the lipocalin-2 assay and to Serge Mostowy for his input in the editing of this manuscript.

## 8. References

1. Twort, F. W. & Ingram, G. L. Y. A Method for Isolating and Cultivating the Mycobacterium enteritidis chronicae pseudotuberculosae bovis, Johne, and some Experiments on the

Preparation of a Diagnostic Vaccine for Pseudo-tuberculous Enteritis of Bovines. *Proceedings of the Royal Society of London Series B* **84**, 517–542 (1912).

2. United States Department of Agriculture. Johne's Disease on U.S. Dairies, 1991–2007. (2008).
3. Imada, J., Kelton, D. & Barkema, H. Epidemiology, Global Prevalence and Economics of Infection. in *Paratuberculosis: Organism, Disease, Control* 1–13 (CABI, 2020).
4. Field, N. L., Mee, J. F. & McAloon, C. G. Characteristics (sensitivity and specificity) of herd-level diagnostic tests for Mycobacterium avium subspecies paratuberculosis in cattle - A systematic review. *Vet J* **279**, 105786 (2022).
5. Rasmussen, P., Barkema, H. W., Mason, S., Beaulieu, E. & Hall, D. C. Economic losses due to Johne's disease (paratuberculosis) in dairy cattle. *J Dairy Sci* **104**, 3123–3143 (2021).
6. Duffy, S. C. & Behr, M. A. Paratuberculosis and crohn's disease. *Paratuberculosis: organism, disease, control* 29–44 (2020) doi:10.1079/9781789243413.0029.
7. Mizzi, R., Plain, K. M., Whittington, R. & Timms, V. J. Global Phylogeny of Mycobacterium avium and Identification of Mutation Hotspots During Niche Adaptation. *Frontiers in Microbiology* **13**, (2022).
8. Alexander, D. C., Turenne, C. Y. & Behr, M. A. Insertion and Deletion Events That Define the Pathogen Mycobacterium avium subsp. paratuberculosis. *J Bacteriol* **191**, 1018–1025 (2009).
9. Coussens, P. *et al.* Host-pathogen interactions and intracellular survival of Mycobacterium avium subsp. paratuberculosis. *Paratuberculosis: organism, disease, control* 120–138 (2020) doi:10.1079/9781789243413.0120.

10. Wu, C. *et al.* Invasion and Persistence of Mycobacterium avium subsp. paratuberculosis during Early Stages of Johne's Disease in Calves. *Infect Immun* **75**, 2110–2119 (2007).
11. Rathnaiah, G. *et al.* Pathogenesis, Molecular Genetics, and Genomics of Mycobacterium avium subsp. paratuberculosis, the Etiologic Agent of Johne's Disease. *Front Vet Sci* **4**, 187 (2017).
12. Whittington, R. J., Marshall, D. J., Nicholls, P. J., Marsh, I. B. & Reddacliff, L. A. Survival and Dormancy of Mycobacterium avium subsp. paratuberculosis in the Environment. *Appl Environ Microbiol* **70**, 2989–3004 (2004).
13. Whittington, R. J., Marsh, I. B. & Reddacliff, L. A. Survival of Mycobacterium avium subsp. paratuberculosis in Dam Water and Sediment. *Appl Environ Microbiol* **71**, 5304–5308 (2005).
14. Talaat, A. M., Wu, C.-W. & Hines II, M. E. Experimental animal models of paratuberculosis. *Paratuberculosis: organism, disease, control* 213–247 (2020) doi:10.1079/9781789243413.0213.
15. Veazey, R. S. *et al.* Histopathology of C57BL/6 mice inoculated orally with Mycobacterium paratuberculosis. *J Comp Pathol* **113**, 75–80 (1995).
16. Cooney, M. A., Steele, J. L., Steinberg, H. & Talaat, A. M. A murine oral model for Mycobacterium avium subsp. paratuberculosis infection and immunomodulation with Lactobacillus casei ATCC 334. *Front Cell Infect Microbiol* **4**, 11 (2014).
17. Eckmann, L. Animal models of inflammatory bowel disease: lessons from enteric infections. *Ann N Y Acad Sci* **1072**, 28–38 (2006).
18. Bouladoux, N., Harrison, O. J. & Belkaid, Y. The mouse model of infection with Citrobacter rodentium. *Curr Protoc Immunol* **119**, 19.15.1-19.15.25 (2017).

19. Barthel, M. *et al.* Pretreatment of Mice with Streptomycin Provides a Salmonella enterica Serovar Typhimurium Colitis Model That Allows Analysis of Both Pathogen and Host. *Infect Immun* **71**, 2839–2858 (2003).
20. Rindi, L. & Garzelli, C. Genetic diversity and phylogeny of Mycobacterium avium. *Infect Genet Evol* **21**, 375–383 (2014).
21. Duffy, S. C. *et al.* Reconsidering Mycobacterium bovis as a proxy for zoonotic tuberculosis: a molecular epidemiological surveillance study. *The Lancet Microbe* **1**, e66–e73 (2020).
22. de la Rua-Domenech, R. Human Mycobacterium bovis infection in the United Kingdom: Incidence, risks, control measures and review of the zoonotic aspects of bovine tuberculosis. *Tuberculosis (Edinb)* **86**, 77–109 (2006).
23. Murphy, K., Papavinasasundaram, K. & Sassetti, C. M. Mycobacterial Recombineering. in *Mycobacterial Protocols* 177–199 (Springer, 2015).
24. Nair, J., Rouse, D. A., Bai, G. H. & Morris, S. L. The rpsL gene and streptomycin resistance in single and multiple drug-resistant strains of Mycobacterium tuberculosis. *Mol Microbiol* **10**, 521–527 (1993).
25. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://doi.org/10.48550/arXiv.1303.3997 (2013).
26. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
27. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178–192 (2013).

28. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at https://doi.org/10.48550/arXiv.1207.3907 (2012).

29. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* **6**, 80–92 (2012).

30. McConnell, E. L., Basit, A. W. & Murdan, S. Measurements of rat and mouse gastrointestinal pH, fluid and lymphoid tissue, and implications for in-vivo experiments. *J Pharm Pharmacol* **60**, 63–70 (2008).

31. Lai, N. Y. *et al.* Gut-Innervating Nociceptor Neurons Regulate Peyer's Patch Microfold Cells and SFB Levels to Mediate Salmonella Host Defense. *Cell* **180**, 33-49.e22 (2020).

32. Rufai, S. B. *et al.* Complete Genome Sequence of Mycobacterium orygis Strain 51145. *Microbiol Resour Announc* **10**, e01279-20 (2021).

33. North, R. J., Ryan, L., LaCource, R., Mogues, T. & Goodrich, M. E. Growth Rate of Mycobacteria in Mice as an Unreliable Indicator of Mycobacterial Virulence. *Infect Immun* **67**, 5483–5485 (1999).

34. Clarke, K.-A. R. *et al.* Experimental inoculation of meadow voles (Microtus pennsylvanicus), house mice (Mus musculus), and Norway rats (Rattus norvegicus) with Mycobacterium bovis. *J Wildl Dis* **43**, 353–365 (2007).

35. Li, L. *et al.* The complete genome sequence of Mycobacterium avium subspecies paratuberculosis. *Proc Natl Acad Sci U S A* **102**, 12344–12349 (2005).

36. McLarren, K. W. *et al.* SHIP-Deficient Mice Develop Spontaneous Intestinal Inflammation and Arginase-Dependent Fibrosis. *Am J Pathol* **179**, 180–188 (2011).

37. Chassaing, B. *et al.* Fecal lipocalin 2, a sensitive and broadly dynamic non-invasive biomarker for intestinal inflammation. *PLoS One* **7**, e44328 (2012).

38. Cangelosi, G. A. *et al.* The Two-Component Regulatory System mtrAB Is Required for Morphotypic Multidrug Resistance in Mycobacterium avium. *Antimicrob Agents Chemother* **50**, 461–468 (2006).

39. Arko-Mensah, J. *et al.* Resistance to mycobacterial infection: a pattern of early immune responses leads to a better control of pulmonary infection in C57BL/6 compared with BALB/c mice. *Vaccine* **27**, 7418–7427 (2009).

40. Schwarz, R., Kaspar, A., Seelig, J. & Künnecke, B. Gastrointestinal transit times in mice and humans measured with 27Al and 19F nuclear magnetic resonance. *Magn Reson Med* **48**, 255–261 (2002).

## 9. Figures

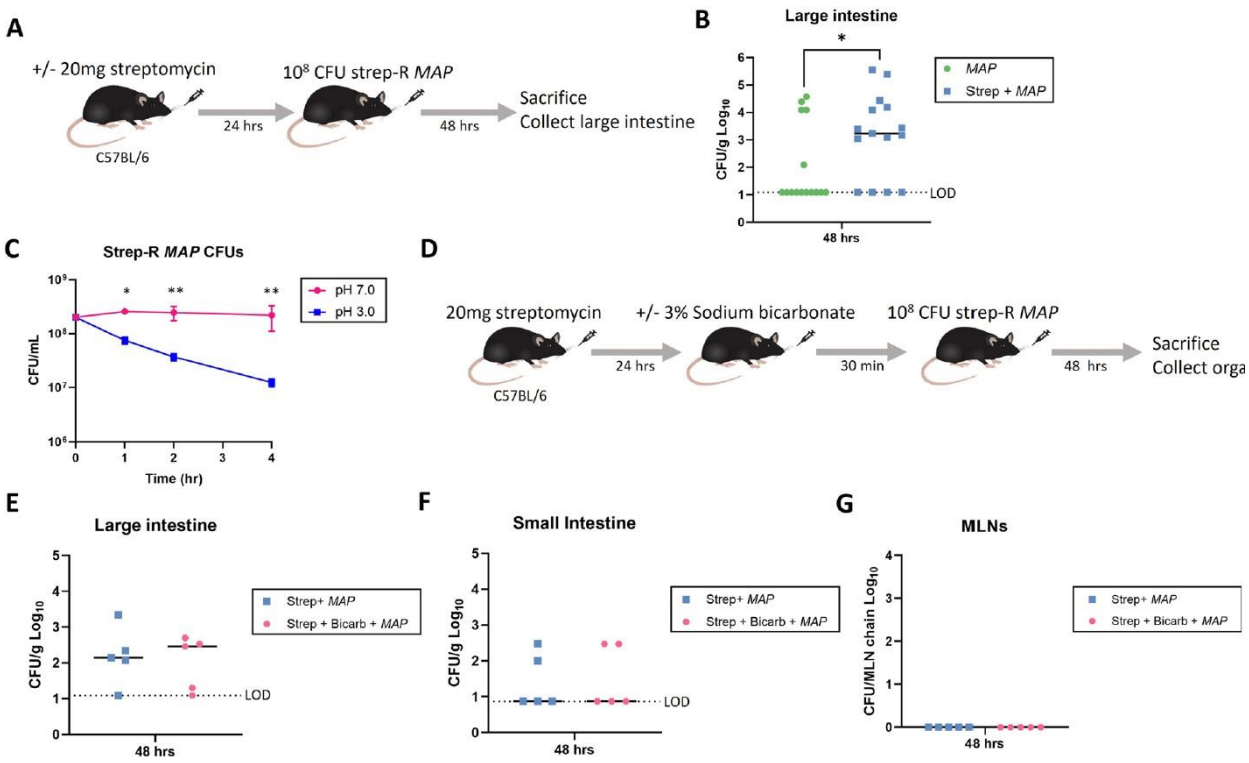



**Figure 1:** The effect of pre-treatment with streptomycin and sodium bicarbonate on *MAP* infection

**A.** C57BL/6 mice were given 20mg of streptomycin or no pre-treatment followed by an oral gavage of $10^8$ CFU strep-R *MAP* 24-hours later. Mice were sacrificed and the large intestines were assessed for CFUs 48-hours post-infection. **B.** Large intestine CFUs were compared between mice given streptomycin pre-treatment or no pre-treatment (*p<0.05). **C.** Strep-R *MAP* was subject to a pH of 7.0 or 3.0 for 1, 2, or 4 hours. The CFUs following incubation at either pH was plotted over time (*p<0.05, **p<0.01). **D.** Mice were orally given 20mg of streptomycin. The next day half of the mice were given sodium bicarbonate followed by a gavage of $10^8$ CFU strep-R *MAP* in all mice 30 minutes later. Mice were sacrificed and their organs were assessed for CFUs 48-hours post-infection. **E-G.** The *MAP* CFUs of the large intestine (**E**), small intestine

(**F**), and MLNs (**G**) were compared between mice given sodium bicarbonate and streptomycin pre-treatment or only streptomycin pre-treatment.

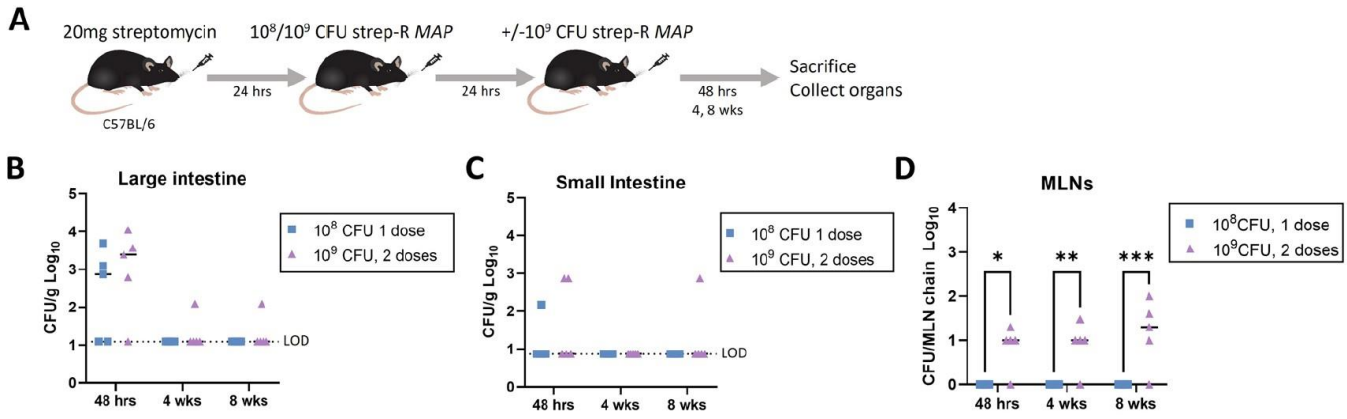



**Figure 2:** Comparison of a single dose of $10^8$ CFU strep-R *MAP* and 2 consecutive doses of $10^9$ CFU strep-R *MAP*

**A.** Mice were pre-treated with 20mg of streptomycin. The following day, mice were given either $10^8$ or $10^9$ CFU strep-R *MAP.* Mice that were inoculated with the higher dose were given a second dose of $10^9$ CFU strep-R *MAP* 24-hours later. Mice were euthanized 48-hours, 4-weeks, or 8-weeks post infection and the organs were assessed for CFUs. **B-D.** Strep-R *MAP* CFUs were compared between the lower and higher dose infection strategies in the large intestine (**B**), small intestine (**C**), and MLNs (**D**) (*p<0.05, **p<0.01, ***p<0.001).

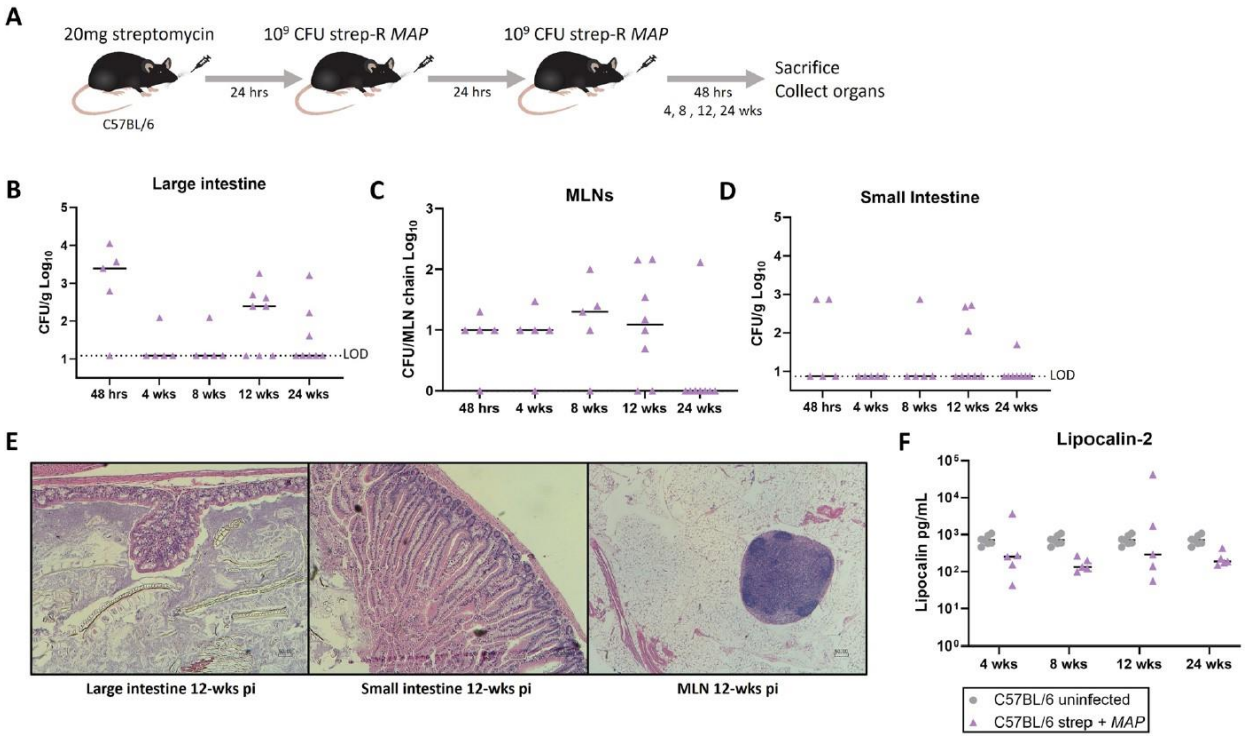



**Figure 3:** Assessment of chronic infection and disease following streptomycin pre-treatment and *MAP* gavage

**A.** Mice were pre-treated with 20mg streptomycin followed by 2 consecutive doses of $10^9$ CFU strep-R *MAP* given 24-hours apart. Mice were euthanized and assessed at 48-hours, 4-, 8-, 12-, and 24-weeks post-infection. **B-D.** Strep-R *MAP* CFUs were assessed in the large intestine (**B**), MLNs (**C**), and small intestine (**D**) throughout the 24-week experiment. **E.** Representative sections of the large intestine (H&E, original magnification x100), small intestine (x100), and MLNs (x40) from mice 12-weeks post *MAP* infection are shown. **F.** Fecal pellets were obtained from uninfected mice and strep-R *MAP* infected mice 4-, 8-, 12-, and 24-weeks post-infection and processed. The amount of lipocalin-2 found in samples from both groups was evaluated by ELISA.

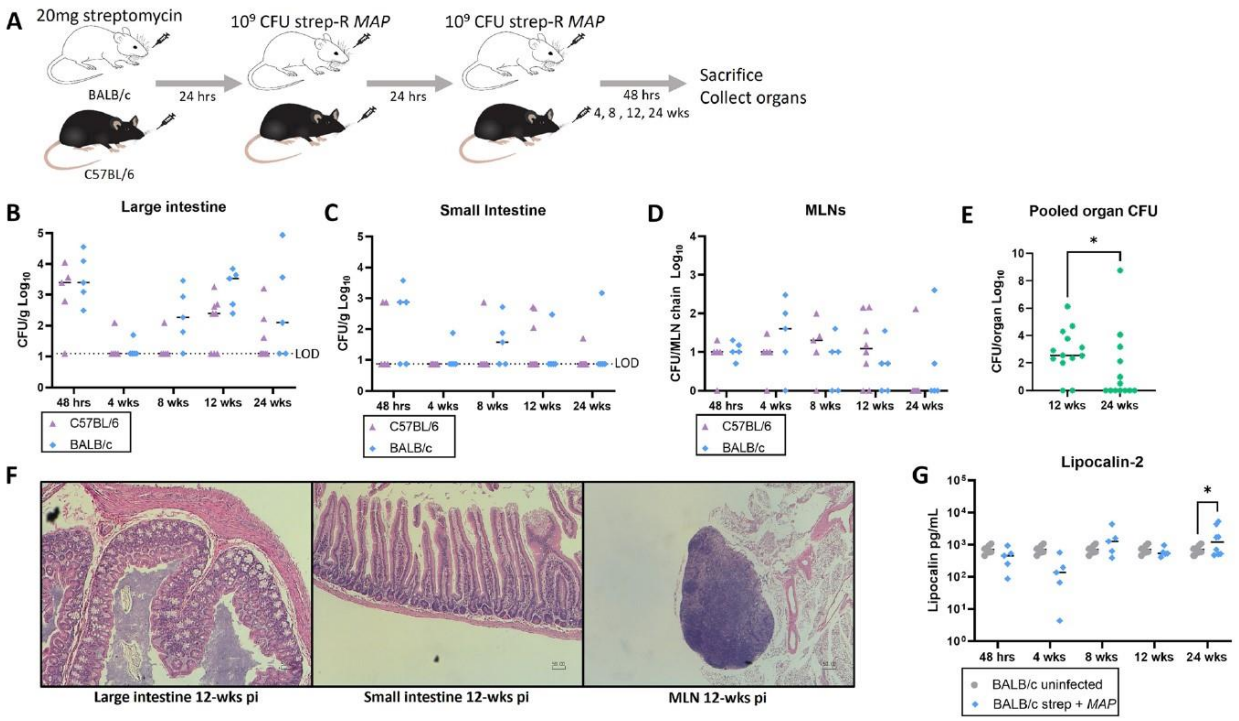



**Figure 4:** Comparison of infection outcomes between C57BL/6 and BALB/c mice

To evaluate whether using a more susceptible mouse strain would induce disease following strep-R *MAP* infection, the infection model was compared with BALB/c mice. **A.** Either C57BL/6 or BALB/c mice were given 20mg streptomycin followed by 2 consecutive doses of $10^9$ CFU strep-R *MAP.* Mice were sacrificed and assessed at 48-hours, 4-, 8-, 12-, and 24-weeks post-infection. **B-D.** Strep-R *MAP* CFUs were compared between the 2 mice strains in the large intestine (**B**), small intestine (**C**), and MLNs (**D**). **E.** Total *MAP* burden was pooled for the large intestine, small intestine, and MLNs of C57BL/6 and BALB/c mice and compared at 12- and 24-weeks post-infection (*$p<0.05$) **F.** Representative sections of the large intestine (H&E, original magnification x100), small intestine (x100), and MLNs (x40) from BALB/c mice 12-weeks post *MAP* infection are shown. **G.** Fecal pellets were obtained from uninfected and strep-R *MAP* infected BALB/c mice 4-, 8-, 12-, and 24-weeks post infection and processed. The amount of lipocalin-2 found in samples from both groups were evaluated by ELISA (*$p<0.05$).

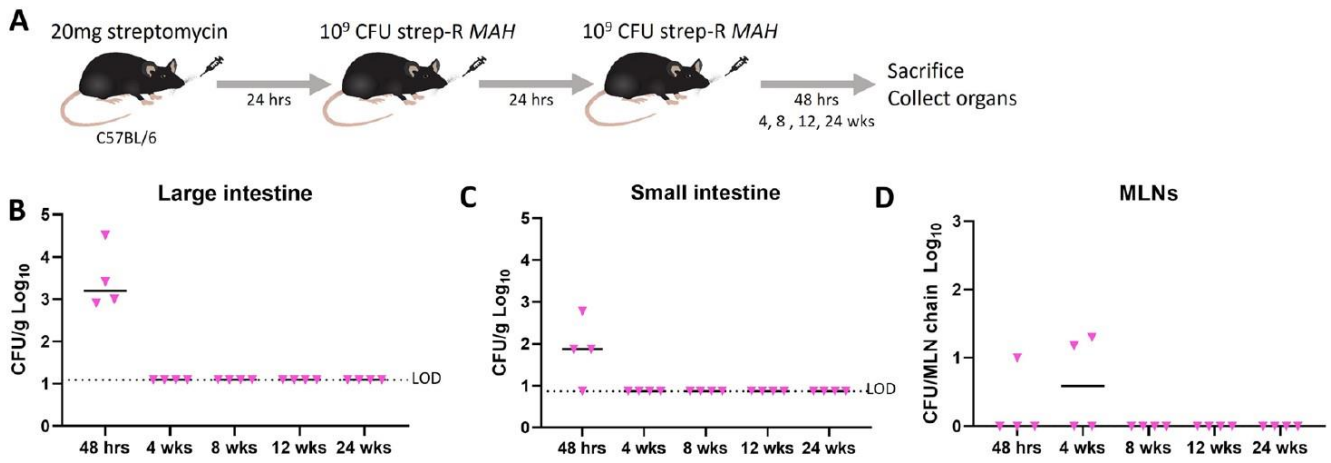



**Figure 5:** Comparison of infection model with *MAH*

To determine how infection outcomes would compare using an environmental mycobacterium, the same infection strategy was repeated using strep-R *MAH.* **A.** C57BL/6 mice were given 20mg of streptomycin followed by 2 consecutive doses of strep-R *MAH* 24-hours apart. Mice were euthanized 48-hours, 4-, 8-, 12-, and 24-weeks post-infection and organs were assessed for CFUs. **B-D.** Strep-R *MAH* CFUs were assessed in the large intestine (**B**), small intestine (**C**), and MLNs (**D**) at each timepoint.

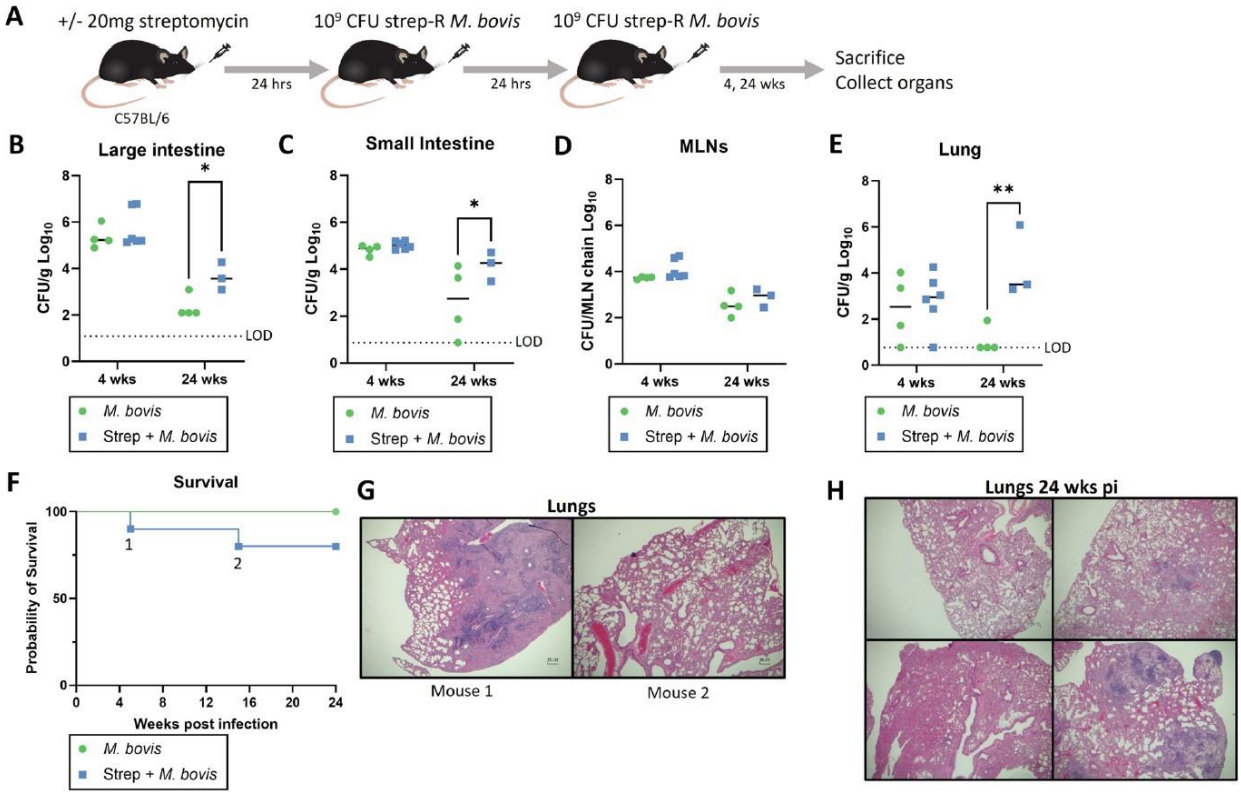
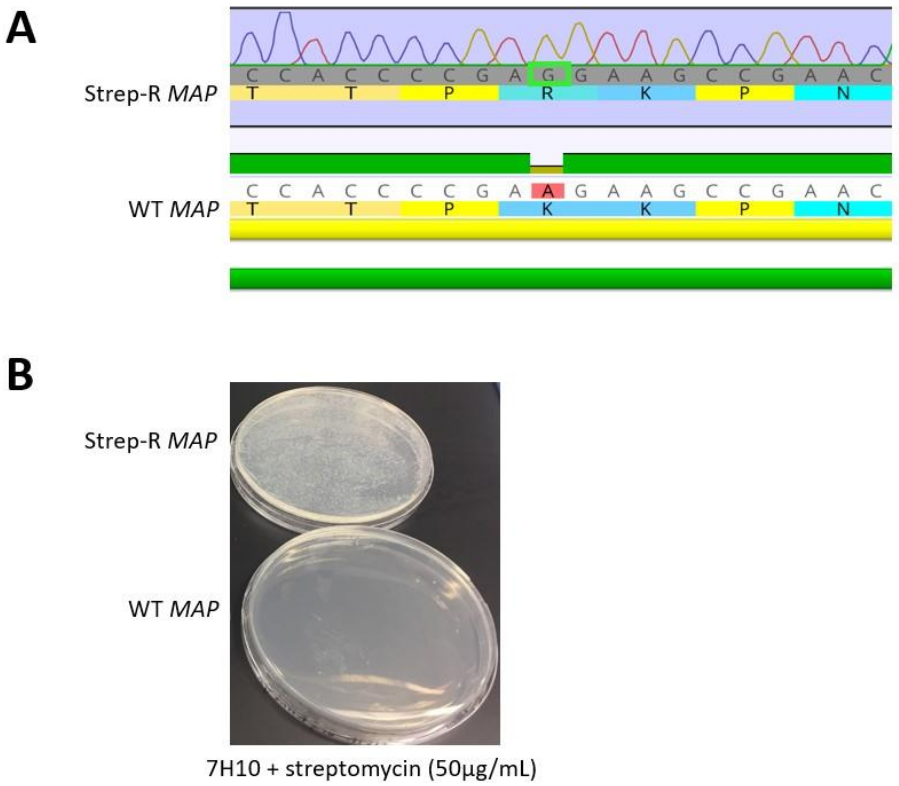



**Figure 6:** Comparison of infection model with *M. bovis*

To determine how infection outcomes may change using a pathogenic tuberculous mycobacterium, the same infection strategy was repeated using strep-R *M. bovis.* **A.** C57BL/6 mice were given 20mg of streptomycin or no pre-treatment followed by 2 consecutive doses of strep-R *M. bovis* 24-hours apart. Mice were sacrificed 4- and 24-weeks post-infection and organs were assessed for CFUs. **B-D.** Strep-R *M. bovis* CFUs were assessed between mice given streptomycin pre-treatment or no pre-treatment in the large intestine (**B**), small intestine (**C**), MLNs (**D**) and lungs (**E**) at each timepoint (*p<0.05, **p<0.01). **F.** Survival was monitored between groups over the 24-week experiment. Two mice were flagged for euthanasia at 5-weeks (mouse 1) and 15-weeks (mouse 2) post-infection **G.** The histopathology of the lungs was assessed for the 2 mice that reached endpoint. Shown are representative images of sections of their lungs (H&E, original magnification x40). **H.** At the experimental endpoint of 24-weeks, the

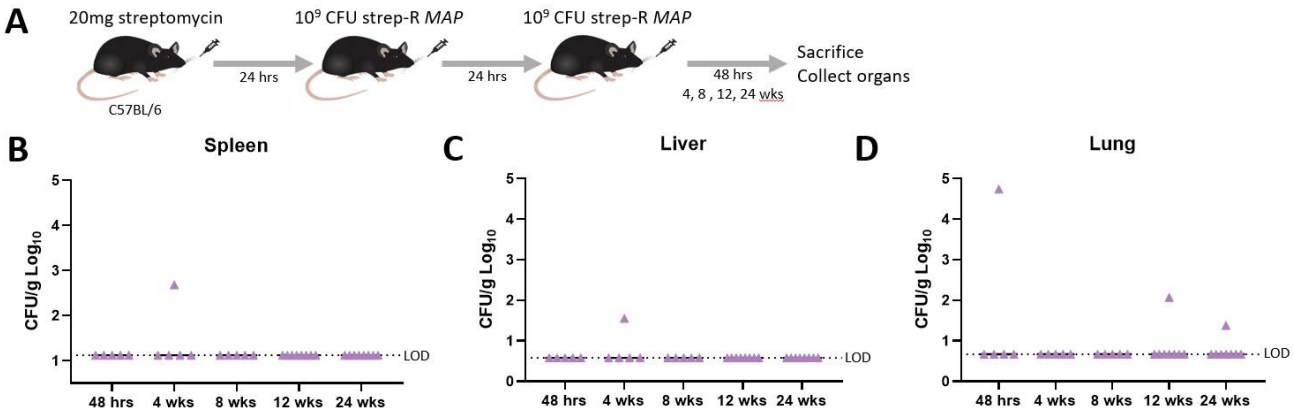
histopathology of the lungs was assessed for the surviving mice. Representative images of the lungs sections (x40) from streptomycin pre-treated mice are pictured.

## 8. Supplementary Figures



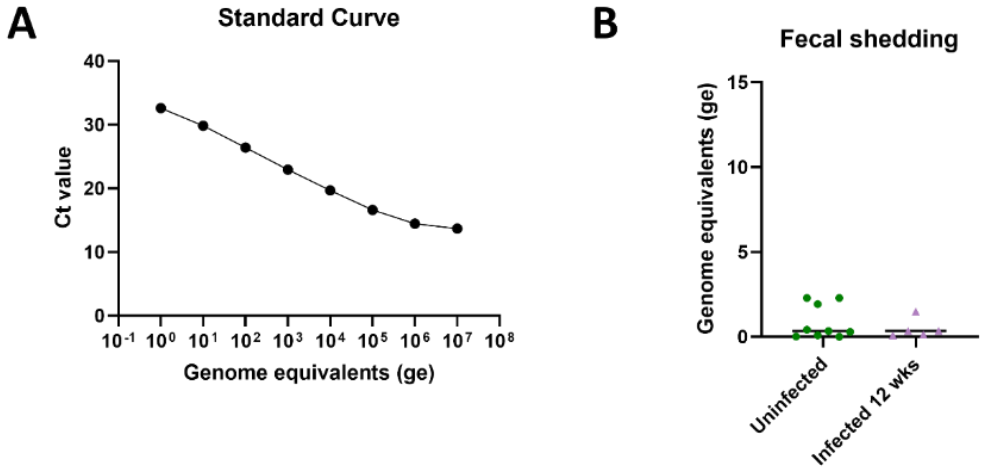

**Supplementary Figure 1:** Generation of strep-R *MAP*

Oligo-mediated recombineering was employed to generate a K43R mutation in *rpsL* of *MAP* which is known to confer resistance to streptomycin. **A.** The sequence of the *rpsL* gene of wildtype (WT) *MAP* and strep-R *MAP* was generated by Sanger sequencing. Amino acids 40-46 of *MAP rpsL* are visualized on Geneious Prime and the point mutation in strep-R *MAP* is highlighted by a green box. **B.** The WT *MAP* and strep-R *MAP* strains were plated on 7H10 agar with streptomycin to compare growth.



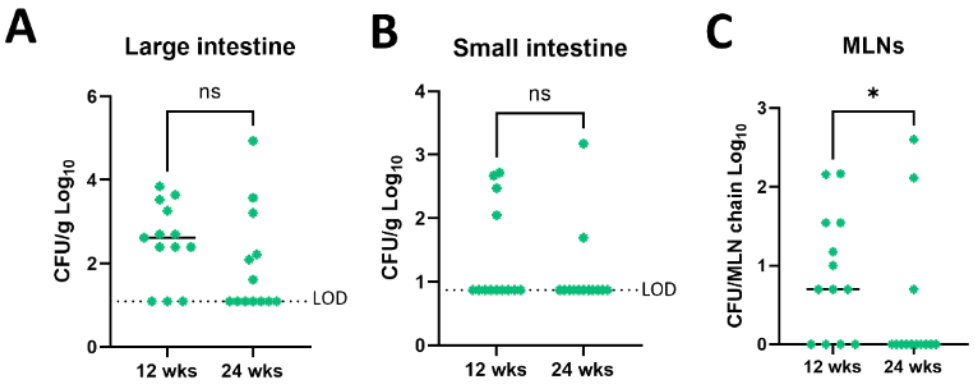

**Supplementary Figure 2:** Dissemination of strep-R *MAP*

**A.** C57BL/6 mice were given 20mg of streptomycin followed by 2 consecutive doses of $10^9$ CFU strep-R *MAP* each 24-hours apart. **B-D.** Dissemination of *MAP* into the spleen (**B**), liver (**C**), and lungs (**D**) was evaluated at 48-hours, 4-, 8-, 12-, and 24-weeks post-gavage.



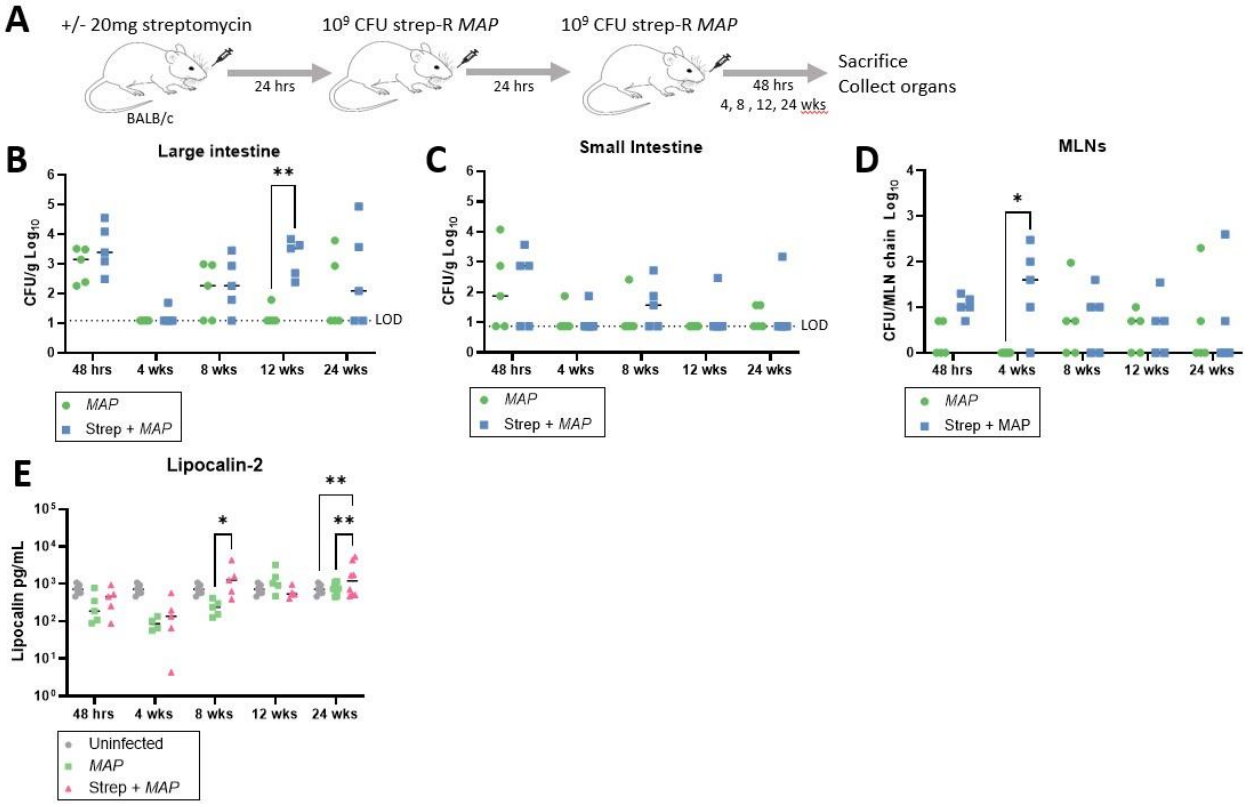

**Supplementary Figure 3:** Fecal shedding assessment

**A.** A standard curve was prepared for quantitative PCR of the F57 gene using *MAP* K10 genomic DNA diluted from $1x10^7$ to 1 genome equivalents in order to interpolate values from fecal samples. **B.** Fecal shedding was assessed in uninfected controls and mice 12-weeks post-gavage with *MAP.*



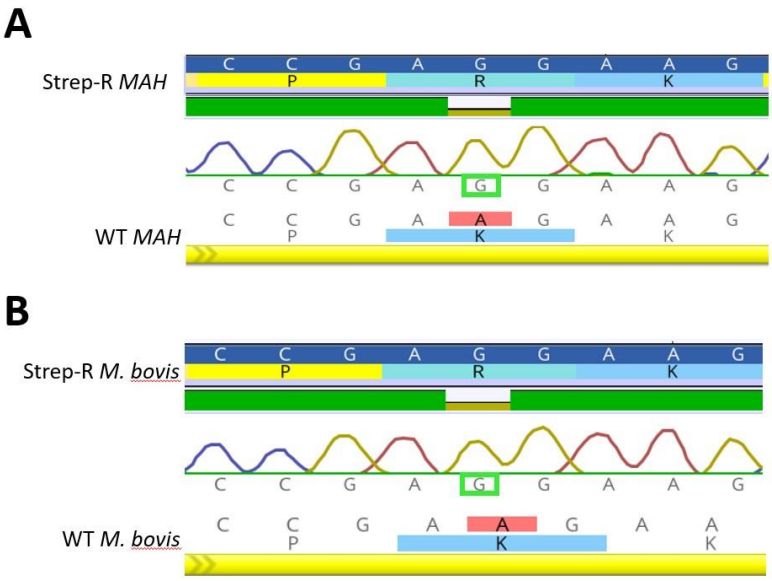

**Supplementary Figure 4:** Pooled organ CFUs of C57BL/6 and BALB/c mice at 12- and 24-weeks post-infection

The organ CFUs of the large intestine (**A**), small intestine (**B**), and MLNs (**C**) were pooled from C57BL/6 and BALB/c mice and compared between 12- and 24-weeks post infection (*$p<0.05$).



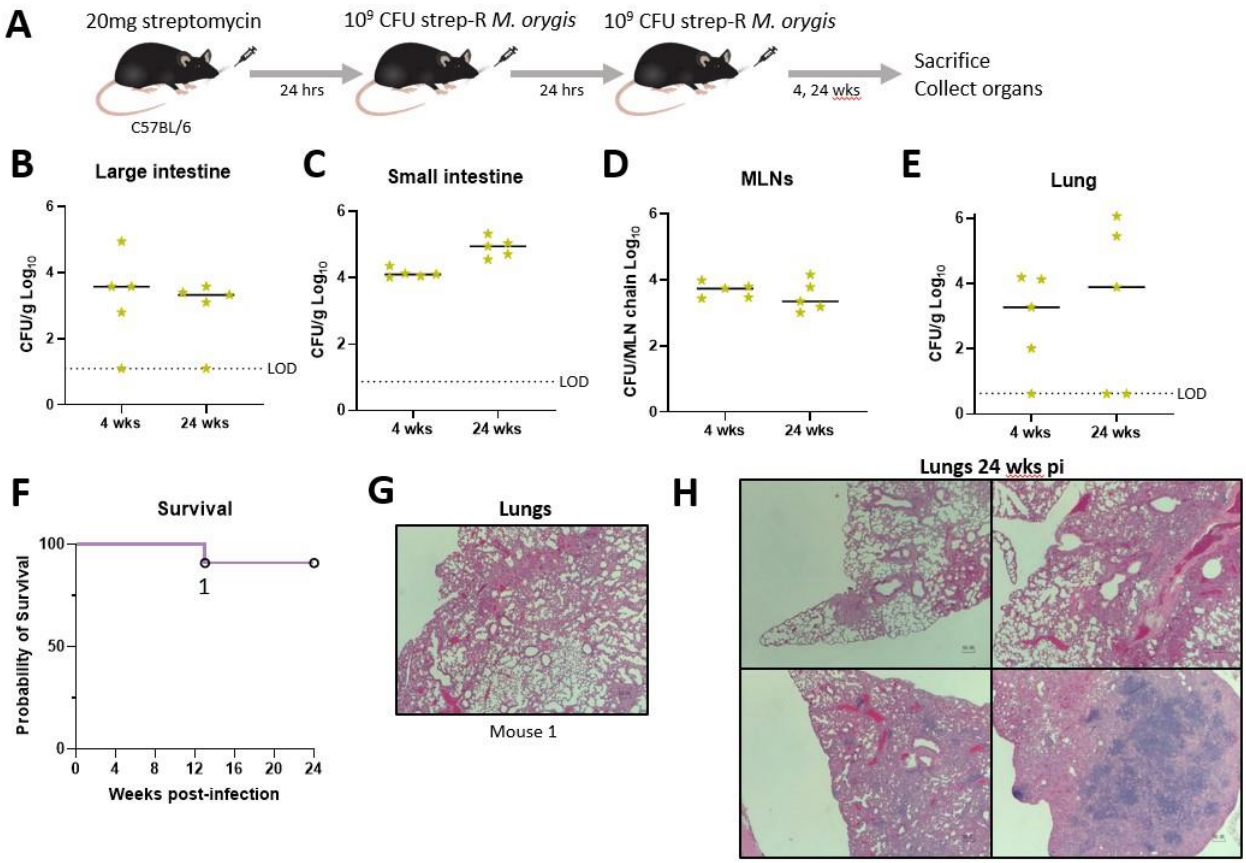

**Supplementary Figure 5:** Comparison of streptomycin pre-treatment and no pre-treatment in BALB/c mice

To determine whether streptomycin pre-treatment would also increase infection in BALB/c mice, infection outcomes were compared between BALB/c mice given streptomycin pre-treatment or no pre-treatment. **A.** BALB/c mice were given 20mg streptomycin or no pre-treatment followed by 2 consecutive doses of $10^9$ CFU strep-R *MAP* each 24-hours apart. **B-D.** The CFUs of the large intestine (**B**), small intestine (**C**), and MLNs (**D**) were compared between mice groups 48-hours, 4-, 8-, 12-, and 24-weeks post-infection. **E.** Fecal pellets were collected and processed from uninfected BALB/c mice and *MAP*-infected BALB/c mice with or without streptomycin pre-treatment. The levels of lipocalin-2 found in the feces of each group were evaluated by ELISA (*p<0.05, **p<0.01).




**Supplementary Figure 6:** Generation of strep-R *MAH* and strep-R *M. bovis*

Oligo-mediated recombineering was employed to generate a K43R mutation in *rpsL* of *MAH* and *M. bovis*. **A.** The sequence of the *rpsL* gene of WT *MAH* and strep-R *MAH* was generated by Sanger sequencing. Amino acids 42-44 of *MAH* rpsL were visualized on Geneious Prime and the point mutation in strep-R *MAH* is indicated by a green box. **B.** The sequence of the *rpsL* gene of WT *M. bovis* and strep-R *M. bovis* was generated by Sanger sequencing. Amino acids 42-44 of *M. bovis* rpsL were visualized on Geneious Prime and the point mutation in strep-R *M. bovis* is highlighted by a green box.




**Supplementary Figure 7:** Comparison of infection model with *M. orygis*

The streptomycin pre-treatment model was repeated with another pathogenic bovine mycobacteria *M. orygis.* **A.** C57BL/6 mice were given 20mg of streptomycin followed by 2 consecutive doses of strep-R *M. orygis* 24-hours apart. Mice were euthanized 4- and 24-weeks post-infection. **B-D.** Strep-R *M. orygis* CFUs were assessed in the large intestine (**B**), small intestine (**C**), MLNs (**D**) and lungs (**E**) at each timepoint. **F.** Survival was monitored over the 24-week experiment. One mouse was flagged for euthanasia at 13-weeks post-infection (mouse 1). **G.** The histopathology of the lungs was assessed for mouse 1. Shown is a representative image of the H&E-stained section of its lungs (original magnification x40). **H.** At the experimental endpoint of 24-weeks, the histopathology of the lungs was assessed for the surviving mice. Shown are representative images of the lungs (x40) sections.

## 9. Supplementary Table

| Name | Sequence | Function |
|---|---|---|
| Kan_F | GAAAGCCACGTTGTGTCTCA | To confirm presence of pNit plasmid |
| Kan_R | ATTCCGACTCGTCCAACATC | To confirm presence of pNit plasmid |
| MAP_rpsL_K43R | CGCGGGCGACCTTCCGAAGCGCCGAGTTCGGCTTC CTCGGGGTGGTGGTGTACACGCGGGTGCATACGCC | To generate strep-R *MAP* |
| MAH_rpsL_K43R | CGCGGGCGACCTTCCGAAGCGCCGAGTTCGGCTTC CTCGGGGTGGTGGTGTACACGCGGGTGCATACGCC | To generate strep-R *MAH* |
| Mbovis_rpsL_K43R | CGCGGGCAACCTTCCGAAGCGCCGAGTTCGGCTTC CTCGGAGTGGTGGTGTACACGCGGGTGCATACACC | To generate strep-R *M. bovis* |
| MACrpsL_F | ACAAGATCGGCAAGGTCAAG | To confirm *rpsL* K43R mutation by Sanger sequencing |
| MACrpsL_R | GGATGATCTTGTACCGCACAC | To confirm *rpsL* K43R mutation by Sanger sequencing |
| MTBCrpsL_F | GCCGACAAACAGAACGTG | To confirm *rpsL* K43R mutation by Sanger sequencing |
| MTBCrpsL_R | GTGACCAACTGCGATCCGTAG | To confirm *rpsL* K43R mutation by Sanger sequencing |
| F57_F | TACCGAATGTTGTTGTCACCG | To test for fecal shedding of *MAP* |
| F57_R | TGGCACAGACGACCATTCAA | To test for fecal shedding of *MAP* |

**Supplementary Table 1:** Primer and oligo sequences

# CHAPTER V

# Discussion

---

## 1. Discussion

In this thesis, I describe my research investigating both the epidemiology and pathogenesis of bovine mycobacteria. I first aimed to generate tools to rapidly identify MTBC subspecies to determine the prevalence and cause of zoonotic TB in South Asia. My work emphasized the importance of *M. orygis,* rather than *M. bovis*, as the primary cause of bovine and zoonotic TB in India and the other countries of South Asia which make up ~38% of the world's TB cases[23]. This study also highlighted the need for accessible and reliable methods of identifying BCG disease, which led to my development of a second assay which was implemented into a clinic in Vellore. Finally, I sought to better understand the pathogenesis of bovine mycobacteria by generating an experimental model which would mimic the disease observed upon infection with *MAP* (paratuberculosis), or with *M. bovis* or *M. orygis* (zoonotic TB). I found that a high dose streptomycin pre-treatment model of oral infection with *MAP* led to persistent localized intestinal infection and its use with *M. bovis* or *M. orygis* led to pulmonary dissemination and disease. Taken together, this thesis demonstrates how epidemiological studies can inform research conducted at the bench, and conversely, how the tools developed at the bench can be translated into clinical application. This research has produced immediate products (novel methods of detecting infection with bovine mycobacteria in the clinic) as well as generating fundamental knowledge regarding the pathogenesis of intestinal mycobacterial infections.

The completion of the work described in chapter II sparked interest in zoonotic TB and initiated several new research opportunities to continue investigations in South Asia. One of the key questions which emerged from this work was where does the transition from *M. orygis* to *M. bovis* occur geographically? If the distribution of *M. bovis* in Africa, Europe, and the Americas and *M. orygis* in South Asia resulted from 2 independent cattle domestication events of *Bos taurus* and *Bos indicus* respectively as has been previously suggested[26], then in what location would one expect to see a transition from *M. bovis* to *M. orygis* infections or a mix of the 2? This question led to a collaboration with Dr. Mamoona Chaudhry and Dr. Rubab Maqsood where we investigated the MTBC subspecies identified from gross TB-like lesions found in buffalo and cattle at a slaughterhouse in Punjab, Pakistan (manuscript in revision). I used the same assays described in chapter II to identify the MTBC subspecies causing disease from DNA extracted from these lesions. Our study identified 20 cases of bovine TB caused by *M. orygis* (n=10) and *M. tuberculosis* (n=10), but again no cases were caused by *M. bovis.* Larger studies on more cases of bovine and zoonotic TB should be conducted in Pakistan to confirm these findings. However, this study suggested that the transition to disease caused by *M. bovis* may be occurring in countries further west such as Iran.

The research outlined in chapter II also led to the formation of the Estimation and Risk Assessment of the burden of Zoonotic Tuberculosis (ERAzTB) group. This is a collection of researchers, clinicians, and veterinarians within South Asia, North America, and Europe who meet weekly to collaborate and present on projects related to bovine and zoonotic TB. One of the key objectives of this group was to obtain a bigger dataset of positive TB cultures from India which was representative of all regions within the country, rather than being obtained from a single clinic. The results of this large collaborative study are upcoming, but they continue to

support that zoonotic TB in India does occur, albeit infrequently, and when it occurs, it is caused by *M. orygis,* not *M. bovis*.

In chapter III, I described my development of an assay to detect BCG disease and differentiate between early and late BCG strains. This assay was designed keeping in mind the lessons learned in chapter II about screening samples in a high-volume clinical laboratory. I found that RD regions can vary amongst isolates leading to the WGS of several isolates in our study which were originally labeled as inconclusive due to their being negative for either RD12 or RD1 but retaining the other RDs consistent with *M. tuberculosis sensu stricto.* I also found that conventional PCR assays can be more laborious and space-consuming due to the need for a gel. A closed system which does not require electrophoresis and photo steps was determined to be a more reliable and practical option. Upon learning that the work described in chapter II had provided a means of identifying BCG disease in patients where previously there was not one available, we set out to develop an improved BCG assay which built upon our findings. I designed a SNP-based real-time PCR assay which avoided the use of RDs and used a rule-in approach to identify BCG while simultaneously providing additional important clinical information (e.g. pyrazinamide resistance, BCG strain).

In chapter IV, I found that zoonotic TB could be modeled by a streptomycin pre-treatment model of gavage with either *M. bovis* or *M. orygis* where oral exposure led to infection of the intestines and mesenteric lymph nodes, following which there is dissemination to the lungs resulting in a slow-progressing lung disease. Considering that the model used 2 very high doses of *M. bovis* and *M. orygis* with streptomycin pre-treatment and there was no evidence of enteritis after 24-weeks, this suggests that the mice strain we used are quite resistant to intestinal disease with these bacteria. It has been previously shown that the same mice infected with *M. bovis* by

aerosol can rapidly develop pulmonary disease and suffer high rates of early mortality when compared to *M. tuberculosis*[297,315]. Unpublished data from the Behr lab has shown that a similar phenomenon occurs post-aerosol with *M. orygis.* Yet following 2 doses by gavage that are 5,000,000 times greater than infections by aerosol, only 1-2 mice died over the course of 24-weeks. Despite there being bacteria in the lungs as early as 4-weeks post-infection, only some of the mice displayed varied degrees of lung disease from mild to more severe at the experimental endpoint of 24-weeks post-infection. This result demonstrated that the route of inoculation affects the outcome of infection and suggests that the location and types of initial interactions between the host and the pathogen affect the host's ability to contain the bacteria and combat disease progression.

Given the need for streptomycin pre-treatment and such high doses of zoonotic TB pathogens to induce zoonotic TB-like disease in mice, one may ask whether additional host or environmental factors must be involved for zoonotic TB to occur in humans. We cannot be certain that the results obtained in mice would translate to humans, therefore we cannot rule out that these results are host specific. However, it seems unlikely that humans are exposed to such high levels of *M. bovis* or *M. orygis* when drinking unpasteurized milk. In addition, not everyone who drinks unpasteurized milk contaminated with these pathogens develops zoonotic TB – in fact, the great majority of people living in parts of the world where there is no bovine TB control program and dairy products are not pasteurized do not develop zoonotic TB. Collectively, this suggests that host-specific factors play a role in who progresses to disease following exposure. In human infection with *M. tuberculosis sensu stricto* only about 5-10% of those infected will develop active TB disease[316]. In the case of zoonotic TB, one can infer that a much smaller proportion of people drinking unpasteurized milk contaminated with *M. orygis* or *M. bovis*

progress to disease. This indirectly suggests that additional factors, whether genetic or acquired, are required for disease to ensue.

Considering that 2 high doses of *M. bovis* or *M. orygis* by gavage led to a large reduction in the severity of infection outcomes compared to aerosol, one of the additional questions that we explored was whether a previous oral exposure could be protective against a future aerosol challenge. I performed an infection where I gave mice a single reduced gavage dose of $10^8$ CFU of strep-R *M. orygis* with or without streptomycin pre-treatment followed by an aerosol challenge of ~200 CFUs of wildtype *M. orygis* 24-weeks post-gavage (Figure 1). Following the initial gavage, all mice survived and gained weight and no mice had clinical signs of disease (Figure 2). When mice were challenged by aerosol, the median survival time of naïve mice was 3.5 weeks whereas the median survival time of those that were previously exposed was 20.5 weeks (Figure 3). When comparing streptomycin pre-treatment with no pre-treatment, the median survival of mice previously given *M. orygis* by gavage without streptomycin pre-treatment was 16 weeks and the median survival of mice previously given *M. orygis* by gavage with streptomycin pre-treatment was 23 weeks (Figure 4). The mice previously exposed to *M. orygis* had a significant increase in survival time compared to naïve (**p=0.002) and within a subset analysis, mice given streptomycin pre-treatment had a significant increase in survival time compared to those given no streptomycin pre-treatment (*p=0.025). Since we have seen that streptomycin pre-treatment has led to a significant increase in organ CFUs, this suggests that an increase in CFUs led to a more robust host response which was more protective against a future challenge. One interesting idea suggested by this data is whether people who drink unpasteurized milk are gaining some protection from contracting TB, if they are exposed to *M. tuberculosis* through aerosols. In the future, it may be interesting to determine whether previous gavage by *M.*

*orygis* or *M. bovis* is protective against aerosol challenge with *M. tuberculosis* and investigate how this compares with the current vaccine strain BCG.

When the streptomycin pre-treatment high dose infection strategy was applied to *MAP,* dissemination did not occur. There was however a persistent infection of the gastrointestinal tract which peaked at 12 weeks and persisted in some animals to 24 weeks. Despite the ongoing infection, there was no evidence of disease, detected either by histopathology or by ELISA of the feces for the inflammatory marker lipocalin-2. These findings begged the question as to whether *MAP* disease could be modeled in mice or whether a change in host species is necessary. To answer this, I have recently repeated the streptomycin pre-treatment and high dose strep-R *MAP* infection in mice which are lacking the receptor for IFN-γ (IFN-γR$^{-/-}$ mice). IFN-γ is a cytokine that is key to mounting proper immune responses to intracellular pathogens such as *MAP*[317]. As previously discussed, the human immunodeficiency disease MSMD is caused by mutations causing improper secretion of IFN-γ and is associated with high susceptibility to MTBC and NTM infections[140]. It is well-known that this mouse strain is also highly susceptible to infection with mycobacteria[318].

The weights of these mice were monitored over the course of 12-weeks before all mice were sacrificed, and their organs were collected for both histopathology and CFUs. We chose a timepoint of 12-weeks since this was found to be the peak of organ infection in C57BL/6 mice. The mice weights were stable throughout the experiment and the mice did not develop any signs of clinical disease (Figure 5). Upon sacrifice, the large intestine of approximately half of the mice appeared enlarged compared to uninfected controls, but otherwise appeared healthy (Figure 6). A qPCR for the F57 gene did not reveal any evidence of fecal shedding of *MAP* in these mice nor was there evidence of inflammation detected by ELISA for lipocalin-2 (Figure 7).

Assessment of the histopathology of the small intestines of infected IFN-γR$^{-/-}$ mice suggested some sites of potential mild inflammation which will continue to be investigated by the Behr lab (Figure 8). Notable differences were observed in the organ CFUs of infected IFN-γR$^{-/-}$ mice (Figure 9). Within the intestines, there was a shift in the location of *MAP* CFUs, where IFN-γR$^{-/-}$ mice primarily had CFUs in the small intestine whereas C57BL/6 mice more consistently had CFUs in the large intestine. There were also significantly higher CFUs in both the small intestine and MLNs of IFN-γR$^{-/-}$ mice compared to C57BL/6. Dissemination of *MAP* to the spleen, liver, and lungs remained minimal. There were no significant differences in the dissemination of *MAP* between C57BL/6 and IFN-γR$^{-/-}$ mice (Figure 10). In animals with paratuberculosis, *MAP* is primarily found in the terminal ileum of the small intestine and the MLNs[168]. Thus, an IFN-γR$^{-/-}$ mouse infection seems to more closely mimic natural *MAP* infection. This model will continue to be developed within the Behr lab to explore whether reproducible *MAP*-induced enteritis can be achieved using this mouse strain.

## 2. Figures

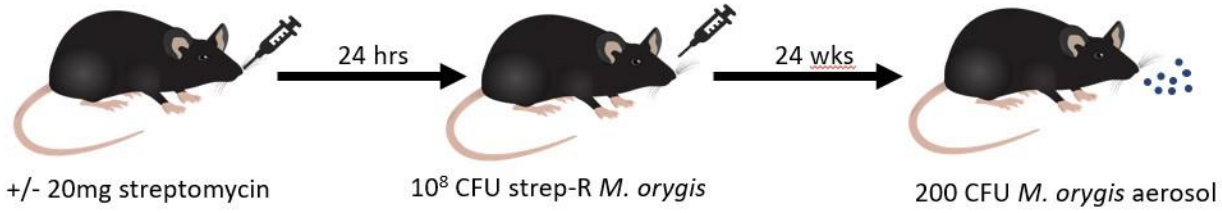



**Figure 1** Oral gavage of mice with strep-R *M. orygis* followed by aerosol challenge with wildtype *M. orygis*

C57BL/6 mice were given either 20mg of streptomycin or no pre-treatment followed 24-hours later by a gavage of $10^8$ CFU of strep-R *M. orygis.* After 24-weeks, all mice were challenged with 200 CFU of wildtype *M. orygis* by aerosol and monitored for survival.

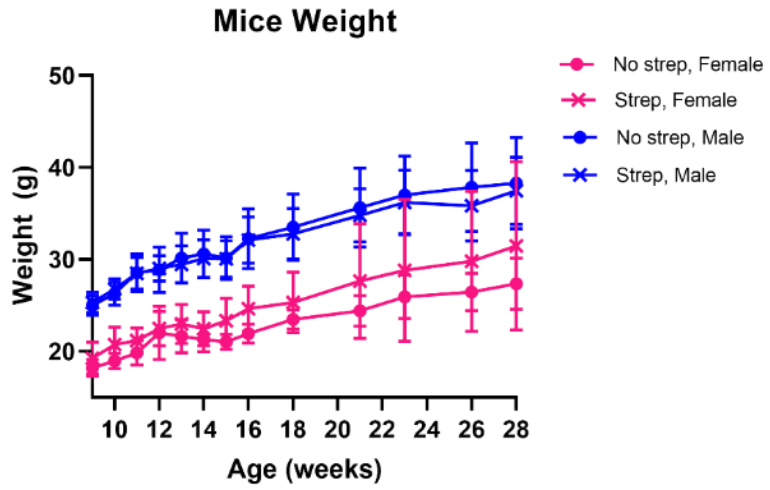



**Figure 2** Mice weights post-gavage with strep-R *M. orygis*

After mice were exposed to strep-R *M. orygis* gavage with or without streptomycin pre-treatment, their weights were monitored over 24-weeks.

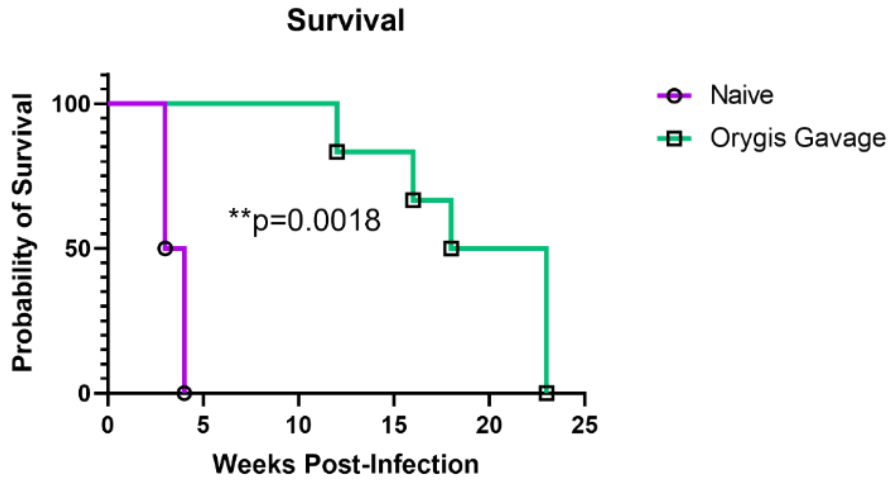



**Figure 3** Survival curve of naïve mice and mice previously-gavaged with strep-R *M. orygis* post-aerosol challenge

The survival curves were compared between naïve mice and all mice given a previous gavage of strep-R *M. orygis* prior to aerosol challenge with wildtype *M. orygis.* The 2 survival curves were compared using a log-rank (Mantel-Cox) test.

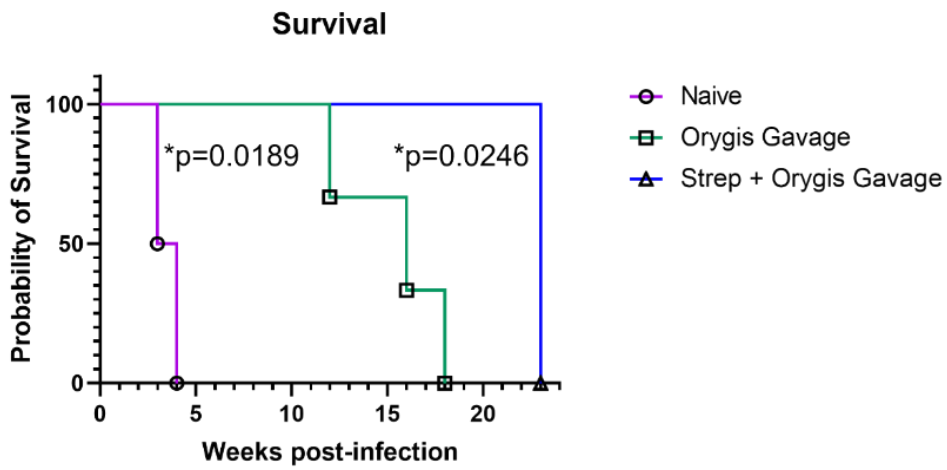



**Figure 4** Survival curve of naïve mice and mice previously-gavaged with strep-R *M. orygis* with or without streptomycin pre-treatment post-aerosol challenge

The survival curves were compared between naïve mice and mice given a previous gavage of strep-R *M. orygis* with or without streptomycin pre-treatment prior to aerosol challenge with wildtype *M. orygis.* Survival curves were compared using a log-rank (Mantel-Cox) test.

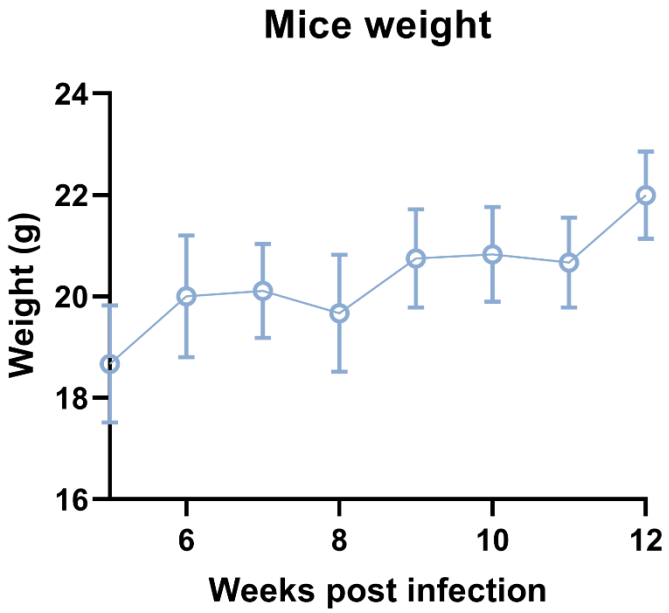



**Figure 5** IFN-γR$^{-/-}$ mice weights following gavage with strep-R *MAP*

The weights of C57BL/6 mice were monitored for 12-weeks following streptomycin pre-treatment and 2 consecutive oral doses of $10^9$ CFU strep-R *MAP* spaced 24-hours apart.

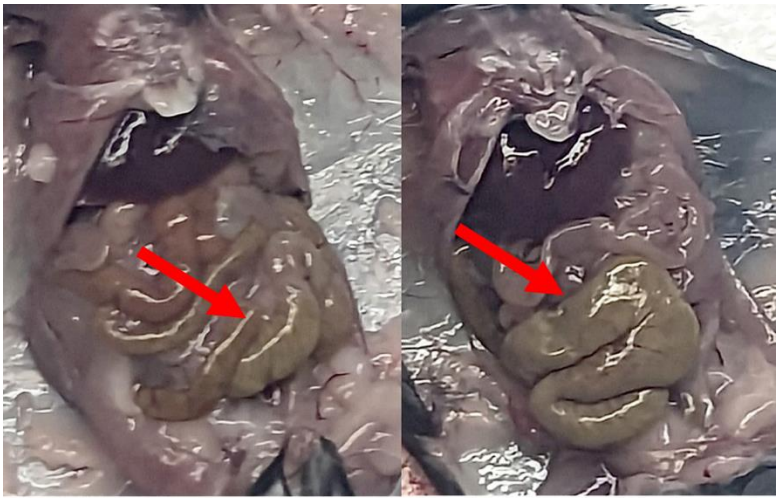



**Figure 6** Comparison of large intestine sizes of IFN-γR$^{-/-}$ mice 12-weeks post-infection with strep-R *MAP*

The intestines of uninfected IFN-γR$^{-/-}$ mice and strep-R *MAP* infected IFN-γR$^{-/-}$ mice were photographed 12-weeks post-gavage. The red arrows indicate the location where differences in intestinal size were observed.

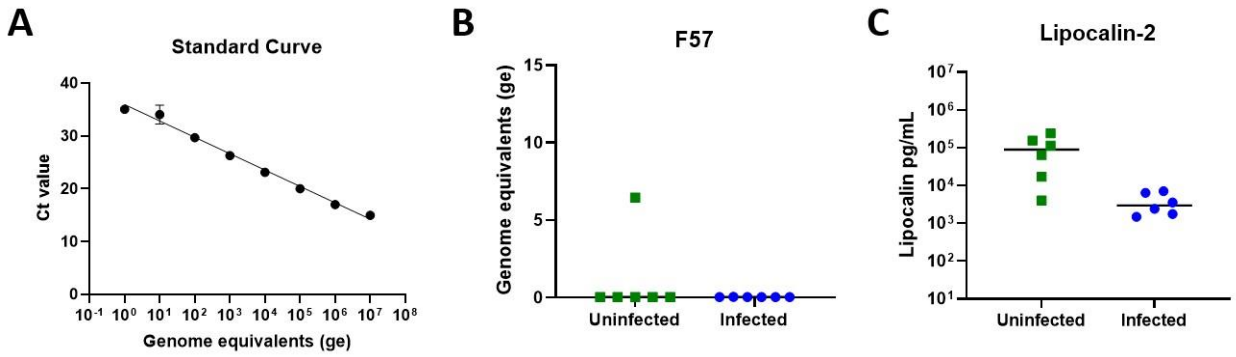



**Figure 7** Fecal shedding and lipocalin-2 assessment of IFNγR$^{-/-}$ mice

**A.** A standard curve was prepared for quantitative PCR of the F57 gene using *MAP* K10 genomic DNA diluted from 1x10$^{7}$ to 1 genome equivalents in order to interpolate values from fecal samples. **B.** The interpolated F57 values assessed fecal shedding in uninfected IFNγR$^{-/-}$ controls and IFNγR$^{-/-}$ mice 12-weeks post-gavage with strep-R *MAP.* **C.** An ELISA was performed to quantify lipocalin-2 in the fecal pellets of uninfected and infected IFNγR$^{-/-}$ mice.

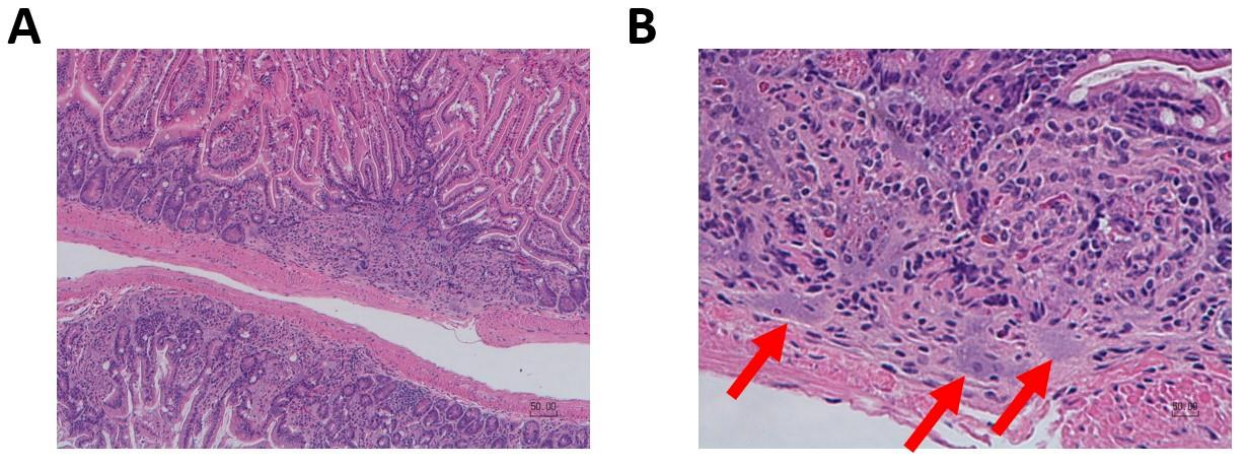



**Figure 8** Assessment of histopathology in *MAP*-infected IFNγR$^{-/-}$ mice

**A-B.** Shown are images of an H&E-stained section of the small intestine in an IFNγR$^{-/-}$ mouse at an original magnification of x100 (**A**) or x400 (**B**). The red arrows indicate histiocytes with paler stained cytoplasm which may represent epithelioid or foamy macrophages.

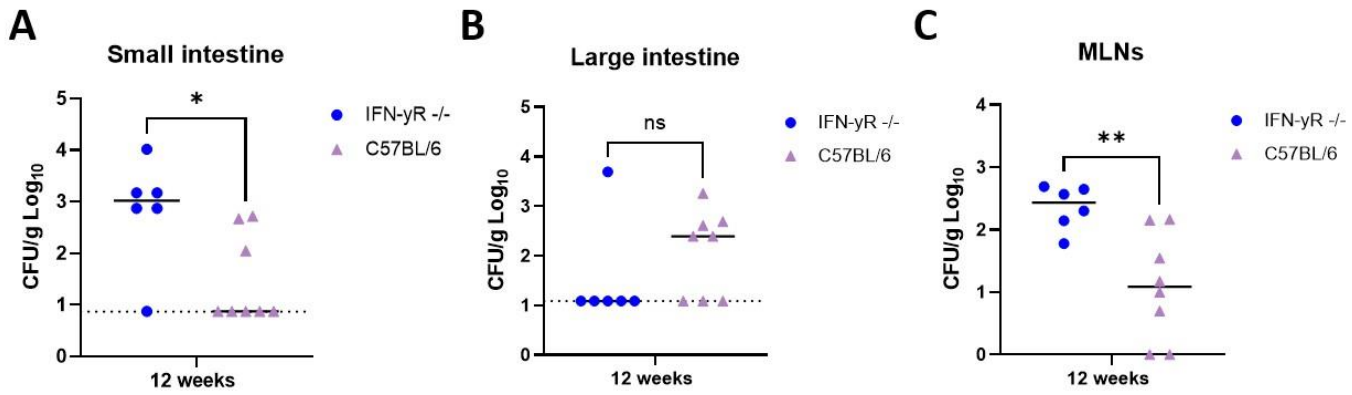



**Figure 9** Comparison of organ infection in strep-R *MAP* infected IFNγR$^{-/-}$ and C57BL/6 mice

Strep-R *MAP* CFUs were compared between IFNγR$^{-/-}$ and C57BL/6 mice in the small intestine (**A**), large intestine (**B**), and MLNs (**C**) at 12-weeks post-infection. Group comparisons were made by an unpaired t-test.

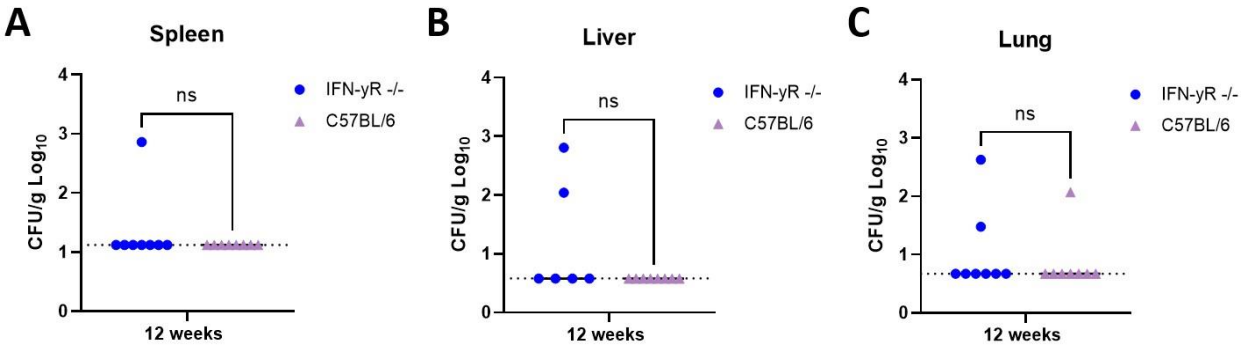



**Figure 10** Comparison of organ dissemination in strep-R *MAP* infected IFNγ$^{-/-}$ and C57BL/6 mice

Strep-R *MAP* CFUs were compared between IFNγR$^{-/-}$ and C57BL/6 mice in the spleen (**A**), liver (**B**), and lungs (**C**) at 12-weeks post-infection. Group comparisons were made by an unpaired t-test.

# SUMMARY AND FINAL CONCLUSIONS

In this thesis, my first aim was to determine the proportion of MTBC subspecies causing disease in the context of South Asia. I generated rapid and reliable assays to screen clinic samples and determined that zoonotic TB in India is a rare event and is caused by *M. orygis* rather than *M. bovis.* BCG disease was also detected in a handful of cases, but diagnostics were found to be either unavailable or limited, so an improved assay was developed. The low number of zoonotic TB cases was perhaps surprising given the number of TB-infected humans and cattle in the region and indirectly suggests that host-specific factors are also involved in developing zoonotic TB disease. The finding that *M. orygis* was the only zoonotic TB pathogen found in humans in our screens of South Asia stresses the need to alter the current definition of zoonotic TB to include other MTBC subspecies other than *M. bovis.* In order to ensure that future investigations are applying the proper assays and searching for the primary zoonotic TB pathogen of that region, it's necessary that this definition reflect the diversity of pathogens capable of causing disease.

My second aim was to study the pathogenesis of bovine mycobacteria by developing representative mouse models of oral infection. To do so, I developed a streptomycin pre-treatment oral model of infection with mycobacteria. This model, when applied to either *M. bovis* or *M. orygis,* led to chronic organ infection along with dissemination and pulmonary disease reminiscent of zoonotic TB. When applied to *MAP,* this infection strategy improved organ infection in the intestines and mesenteric lymph nodes, but unfortunately no evidence of inflammation or disease was observed, indicating either an immune deficient mouse strain or

another host altogether may be required to generate a small animal model of *MAP*-induced enteritis.

Bovine mycobacteria have a substantial impact on animal health, generate considerable global economic losses, and pose a threat to public health due to their zoonotic potential. This thesis provides new tools, spanning molecular detection assays and experimental models of infection, to enable the study of their prevalence and pathogenesis. Future investigations building upon this work may continue to improve our ability to both detect and understand infection with bovine mycobacteria.

## BIBLIOGRAPHY


1. Taylor, L. H., Latham, S. M. & Woolhouse, M. E. Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* **356**, 983–989 (2001).
2. Mackenzie, J. S. & Jeggo, M. The One Health Approach—Why Is It So Important? *Trop Med Infect Dis* **4**, 88 (2019).
3. Marani, M., Katul, G. G., Pan, W. K. & Parolari, A. J. Intensity and frequency of extreme novel epidemics. *Proc Natl Acad Sci U S A* **118**, e2105482118 (2021).
4. Glatter, K. A. & Finkelman, P. History of the Plague: An Ancient Pandemic for the Age of COVID-19. *Am J Med* **134**, 176–181 (2021).
5. Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).
6. Johnson, N. P. A. S. & Mueller, J. Updating the accounts: global mortality of the 1918-1920 'Spanish' influenza pandemic. *Bull Hist Med* **76**, 105–115 (2002).
7. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS Pandemic. *Cold Spring Harb Perspect Med* **1**, a006841 (2011).
8. World Health Organization. HIV. https://www.who.int/data/gho/data/themes/hiv-aids.
9. Marí Saéz, A. *et al.* Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med* **7**, 17–23 (2015).
10. World Health Organization. Ebola outbreak 2014-2016 - West Africa. https://www.who.int/emergencies/situations/ebola-outbreak-2014-2016-West-Africa.
11. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int.

12. Pekar, J. E. *et al.* The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science* **377**, 960–966 (2022).
13. McDaniel, C. J., Cardwell, D. M., Moeller, R. B. & Gray, G. C. Humans and Cattle: A Review of Bovine Zoonoses. *Vector Borne Zoonotic Dis* **14**, 1–19 (2014).
14. Robinson, T. P. *et al.* Mapping the global distribution of livestock. *PLoS One* **9**, e96084 (2014).
15. Lehmann, K. & Neumann, R. *Atlas und Grundriss der Bakeriologie und Lehrbuch der speziellen bakteriologischen Diagnositk*. (1896).
16. Payeur, J. B. Mycobacterium. *Encyclopedia of Food Microbiology 2nd ed.* 841–853 (2014) doi:10.1016/B978-0-12-384730-0.00229-9.
17. Dulberger, C. L., Rubin, E. J. & Boutte, C. C. The mycobacterial cell envelope — a moving target. *Nat Rev Microbiol* **18**, 47–59 (2020).
18. Malone, K. M. & Gordon, S. V. Mycobacterium tuberculosis Complex Members Adapted to Wild and Domestic Animals. *Adv Exp Med Biol* **1019**, 135–154 (2017).
19. World Health Organization. Leprosy. https://www.who.int/news-room/fact-sheets/detail/leprosy.
20. Han, X. Y. *et al.* Comparative Sequence Analysis of Mycobacterium leprae and the New Leprosy-Causing Mycobacterium lepromatosis. *J Bacteriol* **191**, 6067–6074 (2009).
21. To, K., Cao, R., Yegiazaryan, A., Owens, J. & Venketaraman, V. General Overview of Nontuberculous Mycobacteria Opportunistic Pathogens: Mycobacterium avium and Mycobacterium abscessus. *J Clin Med* **9**, 2541 (2020).
22. Ernst, J. D., Trevejo-Nuñez, G. & Banaiee, N. Genomics and the evolution, pathogenesis, and diagnosis of tuberculosis. *J Clin Invest* **117**, 1738–1745 (2007).

23. Geneva: World Health Organization. Global tuberculosis report 2022. https://iris.who.int/bitstream/handle/10665/363752/9789240061729-eng.pdf?sequence=1 (2022).
24. Loddenkemper, R., Lipman, M. & Zumla, A. Clinical Aspects of Adult Tuberculosis. *Cold Spring Harb Perspect Med* **6**, a017848 (2016).
25. Coscolla, M. *et al.* Phylogenomics of Mycobacterium africanum reveals a new lineage and a complex evolutionary history. *Microb Genom* **7**, 000477 (2021).
26. Brites, D. *et al.* A New Phylogenetic Framework for the Animal-Adapted Mycobacterium tuberculosis Complex. *Front Microbiol* **9**, 2820 (2018).
27. Ngabonziza, J. C. S. *et al.* A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. *Nat Commun* **11**, 2917 (2020).
28. Yeboah-Manu, D., de Jong, B. C. & Gehre, F. The Biology and Epidemiology of Mycobacterium africanum. *Adv Exp Med Biol* **1019**, 117–133 (2017).
29. Supply, P. & Brosch, R. The Biology and Epidemiology of Mycobacterium canettii. *Adv Exp Med Biol* **1019**, 27–41 (2017).
30. Luca, S. & Mihaescu, T. History of BCG Vaccine. *Maedica (Bucur)* **8**, 53–58 (2013).
31. Tortoli, E. *et al.* Emended description of Mycobacterium abscessus, Mycobacterium abscessus subsp. abscessus and Mycobacterium abscessus subsp. bolletii and designation of Mycobacterium abscessus subsp. massiliense comb. nov. *Int J Syst Evol Microbiol* **66**, 4471–4479 (2016).
32. Lopeman, R. C., Harrison, J., Desai, M. & Cox, J. A. G. Mycobacterium abscessus: Environmental Bacterium Turned Clinical Nightmare. *Microorganisms* **7**, 90 (2019).

33. Galmés-Truyols, A. *et al.* An outbreak of cutaneous infection due to Mycobacterium abscessus associated to mesotherapy. *Enferm Infecc Microbiol Clin* **29**, 510–514 (2011).

34. P, A. *et al.* Late-onset posttraumatic skin and soft-tissue infections caused by rapid-growing mycobacteria in tsunami survivors. *Clin Infect Dis* **47**, (2008).

35. Sparks, I. L., Derbyshire, K. M., Jacobs, W. R. & Morita, Y. S. Mycobacterium smegmatis: The Vanguard of Mycobacterial Research. *J Bacteriol* **205**, e0033722 (2023).

36. Aronson, J. D. Spontaneous Tuberculosis in Salt Water Fish. *J Infect Dis* **39**, 315–320 (1926).

37. Aubry, A., Mougari, F., Reibel, F. & Cambau, E. Mycobacterium marinum. *Microbiol Spectr* (2017) doi:10.1128/microbiolspec.TNMI7-0038-2016.

38. Stinear, T. P. *et al.* Insights from the complete genome sequence of Mycobacterium marinum on the evolution of Mycobacterium tuberculosis. *Genome Res* **18**, 729–741 (2008).

39. Stinear, T. P. *et al.* Reductive evolution and niche adaptation inferred from the genome of Mycobacterium ulcerans, the causative agent of Buruli ulcer. *Genome Res* **17**, 192 (2007).

40. van der Werf, T. S. *et al.* Mycobacterium ulcerans disease. *Bull World Health Organ* **83**, 785–791 (2005).

41. Lindeboom, J. A., Bruijnesteijn van Coppenraet, L. E. S., van Soolingen, D., Prins, J. M. & Kuijper, E. J. Clinical Manifestations, Diagnosis, and Treatment of Mycobacterium haemophilum Infections. *Clin Microbiol Rev* **24**, 701–717 (2011).

42. Tufariello, J. M. *et al.* The Complete Genome Sequence of the Emerging Pathogen Mycobacterium haemophilum Explains Its Unique Culture Requirements. *mBio* **6**, (2015).

43. van Ingen, J., Turenne, C. Y., Tortoli, E., Wallace, R. J. & Brown-Elliott, B. A. A definition of the Mycobacterium avium complex for taxonomical and clinical purposes, a review. *Int J Syst Evol Microbiol* **68**, 3666–3677 (2018).

44. Kalayjian, R. C. *et al.* Pulmonary disease due to infection by Mycobacterium avium complex in patients with AIDS. *Clin Infect Dis* **20**, 1186–1194 (1995).

45. Rosenzweig, D. Y. Pulmonary mycobacterial infections due to Mycobacterium intracellulare-avium complex. Clinical features and course in 100 consecutive cases. *Chest* **75**, 115–119 (1979).

46. Yu, X. & Jiang, W. Mycobacterium colombiense and Mycobacterium avium Complex Causing Severe Pneumonia in a Patient with HIV Identified by a Novel Molecular-Based Method. *Infect Drug Resist* **14**, 11–16 (2021).

47. Inojosa, W. O. *et al.* Mycobacterium chimaera infections following cardiac surgery in Treviso Hospital, Italy, from 2016 to 2019: Cases report. *World J Clin Cases* **7**, 2776–2786 (2019).

48. Mizzi, R., Plain, K. M., Whittington, R. & Timms, V. J. Global Phylogeny of Mycobacterium avium and Identification of Mutation Hotspots During Niche Adaptation. *Front Microbiol* **13**, (2022).

49. World Health Organization, World Organisation for Animal Health, Food and Agriculture Organization & International Union Against Tuberculosis and Lung Disease. *Zoonotic tuberculosis*. https://cdn.who.int/media/docs/default-source/hq-tuberculosis/zoonotic-tb-factsheet-2017.pdf?sfvrsn=66fdf3a1_3&download=true (2017).

50. Duffy, S. C. *et al.* Reconsidering Mycobacterium bovis as a proxy for zoonotic tuberculosis: a molecular epidemiological surveillance study. *Lancet Microbe* **1**, e66–e73 (2020).

51. Koch, R. Die Aetiologie der Tuberculose. *Berliner Klinische Wochenschrift* **15**, 287–296 (1882).

52. Grange, J. M. & Collins, C. H. Bovine tubercle bacilli and disease in animals and man. *Epidemiol Infect* **99**, 221–234 (1987).

53. Smith, T. A comparative study of bovine tubercle bacilli and of human bacilli from sputum. *J Exp Med* **3**, 451–511 (1898).

54. The transactions of the British congress on tuberculosis, 1901. *Lancet* **160**, 768–771 (1902).

55. Royal Commission on Tuberculosis. Final report of the Royal Commission on Tuberculosis, 1911. *Br Med J* **2**, 122–125 (1911).

56. Karlson, A. G. & Lessel, E. F. Mycobacterium bovis nom.nov. *Int J Syst Bacteriol* **20**, 273–282 (1970).

57. Garnier, T. *et al.* The complete genome sequence of Mycobacterium bovis. *Proc Natl Acad Sci U S A* **100**, 7877–7882 (2003).

58. Griffith, A. S. The Cultural Characteristics and Virulence of Mammalian Tubercle Bacilli and the Circumstances Under Which They May Vary. *The Journal of State Medicine (1912-1937)* **30**, 139–159 (1922).

59. Lebek, G. Dependence of the oxygen optimal concentration of both strains of mammalian Mycobacterium tuberculosis on the nutritive substrate. *Zentralbl Bakteriol* **176**, 530–537 (1959).

60. Scorpio, A. & Zhang, Y. Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med* **2**, 662–667 (1996).

61. Collins, C. H. & Grange, J. M. The bovine tubercle bacillus. *J Appl Bacteriol* **55**, 13–29 (1983).

62. Early, J. & Alling, T. Determination of Compound Kill Kinetics Against Mycobacterium tuberculosis. *Methods Mol Biol* 269–279 (2015) doi:10.1007/978-1-4939-2450-9_16.

63. Rodriguez-Morales, A. J. & Castañeda-Hernández, D. M. Bacteria: Mycobacterium bovis. *Encyclopedia of Food Safety* **1**, 468–475 (2014).

64. Brosch, R. *et al.* A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A* **99**, 3684–3689 (2002).

65. Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A. & Behr, M. A. Genomic deletions suggest a phylogeny for the Mycobacterium tuberculosis complex. *J Infect Dis* **186**, 74–80 (2002).

66. Siala, M. *et al.* First-time detection and identification of the Mycobacterium tuberculosis Complex members in extrapulmonary tuberculosis clinical samples in south Tunisia by a single tube tetraplex real-time PCR assay. *PLoS Negl Trop Dis* **11**, e0005572 (2017).

67. Taylor, G. M., Worth, D. R., Palmer, S., Jahans, K. & Hewinson, R. G. Rapid detection of Mycobacterium bovis DNA in cattle lymph nodes with visible lesions using PCR. *BMC Vet Res* **3**, 12 (2007).

68. Bapat, P. R. *et al.* Prevalence of zoonotic tuberculosis and associated risk factors in Central Indian populations. *J Epidemiol Glob Health* **7**, 277–283 (2017).

69. Zimpel, C. K. *et al.* Global Distribution and Evolution of Mycobacterium bovis Lineages. *Front Microbiol* **11**, (2020).

70. Lomme, J. R., Thoen, C. O., Himes, E. M., Vinson, J. W. & King, R. E. Mycobacterium tuberculosis infection in two East African oryxes. *J Am Vet Med Assoc* **169**, 912–914 (1976).

71. van Soolingen, D. *et al.* Use of various genetic markers in differentiation of Mycobacterium bovis strains from animals and humans and for studying epidemiology of bovine tuberculosis. *J Clin Microbiol* **32**, 2425–2433 (1994).

72. van Ingen, J. *et al.* Characterization of Mycobacterium orygis as M. tuberculosis Complex Subspecies. *Emerg Infect Dis* **18**, 653–655 (2012).

73. Rufai, S. B. *et al.* Complete Genome Sequence of Mycobacterium orygis Strain 51145. *Microbiol Resour Announc* **10**, e01279-20 (2021).

74. Saïd-Salim, B., Mostowy, S., Kristof, A. S. & Behr, M. A. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. *Mol Microbiol* **62**, 1251–1263 (2006).

75. Dawson, K. L. *et al.* Transmission of Mycobacterium orygis (M. tuberculosis Complex Species) from a Tuberculosis Patient to a Dairy Cow in New Zealand. *J Clin Microbiol* **50**, 3136–3138 (2012).

76. Thapa, J., Nakajima, C., Maharjan, B., Poudell, A. & Suzuki, Y. Molecular characterization of Mycobacterium orygis isolates from wild animals of Nepal. *Jpn J Vet Res* **63**, 151–158 (2015).

77. Thapa, J. *et al.* Mycobacterium orygis-Associated Tuberculosis in Free-Ranging Rhinoceros, Nepal, 2015. *Emerg Infect Dis* **22**, 570–572 (2016).

78. Marcos, L. A. *et al.* Mycobacterium orygis Lymphadenitis in New York, USA. *Emerg Infect Dis* **23**, 1749–1751 (2017).

79. Rahim, Z. *et al.* Tuberculosis Caused by Mycobacterium orygis in Dairy Cattle and Captured Monkeys in Bangladesh: a New Scenario of Tuberculosis in South Asia. *Transbound Emerg Dis* **64**, 1965–1969 (2017).

80. Eldholm, V., Rønning, J. O., Mengshoel, A. T. & Arnesen, T. Import and transmission of Mycobacterium orygis and Mycobacterium africanum, Norway. *BMC Infect Dis* **21**, 562 (2021).

81. Refaya, A. K. *et al.* Tuberculosis caused by Mycobacterium orygis in wild ungulates in Chennai, South India. *Transbound Emerg Dis* **69**, e3327–e3333 (2022).

82. Sharma, M. *et al.* Emergence of Mycobacterium orygis–Associated Tuberculosis in Wild Ruminants, India. *Emerg Infect Dis* **29**, 661–663 (2023).

83. Thapa, J., Gordon, S. V., Nakajima, C. & Suzuki, Y. Threat from Mycobacterium orygis-associated tuberculosis in south Asia. *Lancet Microbe* **3**, e641–e642 (2022).

84. Palmer, M. V., Kanipe, C. & Boggiatto, P. M. The Bovine Tuberculoid Granuloma. *Pathogens* **11**, 61 (2022).

85. World Organization for Animal Health. *Bovine Tuberculosis*. https://www.woah.org/fileadmin/Home/eng/Media_Center/docs/pdf/Disease_cards/BOVINE-TB-EN.pdf.

86. Ní Bhuachalla, D., Corner, L. A., More, S. J. & Gormley, E. The role of badgers in the epidemiology of Mycobacterium bovis infection (tuberculosis) in cattle in the United Kingdom and the Republic of Ireland: current perspectives on control strategies. *Vet Med (Auckl)* **6**, 27–38 (2014).

87. Nugent, G., Buddle, B. & Knowles, G. Epidemiology and control of Mycobacterium bovis infection in brushtail possums (Trichosurus vulpecula), the primary wildlife host of bovine tuberculosis in New Zealand. *N Z Vet J* **63**, 28–41 (2015).

88. Ramos, B., Pereira, A. C., Reis, A. C. & Cunha, M. V. Estimates of the global and continental burden of animal tuberculosis in key livestock species worldwide: A meta-analysis study. *One Health* **10**, 100169 (2020).

89. Srinivasan, S. *et al.* Prevalence of Bovine Tuberculosis in India: A systematic review and meta-analysis. *Transbound Emerg Dis* **65**, 1627–1640 (2018).

90. World Organization for Animal Health. Controlling bovine tuberculosis: a One Health challenge. https://bulletin.woah.org/wp-content/uploads/bulletins/panorama-2019-1-en.pdf (2019).

91. Tschopp, R., Zinsstag, J., Conlan, A., Gemechu, G. & Wood, J. Productivity loss and cost of bovine tuberculosis for the dairy livestock sector in Ethiopia. *Prev Vet Med* **202**, 105616 (2022).

92. Charles, G., Christl, D., Glyn, H., Michael, W. & James, W. *Bovine TB Strategy Review*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/756942/tb-review-final-report-corrected.pdf (2018).

93. Currier, R. W. & Widness, J. A. A Brief History of Milk Hygiene and Its Impact on Infant Mortality from 1875 to 1925 and Implications for Today: A Review. *J Food Prot* **81**, 1713–1722 (2018).

94. Roswurm, J. D. & Ranney, A. F. Sharpening the attack on bovine tuberculosis. *Am J Public Health* **63**, 884–886 (1973).

95. Griffith, A. S. Bovine tuberculosis in man. *Tubercle* **18**, 529–543 (1937).

96. Geneva: World Health Organization. *Global tuberculosis report 2020*. https://www.who.int/publications-detail-redirect/9789240013131 (2020).

97. Taye, H. *et al.* Global prevalence of Mycobacterium bovis infections among human tuberculosis cases: Systematic review and meta-analysis. *Zoonoses and Public Health* **68**, 704–718 (2021).

98. Centers for Disease Control. Mycobacterium bovis (Bovine Tuberculosis) in Humans. https://www.cdc.gov/tb/publications/factsheets/general/mbovis.htm (2012).

99. Dürr, S. *et al.* Differences in Primary Sites of Infection between Zoonotic and Human Tuberculosis: Results from a Worldwide Systematic Review. *PLoS Negl Trop Dis* **7**, e2399 (2013).

100. Sumanth, L. J. *et al.* Clinical features of human tuberculosis due to Mycobacterium orygis in Southern India. *J Clin Tuberc Other Mycobact Dis* **32**, 100372 (2023).

101. Norbis, L. *et al.* Challenges and perspectives in the diagnosis of extrapulmonary tuberculosis. *Expert Rev Anti Infect Ther* **12**, 633–647 (2014).

102. Lan, Z., Bastos, M. & Menzies, D. Treatment of human disease due to Mycobacterium bovis: a systematic review. *Eur Respri J* **48**, 1500–1503 (2016).

103. Geneva: World Health Organization. WHO operational handbook on tuberculosis Module 4: Treatment Drug-susceptible tuberculosis treatment. https://iris.who.int/bitstream/handle/10665/354548/9789240050761-eng.pdf?sequence=1 (2022).

104. Majoor, C. J., Magis-Escurra, C., van Ingen, J., Boeree, M. J. & van Soolingen, D. Epidemiology of Mycobacterium bovis Disease in Humans, the Netherlands, 1993–2007. *Emerg Infect Dis* **17**, 457–463 (2011).

105. World Health Organization. The end TB strategy. https://apps.who.int/iris/bitstream/handle/10665/331326/WHO-HTM-TB-2015.19-eng.pdf?sequence=1&isAllowed=y (2015).
106. Palaniyandi, K. *et al.* Isolation and comparative genomics of Mycobacterium tuberculosis isolates from cattle and their attendants in South India. *Sci Rep* **9**, 17892 (2019).
107. Hlokwe, T., Said, H. & Gcebe, N. Mycobacterium tuberculosis infection in cattle from the Eastern Cape Province of South Africa. *BMC Vet Res* **13**, (2017).
108. Macedo Couto, R., Ranzani, O. T. & Waldman, E. A. Zoonotic Tuberculosis in Humans: Control, Surveillance, and the One Health Approach. *Epidemiol Rev* **41**, 130–144 (2019).
109. Couto, R. de M., Santana, G. O., Ranzani, O. T. & Waldman, E. A. One Health and surveillance of zoonotic tuberculosis in selected low-income, middle-income and high-income countries: A systematic review. *PLOS Negl Trop Dis* **16**, e0010428 (2022).
110. Tulu, B. *et al.* Epidemiology of Bovine Tuberculosis and Its Zoonotic Implication in Addis Ababa Milkshed, Central Ethiopia. *Front Vet Sci* **8**, 595511 (2021).
111. Trunz, B. B., Fine, P. & Dye, C. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet* **367**, 1173–1180 (2006).
112. Calmette, A. *L'Infection bacillaire et la tuberculose chez l'homme et chez les animaux : processus d'infection et de défense : étude biologique et expérimentale*. (1920).
113. Abdallah, A. M. & Behr, M. A. Evolution and Strain Variation in BCG. *Adv Exp Med Biol* **1019**, 155–169 (2017).
114. Lewis, K. N. *et al.* Deletion of RD1 from Mycobacterium tuberculosis Mimics Bacille Calmette-Guérin Attenuation. *J Infect Dis* **187**, 117–123 (2003).

115. Hart, P. D. & Sutherland, I. BCG and vole bacillus vaccines in the prevention of tuberculosis in adolescence and early adult life. *Br Med J* **2**, 293–295 (1977).

116. Comstock, G. W. & Palmer, C. E. Long-term results of BCG vaccination in the southern United States. *Am Rev Respir Dis* **93**, 171–183 (1966).

117. Comstock, G. W., Woolpert, S. F. & Livesay, V. T. Tuberculosis studies in Muscogee County, Georgia. Twenty-year evaluation of a community trial of BCG vaccination. *Public Health Rep* **91**, 276–280 (1976).

118. Martinez, L. *et al.* Infant BCG vaccination and risk of pulmonary and extrapulmonary tuberculosis throughout the life course: a systematic review and individual participant data meta-analysis. *Lancet Glob Health* **10**, e1307–e1316 (2022).

119. Mangtani, P. *et al.* Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clin Infect Dis* **58**, 470–480 (2014).

120. Fox, G. J., Orlova, M. & Schurr, E. Tuberculosis in Newborns: The Lessons of the "Lübeck Disaster" (1929–1933). *PLoS Pathog* **12**, e1005271 (2016).

121. World Health Organization. Information sheet observed rate of vaccine reactions bacille Calmette-Guérin (BCG) vaccine. https://cdn.who.int/media/docs/default-source/pvg/global-vaccine-safety/bcg-vaccine-rates-information-sheet.pdf?sfvrsn=8a8f52ba_6&download=true (2012).

122. Roh, E. J. *et al.* Investigation of adverse events following bacille Calmette-Guérin immunization using immunization safety surveillance system in Korea Centers for Disease Control and Prevention. *Clin Exp Vaccine Res* **9**, 133–145 (2020).

123. Schaltz-Buchholzer, F. *et al.* Early Vaccination With Bacille Calmette-Guérin-Denmark or BCG-Japan Versus BCG-Russia to Healthy Newborns in Guinea-Bissau: A Randomized Controlled Trial. *Clin Infect Dis* **71**, 1883–1893 (2020).

124. Biering-Sørensen, S. *et al.* Early BCG-Denmark and Neonatal Mortality Among Infants Weighing <2500 g: A Randomized Controlled Trial. *Clin Infect Dis* **65**, 1183–1190 (2017).

125. Goodridge, H. S. *et al.* Harnessing the beneficial heterologous effects of vaccination. *Nat Rev Immunol* **16**, 392–400 (2016).

126. Old, L. J., Clarke, D. A. & Benacerraf, B. Effect of Bacillus Calmette-Guerin infection on transplanted tumours in the mouse. *Nature* **184(Suppl 5)**, 291–292 (1959).

127. Jiang, S. & Redelman-Sidi, G. BCG in Bladder Cancer Immunotherapy. *Cancers (Basel)* **14**, 3073 (2022).

128. Morales, A., Eidinger, D. & Bruce, A. W. Intracavitary Bacillus Calmette-Guerin in the treatment of superficial bladder tumors. *J Urol* **116**, 180–183 (1976).

129. Mathé, G. *et al.* Active immunotherapy for acute lymphoblastic leukaemia. *Lancet* **1**, 697–699 (1969).

130. Morton, D. L. *et al.* BCG immunotherapy of malignant melanoma: summary of a seven-year experience. *Ann Surg* **180**, 635–643 (1974).

131. Babjuk, M. *et al.* European Association of Urology Guidelines on Non-muscle-invasive Bladder Cancer (TaT1 and Carcinoma In Situ) - 2019 Update. *Eur Urol* **76**, 639–657 (2019).

132. Pittet, L. F. *et al.* Randomized Trial of BCG Vaccine to Protect against Covid-19 in Health Care Workers. *N Engl J Med* **388**, 1582–1596 (2023).

133. Centers for Disease Control. Biosafety in Microbiological and Biomedical Laboratories. https://www.cdc.gov/labs/pdf/CDC-BiosafetymicrobiologicalBiomedicalLaboratories-2009-P.pdf (2009).

134. Bernatowska, E. A. *et al.* Disseminated Bacillus Calmette-Guérin Infection and Immunodeficiency. *Emerg Infect Dis* **13**, 799–801 (2007).

135. Lee, P. P. Disseminated Bacillus Calmette-Guérin and Susceptibility to Mycobacterial Infections-Implications on Bacillus Calmette-Guérin Vaccinations. *Ann Acad Med Singap* **44**, 297–301 (2015).

136. Ong, R. Y. L. *et al.* Disseminated bacillus-Calmette-Guérin infections and primary immunodeficiency disorders in Singapore: a single center 15-year retrospective review. *Int J Infect Dis* **97**, 117–125 (2020).

137. Aelami, M. H. *et al.* Post-Vaccination Disseminated Bacillus Calmette Guerin Infection Among Children in Southern Iran. *Jundishapur J Microbiol* **8**, e25663 (2015).

138. Sharifi Asadi, P. *et al.* Clinical, laboratory and imaging findings of the patients with disseminated bacilli Calmette–Guerin disease. *Allergol Immunopathol (Madr)* **43**, 254–258 (2015).

139. Marques, M. *et al.* Disseminated Bacillus Calmette-Guérin (BCG) infection with pulmonary and renal involvement: A rare complication of BCG immunotherapy. A case report and narrative review. *Pulmonol* **26**, 346–352 (2020).

140. Norouzi, S., Aghamohammadi, A., Mamishi, S., Rosenzweig, S. D. & Rezaei, N. Bacillus Calmette-Guérin (BCG) complications associated with primary immunodeficiency diseases. *J Infect* **64**, 543–554 (2012).

141. Fekrvand, S. *et al.* Primary Immunodeficiency Diseases and Bacillus Calmette-Guérin (BCG)-Vaccine–Derived Complications: A Systematic Review. *J Allergy Clin Immunol Pract* **8**, 1371–1386 (2020).

142. Marciano, B. E. *et al.* BCG vaccination in SCID patients: complications, risks and vaccination policies. *J Allergy Clin Immunol* **133**, 1134–1141 (2014).

143. Li, T. *et al.* Genetic and Clinical Profiles of Disseminated Bacillus Calmette-Guérin Disease and Chronic Granulomatous Disease in China. *Front Immunol* **10**, (2019).

144. Aktan, F. iNOS-mediated nitric oxide production and its regulation. *Life Sci* **75**, 639–653 (2004).

145. Abbas, M., Bakhtyar, A. & Bazzi, R. Neonatal HIV. in *StatPearls* (StatPearls Publishing, 2023).

146. World Health Organization. Mother-to-child transmission of HIV. https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/prevention/mother-to-child-transmission-of-hiv.

147. BCG World Atlas. http://www.bcgatlas.org/index.php.

148. Hesseling, A. C. *et al.* Disseminated bacille Calmette-Guérin disease in HIV-infected South African infants. *Bull World Health Organ* **87**, 505–511 (2009).

149. Rezai, M. S., Khotaei, G., Mamishi, S., Kheirkhah, M. & Parvaneh, N. Disseminated Bacillus Calmette–Guerin Infection after BCG Vaccination. *J Trop Pediatr* **54**, 413–416 (2008).

150. Dash, N. *et al.* Clinical, Microbiological Profile and Treatment Outcomes of Infants with BCG Adenitis: A Retrospective Study. *J Trop Pediatr* **68**, fmac094 (2022).

151. Zeng, Y. *et al.* Clinical and Genetic Characteristics of BCG Disease in Chinese Children: a Retrospective Study. *J Clin Immunol* **43**, 756–768 (2023).

152. Mahdaviani, S. A. *et al.* Effective anti-mycobacterial treatment for BCG disease in patients with Mendelian Susceptibility to Mycobacterial Disease (MSMD): a case series. *Ann Clin Microbiol Antimicrob* **21**, 8 (2022).

153. Ying, W. *et al.* Clinical Characteristics and Immunogenetics of BCGosis/BCGitis in Chinese Children: A 6 Year Follow-Up Study. *PLoS One* **9**, e94485 (2014).

154. Radwan, N. *et al.* Outcome of Hematopoietic Stem Cell Transplantation in patients with Mendelian Susceptibility to Mycobacterial Diseases. *J Clin Immunol* **41**, 1774–1780 (2021).

155. Castagnoli, R., Delmonte, O. M., Calzoni, E. & Notarangelo, L. D. Hematopoietic Stem Cell Transplantation in Primary Immunodeficiency Diseases: Current Status and Future Perspectives. *Front Pediatr* **7**, 295 (2019).

156. Twort, F. W. & Ingram, G. L. Y. A Method for Isolating and Cultivating the Mycobacterium enteritidis chronicae pseudotuberculosae bovis, Johne, and some Experiments on the Preparation of a Diagnostic Vaccine for Pseudo-tuberculous Enteritis of Bovines. *Proceedings of the Royal Society of London Series B* **84**, 517–542 (1912).

157. Johne, H. & Frothingham, J. Ein eigenthuemlicher fall von tuberculose beim rind. *Dtsch Z Tiermed Pathol* **21**, 438–454 (1895).

158. Li, L. *et al.* The complete genome sequence of Mycobacterium avium subspecies paratuberculosis. *Proc Natl Acad Sci U S A* **102**, 12344–12349 (2005).

159. Thorel, M. F., Krichevsky, M. & Lévy-Frébault, V. V. Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of Mycobacterium avium, and description of

Mycobacterium avium subsp. avium subsp. nov., Mycobacterium avium subsp. paratuberculosis subsp. nov., and Mycobacterium avium subsp. silvaticum subsp. nov. *Int J Syst Bacteriol* **40**, 254–260 (1990).

160. Coussens, P. *et al.* Host-pathogen interactions and intracellular survival of Mycobacterium avium subsp. paratuberculosis. in *Paratuberculosis: organism, disease, control* 120–138 (CAB International, 2020).
161. Wu, C. *et al.* Invasion and Persistence of Mycobacterium avium subsp. paratuberculosis during Early Stages of Johne's Disease in Calves. *Infect Immun* **75**, 2110–2119 (2007).
162. Stabel, J. R., Koets, A. & de Silva, K. Immunology of paratuberculosis infection and disease. in *Paratuberculosis: organism, disease, control* 248–265 (CAB International, 2020).
163. Clarke, C. J. & Little, D. The pathology of ovine paratuberculosis: Gross and histological changes in the intestine and other tissues. *J Comp Pathol* **114**, 419–437 (1996).
164. Whittington, R. Cultivation of Mycobacterium avium subsp. paratuberculosis. in *Paratuberculosis: organism, disease, control* 266–304 (CAB International, 2020).
165. Alexander, D. C., Turenne, C. Y. & Behr, M. A. Insertion and Deletion Events That Define the Pathogen Mycobacterium avium subsp. paratuberculosis. *J Bacteriol* **191**, 1018–1025 (2009).
166. Imada, J., Kelton, D. & Barkema, H. Epidemiology, Global Prevalence and Economics of Infection. in *Paratuberculosis: Organism, Disease, Control* 1–13 (CAB International, 2020).

167. United States Department of Agriculture. Johne's Disease on U.S. Dairies, 1991–2007. https://www.aphis.usda.gov/animal_health/nahms/dairy/downloads/dairy07/Dairy07_is_Johnes_1.pdf (2008).

168. Rathnaiah, G. *et al.* Pathogenesis, Molecular Genetics, and Genomics of Mycobacterium avium subsp. paratuberculosis, the Etiologic Agent of Johne's Disease. *Front Vet Sci* **4**, 187 (2017).

169. Corbett, C. S. *et al.* Prevalence of Mycobacterium avium ssp. paratuberculosis infections in Canadian dairy herds. *J Dairy Sci* **101**, 11218–11228 (2018).

170. Rasmussen, P., Barkema, H. W., Mason, S., Beaulieu, E. & Hall, D. C. Economic losses due to Johne's disease (paratuberculosis) in dairy cattle. *J Dairy Sci* **104**, 3123–3143 (2021).

171. Lombard, J. E. *et al.* Herd-level prevalence of Mycobacterium avium subsp. paratuberculosis infection in United States dairy herds in 2007. *Prev Vet Med* **108**, 234–238 (2013).

172. Barkema, H. *et al.* Lessons Learned from the Canadian Johne's Disease Programs. *WCDS Advances in Dairy Technology* **30**, 309–318 (2018).

173. Field, N. L., Mee, J. F. & McAloon, C. G. Characteristics (sensitivity and specificity) of herd-level diagnostic tests for Mycobacterium avium subspecies paratuberculosis in cattle - A systematic review. *Vet J* **279**, 105786 (2022).

174. Whittington, R. J., Marshall, D. J., Nicholls, P. J., Marsh, I. B. & Reddacliff, L. A. Survival and Dormancy of Mycobacterium avium subsp. paratuberculosis in the Environment. *Appl Environ Microbiol* **70**, 2989–3004 (2004).

175. Roda, G. *et al.* Crohn's disease. *Nat Rev Dis Primers* **6**, 1–19 (2020).

176. Ng, S. C. *et al.* Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* **390**, 2769–2778 (2017).

177. Abraham, C. & Cho, J. H. Inflammatory Bowel Disease. *N Engl J Med* **361**, 2066–2078 (2009).

178. Stappenbeck, T. S. *et al.* Crohn disease: a current perspective on genetics, autophagy and immunity. *Autophagy* **7**, 355–374 (2011).

179. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **146**, 1489–1499 (2014).

180. Palmela, C. *et al.* Adherent-invasive Escherichia coli in inflammatory bowel disease. *Gut* **67**, 574–587 (2018).

181. Nishida, A. *et al.* Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin J Gastroenterol* **11**, 1–10 (2018).

182. Kumar, A., Cole, A., Segal, J., Smith, P. & Limdi, J. K. A review of the therapeutic management of Crohn's disease. *Therap Adv Gastroenterol* **15**, 17562848221078456 (2022).

183. M, K. *et al.* Net Remission Rates with Biologic Treatment in Crohn's Disease: A Reappraisal of the Clinical Trial Data. *Clin Gastroenterol Hepatol* **21**, (2023).

184. Dalziel, T. K. Chronic Interstitial Enteritis. *Br Med J* **2**, 1068–1070 (1913).

185. Savarino, E. *et al.* Antimicrobial treatment with the fixed-dose antibiotic combination RHB-104 for Mycobacterium avium subspecies paratuberculosis in Crohn's disease: pharmacological and clinical implications. *Expert Opin Biol Ther* **19**, 79–88 (2019).

186. Sartor, R. B. Does Mycobacterium avium subspecies paratuberculosis cause Crohn's disease? *Gut* **54**, 896–898 (2005).

187. Gitlin, L., Borody, T. J., Chamberlin, W. & Campbell, J. Mycobacterium avium ss paratuberculosis-associated diseases: piecing the Crohn's puzzle together. *J Clin Gastroenterol* **46**, 649–655 (2012).

188. Liverani, E., Scaioli, E., Cardamone, C., Dal Monte, P. & Belluzzi, A. Mycobacterium avium subspecies paratuberculosis in the etiology of Crohn's disease, cause or epiphenomenon? *World J Gastroenterol* **20**, 13060–13070 (2014).

189. Bach, H. What Role Does Mycobacterium avium subsp. paratuberculosis Play in Crohn's Disease? *Curr Infect Dis Rep* **17**, 463 (2015).

190. Selby, W. *et al.* Two-year combination antibiotic therapy with clarithromycin, rifabutin, and clofazimine for Crohn's disease. *Gastroenterology* **132**, 2313–2319 (2007).

191. Chiodini, R. J. Crohn's disease and the mycobacterioses: a review and comparison of two disease entities. *Clin Microbiol Rev* **2**, 90–117 (1989).

192. Chacon, O., Bermudez, L. E. & Barletta, R. G. Johne's disease, inflammatory bowel disease, and Mycobacterium paratuberculosis. *Annu Rev Microbiol* **58**, 329–363 (2004).

193. Zarei-Kordshouli, F., Geramizadeh, B. & Khodakaram-Tafti, A. Prevalence of Mycobacterium avium subspecies paratuberculosis IS 900 DNA in biopsy tissues from patients with Crohn's disease: histopathological and molecular comparison with Johne's disease in Fars province of Iran. *BMC Infect Dis* **19**, 23 (2019).

194. Momotani, E. *et al.* Molecular pathogenesis of bovine paratuberculosis and human inflammatory bowel diseases. *Vet Immunol Immunopathol* **148**, 55–68 (2012).

195. Boulais, C., Wacker, R., Augustin, J.-C., Cheikh, M. H. B. & Peladan, F. Modeling the occurrence of Mycobacterium avium subsp. paratuberculosis in bulk raw milk and the impact of management options for exposure mitigation. *J Food Prot* **74**, 1126–1136 (2011).

196. Okura, H., Toft, N. & Nielsen, S. S. Occurrence of Mycobacterium avium subsp. paratuberculosis in milk at dairy cattle farms: a systematic review and meta-analysis. *Vet Microbiol* **157**, 253–263 (2012).

197. Carvalho, I. A., Pietralonga, P. a. G., Schwarz, D. G. G., Faria, A. C. S. & Moreira, M. a. S. Short communication: Recovery of viable Mycobacterium avium subspecies paratuberculosis from retail pasteurized whole milk in Brazil. *J Dairy Sci* **95**, 6946–6948 (2012).

198. Gerrard, Z. E. *et al.* Survival of Mycobacterium avium subspecies paratuberculosis in retail pasteurised milk. *Food Microbiol* **74**, 57–63 (2018).

199. Lorencova, A., Babak, V., Kralova, A. & Borilova, G. Survival of Mycobacterium avium subsp. paratuberculosis in raw fermented sausages during production and storage. *Meat Sci* **155**, 20–26 (2019).

200. Botsaris, G. *et al.* Rapid detection methods for viable Mycobacterium avium subspecies paratuberculosis in milk and cheese. *Int J Food Microbiol* **141 Suppl 1**, S87-90 (2010).

201. Faria, A. C. S. *et al.* Short communication: Viable Mycobacterium avium subspecies paratuberculosis in retail artisanal Coalho cheese from Northeastern Brazil. *J Dairy Sci* **97**, 4111–4114 (2014).

202. Botsaris, G. *et al.* Detection of viable Mycobacterium avium subspecies paratuberculosis in powdered infant formula by phage-PCR and confirmed by culture. *Int J Food Microbiol* **216**, 91–94 (2016).

203. Acharya, K. R., Dhand, N. K., Whittington, R. J. & Plain, K. M. Detection of Mycobacterium avium subspecies paratuberculosis in powdered infant formula using IS900 quantitative PCR and liquid culture media. *Int J Food Microbiol* **257**, 1–9 (2017).

204. Beumer, A. *et al.* Detection of Mycobacterium avium subsp. paratuberculosis in drinking water and biofilms by quantitative PCR. *Appl Environ Microbiol* **76**, 7367–7370 (2010).

205. Rhodes, G., Henrys, P., Thomson, B. C. & Pickup, R. W. Mycobacterium avium subspecies paratuberculosis is widely distributed in British soils and waters: implications for animal and human health. *Environ Microbiol* **15**, 2761–2774 (2013).

206. Espeschit, I. F., Souza, M. C. C., Lima, M. C. & Moreira, M. A. S. First molecular typing of Mycobacterium avium subspecies paratuberculosis identified in animal and human drinking water from dairy goat farms in Brazil. *Braz J Microbiol* **49**, 358–361 (2018).

207. Rhodes, G. *et al.* Mycobacterium avium Subspecies paratuberculosis: Human Exposure through Environmental and Domestic Aerosols. *Pathogens* **3**, 577–595 (2014).

208. Pierce, E. S. Possible transmission of Mycobacterium avium subspecies paratuberculosis through potable water: lessons from an urban cluster of Crohn's disease. *Gut Pathog* **1**, 17 (2009).

209. Pierce, E. S., Borowitz, S. M. & Naser, S. A. The Broad Street pump revisited: dairy farms and an ongoing outbreak of inflammatory bowel disease in Forest, Virginia. *Gut Pathog* **3**, 20 (2011).

210. Opstelten, J. L. *et al.* Dairy Products, Dietary Calcium, and Risk of Inflammatory Bowel Disease: Results From a European Prospective Cohort Investigation. *Inflamm Bowel Dis* **22**, 1403–1411 (2016).

211. Pistone, D. *et al.* Mycobacterium avium paratuberculosis in Italy: commensal or emerging human pathogen? *Dig Liver Dis* **44**, 461–465 (2012).

212. Jones, P. H., Farver, T. B., Beaman, B., Çetinkaya, B. & Morgan, K. L. Crohn's disease in people exposed to clinical cases of bovine paratuberculosis. *Epidemiol Infect* **134**, 49–56 (2006).

213. Qual, D. A., Kaneene, J. B., Varty, T. J., Miller, R. & Thoen, C. O. Lack of association between the occurrence of Crohn's disease and occupational exposure to dairy and beef cattle herds infected with Mycobacterium avium subspecies paratuberculosis. *J Dairy Sci* **93**, 2371–2376 (2010).

214. Singh, A. V., Singh, S. V., Singh, P. K., Sohal, J. S. & Singh, M. K. High prevalence of Mycobacterium avium subspecies paratuberculosis ('Indian bison type') in animal attendants suffering from gastrointestinal complaints who work with goat herds endemic for Johne's disease in India. *Int J Infect Dis* **15**, e677-683 (2011).

215. Rangel, J. M., Sparling, P. H., Crowe, C., Griffin, P. M. & Swerdlow, D. L. Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982–2002. *Emerg Infect Dis* **11**, 603–609 (2005).

216. Feller, M. *et al.* Mycobacterium avium subspecies paratuberculosis and Crohn's disease: a systematic review and meta-analysis. *Lancet Infect Dis* **7**, 607–613 (2007).

217. Abubakar, I., Myhill, D., Aliyu, S. H. & Hunter, P. R. Detection of Mycobacterium avium subspecies paratuberculosis from patients with Crohn's disease using nucleic acid-based techniques: a systematic review and meta-analysis. *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews* (2008).

218. Waddell, L. A. *et al.* The zoonotic potential of Mycobacterium avium spp. paratuberculosis: a systematic review. *Can J Public Health* **99**, 145–155 (2008).

219. Waddell, L. A., Rajić, A., Stärk, K. D. C. & McEWEN, S. A. The zoonotic potential of Mycobacterium avium ssp. paratuberculosis: a systematic review and meta-analyses of the evidence. *Epidemiol Infect* **143**, 3135–3157 (2015).

220. Weiser, T. G., Forrester, J. D. & Forrester, J. A. Tactics to Prevent Intra-Abdominal Infections in General Surgery. *Surg Infect (Larchmt)* **20**, 139–145 (2019).

221. Zamani, S. *et al.* Mycobacterium avium subsp. paratuberculosis and associated risk factors for inflammatory bowel disease in Iranian patients. *Gut Pathog* **9**, 1 (2017).

222. Banche, G. *et al.* Application of multiple laboratory tests for Mycobacterium avium ssp. paratuberculosis detection in Crohn's disease patient specimens. *New Microbiol* **38**, 357–367 (2015).

223. Carvalho, I. A. *et al.* Presence of Mycobacterium avium subsp. paratuberculosis (MAP) in Brazilian patients with inflammatory bowel diseases and in controls. *Sao Paulo Med J* **134**, 13–19 (2016).

224. Ricanek, P. *et al.* Paucity of mycobacteria in mucosal bowel biopsies from adults and children with early inflammatory bowel disease. *J Crohns Colitis* **4**, 561–566 (2010).

225. Singh, A. V. *et al.* Evaluation of "Indigenous Absorbed ELISA Kit" for the Estimation of Seroprevalence of Mycobacterium avium Subspecies paratuberculosis Antibodies in Human Beings in North India. *ISRN Vet Sci* **2011**, 636038 (2011).

226. Khan, I. A. *et al.* Prevalence and Association of Mycobacterium avium subspecies paratuberculosis with Disease Course in Patients with Ulcero-Constrictive Ileocolonic Disease. *PLoS One* **11**, e0152063 (2016).

227. Parrish, N. M. *et al.* Absence of mycobacterium avium subsp. paratuberculosis in Crohn's patients. *Inflamm Bowel Dis* **15**, 558–565 (2009).

228. Sasikala, M. *et al.* Absence of Mycobacterium avium ss paratuberculosis-specific IS900 sequence in intestinal biopsy tissues of Indian patients with Crohn's disease. *Indian J Gastroenterol* **28**, 169–174 (2009).

229. Mendoza, J. L. *et al.* High prevalence of viable Mycobacterium avium subspecies paratuberculosis in Crohn's disease. *World J Gastroenterol* **16**, 4558–4563 (2010).

230. Lee, A. *et al.* Association of Mycobacterium avium subspecies paratuberculosis with Crohn Disease in pediatric patients. *J Pediatr Gastroenterol Nutr* **52**, 170–174 (2011).

231. Molicotti, P. *et al.* Molecular identification of Mycobacterium avium subspecies paratuberculosis in oral biopsies of Crohn's disease patients. *Gut Pathog* **5**, 18 (2013).

232. Nazareth, N. *et al.* Increased viability but decreased culturability of Mycobacterium avium subsp. paratuberculosis in macrophages from inflammatory bowel disease patients under Infliximab treatment. *Med Microbiol Immunol* **204**, 647–656 (2015).

233. Timms, V. J., Daskalopoulos, G., Mitchell, H. M. & Neilan, B. A. The Association of Mycobacterium avium subsp. paratuberculosis with Inflammatory Bowel Disease. *PLoS One* **11**, e0148731 (2016).

234. Sharp, R. C., Beg, S. A. & Naser, S. A. Role of PTPN2/22 polymorphisms in pathophysiology of Crohn's disease. *World J Gastroenterol* **24**, 657–670 (2018).

235. Jeyanathan, M., Alexander, D. C., Turenne, C. Y., Girard, C. & Behr, M. A. Evaluation of In Situ Methods Used To Detect Mycobacterium avium subsp. paratuberculosis in Samples from Patients with Crohn's Disease. *J Clin Microbiol* **44**, 2942–2950 (2006).

236. Jeyanathan, M. *et al.* Visualization of Mycobacterium avium in Crohn's tissue by oil-immersion microscopy. *Microbes Infect* **9**, 1567–1573 (2007).

237. Magin, W. S., Van Kruiningen, H. J. & Colombel, J.-F. Immunohistochemical search for viral and bacterial antigens in Crohn's disease. *J Crohns Colitis* **7**, 161–166 (2013).

238. Chiodini, R. J., Van Kruiningen, H. J., Thayer, W. R. & Coutu, J. A. Spheroplastic phase of mycobacteria isolated from patients with Crohn's disease. *J Clin Microbiol* **24**, 357–363 (1986).

239. Kirkwood, C. D. *et al.* Mycobacterium avium subspecies paratuberculosis in children with early-onset Crohn's disease. *Inflamm Bowel Dis* **15**, 1643–1655 (2009).

240. Wagner, J. *et al.* Mycobacterium avium subspecies paratuberculosis in children with early-onset Crohn's disease: a longitudinal follow-up study. *Inflamm Bowel Dis* **17**, 1825–1826 (2011).

241. Motiwala, A. S. *et al.* Molecular Epidemiology of Mycobacterium avium subsp. paratuberculosis: Evidence for Limited Strain Diversity, Strain Sharing, and Identification of Unique Targets for Diagnosis. *J Clin Microbiol* **41**, 2015–2026 (2003).

242. Paustian, M. L. *et al.* Comparative genomic analysis of Mycobacterium avium subspecies obtained from multiple host species. *BMC Genomics* **9**, 135 (2008).

243. Singh, S. V., Sohal, J. S., Singh, P. K. & Singh, A. V. Genotype profiles of Mycobacterium avium subspecies paratuberculosis isolates recovered from animals, commercial milk, and human beings in North India. *Int J Infect Dis* **13**, e221-227 (2009).

244. Wynne, J. W. *et al.* Exploring the zoonotic potential of Mycobacterium avium subspecies paratuberculosis through comparative genomics. *PLoS One* **6**, e22171 (2011).

245. Bannantine, J. P. *et al.* Complete Genome Sequence of Mycobacterium avium subsp. paratuberculosis, Isolated from Human Breast Milk. *Genome Announc* **2**, e01252-13 (2014).

246. Timms, V. J., Hassan, K. A., Mitchell, H. M. & Neilan, B. A. Comparative genomics between human and animal associated subspecies of the Mycobacterium avium complex: a basis for pathogenicity. *BMC Genomics* **16**, 695 (2015).

247. Wynne, J. W. *et al.* SNP genotyping of animal and human derived isolates of Mycobacterium avium subsp. paratuberculosis. *Vet Microbiol* **172**, 479–485 (2014).

248. Amin, A. S. *et al.* Ecology and genomic features of infection with Mycobacterium avium subspecies paratuberculosis in Egypt. *Microbiology (Reading)* **161**, 807–818 (2015).

249. Verdier, J. *et al.* Specific IgG response against Mycobacterium avium paratuberculosis in children and adults with Crohn's disease. *PLoS One* **8**, e62780 (2013).

250. Lefrançois, L. H. *et al.* Characterization of the Mycobacterium avium subsp. paratuberculosis laminin-binding/histone-like protein (Lbp/Hlp) which reacts with sera from patients with Crohn's disease. *Microbes Infect* **13**, 585–594 (2011).

251. Xia, A. *et al.* Effect of inflammatory bowel disease therapies on immunogenicity of Mycobacterium paratuberculosis proteins. *Scand J Gastroenterol* **49**, 157–163 (2014).

252. Li, L. *et al.* Early detection of Mycobacterium avium subsp. paratuberculosis infection in cattle with multiplex-bead based immunoassays. *PLoS One* **12**, e0189783 (2017).

253. Sibartie, S. *et al.* Mycobacterium avium subsp. Paratuberculosis (MAP) as a modifying factor in Crohn's disease. *Inflamm Bowel Dis* **16**, 296–304 (2010).

254. Olsen, I. *et al.* Isolation of Mycobacterium avium subspecies paratuberculosis reactive CD4 T cells from intestinal biopsies of Crohn's disease patients. *PLoS One* **4**, e5641 (2009).

255. Olsen, I., Lundin, K. E. & Sollid, L. M. Increased frequency of intestinal CD4+ T cells reactive with mycobacteria in patients with Crohn's disease. *Scand J Gastroenterol* **48**, 1278–1285 (2013).

256. Nakase, H., Tamaki, H., Matsuura, M., Chiba, T. & Okazaki, K. Involvement of mycobacterium avium subspecies paratuberculosis in TNF-α production from macrophage: Possible link between MAP and immune response in Crohn's disease. *Inflamm Bowel Dis* **17**, E140–E142 (2011).

257. Ogura, Y. *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).

258. Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).

259. Coulombe, F. *et al.* Increased NOD2-mediated recognition of N-glycolyl muramyl dipeptide. *J Exp Med* **206**, 1709–1716 (2009).

260. Vinh, D. C. & Behr, M. A. Crohn's as an immune deficiency: from apparent paradox to evolving paradigm. *Expert Rev Clin Immunol* **9**, 17–30 (2013).

261. Singh, S. B., Davis, A. S., Taylor, G. A. & Deretic, V. Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science* **313**, 1438–1441 (2006).

262. Parkes, M. *et al.* Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* **39**, 830–832 (2007).

263. MacMicking, J. D., Taylor, G. A. & McKinney, J. D. Immune control of tuberculosis by IFN-gamma-inducible LRG-47. *Science* **302**, 654–659 (2003).

264. Feng, C. G. *et al.* Mice deficient in LRG-47 display increased susceptibility to mycobacterial infection associated with the induction of lymphopenia. *J Immunol* **172**, 1163–1168 (2004).

265. Cooney, R. *et al.* NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* **16**, 90–97 (2010).

266. Travassos, L. H. *et al.* Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* **11**, 55–62 (2010).

267. Bernstein, C. N., Wang, M.-H., Sargent, M., Brant, S. R. & Collins, M. T. Testing the interaction between NOD-2 status and serological response to Mycobacterium paratuberculosis in cases of inflammatory bowel disease. *J Clin Microbiol* **45**, 968–971 (2007).

268. Fava, V. M. *et al.* A Missense LRRK2 Variant Is a Risk Factor for Excessive Inflammatory Responses in Leprosy. *PLoS Negl Trop Dis* **10**, e0004412 (2016).

269. Lu, Y., Li, X., Liu, S., Zhang, Y. & Zhang, D. Toll-like Receptors and Inflammatory Bowel Disease. *Front Immunol* **9**, 72 (2018).

270. Nunberg, M. Y. *et al.* Impaired IL-10 Receptor-mediated Suppression in Monocyte From Patients With Crohn Disease. *J Pediatr Gastroenterol Nutr* **66**, 779–784 (2018).

271. Wagner, J. *et al.* TLR4, IL10RA, and NOD2 mutation in paediatric Crohn's disease patients: an association with Mycobacterium avium subspecies paratuberculosis and TLR4 and IL10RA expression. *Med Microbiol Immunol* **202**, 267–276 (2013).

272. Cao, B. L., Qasem, A., Sharp, R. C., Abdelli, L. S. & Naser, S. A. Systematic review and meta-analysis on the association of tuberculosis in Crohn's disease patients treated with

tumor necrosis factor-α inhibitors (Anti-TNFα). *World J Gastroenterol* **24**, 2764–2775 (2018).

273. Allen, A. J. *et al.* Experimental infection of a bovine model with human isolates of Mycobacterium avium subsp. paratuberculosis. *Vet Immunol Immunopathol* **141**, 258–266 (2011).

274. Athreya, S. P. K. Azathioprine in controlling type 2 reactions in leprosy: a case report. *Lepr Rev* **78**, 290–292 (2007).

275. Bach, H., Rosenfeld, G. & Bressler, B. Treatment of Crohn's disease patients with infliximab is detrimental for the survival of Mycobacterium avium ssp. paratuberculosis within macrophages and shows a remarkable decrease in the immunogenicity of mycobacterial proteins. *J Crohns Colitis* **6**, 628–629 (2012).

276. Bach, H. *et al.* Immunogenicity of Mycobacterium avium subsp. paratuberculosis proteins in Crohn's disease patients. *Scand J Gastroenterol* **46**, 30–39 (2011).

277. Nazareth, N. *et al.* Prevalence of Mycobacterium avium subsp. paratuberculosis and Escherichia coli in blood samples from patients with inflammatory bowel disease. *Med Microbiol Immunol* **204**, 681–692 (2015).

278. Clancy, R., Ren, Z., Turton, J., Pang, G. & Wettstein, A. Molecular evidence for Mycobacterium avium subspecies paratuberculosis (MAP) in Crohn's disease correlates with enhanced TNF-alpha secretion. *Dig Liver Dis* **39**, 445–451 (2007).

279. Qasem, A., Ramesh, S. & Naser, S. A. Genetic polymorphisms in tumour necrosis factor receptors (TNFRSF1A/1B) illustrate differential treatment response to TNFα inhibitors in patients with Crohn's disease. *BMJ Open Gastroenterol* **6**, e000246 (2019).

280. Campos, N. *et al.* Macrophages from IBD patients exhibit defective tumour necrosis factor-α secretion but otherwise normal or augmented pro-inflammatory responses to infection. *Immunobiology* **216**, 961–970 (2011).

281. Behr, M. A. & Hanley, J. Antimycobacterial therapy for Crohn's disease: a reanalysis. *Lancet Infect Dis* **8**, 344 (2008).

282. Feller, M. *et al.* Long-term antibiotic treatment for Crohn's disease: systematic review and meta-analysis of placebo-controlled trials. *Clin Infect Dis* **50**, 473–480 (2010).

283. Khan, K. J. *et al.* Antibiotic therapy in inflammatory bowel disease: a systematic review and meta-analysis. *Am J Gastroenterol* **106**, 661–673 (2011).

284. Alcedo, K. P., Thanigachalam, S. & Naser, S. A. RHB-104 triple antibiotics combination in culture is bactericidal and should be effective for treatment of Crohn's disease associated with Mycobacterium paratuberculosis. *Gut Pathog* **8**, 32 (2016).

285. Qasem, A., Safavikhasraghi, M. & Naser, S. A. A single capsule formulation of RHB-104 demonstrates higher anti-microbial growth potency for effective treatment of Crohn's disease associated with Mycobacterium avium subspecies paratuberculosis. *Gut Pathog* **8**, 45 (2016).

286. RedHill Biopharma. *RedHill Biopharma Announces Positive Top-Line Results from Phase III Study of RHB-104 in Crohn's Disease*. https://ir.redhillbio.com/news-releases/news-release-details/redhill-biopharma-announces-positive-top-line-results-phase-iii (2018).

287. RedHill Biopharma. *RedHill Biopharma Elaborates on Its Announced Positive Top-Line Results from Phase III Study of RHB-104 in Crohn's Disease*. https://ir.redhillbio.com/news-releases/news-release-details/redhill-biopharma-elaborates-its-announced-positive-top-

line?fbclid=IwAR2k397OF4VusXkvYS5553GPfvURrmZzpqmAKoFp5IIRZbJtpkgBe9Ii1xM (2018).

288. Fitzgerald, S. D. *et al.* Experimental Inoculation of Pigeons (Columba livia) with Mycobacterium bovis. *Avian Dis* **47**, 470–475 (2003).

289. Corner, L. a. L. *et al.* Experimental tuberculosis in the European badger (Meles meles) after endobronchial inoculation of Mycobacterium bovis: I. Pathology and bacteriology. *Res Vet Sci* **83**, 53–62 (2007).

290. Mackintosh, C. G. *et al.* Genetic Resistance to Experimental Infection with Mycobacterium bovis in Red Deer (Cervus elaphus). *Infect Immun* **68**, 1620–1625 (2000).

291. Nedeltchev, G. G. *et al.* Extrapulmonary Dissemination of Mycobacterium bovis but Not Mycobacterium tuberculosis in a Bronchoscopic Rabbit Model of Cavitary Tuberculosis. *Infect Immun* **77**, 598–603 (2009).

292. Cassidy, J. P. *et al.* Early lesion formation in cattle experimentally infected with Mycobacterium bovis. *J Comp Pathol* **119**, 27–44 (1998).

293. Wangoo, A. *et al.* Advanced Granulomatous Lesions in Mycobacterium bovis-infected Cattle are Associated with Increased Expression of Type I Procollagen, γδ (WC1+) T Cells and CD 68+ Cells. *J Comp Pathol* **133**, 223–234 (2005).

294. Waters, W. R. *et al.* Virulence of two strains of mycobacterium bovis in cattle following aerosol infection. *J Comp Pathol* **151**, 410–419 (2014).

295. Serrano, M. *et al.* Different lesion distribution in calves orally or intratracheally challenged with Mycobacterium bovis: implications for diagnosis. *Vet Res* **49**, 74 (2018).

296. Cassidy, J. P. *et al.* Lesions in cattle exposed to Mycobacterium bovis-inoculated calves. *J Comp Pathol* **121**, 321–337 (1999).

297. Bouté, M. *et al.* The C3HeB/FeJ mouse model recapitulates the hallmark of bovine tuberculosis lung lesions following Mycobacterium bovis aerogenous infection. *Vet Res* **48**, 73 (2017).

298. Chambers, M. A., Gavier-Widen, D. & Hewinson, R. G. Histopathogenesis of experimental Mycobacterium bovis infection in mice. *Res Vet Sci* **80**, 62–70 (2006).

299. Chen, X. *et al.* Aerosol inhalation of Mycobacterium bovis can reduce the Th2 dominant immune response induced by ovalbumin sensitization. *Am J Transl Res* **14**, 3430–3438 (2022).

300. Clarke, K.-A. R. *et al.* Experimental inoculation of meadow voles (Microtus pennsylvanicus), house mice (Mus musculus), and Norway rats (Rattus norvegicus) with Mycobacterium bovis. *J Wildl Dis* **43**, 353–365 (2007).

301. Lima, P. *et al.* Enhanced mortality despite control of lung infection in mice aerogenically infected with a Mycobacterium tuberculosis mce1 operon mutant. *Microbes Infect* **9**, 1285–1290 (2007).

302. Talaat, A. M., Wu, C.-W. & Hines II, M. E. Experimental animal models of paratuberculosis. in *Paratuberculosis: organism, disease, control* 213–247 (CAB International, 2020).

303. Vaughan, J. A., Lenghaus, C., Stewart, D. J., Tizard, M. L. & Michalski, W. P. Development of a Johne's disease infection model in laboratory rabbits following oral administration of Mycobacterium avium subspecies paratuberculosis. *Vet Microbiol* **105**, 207–213 (2005).

304. Arrazuria, R. *et al.* Vaccination sequence effects on immunological response and tissue bacterial burden in paratuberculosis infection in a rabbit model. *Vet Res* **47**, 77 (2016).

305. Arrazuria, R. *et al.* Effect of various dietary regimens on oral challenge with Mycobacterium avium subsp. paratuberculosis in a rabbit model. *Res Vet Sci* **101**, 80–83 (2015).

306. Tanaka, S., Itohara, S., Sato, M., Taniguchi, T. & Yokomizo, Y. Reduced formation of granulomata in gamma(delta) T cell knockout BALB/c mice inoculated with Mycobacterium avium subsp. paratuberculosis. *Vet Pathol* **37**, 415–421 (2000).

307. Wang, J., Pritchard, J. R., Kreitmann, L., Montpetit, A. & Behr, M. A. Disruption of Mycobacterium avium subsp. paratuberculosis-specific genes impairs in vivo fitness. *BMC Genomics* **15**, 415 (2014).

308. Eckelt, E. *et al.* FurA contributes to the oxidative stress response regulation of Mycobacterium avium ssp. paratuberculosis. *Front Microbiol* **6**, (2015).

309. Mutwiri, G. K. *et al.* Mycobacterium avium Subspecies Paratuberculosis Triggers Intestinal Pathophysiologic Changes in Beige/Scid Mice. *Comp Med* **51**, 538–544 (2001).

310. Cooney, M. A., Steele, J. L., Steinberg, H. & Talaat, A. M. A murine oral model for Mycobacterium avium subsp. paratuberculosis infection and immunomodulation with Lactobacillus casei ATCC 334. *Front Cell Infect Microbiol* **4**, 11 (2014).

311. Veazey, R. S. *et al.* Histopathology of C57BL/6 mice inoculated orally with Mycobacterium paratuberculosis. *J Comp Pathol* **113**, 75–80 (1995).

312. Mundy, R., Girard, F., FitzGerald, A. J. & Frankel, G. Comparison of colonization dynamics and pathology of mice infected with enteropathogenic Escherichia coli, enterohaemorrhagic E. coli and Citrobacter rodentium. *FEMS Microbiol Lett* **265**, 126–132 (2006).

313. Barthel, M. *et al.* Pretreatment of Mice with Streptomycin Provides a Salmonella enterica Serovar Typhimurium Colitis Model That Allows Analysis of Both Pathogen and Host. *Infect Immun* **71**, 2839–2858 (2003).

314. Rue-Albrecht, K. *et al.* Comparative Functional Genomics and the Bovine Macrophage Response to Strains of the Mycobacterium Genus. *Front Immunol* **5**, (2014).

315. North, R. J., Ryan, L., LaCource, R., Mogues, T. & Goodrich, M. E. Growth Rate of Mycobacteria in Mice as an Unreliable Indicator of Mycobacterial Virulence. *Infect Immun* **67**, 5483–5485 (1999).

316. Jilani, T. N., Avula, A., Zafar Gondal, A. & Siddiqui, A. H. Active Tuberculosis. in *StatPearls* (StatPearls Publishing, 2023).

317. Ghanavi, J., Farnia, P., Farnia, P. & Velayati, A. A. The Role of Interferon-Gamma and Interferon-Gamma Receptor in Tuberculosis and Nontuberculous Mycobacterial Infections. *Int J Mycobacteriol* **10**, 349 (2021).

318. Dalton, D. IFN-γ and IFN-γ Receptor Knockout Mice. in *Cytokine Knockouts* (ed. Fantuzzi, G.) 347–359 (Humana Press, 2003).