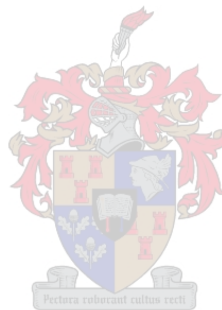
Characterizing the proteomes of selected members of the *Mycobacterium tuberculosis* complex

By

Louise Botha

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Medical Science (Molecular Biology) in the Faculty of Medicine and Health Sciences at Stellenbosch University



Supervisor: Prof R.M. Warren
Co-supervisor: Prof N.C. Gey van Pittius

April 2014

Stellenbosch University http://scholar.sun.ac.za

DECLARATION

I, the undersigned, hereby declare that the work contained in this assignment is my original work and that it has not previously in its entirety or in part been submitted at any university for a degree.

VERKLARING

Ek, die ondergetekende, verklaar hiermee dat die werk in hierdie werkstuk vervat, my oorspronklike werk is en dat dit nie vantevore in die geheel of gedeeltelik by enige universiteit ter verkryging van 'n graad voorgelê is nie.

Signature: .... **Date: ....**

Copyright © 2014 Stellenbosch University

All rights reserved

Stellenbosch University http://scholar.sun.ac.za

Abstract

Mycobacterium tuberculosis is a pathogenic organism that infects a third of the world's population and causes approximately 2 million deaths per year. This pathogen is a member of the *Mycobacterium tuberculosis* complex (MTBC) which constitutes eleven members that share 99.9% similarity at nucleotide level and have near identical 16S rRNA. MTBC members cause Tuberculosis in a variety of host species. *M. bovis* and *M. caprae* form part of the animal-adapted MTBC members that cause disease in a variety of animal hosts (primarily bovidae) and goats, respectively. Extensive genetic analyses have been done to try and explain virulence, phenotype and host-preferences of these members with no success. Recent advances in mass spectrometry techniques enable us to analyse thousands of proteins simultaneously and explore the possible proteomic variation between these members that could contribute to the phenotypic, virulence and host-specificity characteristics of the MTBC members.

In this study, we aimed to characterize the proteomes of *M. bovis* and *M. caprae* by analysing the high and or low abundance proteins, relative to *M. tuberculosis* H37Rv, which could possibly explain virulence mechanisms and host-specificity of these MTBC members. Whole cell lysate protein extracts were extracted from mid-log phase cultures of *M. tuberculosis* H37Rv (A_{600} = 0.7), *M. bovis* (A_{600} = 0.65) and *M. caprae* (A_{600} = 0.7). Proteins were fractionated by SDS-PAGE and in gel reduction/alkylation and trypsin digests were done. Peptides were identified using LC-MS/MS on the Orbitrap Velos mass spectrometer and their corresponding proteins were identified by searching peptide databases. Protein functional groups were assigned according to TubercuList. To provide an integrated overview of the overall network of protein expression (rather than just limit analysis to individual proteins), pathway analysis was done on the differentially expressed proteins of

Stellenbosch University http://scholar.sun.ac.za

M. bovis and *M. caprae* using PATRIC (Pathosystems Resource Integration Center) and pathways were visualized using iTUBY (Interactive Pathway Explorer database).

We detected 2199, 2367 and 2350 proteins for *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* which correlate to 60% of the proposed *M. tuberculosis* proteins being expressed during log-phase. Considering similarities between genomes, it was no surprise that the functional distribution of the detected proteins extracted was similar. Metabolic pathways affected by the proteins which were in higher abundance in *M. bovis* and *M. caprae* included amino acid and lipid metabolism, oxidative phosphorylation and xenobiotic degradation. The over-abundant proteins in *M. bovis* and *M. caprae* were also involved in ribosomal proteins and carbohydrate metabolism, respectively. Lower abundance proteins in these species were found in lipid and pyrimidine metabolism. These affected pathways can be associated with the ability of *M. bovis* and *M. caprae* to adapt to their environment more readily which helps them to survive inside the hosts and cause severe pathogenesis.

In this study the proteomes of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* were characterized and the variation between detected proteins and protein abundances explored in order to describe differences between these closely related strains. Future research on animal-adapted Mycobacterial species will address knowledge gaps that are needed to prevent transmission and spread of the disease. Understanding the mechanisms of virulence and pathogenicity could lead to development of efficient vaccines and diagnostic tests for a variety of animal hosts.

Opsomming

Mycobacterium tuberculosis is 'n patogene organisme wat 'n derde van die wêreld se bevolking infekteer en veroorsaak ongeveer 2 miljoen sterftes per jaar. Hierdie patogeen is 'n lid van die *Mycobacterium tuberculosis* kompleks (MTBK) wat bestaan uit elf lede wat 99,9% ooreenkoms op nukleotiedvlak toon en amper identiese 16S rRNA deel. MTBK lede veroorsaak Tuberkulose in 'n verskeidenheid van gasheerspesies. *M. bovis* en *M. caprae* vorm deel van die MTBK en veroorsaak Tuberkulose in 'n verskeidenheid van diere-gashere (hoofsaaklik Bovidae) en bokke, onderskeidelik. Verskeie genetiese ontledings is al gedoen om virulensie, fenotipe en gasheer-voorkeure van hierdie lede te ondersoek, maar was onsuksesvol. Die onlangse vooruitgang in massa-spektrometriese tegnieke stel ons in staat om duisende proteïene gelyktydig te analiseer en die moontlike proteomiese variasie tussen hierdie lede te identifiseer. Proteomiese analises kan bydra tot die fenotipiese-, virulensie- en gasheer-spesifieke eienskappe van die hierdie lede.

Die doel met hierdie studie was om die proteome van *M. bovis* en *M. caprae* te beskryf deur die proteïene te identifiseer wat differentieel uitgedruk was, in vergelyking met *M. tuberculosis* H37Rv, wat moontlik die virulensie meganismes en gasheer-spesifisiteit van hierdie MTBK lede kan verduidelik. Proteïen ekstraksies is geneem uit die middel-logaritmiese groeifase van *M. tuberculosis* H37Rv (A_{600} = 0.7), *M. bovis* (A_{600} = 0,65) en *M. caprae* (A_{600} = 0,7) kulture. Proteïene is gefraksioneer deur SDS-PAGE en in-jel vermindering/alkilering en tripsien vertering is gedoen. Peptiede is geïdentifiseer met behulp van LC-MS/MS op die Orbitrap Velos massa-spektrometer en die ooreenstemmende proteïene is geïdentifiseer. Proteïen funksionele groepe is toegeken aan proteïene volgens TubercuList. Om 'n geïntegreerde oorsig van die totale netwerk van die proteïen uitdrukking te gee (eerder as ontleding van slegs individuele proteïene), is metaboliese weë analises op die differensieel uitgedrukte proteïene van *M. bovis* en *M. caprae* gedoen, deur gebruik te

maak van PATRIC (Pathosystems Resource Integration Center). Metaboliese weë is gevisualiseer deur iTUBY (Interactive Pathway Explorer databasis).

‘n Totaal van 2199, 2367 en 2350 proteïene is ontdek vir *M. tuberculosis* H37Rv, *M. bovis* en *M. caprae* onderskeidelik, wat ooreenstem met 60% van die voorgestelde *M. tuberculosis* proteïene. A.g.v. genoom ooreenkomste, was dit geen verrassing dat die funksionele verspreiding van die proteïene soortgelyk was nie. Metaboliese weë wat geraak word deur die proteïene wat in hoë-oorvloed in *M. bovis* en *M. caprae* ontdek is, sluit die aminosuur- en lipiedmetabolisme, oksidatiewe fosforilering en xenobiotiese afbreking in. Die hoë-oorvloed proteïene in *M. bovis* en *M. caprae* is ook betrokke by ribosomale funksies en koolhidraatmetabolisme, onderskeidelik. Proteiene wat in laer-oorvloed in hierdie twee spesies geidentifiseer is, speel ‘n rol in lipied- en pirimidienmetabolisme Geaffekteerde metaboliese weë kan geassosieer word met die vermoë van *M. bovis* en *M. caprae* om meer geredelik by hul omgewing aan te pas wat die organimses help om te oorleef in die gasheer en patogenese te ontwikkel.

In hierdie studie is die proteome van *M. tuberculosis* H37Rv, *M. bovis* en *M. caprae* beskryf en die variasie tussen die ontdekte proteïene en proteïen verspreidings ontleed om die verskille tussen hierdie nou verwante spesies te beskryf. Toekomstige navorsing op diere-aangepaste mikobakteriële spesies sal die kennisgapings oorbrug wat nodig is om die oordrag en verspreiding van Tuberkulose te voorkom. Begrip van die meganismes van virulensie en patogenisiteit kan lei tot die ontwikkeling van doeltreffende entstowwe en diagnostiese toetse vir 'n verskeidenheid van diere-gashere.

Stellenbosch University http://scholar.sun.ac.za

Acknowledgements

My wonderful family, how will I ever be able to thank you for your immense support, love and understanding during this degree. To my Mom and Dad, Helene and Gawie, you have taught me the value of believing in yourself and hard work. Thank you for installing a degree of integrity, kindness and ambition in my heart. Without these traits I would not have been able to come this far. I am proud to be your daughter and thank you for your wonderful example in your professional career as well as your personal life. To my sister, Heslé, you have been such an inspiration in my life. Thank you for words of encouragement when needed, keeping me on my toes and making sure I succeed in whatever I decide to do. You are the best. I love you.

To my husband, Raynardt Botha, thank you for being my best friend and understanding my nerdy side. Your love and support has been my rock during this degree. Thank you for listening to my practice presentations, getting excited about my experiment results and understanding late nights working on my articles and thesis. I love you.

To my boss, Prof Paul van Helden, words cannot describe the impact your leadership and guidance has had on my life. Thank you for being my mentor in my scientific career. You have taught me the value of good research, work friendships and team dynamics. I want to thank you for sharing my enthusiasm for this project and helping me during the difficult times where I needed to “step-up” as a scientist. I hope that we will work together for many years, as you have inspired me to be the best I can be.

I would like to thank Prof Rob Warren and Prof Nico Gey van Pittius for their support during my MSc. Your contribution to my development, both as a scientist and an adult, has been invaluable. Thank you for believing in this project and nurturing my love for science.

Stellenbosch University http://scholar.sun.ac.za

Thank you to my wonderful lab friends Mae, Anzaan, Nastassja and Zhuo for helping me with experiments, reading of my thesis/articles and general stimulating conversations at lunch (workwise or otherwise). Most of all, thank you for your kindness and friendship. I look forward to building our scientific careers together.

Thank you to Dr Salomé Smit for the proteomic work done for this project and for going the extra mile in helping me with data analysis and interpretation.

I would like to thank the National Research Foundation for awarding me with a full NRF/DST Innovation Scholarship during my MSc degree. Thank you to Harry Crossley for awarding me project funding for my research.

Table of Contents

Declaration i

Abstract ii

Opsomming iv

Acknowledgements vi

Table of Contents viii

List of Figures xii

List of Tables xiv

List of Abbreviations xv

Chapter 1: Introduction 1

1.1 Introduction 2

1.2 Problem statement 4

1.3 Aims and Objectives 5

Chapter 2: Literature review titled "Mycobacteria and disease in Southern Africa" 6

Article front page 7

2.1 Abstract 8

2.2 Introduction 8

2.3 *Mycobacterium tuberculosis* complex members 11

2.4 Non-tuberculous Mycobacteria 14

Stellenbosch University http://scholar.sun.ac.za

2.5 Geographical distribution and environmental determinants 18

2.6 Hosts of Mycobacteria 20

2.7 Clinical signs and diagnoses 22

2.8 Conclusion 25

Chapter 3: Materials and Methods **26**

3.1 Bacterial strains 27

3.2 Media and culture conditions 27

3.3 Verification of Mycobacterial strains 28

3.3.1 Ziehl-Neelsen staining 28

3.3.2 Blood agar plates 29

3.3.3 Spoligotyping 29

3.3.4 Regions of difference PCR 30

3.3.5 IS*6110* DNA fingerprinting 31

3.4 Growth curves of Mycobacterial strains 31

3.5 Whole cell lysate protein extraction 32

3.6 Protein concentration determination 33

3.7 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis 33

3.7.1 SDS-PAGE sample preparation 33

3.7.2 Electrophoresis and staining 34

3.8 Sample preparation for Mass Spectrometry (MS) analyses 34

3.8.1 Peptide extraction 34

3.8.2 Stage-tip activation 36

3.9 Mass spectrometry and data analysis 37

3.10 Bio-informatics 39

Chapter 4: Results **40**

4.1 *M. tuberculosis* complex species verification 41

4.2 Proteomic analysis of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* 43

4.2.1 Whole cell lysate proteomic analysis 43

4.2.2 Analysis of the abundance of whole cell lysate proteins 49

4.3 Pathway analysis of differentially abundant proteins 55

4.4 RD deletions 60

Chapter 5: Discussion **62**

Chapter 6: Limitations and Future directions **71**

6.1 Limitations 72

6.2 Future directions 73

Stellenbosch University http://scholar.sun.ac.za

Chapter 7: List of References **74**

Appendix A: Media, Reagents and Solutions 85

Appendix B (disk included): Protein datasets 90

Stellenbosch University http://scholar.sun.ac.za

List of Figures

Figure 2.1 The genus Mycobacterium.

Figure 2.2 The phylogeny of the *M. tuberculosis* complex (MTBC) members based on genome differences, which can form the basis of differential diagnostics.

Figure 3.1 Spoligotyping signatures of the *M. tuberculosis* complex members.

Figure 3.8.1 Fractionation of gel lanes.

Figure 4.1.1. Unique spoligotyping signatures of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*.

Figure 4.1.2 IS*6110* RFLP results for *M. bovis* and *M. caprae*.

Figure 4.2.1.1 SDS-PAGE results for (a) *M. tuberculosis* H37Rv, (b) *M. bovis* and (c) *M. caprae*.

Figure 4.2.1.2 Whole cell lysate proteins of *M. tuberculosis* H37Rv grouped in functional categories.

Figure 4.2.1.3 Whole cell lysate proteins of *M. bovis* grouped in functional categories.

Figure 4.2.1.4 Whole cell lysate proteins of *M. caprae* grouped in functional categories.

Figure 4.2.1.5 A Venn diagram showing the number of proteins detected in each of the three species, differentiating those that are common or unique amongst the three species.

Figure 4.2.1.6 Distribution of uniquely detected proteins in *M. tuberculosis* H37Rv (green), *M. bovis* (blue) and *M. caprae* (orange).

Figure 4.2.1.7 Distribution of shared proteins between *M. bovis* and *M. caprae.*

Stellenbosch University http://scholar.sun.ac.za

Figure 4.2.2.1 Functional groups of proteins detected in higher and lower abundance in *M. bovis*, relative to *M. tuberculosis* H37Rv.

Figure 4.2.2.2 Functional groups of proteins detected in higher and lower abundance in *M. bovis*, relative to *M. tuberculosis* H37Rv.

Figure 4.2.2.3 Venn diagrams showing the number of proteins which were detected in (a) higher abundance or (b) lower abundance in *M. bovis* and *M. caprae* relative to *M. tuberculosis*, indicating the number which showed similar expression profiles in both strains relative to *M. tuberculosis.*

Figure 4.2.2.4 Functional categories of over-or under-abundant proteins detected only in *M. bovis* or *M. caprae*.

Figure 4.3.1 Pathway analysis of higher abundance proteins of *M. bovis* compared to *M. tuberculosis* H37Rv.

Figure 4.3.2 Pathway analysis of lower abundance proteins of *M. bovis* compared to *M. tuberculosis* H37Rv.

Figure 4.3.3 Pathway analysis of higher abundance proteins in *M. caprae* compared to *M. tuberculosis* H37Rv.

Figure 4.3.4 Pathway analysis of lower abundance proteins of *M. caprae* compared to *M. tuberculosis* H37Rv.

Stellenbosch University http://scholar.sun.ac.za

List of Tables

Table 2.1 *M. tuberculosis* complex species and their host-association.

Table 3.1 PCR primer sequences for regions of deletion (RD) speciation of *M. tuberculosis* complex members.

Table 4.2.2.1 Summary of higher and lower-abundance proteins detected in only one species in functional categories that differed significantly between *M. bovis* and *M. caprae*.

Table 4.3.1 Summary of pathways affected by the higher abundance proteins in *M. bovis* relative to *M. tuberculosis.*

Table 4.3.2 Summary of pathways affected by the lower abundant proteins in *M. bovis.*

Table 4.3.3 Summary of pathways affected in *M. caprae* using the higher abundant proteins.

Table 4.3.4 Summary of pathways affected by the lower abundant proteins in *M. caprae*.

Table 4.4 Regions of deletion and their associated proteins.

Stellenbosch University http://scholar.sun.ac.za

List of Abbreviations

°C	Degrees Celcius
µg	Microgram
µl	Microliter
7H9	Difco™ Middlebrooks 7H9 Broth
AIDS	Acquired Immune Deficiency Syndrome
ATP	Adenosine triphosphate
ATS/IDSA	American Thoracic Society/ International Disease Society of America
BCG	Bacillus Calmette–Guérin
CFP-10	Culture filtrate protein 10
DOTS	Directly observed treatment, short-course
ESAT-6	Early secretory antigenic target of 6 kDa
et al.	And others
g	Gravitational acceleration
IFN-γ	Interferon-gamma
IGRA	Interferon gamma reaction assay
M.	Mycobacterium
MAC	*Mycobacterium avium* complex

Stellenbosch University http://scholar.sun.ac.za

$MgCl_2$	Magnesium Chloride
min	Minute or minutes
MIRU-VNTR	Mycobacterial interspersed repetitive units-Variable number of tandem repeats
ml	Milliliter
MTBC	*Mycobacterium tuberculosis* complex
mQH_2O	MilliQ water
ng	Nanogram
NTM	Non-tuberculous Mycobacterium
NTMs	Non-tuberculous Mycobacteria
OD	Optical density
PCR	Polymerase chain reaction
PPD	Tuberculin purified protein derivative
RD	Region of Difference
RFLP	Restriction fragment length polymorphism
rRNA	Ribosomal ribonucleic acid
s	Second or seconds
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Stellenbosch University http://scholar.sun.ac.za

Spoligotyping	Spacer oligonucleotide typing
TB	Tuberculosis
TBE	Tris-Borate-EDTA buffer
TST	Tuberculin skin test
Tween-80	Poly-oxy-ethylene sorbian mono-oleate
USA	United States of America
V	Volt
WCL	Whole cell lysate
XT MOPS	3-(N-morpholino)-propanesulfonic acid buffer
ZN	Ziehl-Neelsen

Stellenbosch University http://scholar.sun.ac.za

Chapter 1

Introduction

Stellenbosch University http://scholar.sun.ac.za

1.1 Introduction

In 1993, a global emergency was declared by the World Health Organization (WHO) because tuberculosis (TB) was epidemic in many areas of the world. Now, two decades later, TB remains an important public health problem that results in millions of deaths each year (World Health Organization). Zoonotic TB is present in animals in most developing countries where surveillance and control activities are often inadequate or unavailable; therefore, animal tuberculosis can have devastating effects on national economy and agricultural industries (Cosivi, Grange *et al*. 1998).

Tuberculosis is caused in different hosts by different members of the *Mycobacterium tuberculosis* complex (MTBC) which are characterized by 99.9% similarity at the nucleotide level and near identical 16S rRNA sequences (Parsons, Drewe *et al.* 2013, Alexander, Laver *et al*. 2010). However, these members differ in terms of host specificity, phenotype and pathogenicity (Brosch, Gordon *et al*. 2002). The MTBC includes *M. tuberculosis*, *M. africanum*, *M. microti* and *M. bovis* and newly recognized members include *M. caprae*, *M. pinnipedii* and *M. canettii* and *M. origys* (Huard, Fabre *et al*. 2006, Gey van Pittius, van Helden *et al*. 2012). Rare MTBC variants, the so-called dassie bacillus, *M. mungi*, and a newly identified member isolated from meerkats (*Suricata suricatta)* (Parsons, Drewe *et al*. 2013) have been described (Alexander, Laver *et al*. 2010). The origin and evolution of the MTBC have intrigued researchers since their discovery. It is assumed that the MTBC members are derived from a common ancestor, but interestingly some members, for example *M. tuberculosis*, *M. africanum* and *M. canettii*, are exclusively found in humans whilst other members, like *M. bovis,* are seen as promiscuous, infecting a wide range of hosts (Brosch, Gordon *et al*. 2002). Other members are described as being host-specific, including *M. caprae*, *M. origys* and *M. mungi* which cause disease in goats, Arabian oryxes and banded mongoose, respectively (Huard, Fabre *et al*. 2006).

Standard molecular tools used to delineate and speciate the MTBC include IS*6110*-restriction fragment length polymorphism (RFLP) analysis, spacer oligonucleotide typing (Spoligotyping), MIRU-VNTR typing and regions of difference (RD) analysis. RD analysis of the MTBC members can distinguish between so-called human pathogens and animal-adapted pathogens (Brosch, Gordon *et al.* 2002). All species that are primarily animal-adapted belong to a single lineage that can be distinguished by the deletion of the chromosomal RD9, which is intact in all *M. tuberculosis* strains (Smith, Hewinson *et al.* 2009). These animal-adapted strains include *M. mungi* (mongoose bacillus), dassie bacillus, *M. origys* (Arabian oryx bacillus), *M. microti* (vole bacillus), *M. pinnipedii* (seal bacillus), *M. caprae* (goat bacillus) and *M. bovis* (Smith, Hewinson *et al*. 2009).

Whilst the incidence and prevalence of *M. bovis* cases in humans is far lower than *M. tuberculosis*, this pathogen is problematic in various settings in animal and human populations in a number of countries. Infection of animals by members of the MTBC is gaining greater recognition because these infections contribute to the high TB burden (Skinner, Wedlock *et al*. 2001). One such member from the MTBC is *M. caprae*. *M. caprae* infected goat herds can constitute a reservoir of TB inducing mycobacteria in the field, posing a risk of infection to cattle and wildlife (Cvetnic, Katalinic-Jankovic *et al*. 2007, Rodriguez, Bezos *et al*. 2011). Interestingly, *M. caprae* and *M .bovis* share most of the RD deletions, except for RD4 which is uniquely deleted in *M. bovis* (Smith, Hewinson *et al.* 2009)(Smith, Hewinson *et al*. 2009, Smith, Yatsunenko *et al*. 2013) The similarity between the genomes of these two members with differing disease phenotypes provides an opportunity to study and attempt to explain how *M. bovis* is able to infect a wide range of hosts, whereas *M. caprae* predominantly infects goats. However, genomic analysis of closely related species may not directly explain specific phenotypes, host preferences or pathogenicity.

Stellenbosch University http://scholar.sun.ac.za

Recent advances in mass spectrometry allow the simultaneous analysis of thousands of proteins. This technique enables the investigation of the strain differences between MTBC members at the level of the proteome, to explore possible proteomic variation which contributes to phenotype, host preferences and pathogenicity. The relationship between single nucleotide polymorphisms and protein abundance is not known and thus the analyses of final gene products could inform us what pathways are being used to regulate virulence and pathogenesis during infection in the host.

1.2 Problem statement

M. tuberculosis and *M. bovis* are the primary causative agents of tuberculosis in humans and animals, respectively. These species form part of the *M. tuberculosis* complex which shares 99.9% similarity at nucleotide level as well as near identical 16S rRNA sequences. Extensive genetic resources have not been able to clarify the mechanisms of pathogenicity between these members. We propose that analysing the proteins expressed in selected members of the *M. tuberculosis* complex the human-adapted *M. tuberculosis* H37Rv, the promiscuous animal-adapted *M. bovis* and the host-specific animal–adapted *M. caprae*; will enhance our understanding of what influences host-specificity in these members, as well as contribute to the knowledge of the virulence mechanisms used by these members.

Stellenbosch University http://scholar.sun.ac.za

1.3 Aims of study

1. To characterize the proteomes of selected members of the MTBC, namely *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*.

2. To compare abundances of proteins identified in *M. bovis* and *M. caprae*, relative to *M. tuberculosis* H37Rv.

3. To identify proteins which may be involved in host-specificity and virulence mechanisms in *M. bovis* and *M. caprae*.

Stellenbosch University http://scholar.sun.ac.za

Chapter 2

Literature review

Stellenbosch University http://scholar.sun.ac.za

Mycobacteria and disease in Southern Africa

Botha L[1], Gey van Pittius N. C.[1], van Helden P.D.[1]

[1]DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ Medical Research Council (MRC) Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University, PO Box 19063, Tygerberg, South Africa, 7505

Published in Transboundary and Emerging Diseases (2013), 60:147-156.

2.1 Abstract

The genus Mycobacterium consists of over 120 known species, some of which (e.g. *M. bovis* and *M. tuberculosis*) contribute extensively to the burden of infectious disease in humans and animals, whilst others are commonly found in the environment but may rarely if ever be disease-causing. This paper reviews the Mycobacteria found in Southern Africa, focussing on those in the *M. tuberculosis* complex as well as the non-tuberculous Mycobacteria (NTMs) identifying those found in the area and including those causing disease in humans and animals, and outlines some recent reports describing the distribution and prevalence of the disease in Africa. Difficulties in diagnosis, host preference and reaction, immunology and transmission are discussed.

2.2 Introduction

Tuberculosis (TB) commonly occurs in humans and animals throughout the world, with Africa and particularly South Africa being a very high burden region with the incidence of *M. tuberculosis* in humans averaging 913/100 000 per annum (World Health Organization). Detailed and thorough prevalence surveys are rarely done, but the few that have been done show that infection with *M. tuberculosis*, the causative agent for TB in humans, is far more common than estimated from passive case finding (den Boon, van Lill *et al*. 2007). South Africa currently has the highest global burden of human disease from *M. tuberculosis.* It is not the purpose of this manuscript to reiterate that situation, other than to remark that the situation is complex, given that we know that the genus *Mycobacterium*, to which *M. tuberculosis* belongs, consists of a larger group of species, most of which have been shown to be able to cause different forms of Mycobacterial disease under the appropriate conditions (Petrini 2006, Moore, Kruijshaar *et al*. 2010, Phillips, von Reyn 2001, Simons, van Ingen *et al*. 2011, Warren, Gey van Pittius *et al*. 2006a). The fast growing species (less than 7 days in culture) are often non-pathogenic, whereas slow-growing species (more than 7

days in culture) are mostly pathogenic (see Figure 2.1). The slow growing species include the members of the *M. tuberculosis* complex. Other species outside the complex are known as the non-tuberculous Mycobacteria (NTMs). Prevalence surveys, where Mycobacterial speciation has been done, have shown that non-tuberculous Mycobacteria (NTMs) may be up to three times more commonly found in humans than *M. tuberculosis* (Muyoyeta, de Haas *et al*. 2010) and own unpublished data). Similarly, although *M. bovis* is the most common cause of tuberculosis in animals, it cannot always be assumed to be the causative agent of tuberculosis in animals.

The genus *Mycobacterium* currently consists of around 128 validly-published species and 5 subspecies with at least a further 34 species not fully described or named (van Helden, Parsons *et al*. 2009). They are commonly encountered in the environment, where niches include water, soil, protozoans, domestic and wild animals, invertebrates, and milk and food products (Falkinham 2010, Holland 2001, Michel, de Klerk *et al*. 2007). Some are obligate or opportunistic pathogens, but many are saprophytes. Apart from the common and well-known pathogenic Mycobacteria that are primarily transmitted between hosts, the environment can be a source of exposure to Mycobacteria. We have investigated water as a source for such and found *M. terrae*, *M. vaccae* (or *vanbaalenii*), *M. engbaekii* and *M. thermoresistibile* (Michel, de Klerk *et al*. 2007). Although not extensively investigated, there is clearly regional variation in the occurrence of Mycobacterial species e.g. *M. ulcerans* has not (yet) been found in southern Africa. However, with the movement of animals (and birds) often uncontrolled, one can expect the unexpected (e.g. the occurrence of *M. origys* in South Africa (Gey van Pittius, Perrett *et al*. 2012).

Figure 2.1: The genus Mycobacterium. Well-known pathogens are mostly slow-growers. The species marked with a * are NTMs isolated in South Africa from animals. Figure adapted from (Gey van Pittius, Sampson *et al*. 2006).

2.3 *Mycobacterium tuberculosis* complex members

Mycobacteria grouped in the *Mycobacterium tuberculosis* complex (MTBC) are characterized by what is regarded as more than 99% similarity at the nucleotide level and near identical 16S rRNA sequences. The traditional members of the *M. tuberculosis* complex include the following four species: *M. tuberculosis*, *M. africanum*, *M. microti* and *M. bovis*. Newly recognized additions to the MTBC include *M. caprae*, *M. pinnipedii* and *M. canettii,* as well as potentially rarer or perhaps more geographically restricted MTBC variants, such as the so-called dassie (Parsons, Smith *et al*. 2008) and oryx bacillus, now known as *M. origys*, (Gey van Pittius, van Helden *et al*. 2012, van Ingen, Rahim *et al*. 2012) as well as *M. mungi*, which causes TB in the banded mongoose (Alexander, Laver *et al*. 2010).

Infection of animals or humans by the members of the MTBC, most particularly *M. tuberculosis* and *M. bovis,* is clearly of concern, since these are the commonly recognised causes of the massive global burden of TB. Standard molecular tools used to delineate and speciate the MTBC are IS*6110*-restriction fragment length polymorphism (RFLP) analysis; spacer oligonucleotide typing (Spoligotyping); Regions of difference (RD) analysis and MIRU-VNTR typing. These tools can be used to position the different members of the MTBC into the phylogenetic tree of the complex (Figure 2.2).

Our work shows that by far the majority of human tuberculosis cases (from *M. tuberculosis*) in South Africa originate from active and on-going transmission, and not reactivation disease and that this dynamic is valid for both antibiotic sensitive and resistant *M. tuberculosis* cases. (Hanekom, van der Spuy *et al*. 2007). Despite this enormous infection pressure, there have been very few cases of confirmed anthropozoonotic TB in South Africa specifically, viz. a single case of *M. tuberculosis* in a dog (Parsons, Gous *et al*. 2008) and *M. tuberculosis* in a zoo elephant (Michel, Venter *et al.* 2003, Tordiffe A *et al.*, unpublished).

However, a number of *M. tuberculosis* cases in livestock have been reported in Ethiopia and other African countries (Gumi, Schelling *et al*. 2012).

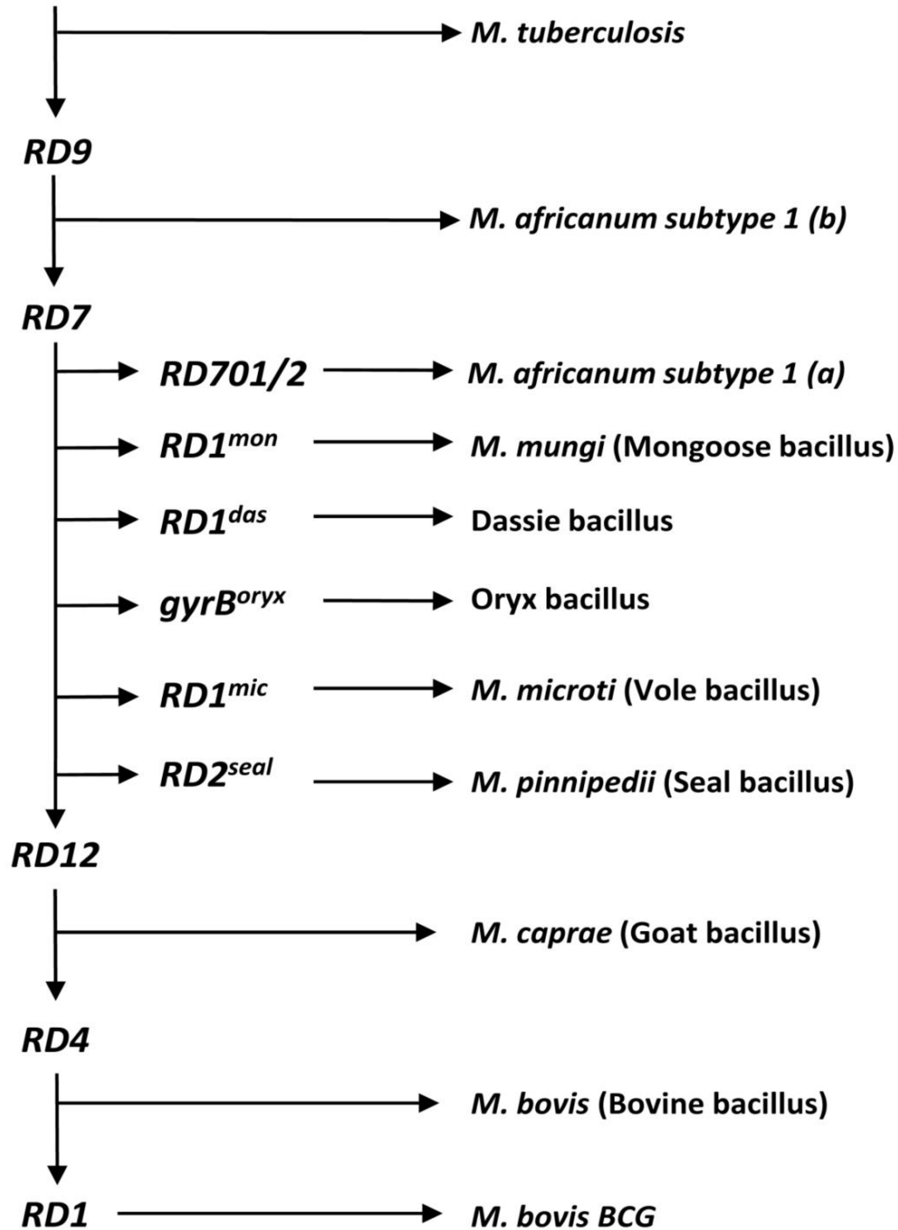


Figure 2.2: The phylogeny of the *Mycobacterium tuberculosis* complex (MTBC) members based on genome differences, which can form the basis of differential diagnostics. Figure adapted from (Gey van Pittius, Perrett *et al*. 2012).

Whilst the incidence and prevalence of *M. bovis* cases in humans is far lower than *M. tuberculosis*, this pathogen is problematic in various settings in animal and human populations in the different countries. However, in southern African countries (here defined as South Africa, Namibia, Zimbabwe, Zambia, Tanzania, and Mozambique) there are some considerable differences. In South Africa, almost all *M. bovis* occurs in the wildlife population, due to good disease control practice in the past, but historical invasion into the African buffalo (Michel, de Klerk *et al*. 2007, Michel, Coetzee *et al*. 2009). Whilst it has spread into at least 12 mammal species, it occurs predominantly in the wildlife parks and there is as yet no evidence to suggest spread across the interface into humans or domestic livestock (Michel, Muller *et al*. 2010). There are sporadic occurrences of *M. bovis* in domestic herds, but these are apparently well-controlled by state veterinarians. Zimbabwe claimed to have no *M. bovis* (in the past), whereas Zambia, Mozambique and Tanzania do. In Zambia, *M. bovis* commonly occurs in wildlife in Kafue (specifically lechwe) (Munyeme, Rigouts *et al*. 2009). In Mozambique, *M. bovis* is found in a number of provinces in livestock, even up to a prevalence of 60% based on skin testing in some districts (A. Machado, personal communication). In Tanzania, there is evidence for *M. bovis* in wildlife and domestic stock (up to 13% positivity) and a few cases of zoonotic *M. bovis* in humans have been reported (Katale, Mbugi *et al*. 2012). Since none of these countries can claim a full diagnostic service with speciation, the true picture with respect to zoonotic TB is unknown.

The social behaviour in captive and free-living animal herds in African wildlife parks provide favourable conditions for *M. bovis* transmission to members of the same herd, for example buffalo in South Africa (Michel, de Klerk *et al*. 2007, Michel, Coetzee *et al*. 2009) and lechwe in Zambia (Munyeme, Rigouts *et al*. 2009). Transmission of the *Mycobacterium tuberculosis* complex members is also effective in species that maintain social or familial groups in underground dens. Drewe *et al.* conducted a study in a South African meerkat

population and found that grooming (both giving and receiving) was more likely than aggression to be correlated with Mycobacterial disease transmission and that groomers were at higher risk of infection than groomees (Drewe, Eames *et al*. 2011).

Although respiratory transmission is probably the most important route of infection in groups of animals that remain in close contact, indirect transmission via food is another important route (Kaneene, Thoen 2004). For oral transmission to occur, an uninfected animal has to consume feed or water contaminated with mucous or nasal secretions, faeces, or urine that contain the infective organisms or receive milk from an infected dam; therefore, the mycobacterium must be able to survive outside an infected host for sufficient time to be ingested by another animal. Research conducted in South Africa and elsewhere suggests that infected buffalo serve as a source of direct oral infection to large predators such as lions and scavenging omnivores such as warthogs (Michel, de Klerk *et al*. 2007, Michel, Coetzee *et al*. 2009).

2.4 Non-tuberculous Mycobacteria

Most of the species included in the genus *Mycobacterium* are distinguished from the *M. tuberculosis* complex members by the fact that they are not obligate pathogens, but are inhabitants of the environment. Although these Mycobacteria are classified as saprophytes and live freely in the environment, they can be opportunistic pathogens, particularly in immunocompromised individuals. This latter factor is of major concern in high HIV prevalence settings, such as those found in many parts of Africa. For such individuals, there may be no such thing as a 'non-pathogenic' Mycobacterium.

Unlike tuberculosis caused by *M. tuberculosis*, which is spread from human-human or human-animal or vice versa, non-tuberculous Mycobacterial infections have not been considered to be particularly contagious (Brown 1985). There is little or no evidence that the

infection can be transmitted from one person to another. Several reports have debated the possibility of co-infection with both *M. tuberculosis* and NTM in individuals (Shamaei, Marjani *et al*. 2010). The role of NTMs, particularly when public health issues are concerned, should nevertheless not be ignored despite the fact that no definitive route of transmission is yet proposed. NTM and the pathologies they cause have received increasing attention world-wide during the past decade (Griffith, Aksamit *et al*. 2007), despite the fact that the reported cases are still relatively few in numbers. Therefore, any disease episode, especially those presenting with clinical manifestations during their involvement merits being reported, along with the major Mycobacterial pathogens, such as *M. tuberculosis, M. leprae, M. avium* and *M. bovis* (Griffith, Aksamit *et al*. 2007).

About one third of the NTMs has thus far been associated with disease in humans and can cause localized disease in the lungs, lymph glands, skin, wounds or bone (Katoch 2004). From a human and clinical perspective, amongst the most important slow-growing species are *M. avium* and *M. intracellulare* [the *M. avium complex* (MAC)].

A recent clinical study in humans (Simons, van Ingen *et al*. 2011) based on data from a number of geographically distant regions in Asia showed that the *M. avium* complex was responsible for 56% of clinically relevant pulmonary disease (range 40-81%), *M. abscessus* for 35%, *M. chelonae* for 31%, *M. kansasii* and *M. scrofulaceum* for 17%, *M. celatum* 9%, *M. szulgai* 6%, *M. fortuitum* 5%, *M. gordonae* 2%, and *M. terrae* 1%. It is important to define what infection means, illustrated by the following example: in Denmark (Andrejak, Thomsen *et al*. 2010) 1,282 adult humans were diagnosed as NTM positive of which 26% had definite disease, 19% possible disease, and 55% colonization only. Five-year mortality after definite NTM disease was 40.1%. Infection with *M. xenopi* was found to be associated with worse prognosis than *M. avium* complex. Among 283 subjects studied in the USA (Yew, Sotgiu *et al*. 2011), 47% of them met ATS/IDSA (American Thoracic Society/ Infectious

Diseases Society of America) pulmonary NTM disease criteria for a minimum overall 2-year period prevalence of 8.6/100,000 persons, and 20.4/100,000 in those at least 50 years of age. The prevalence ranged from 1.4 to 6.6 per 100,000. The prevalence of NTM lung disease globally appears to be increasing (Petrini 2006, Moore, Kruijshaar *et al*. 2010, Phillips, von Reyn 2001, Iseman, Marras 2008, Prevots, Shaw *et al*. 2010) by over 2.8% per year, mostly amongst persons older than 60 years of age, although in most parts of the world there is insufficient data to know how universal this may be. Thus, although the prevalence of disease from NTMs may be low, presence of NTM species is associated with a high likelihood of disease and thereafter mortality.

From such detailed reports in human studies, one may reasonably assume that most if not all of these species could cause TB like disease in animal species and reports suggest that the *M. avium* complex (MAC, which includes *M. intracellulare*), *M. kansasii*, *M. malmoense, M. xenopi, M. ulcerans, M. fortuitum, M. abscessus,* and *M. chelonae* are indeed found in a significant number of animals such as cattle, deer, sheep, and goats, as well as wild and domesticated birds, fish, reptiles, and amphibians (Bercovier, Vincent 2001, Karne, Sangle *et al*. 2012). However, it is important to note that the presence of an NTM does not necessarily imply an active disease process.

There are very few reports of NTMs isolated from animals, especially from Africa. This is partly because of the perception that NTMs rarely cause disease (with the exception of *M. avium* subsp *paratuberculosis*) and partly because identification and speciation of NTMs is relatively complex. However, Mycobacterial species (whether in the MTBC or NTMs) that have been identified to be disease-causing in African animals include *M. goodii, M. origys, M. mungi*, as well as a new species from the meerkat (*Suricata suricatta*), as yet undescribed (Parsons, Drewe *et al*. 2013) (in press).

The following NTMs have been found to cause disease in Africa: *M. avium* is the most frequent bacterial opportunistic infection of AIDS (Karne, Sangle *et al*. 2012). *M. fortuitum* is most frequently isolated from fish and was found to produce a high mortality rate in South African fish farms (Bragg, Huchzermeyer *et al*. 1990); *M. senegalense* has been implicated in tuberculosis-like pathology of cattle in Africa (Bercovier, Vincent 2001). Whether a single agent is responsible for bovine farcy is still unclear, although Chamoiseau considered *M. farcinogenes* var. *senegalense* and *M. farcinogenes* var. *tchadense* to be the sole agents of bovine farcy because of the isolation of these species from the lesions (Chamoiseau 1973). *M. kansasii* has been detected in humans as well as in wild animals such as deer, camels, birds, monkeys and in water (which may be the natural reservoir) and was one of the first species shown to be responsible for non-tuberculous pulmonary infection (Bercovier, Vincent 2001). The impact of this important pathogen in human infection is regularly reported and is considerable; however it is a relatively rare pathogen of animals (Bercovier, Vincent 2001).

Over the past few years, we have prepared various animal samples for Mycobacterial culture analysis when necropsy has been done as part of routine practice. Cultures obtained from bronchial washes, lymph nodes and lung material include the following species identified by 16S rRNA sequence analysis: *M. abcessus*, *M. asiaticum*, *M. avium*, *M. brasilienses*, *M. chelonae*, *M. elephantis, M. engbackii, M. farcinogenes, M. fortuitum, M. gilven, M. gordonae, M. heraklionense, M. hiberniae, M. intracellulare, M. interjectum, M. lentiflavum, M. marseillense, M. moriokaense, M. nonchromogenicum, M. palustre, M. palveris, M. paraffinicum, M. phlei, M. senegalense, M. simiae, M. sherrisii, M. sphagni, M. terrae* and *M. vulneris*.. Animal species examined were lions, rhinoceroses, banded mongooses, cattle, baboons, elephants and monkeys. The presence and isolation of these Mycobacteria does not imply active disease in each case. However, should any animal

become immunocompromised, many species could be pathogenic. Thus, with rare exceptions, we have observed that Mycobacterial species other than *M. tuberculosis*, *M. bovis* and *M. avium subsp paratuberculosis*, seldom cause extensive disease, and very few reports exist in the literature.

2.5 Geographical distribution and environmental determinants

There is a reasonable body of literature regarding the distribution of the MTBC, but little concerning NTMs. These will both be discussed briefly below. Regional variations in the geographical distribution of Mycobacteria in Asia are reported and it is reasonable to expect a similar variation across Africa (Simons, van Ingen *et al*. 2011). Certainly, for example, this is supported by the localised occurrence of Mycobacterial species such as *M. cannetti, M. ulcerans* and subtypes of *M. bovis.* Specifically, *M. ulcerans* and *M. bovis* Af1 subtype are found predominantly in West Africa (Niemann, Richter *et al*. 2000), *M. cannetti* in the Horn of Africa (Pfyffer, Auckenthaler *et al*. 1998), *M. bovis* Af2 subtype in East Africa (Niemann, Richter *et al*. 2000), *M. mungi* in Botswana (Alexander, Laver *et al*. 2010) and the Dassie bacillus and *M. suricattae* thus far in South Africa only (Parsons, Drewe *et al*. 2013, Parsons, Smith *et al*. 2008). Factors playing a role in distribution may include initial host range and topography, host specificity, and host susceptibility. Environmental factors leading to previously unrecorded species appearing sporadically in other regions, can be as a result of travel of animals, humans or birds, for example (De Groote, Huitt 2006)(Griffith, Aksamit *et al*. 2007). Alternatively, the apparent absence of a species from an area may simply be because no survey or relevant diagnostic work has been done there.

With the exception of *M. ulcerans*, these are all members of the *M. tuberculosis* complex and this biodiversity suggests a possible origin for the *M. tuberculosis* complex to be "Out of Africa", with regional evolution to explain the emergence of different strains or

species of Mycobacteria. Further evidence to support this "Out of Africa" suggestion comes from work done in Ethiopia, where a new *M. tuberculosis* lineage has been found which is restricted to the Ethiopian highlands (Berg, Garcia-Pelayo *et al*. 2011).

The distribution of NTMs may be regional or global: for example, the *M. avium* complex, *M. abscessus, M. scrofulaceum, M. marinum* and *M. fortuitum* are encountered globally, whilst *M. malmoense* is found mainly in Scandinavia (Griffith, Aksamit *et al*. 2007) and *M. ulcerans* mainly in Australia, Africa and South-East Asia (the tropics) (Ablordey, Swings *et al*. 2005). In Chad and Nigeria, the *M. fortuitum* complex was most frequently isolated from cattle and pigs (respectively) when they conducted a study to investigate tuberculosis (Bercovier, Vincent 2001). Similarly, pathogenic NTMs isolated from the pastoral ecosystems of Uganda included the *M. avium* complex (Muller, Steiner *et al*. 2008).

Because of the presence of NTMs in the environment, human activities have had direct impacts on their ecology and hence their epidemiology (Kankya, Muwonge *et al*. 2011), since humans have very substantially altered the living environment (Falkinham 2010). This may provide new niches that some Mycobacteria can exploit and thereby increase our risk of exposure. An example of this is water supply systems, where Mycobacteria can easily form biofilms which are almost impossible to remove (Falkinham 2011). In all the habitats where NTMs have been recovered (Michel, de Klerk *et al*. 2007), the Mycobacteria are part of the normal flora, existing as stable, resident, and growing populations. An exception may be *M. avium* subsp. *paratuberculosis,* where growth and persistence in the environment has not been reported, despite being prevalent in bovidae worldwide. Farms or settlements with persistent infection have been described as sources of infection in this case (Whittington, Marsh *et al*. 2005)

Compared to other bacteria, NTMs are relatively disinfectant-, heavy metal-, and antibiotic-resistant. Thus, the use of any antimicrobial agent selects for Mycobacteria. Employment of disinfectants for drinking water treatment leads to selection and enrichment of Mycobacteria in distribution systems in the absence of disinfectant-sensitive competing microorganisms (Falkinham 2010, Falkinham 2011). NTM selection may also occur as a consequence of the presence of antibiotics in drinking water and drinking water sources. Likewise, pollution, large scale agriculture, human and animal movement and a myriad of activities can provide microorganisms in general and Mycobacteria specifically with a rich opportunity to move and exploit their environment (Falkinham 2010, Falkinham 2011).

Discovery that human behaviours lead to selection and proliferation of NTMs in habitats occupied by both humans and NTMs, creates the dilemma that human actions taken to reduce pathogen exposure (i.e., water disinfection), lead to possible increased NTM disease (Petrini 2006, Moore, Kruijshaar *et al*. 2010, Yew, Sotgiu *et al*. 2011, Prevots, Shaw *et al*. 2010).

2.6 Hosts of Mycobacteria

Some individuals (and species or strains of animals) are more susceptible to Mycobacterial disease than others. For example, in cattle it has been shown that various breeds have differing susceptibility to bovine TB (Ameni, Aseffa *et al*. 2007). Additionally, there is evidence to suggest that different populations of hosts may also select for certain pathogen types or subtypes (Hanekom, van der Spuy *et al*. 2007).

Although it is largely *M. tuberculosis* that is found in active TB humans and *M. bovis* in terrestrial animals, Mycobacterium species can cross the species barrier in both the zoonotic and anthropozoonotic directions. Thus, cases of *M. bovis* in humans are well known, but less commonly, cases of *M. tuberculosis* have been reported in animals such as cats, dogs,

and elephants (Parsons, Gous *et al*. 2008, LoBue, Enarson *et al*. 2010, Angkawanish, Wajjwalku *et al*. 2010), where the sources of these infections have commonly been traced to infected humans. However, there are many cases of *M. tuberculosis* in animal collections, where it is arguably an important cause of morbidity and mortality from infectious disease in captive wildlife (Montali, Mikota *et al*. 2001). Transmission of *M. tuberculosis* from animal to animal has not been conclusively shown, although extensive transmission is found with the animal-adapted species of the complex, such as *M. bovis*. Parsons *et al.* showed that *M. tuberculosis* antigens were found in 50% of dogs living in close contact with sputum smear-positive TB patients (Parsons, Warren *et al*. 2012). Since South Africa is rated as a high TB incidence and risk setting, it is not surprising that companion animals living in such environments will be exposed to this pathogen. Whether these animals can be reservoirs for human disease, in turn, is not known. Despite these cases of *M. tuberculosis* in animals, most free-living animals diagnosed as tuberculous are reported or regarded to be infected with *M. bovis*; therefore, infection with *M. bovis* is of public-health and economic importance.

However, the complex members also cause TB; even if they show limited host or geographical association (see Table 2.1). For example, the recently described *M. origys* (Gey van Pittius, van Helden *et al*. 2012, van Ingen, Rahim *et al*. 2012) is apparently host-adapted to the antelope and found almost exclusively in the Arabian oryx (*Oryx leucoryx*) in the Middle East. However, the first isolation of this species in South Africa was from an African buffalo (Gey van Pittius, Perrett *et al*. 2012), however the history of this animal suggests that contact with Arabian oryxes was highly likely at some stage. This finding shows that infected animals with host-adapted pathogens which may at first appear to be host specific can sometimes infect a new host species.

Table 2.1: ***M. tuberculosis*** **complex species and their host-association**

Species	**Host (s)**	**Found in Southern Africa**
M. tuberculosis	Humans	Yes
M. canetti	Humans	No
M. africanum	Humans	No
M. mungi	Banded mongoose	Yes
Dassie bacillus	Dassies	Yes
M. origys	Arabian oryxes	Yes
M. microti	Voles, mice, shrews	No
M. pinnipedii	Sea lions, fur seals	No
M. caprae	Goats	No
M. bovis	Bovids mainly, array of mammalian hosts	Yes

Note: this table is based on published data and is accurate as of June 2012

2.7 Clinical signs and diagnoses

Diagnoses of animals infected with *M. tuberculosis* or *M. bovis* are done using two primary immunological tests, namely the *in vivo* tuberculin skin test (TST) and the *in vitro* interferon-gamma (IFN-γ) release assay (IGRA). These tests typically rely on the detection of antigen-specific T lymphocyte-mediated responses as surrogate markers of infection by the

causative organism. The IGRAs now include proteins largely specific to *M. tuberculosis* and *M. bovis*, such as early secretory antigenic target 6 kDa (ESAT-6) and culture filtrate protein 10 kDa (CFP-10), for better diagnoses. These tests generally indicate infection and not active disease and are also not able to speciate (differentiate) the pathogen or definitively diagnose an infection or disease caused by an NTM. However, once a positive culture of NTMs has been obtained, the Capilia TB assay was found to be a quick and easy test to use for differentiation between the *M. tuberculosis* complex and NTM culture isolates (Muyoyeta, de Haas *et al*. 2010, Kaufmann 2002, de la Rua-Domenech, Goodchild *et al*. 2006).

Because of the high incidence of *M. bovis* detected in a wide variety of animal species (and very high incidence of *M. tuberculosis* in humans) in Africa, it is not surprising that less emphasis is put on the diagnoses of NTM infection. This is the case in animals, even if the animal has tested positive for PPD^{Avium}. It is particularly true in areas of high *M. tuberculosis* complex prevalence. In the face of a high disease burden from *M. tuberculosis* and *M. bovis,* many under-developed nations with inadequate health budgets, do not devote any attention to NTM infection, where the literature mostly reports an NTM incidence rate of less than 5/100 000. However, recent studies in South and southern Africa show that one is likely to isolate NTMs from people suspected to have TB, at a rate at least equal to that with which *M. tuberculosis* is found. In fact, in some communities with a high prevalence of TB, far more NTMs than *M. tuberculosis* have been isolated from subjects (Muyoyeta, de Haas *et al*. 2010, Chihota, Muller *et al*. 2012). From the disease and infection control perspective, it is important to know whether one is dealing with disease due to *M. tuberculosis*, complex or NTMs, since antibiotic susceptibility and control may differ. For example, in animal health work, the discovery of TB-like lesions in a livestock animal during necropsy and the easily derived but possibly incorrect diagnosis of *M. bovis* infection (Muller, de Klerk-Lorist *et al*. 2011) may have devastating and unnecessary consequences for the herd owner.

However, NTMs are increasingly identified due to better detection and speciation by PCR (Warren, Gey van Pittius *et al*. 2006a) or target-gene sequencing. Clinical signs of NTM disease manifestation vary depending on the extent and location of the lesions. NTM disease most commonly presents as pulmonary manifestations (Cook 2010). However, lymph node, skin, soft tissue involvement as well as disseminated disease are of clinical importance. Clinical presentation of pulmonary disease due to NTM may be similar to tuberculosis with the following clinical signs; low fever, night sweats, anaemia and weight loss, malaise, anorexia, diarrhoea and painful adenopathy in humans and in animals (Katoch 2004). Clinical diagnosis of tuberculosis is usually possible only after the disease has reached an advanced stage and is dependent on the site of lesions. At the time of diagnosis, most infected hosts may have shed bacilli and have been or are a potential source of infection for other hosts.

Since it is relatively difficult to isolate, culture and speciate Mycobacteria, as well as decide whether the presence of that organism is simply colonization, or infection and disease, the literature on many of the complex species and for NTMs is not definitive. Isolation may simply represent recent acquisition, colonization without disease, or infection with disease implying clinical manifestation of tissue damage (at either micro or visible scale). Additionally, since NTMs are ubiquitous, great care must be exercised with sampling, since environmental contamination could occur (Hatherill, Hawkridge *et al*. 2006). It is clear that care must be taken in interpretation of diagnostic reports: if the only lab-based test used is a ZN (Ziehl-Neelsen) smear, then acid fastness may be over-interpreted as TB positive (*M. tuberculosis* or perhaps *M. bovis*) and dealt with appropriately, but incorrectly under the DOTS (directly observed therapy, short-course) programme. DOTS treatment of human pulmonary tuberculosis is started only on the basis of sputum microscopy results, which has an inherent possibility of misdiagnosing NTM disease, with subsequent incorrect treatment. For example, it has been noted that some of the cases that are identified as anti-TB treatment

Stellenbosch University http://scholar.sun.ac.za

failure or suspected as drug resistance are actually due to NTMs (Griffith, Aksamit *et al*. 2007).

2.8 Conclusion

Although the burden of disease from *M. tuberculosis* or *M. bovis* is apparently orders of magnitude higher than that from NTMs, NTMs can potentially carry a relatively high burden of morbidity and mortality since they may be difficult to diagnose and treat (van Ingen, Boeree *et al*. 2012). For the purposes of Mycobacterial disease control, care must be exercised where diagnosis is based only on macroscopic examination and perhaps smear or simple culture positivity, or immunological reactivity. We do not yet clearly understand the disease-causing potential of the various Mycobacteria in different animal species or the risk factors and drivers that may promote such disease, as opposed to exposure and infection. Our lack of such knowledge impacts on our ability to control these diseases and generate useful efficacious vaccines. The public health threat of tuberculosis in Africa requires urgent investigation on this topic through collaborative veterinary and medical research programs.

Stellenbosch University http://scholar.sun.ac.za

Chapter 3

Materials and Methods

Stellenbosch University http://scholar.sun.ac.za

Reagents, solutions and media compositions are summarized in Appendix A.

3.1 Bacterial strains

The following species of the *Mycobacterium tuberculosis* complex were selected for proteomic analysis, namely *M. tuberculosis* H37Rv (ATCC 27294), *Mycobacterium bovis* (ATCC SB0267) and *Mycobacterium caprae* (ATCC BAA-824).

3.2 Media and culture conditions

Mycobacterial species were inoculated into Mycobacterial Growth Indicator Tubes (MGIT) (BD Biosciences, USA) and incubated at 37ºC until positive growth was detected using the Bactec 460 TB system (BD Biosciences, USA). Approximately 0.1 ml of the positive MGIT culture was plated out onto supplemented Middlebrook agar (BD Biosciences, USA) containing 0.2% (v/v) glycerol (Merck, USA), 0.05% (v/v) Tween® 80 (Sigma-Aldrich, USA), and 10% (v/v) albumin-dextrose-catalase (0.5% BSA, 0.2% glucose, 0.015% catalase) ($7H10_{ADC}$) and incubated at 37 ºC until single colony forming units (CFU's) were visible.

Stock cultures of each strain were made by inoculating single colonies into 7H9 Middlebrook medium (BD Biosciences, USA) supplemented with 0.2% (v/v) glycerol, 0.1% (v/v) Tween® 80, and 10% (v/v) albumin-dextrose-catalase ($7H9_{ADC}$) and incubated at 37ºC. Once the culture reached an A_{600} of 1.0, freezer stock cultures were made by mixing 300 µl of filter sterilized 50% glycerol with a 700 µl aliquot of Mycobacterial culture in 2 ml cryogenic tubes with O-rings and stored at -80ºC until further use.

For proteomic analysis, freezer stock cultures were inoculated into $7H9_{ADC}$ and incubated at 37ºC.

Stellenbosch University http://scholar.sun.ac.za

Mycobacteria were cultured under strict biosafety level 3 (BSL3) conditions and all samples were decontaminated according to established protocols before removal from the BSL3 facility.

3.3 Verification of Mycobacterial strains

Mycobacterial cultures grown in $7H9_{ADC}$ to A_{600} of 0.4 - 0.6 were tested for contamination using Ziehl-Neelsen staining and growth on blood agar plates. Boiled samples were prepared by boiling one ml of culture for one hour at 95ºC and used for spoligotyping, analysis of Regions of Difference PCR and IS*6110* RFLP to confirm the identity of the Mycobacterial species.

3.3.1 Ziehl-Neelsen staining

Cultures of *M. tuberculosis*, *M. bovis* and *M. caprae* were screened for contamination using Ziehl-Neelsen staining (ZN staining) as described by Kent and Kubica (1985). Cultures were heat-fixed to microscope slides and flooded with ZN Carbol Fuchsin. The microscope slides were heated with a flame until steaming, followed by an incubation period of 5 minutes at room temperature. Slides were rinsed with water and decolourised with 5% acid-alcohol solution for 2 minutes, followed by a wash step with water. Slides were then counterstained using Methylene Blue for 1-2 minutes, rinsed with water and allowed to air dry. Slides were viewed with a light microscope under a 100X oil immersion lens. Uncontaminated cultures of *M. tuberculosis*, *M. bovis* and *M. caprae* appeared as pink acid fast bacilli, whereas contaminated cultures would contain blue-stained organisms.

3.3.2 Blood agar plates

Cultures of *M. tuberculosis*, *M. bovis* and *M. caprae* were plated out on blood agar plates and incubated at 37ºC for 24 hours. No growth on the plates indicated no contamination of cultures.

3.3.3 Spoligotyping

Spoligotyping of each selected Mycobacterial species was done by Dr Elizabeth Streicher, according to the spoligotyping method previously described by Kamerbeek *et al.* 1997 (Kamerbeek, Schouls *et al*. 1997), to genotypically identify each species. Each member of the *M. tuberculosis* complex has a specific spoligotyping signature (figure 3.1) (Streicher, Victor *et al*. 2007).

PGG[a]	Species (strain family[b])	ST no.[c]	Core spoligotype pattern	International family name [reference(s)][d]
1	*M. africanum*	181		AFRI1 (4, 8)*
	M. africanum	331		AFRI2 (4, 8)*
	M. africanum	438		AFRI3 (4, 8)*
	M. bovis	482		BOV1/*M. bovis*-BCG (4, 8)*
	M. caprae	647		BOV4 (4)
	M. pinnipedii	593		PINI1 (4)
	M. pinnipedii	637		PINI2 (4)
	M. canettii	592		canettii (4)
	M. microti	539		microti (4)
	M. tuberculosis (F27; *n* = 6)	1		Beijing (12)*
	M. tuberculosis (F29; *n* = 318)	1		Beijing (12)*
	M. tuberculosis (F31; *n* = 31)	1		Beijing (12)*
	M. tuberculosis (F25; *n* = 12)	21		CAS1/Kili (2, 4, 8, 24)
	M. tuberculosis (F20; *n* = 9)	26		CAS1/Delhi (2, 4, 8, 24)
	M. tuberculosis	288		CAS2 (2, 4, 24)*
	M. tuberculosis	19		Manila/EAI2 (4, 6)*
	M. tuberculosis	236		Ancestral Mtb/EAI (26)

Figure 3.2: Spoligotyping signatures of the *Mycobacterium tuberculosis* complex members. Figure adapted from Streicher *et al.* 2007 (Streicher, Victor *et al*. 2007).

Stellenbosch University http://scholar.sun.ac.za

3.3.4 Regions of difference PCR

Differentiation of the selected members of the MTBC was done by a two-step, multiplex PCR method based on the regions of difference (RD) in the Mycobacterial genomes (Warren, Gey van Pittius *et al*. 2006b). These regions include RD1, RD1^{mic}, RD2^{seal}, RD4, RD9 and RD12. Primers were designed according to Warren *et al.* 2006 (Table 3.1) Each PCR reaction contained: 1 µl of DNA template, 2.5 µl of 10x buffer, 2 µl 25 mM $MgCl_2$, 4 µl 10 mM dNTPs, 5 µl Q-buffer, 0.5 µl of each primer (50 pmol/µl), 0.125 µl HotStarTaq DNA polymerase (Qiagen, Germany) in a 25 µl reaction with (Warren, Gey van Pittius *et al*. 2006b). A negative control (no template) and a positive control (DNA template from *M. tuberculosis* H37Rv) were included to assay for contamination of the reagents and successful PCR amplification, respectively. Amplification was done by activating the Taq polymerase at 95ºC for 15 min, followed by 45 cycles of 94ºC for 1 min, 62ºC for 1 min and 72ºC for 1 min., followed by a final elongation step at 72ºC for 10 min. The amplified products were fractionated on a 3.0% agarose gel in 1x TBE buffer, pH 8.3 at 6 V/cm for 4 hrs. Amplification products were visualized under UV light after staining with ethidium bromide.

Table 3.1 PCR primer sequences for regions of deletion (RD) speciation of *Mycobacterium tuberculosis* complex members.

RD	Primer sequence (5' to 3')
1	AAGCGGTTGCCGCCGACCGACC
1	CTGGCTATATTCCTGGGCCCGG
1	GAGGCGATCTGGCGCGTTTGGGG
4	ATGTGCGAGCTGAGCGATC
4	TGTACTATGCTGACCCATGCG
4	AAAGGAGCACCATCGTCCAC
9	CAAGTTGCCGTTTCGAGCC
9	CAATGTTTGTTGCGCTGC
9	GCTACCCTCGACCAAGTGTT
12	GGGAGCCCAGCATTTACCTC
12	GTGTTGCGGGAATTACTCGG
12	AGCAGGAGCGGTTGGATATTC
1^{mic}	CGGTTCGTCGCTGTTCAAAC
1^{mic}	CGCGTATCGGAGACGTATTTG
1^{mic}	CAATCAGCCAAGACGAGGTTG
2^{seal}	TCAGCGGTCTCATAGCATTGC
2^{seal}	CGGGTTGGGAATGTCAGAAAC
2^{seal}	GCGGCAAGGTACGTCAGAAC

3.3.5 IS*6110* DNA fingerprinting

Selected strains were genotyped using IS*6110* DNA fingerprinting, by Ms Ruzayda van Aarde, according to internationally standardized methods (van Embden, Cave *et al.* 1993).

3.4 Growth curves of Mycobacterial strains

Growth curves were done to determine the log phase for each selected Mycobacterial species. Mid-log phase was used for proteomic analyses to ensure that the species were under the most similar growth conditions and to limit changes in protein expression due to growth phase. Stock cultures of each strain were inoculated into 5 ml of 7H9 Middlebrook medium

(BD Biosciences, USA) supplemented with 0.2% (v/v) glycerol, 0.1% Tween® 80, and 10% dextrose-catalase ($7H9_{DC}$) and incubated at 37ºC. Once the starter culture reached an A_{600} of 1.0, 1 ml of the culture was inoculated into 50 ml of $7H9_{DC}$. Absorbancy readings were taken every few days, the log phase growth curve plotted and mid-log Absorbances for *M. tuberculosis*, *M. bovis* and *M. caprae* were calculated.

3.5 Whole cell lysate protein extraction

One ml freezer stock cultures of *M. tuberculosis*, *M. bovis* and *M. caprae* were inoculated into 10 ml of $7H9_{DC}$ and incubated at 37ºC until an A_{600} of 0.6 - 0.7 was reached. One ml was re-inoculated into 50 ml of $7H9_{DC}$ and incubated at 37ºC until the calculated mid-log phase for each selected strain was reached. Each Mycobacterial species was cultured in duplicate (biological replicates) and whole cell lysate proteins extracted in duplicate from each biological replicate (technical duplicates).

Each culture was divided into two equal aliquots (25 ml) and the Mycobacterial cells were collected by centrifugation (10 min at 2500 x *g*) at 4ºC. The pellets were combined and re-suspended in 1 ml cold lysis buffer containing 10 mM Tris-HCl, pH 7.4 (Merck, Germany), 0.1% (v/v) Tween® 80, and 20 µl/ml protease inhibitor mixture (Protease Inhibitor set III, Merck, Germany). Re-suspended cells were transferred to 2 ml cryogenic tubes with O-rings and centrifuged (2x 1 min at 13 000 x *g*) at room temperature, with 1 min on ice between centrifugation steps. The supernatants were discarded and equal volumes of 0.1 mm glass beads (Biospec Products Inc., OK) were added to the pelleted cells. Additionally, 300 µl of cold lysis buffer and 10 µl of 2 units/ml RNase-free DNase I (New England Biolabs) were added to each tube. The cells were lysed mechanically by bead beating for 20s in a Ribolyser (BIO101 Savant, Vista, CA) at the speed of 4.0 $m.s^{-1}$.

The respective tubes were cooled on ice for 1 min after ribolysing and this procedure was repeated six times. The lysates were centrifuged (2 min at 13 000 x *g*) at room temperature. Thereafter the tubes containing the separated lysate were immediately put on ice. Each supernatant (containing the whole cell lysate) was aspirated by passing through a 0.22 µm millipore syringe sterile filter (Millipore Ireland Ltd., Carrigtwohill, CO. CORK IRL). The filtering step was repeated to ensure sterility. Each whole cell lysate was frozen overnight at -20ºC and stored at -80ºC.

3.6 Protein concentration determination

The Biorad RC-DC assay was used to determine the protein concentration of each whole cell lysate extraction according to the manufacturer's instructions. A standard curve ranging from 0 mg/ml to 2 mg/ml was generated using the Quick StartTM Bovine Serum Albumin (BSA) Standard set (2 mg/ml). Protein absorbance was measured spectrophotometrically at 595 nm using the Ultrospec 4051(LKB Biochrom). The BSA absorbance readings were plotted as a standard curve and the whole cell lysate concentrations were determined from the curve.

3.7 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

All reagents and buffers used for 1-D SDS-PAGE are described in Appendix A. Technical duplicates of each whole cell lysate biological replicate of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* were separated on 4%- 12% gradient, 1.0 mm NuPage gels (Thermo Scientific, Germany) according to the manufacturer's instructions.

3.7.1 SDS-PAGE sample preparation

A total of 60 µg of each whole cell lysate protein extract (technical duplicate) was added to 5 µl of 4x Laemmli sample buffer. The samples were mixed, pulse-centrifuged and

boiled at 95ºC for 5 min to denature the proteins. Care was taken to not let samples cool down after denaturing.

3.7.2 Electrophoresis and staining

The gradient SDS-PAGE gels were placed in the running chamber and covered with 0.2% XT MOPS running buffer (Biorad, USA). The combs were carefully removed and 6 µl of the pre-stained PageRuler^{TM} protein ladder (Fermentas, Canada) was loaded as the molecular weight marker. Prepared whole cell lysate samples were loaded into the wells and separated at 150 Volts at 4 ºC.

The gel was removed from the tank when the dye front reached the end of the gel and stained for 2 hrs with Imperial protein stain (Biorad, USA) on a shaker at room temperature. De-staining of the gel with milliQ (mQ) H_2O took place overnight on a shaker at room temperature.

3.8 Sample preparation for Mass Spectrometry (MS) analysis

All reagents and buffers used for sample preparation of MS analysis are described in Appendix A. Technical duplicates of each biological replicate of each species were prepared for MS analysis.

3.8.1 Peptide extraction

Each gel lane, containing biological/technical replicates of whole cell lysate proteins, was fractionated into 10 fractions by excising the bands with a sterile scalpel on a sterile glass surface. For data analysis purposes, each gel lane was fractionated in the same way (Figure 3.2).

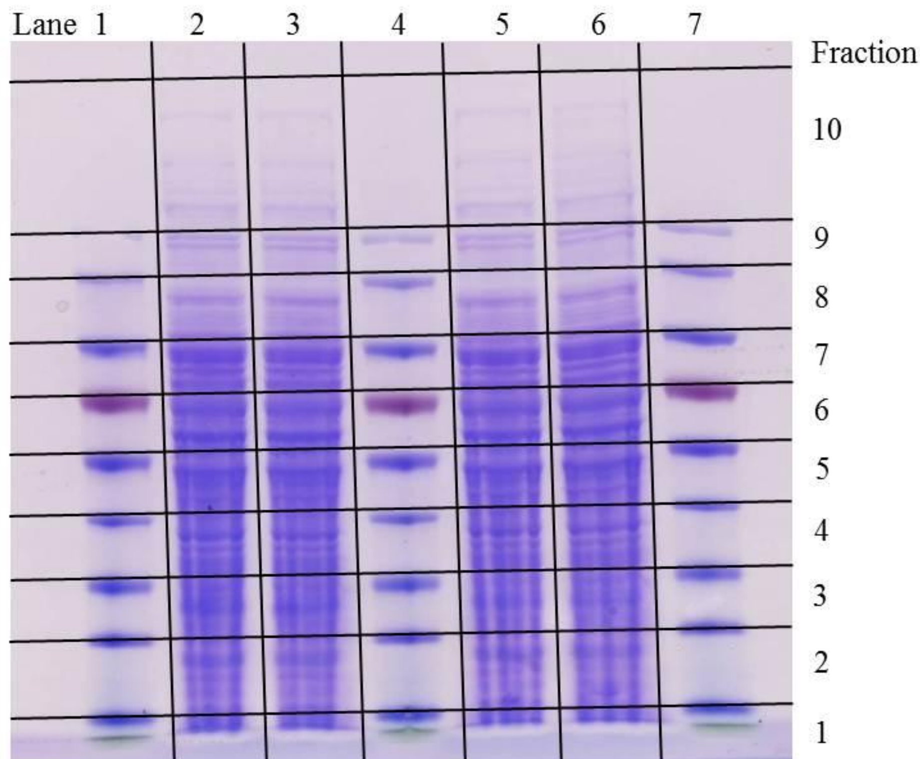


Figure 3.8.1: Fractionation of gel lanes. Lane 2 and 3, and 5 and 6 represent technical duplicates of biological duplicates of whole cell lysates. Lane 1, 4 and 7 represent the pre-stained PageRuler[TM] protein ladder. NuPage 1.0mm, 4%-12% gradient were used with 0.2% XT MOPS running buffer.

Each fraction was cut into smaller pieces of approximately 1 mm^2 and transferred to a sterile micro-centrifuge tube. The gel pieces were washed by adding 300 µl of mQH_2O and incubating for 10 min at room temperature. Thereafter the mQH_2O was aspirated and the washing step was repeated. Subsequently, the gel pieces were de-stained by alternating wash steps with 300 µl 50% acetonitrile (ACN) and 50mM ammonium bicarbonate (ABC) for 10 min each, until the gel pieces were decolourised. Gel pieces were dehydrated by adding 100% ACN for 10 min at room temperature. When the pieces had turned white, the ACN was aspirated and the gel pieces were dried in a speedy-vac for 20 min at room temperature. Gel pieces were rehydrated in 120 µl of 10 mM dithiothreitol (DTT) and incubated at 56 ºC for 60 min to reduce any disulphide bonds. Thereafter the excess DTT was discarded by aspiration and the proteins were alkylated with the addition of 120 µl of 55 mM iodoacetic acid (IAA) followed by incubation in the dark at room temperature for 60 min. The excess

Stellenbosch University http://scholar.sun.ac.za

IAA was then discarded and the gel pieces were washed with 300 µl of 50 mM ABC for 10 min followed by a second wash with 300 µl of 50% ACN for 20 min. After aspirating the 50% ACN, the proteins were in-gel digested with 80 µl of 10 ng/µl sequencing grade trypsin solution (Promega, USA) at 37^{0}C for 17 hours. The reaction was stopped and peptides were eluted from the gel pieces with 70% ACN containing 0.1% formic acid. Hundred percent ACN was added to the gel pieces to elute the remaining peptides.

3.8.2 Stage-tip activation

Prior to mass spectrometry analysis, the peptide mixtures were cleaned and concentrated using self-made stop-and-go-extraction-tips (Stage-tips). Desalting columns were prepared by punching out discs of C18 empore filter (Merck, Germany) with a needle and ejecting the disks into sterile P200 pipette tips. The discs were securely wedged at the bottom of the tip. The supernatant containing the eluted peptides was dried in the speedy vac (Merck, Germany) to approximately 30 µl of sample. Samples were acidified by adding 20 µl of 5% formic acid. Stage-tips were activated by forcing 50 µl of MeOH through the disk. The disk was washed with 50 µl of 5% formic acid. Acidified samples containing the peptides were run through the column twice, where after the disk was washed with 50 µl 5% formic acid. Peptides were eluted from the column with 50 µl of 80% ACN (with 5% formic acid) and dried in the speedy vac (Merck, Germany). Peptides were stored at -20 ºC until MS analysis.

3.9 Mass spectrometry and data analysis

Mass spectrometry work was done by Dr Salome Smit from the Central Analytical Facility, Stellenbosch University. All experiments were performed on a Thermo Scientific EASY-nLC II connected to a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Germany), equipped with a nano-electrospray source.

Dried peptides were dissolved in 5% acetonitrile in 0.1% formic acid and 10 µl injections were made for nano-LC chromatography. For liquid chromatography, separation was performed on an EASY-Column (2 cm, ID 100µm, 5 µm, C18) pre-column followed by XBridge BEH130 NanoEase column (15 cm, ID 75 µm, 3.5 µm, C18) with a flow rate of 300 nl/min. The gradient used was from 5-17% B in 5 min, 17-25% B in 90 min, 25-60% B in 10 min, and 60-80% B in 5 min and kept at 80% B for 10 min. Solvent A was 100% water in 0.1 % formic acid, and solvent B was 100 % acetonitrile in 0.1% formic acid.

The mass spectrometer was operated in data-dependent mode to automatically switch between Orbitrap-MS and LTQ-MS/MS acquisition. Data were acquired using the Xcaliber software package (Thermo Scientific, Germany). The precursor ion scan MS spectra (*m/z* 400 – 2000) were acquired in the Orbitrap with resolution R = 60000 with the number of accumulated ions being 1×10^6. The 20 most intense ions were isolated and fragmented in linear ion trap (number of accumulated ions 1.5×10^4) using collision induced dissociation. The lock mass option (polydimethylcyclosiloxane; *m/z* 445.120025) enabled accurate mass measurement in both the MS and MS/MS modes. In data-dependent LC-MS/MS experiments, dynamic exclusion was used with 60 s exclusion duration. Mass spectrometry conditions were 1.8 kV, capillary temperature of 250°C, with no sheath and auxiliary gas flow. The ion selection threshold was 500 counts for MS/MS and an activation Q-value of 0.25 and activation time of 10 ms were also applied for MS/MS.

MaxQuant 1.2.2.5 was used to identify proteins via an automated database of all tandem mass spectra against the Uniprot solanum lycopersicum database (Cox, Mann 2008). A total of 80 RAW files which accounted for a total of 180 hours of acquisition time were submitted for protein identification using the TbDB *M. tuberculosis* H37Rv (GB:AL123456), TbDB *M. bovis* (AF2122/97) and an in-house custom database (de Souza, Arntzen *et al.* 2011). Carbamidomethyl cysteine was set as fixed modification, and oxidized methionine, N-terminal acetylation (protein), deamidation (NQ), pyro-Glu(Gln) and pyro-Glu(Glu) as variable modifications. The precursor mass tolerance was set to 20 ppm, and fragment mass tolerance set to 0.8 Da. Two missed tryptic cleavages were allowed. The following criteria were applied: peptide and protein false discovery rate of 1% (0.01), minimal peptide length of six, and a protein posterior error probability (PEP) of 0.01 and peptide PEP of 0.15. These extremely stringent parameters guarantee that proteins are identified with high confidence.

Statistical analysis was done using Perseus. Data were transformed (log 2) and imputated (width 0.3, downshift 1.8) to replace missing values. Two samples ANOVA testing was performed using p-values with a threshold of 0.05. Proteins were only considered to be present if they were detected in all replicates that were done. Differential abundance was only considered if ≥ 2 fold abundance was observed ant the t-test value was $p<0.05$. For unique proteins, the proteins were present in all replicates of one species and absent in all replicates of another species with % CV of less than 5%.

Stellenbosch University http://scholar.sun.ac.za

3.10 Bio-informatic analysis

Each protein identified for *M. tuberculosis*, *M. bovis* and *M. caprae* was assigned to a functional category of proteins as defined by Cole *et al.* and summarised on the TubercuList knowledge base (Cole, Brosch *et al*. 1998). Pathway analysis for selected proteins was done using the Pathosystems Resource Integration Center (PATRIC) (Gillespie, Wattam *et al*. 2011) and the Kyoto Encyclopedia of Genes and Genomes database (KEGG) (Kanehisa, Goto 2000).

Stellenbosch University http://scholar.sun.ac.za

Chapter 4

Results

4.1 *M. tuberculosis* complex species verification

Members of the *Mycobacterium tuberculosis* complex share 99.9% sequence identity at the nucleotide level and have near identical 16S rRNA sequences. Extensive verification steps were performed on the closely related bacteria studied here to ensure that proteomic analysis was done on the correct members. *Mycobacterium caprae* was cultured in 7H9 media for the first time in our laboratory.

Ziehl-Neelsen staining confirmed that cultures of *M. tuberculosis* H37Rv, *Mycobacterium bovis* (ATCC SB0267) and *M. caprae* (ATCC SB0416) contained only acid fast bacilli and were not contaminated. The absence of contamination was confirmed by the fact that no growth was detected after 48 hours on blood agar plates for all three cultures.

Cultured *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* species were genotypically differentiated by spoligotyping, mapping of regions of difference (RD) PCR and IS*6110* RFLP fingerprinting. Figure 4.1.1 shows the spoligotyping signature for each member. The phylogenetic lineage representing *M. bovis* and *M. caprae* were characterized by an absence of direct variable repeats 39 to 43. *M. caprae* showed additional deletions of direct variable repeats 1, 3-4, 6-8, 28 and 30-33 when compared to *M. bovis*.

Figure 4.1.1: Unique spoligotyping signatures of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*. *M. bovis* BCG was used as the positive control in this experiment.

Stellenbosch University http://scholar.sun.ac.za

Analysis of the regions of difference (RD) by PCR mapping (not shown) showed, as expected, that RD1, RD4, RD9 and RD12 were present in *M. tuberculosis* H37Rv. RD1 and RD4 were present in *M. caprae* while RD9 and RD12 were absent. RD1 was present in *M. bovis* while RD4, RD9 and RD12 were absent. These results correlate with previously published RD defined genomic comparisons.

IS*6110* DNA fingerprinting of *M. bovis* and *M. caprae* identified 1 and 6 copies of the IS*6110* sequence, respectively (Figure 4.1.2). This differs from *M. tuberculosis* H37Rv which contains 16 copies of this insertion element (Cole, Brosch *et al*. 1998).

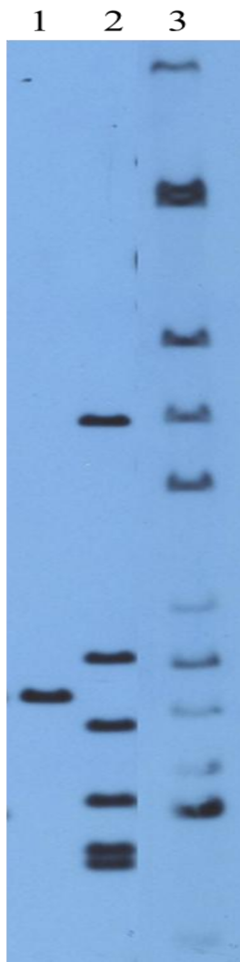


Figure 4.1.2: IS*6110* RFLP results for *M. bovis* and *M. caprae*. Lane 1, 2 and 3 show the RFLP results for *M. bovis*, *M. caprae* and the *M. tuberculosis* reference strain, respectively.

4.2 Proteomic analysis of *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*

4.2.1 Whole cell lysate proteomic analysis

Datasets of proteins identified in this study can be found in Appendix B (disk included), unless otherwise stated. Whole cell lysate protein extracts were harvested from mid-log phase cultures of *M. tuberculosis* H37Rv (A_{600} = 0.7), *M. bovis* (A_{600} = 0.65) and *M. caprae* (A_{600} = 0.7). Proteins were fractionated by SDS-PAGE (Figure 4.2.1.1) and in gel reduction/alkylation and trypsin digests were done. Proteins were identified using LC-MS/MS on the Orbitrap Velos mass spectrometer.

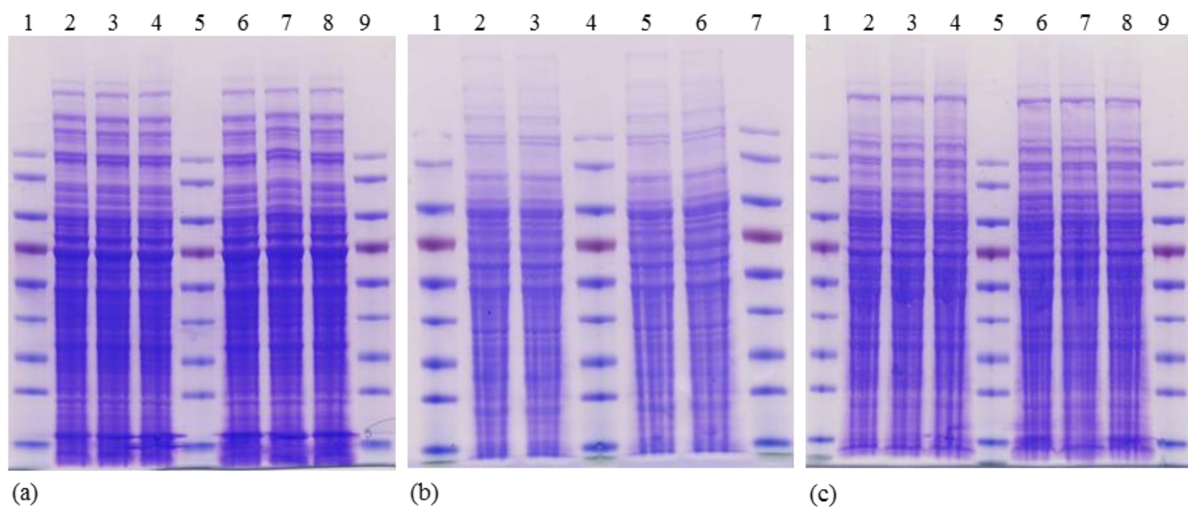


Figure 4.2.1.1: SDS-PAGE results for (a) *M. tuberculosis* H37Rv, (b) *M. bovis* and (c) *M. caprae*. Lanes 1, 5 and 9 in (a) and (c) and lanes 1, 4 and 7 in (b) represent the pre-stained PageRuler^{TM} protein ladder. Whole cell lysate protein samples are represented by lanes 2-4 and 6-8 in (a) and (c) and lanes 2-3 and 5-6 in (b). NuPage 1.0mm, 4%-12% gradient were used with 0.2% XT MOPS running buffer.

Functional groups were assigned using the 10 functional categories described by TubercuList, namely: cell wall and cell wall processes, information pathways, regulatory, intermediary metabolism and respiration, lipid metabolism, virulence, detoxification and adaptation, insertion sequences and phages, PE/PPE families; conserved hypothetical and hypothetical proteins.

***M. tuberculosis* H37Rv**

A total of 13281 *M. tuberculosis* H37Rv peptides were identified by mass spectrometry, which corresponded to 2199 uniquely detected proteins based on these proteins being identified in both the biological and technical replicates. Seventy percent of the proteins detected were classified as intermediary metabolism and respiration, conserved hypotheticals and cell wall and cell wall processes (Figure 4.2.1.2).

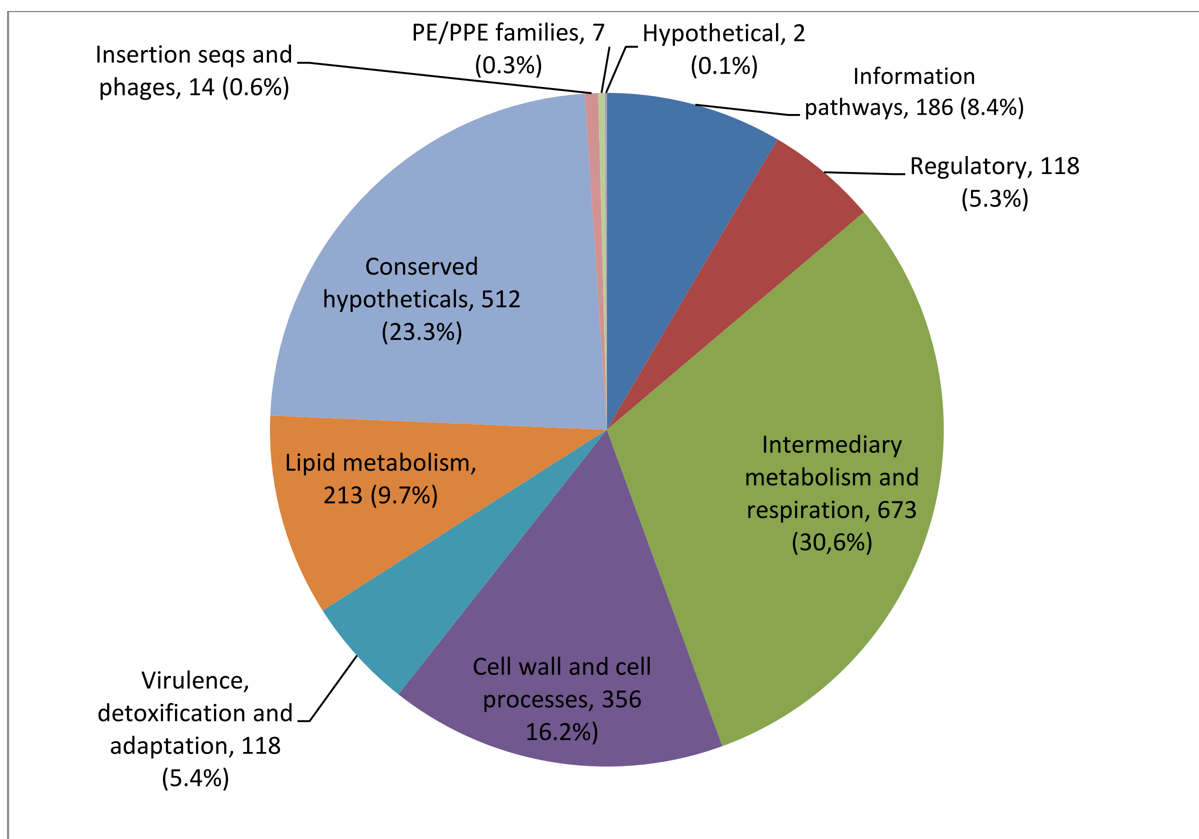


Figure 4.2.1.2: Whole cell lysate proteins of *M. tuberculosis* H37Rv grouped in functional categories.

M. bovis

Mass spectrometry identified 16288 peptides which correlated to 2367 uniquely detected proteins. The detected proteins showed a similar distribution to *M. tuberculosis* H37Rv and the majority were assigned to three categories: intermediary metabolism and respiration, conserved hypotheticals and cell wall and cell wall processes (Figure 4.2.1.3).

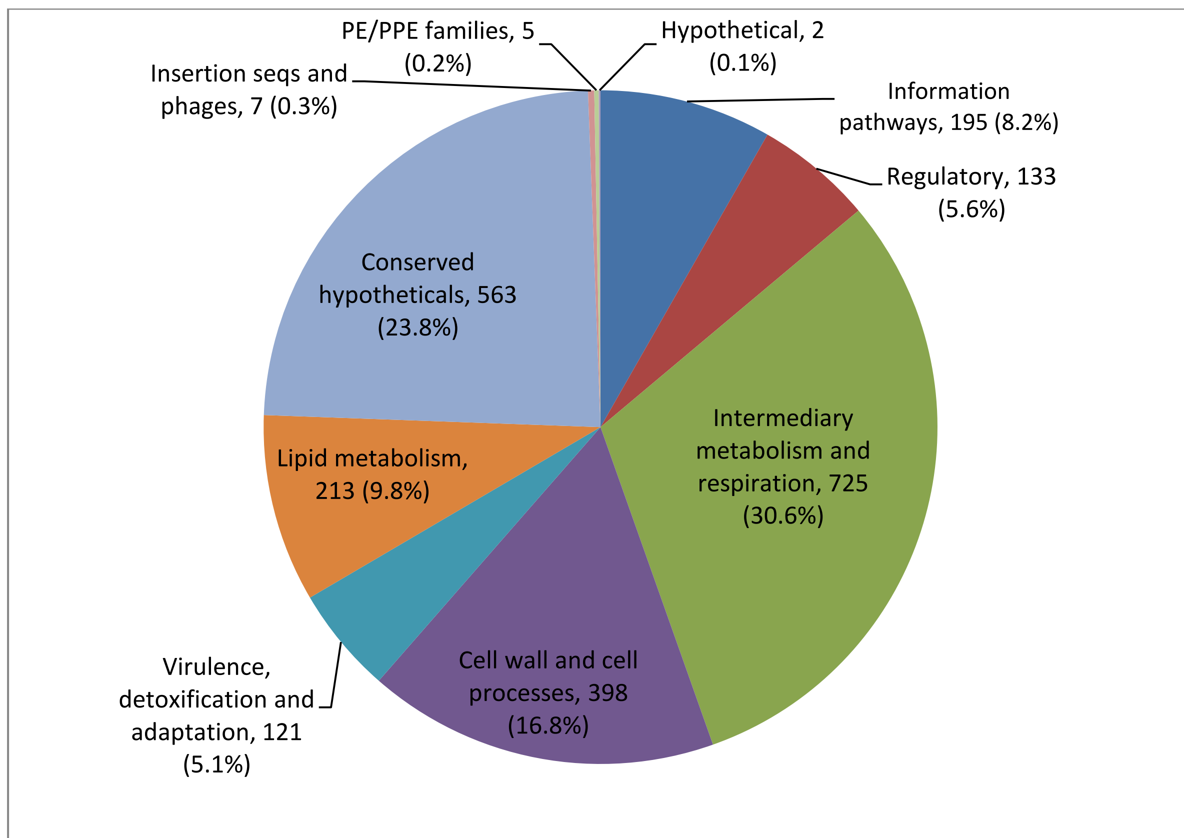


Figure 4.2.1.3: Whole cell lysate proteins of *M. bovis* grouped in functional categories.

M. caprae

Using the annotated genomes of *M. tuberculosis* H37Rv and *M. bovis* (the genome of *M. caprae* has not yet been sequenced or annotated), we identified a total of 14070 peptides using mass spectrometry that correlated with 2350 uniquely detected proteins. The

distribution of proteins in their respective functional categories was similar to that of *M. tuberculosis* H37Rv and *M. bovis* (Figure 4.2.1.4).

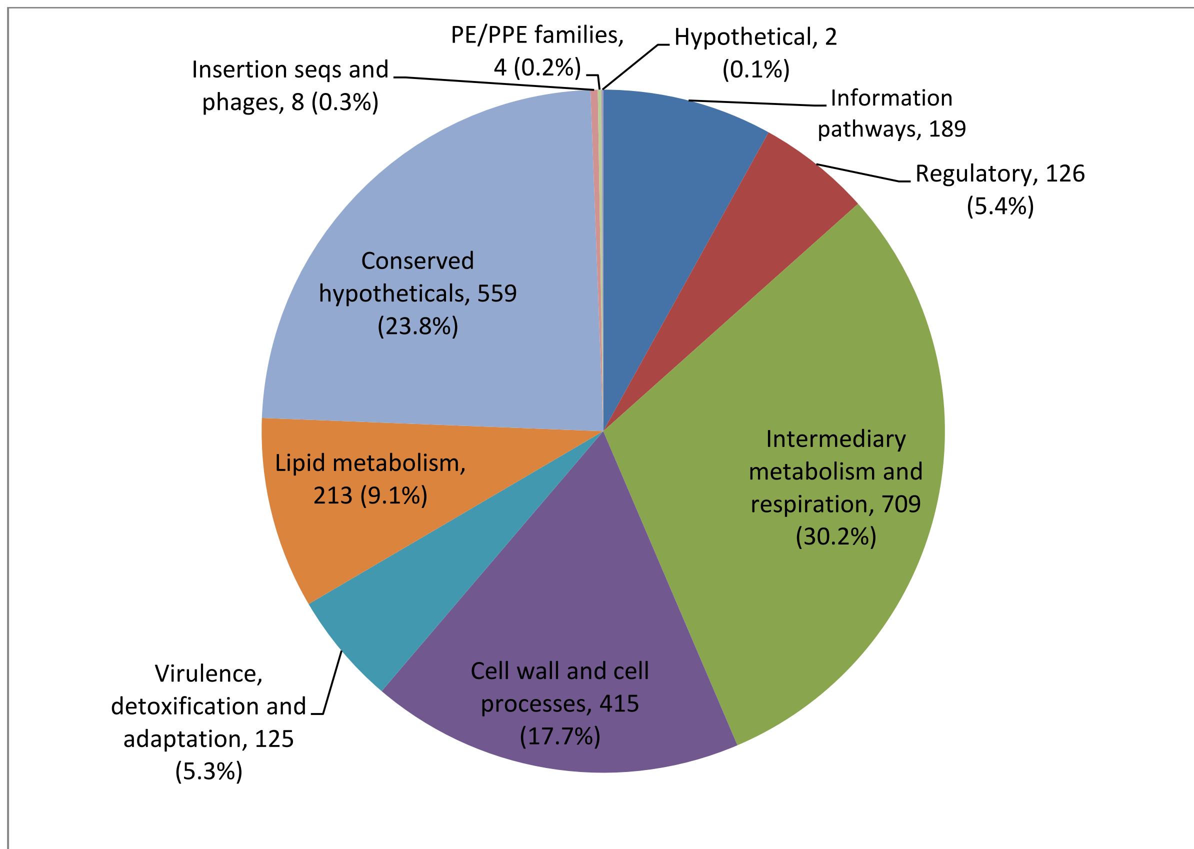


Figure 4.2.1.4: Whole cell lysate proteins of *M. caprae* grouped in functional categories.

Comparison of proteins which differentiate between M. tuberculosis H37Rv, M. bovis and M. caprae

The whole cell lysate proteins detected in *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* were analysed to determine which proteins were uniquely detected in each species. A total of 1804 proteins were detected in all three strains, while 162, 131 and 183 proteins were uniquely detected in *M. bovis*, *M. caprae* and *M. tuberculosis* H37Rv, respectively (Figure 4.2.1.5). The remaining proteins were present in only two of the members tested.

Stellenbosch University http://scholar.sun.ac.za

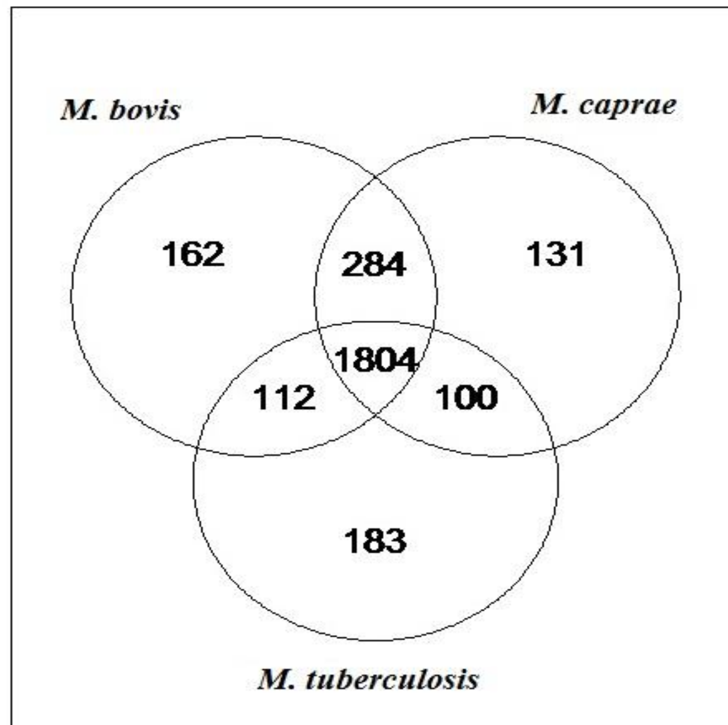


Figure 4.2.1.5: A Venn diagram showing the number of proteins detected in each of the three species, differentiating those that are common or unique amongst the three species.

The uniquely detected proteins for each of the members were divided into functional groups (Figure 4.2.1.6). The distribution of the proteins in each species follows the same distribution pattern as the full constituent of detected proteins, with the largest numbers of unique proteins classified as intermediary metabolism and respiration, cell wall and cell wall processes and conserved hypotheticals. *M. tuberculosis* H37Rv had more than twice the number of uniquely detected proteins identified in the functional categories lipid metabolism and insertion sequences and phages, in comparison with *M. bovis* and *M. caprae*.

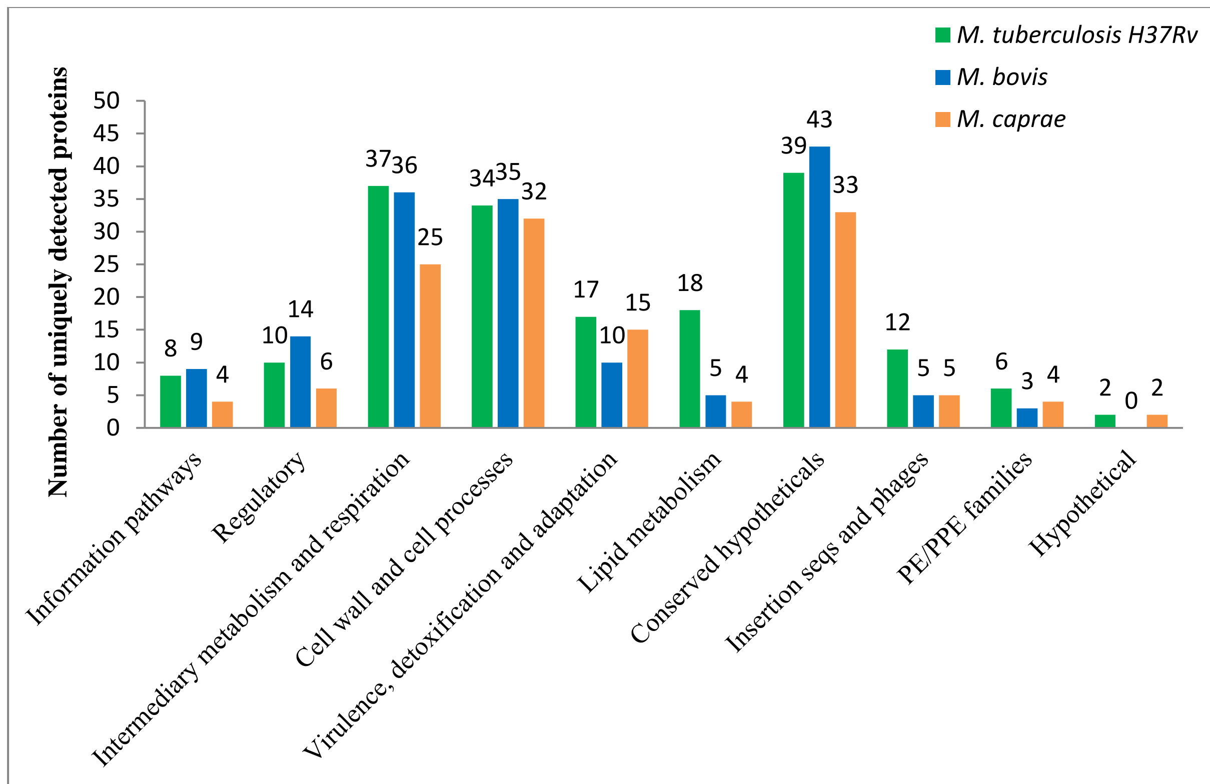


Figure 4.2.1.6: Distribution of uniquely detected proteins in *M. tuberculosis* H37Rv (green), *M. bovis* (blue) and *M. caprae* (orange).

M. bovis and *M. caprae* shared 284 proteins, which differentiate them from *M. tuberculosis* H37Rv. The 284 shared proteins between *M. bovis* and *M. caprae* were divided into functional groups (Figure 4.2.1.7). The distribution of the proteins follows the same distribution pattern as the uniquely detected proteins in each species, with the largest numbers of shared proteins classified as intermediary metabolism and respiration, cell wall and cell wall processes and conserved hypotheticals.

Proteins which were unique to *M. tuberculosis*, *M. bovis* or *M. caprae* and proteins shared between two species only were included in the protein abundance datasets for functional group and pathway analysis to further characterize the proteomes of these species.

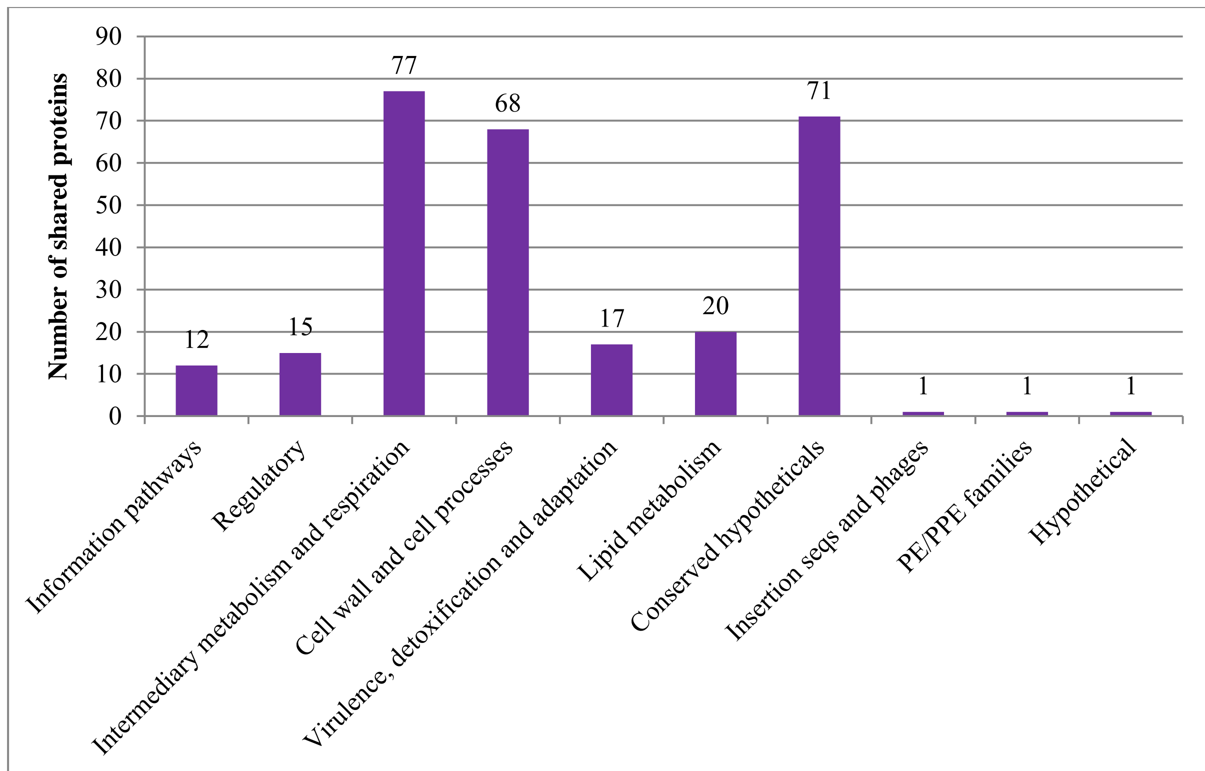


Figure 4.2.1.7: Distribution of shared proteins between *M. bovis* and *M. caprae*.

4.2.2 Analysis of the abundance of whole cell lysate proteins

For analysis of protein abundance, the protein dataset of *M. tuberculosis* H37Rv was used as the reference for protein expression to which the abundance of proteins identified in *M. bovis* and *M. caprae* were compared. A two fold difference in protein abundance with a p-value of <0.05 was considered significant. Only proteins detected in all four replicates of each species were used for further analysis.

M. bovis

We identified that 409 proteins were more highly abundant and 211 proteins were in lower abundance in *M. bovis* relative to *M. tuberculosis* H37Rv. Figure 4.2.2.1 shows the number of proteins from each functional group which were detected in greater or lesser abundance than in *M. tuberculosis* H37Rv. Functional groups intermediary metabolism and respiration; conserved hypotheticals and cell wall and cell processes had the highest number

of over-abundant proteins relative to *M. tuberculosis* H37Rv. The highest numbers of low abundance proteins were found in the following categories: conserved hypotheticals, intermediary metabolism and respiration and lipid metabolism.

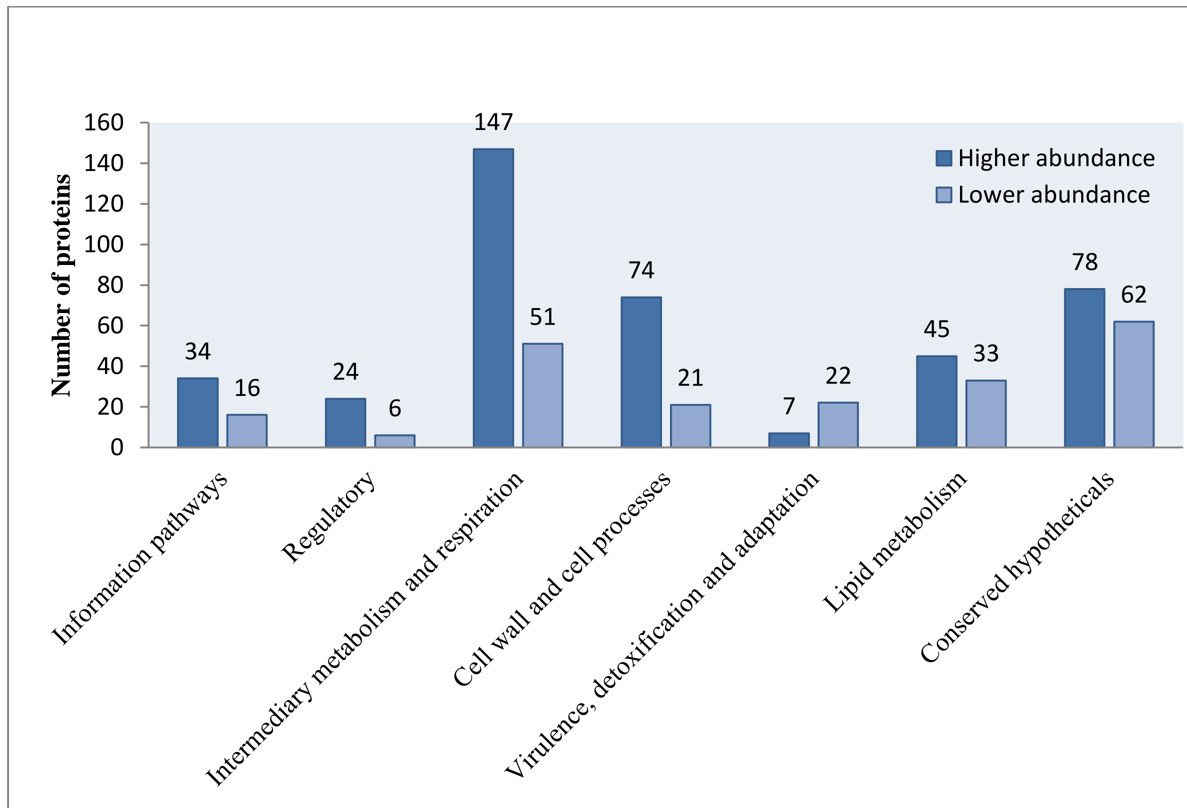


Figure 4.2.2.1: Functional groups of proteins detected in higher and lower abundance in *M. bovis*, relative to *M. tuberculosis* H37Rv. No significant differences were detected for functional groups Insertion sequences and phages, PE/PPE families and Hypothetical proteins in *M. bovis*, relative to *M. tuberculosis* H37Rv.

M. caprae

A total of 393 proteins were over-abundant by more than 2 fold and a total of 175 proteins were under-abundant by more than 2 fold, in *M. caprae* relative to *M. tuberculosis*. Functional categories were assigned to the differentially expressed proteins (Figure 4.2.5) and the categories containing the largest numbers of over- and under-abundant proteins were

intermediary metabolism and respiration, cell wall and cell processes, conserved hypotheticals and lipid metabolism.

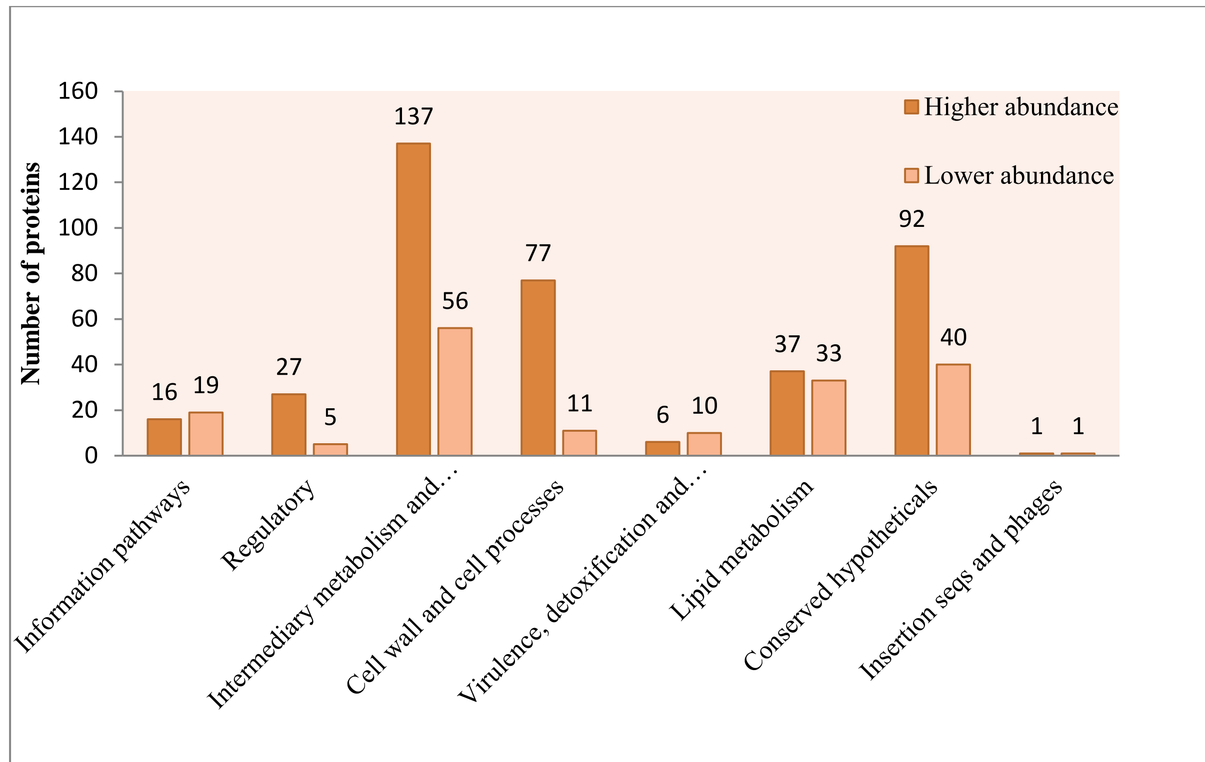


Figure 4.2.2.2: Functional groups of proteins detected in higher and lower abundance in *M. caprae*, relative to *M. tuberculosis* H37Rv. No significant differences were detected for functional groups PE/PPE families and Hypothetical proteins in *M. caprae*.

Comparison of differentially proteins detected in M. bovis and M. caprae

Proteins which were over- and under-abundant in *M. bovis* and *M. caprae* were compared. A total of 222 over-abundant proteins were shared between these two species and 187 and 172 proteins were uniquely over-abundant in *M. bovis* and *M. caprae,* respectively (Figure 4.2.2.3.a). Under-abundant proteins were distributed as 201 shared proteins, and 112 and 75 proteins in *M. bovis* and *M. caprae* respectively (Figure 4.2.2.3.b). The 222 over-

abundant and 201 under-abundant proteins in both *M. bovis* and *M. caprae*, had approximately the same fold difference in abundance, relative to *M. tuberculosis* H37Rv.

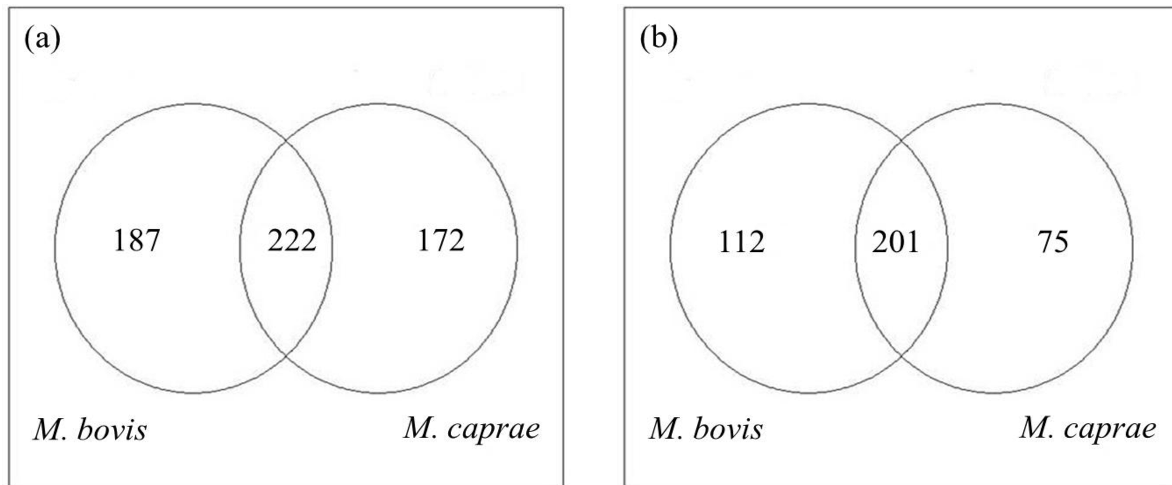
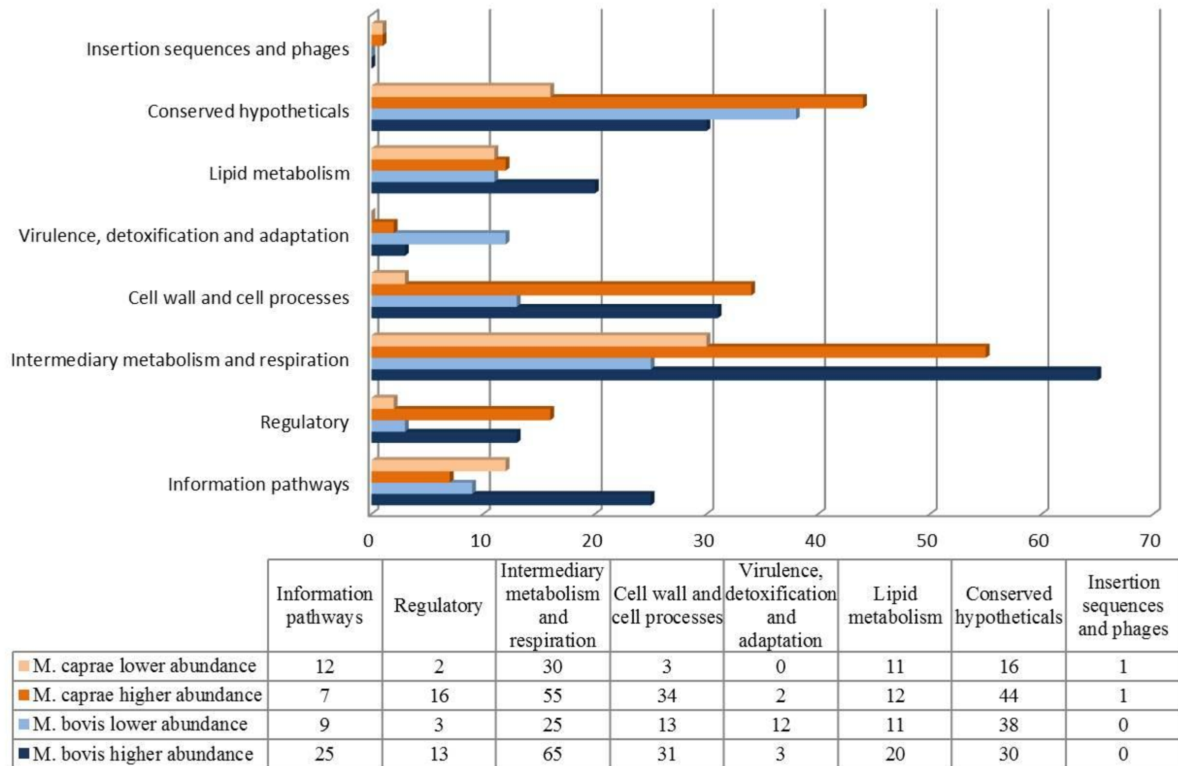


Figure 4.2.2.3: Venn diagrams showing the number of proteins which were detected in (a) higher abundance or (b) lower abundance in *M. bovis* and *M. caprae* relative to *M. tuberculosis*, indicating the number which showed similar expression profiles in both strains relative to *M. tuberculosis.*

Over- or under-abundant proteins detected only in one species were divided into functional groups (Figure 4.2.2.4). Functional groups that differed between these two species were information pathways, cell wall and cell processes and virulence, detoxification and adaptation. *M. bovis* had 25 higher abundant proteins in information pathways compared to 7 in *M. caprae*. Thirteen less abundant proteins were indentified in cell wall and cell processes in *M. bovis* compared to 3 proteins in *M. caprae*. A total of 12 proteins were lower in abundance in *M. bovis* relative to *M. caprae* in virulence, detoxification and adaptation. Proteins detected in these over- and under abundant functional categories are summarized in Table 4.2.2.1.



	Information pathways	Regulatory	Intermediary metabolism and respiration	Cell wall and cell processes	Virulence, detoxification and adaptation	Lipid metabolism	Conserved hypotheticals	Insertion sequences and phages
M. caprae lower abundance	12	2	30	3	0	11	16	1
M. caprae higher abundance	7	16	55	34	2	12	44	1
M. bovis lower abundance	9	3	25	13	12	11	38	0
M. bovis higher abundance	25	13	65	31	3	20	30	0

Figure 4.2.2.4: Functional categories of over-or under-abundant proteins detected only in *M. bovis* or *M. caprae*. Proteins detected in *M. caprae* in lower abundance (light orange) and higher abundance (dark orange). For *M. bovis*, the proteins are shown in light and dark blue for lower and higher abundant proteins, respectively. No proteins were only found in the PE/PPE families, as well as the Insertion sequences and phages category for *M. bovis*.

Table 4.2.2.1: Summary of higher and lower-abundance proteins detected in only one species in functional categories that differed significantly between *M. bovis* and *M. caprae*.

	Protein of interest	Functional group	Fold difference	Gene name	Proposed function
M. bovis higher abundance	Rv0003	Information pathways	3.0	recF	DNA recombination and repair protein RecF
	Rv0640	Information pathways	2.9	rplK	LSU ribosomal protein L11p (L12e)
	Rv0682	Information pathways	2.3	rpsL	SSU ribosomal protein S12p (S23e)
	Rv0700	Information pathways	2.4	rpsJ	SSU ribosomal protein S10p (S20e)
	Rv0701	Information pathways	3.6	rplC	LSU ribosomal protein L3p (L3e)
	Rv0704	Information pathways	2.9	rplB	LSU ribosomal protein L2p (L8e)
	Rv0705	Information pathways	3.4	rpsS	SSU ribosomal protein S19p (S15e)
	Rv0706	Information pathways	3.5	rplV	LSU ribosomal protein L22p (L17e)
	Rv0709	Information pathways	3.9	rpmC	LSU ribosomal protein L29p (L35e)
	Rv0714	Information pathways	2.6	rplN	LSU ribosomal protein L14p (L23e)
	Rv0715	Information pathways	4.2	rplX	LSU ribosomal protein L24p (L26e)
	Rv0720	Information pathways	4.1	rplR	LSU ribosomal protein L18p (L5e)
	Rv0722	Information pathways	6.6	rpmD	LSU ribosomal protein L30p (L7e)
	Rv0979A	Information pathways	2.3	rpmF	LSU ribosomal protein L32p
	Rv1643	Information pathways	2.3	rplT	LSU ribosomal protein L20p
	Rv2228c	Information pathways	2.0	Rv2228c	FIG006762: Phosphoglycerate mutase family
	Rv2357c	Information pathways	2.0	glyS	Glycyl-tRNA synthetase (EC 6.1.1.14)
	Rv2412	Information pathways	2.9	rpsT	SSU ribosomal protein S20p
	Rv2441c	Information pathways	5.8	rpmA	LSU ribosomal protein L27p
	Rv2442c	Information pathways	4.4	rplU	LSU ribosomal protein L21p
	Rv2534c	Information pathways	2.7	efp	Translation elongation factor P
	Rv2985	Information pathways	4.8	mutT1	Possible hydrolase mutT1
	Rv2986c	Information pathways	2.5	hupB	DNA-binding protein HU / low-complexity
	Rv3456c	Information pathways	3.5	rplQ	LSU ribosomal protein L17p
	Rv3580c	Information pathways	2.6	cysS	Cysteinyl-tRNA synthetase
M. caprae higher abundance	Rv0949	Information pathways	2.0	uvrD1	Probable ATP-dependent DNA helicase II
	Rv1108c	Information pathways	2.6	xseA	Exonuclease VII large subunit
	Rv1640c	Information pathways	2.7	lysX	Possible lysyl-tRNA synthetase 2
	Rv2528c	Information pathways	4.7	mrr	Probable restriction system protein
	Rv3296	Information pathways	2.1	lhr	Probable ATP-dependent helicase LHR
	Rv3597c	Information pathways	2.0	lsr2	Iron regultaed H-NS-like protein
	Rv3834c	Information pathways	2.0	serS	Seryl-tRNA synthetase
M. bovis lower abundance	Rv0290	Cell wall and cell processes	2.1	eccD3	Esx conserved component eccD3
	Rv0394c	Cell wall and cell processes	2.1	Rv0394c	Possible secreted protein
	Rv0411c	Cell wall and cell processes	3.2	glnH	Probable glutamine-binding lipoprotein
	Rv0425c	Cell wall and cell processes	3.6	ctpH	Possible metal cation transporting P-type ATPase
	Rv0482	Cell wall and cell processes	2.3	murB	UDP-N-Acetylenolpyruvoylglucosamine reductase
	Rv0638	Cell wall and cell processes	2.4	secE1	Probable preprotein translocase
	Rv0875c	Cell wall and cell processes	3.0	Rv0875c	Possible conserved exported protein
	Rv2115c	Cell wall and cell processes	2.2	mpa	Mycobacterial proteasome ATPase
	Rv3193c	Cell wall and cell processes	4.5	Rv3193c	Probable conserved transmembrane protein
	Rv3809c	Cell wall and cell processes	2.6	glf	UDP-galactopyranose mutase
	Rv3865	Cell wall and cell processes	2.1	espF	Esx-1 secretion associated protein
	Rv3919c	Cell wall and cell processes	4.4	gid	Probable glucose-inhibited division protein B
	Rv3921c	Cell wall and cell processes	6.5	Rv3921c	Probable conserved transmembrane protein
	Rv0251c	Virulence, detoxification and adaptation	5.6	hsp	Heat shock protein
	Rv0634c	Virulence, detoxification and adaptation	4.3	Rv0634c	Possible glyoxalase II
	Rv1114	Virulence, detoxification and adaptation	4.6	vapC32	Possible toxin vapC32
	Rv1608c	Virulence, detoxification and adaptation	3.4	bcpB	Probable peroxidoxin
	Rv1901	Virulence, detoxification and adaptation	27.6	cinA	Probable cinA-like protein
	Rv1932	Virulence, detoxification and adaptation	4.8	tpx	Probable thiol peroxidase
	Rv1962c	Virulence, detoxification and adaptation	2.1	vapC35	Possible toxin vapC35
	Rv2416c	Virulence, detoxification and adaptation	33.0	eis	Enhanced intracellular survival protein
	Rv2428	Virulence, detoxification and adaptation	82.8	ahpC	Alkyl hydroperoxide reductase C protein
	Rv2494	Virulence, detoxification and adaptation	3.0	vapC38	Possoble toxin vapC38
	Rv3100c	Virulence, detoxification and adaptation	6.2	smpB	Probable SSRA-binding protein
	Rv3661	Virulence, detoxification and adaptation	2.6	Rv3661	Conserved hypothetical protein
M. caprae lower abundance	Rv0934	Cell wall and cell processes	7.3	pstS1	Periplasmic phosphate-binding lipoprotein
	Rv2150c	Cell wall and cell processes	8.7	ftsZ	Cell division protein
	Rv2916c	Cell wall and cell processes	2.1	ffh	Probable signal recognition particle protein

Stellenbosch University http://scholar.sun.ac.za

4.3 Pathway analysis of differentially abundant proteins detected in *M. bovis* and *M. caprae*

To provide an integrated overview of the overall network of protein expression (rather than just limit analysis to individual proteins), pathway analysis was done on the differentially expressed proteins of *M. bovis* and *M. caprae*. Proteins were assigned pathway categories in PATRIC (Pathosystems Resource Integration Center) and visualized using iTUBY (Interactive Pathway Explorer database). Pathways highlighted in the figures indicate altered protein abundance relative to *M. tuberculosis* H37Rv.

M. bovis

PATRIC analysis showed that most of the 409 higher abundant proteins in *M. bovis* contribute to/function in valine, leucine and isoleucine degradation, fatty acid metabolism and oxidative phosphorylation (Table 4.3.1). Pathway results correlated with the analysis and visualization of these proteins in iTUBY and also showed that Xenobiotic degradation and ribosomal protein pathways were influenced (Figure 4.3.1). The 187 over-abundant proteins that were only detected in *M. bovis* follow the same functional group and pathway distribution (data not shown). Analysis on PATRIC and iTUBY showed that the 211 lower abundant proteins, which include the 112 lower abundance proteins only detected in *M. bovis*, were mostly involved in Lipid metabolism (Table 4.3.2 and Figure 4.3.2).

Table 4.3.1: Summary of pathways affected by the higher abundance proteins in *M. bovis* relative to *M. tuberculosis.*

Pathway Name	Number of proteins
Valine, leucine and isoleucine degradation	18
Fatty acid metabolism	17
Oxidative phosphorylation	16
Arginine and proline metabolism	15
Butanoate metabolism	15
Methane metabolism	14
Tryptophan metabolism	14
Phenylalanine metabolism	13
Reductive carboxylate cycle (CO2 fixation)	12
Propanoate metabolism	12

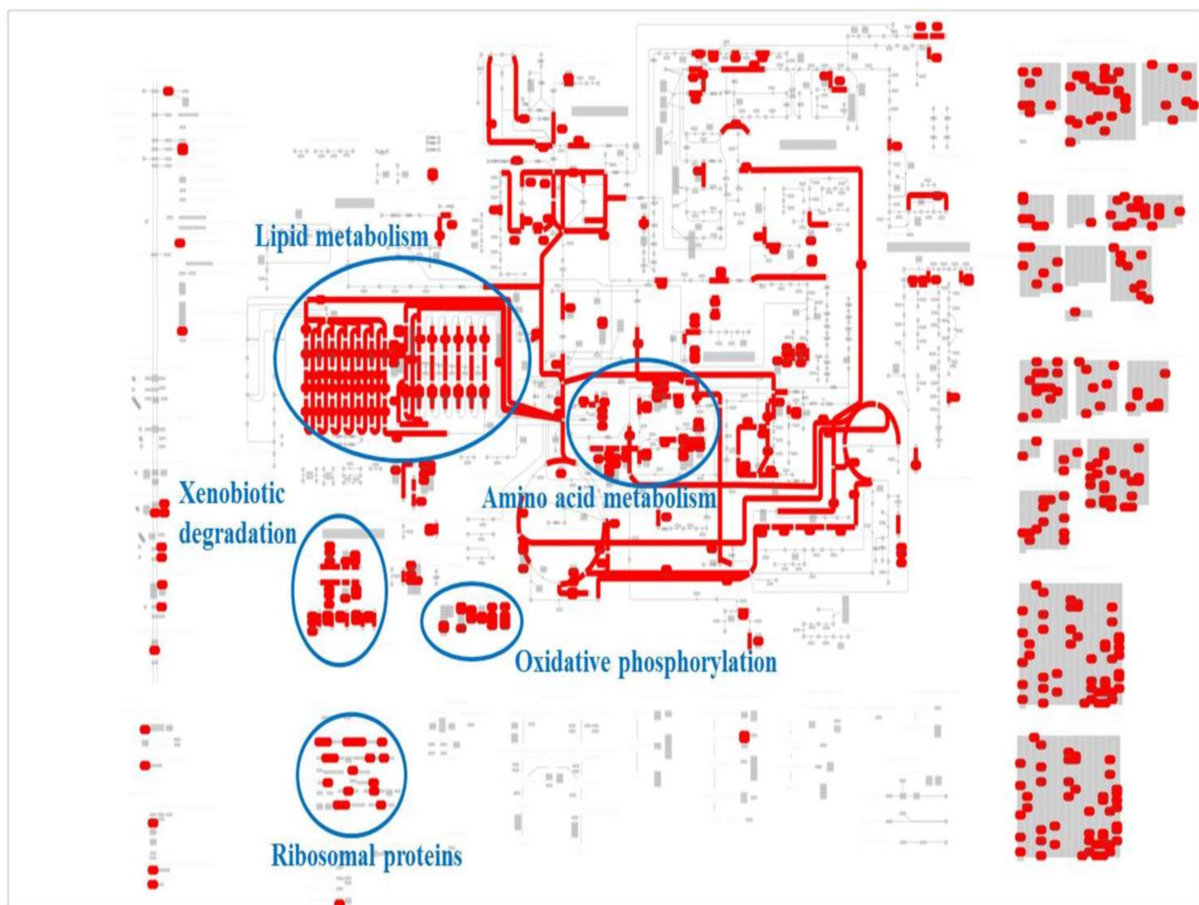


Figure 4.3.1: Pathway analysis of higher abundance proteins of *M. bovis* compared to *M. tuberculosis* H37Rv. Lipid metabolism, amino acid metabolism, xenobiotic degradation, oxidative phosphorylation and ribosomal protein pathways are noticeably affected. Proteins that mapped to the pathways are illustrated in red.

Table 4.3.2: Summary of pathways affected by the lower abundant proteins in *M. bovis,* relative to *M. tuberculosis.*

Pathway Name	Number of proteins
Fatty acid metabolism	13
Biotin metabolism	6
Valine, leucine and isoleucine degradation	6
Fatty acid biosynthesis	5
alpha-Linolenic acid metabolism	5
Pyrimidine metabolism	5
Benzoate degradation via hydroxylation	5
Phenylalanine, tyrosine and tryptophan biosynthesis	4
Glutathione metabolism	4

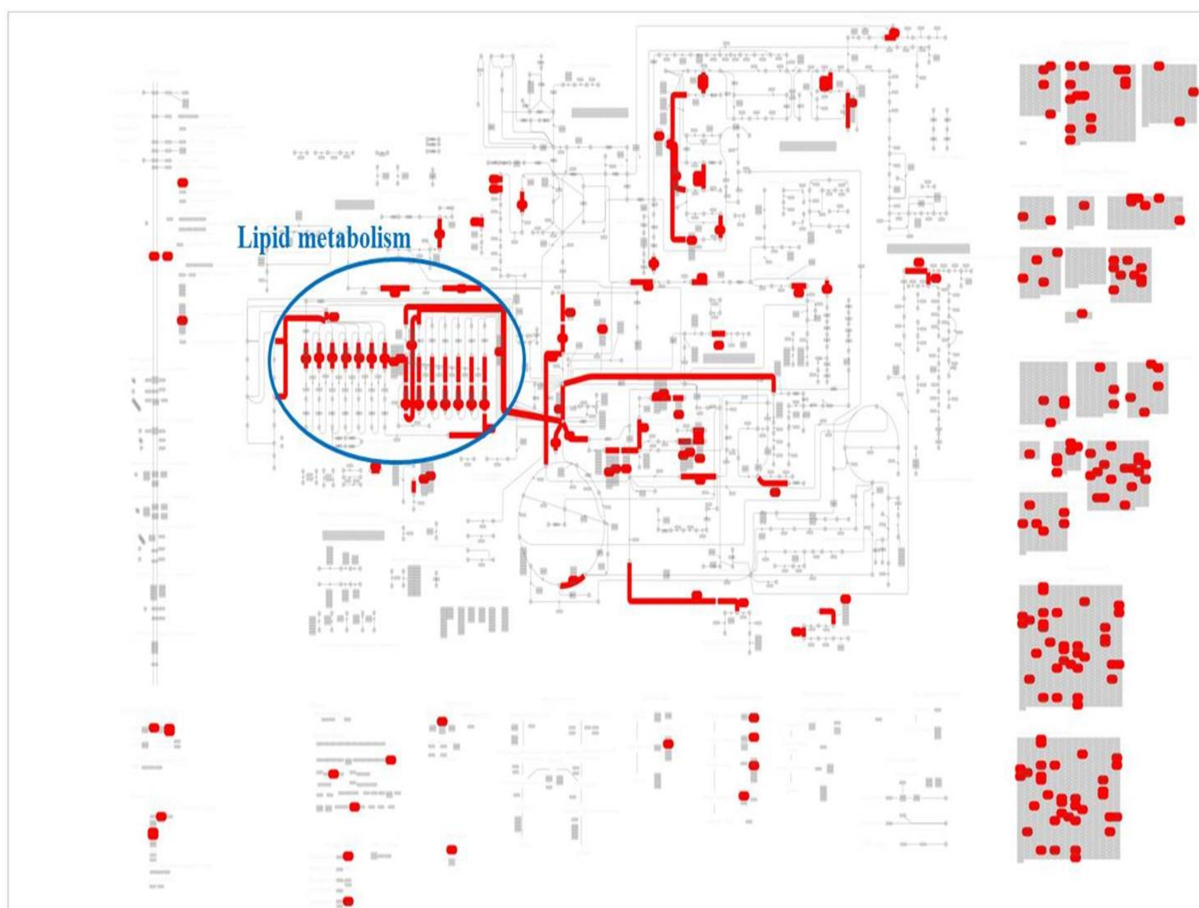


Figure 4.3.2: Pathway analysis of lower abundance proteins of *M. bovis* compared to *M. tuberculosis* H37Rv. The lipid metabolism pathway is highly affected. Proteins that mapped to the pathways are illustrated in red.

M. caprae

Pathway analysis using the PATRIC database concluded that the 394 higher abundance proteins were mostly involved in oxidative phosphorylation, methane metabolism, valine, leucine and isoleucine degradation and glycolysis (Table 4.3.3). Analysis of these proteins using iTUBY showed that carbohydrate metabolism, lipid metabolism, amino acid metabolism, xenobiotic degradation and oxidative phosphorylation were affected (Figure 4.3.3). Most of the 172 over-abundant proteins only detected in *M. caprae* where found in these pathways. The 175 down-regulated proteins, which include the 75 proteins only detected in *M. caprae*, were mostly found in the pyrimidine metabolism, purine metabolism and lipid metabolism pathways (Table 4.3.4 and Figure 4.3.4).

Table 4.3.3: Summary of pathways affected in *M. caprae* using the higher abundant proteins, relative to *M. tuberculosis*.

Pathway Name	Number of Proteins
Oxidative phosphorylation	18
Methane metabolism	15
Valine, leucine and isoleucine degradation	15
Propanoate metabolism	13
Glycolysis / Gluconeogenesis	12
Fatty acid metabolism	12
Pyruvate metabolism	11
Nitrogen metabolism	10
Reductive carboxylate cycle (CO2 fixation)	10
Arginine and proline metabolism	10

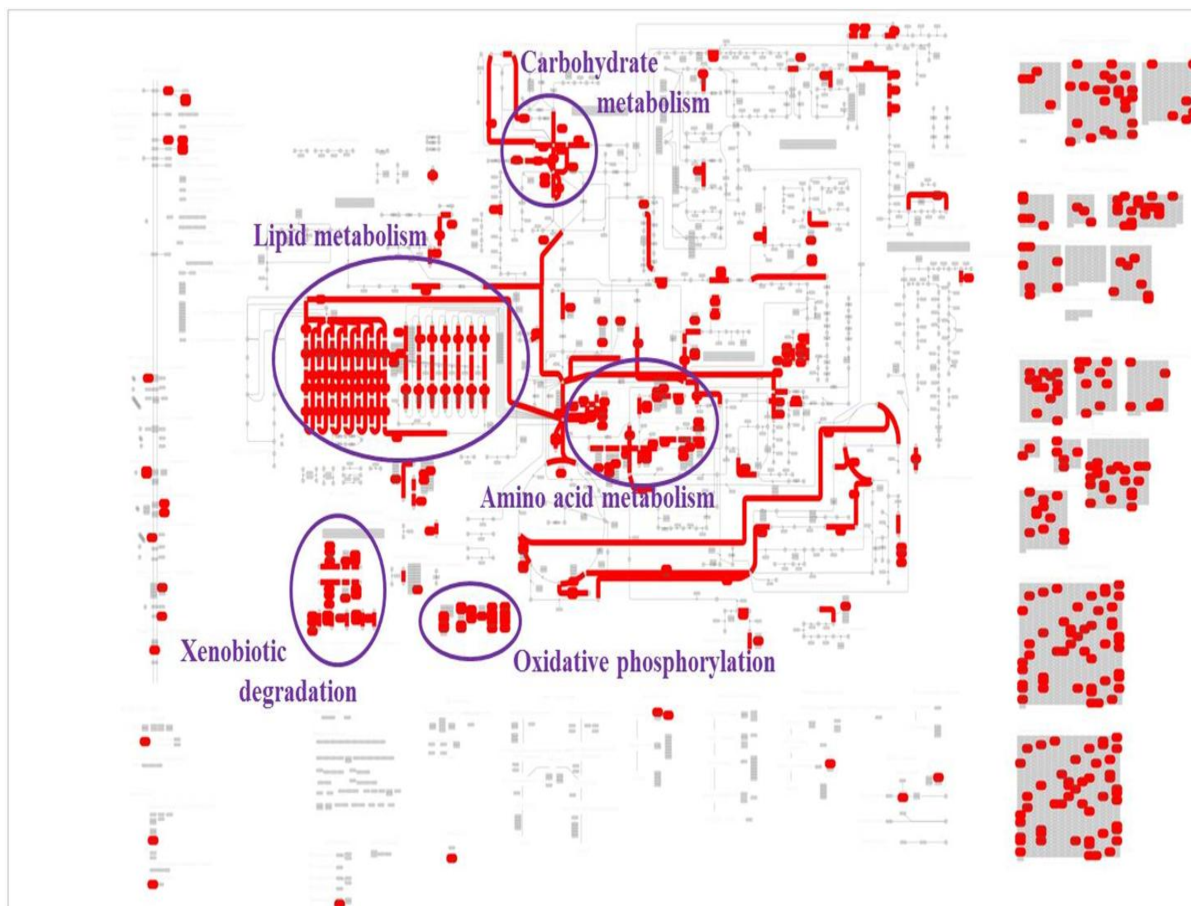


Figure 4.3.3: Pathway analysis of higher abundance proteins in *M. caprae* compared to *M. tuberculosis* H37Rv. Carbohydrate metabolism, lipid metabolism, amino acid metabolism, xenobiotic degradation and oxidative phosphorylation pathways were affected. Proteins that mapped to the pathways are illustrated in red.

Table 4.3.4: Summary of pathways affected by the lower abundant proteins in *M. caprae,* relative to *M. tuberculosis.*

Pathway Name	Number of Proteins
Purine metabolism	9
Fatty acid metabolism	8
Pyrimidine metabolism	6
Lysine biosynthesis	4
Glycerophospholipid metabolism	4
Porphyrin and chlorophyll metabolism	4
Citrate cycle (TCA cycle)	3
Cysteine and methionine metabolism	3
Amino sugar and nucleotide sugar metabolism	3
Pyruvate metabolism	3

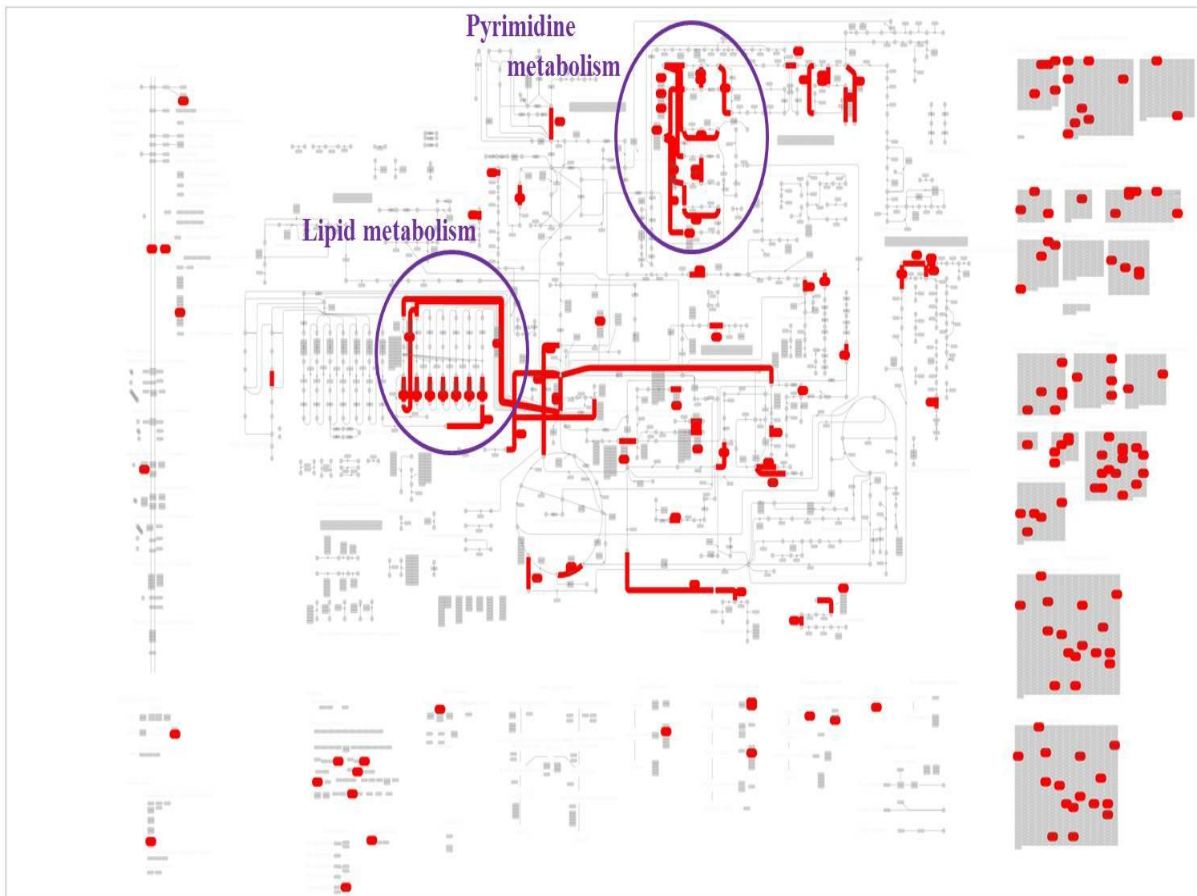


Figure 4.3.4: Pathway analysis of lower abundance proteins of *M. caprae* compared to *M. tuberculosis* H37Rv. Lipid metabolism and pyrimidine metabolism pathways were highly affected. Proteins that mapped to the pathways are illustrated in red.

4.4 RD deletions

M. tuberculosis complex members can be distinguished based on RD deletions. *M. bovis* and *M. caprae* share RD5, RD7, RD8, RD9, RD12 and RD13, but can be distinguished based on RD4 which is absent in *M. bovis*. Proteins associated with the RD5, RD7, RD8, RD9, RD12 and RD13 deletion regions (Table 4.4) were not found in the protein datasets for *M. bovis* or *M. caprae*, confirming the deletion of these regions. Proteins associated with RD4 were found in *M. caprae*.

Table 4.4: Regions of deletion and their associated proteins. *M. bovis* and *M. caprae* share the deletions RD5, RD7, RD8, RD9, RD12 and RD13. RD4 is only absent in *M. bovis*.

Region of deletion	Protein	Gene	Functional category	Proposed function of protein
	Rv1505c	Rv1505c	Conserved hypotheticals	Conserved hypothetical protein
	Rv1507A	Rv1507A	Unknown	Hypothetical protein
	Rv1507c	Rv1507c	Conserved hypotheticals	Conserved hypothetical protein
	Rv1508c	Rv1508c	Cell wall and Cell processes	Probable membrane protein
	Rv1508A	Rv1508A	Conserved hypotheticals	Conserved hypothetical protein
	Rv1509	Rv1509	Unknown	Hypothetical protein
RD4	Rv1510	Rv1510	Cell wall and Cell processes	Conserved Probable membrane protein
	Rv1511	gmdA	Intermediary metabolism and respiration	GDP-D-mannose dehydratase
	Rv1512	epiA	Intermediary metabolism and respiration	Probable nucleotide-sugar epimerase
	Rv1513	Rv1513	Conserved hypotheticals	Conserved hypothetical protein
	Rv1514c	Rv1514c	Conserved hypotheticals	Conserved hypothetical protein
	Rv1515c	Rv1515c	Conserved hypotheticals	Conserved hypothetical protein
	Rv1516c	Rv1516c	Intermediary metabolism and respiration	Probable sugar transferase
	Rv3117	cysA3	Intermediary metabolism and respiration	Probable thiosulphate sulfurtransferase
	Rv3118	sseC1	Intermediary metabolism and respiration	Conserved hypothetical protein
RD5	Rv3119	moaE1	Intermediary metabolism and respiration	Probable molybdenum cofactor biosynthesis protein
	Rv3120	Rv3120	Conserved hypothetical	Conserved hypothetical protein
	Rv3121	cyp141	Intermediary metabolism and respiration	Ptobable cytochrome P450
	Rv1964	yrbE3A	Virulence, detoxification, adaptation	Conserved hypothetical integral membrane protein YRBE3A
	Rv1965	yrbE3B	Virulence, detoxification, adaptation	Conserved hypothetical integral membrane protein YRBE3B
	Rv1966	mce3A	Virulence, detoxification, adaptation	MCE-family protein MCE3A
	Rv1967	mce3B	Virulence, detoxification, adaptation	MCE-family protein MCE3B
	Rv1968	mce3C	Virulence, detoxification, adaptation	MCE-family protein MCE3C
	Rv1969	mce3D	Virulence, detoxification, adaptation	MCE-family protein MCE3D
RD7	Rv1970	lprM	Cell wall and Cell processes	Possible MCE-family lipoprotein
	Rv1971	mce3F	Virulence, detoxification, adaptation	MCE-family protein MCE3F
	Rv1972	Rv1972	Cell wall and Cell processes	Probable conserved MCE associated membrane protein
	Rv1973	Rv1973	Cell wall and Cell processes	Possible conserved MCE associated membrane protein
	Rv1974	Rv1974	Cell wall and Cell processes	Probable conserved membrane protein
	Rv1975	Rv1975	Conserved hypotheticals	Conserved hypothetical protein
	Rv1976c	Rv1976c	Conserved hypotheticals	Conserved hypothetical protein
	Rv1977	Rv1977	Conserved hypotheticals	Conserved hypothetical protein
	Rv3617	ephA	Virulence, detoxification, adaptation	Probable epoxide hydrolase
	Rv3618	Rv3618	Intermediary metabolism and respiration	Possible monooxygenase
	Rv3619c	esxV	Cell wall and Cell processes	Putative ESAT-6 like protein ESXV (ESAT-6 like protein 1)
RD8	Rv3620c	esxW	Cell wall and Cell processes	Putative ESAT-6 like protein ESXW (ESAT-6 like protein 10)
	Rv3621c	PPE65	PE/PPE	PPE family protein PPE65
	Rv3622c	PE32	PE/PPE	PE family protein PE32
	Rv3623	lpqG	Cell wall and Cell processes	Probable Conserved lipoprotein LPQG
	Rv2072c	cobL	Intermediary metabolism and respiration	Precorrin-6Y C(5,15)-methyltransferase
RD9	Rv2073c	Rv2073c	Intermediary metabolism and respiration	Probable shortchain dehydrogenase
	Rv2074	Rv2074	Intermediary metabolism and respiration	Possible pyridoxamine 5'-phosphate oxidase
	Rv2075c	Rv2075c	Cell wall and Cell processes	Possible hypothetical exported or envelop protein
	Rv0221	Rv0221	Lipid metabolism	Possible triacylglycerol synthase
RD10	Rv0222	echA1	Lipid metabolism	Probable enoyl-CoA hydratase
	Rv0223c	Rv0223c	Intermediary metabolism and respiration	Probable aldehyde dehydrogenase
	Rv0866	moaE2	Intermediary metabolism and respiration	Probable molybdenum cofactor biosynthesis
RD12	Rv3119	moaE1	Intermediary metabolism and respiration	Probable molybdenum cofactor biosynthesis
	Rv3120	Rv3120	Conserved hypotheticals	Conserved hypothetical protein
	Rv3121	cyp141	Intermediary metabolism and respiration	Probable cytochrome P450
	Rv1255c	Rv1255c	Regulatory proteins	Probable transcriptional regulatory protein
RD13	Rv1256c	cyp130	Intermediary metabolism and respiration	Probable cytochrome P450
	Rv1257c	Rv1257c	Intermediary metabolism and respiration	Probable oxyreductase

Stellenbosch University http://scholar.sun.ac.za

Chapter 5

Discussion

In this study we aimed to characterize the proteomes of the closely related *Mycobacterium tuberculosis* complex members in relation to *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*. *M. bovis* and *M. caprae* are animal-adapted species and cause disease mainly in bovidae and goats, respectively. *M. bovis* is a promiscuous species and has also been isolated from an array of mammalian species such as humans, deer, llamas, pigs, domesticated animals and wildlife. The relative abundances of the proteins identified in the animal-adapted strains, *M. bovis* and *M. caprae*, were compared to *M. tuberculosis* H37Rv. The metabolic pathways influenced by the proteins which were uniquely detected in a defined member, or were differentially abundant, were analyzed to explore their possible roles in host-specificity and virulence mechanisms.

Genomic analysis versus proteomic analysis

The 4.4 Mb *M. tuberculosis* H37Rv genome and 4.3 Mb *M. bovis* AF2122/97 genome share >99.95% sequence identity (Garnier, Eiglmeier *et al*. 2003). Importantly, there are no unique genes in *M. bovis,* suggesting that differential gene expression and subsequent protein abundances, in conjunction with absence of the genes encoded in the RDs are important determinants in host-specificity and virulence (Garnier, Eiglmeier *et al*. 2003). *M. caprae* is closely to *M. bovis* although its genome has not been sequenced, and therefore it may be assumed that genome sequence of this member would show sequence identity with exception of unique polymorphisms and deletions

Considering the similarity between genomes, it was no surprise that the functional distribution of the 2199, 2367 and 2350 detected proteins extracted during log-phase growth was similar for *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* respectively. Cole *et al.* identified 3924 predicted protein-encoding open reading frames in the *M. tuberculosis* H37Rv genome and later added another 82 when the genome was re-annotated (Cole, Brosch *et al*. 1998). This implies that 54.8%, 59.1% and 58.7% of the proposed *M. tuberculosis*

Stellenbosch University http://scholar.sun.ac.za

proteins are expressed and detectable in *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae*, respectively.

M. bovis and *M. caprae* share the RD5, RD7, RD8, RD9, RD12 and RD13 deletions, but can be distinguished based on RD4 which is deleted only in *M. bovis*. However, most of the proteins encoded within RD4 are classified as conserved hypothetical proteins and thereby their functions remain unknown. Analysis of the functional group distribution of proteins detected and pathways affected by higher or lower abundant proteins in these two species, it was evident that the RD4 deletion does not play a major role in the differentiation of these members. RD5, RD7 and RD8 encode mammalian cell entry (MCE) proteins and ESAT-6 family proteins, which are part of the virulence mechanisms of *M. tuberculosis* H37Rv (Garnier, Eiglmeier *et al*. 2003). The consequence of the deletion of these genes and subsequent absence of these proteins in *M. bovis* and *M. caprae* are difficult to predict. Deletion of the MCE proteins could explain why *M. bovis* and *M. caprae* prefer specific hosts. ESAT-6 family proteins which have been deleted have been said to have an impact on the antigen load either singly or in combination (Sørensen, Nagai *et al*. 1995). It has been suggested that strong selective pressure, imposed by the host immune system, may have resulted in the loss of some genomic domains (Garnier, Eiglmeier *et al*. 2003).

Metabolic pathway analysis of differentially abundant proteins

Metabolic pathways affected by the proteins which were in higher abundance in *M. bovis* relative to *M. tuberculosis* H37Rv included amino acid metabolism, lipid metabolism, oxidative phosphorylation, ribosomal proteins and xenobiotic degradation. The over-abundant proteins in *M. caprae* were also involved in amino acid metabolism, lipid metabolism, oxidative phosphorylation and xenobiotic degradation, as well as carbohydrate metabolism. Lower abundance proteins in *M. bovis* and *M. caprae*, relative to *M. tuberculosis* H37Rv, were found in lipid metabolism and pyrimidine metabolism. The affected metabolic pathways and their possible involvement in host-specificity are discussed below.

Amino acid metabolism

Proteins involved in valine, leucine and isoleucine metabolism were over-abundant in *M. bovis* and *M. caprae*. Mycobacteria have the capacity to synthesize these branched-chain amino acids (Grandoni, Marta *et al*. 1998), and the over-abundance of these proteins may indicate increased synthesis of these amino acids in *M. bovis* and *M. caprae* relative to *M. tuberculosis*. Leucine auxotrophs of *M. bovis* BCG were found to have a reduced ability to survive in the spleens and lungs of mice (McAdam, Weisbrod *et al*. 1995), suggesting that leucine metabolism is required for virulence of *M. bovis* (Grandoni, Marta *et al*. 1998), and similarly in *M. caprae*. Mammalian hosts are unable to produce branched chain amino acids; therefore increased synthesis of these amino acids may be advantageous to these pathogens during infection. This could explain why *M. bovis* is able to infect a variety of hosts and seeing as the intermediates formed during valine, leucine and isoleucine metabolism and the enzymes responsible for these reactions could be targets for preventing pathogenesis of these species (Grandoni, Marta *et al*. 1998). Since mammals do not produce enzymes involved in branched-chain amino acid synthesis; these enzymes could be targeted for anti-tuberculosis drugs with little toxicity to the animal host (Grandoni, Marta *et al*. 1998).

Lipid and fatty acid metabolism

An estimated 8% of the genome of *M. tuberculosis* H37Rv is dedicated to the activities of lipid and fatty acid metabolism (Cole, Brosch *et al*. 1998). *M. tuberculosis* complex members are unique among bacterial pathogens in that they display complex lipids and lipoglycans on their cell surface (Converse, Mougous *et al*. 2003). Cell wall lipids are required for virulence because they interact with the host cell receptors leading to cell invasion (Brennan, Nikaido 1995), provide protection against host damage (Rousseau, Winter *et al*. 2004) and modulate the immune response to infection (Reed, Domenech *et al*. 2004). Several cell wall lipoproteins, trans-membrane proteins and cell wall transfer proteins were over-abundant in *M. bovis* and *M. caprae* relative to *M. tuberculosis* H37Rv. This could imply that *M. bovis* and *M caprae* have a unique cell wall composition which better suits them for survival in their specific hosts.

M. bovis and *M. caprae* showed an increased abundance of lipid and fatty acid metabolism-related proteins, relative to *M. tuberculosis* H37Rv. Amongst the over-abundant proteins are several fatty acyl CoA synthases and dehydrogenases. It has been proposed that fatty acid CoA synthases and dehydrogenases together with polyketide synthases, which synthesize secondary metabolites such as antibiotics in other Actinomycetes, are involved in the synthesis of complex lipids (Kolattukudy, Fernandes *et al*. 1997). Experiments with mutant strains of *M. tuberculosis* have shown that a deficiency in complex lipid biosynthesis can lead to hypo-virulence of previously ‘hyper-lethal’ behavior in murine infection models (Reed, Domenech *et al*. 2004), reduced growth in the lungs of infected mice (Cox, Chen *et al*. 1999) and reduced translocation of lipids to the outer cell membrane (Camacho, Constant *et al*. 2001). Thus, the up-regulation of this pathway could increase the virulence and disease manifestation potential of *M. bovis* and *M. caprae* in their respective hosts.

M. tuberculosis complex members rely on the host cell for a carbon source for survival and therefor the ability to adapt to their immediate environment is important. During infection, *M. tuberculosis* uses fatty acids and cholesterol instead of carbohydrates as a main carbon source (Bloch, Segal 1956). However, in this study these species were cultured in $7H9_{DC}$ media, containing no primary source of fatty acids or cholesterol, which may imply that there was either an innate increase in the expression of these proteins, or that these strains are less equipped to use other carbon sources and are increasing the expression of these proteins in an attempt to obtain more carbon from these “scarce” carbon sources. Proteins identified as under-abundant in *M. bovis* and *M. caprae*, compared to *M. tuberculosis* H37Rv in these pathways, form part of the feedback and regulation loops to keep lipid and fatty acid metabolism balanced. Lower abundance of these proteins contributes to the up-regulation of these pathways in the animal-adapted species. The over-abundance of fatty acid metabolism pathway proteins relative to *M. tuberculosis* may imply that *M. bovis* and *M. caprae* are better adapted to utilize fatty based carbon sources for energy purposes and biosynthesis of virulence-associated lipids during infection.

Xenobiotic degradation

Xenobiotic degradation is also known as drug metabolism and the reactions in these pathways are of particular interest relating to drug resistance in infectious diseases. Increased abundance of the proteins in this metabolic pathway was seen in *M. bovis* and *M. caprae*, relative to *M. tuberculosis* H37Rv. This could lead to *M. bovis* and *M. caprae* being able to readily get rid of foreign substances which target them within the host, for example anti-tuberculosis drugs. *M. bovis* is intrinsically resistant to pyrazinamide (PZA), whereas *M. caprae* is sentivitive to PZA (Niemann, Richter *et al*. 2000). *M. bovis* and *M. caprae* could also be better suited for the harsh environment of the macrophages during infection, seeing as

this pathway would be advantageous in detoxification of stress induced substances such as nitric oxide (NO) (Oliveira, Singh *et al*. 2012).

Oxidative phosphorylation

Cultures were grown in aerobic conditions which lead to ATP generation by oxidative phosphorylation in the electron transport chain. Host-specificity could be explained by the over-abundance of proteins involved in this pathway in *M. bovis* and *M. caprae*, seeing as infection can range from areas within the host of high to low aerobic conditions. When infection is in the lungs specifically, these species will have to adapt to their environment where they will have to compete for oxygen or adapt to the micro-aerophilic or anaerobic environment of the granuloma. As lower activity was observed in this pathway for *M. tuberculosis*, one might speculate that it is better suited to the anaerobic environment of the human lung and is therefore more virulent.

Ribosomal proteins

Ribosomal proteins were found to be over-abundant and uniquely detected in *M. bovis*, compared to *M. caprae* and *M. tuberculosis* H37Rv. Increased abundance of these proteins could relate to the fact that whole cell lysate proteins were harvested in mid-log growth phase, during which time *M. bovis* could have been growing and replicating at a faster rate relative to *M. caprae* and *M. tuberculosis*. Understanding the mechanisms whereby the mRNA code is translated into peptides in bacteria is of great importance in fighting infectious diseases. Mycobacterial tolerance to antibiotics is attributed mostly to the unique properties of the highly impermeable cell wall and has been the main topic of studying virulence mechanisms (Honore, Cole 1994, Martin, Timm *et al*. 1990). The structures of the ribosome differ in human and bacterial cells, thus animal cells too, and this has led to some antibiotics to kill bacteria only. Many anti-tuberculosis drugs, such as streptomycin target translation or translation machinery (Kenney, Churchward 1994). By identifying ribosomal proteins only

detected during the actively growing and replicating phase of *M. bovis*, relative to *M. tuberculosis* and *M. caprae*, one can investigate the possibilities of targeting the mechanisms in the different steps of translation to prevent disease manifestation in hosts. In addition, structural studies of the complex ribosome could also lead to the understanding of the unique properties of Mycobacteria for *e.g.* slow growth (Lonnroth, Ridell 1985).

Carbohydrate metabolism

Proteins involved in carbohydrate metabolism were more abundant in *M. caprae*, relative to *M. tuberculosis* H37Rv and *M. bovis*. As previously mentioned, pathogenic Mycobacteria prefer fatty acids and cholesterol to carbohydrates as a carbon source. (BLOCH, SEGAL 1956)(Munoz-Elias, McKinney 2005) Estimates of the concentrations of potential substrates available to a pathogen in host tissues suggest that lipids and sterols are more abundant than carbohydrates (Cole, Brosch *et al*. 1998)(Marrero, Trujillo *et al*. 2013). However, Mycobacteria can co-metabolize multiple carbon sources including carbohydrates. The importance of carbohydrate metabolism in *in vivo* growth and persistence of Mycobacteria remains unclear (Marrero, Trujillo *et al*. 2013). In an attempt to understand the role of carbohydrate metabolism during infection, Marrero *et al.* showed that mutations in the enzymes responsible for the phosphorylation of glucose lead to *M. tuberculosis* H37Rv not being able to persist in mouse lungs (Marrero, Trujillo *et al*. 2013). Up-regulation of this pathway in *M. caprae* could give this species an advantage of utilizing carbohydrates supplied by the host during infection and could lead to better survival in goats.

Pyrimidine metabolism

Proteins involved in pyrimidine metabolism were less abundant in *M. caprae*. *M. tuberculosis* recycles bases and/or nucleosides to survive in the hostile environment imposed by the host via a pyrimidine salvage pathway (Villela, Basso *et al*. 2013)(Moffatt, Ashihara 2002). Enzymes involved in the pyrimidine salvage pathway might be attractive

Stellenbosch University http://scholar.sun.ac.za

targets for rational drug design against TB, since this pathway is vital for all bacterial cells, plays an important role in the latent state and is composed of enzymes considerably different from those present in humans (Villela, Basso *et al*. 2013). Down-regulation can be attributed to a non-hostile environment in the culture media to which *M. caprae* could be more sensitive to or slower pyrimidine salvaging *in vivo* which could lead to slower conversion to latent phase TB in goats.

Conclusion

In this study the proteomes of the *M. tuberculosis* complex members, *M. tuberculosis* H37Rv, *M. bovis* and *M. caprae* were characterized and the variation between detected proteins and protein abundances explored in order to describe differences between these closely related strains. The influence of these proteins on various metabolic processes is explored and correlated with the specific or broad host range of these Mycobacteria. By changing the rate of expression of some of these metabolic pathways, they are able to adapt to their environment more readily which helps them to survive inside the hosts and cause severe pathogenesis if needed.

Stellenbosch University http://scholar.sun.ac.za

Chapter 6

Limitations and Future directions

Stellenbosch University http://scholar.sun.ac.za

6.1 Limitations

A major limitation of this study is our limited knowledge on how *in vitro* culture conditions translate to *in vivo* condition where infection, virulence and pathogenesis determinants may be expressed. Thus, it is possible that the proteins expressed under optimal culture conditions during log-growth phase will differ from the proteins expressed during growth within the macrophage where the pathogen experiences a multitude of stressors.

The genome of *M. caprae* has not yet been annotated, and thus the proteins identified in the whole cell lysate and their associated functions could differ slightly from those in *M. tuberculosis* H37Rv and *M. bovis*. Non-synonymous single nucleotide polymorphisms (nsSNPs) in the genes encoding proteins in *M. caprae* would not be detected by MaxQuant as these peptides would not be present in the searched peptide databases, *M. tuberculosis* H37Rv and *M. bovis*.

The LTQ Orbitrap Velos mass spectrometer has a detection limit and cannot detect proteins of extremely low abundance implying that unique low abundant proteins for each MTBC member may be missed. These proteins could form part of a pathway which is associated with virulence, host-specificity and pathogenesis.

6.2 Future directions

Whole cell lysate protein extracts can be taken at different time points during growth phases (*in vivo* and *in vitro*) to shed light on how the proteome changes during as a function of the nutrient status of the medium. This will contribute to understanding how the proteome adapts to a changing environment.

Whole genome sequencing of *M. caprae* will contribute to differentiation of this member from the other members in the MTBC, by identification of the SNPs, RDs and large nucleotide polymorphisms (LSPs). Annotation of the genome may lead to the identification of new proteins specific to *M. caprae*. The genome of an *M. caprae* isolate will be sequenced on the Illumina HiSeq2000 platform using a paired-end approach, with 500 base fragment sizes resulting in sequences of between 350 and 550 bases. Whole genome sequencing will be done in collaboration with the King Abdullah University of Technology (KAUST) location. The depth of coverage for the sequence will be at least or above 100, this, together with the excellent quality of the sequencing data, will ensure a high level of confidence for identifying variation in the genome, relative to *M. tuberculosis* and *M. bovis*. The consensus sequence that will be obtained will be automatically annotated using the Rapid Annotation using Subsystem Technology (RAST) server.

Future research on animal-adapted Mycobacterial species will contribute to preventing transmission and spread of disease to domesticated animals, free and captive wildlife species. Understanding the mechanisms of virulence and pathogenicity could lead to development of efficient vaccines and diagnostic tests for a variety of animal hosts. Better control regimes of zoonotic diseases could have a positive influence on the economics and agricultural industries of a country.

Stellenbosch University http://scholar.sun.ac.za

Chapter 7

List of References

ABLORDEY, A., SWINGS, J., HUBANS, C., CHEMLAL, K., LOCHT, C., PORTAELS, F. and SUPPLY, P., 2005. Multilocus variable-number tandem repeat typing of *Mycobacterium ulcerans*. *Journal of clinical microbiology,* **43**(4), pp. 1546-1551.

ALEXANDER, K.A., LAVER, P.N., MICHEL, A.L., WILLIAMS, M., VAN HELDEN, P.D., WARREN, R.M. and GEY VAN PITTIUS, N.C., 2010. Novel *Mycobacterium tuberculosis* complex pathogen, M. mungi. *Emerging infectious diseases,* **16**(8), pp. 1296-1299.

AMENI, G., ASEFFA, A., ENGERS, H., YOUNG, D., GORDON, S., HEWINSON, G. and VORDERMEIER, M., 2007. High prevalence and increased severity of pathology of bovine tuberculosis in Holsteins compared to zebu breeds under field cattle husbandry in central Ethiopia. *Clinical and vaccine immunology : CVI,* **14**(10), pp. 1356-1361.

ANDREJAK, C., THOMSEN, V.O., JOHANSEN, I.S., RIIS, A., BENFIELD, T.L., DUHAUT, P., SORENSEN, H.T., LESCURE, F.X. and THOMSEN, R.W., 2010. Nontuberculous pulmonary mycobacteriosis in Denmark: incidence and prognostic factors. *American journal of respiratory and critical care medicine,* **181**(5), pp. 514-521.

ANGKAWANISH, T., WAJJWALKU, W., SIRIMALAISUWAN, A., MAHASAWANGKUL, S., KAEWSAKHORN, T., BOONSRI, K. and RUTTEN, V.P., 2010. *Mycobacterium tuberculosis* infection of domesticated Asian elephants, Thailand. *Emerging infectious diseases,* **16**(12), pp. 1949-1951.

BERCOVIER, H. and VINCENT, V., 2001. Mycobacterial infections in domestic and wild animals due to *Mycobacterium marinum*, *M. fortuitum, M. chelonae, M. porcinum, M. farcinogenes, M. smegmatis, M. scrofulaceum, M. xenopi, M. kansasii, M. simiae* and *M. genavense*. *Revue scientifique et technique (International Office of Epizootics),* **20**(1), pp. 265-290.

BERG, S., GARCIA-PELAYO, M.C., MULLER, B., HAILU, E., ASIIMWE, B., KREMER, K., DALE, J., BONIOTTI, M.B., RODRIGUEZ, S., HILTY, M., RIGOUTS, L., FIRDESSA, R., MACHADO, A., MUCAVELE, C., NGANDOLO, B.N., BRUCHFELD, J., BOSCHIROLI, L., MULLER, A., SAHRAOUI, N., PACCIARINI, M., CADMUS, S., JOLOBA, M., VAN SOOLINGEN, D., MICHEL, A.L., DJONNE, B., ARANAZ, A., ZINSSTAG, J., VAN HELDEN, P., PORTAELS, F., KAZWALA, R., KALLENIUS, G., HEWINSON, R.G., ASEFFA, A., GORDON, S.V. and SMITH, N.H., 2011. African 2, a clonal complex of *Mycobacterium bovis* epidemiologically important in East Africa. *Journal of Bacteriology,* **193**(3), pp. 670-678.

BLOCH, H. and SEGAL, W., 1956. Biochemical differentiation of *Mycobacterium tuberculosis* grown in vivo and in vitro. *Journal of Bacteriology,* **72**(2), pp. 132-141.

BRAGG, R.R., HUCHZERMEYER, H.F. and HANISCH, M.A., 1990. *Mycobacterium fortuitum* isolated from three species of fish in South Africa. *The Onderstepoort journal of veterinary research,* **57**(1), pp. 101-102.

BRENNAN, P.J. and NIKAIDO, H., 1995. The envelope of mycobacteria. *Annual Review of Biochemistry,* **64**, pp. 29-63.

Stellenbosch University http://scholar.sun.ac.za

BROSCH, R., GORDON, S.V., MARMIESSE, M., BRODIN, P., BUCHRIESER, C., EIGLMEIER, K., GARNIER, T., GUTIERREZ, C., HEWINSON, G., KREMER, K., PARSONS, L.M., PYM, A.S., SAMPER, S., VAN SOOLINGEN, D. and COLE, S.T., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proceedings of the National Academy of Sciences of the United States of America,* **99**(6), pp. 3684-3689.

BROWN, T.H., 1985. The rapidly growing mycobacteria--*Mycobacterium fortuitum* and *Mycobacterium chelonei*. *Infection control : IC,* **6**(7), pp. 283-288.

CAMACHO, L.R., CONSTANT, P., RAYNAUD, C., LANEELLE, M.A., TRICCAS, J.A., GICQUEL, B., DAFFE, M. and GUILHOT, C., 2001. Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *The Journal of biological chemistry,* **276**(23), pp. 19845-19854.

CHAMOISEAU, G., 1973. "*Mycobacterium farcinogenes*" causal agent of bovine farcy in Africa (author's transl). *Annales de Microbiologie,* **124**(2), pp. 215-222.

CHIHOTA, V.N., MULLER, B., MLAMBO, C.K., PILLAY, M., TAIT, M., STREICHER, E.M., MARAIS, E., VAN DER SPUY, G.D., HANEKOM, M., COETZEE, G., TROLLIP, A., HAYES, C., BOSMAN, M.E., GEY VAN PITTIUS, N.C., VICTOR, T.C., VAN HELDEN, P.D. and WARREN, R.M., 2012. Population structure of multi- and extensively drug-resistant *Mycobacterium tuberculosis* strains in South Africa. *Journal of clinical microbiology,* **50**(3), pp. 995-1002.

COLE, S.T., BROSCH, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S.V., EIGLMEIER, K., GAS, S., BARRY, C.E.,3RD, TEKAIA, F., BADCOCK, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K., FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL, M.A., RAJANDREAM, M.A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J., SQUARES, R., SQUARES, S., SULSTON, J.E., TAYLOR, K., WHITEHEAD, S. and BARRELL, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature,* **393**(6685), pp. 537-544.

CONVERSE, S.E., MOUGOUS, J.D., LEAVELL, M.D., LEARY, J.A., BERTOZZI, C.R. and COX, J.S., 2003. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proceedings of the National Academy of Sciences of the United States of America,* **100**(10), pp. 6121-6126.

COOK, J.L., 2010. Nontuberculous mycobacteria: opportunistic environmental pathogens for predisposed hosts. *British medical bulletin,* **96**, pp. 45-59.

COSIVI, O., GRANGE, J.M., DABORN, C.J., RAVIGLIONE, M.C., FUJIKURA, T., COUSINS, D., ROBINSON, R.A., HUCHZERMEYER, H.F., DE KANTOR, I. and MESLIN, F.X., 1998. Zoonotic tuberculosis due to *Mycobacterium bovis* in developing countries. *Emerging infectious diseases,* **4**(1), pp. 59-70.

Stellenbosch University http://scholar.sun.ac.za

COX, J. and MANN, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology,* **26**(12), pp. 1367-1372.

COX, J.S., CHEN, B., MCNEIL, M. and JACOBS, W.R.,JR, 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature,* **402**(6757), pp. 79-83.

CVETNIC, Z., KATALINIC-JANKOVIC, V., SOSTARIC, B., SPICIC, S., OBROVAC, M., MARJANOVIC, S., BENIC, M., KIRIN, B.K. and VICKOVIC, I., 2007. *Mycobacterium caprae* in cattle and humans in Croatia. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease,* **11**(6), pp. 652-658.

DE GROOTE, M.A. and HUITT, G., 2006. Infections Due to Rapidly Growing Mycobacteria. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America,* **42**, pp. 1756-1763.

DE LA RUA-DOMENECH, R., GOODCHILD, A.T., VORDERMEIER, H.M., HEWINSON, R.G., CHRISTIANSEN, K.H. and CLIFTON-HADLEY, R.S., 2006. Ante mortem diagnosis of tuberculosis in cattle: a review of the tuberculin tests, gamma-interferon assay and other ancillary diagnostic techniques. *Research in veterinary science,* **81**(2), pp. 190-210.

DE SOUZA, G.A., ARNTZEN, M.O., FORTUIN, S., SCHURCH, A.C., MALEN, H., MCEVOY, C.R., VAN SOOLINGEN, D., THIEDE, B., WARREN, R.M. and WIKER, H.G., 2011. Proteogenomic analysis of polymorphisms and gene annotation divergences in prokaryotes using a clustered mass spectrometry-friendly database. *Molecular & cellular proteomics : MCP,* **10**(1), pp. M110.002527.

DEN BOON, S., VAN LILL, S.W., BORGDORFF, M.W., ENARSON, D.A., VERVER, S., BATEMAN, E.D., IRUSEN, E., LOMBARD, C.J., WHITE, N.W., DE VILLIERS, C. and BEYERS, N., 2007. High prevalence of tuberculosis in previously treated patients, Cape Town, South Africa. *Emerging infectious diseases,* **13**(8), pp. 1189-1194.

DREWE, J.A., EAMES, K.T., MADDEN, J.R. and PEARCE, G.P., 2011. Integrating contact network structure into tuberculosis epidemiology in meerkats in South Africa: Implications for control. *Preventive veterinary medicine,* **101**(1-2), pp. 113-120.

FALKINHAM, J.O., 2010. Impact of human activities on the ecology of nontuberculous mycobacteria. *Future microbiology,* **5**(6), pp. 951-960.

FALKINHAM, J.O.,3RD, 2011. Nontuberculous mycobacteria from household plumbing of patients with nontuberculous mycobacteria disease. *Emerging infectious diseases,* **17**(3), pp. 419-424.

GARNIER, T., EIGLMEIER, K., CAMUS, J.C., MEDINA, N., MANSOOR, H., PRYOR, M., DUTHOY, S., GRONDIN, S., LACROIX, C., MONSEMPE, C., SIMON, S., HARRIS, B., ATKIN, R., DOGGETT, J., MAYES, R., KEATING, L., WHEELER, P.R., PARKHILL, J., BARRELL, B.G., COLE, S.T., GORDON, S.V. and HEWINSON, R.G., 2003. The

Stellenbosch University http://scholar.sun.ac.za

complete genome sequence of *Mycobacterium bovis*. *Proceedings of the National Academy of Sciences of the United States of America,* **100**(13), pp. 7877-7882.

GEY VAN PITTIUS, N.C., PERRETT, K.D., MICHEL, A.L., KEET, D.F., HLOKWE, T., STREICHER, E.M., WARREN, R.M. and VAN HELDEN, P.D., 2012. Infection of African buffalo (*Syncerus caffer*) by oryx bacillus, a rare member of the antelope clade of the *Mycobacterium tuberculosis* complex. *Journal of wildlife diseases,* **48**(4), pp. 849-857.

GEY VAN PITTIUS, N.C., SAMPSON, S.L., LEE, H., KIM, Y., VAN HELDEN, P.D. and WARREN, R.M., 2006. Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. *BMC evolutionary biology,* **6**, pp. 95.

GEY VAN PITTIUS, N.C., VAN HELDEN, P.D. and WARREN, R.M., 2012. Characterization of *Mycobacterium orygis*. *Emerging infectious diseases,* **18**(10), pp. 1708-1709.

GILLESPIE, J.J., WATTAM, A.R., CAMMER, S.A., GABBARD, J.L., SHUKLA, M.P., DALAY, O., DRISCOLL, T., HIX, D., MANE, S.P., MAO, C., NORDBERG, E.K., SCOTT, M., SCHULMAN, J.R., SNYDER, E.E., SULLIVAN, D.E., WANG, C., WARREN, A., WILLIAMS, K.P., XUE, T., YOO, H.S., ZHANG, C., ZHANG, Y., WILL, R., KENYON, R.W. and SOBRAL, B.W., 2011. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. *Infection and immunity,* **79**(11), pp. 4286-4298.

GRANDONI, J.A., MARTA, P.T. and SCHLOSS, J.V., 1998. Inhibitors of branched-chain amino acid biosynthesis as potential antituberculosis agents. *The Journal of antimicrobial chemotherapy,* **42**(4), pp. 475-482.

GRIFFITH, D.E., AKSAMIT, T., BROWN-ELLIOTT, B.A., CATANZARO, A., DALEY, C., GORDIN, F., HOLLAND, S.M., HORSBURGH, R., HUITT, G., IADEMARCO, M.F., ISEMAN, M., OLIVIER, K., RUOSS, S., VON REYN, C.F., WALLACE, R.J.,JR, WINTHROP, K., ATS MYCOBACTERIAL DISEASES SUBCOMMITTEE, AMERICAN THORACIC SOCIETY and INFECTIOUS DISEASE SOCIETY OF AMERICA, 2007. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *American journal of respiratory and critical care medicine,* **175**(4), pp. 367-416.

GUMI, B., SCHELLING, E., BERG, S., FIRDESSA, R., ERENSO, G., MEKONNEN, W., HAILU, E., MELESE, E., HUSSEIN, J., ASEFFA, A. and ZINSSTAG, J., 2012. Zoonotic transmission of tuberculosis between pastoralists and their livestock in South-East Ethiopia. *EcoHealth,* **9**(2), pp. 139-149.

HANEKOM, M., VAN DER SPUY, G.D., GEY VAN PITTIUS, N.C., MCEVOY, C.R., NDABAMBI, S.L., VICTOR, T.C., HOAL, E.G., VAN HELDEN, P.D. and WARREN, R.M., 2007. Evidence that the spread of *Mycobacterium tuberculosis* strains with the Beijing genotype is human population dependent. *Journal of clinical microbiology,* **45**(7), pp. 2263-2266.

Stellenbosch University http://scholar.sun.ac.za

HATHERILL, M., HAWKRIDGE, T., WHITELAW, A., TAMERIS, M., MAHOMED, H., MOYO, S., HANEKOM, W. and HUSSEY, G., 2006. Isolation of non-tuberculous mycobacteria in children investigated for pulmonary tuberculosis. *PloS one,* **1**, pp. e21.

HOLLAND, S.M., 2001. Nontuberculous mycobacteria. *The American Journal of the Medical Sciences,* **321**(1), pp. 49-55.

HONORE, N. and COLE, S.T., 1994. Streptomycin resistance in mycobacteria. *Antimicrobial Agents and Chemotherapy,* **38**(2), pp. 238-242.

HUARD, R.C., FABRE, M., DE HAAS, P., LAZZARINI, L.C., VAN SOOLINGEN, D., COUSINS, D. and HO, J.L., 2006. Novel genetic polymorphisms that further delineate the phylogeny of the *Mycobacterium tuberculosis* complex. *Journal of Bacteriology,* **188**(12), pp. 4271-4287.

ISEMAN, M.D. and MARRAS, T.K., 2008. The importance of nontuberculous mycobacterial lung disease. *American journal of respiratory and critical care medicine,* **178**(10), pp. 999-1000.

KAMERBEEK, J., SCHOULS, L., KOLK, A., VAN AGTERVELD, M., VAN SOOLINGEN, D., KUIJPER, S., BUNSCHOTEN, A., MOLHUIZEN, H., SHAW, R., GOYAL, M. and VAN EMBDEN, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *Journal of clinical microbiology,* **35**(4), pp. 907-914.

KANEENE, J.B. and THOEN, C.O., 2004. Tuberculosis. *Journal of the American Veterinary Medical Association,* **224**(5), pp. 685-691.

KANEHISA, M. and GOTO, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research,* **28**(1), pp. 27-30.

KANKYA, C., MUWONGE, A., DJONNE, B., MUNYEME, M., OPUDA-ASIBO, J., SKJERVE, E., OLOYA, J., EDVARDSEN, V. and JOHANSEN, T.B., 2011. Isolation of non-tuberculous mycobacteria from pastoral ecosystems of Uganda: public health significance. *BMC public health,* **11**, pp. 320-2458-11-320.

KARNE, S.S., SANGLE, S.A., KIYAWAT, D.S., DHARMASHALE, S.N., KADAM, D.B. and BHARDWAJ, R.S., 2012. *Mycobacterium avium-intracellulare* brain abscess in HIV-positive patient. *Annals of Indian Academy of Neurology,* **15**(1), pp. 54-55.

KATALE, B.Z., MBUGI, E.V., KENDAL, S., FYUMAGWA, R.D., KIBIKI, G.S., GODFREY-FAUSSETT, P., KEYYU, J.D., VAN HELDEN, P. and MATEE, M.I., 2012. Bovine tuberculosis at the human-livestock-wildlife interface: Is it a public health problem in Tanzania? A review. *The Onderstepoort journal of veterinary research,* **79**(2), pp. E1-8.

KATOCH, V.M., 2004. Infections due to non-tuberculous mycobacteria (NTM). *The Indian journal of medical research,* **120**(4), pp. 290-304.

KAUFMANN, S.H., 2002. Protection against tuberculosis: cytokines, T cells, and macrophages. *Annals of the Rheumatic Diseases,* **61 Suppl 2**, pp. ii54-8.

KENNEY, T.J. and CHURCHWARD, G., 1994. Cloning and sequence analysis of the rpsL and rpsG genes of *Mycobacterium smegmatis* and characterization of mutations causing resistance to streptomycin. *Journal of Bacteriology,* **176**(19), pp. 6153-6156.

KOLATTUKUDY, P.E., FERNANDES, N.D., AZAD, A.K., FITZMAURICE, A.M. and SIRAKOVA, T.D., 1997. Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. *Molecular microbiology,* **24**(2), pp. 263-270.

LOBUE, P.A., ENARSON, D.A. and THOEN, C.O., 2010. Tuberculosis in humans and animals: an overview. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease,* **14**(9), pp. 1075-1078.

LONNROTH, L. and RIDELL, M., 1985. Analyses of ribosomal proteins from mycobacteria with particular reference to the beta antigen. *FEMS Microbiology Letters,* **27**, pp. 203-207.

MARRERO, J., TRUJILLO, C., RHEE, K.Y. and EHRT, S., 2013. Glucose phosphorylation is required for *Mycobacterium tuberculosis* persistence in mice. *PLoS pathogens,* **9**(1), pp. e1003116.

MARTIN, C., TIMM, J., RAUZIER, J., GOMEZ-LUS, R., DAVIES, J. and GICQUEL, B., 1990. Transposition of an antibiotic resistance element in mycobacteria. *Nature,* **345**(6277), pp. 739-743.

MCADAM, R.A., WEISBROD, T.R., MARTIN, J., SCUDERI, J.D., BROWN, A.M., CIRILLO, J.D., BLOOM, B.R. and JACOBS, W.R.,JR, 1995. In vivo growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. *Infection and immunity,* **63**(3), pp. 1004-1012.

MICHEL, A.L., COETZEE, M.L., KEET, D.F., MARE, L., WARREN, R., COOPER, D., BENGIS, R.G., KREMER, K. and VAN HELDEN, P., 2009. Molecular epidemiology of *Mycobacterium bovis* isolates from free-ranging wildlife in South African game reserves. *Veterinary microbiology,* **133**(4), pp. 335-343.

MICHEL, A.L., DE KLERK, L.M., GEY VAN PITTIUS, N.C., WARREN, R.M. and VAN HELDEN, P.D., 2007. Bovine tuberculosis in African buffaloes: observations regarding *Mycobacterium bovis* shedding into water and exposure to environmental mycobacteria. *BMC veterinary research,* **3**, pp. 23.

MICHEL, A.L., MULLER, B. and VAN HELDEN, P.D., 2010. *Mycobacterium bovis* at the animal-human interface: a problem, or not? *Veterinary microbiology,* **140**(3-4), pp. 371-381.

MOFFATT, B.A. and ASHIHARA, H., 2002. Purine and pyrimidine nucleotide synthesis and metabolism. *The Arabidopsis book / American Society of Plant Biologists,* **1**, pp. e0018.

MONTALI, R.J., MIKOTA, S.K. and CHENG, L.I., 2001. *Mycobacterium tuberculosis* in zoo and wildlife species. *Revue scientifique et technique (International Office of Epizootics),* **20**(1), pp. 291-303.

MOORE, J.E., KRUIJSHAAR, M.E., ORMEROD, L.P., DROBNIEWSKI, F. and ABUBAKAR, I., 2010. Increasing reports of non-tuberculous mycobacteria in England, Wales and Northern Ireland, 1995-2006. *BMC public health,* **10**, pp. 612-2458-10-612.

MULLER, B., DE KLERK-LORIST, L.M., HENTON, M.M., LANE, E., PARSONS, S., GEY VAN PITTIUS, N.C., KOTZE, A., VAN HELDEN, P.D. and TANNER, M., 2011. Mixed infections of *Corynebacterium pseudotuberculosis* and non-tuberculous mycobacteria in South African antelopes presenting with tuberculosis-like lesions. *Veterinary microbiology,* **147**(3-4), pp. 340-345.

MULLER, B., STEINER, B., BONFOH, B., FANE, A., SMITH, N.H. and ZINSSTAG, J., 2008. Molecular characterisation of *Mycobacterium bovis* isolated from cattle slaughtered at the Bamako abattoir in Mali. *BMC veterinary research,* **4**, pp. 26-6148-4-26.

MUNOZ-ELIAS, E.J. and MCKINNEY, J.D., 2005. *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nature medicine,* **11**(6), pp. 638-644.

MUNYEME, M., RIGOUTS, L., SHAMPUTA, I.C., MUMA, J.B., TRYLAND, M., SKJERVE, E. and DJONNE, B., 2009. Isolation and characterization of *Mycobacterium bovis* strains from indigenous Zambian cattle using Spacer oligonucleotide typing technique. *BMC microbiology,* **9**, pp. 144-2180-9-144.

MUYOYETA, M., DE HAAS, P.E., MUELLER, D.H., VAN HELDEN, P.D., MWENGE, L., SCHAAP, A., KRUGER, C., GEY VAN PITTIUS, N.C., LAWRENCE, K., BEYERS, N., GODFREY-FAUSSETT, P. and AYLES, H., 2010. Evaluation of the Capilia TB assay for culture confirmation of *Mycobacterium tuberculosis* infections in Zambia and South Africa. *Journal of clinical microbiology,* **48**(10), pp. 3773-3775.

NIEMANN, S., RICHTER, E. and RUSCH-GERDES, S., 2000. Differentiation among members of the *Mycobacterium tuberculosis* complex by molecular and biochemical features: evidence for two pyrazinamide-susceptible subtypes of *M. bovis*. *Journal of clinical microbiology,* **38**(1), pp. 152-157.

OLIVEIRA, A., SINGH, S., BIDON-CHANAL, A., FORTI, F., MARTI, M.A., BOECHI, L., ESTRIN, D.A., DIKSHIT, K.L. and LUQUE, F.J., 2012. Role of PheE15 gate in ligand entry and nitric oxide detoxification function of *Mycobacterium tuberculosis* truncated hemoglobin N. *PloS one,* **7**(11), pp. e49291.

PARSONS, S.D.C., DREWE, J.A., GEY VAN PITTIUS, N.C., WARREN, R.M. and VAN HELDEN, P.D., 2013. A novel pathogen, *Mycobacterium suricattae*, is the cause of tuberculosis in meerkats (*Suricata suricatta*) in South Africa. *Emerging infectious diseases,* 19(12), pp. 2004-2007.

PARSONS, S., SMITH, S.G., MARTINS, Q., HORSNELL, W.G., GOUS, T.A., STREICHER, E.M., WARREN, R.M., VAN HELDEN, P.D. and GEY VAN PITTIUS, N.C., 2008. Pulmonary infection due to the dassie bacillus (*Mycobacterium tuberculosis* complex sp.) in a free-living dassie (rock hyrax-*Procavia capensis*) from South Africa. *Tuberculosis (Edinburgh, Scotland),* **88**(1), pp. 80-83.

Stellenbosch University http://scholar.sun.ac.za

PARSONS, S.D., GOUS, T.A., WARREN, R.M. and VAN HELDEN, P.D., 2008. Pulmonary *Mycobacterium tuberculosis* (Beijing strain) infection in a stray dog. *Journal of the South African Veterinary Association,* **79**(2), pp. 95-98.

PARSONS, S.D., WARREN, R.M., OTTENHOFF, T.H., GEY VAN PITTIUS, N.C. and VAN HELDEN, P.D., 2012. Detection of *Mycobacterium tuberculosis* infection in dogs in a high-risk setting. *Research in veterinary science,* **92**(3), pp. 414-419.

PETRINI, B., 2006. *Mycobacterium abscessu*s: an emerging rapid-growing potential pathogen. *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* **114**(5), pp. 319-328.

PFYFFER, G.E., AUCKENTHALER, R., VAN EMBDEN, J.D. and VAN SOOLINGEN, D., 1998. *Mycobacterium canettii*, the smooth variant of *M. tuberculosis*, isolated from a Swiss patient exposed in Africa. *Emerging infectious diseases,* **4**(4), pp. 631-634.

PHILLIPS, M.S. and VON REYN, C.F., 2001. Nosocomial infections due to nontuberculous mycobacteria. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America,* **33**(8), pp. 1363-1374.

PREVOTS, D.R., SHAW, P.A., STRICKLAND, D., JACKSON, L.A., RAEBEL, M.A., BLOSKY, M.A., MONTES DE OCA, R., SHEA, Y.R., SEITZ, A.E., HOLLAND, S.M. and OLIVIER, K.N., 2010. Nontuberculous mycobacterial lung disease prevalence at four integrated health care delivery systems. *American journal of respiratory and critical care medicine,* **182**(7), pp. 970-976.

REED, M.B., DOMENECH, P., MANCA, C., SU, H., BARCZAK, A.K., KREISWIRTH, B.N., KAPLAN, G. and BARRY, C.E.,3RD, 2004. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature,* **431**(7004), pp. 84-87.

RODRIGUEZ, S., BEZOS, J., ROMERO, B., DE JUAN, L., ALVAREZ, J., CASTELLANOS, E., MOYA, N., LOZANO, F., JAVED, M.T., SAEZ-LLORENTE, J.L., LIEBANA, E., MATEOS, A., DOMINGUEZ, L., ARANAZ, A. and SPANISH NETWORK ON SURVEILLANCE AND MONITORING OF ANIMAL TUBERCULOSIS, 2011. *Mycobacterium caprae* infection in livestock and wildlife, Spain. *Emerging infectious diseases,* **17**(3), pp. 532-535.

ROUSSEAU, C., WINTER, N., PIVERT, E., BORDAT, Y., NEYROLLES, O., AVE, P., HUERRE, M., GICQUEL, B. and JACKSON, M., 2004. Production of phthiocerol dimycocerosates protects *Mycobacterium tuberculosis* from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection. *Cellular microbiology,* **6**(3), pp. 277-287.

SHAMAEI, M., MARJANI, M., FARNIA, P., TABARSI, P. and MANSOURI, D., 2010. Human infections due to *Mycobacterium lentiflavum*: first report in Iran. *Iranian journal of microbiology,* **2**(1), pp. 27-29.

SIMONS, S., VAN INGEN, J., HSUEH, P.R., VAN HUNG, N., DEKHUIJZEN, P.N., BOEREE, M.J. and VAN SOOLINGEN, D., 2011. Nontuberculous mycobacteria in respiratory tract infections, eastern Asia. *Emerging infectious diseases,* **17**(3), pp. 343-349.

SKINNER, M.A., WEDLOCK, D.N. and BUDDLE, B.M., 2001. Vaccination of animals against Mycobacterium bovis. *Revue scientifique et technique (International Office of Epizootics),* **20**(1), pp. 112-132.

SMITH, M.I., YATSUNENKO, T., MANARY, M.J., TREHAN, I., MKAKOSYA, R., CHENG, J., KAU, A.L., RICH, S.S., CONCANNON, P., MYCHALECKYJ, J.C., LIU, J., HOUPT, E., LI, J.V., HOLMES, E., NICHOLSON, J., KNIGHTS, D., URSELL, L.K., KNIGHT, R. and GORDON, J.I., 2013. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science (New York, N.Y.),* **339**(6119), pp. 548-554.

SMITH, N.H., HEWINSON, R.G., KREMER, K., BROSCH, R. and GORDON, S.V., 2009. Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nature reviews.Microbiology,* **7**(7), pp. 537-544.

SØRENSEN, A.L., NAGAI, S., HOUEN, G., ANDERSEN, P. and ANDERSEN, A.B., 1995. Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*.. *Infection and Immunity,* **63**(5), pp. 1710-1717.

STREICHER, E.M., VICTOR, T.C., VAN DER SPUY, G., SOLA, C., RASTOGI, N., VAN HELDEN, P.D. and WARREN, R.M., 2007. Spoligotype signatures in the *Mycobacterium tuberculosis* complex. *Journal of clinical microbiology,* **45**(1), pp. 237-240.

VAN EMBDEN, J.D., CAVE, M.D., CRAWFORD, J.T., DALE, J.W., EISENACH, K.D., GICQUEL, B., HERMANS, P., MARTIN, C., MCADAM, R. and SHINNICK, T.M., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *Journal of clinical microbiology,* **31**(2), pp. 406-409.

VAN HELDEN, P.D., PARSONS, S.D. and GEY VAN PITTIUS, N.C., 2009. 'Emerging' mycobacteria in South Africa. *Journal of the South African Veterinary Association,* **80**(4), pp. 210-214.

VAN INGEN, J., BOEREE, M.J., VAN SOOLINGEN, D. and MOUTON, J.W., 2012. Resistance mechanisms and drug susceptibility testing of nontuberculous mycobacteria. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy,* **15**(3), pp. 149-161.

VAN INGEN, J., RAHIM, Z., MULDER, A., BOEREE, M.J., SIMEONE, R., BROSCH, R. and VAN SOOLINGEN, D., 2012. Characterization of *Mycobacterium orygis* as *M. tuberculosis* complex subspecies. *Emerging infectious diseases,* **18**(4), pp. 653-655.

VILLELA, A.D., BASSO, L.A. and SANTOS, D.S., 2013. Biochemical Characterization of Uracil Phosphoribosyltransferase from *Mycobacterium tuberculosis*. *PLoS One,* **8**(2), pp. e56445.

WARREN, R.M., GEY VAN PITTIUS, N.C., BARNARD, M., HESSELING, A., ENGELKE, E., DE KOCK, M., GUTIERREZ, M.C., CHEGE, G.K., VICTOR, T.C., HOAL, E.G. and VAN HELDEN, P.D., 2006a. Differentiation of *Mycobacterium tuberculosis* complex by PCR amplification of genomic regions of difference. *The international journal of*

Stellenbosch University http://scholar.sun.ac.za

tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease, **10**(7), pp. 818-822.

WARREN, R.M., GEY VAN PITTIUS, N.C., BARNARD, M., HESSELING, A., ENGELKE, E., DE KOCK, M., GUTIERREZ, M.C., CHEGE, G.K., VICTOR, T.C., HOAL, E.G. and VAN HELDEN, P.D., 2006b. Differentiation of *Mycobacterium tuberculosis* complex by PCR amplification of genomic regions of difference. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease,* **10**(7), pp. 818-822.

WHITTINGTON, R.J., MARSH, I.B. and REDDACLIFF, L.A., 2005. Survival of *Mycobacterium avium subsp. paratuberculosis* in dam water and sediment. *Applied and Environmental Microbiology,* **71**(9), pp. 5304-5308.

YEW, W.W., SOTGIU, G. and MIGLIORI, G.B., 2011. Update in tuberculosis and nontuberculous mycobacterial disease 2010. *American journal of respiratory and critical care medicine,* **184**(2), pp. 180-185.

Stellenbosch University http://scholar.sun.ac.za

Appendix A

Media, Reagents and Solutions

Stellenbosch University http://scholar.sun.ac.za

Culture media composition

Table A1: Difco™ Middlebrook 7H9 medium supplemented with ADC

Composition	Source
0.47% (w/v) Middlebrook 7H9 0.2% (v/v) glycerol 0.1% (v/v) Tween®-80 10.0% (v/v) ADC* 90.0 ml milliQ water	Beckton-Dickson (BD) Biosciences, USA Merck, Germany Sigma-Aldrich, USA

**See Table A4 for composition.*

Table A2: Difco™ Middlebrook 7H9 medium supplemented with DC

Composition	Source
0.47% (w/v) Middlebrook 7H9 0.2% (v/v) glycerol 0.1% (v/v) Tween®-80 10.0% (v/v) DC* 90.0 ml milliQ water	Beckton-Dickson (BD) Biosciences, USA Merck, Germany Sigma-Aldrich, USA

**See Table A4 for composition.*

Table A3: Difco™ Middlebrook 7H10 base supplemented with ADC

Composition	Source
1.9% (w/v) Middlebrook 7H10 0.5% (v/v) glycerol 10.0% (v/v) ADC* 90.0 ml milliQ water	Beckton-Dickson (BD) Biosciences, USA Merck, Germany

**See Table A4 for composition.*

Stellenbosch University http://scholar.sun.ac.za

Table A4: Supplements of culture media

Supplement	Stock concentration	Sterilization	Storage	Source
Glycerol	100% (v/v)	Filtered	4ºC	Merck, Germany
Tween®-80	20% (v/v)	Filtered	4ºC	Sigma-Aldrich, USA
Albumin-Dextrose-Catalase (ADC)	0.5% (v/v) albumin 0.2% (v/v) glucose 0.015% catalase	Filtered	4ºC	Sigma-Aldrich, USA
Dextrose-Catalase (DC)	0.2% (v/v) glucose 0.015% catalase	Filtered	4ºC	Sigma-Aldrich, USA

Whole cell lysate protein extraction solution

Table A5: Lysis buffer

Composition	Source
50 µl 1M Tris (pH 7.4) 3 µl 20% Tween-80 200 µl 20 µl/1 ml Protease inhibitors set III	Merck, Germany Sigma-Aldrich, USA Merck, Germany

Lysis buffer was filter sterilized and stored at 4ºC.

SDS-PAGE solutions

Table A6: Laemmli sample buffer (4x)

Composition	Source
2.4 ml 1M Tris (pH 6.8)	Merck, Germany
0.8 g SDS	Sigma-Aldrich, USA
4.0 ml 100% glycerol	Merck, Germany
0.01% bromophenol blue	Sigma-Aldrich, USA
1.0 ml β-mercapto-ethanol	Sigma-Aldrich, USA
2.8 ml milliQ water	

Laemmli buffer was stored at room temperature and added to samples in a flow-hood.

In-gel trypsin digest solutions

The following solutions were freshly prepared for each subset of 40 samples.

Table A7: In-gel trypsin digest solutions

Solution	Composition	Source
50% ACN	10 ml 100% ACN	Sigma-Aldrich, USA
	10 ml milliQ water	
25 mM ABC	0.08 g ABC	Sigma-Aldrich, USA
	40 ml milliQ water	
50 mM ABC	0.12 g ABC	Sigma-Aldrich, USA
	30 ml milliQ water	Sigma-Aldrich, USA
10 mM DTT	0.031 g DTT	Sigma-Aldrich, USA
	20 ml 25 mM ABC	Sigma-Aldrich, USA
55 mM IAA	0.2 g IAA	Sigma-Aldrich, USA
	20 ml 25mM ABC	Sigma-Aldrich, USA
10 ng/ µl Trypsin	20 µl 1µl/10 µl Trypsin	Promega, USA
	180 µl 50 mM ABC	Sigma-Aldrich, USA

Stage-tip activation solutions

The following solutions were freshly prepared for each subset of 40 samples.

Table A8: Stage-tip activation solutions

Solution	Composition	Source
5% Formic acid (FA)	250 µl 100% Formic acid	Sigma-Aldrich, USA
	4,750 ml milliQ water	
80% ACN, 5% FA	12.0 ml ACN	Sigma-Aldrich, USA
	3.0 ml milliQ water	
	750 µl 100% FA	Sigma-Aldrich, USA

Stellenbosch University http://scholar.sun.ac.za

Appendix B