Comparative genomics of mycobacteria: Identification of novel anti-tubercular drug targets

By

Mari van Wyk

Thesis submitted in fulfilment of the requirements of the degree

Doctor of Health Sciences in Biomedical Technology

Department of Health Sciences

Central University of Technology, Free State

20 June 2021

Promoter: Prof Khajamohiddin Syed

Co-Promoter: Prof Samson Sitheni Mashele

Co-Promoter External: Dr Wanping Chen

DECLARATION

I, **MARI VAN WYK**, hereby certify that the dissertation submitted by me for the degree DOCTOR OF HEALTH SCIENCES IN BIOMEDICAL TECHNOLOGY, is my own independent work; and complies with the Code of Academic Integrity, as well as other relevant policies, procedures, rules and regulations of the Central University of Technology (Free State). I hereby declare that this research project has not been previously submitted before to any university or faculty for the attainment of any qualification. I further waive copyright of the dissertation in favour of the Central University of Technology (Free State).

21 February 2021

MARI VAN WYK

DATE

ABSTRACT

Mycobacteria are aerobic, unicellular, highly pathogenic organisms, responsible for lethal diseases in humans and other animals. One of the most investigated of the Mycobacteria is *Mycobacterium tuberculosis* (*M. tuberculosis*), a deadly pathogen causing tuberculosis in humans, a continuously growing threat to mankind. The development of multidrug-resistant, extensively drug-resistant, and totally drug-resistant strains of *M. tuberculosis*, combined with the scarcity of new drug targets, suggests that new research is needed to explore the biology of *M. tuberculosis* to investigate new therapeutic measures against tuberculosis and set forth novel potential drug targets.

The genome sequence of *M tuberculosis* has been available for a while now on public databases, and a genome-based method can be used to help understand the adaptation/evolution of this pathogen. Comparative genomics provides us with the opportunity to attempt to understand the mechanism of pathogenesis, including pathways and functions shared among species, and also the physiologic differences, which can help address the pathogenic properties of *M. tuberculosis*. Comparing the genome of the pathogenic *M. tuberculosis* to that of a non-pathogenic mycobacterium, e.g. *M. vanbaalenii*, can shed some light on what unique characteristic it is that makes *M. tuberculosis* more virulent and prone to developing resistance. Recognising any such uniqueness that can possibly be a drug target may ultimately lead to the development of a new drug in the fight against tuberculosis.

This study was performed to see if the use of comparative genomics can be used to identify unique genes that can possibly become novel drug targets. The whole genomes of 8 mycobacterial species belonging to six different niches (*M. tuberculosis* H37Rv, *M. africanum* GM041182, *M. bovis* AF2122/97, *M. abscessus* ATCC 19977, *M. avium* 104,

M. leprae TN, *M. marinum*, and *M. vanbalenii*) were downloaded from TB database and compared using BLAST software. Genes unique to each of the species were identified.

The genome comparison of 8 different mycobacterial species belonging to 6 different niches revealed 200 genes unique to *M. tuberculosis* H37Rv. The uniqueness of 200 genes identified for *M. tuberculosis* H37Rv by genome comparison were further investigated by comparing the genes with available amino acid sequences to the genomes of 96 mycobacterial species. Functional classification of these genes was done using genome mapping with each of these genes with 5 upstream and 5 downstream proteins.

The study results indicated that genome comparison is a viable method to identify unique genes that in the quest to identify novel potential drug targets.

DEDICATION

This research is dedicated to my Creator for his unlimited grace and mercy. This is also dedicated to my friends, family and promotors for their unmatched support and insight.

ACKNOWLEDGEMENTS

I would like to extend my sincere and heartfelt gratitude to the following people and organizations. The completion and success of my Doctoral study would not have been possible without any of them.

- My promotor, Prof Khajamohiddin Syed (University of Zululand), for the opportunity to be part of his research team, sharing insight on bioinformatics, constant guidance, mentorship, invaluable constructive criticism, unrivalled support and for allowing me to grow and become an autonomous scientist. Without your incredible patience and timely wisdom and counsel, this thesis would not have been.

- Prof. Samson Setheni Mashele for his mentorship, guidance and support.

- Dr Wanping Chen at the Huazhong Agricultural University, China, for the invitation extended to me to learn Bioinformatic techniques from him, and his hospitality at receiving me at the University.

- Research and Innovation Fund, Central University of Technology, Free State for funding my studies and enabling dreams to be realised.

- My colleagues at the Faculty of Health and Environmental Sciences, especially the Department of Clinical Sciences and Department of Biomedical Technology at the Central University of Technology for their continuous support, assistance and encouragement. I am very blessed, and the list is too long to mention, I am grateful to all of you.

- To my family for always believing in me, being my pillars of support, loving me unconditionally and respecting my decisions. Thank you for nurturing me into who I am today. I am blessed to have all of you.

- To my dearest friends and extended family, for being my support structure. Their continuous faith in me, prayers, moral support and motivation, gave me strength and inspiration during my studies. I am grateful for your love and support.

- To my Creator for His guidance, gracious blessings, loving protection, for instilling in me this hunger for knowledge and putting the right people on my path to help me reach my goals.

RESEARCH OUTPUTS

- **Articles**

1. **Van Wyk M**, Mashele SS, Chen W, Syed K. 2021. Comparative genomics of mycobacteria: Identification of novel anti-tubercular drug targets. (In preparation for submission to *International Journal of Molecular Sciences*).
2. Van Wyk R, **Van Wyk M**, Mashele SM, Nelson DR, Syed K. Comprehensive comparative analysis of cholesterol catabolic genes/proteins in mycobacterial species. *International Journal of Molecular Science*; Vol 20, Issue 5, March 2019. https://doi.org/10.3390/ijms20051032.
3. Matowane RG, Wieteska L, Bamal HD, Kgosiemang I, **Van Wyk M**, Manumane NA, Abdalla SMH, Mashele SS, Gront D; BBA, Syed K. *In silico* analysis of cytochrome P450 monooxygenases in chronic granulomatous infectious *Sporothrix schenckii*: special focus on CYP51. *Proteins and Proteomics*; Vol 1886, pages 166-177, January 2018. https://doi.org/10.1016/j.bbapap.2017.10.003.
4. Parvez M, Qhanya L, Mthakathi T, Kgosiemang I, Bamal H, Pagadala N, Xie T, Yang Y, Chen H, Theron C, Monyaki M, Raselemane S, Salewe V, Mongale B, Matowane R, Abdalla S, Booi W, **Van Wyk M**, Olivier D, Boucher C, Nelson D, Tuszynski J, Blackburn J, Yu J, Mashele S, Chen W, Syed K. Molecular evolutionary dynamics of cytochrome P450 monooxygenases across kingdoms: Special focus on mycobacterial P450s. *Scientific Reports*; Vol 6; 12 Sep 2016. https://doi.org/10.1038/srep33099.

- **Conference presentations (International)**

van Wyk, M., Nelson, DR., Chen, W., Syed, K. *In silico* analysis of cytochrome P450 monooxygenases in the genus *Mycobacterium tuberculosis*. 20th International Conference on Cytochrome P450 Biochemistry, Biophysics and Biotechnology, 27 – 31 August 2017, Düsseldorf, Germany. S01-P08.

- **Conference presentation (Local)**

1. van Wyk, R., **van Wyk, M.,** Mashele, SS., Chen, W., Syed, K. In silico Analysis of Cholesterol Catabolic Genes/Proteins in the Genus Mycobacterium. The South African Annula Pharmacology Conference (SAPHARM-2017). P08.
2. **M van Wyk,** EM Letsimo, LB Qhanya, SS Mashele, **K Syed**. Analysis of cytochrome P450 monooxygenases in skin infectious yeasts *Malassezia globosa* and *M. sympodialis*. International Symposium on Methods for Studying Drug Metabolism and Transport, and African Traditional Medicines (METHODS-2015) from 23-25 November 2015 at St Georges Hotel and Conference Centre, Pretoria, South Africa.
3. WI Booi, P Nkolanyane, L Wieteska, M Parvez, **M van Wyk,** SS Mashele, D Grontand, K Syed. In silico approach to understand the molecular basis of ketoconazole-resistance in chronic granulomatous infectious fungus Sporothrix schenckii. International Symposium on Methods for Studying Drug Metabolism and Transport, and African Traditional Medicines (METHODS-2015) from 23-25 November 2015 at St Georges Hotel and Conference Centre, Pretoria, South Africa.

ABBREVIATIONS

AIDS	Acquired immune deficiency syndrome
ATP	Adenosine triphosphate
BCG	Bacille Calmette-Guérin
DNA	Deoxyribonucleic acid
EC	Enzyme commission
EGO	Eukaryotic gene Orthologs
HIV	Human immunodeficiency virus
MAC	Mycobacterium avium complex
MCAC	Mycobacterium chelonae abscessus complex
MCL	Mycobacterium causing leprosy
MDR	Multi-drug resistant
MTBC	*Mycobacterium tuberculosis* complex
NTM	Non-tuberculosis mycobacterium
PAH	Polycyclic aromatic hydrocarbon
PCR	Polymerase chain reaction
PIM	Phosphatidylinositol-mannosides
SAP	Saprophyte
TB	Tuberculosis

VC	Vinyl chloride
WHO	World Health Organization
XDR	Extremely drug resistant

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Mycobacteria are aerobic, unicellular, highly pathogenic organisms, responsible for lethal diseases in humans and other animals (Van Ingen, 2017). Mycobacteria include both obligate and opportunistic pathogens (Van Ingen, 2017). Mycobacteria are naturally resistant to antibiotics that usually break down cell wall biosynthesis and they can survive even when exposed to acids, alkalis, detergents and complement lysis. This endurance is due to its characteristic cell wall (Chiaradia *et al*., 2017), which is neither Gram positive, nor Gram negative (De Voss *et al.*, 2000).

One of the most investigated of the Mycobacteria is *Mycobacterium tuberculosis* (*M. tuberculosis*), a deadly pathogen causing Tuberculosis (TB) in humans (Cambau *et al*., 2014). *M. tuberculosis* is an acid-fast, aerobic, chemoorganotrophic organism (Van Ingen, 2017). Tuberculosis continues to be an enormous threat to mankind, being a leading cause of death worldwide (WHO Global tuberculosis report, 2016; Quan *et al*., 2017; TB Alliance), in part due to its growing resistance to known antibiotics and treatment regimens (Migliori *et al*., 2012). The development of multidrug-resistant, extensively drug-resistant, and totally drug-resistant strains of *M. tuberculosis*, combined with the scarcity of new drug targets, suggests that new research is needed to set forth novel potential drug targets (Migliori *et al*., 2012). The growing drug resistance has made many of the currently available drugs ineffective, demanding identification of new drug targets, otherwise TB threatens to become incurable (Zhong *et al.*, 2013). Even now, the mechanism of drug resistance is poorly understood (Zhong *et al.*, 2013; Zhang, 2015). Information from the sequencing and analysis of 161 isolates of Mycobacteria revealed that the genetic basis of drug resistance is more complex

than initially thought, providing a strong basis for clarifying unknown drug resistance mechanisms (Zhong *et al.*, 2013; Zhang, 2015).

For the purpose of this study, the genome analysis of eight mycobacterial species from six different categories (ecological niches) (Parvez *et* al., 2016) will be analysed and compared. Following is some basic background on the mycobacterial species chosen for the study:

M. tuberculosis H37Rv, from the category *Mycobacterium tuberculosis* complex (MTBC): *M. tuberculosis*, as mentioned earlier, causes TB, which is a global epidemic and needs immediate attention (WHO Global tuberculosis report, 2016; Quan *et al*., 2017; TB Alliance). Genomic analysis of the organism has shed light on the pathogenic mechanism of *M. tuberculosis* and helped in the formulation of therapeutic strategies (Cole *et al.*, 1998). One third of the world's population is affected by TB (WHO Global tuberculosis report, 2016; Quan *et al*., 2017; TB Alliance) and it is only second to human immunodeficiency virus (HIV) as a leading cause of death by a single agent. Co-occurrence of TB with HIV/AIDS has further worsened the treatment of TB or HIV/AIDS and many patients are dying because of TB, rather than the initial viral infection. (Ghandi *et al*., 2006). The high mortality rate revives the quest to find new drugs that act with different mechanisms, leading to shorter treatment periods, fewer side effects and compatibility with HIV/ Acquired immune deficiency syndrome (AIDS) treatment. TB can involve any organ, but most infections in immune competent individuals are restricted to the lungs. The likelihood of infection progressing to active disease depends on infectious dose and immune competence (Nathavitharana *et al*., 2018).

M. africanum, also from the category MTBC: Infection with *M. africanum* is comparable to *M. tuberculosis* (De Jong *et al.*, 2010 [a]). It's an opportunistic pathogen that affects immune compromised patients and is more likely to progress from infection to active disease in patients with HIV (De Jong *et al.*, 2010 [a]). Despite the similarities with *M. tuberculosis*, they differ in host ranges, geographical prevalence and pathogenesis and therefore it remains necessary to accurately differentiate clinical isolates of species, for epidemiological and public health purposes (Niemann *et al.*, 2004, De Jong *et al*., 2010 [b]).

M. abscessus ATCC 19977 from the category *Mycobacterium chelonae abscessus* complex (MCAC): It is one of the few rapidly growing mycobacteria causing infection in humans, and one of the most difficult of this group to combat (Ripoll *et al.*, 2009; Kasperbauer *et al*., 2015; Sfeir, 2018). Common infection sites for the organism include the skin, soft tissue and lungs (Ripoll *et al.*, 2009). The bacterium is commonly found in contaminated water, soil and dust (Catherinol *et al.*, 2007; Ripoll *et al.*, 2009). It is the most antibiotic resistant of the rapidly growing mycobacteria, making infection serious (Ripoll *et al.*, 2009, Choo *et al*, 2012, Kasperbauer *et al*., 2015). The bacterium is resistant to most disinfectants and biocide and can survive in the most hostile environments, including medical equipment in hospitals (Ripoll *et al.*, 2009; Choo *et al.*, 2012; Kasperbauer *et al*., 2015). However, it is not clear how the genomic composition influences pathogenicity, virulence and response to therapy (Choo *et al.*, 2012, Kasperbauer *et al*., 2015).

M. avium 104, from the category *Mycobacterium avium* complex (MAC): It is a positive, aerobic, non-motile, non-sporulating, mesophilic, rod shaped bacteria that is an opportunistic pathogen (Horan *et al.*, 2006). The organism is most likely environmentally acquired, through contaminated water or soil (https://www.cdc.gov/), *via* ingestion or inhalation. Infection is more common in immunocompromised patients (Wu *et al.*, 2006).

Infection before co-occurrence with AIDS was typically asymptomatic and when symptoms similar to TB were observed, it was usually in patients with compromised pulmonary function (https://www.cdc.gov/).

M. leprae TN, from the category Mycobacterium causing leprosy (MCL): A lethal pathogen causing leprosy. It is a gram positive, intracellular, pleomorphic, acid-fast bacillus in a waxy covering (Monot *et al.*, 2009). Manifestations of leprosy, also called Hansen's disease, depend on the patient's immunity (Van Ingen, 2017). Spread from person to person is through respiratory droplets (https://www.cdc.gov/). The organism has never been grown on culture media or cell culture but in mouse foot pads. Infection mainly involves the skin, nerves and mucous membranes (Van Ingen, 2017).

M. marinum, from the category Non-tuberculosis mycobacterium (NTM): Infection mainly occurs in marine species (Stinear *et al.*, 2008). It is found commonly in bodies of salt or fresh water (Parrish *et al*., 2011). The bacterium infects the exterior of the skin and damaged soft tissue of aquatic species (Arend *et al.*, 2002, Van Ingen, 2017). In humans it is an opportunistic pathogen which causes granulomatous skin lesions (Stinear *et al.*, 2008), and exposure is commonly through breaks in the skin (Parrish *et* al., 2011; Van Ingen, 2017).

M. vanbaalenii PYR-1, from the saprophyte (SAP) category: It is an eco-friendly, non-pathogenic strain of the *Mycobacterium*. The organism can use polycyclic aromatic hydrocarbons (PAH), and as such has potential in the biodegradation industry. Its genome has recently been sequenced, allowing some insight into molecular basis for the degradation of PAH's (Kim *et al.*, 2008).

1.1. Study rationale

Tuberculosis is and will remain a major threat to global health and exploring the biology of, and new therapeutic measures against, *M. tuberculosis* has become a priority. The genome sequence of M tuberculosis has been available for a while now, but the biology is poorly understood (McGuire et al., 2012). A genome-based method can be used to help understand the adaptation/evolution of the pathogen (Prasanna et al., 2013). Genes responsible for causing disease have been identified but understanding the mechanism of action of pathogenesis is an active area of research (Prasanna et al., 2013). Comparative genomics provides us with the opportunity to attempt to understand the mechanism of pathogenesis, including pathways and functions shared among species (Malhotra et al., 2017), and also the physiologic differences, which can help address the pathogenic properties of M tuberculosis (Long et al., 2012; McGuire et al., 2012). Comparing the genome of the pathogenic M tuberculosis to that of non-pathogenic mycobacterium, e.g. M. vanbaalenii, can provide some useful insight into the evolutionary relationship between these bacteria (Prasanna et al., 2013; Malhotra et al., 2017) and shed some light on what unique characteristic it is that makes M. tuberculosis more virulent and prone to developing resistance. Recognising any such uniqueness that can possibly be a drug target may ultimately lead to the development of a new drug in the fight against TB (McGuire et al., 2012; Malhotra et al., 2017).

1.2. Problem statement

In today's age of advanced scientific research, pathogenic mechanisms of Mycobacteria and development of effective drug targets seem to be poorly understood. A key trait of Mycobacteria is its innate resistance to most antibiotics, making development of treatment regimens very challenging. Development of multi drug-resistance (Migliori *et al*., 2012) and lack of understanding in the mechanism of pathogenicity has rekindled the need for research

that will provide new insight into the mechanism of pathogenicity and the need to discover novel drug targets.

Genomes sequencing of mycobacterial species from different categories are readily available. The available genome sequencing data presents the opportunity to use the data to do a comparative study and possibly identify unique proteins that may shed some light on their individual mechanisms of pathogenicity and possible discovery of new drug targets.

1.3. Aims and objectives

1.3.1. Aim of the study

The aim of this study includes the identification of unique genes/proteins in *M. tuberculosis* H37Rv *via* comparative analysis of mycobacterial species belonging to six different ecological niches.

1.3.2. Objectives of the study

- Comparative analysis of genes among mycobacterial species belonging to six different ecological niches
- Identification of common and unique proteins in Mycobacterial species belonging to each category especially in *M. tuberculosis* H37Rv
- Functional classification of unique proteins identified in *M. tuberculosis* H37Rv.

1.3.3 Thesis overview

This thesis is divided into 5 chapters. Chapter 1 provides an introduction and brief overview of the work, problem statement, aims and objectives of the study. Chapter 2 consists of a literature review. Chapter 3 details the methodology used in this study including the species

used in the study, genome comparison and gene mapping of identified genes of *M. tuberculosis*. The results of the methods outlined in Chapter 3 are presented in Chapter 4 and an overall conclusion is presented in Chapter 5. Due to the nature of data presented in the thesis, two tables (Tables 13 and 14) are provided as attachment files.

1.4. References

Van Ingen (2017) Mycobacteria. **Infectious Diseases (4th edition)**, 2:1645-59.e2

Chiaradia L, Lefebvre C, Parra J, Marcoux J, Burlet-Schiltz O, Etienne G, et al (2017) Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane. **Scientific Reports**, 7: 10.1038/s41598-017-12718-4

De Voss, JJ, Rutter, K, Schroeder, BG, Su, H, Zhu, YQ, Barry, CE (2000) The Salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. **Proceedings of the National Academy of Sciences**, 97: 1252 – 1257

Cambau E, Drancourt M (2014) Steps towards the discovery of *Mycobacterium tuberculosis* by Robert Koch, 1882. **Clinical Microbiology and Infection**, 20(3): 196-201

WHO tuberculosis report (2016) Available at: http://www.who.int/tb/publications/global_report/en/

Quan D, Nagalingam G., Payne R. & Triccas J A (2017) New tuberculosis drug leads from naturally occurring compounds. International Journal of Infectious Diseases, 56: 212 – 20

TB Alliance. Available at: https://www.tballiance.org/why-new-tb-drugs/global-pandemic

Migliori G B, Centis R, Ambrosio L, Spanevello A, Borroni E, Cirillo D M, et al (2012) totally drug-resistant and extremely drug-resistant tuberculosis: the same disease? **Clinical infectious diseases**, 54(9): 1379 – 80

Zhong H, Li D, Zhao L, Fleming J, Lin N, Wong T, Liu Z et al. (2013) Genome sequencing of 161 Mycobacterium tuberculosis isolates form China identifies genes and intergenic regions associated with drug resistance. **Nature genetics**, 45: 1255 – 60

Zhang Yew W W (2015) Mechanism of drug resistance in Mycobacterium tuberculosis: update 2015. **PubMed**, 19(11): 1276 – 89

Parvez M, Qhanya L B, Mthakathi N T, Kgosiemand I K, Bamal H D, Pagadala N S, et al (2016) Molecular evolutionary dynamics of cytochrome P450 monooxygenases across kingdoms: Special focus on mycobacterial P450s. **Scientific Reports**, 6: 33099

Cole, ST, Bosch, R, Parkhill, J, Garnier, T, Churcher, C, Harris, D, Gordon, SV et al. (1998) Deciphering the bilogy of Mycobacterium tuberculosis from the complete genome sequence. **Nature**, 393: 537 – 544

Gandhi N R, Moll A., Sturm A W, Pawinski R, Govender T, Lalloo U, et al (2006) Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. **The Lancet**, 368(9547): 1575 – 80

Nathavitharana R, Pai M (2018) Extrapulmonary Tuberculosis: New Diagnostics and New Policies. **Let's talk TB**, 3rd edition, 52 – 6

De Jong B C, Antonio M, Gagneux S (2010) Mycobacterium africanum – review of an important cause of human tuberculosis in West Africa. **PLoS neglected tropical diseases**, 4(9): e744 (a)

De Jong B C, Adetifa I, Walther B, Hill P C, Antonio M, Ota M, Adegbola R A (2010) Differences between TB cases infected with *M. africanum*, West-African type 2, relative to

Euro-American *M. tuberculosis*- an update. **FEMS Immunology Medical Microbiology**, 58(1): 102 – 5 (b)

Niemann S, Kubica T, Bonge F C, Adjei O, Browne, E N et al. (2004) The species Mycobacterium africanum in the light of new molecular markers. **Journal of Clinical Microbiology**, 42(9), 3958 – 62

Sfeir M, Walsh M, Rosa R, Aragon L, Liu S Y, Cleary T, et al (2018) *Mycobacterium abscessus* Complex Infections: A Retrospective Cohort Study. **Open Forum Infectious Disease**, 5(2): ofy022

Kasperbauer S H, De Groote M A (2015) The treatment of rapidly growing mycobacterial infections. **Clinical Chest Medicine**, 36(1): 67 – 78

Ripoll F, Pasek S, Schenowits C, Dossat C, Barbe V, Rottman M et al. (2009) Non-mycobacterial virulence genes in the genome of the emerging pathogen Mycobacterium abscessus. **PLoS ONE**, 4: e5660

Catherinol E, Clarrisou J, Etienne G, Ripoll F, Emile J F, Daffe M et al. (2007) Hypervirulence of a rough variant of the Mycobacterium abscessus type strain. **Infection and Immunity**, 75, 1055 – 58

Choo S W, Wong Y L, Yusoff A M, Leong M L, Wong G J, Ong C S et al. (2012) Genome sequence of the Mycobacterium abscessus strain M93. **Journal of Bacteriology**, 194(12): 3278

Horan K L, Freeman R, Wergel K, Semiet M, Pfaller S, Covert T C et al. (2006) Isolation of the genome sequence strain *Mycobacterium avium* 104 from multiple patients over a 17 year period. **Journal of Clinical Microbiology**, 44: 783 – 9

Wu C, Glasne, J, Collin, M, Nase, S, Talaat A (2006) Whole Genome Plasticity among Mycobacterium avium Subspecies: Insights from Comparative Genomic Hybridizations. **Journal of Bacteriology**, 711-23

Monot M, Honorã N, Garnier T, Zidone N, Sherafi D, Paniz-Mondolfi A et al. (2009) Comparative genomic and phylogeographic analysis of Mycobacterium leprae. **Nature genetics**, 41: 1282 – 89

Stinear T P, Seemann T, Harrison P F, Jenkin G A, Darvies J K, Johnson P D R et al. (2008) Insights from the complete genome sequence of Mycobacterium marinum on the evolution of Mycobacterium tuberculosis. **Genome research**, 18: 729 – 41

Parrish N, Luethke R, Dionne K, Carroll K, Riedel S (2011) Case of *Mycobacterium marinum* Infection with Unusual Patterns of Susceptibility to Commonly Used Antibiotics. **Journal of Clinical Microbiology**, 2056 – 8

Arend S M, Van Meijgarden K E, De Boer K, De Palou E C, Van Soolingen D, Okkenhoff T H M et al. (2002) Tuberculin skin testing and in vitro Tcell responses to ESAT-6 and culture filtrate protein 10 after infection with Mycobacterium marinum or Mycobacterium kansasii. **Journal of Infectious diseases**, (186): 1797 – 807

Kim S J, Kweon O, Jones R C, Edmondson R D & Cerniglia C E (2008) Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. **Biodegradation**, 19, 859 – 881

Long Q, Zhou Q, Ji L, Wu J, Wang W, Xie J (2012) *Mycobacterium smegmatis* Genomic Characteristics Associated With its Saprophyte Lifestyle. **Journal of Cellular Biochemistry**, 113: 3051 – 55

McGuire A M, Weiner B, Park S T, Wapinski I, Raman S, Dolganov G, et al (2012) Comparative analysis of Mycobacterium and related Actinomycetes yields insight into the evolution of *Mycobacterium tuberculosis* pathogenesis. **BMC Genomics**, 13: 1 - 20

Prasanna A N, Mehra S (2013) Comparative Phylogenomics of Pathogenic and Non-Pathogenic Mycobacterium. **PLOS ONE**, 8(8): e71248

Malhotra S, Vedithi S C, Blundell T L (2017) Decoding the similarities and differences among mycobacterial species. **PLOS Neglected Tropical Diseases**, 11(8): e0005883

CHAPTER 2

LITERATURE REVIEW

2.1. Mycobacteria

Mycobacteria consists of a group of over one hundred species, mostly environmental, ranging from the saprophytic organisms that rarely causes infection/disease in humans to major human pathogens (Ripoll *et al*., 2009; Gillespie & Bamford, 2012). The organisms are further divided into slow growing mycobacteria and the rapidly growing mycobacteria (Ripoll *et al.*, 2009). Mycobacteria are generally intracellular pathogens, occupying macrophages as a preferred niche, and are transmitted mainly *via* the respiratory route (Vissa & Brennan, 2001). There are some mycobacterial species (e.g. NTM) that may cause localised or disseminated disease in immune compromised patients, for example infection of prosthetic devices (Gillespie & Bamford, 2012).

Mycobacteria possess a lipid rich cell wall (Gillespie & Bamford, 2012). Several antibodies cross-react among the different species, implying similar protein composition amongst the species, as well as a similar cell wall structure (Vissa & Brennan, 2001). An extensive study of the cell wall structure has revealed a few properties that are common in all mycobacteria (Figure 1) (Vissa & Brennan, 2001):

- An innermost electron-dense layer that makes up the rigid layer just beyond the plasma membrane, consisting of peptidoglycan and arabinogalactin
- An intermediate electron-transparent layer, i.e. the mycolate layer
- An outermost electron-dense layer composed of a variety of lipoglycans, free polysaccharides, glycolipids and phospholipids

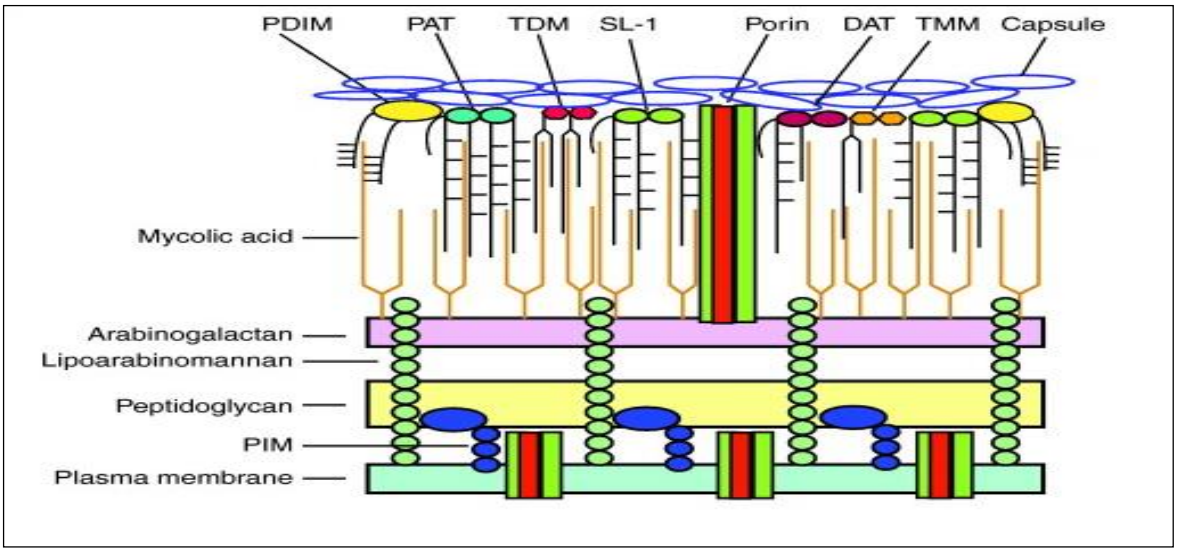


Figure 1 Schematic representation of mycobacterial cell envelope (taken from Ouellet et al., 2011). *The cell envelope from the inside consists of the plasma membrane, the cell wall made up of three covalently linked macromolecules – peptidoglycan, arabinogalactan and mycolic acids – and noncovalently linked lipids and proteins, and the capsule composed of polysaccharides, proteins and lipids.*

The core of all mycobacterial cell walls consists of peptidoglycan, covalently attached by a linker unit to arabinogalactan, differentiated by furanose sugars (Figure 1) (Vissa & Brennan, 2001). The mycolic acids are mainly responsible for the mycobacteria's characteristic ability to proliferate in lipids displayed by mycobacteria (Vissa & Brennan, 2001). Some factors that are also present in the cell wall include PIM's, cord factor/trehalose dimycolate, sulfolipids and other proteins, but the arrangement (Figure 1) and purpose of the arrangement is poorly understood (Vissa & Brennan, 2001). All Mycobacteria contain coding sequences coding for lipoproteins, and the variation in the lipoprotein content reflect the different host ranges of each (Stinear et al., 2008).

Mycobacteria have initially been assigned to different species based on morphological, biochemical, and phenotypic characteristics but later challenges arising from taxonomic classification have led to the recognition of imperfections in this method and lead to the exploration of genome investigations (Mostowy et al., 2004, Vasconcellos et al., 2010).

Identification of differences at a deoxyribonucleic acid (DNA) sequence level has improved the capacity for performing and understanding molecular epidemiology, phylogenetic structuring and species classification (Vasconcellos et al., 2010). Comparative genomics is useful in the identification of similar pathways in the different species, as well as pathways that are unique to a species (Vissa & Brennan, 2001). Genetic exploration and sequence analysis have superseded the previous methods that may be time consuming, troublesome, imprecise, have decreased reproducibility and may not give clear-cut results in every case (Vasconcellos et al., 2010). The other downside of the previous methods was that it wasn't always possible to do in every clinical microbiology laboratory (Vasconcellos et al., 2010).

Some of the more recently applied methods, like polymerase chain reaction (PCR) and sequence analysis have been useful in identifying single-nucleotide polymorphisms and differences in chromosomal regions, such as insertions, deletions or mutations (Vasconcellos et al., 2010). These genetic differences are usually lineage-, strain- or species specific (Vasconcellos et al., 2010).

2.2. Mycobacterial species

A large number of mycobacterial species can be found in nature. Most of these mycobacterial species can be grouped into six different categories based on their ecological niches (Parvez et al., 2016) (Table 1).

Table 1 Classification of mycobacterial species based on different ecological niches/pathogenic profiles

Ecological niche	Mycobacterial species
Mycobacterium tuberculosis complex (MTBC)	*Mycobacterium africanum* GM41182
	Mycobacterium tuberculosis F11
	Mycobacterium tuberculosis H37Ra
	Mycobacterium tuberculosis H37Rv
	Mycobacterium tuberculosis Haarlem
	Mycobacterium tuberculosis KZN1435
	Mycobacterium tuberculosis KSN 605
	Mycobacterium tuberculosis KZN 4207
	Mycobacterium tuberculosis RGTB327
	Mycobacterium tuberculosis CDC 1551
	Mycobacterium tuberculosis strains CCDC5079
	Mycobacterium tuberculosis 7199 -99
	Mycobacterium tuberculosis Beijing/NITR203
	Mycobacterium tuberculosis CAS/NITR204
	Mycobacterium tuberculosis EAI5
	Mycobacterium tuberculosis EAI5/NITR206
	Mycobacterium tuberculosis Erdman 35801
	Mycobacterium tuberculosis UT205
	Mycobacterium canettii
	Mycobacterium bovis AF 2122/97
	Mycobacterium bovis BCG Pasteur 1137P2
	Mycobacterium bovis BCG Korea 1168P
	Mycobacterium bovis BCG Mexico
	Mycobacterium bovis BCG Tokyo 172
Mycobacterium chelonae-abscessus complex (MAC)	*Mycobacterium abscessus* ATCC 19977
	Mycobacterium abscessus subsp. bolletii 50594
Mycobacterium avium complex (MAC)	*Mycobacterium avium* 104
	Mycobacterium avium subsp. paratuberculosis K10
	Mycobacterium avium subsp. paratuberculosis MAP4
	Mycobacterium intracellulare ATCC 13950
	Mycobacterium intracellulare MOTT -02
	Mycobacterium intracellulare MOTT -64
	Mycobacterium indicus pranii MTCC 9506
Mycobacterium causing Leprosy (MCL)	*Mycobacterium leprae* Br4923
	Mycobacterium leprae TN
Nontuberculosis mycobacteria (NTM)	*Mycobacterium* sp. JDM601
	Mycobacterium liflandii 128FXT
	Mycobacterium ulcerans Agy99
	Mycobacterium marinum
	Mycobacterium kansasii ATCC 12478
Saprophytes (SAP)	*Mycobacterium* sp. JLS
	Mycobacterium sp. KMS

	Mycobacterium sp. MCS
	Mycobacterium vanbaalenii PYR-1
	Mycobacterium smegmatis MC2 155
	Mycobacterium smegmatis JS623
	Mycobacterium chubuense NBB4
	Mycobacterium gilvum PYR-GCK
	Mycobacterium gilvum Spyr1
	Mycobacterium rhodesiae NBB3
	Mycobacterium neoaurum VKM Ac- 1815D

Information on the different mycobacterial species belonging to the different categories are presented below:

2.2.1. *Mycobacterium tuberculosis* complex

2.2.1.1. *Mycobacterium africanum* GM041182

M. africanum was first isolated from a non-infected, HIV positive patient in 2004 with a positive smear for pulmonary tuberculosis (Bentley *et al*., 2012). In contrast to other species in this category, *M. africanum* can posess all the biochemical properties normally used to differentiate between the different species, especially between *M. tuberculosis* and *M. bovis* (Mostowy *et al*., 2004). Due to this trait, *M. africanum* is an unspecified classification and is generally applied to isolates in the category that differs in their phenotypic presentation and genomic content (Mostowy *et al*., 2004).

M. africanum is also known for causing TB in humans, especially in the West African countries (Vasconcellos *et al*., 2010; Bentley *et al*., 2012). Symptoms of infection are very similar to those of infection caused by *M. tuberculosis*, and predisposing factors include immune compromised states, such as patients with HIV (Galagan *et al*., 2010). Although the host range of *M. africanum* is believed to be human, the organism is isolated with much lower frequency than *M. tuberculosis* (Mostowy *et al*., 2004). *M. africanum* is found commonly in Northern, Western and Central Africa (Washington *et al*., 2006; Galagan *et al*.,

2010; Bentley *et al*., 2012). *M. africanum* was subdivided into two major subgroups by Mostowy *et al*. (2004) based on geographic origin and biochemical properties:

- Subtype I, which originates from West Africa and has *M. bovis* like properties, and
- Subtype II, which originates from East Africa and has *M. tuberculosis* like properties.

The challenge is that even genomic analysis has been unable to distinguish between subtype II and the modern *M. tuberculosis*, suggesting the *M. africanum* may be a phenotypically atypical strain of *M. tuberculosis*, and previously classified strains of *M. africanum* subtype II may actually be *M. tuberculosis* (Mostowy *et al*., 2004, Vasconcellos *et al*., 2010). There were however genomic deletions that allowed for the differentiation of subtype I from *M. tuberculosis* (Mostowy *et al*., 2004). At the genome level all the species of the MTBC exhibit near identical genes and are differentiated only by large sequence polymorphism (Mostowy *et al*., 2004).

Vasconcellos *et al*. (2010) concluded that although Subtype II is hardly distinguishable from *M. tuberculosis*, Subtype I is easily distinguishable and can be divided into two genealogical clades based on the genome-level sequence differences:

- *M. africanum* West African-1 (clade 1), which possess a long sequence polymorphism RD713
- *M. africanum West African-2 (clade 2), which carries the defining long sequence polymorphism RD701 and RD702*

2.2.1.2. *Mycobacterium tuberculosis*

M. tuberculosis is a pathogenic, aerobic, chemoorganotrophic, rod-shaped, non-motile bacterium, that causes TB in humans (Cole *et al*., 1998; Vasconcellos *et al*., 2010). *M. tuberculosis* is typically a slow-growing, dormant, intracellular pathogen with a complex cell envelope and genetic homogenecity (Cole *et al*., 1998). The exact mechanism by which *M. tuberculosis* causes disease was poorly understood until some significant advances were made using investigation of the genome sequence (Zheng *et al*., 2008).

M. tuberculosis, although primarily affecting the lungs, can affect almost every organ in the body, producing symptoms imitating both inflammatory and malignant disease (Gillespie & Bamford, 2012). The presence of *M. tuberculosis* in a clinical specimen is almost always associated with infection (Washington *et al*., 2006). The need for prompt and effective anti-mycobacterial drugs has become a necessity since the emergence of multi-drug resistant (MDR) strains of *M. tuberculosis* (Washington *et al*., 2006). The organism has evolved to become highly successful in evading host immunity and known antibiotic treatment (Jena *et al*., 2013).

M. tuberculosis F11

This strain of *M. tuberculosis* is found in large clusters on at least four different continents and 25 countries in the world and is a major contributing factor to the TB epidemic in South Africa (Viktor *et al*., 2004). This strain was isolated in the 1990's in a TB epidemic in the Western Cape, South Africa (Viktor *et al*., 2004).

M. tuberculosis H37Ra

M. tuberculosis H37Ra is an avirulant strain derived from the H37 strain (Steenken *et al*., 1934). The strain has lost some of its virulence characteristics, including the ability to survive within the macrophage or in hypoxic conditions (Steenken *et al*., 1934).

M. tuberculosis H37Rv

The genome sequence of *M. tuberculosis* H37Rv is probably the best characterised strain and has been investigated and analysed in an attempt to improve our understanding of its biology and to help in the discovery of new prophylactic and therapeutic interventions (Cole *et al*., 1998, Camus *et al*., 2002).

M. tuberculosis Haarlem

The first isolate of this strain was found in Haarlem in the Netherlands, but today can be found in many places in the world (Kremer *et al*., 1999). This strain has been found to be MDR and infection in immune competent and non-hospitalised patients are reported more frequently (Kremer *et al*., 1999).

M. tuberculosis KZN1435

This strain of *M. tuberculosis* was first isolated from patients in Kwazulu Natal, South Africa (Ioerger *et al*., 2009). This particular strain was found to be MDR (resistant to both isoniazid and rifampicin) (Ioerger *et al*., 2009).

M. tuberculosis KZN605

This strain, also isolated in Kwazulu Natal, is extremely drug resistant (XDR) and is very difficult to treat (Ioerger *et al*., 2009). Treatment of infection from this strain is very challenging as it is resistant the 1st line anti-TB drugs, at least one fluoroquinolone, and at least one 2nd line injectable drug (Ioerger *et al*., 2009). A high mortality rate has been reported for patients infected with XDR-TB in the Kwazulu Natal area (Ioerger *et al*., 2009).

M. tuberculosis KZN4207

This strain from Kwazulu Natal was found to be fully sensitive to anti-TB drugs (Ioerger *et al*., 2009).

M. tuberculosis RGTB327

This strain was first isolated from sputum samples in Kerala, a state in South India (Madhavilatha *et al*., 2012). The strain was found as part of a repository field strain that formed part of a drug screening program (Madhavilatha *et al*., 2012).

M. tuberculosis CDC1551

This strain, also known as the "Oshkosh" was isolated in the Kentucky, Tennessee region in the United States in 1996 (Manca et al., 1999). This strain was found to be highly infectious and, although a human pathogen, studies also found that it was highly virulent in mice (Manca et al., 1999). Fortunately, this strain hasn't caused widespread infection in humans as it is pan-drug sensitive (Manca et al., 1999).

M. tuberculosis CCDC5079

Zhang et al (2011) did a study comparing the genomes of M. tuberculosis strains CCDC5079 and strain CCDC5080. Strains were collected from patients with secondary pulmonary tuberculosis in the Fujian Province in China in 2004 (Zhang et al., 2011). Both these strains belong to the Beijing family (Zhang et al., 2011). M. tuberculosis strain CCDC5079 is sensitive to all the 1st line drugs used for the treatment of TB (Zhang et al., 2011). With the genome analysis it was found that strain CCDC5079 featured high GC-repetitive DNA, which may be responsible for the difficulty experienced in sequencing this strain (Zhang *et al*., 2011). Genome annotation revealed three kinds of virulence factors present (Zhang *et al*., 2011).

- One kind is involved in *in-vivo* growth, playing a critical role in lipid metabolism and signal pathways
- The second kind includes cell envelope- and cell wall- associated genes, which assists in the pathogen's ability to escape or invade the macrophage

- The third kind is associated to the PE/PPE family with related functions.

***M. tuberculosis* 7199-99**

Researchers estimated that this strain was responsible for a TB outbreak in Hamburg, Northern Germany, between 1993 and 1997 (Roetzer *et al*., 2013).

***M. tuberculosis* Beijing/NITR203**

This strain of *M. tuberculosis* was first isolated from the Tamil Nadu state in the south of India (Narayanan & Deshpande, 2013). Although infection of TB with this strain occurs worldwide, the strain appears to be endemic in Asian countries (Coker *et al*., 2014). Strains of *M. tuberculosis* belonging to the Beijing family are found to be resistant to most of the potent anti-TB drugs (Coker *et al*., 2014). Strains isolated from a TB outbreak in Thailand were resistant isoniazid, rifampicin, ethambutol and streptomycin (Coker *et al*., 2014).

***M. tuberculosis* CAS/NITR204**

This strain was also isolated from the Tamil Nadu state in the south of India (Narayanan & Deshpande, 2013). This strain is from the Central Asian lineage circulating in the Indian population (Al Rashdi *et al*., 2014).

***M. tuberculosis* EAI5**

A study was conducted on TB outbreaks in Mumbai and neighbouring rural areas and it was found that four strains of *M. tuberculosis* were responsible, including the East African Indian (EAI5) strains (Al Rashdi *et al*., 2014).

***M. tuberculosis* EAI5/NITR206**

This strain was isolated from the Tamil Nadu state in the south of India (Narayanan & Deshpande, 2013). This is an East African Indian strain of the organism (Al Rashdi *et al*., 2014).

***M. tuberculosis* Erdman=ATCC35801**

This strain was first isolated in Rochester, Minnesota at the Mayo Clinic from human sputum (Miyoshi-Aliyama *et al*., 2012). This strain is highly virulent and as such is used frequently in the laboratory for virulence and immunization studies (Miyoshi-Akiyama *et al*., 2012).

***M. tuberculosis* UT205**

This virulent Colombian strain was isolated from a 33-year-old male who was diagnosed with TB shortly before isolation (Isaza *et al*., 2012).

2.2.1.3 *Mycobacterium canetii*

The organism was first isolated in 1969 by G Canetti form a 20-year-old French farmer suffering from pulmonary tuberculosis (www.uniprot.org, 2015). It seems that the organism preferentially affects children and foreigners in the Horn of Africa, with the Republic of Djibouti being a common place for TB caused by *M. canettii* (www.uniprot.org, 2015). *M canettii* is intrinsically resistant to pyrazinimide and pyrazinoic acid, drugs used in the treatment of susceptible TB (Feuerriegel *et al*., 2013).

2.2.1.4 *Mycobacterium bovis*

M. bovis belongs to the MTBC (Centre for Food Security & Public Health, 2009; Vasconcellos *et al*., 2010). Identification of *M. bovis*, when compared to other species in this category, is based on clear cut differences in phenotypic characteristic and biochemical properties (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html). It is a gram-positive, chemoorganotrophic, non-spore forming, non-motile, slightly curved, aerobic, slow-growing bacteria (Centre for Food Security & Public Health, 2009). The organism can survive for long periods of time in the environment, especially in cold, dark and moist conditions (Centre for Food Security & Public Health, 2009).

M. bovis causes bovine and human TB and avirulent strains of the organism are used for the Bacille Calmette-Guérin (BCG) vaccine (Garnier *et al*., 2003; Washington *et al*., 2006). Although the organism mainly causes TB in cattle, it may infect other animals, including dogs, cats, swine, rabbits, antelope, lions, buffalo and sometimes certain birds of prey (Vissa & Brennan, 2001; Washington *et al*., 2006; Centre for Food Security & Public Health, 2009). Bovine TB is still common in less developed countries, causing severe economic losses from livestock deaths, chronic illness and trade restrictions. The organism may also be a threat to endangered animals (Centre for Food Security & Public Health, 2009).

Transmission from animals to animals occur, especially among cattle as the infected cattle shed the organism in respiratory secretions, faeces, milk or even urine, especially in the late stages of infections (Centre for Food Security & Public Health, 2009). The organism can be transmitted from animals to humans (Washington *et al*., 2006) by eating or drinking contaminated or unpasteurised dairy products or from direct contact with a wound (slaughtering or hunting), and inhalation, although transmission through air is very rare from animals to humans (CDC: Division of Tuberculosis Elimination; Centre for Food Security & Public Health, 2009). Human to human spread is very uncommon in immune-competent individuals but has been reported in alcoholics and individuals with HIV (Centre for Food Security & Public Health, 2009). Individuals at risk include people who work with animals susceptible to infection or products of such animals, including occupations such as ranching, dairy farming, butchery, and hunting (CDC: Division of Tuberculosis Elimination).

Human bovine pulmonary TB resembles infection caused by *M. tuberculosis*, with symptoms such as fever, night sweats and weight loss (Washington *et al*., 2006). Other symptoms may occur, but they depend on the affected area. Not all *M. bovis* infections

progress to active disease, so patients may be asymptomatic (CDC: Division of Tuberculosis Elimination). Symptoms usually take months to develop and some infections can remain dormant for years before it is reactivated by old age, stress or low immunity (Centre for Food Security & Public Health, 2009). Infection may involve lymph nodes, skin, bones and joints, genitourinary system, meninges, or the respiratory system (Centre for Food Security & Public Health, 2009). In cattle the disease can be acute and rapidly progressive, with symptoms of emaciation, low grade fluctuating fever, weakness and lack of appetite, and if there is pulmonary involvement, a moist cough, dyspnea and tachycardia may also be present (Centre for Food Security & Public Health, 2009).

Treatment is similar to infection with *M. tuberculosis* (Washington *et al*., 2006) and if left untreated the patient can die (CDC: Division of Tuberculosis Elimination). *M. bovis* has an innate resistance to the first line anti-TB drug pyrazinimide (Vasconcellos *et al*., 2010). Due to this trait it is important to correctly identify the causative organism in a case of TB, not only to ensure correct treatment but also to properly assist in the collection of epidemiologic information, and if necessary, to implement public health interventions (Vasconcellos *et al*., 2010).

In a study by He et al. (2003) distinctly different genomic regions were found between the virulent *M. tuberculosis*, *M. bovis* and the attenuated BCG. Deletion of large chromosomal regions has led to the attenuation of *M. bovis* resulting in the development of the BCG vaccine, the only known vaccine against TB (Garnier et al., 2003; Ripoll et al., 2009). The genome sequence of the organism strengthened the theory that *M. bovis* has evolved from a progenitor of the MTBC as a clone with a distinct host preference (Garnier et al., 2003).

***M bovis* AF2122/97**

M. bovis AF2122/97 is a highly virulent strain isolated in Great Britain in 1997 from a diseased cow suffering from caseous lesions on the lungs, and bronchomediastinal lymph nodes (Garnier *et al*., 2003).

***M. bovis* BCG Pasteur 1173P2**

This strain is a live vaccine strain for TB (Keller *et al*., 2008).

***M. bovis* BCG Korea 1168P**

M, bovis BCG Korea 1168P is a derivative of the Pasteur 1139P strain and is used for the production of the Korean TB vaccine (Lee *et al*., 2014).

***M. bovis* BCG Mexico**

The BCG substrains are subdivided into four groups based on their regions of difference and tandem duplication markers (Orduña *et al*., 2011). *M. bovis* BCG Mexico used to be one of the most widely used strains for vaccine production in Mexico, and it may be used again for the development of a new recombinant BCG vaccine (Orduña *et al*., 2011). This strain presented with a biochemical profile more similar to *M. bovis* than any other BCG strain (Orduña *et al*., 2011). In 1970 the Mexico strain was replaced by the Danish 1331 strain for vaccine production in Mexico, however, vaccine production in Mexico stopped in 1998 and the country has relied on imported vaccines since then (Orduña *et al.*, 2011).

***M. bovis* Tokyo 172**

This strain is used to produce BCG vaccine in Japan (Seki *et al*., 2009).

2.2.2 *Mycobacterium chelonae-abscessus* complex

2.2.2.1 *Mycobacterium abscessus*

M. abscessus is one of the most clinically important of the rapidly growing mycobacteria (Washington *et al*., 2006; Ripoll *et al*., 2009). It is a common water and soil contaminant

(Ripoll *et al*., 2009). *M. abscessus* commonly causes skin infection, often evolving into draining subcutaneous abscesses, particularly in patients with compromised immunity (Washington *et al*., 2006). The organism, although present in many water sources, is also an important causative agent of pulmonary disease in patients with chronic respiratory disease (Washington *et al*., 2006; Ripoll *et al*., 2009; Choo *et al*., 2012). Other predisposing factors for infection include organ transplantation, rheumatoid arthritis, trauma and invasive medical procedures (Washington *et al*., 2006). Lung infections with *M. abscessus* in patients with cystic fibrosis are becoming an increasingly frequent phenomenon (Ripoll *et al*., 2009).

Pulmonary infection with *M. abscessus* is considered very serious with a much higher fatality rate than any other rapidly growing mycobacterium (Washington *et al*., 2006, Ripoll *et al*., 2009). Although considered an environmental contaminant, the organism is frequently found on hospital equipment and has been associated with nosocomial and community acquired infections, ranging from superficial skin and soft tissue infection to disseminated disease (Choo *et al*., 2012). The frequent observation of the organism in healthcare-associated disease is because the organism is resistant to most disinfectants and biocides and also has the ability to survive in harsh environments (Ripoll *et al*., 2009).

The rapidly growing mycobacteria have shown to be susceptible to amikacin, cefotoxin, doxycycline, and erythromycin when tested *in vitro* (Washington *et al*., 2006). More recently, imipenem-cilastatin, amoxicillin-clavulanate and ciprofloxacin have also shown some activity against these mycobacteria when tested *in vitro* (Washington *et al*., 2006). Despite the available antimicrobial agents, *M. abscessus* is notoriously known as one of the most drug-resistant of the rapidly growing mycobacteria, and therefore treatment of infection may be difficult (Ripoll *et al*., 2009; Choo *et al*., 2012). This drug resistance is ascribed not only to the weak permeability of the cell wall, but also to the presence of drug-modifying enzymes

like Ambler class A beta-lactamase and rifampicin ADP-ribosyl transferase, to name just a few (Ripoll *et al*., 2009). Ripoll *et al*. (2009) also found that the genome of *M. abscessus* encodes for proteins involved in drug efflux systems.

***M. abscessus* ATCC 19977**

This strain was first isolated in 1953 in a subcutaneous like lesion from a knee infection in a human (Choo *et al*., 2012).

***M. abscessus* subspecies *bolletii* 50594**

This strain was isolated from a Korean patient with a pulmonary infection (Kim *et al*., 2013).

2.2.3. *Mycobacterium avium* complex

2.2.3.1. *Mycobacterium avium*

M. avium, from the MAC, is a natural pathogen of birds (Wu *et al*., 2006; Gillespie & Bamford, 2012), and an opportunistic pathogen of humans, although very little epidemiologic information is available regarding disease in humans (Horan *et al*., 2006). It is a gram-positive, non-spore forming, slightly curved, aerobic, slow growing bacteria (Galagan *et al*., 2010). They commonly cause mycobacterial lymphadenitis in children, osteomyelitis in immune compromised patients and chronic pulmonary infection in the elderly (Gillespie & Bamford, 2012). In advanced stages of AIDS, the organism can cause a disseminated infection or bacterial sepsis (Vissa & Brennan, 2001; Washington *et al*., 2006; Gillespie & Bamford, 2012).

Patients at risk for infection with *M. avium* include those suffering or affected by chronic obstructive pulmonary disease by whatever primary cause, bronchiectasis, chronic aspiration or recurrent pneumonia, TB, pneumoconiosis and bronchogenic carcinoma (Washington *et al*., 2006). There has been some correlation reported between infection with

M. avium and cystic fibrosis (Washington *et al*., 2006). In non-immune compromised patients, the infection presents like infection with *M. tuberculosis* (Washington *et al*., 2006), and occasionally patients are asymptomatic.

Treatment usually involves multidrug regimens as the organism is naturally resistant to most of the anti-TB drugs (Gillespie & Bamford, 2012). Infection can be treated with a combination of drugs, such as Rifabutin, ethambutol & Clarithromycin for an optimal survival rate (Shafran *et al*., 1996).

Wu et al. (2006) examined the genome of the subspecies in the MAC to try and identify the diversity in the different subspecies to better understand the evolution of the strains regarding host preference and pathogenesis. Not only will the results of such a study improve our understanding of the evolution of the organisms, but also increase the chance of developing vaccines (Wu et al., 2006).

***M. avium* 104**

Horan *et al*. (2006) revealed 4 480 open reading frames in their draft sequence of *M. avium* 104 and found genes common with other virulent mycobacteria, as well as some genes unique to MAC. The organism is widely distributed in the environment and, being an opportunistic pathogen, is known to cause infection in immune compromised patients (Wu *et al*., 2006). The first of this strain was isolated in 1983 in South California from an adult patient infected with AIDS (KEGG Database: http://www.genome.jp/dbget-bin/www_bget?genome:T00433).

***M. avium* subspecies *paratuberculosis* K10**

M. avium subspecies *paratuberculosis* is an obligate pathogen of animals such as cattle, sheep, antelopes, deer and giraffes, causing disease characterised by enteritis (Vissa & Brennan, 2001; Wu *et al*., 2006). Infection in animals causes major economic losses,

especially in the dairy industry (Wu *et al*., 2006). *M. avium* subspecies *paratuberculosis* has also been associated with Chron's disease in humans, with symptoms like the intestinal symptoms caused by intestinal infection in cattle (Wu *et al*., 2006). This strain can only be treated with a combination of antibiotics such as Rifabutin and a macrolide such as Clarithromycin, as it is not susceptible to a lot of the anti-TB, making treatment very difficult. The genome was sequenced and a total of 4 350 genes was found (Wynne *et al*., 2010).

***M. avium* subspecies *paratuberculosis* MAP4**

This strain is the causative agent for Johne's disease in animals like cattle, sheep, goat, deer, giraffes, antelopes and camels (You *et al*., *2012;* Bannantine *et al*., 2014), and is associated with Chron's disease in humans (Bannantine *et al*., 2014). Johne's disease is a chronic granulomatous infection in the small intestine of these animals with a worldwide distribution (Bannantine *et al*., 2014). Johne's disease takes a great economic toll on the livestock industry (Bannantine *et al*., 2014). This strain was isolated in 2000 from the breast milk of a patient with Chron's disease (KEGG database: http://www.genome.jp/dbget-bin/www_bget?genome:T02662).

The organism has a complex cell wall structure with mycolic acids and several lipids that are very similar to other strains in this genus, but it is the most slow-growing of these organisms (Bannantine *et al*., 2014). The organism requires 8 to 16 weeks before colonies are visible and a siderophore mycobactim is needed in the laboratory medium for growth (Bannantine *et al*., 2014).

2.2.3.2. *Mycobacterium intracellulare*

M. intracullulare is commonly found in the environment and, being an opportunistic pathogen, individuals at risk for infection are immune compromised patients (Wu *et al*.,

2006). Kim *et al*. (2012) stated that *M. intracellulare* is isolated more frequently from South Korea than *M. avium*, which is the most frequently isolated of the NTM.

***M. intracellulare* ATCC 13950**

This strain is the causative agent for extrinsic allergic alveolitis, and is the most frequently isolated in clinical settings of all the NTM (KEGG database: http://www.genome.jp/dbget-bin/www_bget?genome:T02158). This strain belongs to the INT2 genotype (Kim *et al*., 2012).

***M. intracellulare* MOTT -02**

This strain is known for causing extrinsic allergic alveolitis and other nontuberculosis infections (KEGG database: http://www.genome.jp/kegg-bin/show_organism). This strain also belongs to the INT2 genotype (Kim *et al*., 2012).

***M. intracellulare* MOTT -64**

It is the most frequently encountered genotype in South Korea (Kim *et al*., 2012).

***M. Indicus pranii* MTCC 9506**

This strain is placed higher up than *M. tuberculosis* on an evolutionary scale and is considered non-pathogenic (Singh *et al*., 2014). Findings by Singh e*t al*. (2014) suggest a possibility that this strain has the potential to be used as an effective vaccine against TB.

2.2.4. Mycobacterium causing Leprosy

2.2.4.1. *Mycobacterium leprae*

M. leprae is gram-positive, aerobic and surrounded by the unique mycobacterial waxy coating (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_leprae_TN). The organism was identified in 1873 before *M. tuberculosis*, making it the first human pathogenic

mycobacterium to be identified (Vissa & Brennan, 2001). Up to date humans are the only known host (Monot *et al*., 2009).

The bacterium causes leprosy, or Hansen's disease, in humans. Leprosy is a chronic infectious and primarily granulomatous disease of the peripheral nerves, mucosa and upper respiratory tract, producing skin lesions (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_leprae_TN). The organism initially attacks peripheral nerves, followed by digital destruction and deformity, often leaving the patient severely disabled (Monot *et al*., 2009; Gillespie & Bamford, 2012). Patients particularly at risk are immune compromised patients (Gillespie & Bamford, 2012).

Treatment originally involved dapsone and its derivatives, but after the development of resistance to the treatment, we now rely on multidrug therapy for successful treatment http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_leprae_TN). Treatment can render the patient non-infectious, but unfortunately it cannot rectify the nerve damage or resultant deformity, which can only be managed by remedial surgery (Gillespie & Bamford, 2012). Vissa & Brennan (2001) noted a total of 3 268 203 base pairs that encode for 1 605 genes, as well as 1 116 pseudogenes.

***M. leprae* Br4923**

This strain of the organism was first isolated from a patient in Brazil, the country with the second highest incidence of leprosy (Monot *et al*., 2009).

***M. leprae* TN**

This strain was first isolated from an armadillo in Tamil Nadu, India, infected with *M. leprae* from a lepromatous patient's biopsy material (Cole *et al*., 2001).

2.2.5. Non-tuberculosis Mycobacteria

2.2.5.1. *Mycobacterium sp*. JDM601

This is a slow growing strain of mycobacteria with its genes most similar to that of *M. marinum* (Zhang *et al*., 2011).

2.2.5.2. *Mycobacterium liflandii* 128FXT

This strain, of Mycobacterium, resembling *M. ulcerans*, was originally isolated from an African clawed frog (Xenopus tropicalis) (Tobias *et al*., 2013). Like *M. ulcerans*, this strain can also make mycolactone, as it also contains the pMUM mycolactone plasmid in its genome (Tobias *et al*., 2013). This organism has undergone reductive evolution from the ancestral genomes of *M. ulcerans* Agy99 and *M. marinum* (Tobias *et al*., 2013). Although this strain contains many similar genes as *M. ulcerans*, it resembles *M. marinum* more regarding genome length, sequence and architecture (Tobias *et al*., 2013).

2.2.5.3. *Mycobacterium ulcerans* Agy99

This organism causes the devastating Buruli ulcer in humans and is believed to have derived from *M. marinum* (Stinear *et al*., 2007; Stinear *et al*., 2008, Doig *et al*., 2012). Burulis ulcer is a necrotic disease of subcutaneous tissue and is widespread in West and Central Africa (Stinear *et al*., 2007). The mode of transmission is still uncertain (Doig *et al*., 2012). This is the only pathogenic mycobacterium that does not have an intracellular lifestyle in the granulomas, as its polyketide toxin gives the bacteria the ability to lyse cells and survive extracellulary and provoke necrosis (Stinear *et* al., 2007; Stinear *et al*., 2008, Doig *et al*., 2012). The toxin produced by the organism, mycolactone, is described as a macrolide toxin (Stinear *et al*., 2007, Doig *et al*., 2012). The slow growth, ability to produce immune suppressors, cell wall remodelling and modification, and biofilm-forming capability is what gives the organism its survival ability (Doig *et al*., 2012).

2.2.5.4. *Mycobacterium marinum*

M. marinum is a facultative anaerobic, fast growing organism with a growth rate of 4 to 6 hours (Stinear *et al*., 2008). Unlike *M. tuberculosis*, *M. marinum* is unable to reduce nitrate and it produces a bright yellow carotenoid pigment when exposed to light, which protects it from UV damage (Stinear *et al*., 2008). The organism can replicate in both single celled organisms, such as amoeba, and cultured mammalian macrophages, meaning the organism can survive in both intracellular and unpredictable extracellular environments (Stinear *et al*., 2008).

It is mainly a pathogen of fish and amphibians (Stinear *et al*., 2008), and is found in aquatic environments such as swimming pools and drinking water. The systemic granulomatous infection in amphibians and fish has histologic similarities to lesions of TB in humans (Stinear *et al*., 2008). In humans it causes a granulomatous infection of the skin and it is mainly acquired from contaminated rivers, poorly maintained swimming pools and in fish tanks through cuts and scratches in the skin (Washington *et al*., 2006; Stinear *et al*., 2008; Gillespie & Bamford, 2012). Infection with *M. marinum*, commonly referred to as fish tank or aquarium tank granuloma (Stinear *et al*., 2008), has been reported in lifeguards, fishermen and people working in the aquatic industry (Washington *et al*., 2006). The granulomatous infection is generally limited to the skin and soft tissue extrimities at the point of inoculation and pathologically it resembles dermal disease caused by *M. tuberculosis* (Stinear *et al*., 2008). *M. marinum* has an optimal growth temperature at 35^0C in Middlebrook 7H9 medium and grows poorly at 37^0C, which explains its tendency to infect poikilotherms and a superficial disease restricted to cooler extremities of the human body (Stinear *et al*., 2008).

This organism was first isolated in 1992 in San Francisco from a patient with fish tank granuloma (Stinear *et al*., 2008). Fish tank granuloma can heal spontaneously but due to the *in vivo* development of antibiotic resistance, long term antibiotic treatment is frequently required (Stinear *et al*., 2008). Treatment consists of chemotherapy with a combination of drugs, including rifampicin and ethambutol and the quinilones, as well as doxycycline and clarithromycin (Washington *et al*., 2006; Stinear *et al*., 2008). The treatment may be effective, but frequently surgical intervention, i.e. resecting the primary lesion, may be required (Washington *et al*., 2006). When chronic lymphatic spread is evident, anti-tuberculosis therapy may be required (Washington *et al*., 2006). Most strains are susceptible to rifampicin and ethambutol, but resistance to isoniazid and streptomycin have been reported (Washington *et al*., 2006).

Some patients present with sporotrichiosis-like lesions, with infection spreading from the primary site of infection through the lymphatics (Washington *et al*., 2006). More commonly, however, the lesions present as tender, red or blue-red subcutaneous nodules, typically on the elbow, knee, toe or finger. Involvement of the subcutaneous bursae, tendon sheaths, joints and bones have been reported (Washington *et al*., 2006).

2.2.5.5. *Mycobacterium kansasii* ATCC 12478

M. kansassii ATCC 12478 is resistant to high levels of ρ-Aminosalicylic acid (ATCC Product Sheet: www.atcc.org). This strain is used as a CLSI control for the susceptibility testing of clinical isolates of *M. kansasii* to rifampicin (ATCC Product Sheet: www.atcc.org).

2.2.6. Saprophytes

2.2.6.1. *Mycobacterium sp*. JLS

Mycobacterium sp. JLS is a fast growing, Gram negative organism with mycolic acid derivatives on its surface. The organism was first isolated from soil in a wood preservative-contaminated area where PAH degradation was occurring (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#_Mycobacterium_JLS_).

2.2.6.2. *Mycobacterium sp.* KMS

Mycobacterium sp. KMS is an aerobic, obligate mycobacterium that is non-motile (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_sp._KMS). The organism was also isolated from a wood preservative-contaminated land-treatment where bioremediation of PAH was taking place.

2.2.6.3. *Mycobacterium sp.* MCS

This strain is described as fast growing, Gram positive, pleomorphic, non-motile and belongs to the order Actinomycetales http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_sp._MCS). It is characterised by its cell surface mycolic acid derivatives. Like the KMS strain, it was isolated from soil in a wood preservative-contaminated land-treatment where bioremediation of PAH was taking place.

2.2.6.4. *Mycobacterium vanbaalenii* PYR-1

M. vanbaalenii is a rod shaped, non-motile, non-spore forming bacteria (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_vanbaalenii_PYR-1). *M. vanbaalenii* plays a role in bioremediation, as the organism can degrade a wide range of environmentally toxic chemicals, including high-molecular-weight PAH, like pyrene, as a sole source of carbon and energy (Kim *et al*., 2008; Ripoll *et al*., 2009). This strain was first isolated at a site close to the Harbor Island oil tank farm in the watershed of Redfish Bay

(http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_vanbaalenii_PYR-1).

Polycyclic aromatic hydrocarbon is produced by the combustion of fossil fuels, waste incineration and industrial processes, and naturally because of forest fires (Kim *et al*., 2008). Pyrene is a common pollutant of water, air and soil, is harmful to aquatic micro-invertabrates, and its metabolites is toxic to other organisms in the environment (Kim *et al*., 2008). The metabolic diversity of *M. vanbaalenii* makes it a potential source for bioremediation of PAH-contaminated areas (Kim *et al*., 2008). The genome sequence of *M. vanbaalenii* was done and found to contain 6 012 protein-coding sequences, many of which are putative genes involved in the catabolism of aromatic compounds (Kim *et al*., 2008).

2.2.6.5. *Mycobacterium smegmatis*

M. smegmatis is an aerobic, chemoorganotrophic, non-motile human pathogen that is unable to replicate intracellularly (Stinear *et al*., 2008). Infection is associated with soft tissue lesions following either trauma or surgery, and the organism is associated with penile carcinogenesis (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_smegmatis_MC2155). Despite its large genome *M. smegmatis* has only three ESX loci and lacks ESX -1 (Stinear *et al*., 2008). Although the organism was first isolated from humans it is mostly non-pathogenic and is used as a model for studying mycobacterial physiology, since it shares so many similar traits to other pathogenic mycobacteria (He & De Buck, 2010). The organism has a cell wall that not only gives it shape, but also helps to protect the organism against differences in osmotic pressure and other physical and chemical onslaughts in its environment (He & De Buck, 2010).

M. smegmatis is classified as a saprophytic species that rarely causes disease and isn't dependent on living in an animal, unlike some pathogenic Mycobacterium (Long *et al*., 2012 ;https://microbewiki.kenyon.edu/index.php/Mycobacterium_smegmatis). Hence in this study, *M. smegmatis* is included under Saprophytes.

***M. smegmatis* MC2155**

M. smegmatis MC2155 is a human pathogen and was first isolated in 1990 (http://genome.tbdb.org/tbdb_sysbio/GenomesIndex.html#Mycobacterium_smegmatis_MC2155).

***M. smegmatis* JS623**

M. smegmatis JS623, an ethane oxidising organism, can grow on vinyl chloride (VC) as a source of carbon and energy (Jin *et al*., 2010). Vinyl chloride is a toxic ground water pollutant associated with plastic manufacture and chlorinated solvent used is a known human carcinogen (Jin *et al*., 2010). This trait makes it potentially useful in the bioremediation of such pollutants (Jin *et al*., 2010). This strain is also a useful model system for studying microbial enzyme evolution in response to xenobiotic compounds (Jin *et al*., 2010).

2.2.6.6 *Mycobacterium chubuense* NBB4

This strain was isolated as part of a bioprospecting study, while searching for novel monooxygenase enzymes related to bioremediation (Coleman *et al*., 2011). In their investigation, Coleman *et al*. (2011) found that *M. chubuense* NBB4 contained an unprecedented variety of monooxygenase genes, making it a versatile hydrocarbon degraded (Coleman *et al*., 2011).

2.2.6.7. *Mycoacterium gilvum*

M. gilvum is a soil bacterium capable of degrading pyrene and other aromatic hydrocarbons (DeanRoss & Cerniglia, 1996).

***M. gilvum* PYR-GCK**

This strain was isolated from the sediment of the Grand Calumet River in Northwestern Indiana and can utilize the toxic polycyclic hydrocarbon pyrene for growth (Badejo *et al*., 2014). It has been studied to try and gleam some insight into mechanisms related to its exceptional bioremediating abilities (Badejo *et al*., 2014).

***M. gilvum* Spyr1**

This strain was isolated from a creosote polluted site in Epirus, Greece, where a wood preserving industry was operating for over 30 years (Kallimanis *et al*., 2011). It can degrade a wide range of PAH substrates as a carbon source including: pyrene, fluoranthene, fluorene, anthracene and acenapthene (Kallimanis *et al*., 2011).

2.2.6.8. *Mycobacterium. rhodesiae* NBB3

M. rhodesiae was originally isolated in Rhodesia, now Zimbabwe, from a patient with pulmonary TB (www.uniprot.org). The strain NBB3 was isolated from creosote-contaminated soil in estuanne sediment from Australia (www.uniprot.org). The strain is not only known for its ability to degrade PAH, but also for its ability to develop pigment in both dark and light (www.uniprot.org).

2.2.6.9. *Mycobacterium neoaurum* VKM Ac-1815D

M. neoaurum VKM Ac-1815D can synthesise the steroid precursor 4-androstene-3,17-dione as a major product from sitosterol, a drug used for treatment of high cholesterol, an enlarged prostate and possibly for TB (Shtratnikova *et al*., 2014). This strain is considered important for the development of new biotechnological applications for mycobacterial conversion of bioactive steroids (Shtratnikova *et al*., 2014).

2.3. Tuberculosis: The major threat to human health

Tuberculosis is caused by *M. tuberculosis* (He *et al*., 2003; Zheng *et al*., 2008; Vasconcellos *et al*., 2010; Jena *et al*., 2013). It is a chronic infectious, highly communicable disease (Cole *et al*., 1998; Washington *et al*., 2006), and is spread from human to human via the airborne route (Price & Wilson, 2003; Gillespie & Bamford, 2012). The lung is usually the primary site of infection, especially in children, and most infections resolve with only local scarring (Washington *et al*., 2006). In about 5% the infection is not controlled, and it spreads from the primary focus throughout the body, which can resolve spontaneously or progress into other localised infections, for example meningitis (Gillespie & Bamford, 2012).

Pulmonary infection symptoms include chronic cough, haemoptysis, fever, night sweats, weight loss and chest pain (Price & Wilson, 2003; Washington *et al*., 2006; Gillespie & Bamford, 2012), to name a few, and other organs or systems possibly affected include the renal system, bones, joints and the abdominal system (Gillespie & Bamford, 2012). Disseminated, or miliary, infections are common in individuals who suffer from malnutrition, immunosuppression, or other chronic debilitating diseases (Washington *et al*., 2006). Reactivation, commonly referred to as adult type infection, is a slowly progressive inflammatory process in the lungs characterised by intense granulomatous inflammation, with formation of many Langhans-type giant cells, necrosis, and caseation, with a tendency to erode into the bronchi (Washington *et al*., 2006).

Countless people have died from TB across the globe (Cole *et al*., 1998) despite the availability of effective chemotherapeutic drugs and the BCG vaccine (Cole *et al*., 1998; Zheng *et al*., 2008). Outbreaks of the infection in closed communities/populations, such as schools, ships and crowded family groups are quite common (Washington *et al*., 2006).

The World Health Organisation (WHO) declared TB a global emergency to increase awareness of the disease, both publicly and politically (Cole *et al*., 1998; Washington *et al*., 2006). Currently a third of the world's population is affected by infection with *M. tuberculosis* and the epidemic is growing, with approximately 9 million new cases and upwards of 2 million deaths each year across the world (He *et al*., 2003; Washington *et al*., 2006; Zheng *et al*., 2008; Jena *et al*., 2013). The incidence of HIV virus significantly contributes to the spread of TB and projections indicate that between 2002 and 2020 1 billion people will newly acquire infection (Washington *et al*., 2006; Jena *et al*., 2013). Of these, 150 million will get sick, and approximately 36 million will die (Washington *et al*., 2006). Given the scale and the global threat of tuberculosis, vaccination should not only be a priority, but could possibly be the only feasible intervention that may make a difference in the fight against TB (Cole *et al*., 1998).

The rate of increase is even greater in developing countries because of the high rate of endemnicity, declining socioeconomic conditions in highly populated cities and the increasing number of HIV infected individuals (Washington *et al*., 2006). Even in many developed countries TB poses a considerable risk, despite research on vaccines and treatments (Jena *et al*., 2013). To prevent the grim predictions of the WHO from becoming a reality, radical measures must be taken (Cole *et al*., 1998). Vaccination development research shows some promise especially in the areas of "DNA vaccination, use of secreted surface-exposed proteins as immunogens, recombinant forms of BCG and rational attenuation of *M. tuberculosis*" (Cole *et al*., 1998).

The dormancy characteristic of *M. tuberculosis* may contribute to the chronic nature of the disease and the lengthy treatment regime required (Cole *et al*., 1998). The state of dormancy in which the organism remains inactive in the infected tissue may reflect the

metabolic shutdown resulting from the action of a cell-mediated response that can suppress, but not eradicate the infection (Cole *et al*., 1998). Reactivation of dormant bacteria can be caused by waning immunity, even long after the initial infection (Cole *et al*., 1998).

The WHO has estimated that at least a third of the people with HIV also have TB (Washington *et al*., 2006). HIV infection and TB are synergistic diseases on every level, from the molecular to the epidemiologic, as the presence of HIV promotes the intracellular replication of the tubercle bacillus, promoting the transmission of *M. tuberculosis* to other individuals (Cole *et al*., 1998; Washington *et al*., 2006).

In patients with AIDS, TB tends to follow the pattern of primary progressive disease, and is characterised by more rapid progression, septicaemia, and miliary dissemination to involve virtually every organ in the body, and less focal fibrosis and caseation (Washington et al., 2006).

2.4. Tuberculosis: Treatment and drug resistance

Treatment of TB is a lengthy process and aside from it being toxic, the treatment is expensive and the success rate up to now is unsatisfactory (Ambrosio *et al*., 2015). Prevention and control measures are aimed at early detection and treatment of cases and sources of infection (Price & Wilson, 2003). The purpose of early detection is mainly to identify individuals who might benefit from treatment to prevent development of active disease, benefitting not only the individual but the community as well (Price & Wilson, 2003).

There is a rise in the incidence of MDR strains of *M. tuberculosis* (Migliori *et al*., 2012). Ambrosio *et al*. (2015) reported that over 480 000 of the 9 million new TB cases every year are cases of MDR-TB. Of these, approximately 9% are cases of XDR-TB (Ambrosio *et al*., 2015). Drug resistant TB is a worldwide phenomenon and in a survey in 2000 of 35

countries it was reported that at least 12.6% of isolated strains are resistant to at least one drug and 2.2% were resistant to both of the primary drugs (Price & Wilson, 2003).

Drug resistance develops because of sub-optimal treatment regimes, either as a result of individuals being infected by an already resistant strain of *M. tuberculosis*, or as a result of inadequate treatment or drug non-compliance by the patient (Price & Wilson, 2003). Clinicians, managing the treatment of these cases, face challenges related to adverse events, patient compliance, lack of clinical experience and limited availability of adequate diagnostics and 2nd line drugs every day, making development of further drug resistance a real possibility (Ambrosio *et al*., 2015).

Treatment of TB consists of a multidrug regimen, including rifampicin and isoniazid for at least 6 months, combined with ethambutol and pyrazinamide for the first two months (Gillespie & Bamford, 2012). Treatment of MDR-TB is even more complex, requiring a combination of second-line anti-TB drugs, such as aminoglycosides, fluoroquinolones, ethiomide or cycloserine (Gillespie & Bamford, 2012). Ambrosio *et al* (2015) summarised a MDR strain as one that is resistant to at least isoniazid and rifampicin, and an XDR strain as being resistant to at least one fluoroquinolone and one injectable 2nd line anti-TB drug. The table below (Table 2) shows the grouping of anti-TB drugs based on the treatment stages.

Table 2 World Health Organization grouping of anti-tuberculosis (taken from Ambrosio et al., 2015)

Group	Anti-tuberculosis drug
1st line oral anti-tuberculosis drugs	Isoniazid Rifampicin Ethambutol Pyrazinamide

	Rifabuton Rifapentime
Injectable anti-tuberculosis drugs (Injectable or parenteral agents)	Streptomycin Kanamycin Amikacin Capreomycin
Fluoroquinolones (FQ's)	Levofloxacin Moxifloxacin Gatifloxacin Ofloxacin
Oral bacteriostatic 2nd line anti-tuberculosis drugs	Ethionamide Prothionamide Cycloserine Terizidane ρ-Aminosalicylic acid (PAS) ρ- Aminosalicylate sodium (PAS-Na)
Anti-tuberculosis drugs with limited data on efficacy and/or long-term safety in the treatment drug-resistant TB	Bedaquiline Delamanid Linezolid* Clofazimine Amoxicillin/Clavulanale Imipenim/Cilostatin Meropenem High-dose Isoniazid Theocatezone Clarithromycin

**Found to be effective in difficult to treat resistant cases, but the adverse effects, including peripheral neuropathym,optical neuropathy, gastrointestinal disorders and myelosuppression, limits long term use.*

The WHO recommends the following treatment regime for new 1st time diagnosed patients (Ambrosio *et al*., 2015):

- In the intensive phase, lasting two months: Ethambutol, Isoniazid, Rifampicin and Pyrozinamide

- In the continuous phase, lasting a further four months: Isoniazid, Rifampicin and Ethambutol three times a week

Treatment for previously treated patients, pending results of drug susceptibility testing and low probability of the strain being MDR, as recommended by the WHO (Ambrosio *et al*., 2015):

- In the intensive phase, for two months: Ethambutol, Isoniazid, Rifampicin, Pyrazinamide and Streptomycin

- Another one-month intensive phase: Ethambutol, Isoniazid, Pyrazinamide

The WHO recommends using drugs form the five groups (Table 2) based on effectiveness and safety (Ambrosio *et al*., 2015). Drug susceptibility testing is preferred but if it is not available, these standardised regimes can be followed (Ambrosio *et al*., 2015).

For the treatment of a MDR case, a minimum of four active drugs must be taken from the five groups with the following recommendation by the WHO (Ambrosio *et al*., 2015):

- Choose an injectable – based on drug susceptibility testing and treatment history

- Choose a higher generation fluoroquinolone – if strain is resistant to Levofloxacin, use Moxifloxacin but avoid using Moxifloxacin when using Bedaquiline

- Add group four drugs – add two or more until at least four 2nd line drugs are likely to be effective. Ethionamide or Prothinamide seems to be the most effective. Consider the treatment history, side effects and the cost involved

- Add drugs from group one – Pyrazinamide is routinely added and Ethambutol if it is fully sensitive. Isoniazid is not used until drug susceptibility results are available

- Add drugs from group five – only if four drugs from group two to four are not effective. Add two or more of necessary. Drug susceptibility testing is not standard practice in this group

When using injectables, it should be used for a minimum of eight months and the continuation phase usually lasts another 12 to 18 months, making the length of treatment up to 20 months (Ambrosio *et al*., 2015). The consensus for estimating treatment duration is an added 18 months to the date of the 1st negative culture (Ambrosio *et al*., 2015).

2.5. The search for novel anti-TB drug targets

Despite the advances already made in the approach to therapy and development of vaccines, TB remains a significant threat. Although this infectious disease can be cured, the treatment required is lengthy one. The length of the treatment brings with it the threat or complication of patient non-complicance, for whatever reason, the dangers of adverse effects, and the development of drug resistance (Mdluli & Spigelman, 2006). The ongoing problem of drug resistance is a major contributor to the major threat that is TB, and it emphasises the need for the identification of new drug targets for new, more effective drugs to be developed.

Since the determination of the *M. tuberculosis* genome (Cole *et al*., 1998), the information has become available and can be used for the identification and validation of

drug targets as the basis of development and success rates of new anti-TB drugs with increased frequency (Monaghan & Barrett, 2005; Mdluli & Spigelman, 2006; Zhang *et al*., 2006). Genomics provide the opportunity to better understand factors such as resistance development, movement and transport out of the cell, and interaction between the therapeutic compound and the pathogen (Monaghan & Barrett, 2005). Genomics also offers an insight into bacterial niche, adaptation ability, host susceptibility, pharmacokinetics and microbial genesis of chronic diseases (Monaghan & Barrett, 2005).

The validation of new targets as per Mdluli & Spigelman (2006) may include:

- The demonstration of the biochemical activity of an enzyme
- The determination of an enzyme's structure when in complex with a substrate or an inhibitor
- Confirmation of importance
- Identification of growth inhibitors

If the validation and identification of inhibition can lead to a better understanding of the disease biology, it can open the possibility for the development of new drugs that can shorten the duration of therapy, prevent development of resistance and also possibly eliminate latent disease (Mdluli & Spigelman, 2006). Pathogenesis is an important factor to consider, e.g. liquefaction from a solid necrotic lesion to cavity formation enables the spread of the disease through droplets (Zhang *et al*., 2006). Inhibition of this liquefaction to cavity formation will mean that bacilli cannot be coughed up and spread to others. A novel approach would be to develop drugs that prevent transmission (Zhang *et al*., 2006).

Current drug regimens are focused on targeting the following pathways (Mdluli & Spigelman, 2006; Zhang *et al*., 2006):

- **Cell wall biosynthesis:** Cell walls are made up of peptidoglycan, arabinogalactan and mycolic acids. Some of the current drugs, including isoniazid inhibit these three components. The drug, however, does not shorten the length of treatment of active disease.

- **Fatty acid biosynthesis:** *M. tuberculosis* uniquely has both type I and type II biosynthetic pathways for fatty acids. Inhibition of one of these pathways makes for an attractive drug target.

- **Amino acid biosynthesis:** Inhibition of amino acid biosynthesis prevents the organism from hunting for nutrients in the human host, and thus will inhibit the growth of the organism.

- **Cofactor biosynthesis:** Folate derivatives are cofactors for the biosynthesis of essential molecules such as purines, pyremidines and other amino acids. Trimethoprim targets two enzymes in the folate biosynthesis pathway. Another target is Co-enzyme A, which is involved in lipid synthesis in the organism. Lumazine synthase is an enzyme that catalyses the final step in the riboflavin synthesis, an essential pathway for survival of *M. tuberculosis*, and thus is another attractive pathway.

- **Mycothiol biosynthesis:** This pathway is essential for protection against the effects of reactive oxygen.

- **Terpenoid biosynthesis:** *M. tuberculosis* uses a non-mevalonate pathway for terpenoid synthesis. Terpenoids are described as modified terpenes, a group of naturally occurring organic chemicals, with added functional groups. Inhibition of the pathway of the synthesis of these terpenoids is an attractive target.

- **DNA synthesis:** Many of the enzymes involved in the synthesis of DNA, which is essential for bacterial growth, are already drug targets.

- **The glyoxylate cycle:** Targeting enzymes involved in this pathway will prevent or inhibit cellular replication and enhance elimination of the bacterial from the lungs.

- **Regulatory proteins:** Successfully inhibiting a regulatory protein is very effective as it will disrupt the whole network of proteins under the regulation or influence of that protein/regulator.

- **Menaquinone biosynthesis:** Inhibition of this pathway will prevent growth, as the biosynthesis of this compound is essential for bacterial growth.

- **Stringent response enzymes:** This can be described as a stress response in reaction to conditions such as starvation, fatty acid limitation and iron limitation. In other words, a response that plays a critical role in the organism's ability to adapt to various in vivo conditions, and if the enzymes involved in this response is inhibited or destroyed the adaptation ability will be significantly reduced.

- **Adenosine Triphosphate (ATP) biosynthesis:** ATP transports chemical energy necessary for metabolism in the cells. This pathway has been identified as a possible drug target in recent years.

The WHO launched their "End TB strategy", encompassing their vision of a TB-free world, emphasizing the need for development of new drug regimens (World Health Organization: Global Tuberculosis Report, 2012; Ambrosio *et al*., 2015). The currently available drugs target a small number of essential functions in the organism and identification of more pathways and functions required for bacterial growth and survival may provide more targets for the design and development of effective therapeutic drugs (Mdluli & Spigelman, 2006). The identification of new drug targets should be focused on vital aspects for bacterial growth, metabolism and viability and pathogenesis of the organism (Zhang *et al*., 2006). Virulence factors should also be considered for targets (Zhang *et al*., 2006). New drugs should not only be focused on resistant strains, but also on persistent strains and it should also be aimed at shortening the length of the treatment, which is underlying to the resistance problem due to poor compliance (Zhang *et al*., 2006).

2.6 Comparative genomics

Comparative genomics is the comparison of sequenced genomes using various tools to compare the complete genome sequences of different species and is very helpful in identifying genotypic differences that is responsible for the differences observed between strains or species (Comparative Genomics Fact Sheet, 2015; Sullivan *et* al., 2011; Touchman, 2010). Whole genome comparison offers a highly detailed view of how organisms are related to one another on a genetic level and enables us to see the remarkable unity of all living things, as well as the prominent diversity (Koonin *et al*., 2000; Touchman, 2010). This information will enhance our understanding of the structure and function of the genes and assist in the development of new strategies to combat disease (Comparative Genomics Fact Sheet, 2015; Touchman, 2010), while also giving us a unique understanding of the biochemistry, physiology and pathogenesis of a microorganism (Brosch, 2001). A basic

comparison of the general features of genomes that can serve as a basic entry into genomic analysis include comparison of general features such as genome size and number of genes (Touchman, 2010). Comparative genomics has given us a glimpse into the amazing plasticity of the genome (Koonin *et al*, 2000).

Comparative genomics can be applied to study genetic variability within populations of pathogens to provide insight into their evolution and pathogenesis (Kato-Maeda *et al*., 2001). Genomics in combination with bioinformatics can potentially unable researchers to generate the knowledge to develop new therapeutic interventions (Cole *et al*., 1998).

2.7. References

Al Rashdi, ASA, Jadhay, BL, Deshpande, T & Deshpande, U (2014) Whole-Genome Sequencing and Annotation of a Clinical Isolate of Mycobacterium tuberculosis from Mumbai, India. **Genome Announcements**. 2(2)

Ambrosio, LD, Centis, R, Sotgiu, G, Pontali, E, Spanevello, A & Migliori, GB (2015) New anti-tuberculosis drugs and regimens: 2015 update. **ERJ open research**. Vol 1, Issue 1. Doi: 10.1183/23120541.00010-2015

Badejo, AC, Chung, WH, Kim, NS, Chai, JC, Lee, YS, Jung, KH, Kim, HJ & Chai, YG (2014) Energy Metabolism in Mycobacterium gilvum PYR-GCK: Insights from Transcript Expression Analyses Following Two States of Induction. **PLOS ONE**. http://dx.doi.org/10.1371/journal.pone.0099464

Bannantine, JP, Li, L, Mwangi, M, Cote, R, Garay, JAR & Kapur, V (2014) Complete Genome Sequence of Mycobacterium avium supsp. Paratuberculosis, Isolated from Human Breast Milk. **Genome Announcement**. 2

Bentley, SD, Comas, I, Bryant, JM, Walker, D, Smith, NH, Harris, SR, et al. (2012) The Genome of Mycobacterium africanum West African 2 Reveals a Lineage Specific Locus and Genome Erosion Common to M tuberculosis Complex. **PloS Neglected Tropical Diseases**. 6(2): e1552. Doi: 10.1371/journal pntd 0001552

Brosch, R, Pym, AS, Gordon, SV & Cole, ST (2001) The evolution of mycobacterial pathogenicity: clues from comparative genomics. **TRENDS in Microbiology**. 9(9): 452 – 458

Camus, JC, Pryor, MJ, Médigue, C & Cole, ST (2002) Re-annotation of the genome sequence

of Mycobacterium tuberculosis H37Rv. Microbiology. 148, 2967 - 2973

Cardoso, AM, De Sousa, EM, Viana-Niero, C, De Bortoli, FB, Das Neves, ZCP, Leao, SC, et al. (2008) Emergance of nosocomial Mycobacterium massiliense infection in Goias, Brazil. Microbes and Infection, Vol10, Issue 14 – 15, P 1552 – 1557

CDC: Division of Tuberculosis Elimination. Available from: www.cdc.gov/tb/publications/factsheets/general/mbovis.pdf

Centre for Food Security & Public Health (2009) Bovine Tuberculosis. Institute for International Cooperation in Animal Biologics. Iowa State University. Available from: www.cfsph.iastate.edu/factsheet/pdfs/bovine_tuberculosis

Chan, J, Halachev, M, Yales, E, Smith, G & Pallen, M (2012) Whole-Genome Sequence of the Emerging Pathogen Mycobacterium abscessus Strain 47J26. **Journal of Bacteriology**. 194(2): 549.

Choo, SW, Wong, YL, Yusoff, AM, Leong, ML, Wong, GJ, Ong, CS, Ng, KP & Ngeow, YF (2012) Genome sequence of the Mycobacterium abscessus Strain M93. **Journal of Bacteriology**. 3278

Coker, OO, Regmi, SM, Suriyaphol, P, Chininmanu, K, Prommananon, T & Chaipraser, A (2014) Whole-Genome Sequence of a Multidrug-Resistant Mycobacterium tuberculosis Sequence Type Isolate from an Outbreak in Thailand. **Genome Announc**. (4)

Cole, ST, Bosch, R, Parkhill, J, Garnier, T, Churcher, C, Harris, D, Gordon, SV et al. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. **Nature**. 393, 537 – 544

Cole, ST, Eiglmeier, K, Parkhill, J, James, KD, Thomson, NR, Wheeler, PR, Honorè, N, et al. (2001) Massive gene decay in the leprosy bacillus. **Nature**. 409(6823): 1007 – 1011

Cole, ST. 2002. Comparative and functional genomics of the *Mycobacterium tuberculosis* complex. **Microbiology**. 148: 2919 - 2928

Coleman, NV, You, S, Wilson, NL, Nolan, LM, Migocki, MD, Ly, MA, Crossett, B & Holmes, AJ (2011) Untangling the multiple monooxygenases of Mycobacterium chubuense strain NBB4, a versatile hydrocarbon degrader. **PubMed**

Comparative Genomics Fact Sheet (2015) Available from: http://www.genome.gov/11509542/comparative-genomis-fact-sheet/

De Voss, JJ, Rutter, K, Schroeder, BG, Su, H, Zhu, YQ & Barry, CE (2000) The Salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages. Proceedings of the National Academy of Sciences, 97, 1252 – 1257

DeanRoss, D & Cerniglia, CE (1996) Degradation of pyrene by Mycobacterium flavescens. **Appl Microbiol Biot**. 46(3): 307 – 312. Doi: 10.100/-s002530050822

Doig, KD, Holt, KE, Fyfe, JAM, Lavender, CJ, Eddyani, M, Portaels, F, Yeboah-Manu, D, et al. (2012) On the origin of Mycobacterium ulcerans, the causative agent of Buruli ulcer. **BMC genomics**. 13:258. Doi: 10.1186/1471-2164-13-258

Feuerriegel, S, Köser, CU, Richter, E & Niemann, S (2013) Mycobacterium canettii is intrinsically resistant to both pyrazinimide and pyrazinoic acid. **Journal of Antimicrobial Chemotherapy**

Fleischmann, RD, Alland, D, Eisen, JA, Carpenter, L, White, O, Peterson, J, DeBoy, R, et al.

(2002) Whole-Genome Comparison of Mycobacterium tuberculosis Clinical and Laboratory Strains. **Journal of Bacteriology**. 184(19): 5479 – 5490

Galagan, JE, Sisk, P, Stolte, C, Weiner, B, Koehrsen, M, Wymore, F, Reddy, TB, et al. TB database (2010) overview and update. **Tuberculosis**. 90(4): 225-35. Doi: 10.1016/j.tube2010.03.010

Garnier, T, Eiglemeier, K, Camus, JC, Medina, N, Mansoor, H, Pryor, M, Duthoy, S, et al. (2003) The complete genome sequence of Mycobacterium bovis. **PNAS**. 100(13): 7877 – 7882

Gaude, R & Parish, T (2008) The genetics of cell wall biosynthesis in Mycobacterium tuberculosis. **Future Microbiology**. 3: 299 – 313

Gillespie, SH & Bamford, KB (2012) Medical Microbiology and Infection at a glance. 4th ed. Wiley-Blackwell. UK

He, XY, Zhuang, YH, Zhang, XG & Li, GL (2003) Comparative proteome analysis of culture supernatant proteins of Mycobacterium tuberculosis H37Rv and H37Ra. **Microbes and Infection**. 5: 851 – 856

He, Z & De Buck, J (2010) Cell wall proteome analysis of Mycobacterium smegmatis strain MC2155. **BMC Microbiology**. 10:121. Doi: 10.1186/1471-2180-10-121

Horan, KL, Freeman, R, Weigal, K, Semret, M, Pfaller, S, Covert, TC, Van Soolingen, D, et al. (2006) Isolation of the Genome Sequence Strain Mycobacteirum avium 104 from Multiple Patient over a 17-Year Period. **Journal of Clinical Microbiology**. 44: 783 – 789

Ioerger, TR, Koo, S, No, E, Chen, X, Larsen, M, Jacobs Jr, WR, Pillay, M, Sturm, AW &

Sacchettini, JC (2009) Genome Analysis of Multi- and Extensively-Drug-Resistant Tuberculosis from Kwazulu-Natal, South Africa. **PLoS One**. http://dx.doi.org/10.13471/journal.pone.0007778

Isaza, JP, Duque, C, Gomez, V, Robledo, J, Barrera, LF & Alzate, F (2012) Whole genome shotgun sequencing of one Colombian clinical isolate of Mycobacterium tuberculosis reveals DesR regular gene deletions. **PubMed**. Available from: www.ncbi.nlm.nih.gov/pubmed/22404577

Jena, L, Kahikar, S, Kumar, S & Harinath, BC (2013) Comparative proteomic analysis of Mycobacterium tuberculosis strain H37Rv versus H37Ra. **International Journal of Mycobacteriology**. 2, 220 – 226

Jin, YO, Cheung, S, Coleman, NV & Mattes, TE (2010) Association of Missense Mutations Epoxyalkane Coenzyme M Transferase with Adapation of Mycobacterium sp. JS623 to Growth of Vinyl Chloride. **American Society for Microbiology**

Joung, SM, Jeon, SJ, Lim, YJ, Lim, JS, Choi, BS, Choi, IY, Yu, JH et al. (2013) Complete Genome Sequence of Mycobacterium bovis BCG Korea, the Korean Vaccine Strain for Substantial production. **Genome Announcements**. 1(2)

Kallimanis, A, Karabika, E, Mavromatis, K, Lapidus, A, LAButti, KM, Liolios, K, Ivanova, N, et al. (2011) Complete genome sequence of Mycobacterium sp. Strain (Spyr1) and reclassification to Mycobacterium gilvum Spyr1. **Stand Genomic Sci**. 5(1): 144 – 153. Doi: 10.4056/sigs.2265047

Kato-Maedo, M, Rhee, JT, Cingerase, TR, Salamon, H, Drenkow, J, Smittipat, N & Small, PM (2001) Comparing Genomes within the Species *Mycobacterium tuberculosis*. **Genome**

Research. 11: 547 - 554

Keller, PM, Böttger, EC & Sander, P (2008) Tuberculosis vaccine strain Mycobacterium bovis BCG Russia is a natural recA mutant. **BioMedCentral**. Doi: 10.1186/1471-2180-8-120

Kim, SJ, Kweon, O, Jones, RC, Edmondson, RD & Cerniglia, CE (2008) Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. **Biodegradation**, 19, 859 – 881

Kim, BJ, Choi, BS, Lim, JS, Choi, IJ, Kook, YH & Kim, BJ (2012) Complete Genome Sequence of Mycobacterium intracellulare Clinical Strain MOTT-64, Belonging to the INT1 genotype. **Journal of Bacteriology**. 194(12)

Kim, BJ, Kim, BR, Hong, SH, Seok, SH, Kook, YH & Kim, BJ (2013) Complete Genome Sequence of Mycobacterium massiliense Clinical Strain Asan 50594, Belonging to the Type II Genotype. **Genome Announcement** (4)

Koonin, EV, Aravind, L & Kondrashov, AS (2000) The impact of Comparative Genomics on our Understanding of Evolution. **Cell Press**. 101(6): 573 – 576

Kremer, K, Van Soolingen, D, Frothingham, R, Haas, WH, Hermans, OWM, Martin, C, Palittapongarnpim, P, et al. (1999) Comparison of Methods Based on Different Molecular Epidemiological Markers for Typing Mycobacterium tuberculosis Complex Strains: Interlaboratory Study of Discriminatory Power and Reproducibility. **Journal of Clinical Microbiology**. 37 (8): 2607 – 2618

Lee, WK, Baek, JH, Ryoo, SW & Yu, YG (2014) SWATH-based Comparative Proteomic Analysis of the Mycobacterium bovis BCG – Korea Strain. **kPubs**. Available from:

www.kpubs.org

Long, Q, Zhou, Q, Ji, L, Wu, J, Wang, W, Xie, J (2012) Mycobacterium smegmatis genomic characteristics associated with its saprophyte lifestyle. **Journal of Cell Biochemistry**. 113(10):3051-5

Madhavilatha, GK, Joseph, BV, Paul, LK, Kumar, RA, Hariharan, R & Mundayoor, S (2012) Whole-Genome Sequence of Two Clinical Isolates of Mycobacterium tuberculosis form Kerala, South India. **Journal of Bacteriology**. Vol 194 Number 16: 4430

Manca, C, Tsenoya, L, Barry III, CE, Bergtold, A, Freeman, S, Haslett, PAJ, Musser, JM, et al. (1999) Mycobacterium tuberculosis CDC1551 Induces a More Vigorous Host Response In Vivo and In Vitro But Is Not More Virulent Than Other Clinical Isolates. **Journal of Immunology**. 162(11): 6740 – 6746

Mdluli, K & Spigelman, M (2006) Novel targets for tuberculosis drug discovery. **Current opinion in Pharmacology**. 6: 459 – 667

Migliori, GB, Langendam, MW, D'Ambrosio, L, Centis, R, Blasi, F, Huitric, E, Manissero, D, van der Werf, MJ (2012) Protecting the tuberculosis drug pipeline: stating the case for the rational use of fluoroquinolones. **Clin Infect Dis**. 54:1379–80

Miyoshi-Akiyama, T, Matsumura, K, Iwai, H, Funatogawa, K & Kirikae, T (2012) Complete Annotated Genome Sequence of Mycobacterium tuberculosis Erdman. **Journal of Bacteriology**. 194: 2770

Monaghan, RL & Barrett, JF (2005) Antibacterial drug discovery – Then, now and the genomics future. **Biochemical Pharmacology**. 71(7): 901 – 909

Monot, M, Honoré, N, Garnier, T, Zidane, N, Sheerafi, D, Paniz-Mondolfi, A, Matsuoka, M et al. (2009) Comparative genomic and phylogeographic analysis of Mycobacterium leprae. **Nature genetics**. 41: 1282 – 1288

Mostowy, S, Onipede, A, Cagneux, S, Niemann, S, Kremer, K, Desmond, EP, Kato-Maedo, M & Behr, M (2004) Genomic Analyusis Distinguishes Mycobacterium africanum. Journal of **Clinical Mycrobiology**. 42(8): 3594 – 3599

Narayanan, S & Deshpande, U (2013) Whole-Genome Sequences of Four Clinical Isolates of Mycobacterium tuberculosis from Tamil Nadu, South India. **Genome Announcements**. Doi: 10.1128/genomeA.00186-13

Orduña, P, Cevallos, MA, De León, SP, Arvizu, A, Hernãndes-Gonzãles, IL, Mendoza-Hernãndes, G & López-Vidal, Y (2011) Genomic and proteomic analysis of Mycobacterium bovis BVG Mexico 1931 reveal a diverse immunogenic repertoire against tuberculosis infection. **BMC Genomics**. Available from: www.biomedcentral.com

Ouellet, H, Johnston, JB & Ortiz, PR (2011) Cholesterol catabolism as a therapeutic target in Mycobacterium tuberculosis. **Trends in Microbiology**. 19(11)

Price, SA & Wilson, LM (2003) Pathophysiology Clinical Concepts of Disease processes. 6th ed. Mosby. USA

Ripoll, F, Pasek, S, Schenowitz, C, Dossat, C, Barbe, V, Rottman, M, Macheras, E, Heym, B, et al. (2009) Non Mycobacterial Virulence Genes in the Genome of the Emerging Pathogen Mycobacterium abscessus. **PloS ONE**. 4(6): e5660

Roetzer, A, Diel, R, Kohl, TA, Rückert, C, Nübel, U, Blom, J, Wirth, T, et al. (2013) Whole

Genome Sequencing versus Traditional Genotyping for Investigation of a Mycobacterium tuberculosis Outbreak: A Longitudinal Molecular Epidemiological Study. **PLOS Medicine**. http://dx.doi.org/10.1371/journal.pmed.1001387

Seki, M, Honda, I, Fujita, I, Yano, I, Yamamoto, S & Koyama, A (2009) Whole genome sequence analysis of Mycobacterium bovis bacillus Calmette-Guèrin (BCG) Tokyo 172: a comparative study of BCG vaccine substrains. **PubMed**. http://www.ncbi.nlm.nih.gov/pubmed/19200449

Shafran, SD, Singer, J, Zarowny, DP, Phillips, P, Salit, I, Walmsley, SL, Fong, IW, et al. (1996) A comparison of two regimens for the treatment of Mycobacterium avium complex bacteremia in AIDS: rifabutin, ethambutol, and clarithromycin versus rifampin, ehtambutol, clofaziminem, and ciprofloxacin. Canadian HIV Trials Network Protocol 010 Study Group. **PubMed**. 335(6): 377 – 383

Shtratnikova, VY, Bragin, EY, Dovbnya, DB, Pekov, YA, Schelkunov, MI, Stritzhav, N, Ivashina, TV, et al. (2014) Complete Genome Sequence of Sterol-Transforming Mycobacterium neoaurum Strain VKM Ac- 1815D. **Genome Announcement**. 12(1)

Singh, Y, Kahli, S, Sowpati, DT, Rahman, SA Tyagi, A & Hasnain, SE (2014) Gene cooption in Mycobacteria and search for virulence attributes: Comparative proteomic analysis of Mycobacterium tuberculosis, Mycobacterium indicus pranii and other mycobacteria. **International Journal of Medical Microbiology** 304(5-6): 742 – 748

Steenken, Jr, W, Oatway, WH & Petroff, SA (1934) BIOLOGICAL STUDIES OF THE TUBERCLE BACILLUS III. DISSOCIATION AND PATHOGENECITY OF THE R AND S VARIANTS OF THE HUMAN TUBERCLE BACILLUS (H37). **J Exp Med**. 60(4): 515 –

540

Stinear, TP, Seemann, T, Harrison, PF, Jenkin, GA, Davies, JK, Johnson, PDR, Abdellah, Z, et al. (2008) Insights from the complete genome sequence of Mycobacteirum marinum on the evolution of Mycobacteirum tuberculosis. **Genome Research**. 18: 729 – 741

Stinear, TP, Seemann, T, Pidot, S, Frigui, W, Reysset, G, Garnier, T, Meurice, G, et al. (2007) Reductive evolution and niche adaptation inferred from the genome of Mycobacterium ulcerans, the causative agent of Buruli ulcer. **Genome Res**. 17(2): 192 – 200

Subhraveti, P, Ong, Q, Keseler, I, Holland, T, Kothari, A, Caspi, R & Karp, PD (2015) Summary of Mycobacterium africanum, Strain GM041182, version 19.1. BioCyc Database Collection. Available from: http://biocyc.org/MAFR572418/organism-summary?object=MAFR572418

Sullivan, MJ, Petty, NK & Beatson SA (2011) Easyfig: a genome comparison visualizer. **Bioinformatics**. 27(7): 1009 - 2010

Tobias, NJ, Doig, KD, Medema, MH, Chen, H, Haring, V, Moore, R, Seemann, T & Stinear, TP (2013) Complete genome sequence of the frog pathogen Mycobacterium ulcerans ecovar Liflandii. **Journal of Bacteriology**. 195(3): 556 – 564. Doi: 10.1128/JB.02132-12

Touchman, J (2010) Comparative genomics. **Nature Education Knowledge**. 3(10): 13

Vasconcellos, SEG, Huard, RC, Niemann, S, Kremer, K, Santos, AR, Suffys, PN & Ho, JI (2010) Direct genotypic profiles of the tow major clades of Mycobacterium africanum. **BMC Infectious Diseases**

Viktor, TC, De Haas, PE, Jordaan, AM, Van der Spuy, GD, Richardson, M, Van Soolingen,

D, Van Helden, PD & Warren, R (2004) Molecular Characteristics and Global Spread ofMycobacterium tuberculosis with a Western Cape F11 Genotype. **Journal of Clinical Microbiology**. 42(2): 769 – 772

Vissa, VD & Brennan, PJ (2001) The genome of Mycobacterium leprae: a minimal mycobacterial gene set. **Genome Biology**. 2(8), 1023.1 – 1023.5

Washington, W, Allen, S, Janda, W, Koneman, E, Procop, G, Schreckenberger, P and Woods, G (2006) Koneman's Color Atlas and Textbook of Diagnostic Microbiology. 6th ed. Lippincott Williams & Wilkins. USA

World Health Organization (2011). 2011/2012. Tuberculosis Global Facts. http://www.who.int/tb/publications/2011/factsheet_tb_2011.pdf

World Health Organization: Global Tuberculosis Report (2012) www.who.int/iris/bitstream/10665/75938/1/9789241564502_eng.pdf

Wu, C, Glasner, J, Collins, M, Naser, S & Talaat, A (2006) Whole Genome Plasticity among Mycobacterium avium Subspecies: Insights from Comparative Genomic Hybridizations †. **Journal of Bacteriology**. 711-723

Wynne, J, Seemann, T, Bulach, DM, Coutts, SA, Talaat, AM, Michalski, WP (2010) Resequencing the Mycobacterium avium subsp. paratuberculosis K10 genome: Improved annotation and revised genome sequence. **Journal of Bacteriology**. 192(23): 6319-6320

www.uniprot.org (2015) Article available from: www.uniprot.org/proteomes/UP000008896

You, Q, Verschoor, CP, Pant, SD, Macri, J, Kirby, GM & Karrow, NA (2012) Proteomic analysis of plasma form Holstein cows testing positive for Mycobacterium avium subsp

Paratuberculosis (MAP). **Veterinary Immunology and Immunopathology**. 148(3 – 4): 243 – 251

Zhang, Y, Chen, C, Liu, J, Deng, H, Pan, A, Zhang, L, Zhao, X, Huang, M, et al. (2011) Complete Genome Sequences of Mycobacterium tuberculosis Strains CCDC5079 and CCDC5080, Which belong to the Beijing family. **Journal of Bacteriology**. 193(19): 5591-5592

Zhang, Y, Post-Martins, K & Denkin, S (2006) New drug candidates and therapeutic targets for tuberculosis therapy. **Drug Discovery Today**. 11: 21 – 27

Zhang, ZY, Sun, SQ, Wang, ZL, Wen, ZL, Sun, QW, Zhu, ZQ, Song, YZ, et al. (2011) Complete Genome Sequence of a Novel Clinical Isolate, the Nontuberculous Mycobacterium Strain JDM601. **Journal of Bacteriology**. Doi: 10.1128/JB.0529-11

Zheng, H, Lu, L, Wang, B, Pu, S, Zhong, X, Zhu, G, Shi, W et al. (2008) Genetic Basis of Virulence Attenuation Revealed by Comparative Genomic Analysis of Mycobacterium tuberculosis Strain H37Ra versus H37Rv. **PLoS ONE**. l3(6): e2375 1 – 11

Zumla, AL & Grange, J (2002) Non-tuberculosis mycobacterial pulmonary infections. **Clinics in chest medicine**. (23): 369 – 376

CHAPTER 3

METHODOLOGY

3.1. Species and their databases used in this study

The whole genomes of 8 mycobacterial species belonging to six different categories were downloaded from TB database (http://genome.tbdb.org/tbdb_sysbio/MultiHome.html) (Table 3) and used in the study.

Table 3 Mycobacterial species used in the study for whole genome comparison*

No	Mycobacterial species	Reference
Mycobacterium tuberculosis complex (MTBC)		
1	*Mycobacterium tuberculosis H37Rv*	Zheng *et al.*, 2008
2	*Mycobacterium africanum* GM041182	Niemann *et al.*, 2004
3	*Mycobacterium bovis* AF2122/97	Garnier *et al.*, 2003
Mycobacterium chelonae-abscessus complex (MCAC)		
4	*Mycobacterium abscessus* ATCC 19977	Ripoll *et al.*, 2009
Mycobacterium avium complex (MAC)		
5	*Mycobacterium avium* 104	Horan *et al.*, 2006
Mycobacterium causing leprosy (MCL)		
6	*Mycobacterium leprae* TN	Murray *et al.*, 2003
Nontuberculosis Mycobacterium (NTM)		
7	*Mycobacterium marinum*	Stinear *et al.*, 2008
Saprophytes		
8	*Mycobacterium vanbaalenii* PYR-1	Kim *et al.*, 2008

**Species were grouped into different categories as described by Parvez et al. (2006).*

3.2. Genome comparison

3.2.1. Overview of gene homology

Comparative genomic approaches are increasingly employed to help both functional and evolutionary analyses (Li *et al*., 2003). The identification of orthologous groups in prokaryotes has allowed for cross referencing of genes from multiple species, which helps with genome annotation, protein family classification, studies of evolution and more importantly, with the identification of possible drug targets. Gene homology was carried out by OrthoMcl programs (Li *et al*., 2003). OrthoMCL provides a scalable method for constructing orthologous groups across multiple eukaryotic taxa and prokaryotes (Li *et al*., 2003). OrthoMCL (Li *et al.*, 2013) is one of the most widely used algorithms for predicting orthologous genes across multiple genomes. Similar to many other orthology prediction algorithms (Gabaldón, 2013; Kuzniar *et al.*, 2008], OrthoMCL is based on reciprocal best hits in all-against-all BLAST searches (Altschul *et al.*, 1990) of complete proteomes of the genomes followed by applying the Markov Clustering algorithm (Enright *et al.*, 2002) to a weighted graph constructed based on these best hits (Enright *et al.*, 2002; Dongen, 2000).

3.2.2. Software

BLAST searches were carried out using WU-BLAST 2.0 http://blast.wustl.edu/. INPARANOID software was obtained from http://kisac.cgb.ki.se/cgb/inparanoid/. MCL software was obtained fromhttp://micans.org/mcl/mcl-02-063/.

3.2.3. Identification of ortholog groups using INPARANOID

The source code of INPARANOID was modified to be applied to WU-BLAST results, with $-\log_{10}$(p-value) as the similarity score. Based on the distribution of p-values from BLAST

results, a score of 350 was assigned to a p-value of 0, and the similarity score between a sequence and itself was defined as 350. No match length coverage was required for a p-value of 1e-5, the cut-off p-value. The cut-off confidence value for inparalogs was 0.01.

3.2.4. Construction of "EGO Subset"

Orthologous groups containing gene index sequences from eukaryotes were extracted from the Eukaryotic gene Orthologs (EGO) database, and gene index sequences from other Gene Indices in these groups were discarded. Each gene index sequence was substituted with its best match in the proteome, determined by comparing with their respective proteomes using BLASTP. If more than one gene sequence was associated with a protein or if sequences were identical, the redundant sequences were removed so that each transformed group contained a distinct set of protein sequences. The "EGO subset" was constructed from the groups containing sequences from at least two of the species.

3.2.5. Evaluation of consistency with enzyme commission assignments

SWISS PROT proteins associated with enzyme commission (EC) number annotations were cross-referenced with sequences from the relevant protein using BLASTP (> 98% identity over > 98% of query sequence length). Only orthologous groups containing at least two sequences with EC assignments were investigated. Only if all EC-annotated sequences in this groups were assigned with the same EC number, was the group considered consistent with EC annotation.

3.3. Gene mapping

A total of 120 of the unique proteins identified in *M. tuberculosis* H37Rv were used in the study.

3.3.1. Genome mapping

Genome localization of the identified proteins was carried out using protein IDs at *M tuberculosis* database located at the KEGG website (http://www.genome.jp/kegg-bin/show_organism?category=Mycobacterium). Five proteins upstream of the protein were selected and five proteins downstream were selected and put in a table along with their KEGG gene codes. The genome map was downloaded, and all results were presented in a table.

3.3.2. Representing the gene maps

pDRAW32 DNA analysis software by AcaClone (http://www.acaclone.com/) was used to generate gene maps. The DNA sequence of the proteins with the up-stream and down-stream parts were downloaded from KEGG and used to deduce maps. The proteins were then annotated according to their size by analysing the whole downloaded sequence. The proteins were assigned to different colours according to KEGG colour codes.

3.4. References

Altschul, S, Gish, W, Miller, W, Myers, E & Lipman D (1990) Basic local alignment search tool. **J Mol Biol**. 215: 403 – 410

Dongen S (2000) Graph clustering by flow simulation. Centers Math. **Comput. Sci**. (CWI)

Enright, AJ, Dongen, SV, Ouzounis, CA (2002) An efficient algorithm for large-scale detection of protein families. **Nucleic Acids Res**. 30: 1575 – 1584. [Internet] Oxford University Press

Gabaldón, T, Koonin, EV (2013) Functional and evolutionary implications of gene orthology.

Nat Rev Genet Nat Res. 14: 360 – 366.

Garnier, T, Eiglemeier, K, Camus, JC, Medina, N, Mansoor, H, Pryor, M, Duthoy, S, et al. (2003) The complete genome sequence of Mycobacterium bovis. **PNAS**. 100(13): 7877 – 7882

Horan, KL, Freeman, R, Wergel, K, Semiet, M, Pfaller, S, Covert, TC *et al.* (2006) Isolation of the genome sequence strain Mycobacterium avium 104 from multiple patients over a 17 year period. **Journal of Clinical Microbiology,** 44: 783 – 789

Kim, SJ, Kweon, O, Jones, RC, Edmondson, RD & Cerniglia, CE (2008) Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. **Biodegradation**, 19: 859 – 881

Kuzniar, A, Van Ham, RCHJ, Pongor, S, & Leunissen, JAM (2008) The quest for orthologs: finding the corresponding gene across genomes. **Trends Genet**. 24: 539– 551

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. **Genome Res**. 13:2178–89

Murray, PR, Rosenthal, KS, Kobayashi, GS & Pfaller, MA (2003) *Medical Microbiology.* 4th ed. Mosby Inc. USA

Niemann, S, Kubica, T, Bonge, FC, Adjei, O, Browne, EN *et al.* (2004) The species *Mycobacterium africanum* in the light of new molecular markers. **Journal of Clinical Microbiology,** Vol 42, No 9, 3958 – 3962

Ripoll, F, Pasek,S, Schenowits, C, Dossat, C, Barbe, V, Rottman, M *et al.* (2009) Non mycobacterial virulence genes in the genome of the emerging pathogen *Mycobacterium abscessus*. **PLoS ONE,** 4. e5660

Stinear, TP, Seemann, T, Harrison, PF, Jenkin, GA, Davies, JK, Johnson, PDR, Abdellah, Z, et al. (2008) Insights from the complete genome sequence of Mycobacteirum marinum on the evolution of Mycobacteirum tuberculosis. **Genome Research**. 18: 729 – 741

Zheng, H, Lu, L, Wang, B, Pu, S, Zhong, X, Zhu, G, Shi, W et al. (2008) Genetic Basis of Virulence Attenuation Revealed by Comparative Genomic Analysis of Mycobacterium tuberculosis Strain H37Ra versus H37Rv. **PLoS ONE**. l3(6): e2375 1 – 11

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Genome comparison

Genomic comparison between 8 mycobacterial species form 6 different categories (Table 3) revealed genes unique to each of the 8 species when compared, as presented in Table 4.

Table 4 Number of genes unique to each of the mycobacterial species revealed by genomic comparison

No	Mycobacterial species	Number of unique genes*
Mycobacterium tuberculosis complex MTBC)		
1	*Mycobacterium tuberculosis* H37Rv	200
2	*Mycobacterium africanum* GM041182	72
3	*Mycobacterium bovis* AF2122/97	115
Mycobacterium chelonae-abscessus complex (MCAC)		
4	*Mycobacterium abscessus* ATCC 19977	1 519
Mycobacterium avium complex (MAC)		
5	*Mycobacterium avium* 104	1 154
Mycobacterium causing leprosy (MCL)		
6	*Mycobacterium leprae* TN	117
Nontuberculosis Mycobacterium (NTM)		
7	*Mycobacterium marinum*	1 087
Saprophytes		
8	*Mycobacterium marinum*	1 828

**Number of unique genes in each of the 8 species when compared with the method described in Section 3.2.*

The genes unique to each of the mycobacterial species are presented in Table 5 to Table 12.

Table 5 Genes unique to *Mycobacterium tuberculosis* when the genome was compared to the genomes of 7 other mycobacterial species

RVBD_0007Ac	RVBD_0908A	RVBD_1651c	RVBD_2346c	RVBD_3217Ac
RVBD_0007Bc	RVBD_0917A	RVBD_1674A	RVBD_2371A	RVBD_3233c
RVBD_0064A	RVBD_0946Ac	RVBD_1702Ac	RVBD_2396	RVBD_3312Xc
RVBD_0078Bc	RVBD_0959A	RVBD_1706Xc	RVBD_2427Ac	RVBD_3324A
RVBD_0109	RVBD_0962Ac	RVBD_1719A	RVBD_2490Ac	RVBD_3333A
RVBD_0124	RVBD_1004c	RVBD_1733A	RVBD_2490c	RVBD_3337
RVBD_0192A	RVBD_1028X	RVBD_1735Ac	RVBD_2512A	RVBD_3344c
RVBD_0193A	RVBD_1040Ac	RVBD_1755c	RVBD_2591	RVBD_3345c
RVBD_0229Ac	RVBD_1048Ac	RVBD_1759c	RVBD_2615c	RVBD_3347Ac
RVBD_0277Ac	RVBD_1051Ac	RVBD_1768	RVBD_2633Ac	RVBD_3367
RVBD_0278c	RVBD_1053A	RVBD_1792	RVBD_2634c	RVBD_3372A
RVBD_0279A	RVBD_1067c	RVBD_1792A	RVBD_2645	RVBD_3379A
RVBD_0279c	RVBD_1068A	RVBD_1818c	RVBD_2646	RVBD_3388
RVBD_0297	RVBD_1068c	RVBD_1840c	RVBD_2647	RVBD_3401Ac
RVBD_0378	RVBD_1087	RVBD_1886A	RVBD_2651c	RVBD_3425Ac
RVBD_0397A	RVBD_1091	RVBD_1894A	RVBD_2652c	RVBD_3426
RVBD_0454Ac	RVBD_1139Ac	RVBD_1961A	RVBD_2653c	RVBD_3426A
RVBD_0456Bc	RVBD_1147A	RVBD_1962Ac	RVBD_2654c	RVBD_3430Dc
RVBD_0500B	RVBD_1158c	RVBD_1963Ac	RVBD_2658c	RVBD_3507
RVBD_0532	RVBD_1170A	RVBD_1981Ac	RVBD_2741	RVBD_3508
RVBD_0571Ac	RVBD_1233c	RVBD_1982Ac	RVBD_2771Ac	RVBD_3511
RVBD_0578c	RVBD_1243c	RVBD_1991Ac	RVBD_2815Ac	RVBD_3512
RVBD_0590A	RVBD_1295A	RVBD_2063A	RVBD_2815Bc	RVBD_3514
RVBD_0597A	RVBD_1325c	RVBD_2077Bc	RVBD_2853	RVBD_3528A
RVBD_0609Bc	RVBD_1357Ac	RVBD_2098c	RVBD_2862A	RVBD_3590c
RVBD_0616A	RVBD_1396c	RVBD_2099c	RVBD_2873A	RVBD_3595c
RVBD_0666A	RVBD_1435c	RVBD_2104Ac	RVBD_2882Ac	RVBD_3653
RVBD_0668Ac	RVBD_1439A	RVBD_2126c	RVBD_2964B	RVBD_3654c
RVBD_0691A	RVBD_1441c	RVBD_2132Ac	RVBD_2975c	RVBD_3666A
RVBD_0742	RVBD_1443A	RVBD_2142Ac	RVBD_2986c	RVBD_3666B
RVBD_0744Ac	RVBD_1443Bc	RVBD_2160A	RVBD_3000Ac	RVBD_3697Ac
RVBD_0746	RVBD_1450c	RVBD_2162c	RVBD_3018Bc	RVBD_3705A
RVBD_0747	RVBD_1452c	RVBD_2231Ac	RVBD_3053Ac	RVBD_3706Ac
RVBD_0781	RVBD_1468c	RVBD_2231Bc	RVBD_3089Ac	RVBD_3724A
RVBD_0833	RVBD_1510A	RVBD_2261c	RVBD_3103c	RVBD_3725Ac
RVBD_0834c	RVBD_1520Ac	RVBD_2274Ac	RVBD_3135	RVBD_3749Ac
RVBD_0835Ac	RVBD_1549	RVBD_2292c	RVBD_3178A	RVBD_3845Ac
RVBD_0872c	RVBD_1551A	RVBD_2308Ac	RVBD_3190A	RVBD_3845Bc

RVBD_0887B	RVBD_1581Ac	RVBD_2308X	RVBD_3190B	RVBD_3852
RVBD_0891A	RVBD_1635A	RVBD_2309Xc	RVBD_3192A	RVBD_3901A

Table 6 Genes unique to *Mycobacterium africanum* when the genome was compared to the genomes of 7 other mycobacterial species

MAF_01100	MAF_10520	MAF_15230	MAF_26080	MAF_36120
MAF_02800	MAF_10800	MAF_15280	MAF_26340	MAF_36590
MAF_02810	MAF_10810	MAF_15990	MAF_26530	MAF_36600
MAF_02990	MAF_11000	MAF_16640	MAF_26890	MAF_36610
MAF_03800	MAF_11060	MAF_17900	MAF_28580	MAF_37140
MAF_03890	MAF_11670	MAF_18400	MAF_29910	MAF_37610
MAF_05070	MAF_11750	MAF_18620	MAF_30250	MAF_37620
MAF_05850	MAF_12520	MAF_20370	MAF_31100	MAF_37630
MAF_06100	MAF_12620	MAF_21120	MAF_33600	
MAF_07520	MAF_13490	MAF_21130	MAF_33820	
MAF_07560	MAF_14180	MAF_21380	MAF_34010	
MAF_07570	MAF_14570	MAF_21750	MAF_34430	
MAF_08340	MAF_14630	MAF_23600	MAF_35200	
MAF_08810	MAF_14720	MAF_23660	MAF_35240	
MAF_10140	MAF_14740	MAF_24100	MAF_36030	
MAF_10390	MAF_14900	MAF_25050	MAF_36080	

Table 7 Genes unique to *Mycobacterium bovis* when the genome was compared to the genomes of 7 other mycobacterial species

Mb0074	Mb1031c	Mb1743	Mb2368	Mb3088c
Mb0098c	Mb1057	Mb1789c	Mb2375c	Mb3130c
Mb0113	Mb1096c	Mb1790c	Mb2376c	Mb3151c
Mb0125c	Mb1097c	Mb1797	Mb2416	Mb3201c
Mb0129	Mb1116	Mb1831c	Mb2418	Mb3356
Mb0285c	Mb1121	Mb1838	Mb2448c	Mb3357
Mb0286c	Mb1189c	Mb1847	Mb2517c	Mb3377c
Mb0287c	Mb1265c	Mb1849c	Mb2518c	Mb3385c
Mb0305	Mb1275c	Mb1871c	Mb2541	Mb3402
Mb0385	Mb1291	Mb1908	Mb2607	Mb3420
Mb0545	Mb1346c	Mb1928	Mb2622	Mb3510c
Mb0593c	Mb1351c	Mb1953c	Mb2648c	Mb3537
Mb0609	Mb1360c	Mb1999	Mb2667c	Mb3538
Mb0718	Mb1431c	Mb2055c	Mb2683c	Mb3541
Mb0731	Mb1470c	Mb2099c	Mb2761	Mb3543

Mb0763	Mb1476c	Mb2125c	Mb2762c	Mb3621c
Mb0767	Mb1485c	Mb2150c	Mb2784c	Mb3626c
Mb0768	Mb1487c	Mb2184c	Mb2802	Mb3676
Mb0807	Mb1503c	Mb2186c	Mb2835	Mb3677
Mb0856	Mb1540	Mb2209	Mb2860c	Mb3678c
Mb0857c	Mb1590c	Mb2251	Mb2878	Mb3732c
Mb0896c	Mb1601	Mb2307c	Mb3010c	Mb3802c
Mb0957	Mb1679c	Mb2312	Mb3044c	Mb3827

Table 8 Genes unique to *Mycobacterium abscessus* when the genome was compared to the genomes of 7 other mycobacterial species

MAB_0007	MAB_0846	MAB_1788	MAB_2889	MAB_4109c
MAB_0008	MAB_0848c	MAB_1789	MAB_2899c	MAB_4122
MAB_0009	MAB_0849c	MAB_1790	MAB_2903	MAB_4125
MAB_0010c	MAB_0856c	MAB_1791	MAB_2907	MAB_4134
MAB_0011c	MAB_0869c	MAB_1792	MAB_2915c	MAB_4142c
MAB_0012c	MAB_0873	MAB_1793	MAB_2916	MAB_4143c
MAB_0014	MAB_0880	MAB_1794	MAB_2917	MAB_4171c
MAB_0015	MAB_0884c	MAB_1795	MAB_2918c	MAB_4176c
MAB_0017	MAB_0886c	MAB_1796	MAB_2919c	MAB_4190
MAB_0018c	MAB_0887c	MAB_1797	MAB_2920c	MAB_4191
MAB_0022c	MAB_0888c	MAB_1799	MAB_2929c	MAB_4192
MAB_0030	MAB_0889c	MAB_1800	MAB_2938	MAB_4193
MAB_0039c	MAB_0891c	MAB_1801	MAB_2941	MAB_4194c
MAB_0042	MAB_0892c	MAB_1802	MAB_2943c	MAB_4201c
MAB_0043	MAB_0893	MAB_1803	MAB_2944c	MAB_4202c
MAB_0045	MAB_0894c	MAB_1804	MAB_2945c	MAB_4211c
MAB_0050c	MAB_0897c	MAB_1805	MAB_2953	MAB_4227c
MAB_0051	MAB_0898	MAB_1806	MAB_2954	MAB_4228c
MAB_0052	MAB_0899c	MAB_1807	MAB_2956	MAB_4239
MAB_0054c	MAB_0900c	MAB_1808	MAB_2960	MAB_4241c
MAB_0055c	MAB_0902	MAB_1809	MAB_2964	MAB_4247
MAB_0057	MAB_0903	MAB_1810	MAB_2966c	MAB_4262
MAB_0059c	MAB_0904	MAB_1811	MAB_2967c	MAB_4263
MAB_0065c	MAB_0906	MAB_1812c	MAB_2969	MAB_4266c
MAB_0068	MAB_0907	MAB_1813	MAB_2970c	MAB_4267c
MAB_0083	MAB_0908	MAB_1814	MAB_2974c	MAB_4268c
MAB_0090c	MAB_0909	MAB_1815	MAB_2980	MAB_4274c
MAB_0093	MAB_0911	MAB_1818	MAB_2981c	MAB_4278
MAB_0099	MAB_0912	MAB_1819c	MAB_2982c	MAB_4279

MAB_0100	MAB_0913c	MAB_1821	MAB_2983c	MAB_4284c
MAB_0101	MAB_0914c	MAB_1822	MAB_2989	MAB_4287
MAB_0102	MAB_0915c	MAB_1823c	MAB_3010	MAB_4288
MAB_0108c	MAB_0916c	MAB_1825c	MAB_3011	MAB_4289
MAB_0109c	MAB_0917c	MAB_1826	MAB_3015	MAB_4290c
MAB_0112	MAB_0919	MAB_1828	MAB_3020c	MAB_4298c
MAB_0116	MAB_0920	MAB_1829	MAB_3021c	MAB_4302
MAB_0119c	MAB_0923c	MAB_1830	MAB_3022c	MAB_4303
MAB_0124	MAB_0940	MAB_1831	MAB_3031c	MAB_4308c
MAB_0125c	MAB_0941	MAB_1832	MAB_3049c	MAB_4309c
MAB_0127c	MAB_0958c	MAB_1834	MAB_3055c	MAB_4312
MAB_0134c	MAB_0963c	MAB_1836	MAB_3057	MAB_4313
MAB_0135	MAB_0965	MAB_1837c	MAB_3058	MAB_4314
MAB_0137	MAB_0967	MAB_1838	MAB_3064	MAB_4315c
MAB_0138	MAB_0968c	MAB_1839	MAB_3065c	MAB_4317
MAB_0141c	MAB_0971	MAB_1841c	MAB_3073	MAB_4318
MAB_0142c	MAB_0974	MAB_1844	MAB_3076	MAB_4321
MAB_0143c	MAB_0978	MAB_1845c	MAB_3079c	MAB_4324c
MAB_0144c	MAB_0979	MAB_1848	MAB_3092c	MAB_4325c
MAB_0145c	MAB_0981c	MAB_1851	MAB_3097	MAB_4337
MAB_0147c	MAB_0988c	MAB_1856	MAB_3101c	MAB_4342c
MAB_0148c	MAB_0993c	MAB_1866c	MAB_3102c	MAB_4343c
MAB_0149c	MAB_0995	MAB_1867	MAB_3104c	MAB_4350c
MAB_0152	MAB_0997c	MAB_1872c	MAB_3111	MAB_4354
MAB_0155	MAB_1000	MAB_1874	MAB_3112	MAB_4355
MAB_0158c	MAB_1001	MAB_1875c	MAB_3114	MAB_4356c
MAB_0161	MAB_1002	MAB_1882c	MAB_3115	MAB_4358c
MAB_0162c	MAB_1013	MAB_1883c	MAB_3126c	MAB_4360c
MAB_0169c	MAB_1014c	MAB_1888	MAB_3137	MAB_4361
MAB_0182c	MAB_1016c	MAB_1889	MAB_3141	MAB_4362
MAB_0184c	MAB_1017c	MAB_1890c	MAB_3145	MAB_4363c
MAB_0187	MAB_1018c	MAB_1891	MAB_3146c	MAB_4364c
MAB_0188c	MAB_1019c	MAB_1892	MAB_3152c	MAB_4366c
MAB_0194	MAB_1027	MAB_1895c	MAB_3153c	MAB_4368c
MAB_0198	MAB_1036	MAB_1896c	MAB_3154c	MAB_4371
MAB_0209	MAB_1040	MAB_1911c	MAB_3172c	MAB_4388
MAB_0210	MAB_1044c	MAB_1912c	MAB_3173c	MAB_4391
MAB_0211	MAB_1050c	MAB_1913	MAB_3182c	MAB_4396c
MAB_0212	MAB_1052c	MAB_1922	MAB_3190c	MAB_4398c
MAB_0221c	MAB_1058	MAB_1923c	MAB_3191	MAB_4399c
MAB_0222c	MAB_1066	MAB_1924	MAB_3194c	MAB_4400c

MAB_0223	MAB_1067	MAB_1925	MAB_3198c	MAB_4403c
MAB_0224	MAB_1091	MAB_1932c	MAB_3199c	MAB_4404
MAB_0225	MAB_1092	MAB_1935	MAB_3200	MAB_4405
MAB_0226	MAB_1093c	MAB_1936	MAB_3203	MAB_4406
MAB_0227	MAB_1094	MAB_1963	MAB_3206c	MAB_4414
MAB_0228	MAB_1095	MAB_1972c	MAB_3207	MAB_4415
MAB_0229	MAB_1096	MAB_1977	MAB_3208c	MAB_4416c
MAB_0230	MAB_1097	MAB_1989c	MAB_3209c	MAB_4417c
MAB_0232	MAB_1098	MAB_1993	MAB_3215c	MAB_4418
MAB_0233	MAB_1099c	MAB_2014c	MAB_3216c	MAB_4419
MAB_0234	MAB_1100c	MAB_2015c	MAB_3217	MAB_4424
MAB_0235	MAB_1101	MAB_2018	MAB_3218	MAB_4425
MAB_0236	MAB_1102	MAB_2021c	MAB_3219	MAB_4429
MAB_0238	MAB_1103	MAB_2022	MAB_3220	MAB_4431
MAB_0239	MAB_1104c	MAB_2024	MAB_3231	MAB_4432
MAB_0241	MAB_1105c	MAB_2026	MAB_3232c	MAB_4433c
MAB_0242	MAB_1106c	MAB_2027	MAB_3233	MAB_4439
MAB_0243	MAB_1107c	MAB_2028	MAB_3235c	MAB_4440c
MAB_0244	MAB_1108c	MAB_2029	MAB_3245	MAB_4447
MAB_0245c	MAB_1109c	MAB_2030	MAB_3247	MAB_4450
MAB_0246	MAB_1110	MAB_2031	MAB_3249	MAB_4451c
MAB_0247c	MAB_1112c	MAB_2032	MAB_3261c	MAB_4457
MAB_0249	MAB_1113c	MAB_2034	MAB_3262c	MAB_4459c
MAB_0251	MAB_1114	MAB_2036	MAB_3264c	MAB_4460
MAB_0253	MAB_1115	MAB_2038	MAB_3274	MAB_4461
MAB_0254c	MAB_1116	MAB_2042c	MAB_3275	MAB_4462c
MAB_0256c	MAB_1117c	MAB_2061c	MAB_3276	MAB_4463
MAB_0257c	MAB_1118c	MAB_2062c	MAB_3282c	MAB_4465
MAB_0259c	MAB_1125c	MAB_2063c	MAB_3297	MAB_4466
MAB_0269c	MAB_1126c	MAB_2065c	MAB_3305c	MAB_4467
MAB_0270c	MAB_1133c	MAB_2074	MAB_3306c	MAB_4470c
MAB_0271	MAB_1136	MAB_2085	MAB_3310c	MAB_4479
MAB_0272	MAB_1137c	MAB_2089	MAB_3314	MAB_4486
MAB_0273c	MAB_1138c	MAB_2090	MAB_3316	MAB_4488c
MAB_0283c	MAB_1146	MAB_2091	MAB_3318	MAB_4495
MAB_0284c	MAB_1150	MAB_2092	MAB_3337c	MAB_4497
MAB_0285	MAB_1151	MAB_2093	MAB_3344	MAB_4500
MAB_0287	MAB_1152	MAB_2094	MAB_3347	MAB_4501c
MAB_0289	MAB_1153	MAB_2112	MAB_3348	MAB_4503c
MAB_0291	MAB_1154	MAB_2113	MAB_3349	MAB_4509c
MAB_0292c	MAB_1155	MAB_2114c	MAB_3355	MAB_4510c

MAB_0294	MAB_1159	MAB_2118	MAB_3368c	MAB_4523
MAB_0295	MAB_1160	MAB_2125c	MAB_3370	MAB_4527
MAB_0296	MAB_1169	MAB_2127c	MAB_3373c	MAB_4528c
MAB_0297	MAB_1170	MAB_2149	MAB_3374c	MAB_4531
MAB_0298	MAB_1171c	MAB_2150c	MAB_3377	MAB_4532c
MAB_0301	MAB_1172c	MAB_2151c	MAB_3385	MAB_4533
MAB_0302	MAB_1173c	MAB_2156	MAB_3391	MAB_4546c
MAB_0303	MAB_1178c	MAB_2174	MAB_3392	MAB_4547
MAB_0304	MAB_1179c	MAB_2175	MAB_3393	MAB_4554c
MAB_0306c	MAB_1180c	MAB_2179	MAB_3402	MAB_4555c
MAB_0307c	MAB_1199	MAB_2180c	MAB_3407	MAB_4556c
MAB_0308c	MAB_1207	MAB_2182c	MAB_3411	MAB_4558c
MAB_0312c	MAB_1211c	MAB_2190	MAB_3412	MAB_4563c
MAB_0313c	MAB_1224	MAB_2202	MAB_3419	MAB_4564c
MAB_0314c	MAB_1238	MAB_2209c	MAB_3426c	MAB_4565c
MAB_0321c	MAB_1241c	MAB_2211c	MAB_3432c	MAB_4566c
MAB_0322c	MAB_1242c	MAB_2212	MAB_3434c	MAB_4567c
MAB_0325c	MAB_1243c	MAB_2213	MAB_3436	MAB_4576c
MAB_0327	MAB_1244c	MAB_2214c	MAB_3437c	MAB_4581c
MAB_0328	MAB_1245c	MAB_2222c	MAB_3439	MAB_4583c
MAB_0338c	MAB_1246c	MAB_2236c	MAB_3440c	MAB_4586c
MAB_0339	MAB_1247c	MAB_2246	MAB_3443	MAB_4590
MAB_0340	MAB_1249	MAB_2251	MAB_3445c	MAB_4594c
MAB_0341	MAB_1258c	MAB_2252	MAB_3446	MAB_4595c
MAB_0345	MAB_1261	MAB_2253	MAB_3447c	MAB_4596c
MAB_0348	MAB_1262	MAB_2255	MAB_3457	MAB_4597c
MAB_0349c	MAB_1263	MAB_2258	MAB_3461c	MAB_4598c
MAB_0350	MAB_1264	MAB_2259	MAB_3464	MAB_4599c
MAB_0352	MAB_1267	MAB_2260	MAB_3466c	MAB_4600c
MAB_0355	MAB_1268c	MAB_2268c	MAB_3469	MAB_4601c
MAB_0357c	MAB_1269	MAB_2275	MAB_3472	MAB_4611c
MAB_0358	MAB_1270	MAB_2277	MAB_3476c	MAB_4614
MAB_0360	MAB_1272	MAB_2278	MAB_3480	MAB_4620c
MAB_0361	MAB_1276c	MAB_2279	MAB_3488	MAB_4621c
MAB_0363c	MAB_1277	MAB_2280	MAB_3489c	MAB_4624
MAB_0364c	MAB_1280c	MAB_2281	MAB_3491	MAB_4627
MAB_0375	MAB_1281c	MAB_2284	MAB_3493	MAB_4628c
MAB_0376	MAB_1282	MAB_2286	MAB_3496	MAB_4631c
MAB_0377	MAB_1283c	MAB_2287	MAB_3497	MAB_4635
MAB_0379	MAB_1286c	MAB_2288	MAB_3500	MAB_4639
MAB_0380	MAB_1295c	MAB_2289	MAB_3505c	MAB_4642c

MAB_0381	MAB_1299c	MAB_2290	MAB_3506c	MAB_4643c
MAB_0396c	MAB_1300c	MAB_2292c	MAB_3509c	MAB_4644c
MAB_0397c	MAB_1301c	MAB_2311	MAB_3529	MAB_4645
MAB_0403	MAB_1303c	MAB_2313	MAB_3553	MAB_4648
MAB_0404c	MAB_1309c	MAB_2325	MAB_3555	MAB_4649c
MAB_0405c	MAB_1311	MAB_2327c	MAB_3556	MAB_4656
MAB_0415	MAB_1314	MAB_2332c	MAB_3557	MAB_4657c
MAB_0434c	MAB_1320c	MAB_2349	MAB_3558	MAB_4658
MAB_0436	MAB_1321	MAB_2350c	MAB_3559c	MAB_4662c
MAB_0437c	MAB_1322	MAB_2353	MAB_3567c	MAB_4664
MAB_0438c	MAB_1323	MAB_2355c	MAB_3568c	MAB_4666c
MAB_0439	MAB_1334c	MAB_2374c	MAB_3569c	MAB_4668c
MAB_0445	MAB_1338	MAB_2375	MAB_3570c	MAB_4670c
MAB_0446c	MAB_1340	MAB_2376c	MAB_3571c	MAB_4672c
MAB_0447c	MAB_1347c	MAB_2377	MAB_3572c	MAB_4673c
MAB_0448c	MAB_1353c	MAB_2378c	MAB_3573c	MAB_4674c
MAB_0452c	MAB_1370c	MAB_2380c	MAB_3574c	MAB_4675c
MAB_0455	MAB_1377c	MAB_2381	MAB_3575c	MAB_4678
MAB_0456	MAB_1379c	MAB_2383	MAB_3577c	MAB_4679
MAB_0457	MAB_1380	MAB_2385	MAB_3578c	MAB_4682c
MAB_0459c	MAB_1381	MAB_2387	MAB_3579c	MAB_4683c
MAB_0460	MAB_1382	MAB_2390	MAB_3582	MAB_4684
MAB_0461	MAB_1383	MAB_2391	MAB_3584	MAB_4685
MAB_0464	MAB_1390c	MAB_2392	MAB_3592c	MAB_4686
MAB_0465	MAB_1391	MAB_2393c	MAB_3600c	MAB_4687
MAB_0466	MAB_1392c	MAB_2394	MAB_3610c	MAB_4696c
MAB_0467	MAB_1399	MAB_2396	MAB_3621c	MAB_4706c
MAB_0468	MAB_1400	MAB_2416c	MAB_3622c	MAB_4707c
MAB_0469c	MAB_1405	MAB_2419c	MAB_3623	MAB_4710c
MAB_0470c	MAB_1408c	MAB_2420c	MAB_3624c	MAB_4711
MAB_0475	MAB_1412	MAB_2421c	MAB_3629c	MAB_4715c
MAB_0476c	MAB_1418	MAB_2422c	MAB_3630c	MAB_4719c
MAB_0477	MAB_1423	MAB_2437c	MAB_3638c	MAB_4720c
MAB_0479	MAB_1425	MAB_2439	MAB_3640c	MAB_4721
MAB_0480	MAB_1430c	MAB_2441	MAB_3645	MAB_4723c
MAB_0481c	MAB_1431	MAB_2442c	MAB_3663	MAB_4726
MAB_0482c	MAB_1432	MAB_2444c	MAB_3664	MAB_4730
MAB_0483c	MAB_1439c	MAB_2445	MAB_3665	MAB_4732c
MAB_0484	MAB_1440c	MAB_2448	MAB_3666	MAB_4736
MAB_0485	MAB_1466c	MAB_2451c	MAB_3679c	MAB_4739
MAB_0492	MAB_1473c	MAB_2459	MAB_3690	MAB_4741c

MAB_0493	MAB_1486c	MAB_2461	MAB_3695	MAB_4744c
MAB_0494	MAB_1496c	MAB_2462	MAB_3696c	MAB_4757
MAB_0495c	MAB_1501	MAB_2463	MAB_3697	MAB_4759
MAB_0506	MAB_1502	MAB_2464	MAB_3700c	MAB_4767
MAB_0512	MAB_1503	MAB_2468	MAB_3704	MAB_4772
MAB_0514	MAB_1504	MAB_2469	MAB_3706	MAB_4773
MAB_0517	MAB_1505	MAB_2475c	MAB_3707c	MAB_4774c
MAB_0524c	MAB_1507	MAB_2478c	MAB_3708c	MAB_4778c
MAB_0531c	MAB_1510	MAB_2479c	MAB_3709c	MAB_4779
MAB_0532	MAB_1519	MAB_2481	MAB_3714c	MAB_4783
MAB_0539	MAB_1525c	MAB_2483c	MAB_3715c	MAB_4784
MAB_0547c	MAB_1526	MAB_2484c	MAB_3716	MAB_4786c
MAB_0548c	MAB_1527	MAB_2485c	MAB_3717	MAB_4789c
MAB_0549c	MAB_1529c	MAB_2486c	MAB_3728c	MAB_4790
MAB_0565c	MAB_1535c	MAB_2487	MAB_3729c	MAB_4791c
MAB_0566c	MAB_1536	MAB_2492	MAB_3730c	MAB_4793c
MAB_0578c	MAB_1539c	MAB_2493c	MAB_3742c	MAB_4795c
MAB_0581c	MAB_1541	MAB_2500	MAB_3745c	MAB_4796
MAB_0588c	MAB_1542c	MAB_2506c	MAB_3746c	MAB_4797
MAB_0589	MAB_1545c	MAB_2514c	MAB_3754c	MAB_4798
MAB_0590c	MAB_1550	MAB_2516c	MAB_3755c	MAB_4799c
MAB_0601	MAB_1551c	MAB_2517c	MAB_3764	MAB_4802
MAB_0602c	MAB_1553c	MAB_2518	MAB_3767c	MAB_4803
MAB_0619	MAB_1555	MAB_2521c	MAB_3786c	MAB_4804
MAB_0620c	MAB_1556	MAB_2525c	MAB_3790	MAB_4805
MAB_0627c	MAB_1564c	MAB_2526c	MAB_3791	MAB_4808
MAB_0628	MAB_1576	MAB_2527	MAB_3801c	MAB_4809c
MAB_0629	MAB_1578c	MAB_2532	MAB_3802c	MAB_4810c
MAB_0648	MAB_1586c	MAB_2533	MAB_3809c	MAB_4811c
MAB_0652	MAB_1589	MAB_2536	MAB_3822	MAB_4812c
MAB_0664	MAB_1596	MAB_2538c	MAB_3823	MAB_4813c
MAB_0677c	MAB_1600c	MAB_2541c	MAB_3864c	MAB_4814
MAB_0679c	MAB_1609c	MAB_2543c	MAB_3865	MAB_4815
MAB_0683c	MAB_1614	MAB_2544c	MAB_3867	MAB_4817
MAB_0704c	MAB_1615	MAB_2547c	MAB_3872c	MAB_4818
MAB_0705	MAB_1616	MAB_2549	MAB_3873c	MAB_4819c
MAB_0706c	MAB_1631c	MAB_2553	MAB_3889	MAB_4820
MAB_0712c	MAB_1638	MAB_2554	MAB_3905	MAB_4821c
MAB_0714c	MAB_1642	MAB_2556c	MAB_3906c	MAB_4823c
MAB_0715c	MAB_1647c	MAB_2563	MAB_3909	MAB_4824c
MAB_0717	MAB_1648c	MAB_2564	MAB_3910	MAB_4825c

MAB_0719	MAB_1649c	MAB_2568	MAB_3911	MAB_4826
MAB_0723c	MAB_1650c	MAB_2569c	MAB_3916c	MAB_4827c
MAB_0733	MAB_1657c	MAB_2571c	MAB_3918c	MAB_4828c
MAB_0734	MAB_1659c	MAB_2572c	MAB_3919c	MAB_4829c
MAB_0735	MAB_1660c	MAB_2574c	MAB_3921c	MAB_4831c
MAB_0736	MAB_1673	MAB_2580c	MAB_3922c	MAB_4832c
MAB_0737c	MAB_1693	MAB_2583c	MAB_3923	MAB_4833c
MAB_0755c	MAB_1701	MAB_2585	MAB_3924c	MAB_4834
MAB_0760c	MAB_1705c	MAB_2588	MAB_3936c	MAB_4835
MAB_0761c	MAB_1711c	MAB_2592c	MAB_3937	MAB_4836c
MAB_0762	MAB_1720	MAB_2593	MAB_3944	MAB_4838c
MAB_0763c	MAB_1721	MAB_2594	MAB_3953	MAB_4843
MAB_0764c	MAB_1722	MAB_2599	MAB_3955	MAB_4845
MAB_0765	MAB_1724c	MAB_2600	MAB_3957	MAB_4846
MAB_0766	MAB_1728	MAB_2601	MAB_3959c	MAB_4847c
MAB_0769c	MAB_1729	MAB_2611	MAB_3960c	MAB_4848c
MAB_0770	MAB_1730	MAB_2612	MAB_3962	MAB_4855c
MAB_0771c	MAB_1731	MAB_2613	MAB_3963c	MAB_4856c
MAB_0772c	MAB_1732	MAB_2628c	MAB_3964	MAB_4860c
MAB_0773	MAB_1733	MAB_2634	MAB_3965	MAB_4861c
MAB_0774c	MAB_1734	MAB_2637c	MAB_3966c	MAB_4865
MAB_0775	MAB_1735	MAB_2641c	MAB_3967	MAB_4866c
MAB_0776	MAB_1736	MAB_2649	MAB_3968	MAB_4867c
MAB_0777	MAB_1737	MAB_2652c	MAB_3969	MAB_4868c
MAB_0778	MAB_1738	MAB_2653c	MAB_3979c	MAB_4869
MAB_0779	MAB_1739	MAB_2654	MAB_3981c	MAB_4870c
MAB_0780	MAB_1740	MAB_2655	MAB_3983c	MAB_4871
MAB_0781	MAB_1741	MAB_2656	MAB_3985c	MAB_4872
MAB_0782	MAB_1742	MAB_2658	MAB_3986c	MAB_4873c
MAB_0785	MAB_1743	MAB_2660	MAB_3988c	MAB_4878c
MAB_0786	MAB_1744	MAB_2679	MAB_4000	MAB_4880c
MAB_0787	MAB_1745	MAB_2680c	MAB_4008c	MAB_4881c
MAB_0788	MAB_1746	MAB_2681	MAB_4009c	MAB_4883c
MAB_0789	MAB_1747	MAB_2693	MAB_4012c	MAB_4884c
MAB_0790	MAB_1748	MAB_2694	MAB_4015c	MAB_4885
MAB_0791	MAB_1749	MAB_2697c	MAB_4016c	MAB_4886
MAB_0792	MAB_1750	MAB_2708	MAB_4017c	MAB_4887
MAB_0793	MAB_1751	MAB_2709c	MAB_4022c	MAB_4888
MAB_0794	MAB_1752	MAB_2733c	MAB_4023c	MAB_4889
MAB_0795	MAB_1753	MAB_2734c	MAB_4025c	MAB_4890
MAB_0796c	MAB_1754	MAB_2738c	MAB_4026c	MAB_4892c

MAB_0797c	MAB_1755	MAB_2740c	MAB_4027	MAB_4893c
MAB_0798c	MAB_1756	MAB_2741	MAB_4029	MAB_4903
MAB_0799c	MAB_1757	MAB_2742	MAB_4030	MAB_4908c
MAB_0800	MAB_1758	MAB_2756c	MAB_4031	MAB_4912c
MAB_0801	MAB_1759	MAB_2764c	MAB_4033	MAB_4914c
MAB_0802	MAB_1760	MAB_2765	MAB_4035	MAB_4915c
MAB_0803	MAB_1761	MAB_2766	MAB_4036	MAB_4921c
MAB_0804	MAB_1762	MAB_2767	MAB_4037	MAB_4925
MAB_0808c	MAB_1763	MAB_2770	MAB_4038c	MAB_4926
MAB_0809c	MAB_1764	MAB_2771c	MAB_4039c	MAB_4927
MAB_0810c	MAB_1765	MAB_2772c	MAB_4040c	MAB_4931
MAB_0811	MAB_1766	MAB_2773c	MAB_4041c	MAB_4933
MAB_0812	MAB_1767	MAB_2775c	MAB_4042c	MAB_4943c
MAB_0813c	MAB_1768	MAB_2776	MAB_4043c	MAB_4944c
MAB_0815	MAB_1769	MAB_2785	MAB_4044c	MAB_4945c
MAB_0816	MAB_1770	MAB_2786	MAB_4045c	MAB_4946
MAB_0817	MAB_1771	MAB_2787c	MAB_4046c	MAB_4955c
MAB_0818	MAB_1773c	MAB_2789c	MAB_4050c	MAB_p05c
MAB_0828	MAB_1774	MAB_2815c	MAB_4053c	MAB_p06
MAB_0829	MAB_1775	MAB_2816c	MAB_4058c	MAB_p07
MAB_0836c	MAB_1776	MAB_2817c	MAB_4063c	MAB_p08
MAB_0837c	MAB_1777	MAB_2825	MAB_4064	MAB_p09
MAB_0838c	MAB_1778	MAB_2846	MAB_4068	MAB_p16c
MAB_0839c	MAB_1779	MAB_2847c	MAB_4071	MAB_p17
MAB_0840c	MAB_1780	MAB_2852c	MAB_4079	MAB_p18c
MAB_0841	MAB_1781	MAB_2855c	MAB_4091	MAB_p19c
MAB_0842	MAB_1782	MAB_2870	MAB_4092	MAB_p20c
MAB_0843	MAB_1783	MAB_2886c	MAB_4093	
MAB_0844	MAB_1787	MAB_2887	MAB_4101	

Table 9 **Genes unique to *Mycobacterium avium* when the genome was compared to the genomes of 7 other mycobacterial species**

MAV_0026	MAV_0920	MAV_1942	MAV_2667	MAV_4014
MAV_0031	MAV_0921	MAV_1943	MAV_2669	MAV_4020
MAV_0032	MAV_0936	MAV_1944	MAV_2670	MAV_4049
MAV_0036	MAV_0943	MAV_1948	MAV_2672	MAV_4051
MAV_0047	MAV_0955	MAV_1949	MAV_2674	MAV_4055
MAV_0063	MAV_0956	MAV_1950	MAV_2677	MAV_4064
MAV_0064	MAV_0957	MAV_1952	MAV_2678	MAV_4065
MAV_0065	MAV_0968	MAV_1953	MAV_2689	MAV_4072

MAV_0078	MAV_0974	MAV_1955	MAV_2693	MAV_4073
MAV_0085	MAV_0982	MAV_1956	MAV_2697	MAV_4074
MAV_0086	MAV_1008	MAV_1958	MAV_2706	MAV_4075
MAV_0090	MAV_1024	MAV_1959	MAV_2709	MAV_4076
MAV_0092	MAV_1025	MAV_1961	MAV_2713	MAV_4077
MAV_0093	MAV_1036	MAV_1962	MAV_2714	MAV_4081
MAV_0103	MAV_1037	MAV_1963	MAV_2717	MAV_4083
MAV_0106	MAV_1038	MAV_1964	MAV_2719	MAV_4084
MAV_0107	MAV_1039	MAV_1966	MAV_2720	MAV_4085
MAV_0113	MAV_1040	MAV_1967	MAV_2721	MAV_4086
MAV_0114	MAV_1053	MAV_1968	MAV_2722	MAV_4087
MAV_0115	MAV_1054	MAV_1971	MAV_2724	MAV_4088
MAV_0116	MAV_1055	MAV_1973	MAV_2726	MAV_4090
MAV_0117	MAV_1070	MAV_1975	MAV_2728	MAV_4091
MAV_0118	MAV_1082	MAV_1976	MAV_2729	MAV_4092
MAV_0125	MAV_1107	MAV_1977	MAV_2730	MAV_4096
MAV_0127	MAV_1110	MAV_1978	MAV_2731	MAV_4097
MAV_0129	MAV_1111	MAV_1979	MAV_2732	MAV_4103
MAV_0131	MAV_1112	MAV_1980	MAV_2733	MAV_4104
MAV_0138	MAV_1117	MAV_1981	MAV_2734	MAV_4111
MAV_0140	MAV_1137	MAV_1982	MAV_2735	MAV_4115
MAV_0141	MAV_1139	MAV_1983	MAV_2739	MAV_4116
MAV_0143	MAV_1140	MAV_1985	MAV_2744	MAV_4117
MAV_0148	MAV_1141	MAV_1988	MAV_2760	MAV_4120
MAV_0150	MAV_1143	MAV_1990	MAV_2763	MAV_4122
MAV_0151	MAV_1144	MAV_1991	MAV_2771	MAV_4123
MAV_0165	MAV_1174	MAV_1992	MAV_2780	MAV_4124
MAV_0176	MAV_1183	MAV_1993	MAV_2785	MAV_4127
MAV_0177	MAV_1231	MAV_1994	MAV_2789	MAV_4128
MAV_0183	MAV_1233	MAV_1996	MAV_2792	MAV_4133
MAV_0190	MAV_1236	MAV_1997	MAV_2794	MAV_4134
MAV_0201	MAV_1243	MAV_1998	MAV_2795	MAV_4135
MAV_0213	MAV_1245	MAV_1999	MAV_2815	MAV_4157
MAV_0219	MAV_1246	MAV_2001	MAV_2824	MAV_4174
MAV_0226	MAV_1266	MAV_2005	MAV_2825	MAV_4176
MAV_0228	MAV_1267	MAV_2006	MAV_2830	MAV_4179
MAV_0254	MAV_1268	MAV_2017	MAV_2866	MAV_4187
MAV_0255	MAV_1269	MAV_2027	MAV_2877	MAV_4189
MAV_0256	MAV_1271	MAV_2030	MAV_2883	MAV_4205
MAV_0257	MAV_1272	MAV_2043	MAV_2904	MAV_4229
MAV_0258	MAV_1273	MAV_2057	MAV_2906	MAV_4236

MAV_0259	MAV_1274	MAV_2063	MAV_2907	MAV_4259
MAV_0260	MAV_1297	MAV_2064	MAV_2912	MAV_4270
MAV_0261	MAV_1302	MAV_2065	MAV_2913	MAV_4272
MAV_0262	MAV_1313	MAV_2066	MAV_2927	MAV_4276
MAV_0263	MAV_1315	MAV_2068	MAV_2929	MAV_4290
MAV_0264	MAV_1319	MAV_2069	MAV_2930	MAV_4293
MAV_0265	MAV_1322	MAV_2070	MAV_2937	MAV_4303
MAV_0266	MAV_1323	MAV_2073	MAV_2939	MAV_4320
MAV_0267	MAV_1336	MAV_2077	MAV_2941	MAV_4347
MAV_0268	MAV_1338	MAV_2094	MAV_2943	MAV_4362
MAV_0269	MAV_1341	MAV_2095	MAV_2945	MAV_4363
MAV_0270	MAV_1372	MAV_2102	MAV_2949	MAV_4378
MAV_0271	MAV_1382	MAV_2105	MAV_2950	MAV_4379
MAV_0272	MAV_1389	MAV_2107	MAV_2951	MAV_4391
MAV_0273	MAV_1390	MAV_2113	MAV_2953	MAV_4408
MAV_0274	MAV_1391	MAV_2115	MAV_2954	MAV_4411
MAV_0275	MAV_1392	MAV_2116	MAV_2955	MAV_4426
MAV_0276	MAV_1393	MAV_2125	MAV_2966	MAV_4440
MAV_0277	MAV_1394	MAV_2128	MAV_2967	MAV_4460
MAV_0278	MAV_1395	MAV_2134	MAV_2972	MAV_4473
MAV_0279	MAV_1396	MAV_2135	MAV_2975	MAV_4480
MAV_0280	MAV_1399	MAV_2136	MAV_2981	MAV_4483
MAV_0281	MAV_1412	MAV_2137	MAV_2983	MAV_4484
MAV_0282	MAV_1422	MAV_2138	MAV_2986	MAV_4493
MAV_0283	MAV_1442	MAV_2141	MAV_2988	MAV_4509
MAV_0284	MAV_1444	MAV_2148	MAV_2989	MAV_4535
MAV_0285	MAV_1449	MAV_2155	MAV_2998	MAV_4536
MAV_0286	MAV_1451	MAV_2159	MAV_3002	MAV_4558
MAV_0287	MAV_1453	MAV_2162	MAV_3003	MAV_4562
MAV_0288	MAV_1456	MAV_2163	MAV_3004	MAV_4570
MAV_0289	MAV_1461	MAV_2164	MAV_3005	MAV_4572
MAV_0290	MAV_1464	MAV_2165	MAV_3006	MAV_4573
MAV_0291	MAV_1465	MAV_2168	MAV_3007	MAV_4585
MAV_0292	MAV_1466	MAV_2170	MAV_3008	MAV_4594
MAV_0293	MAV_1467	MAV_2200	MAV_3011	MAV_4607
MAV_0294	MAV_1469	MAV_2219	MAV_3013	MAV_4613
MAV_0295	MAV_1470	MAV_2225	MAV_3014	MAV_4622
MAV_0296	MAV_1471	MAV_2226	MAV_3023	MAV_4634
MAV_0297	MAV_1474	MAV_2231	MAV_3034	MAV_4650
MAV_0298	MAV_1476	MAV_2232	MAV_3040	MAV_4686
MAV_0299	MAV_1477	MAV_2233	MAV_3041	MAV_4688

MAV_0303	MAV_1478	MAV_2236	MAV_3046	MAV_4693
MAV_0313	MAV_1480	MAV_2242	MAV_3051	MAV_4697
MAV_0327	MAV_1483	MAV_2248	MAV_3057	MAV_4706
MAV_0328	MAV_1485	MAV_2249	MAV_3058	MAV_4708
MAV_0330	MAV_1486	MAV_2250	MAV_3060	MAV_4712
MAV_0331	MAV_1487	MAV_2251	MAV_3061	MAV_4728
MAV_0333	MAV_1488	MAV_2252	MAV_3090	MAV_4729
MAV_0337	MAV_1489	MAV_2253	MAV_3091	MAV_4734
MAV_0339	MAV_1490	MAV_2256	MAV_3092	MAV_4735
MAV_0347	MAV_1491	MAV_2257	MAV_3094	MAV_4753
MAV_0350	MAV_1499	MAV_2259	MAV_3095	MAV_4757
MAV_0352	MAV_1500	MAV_2260	MAV_3098	MAV_4759
MAV_0354	MAV_1501	MAV_2263	MAV_3101	MAV_4764
MAV_0357	MAV_1502	MAV_2265	MAV_3102	MAV_4766
MAV_0361	MAV_1503	MAV_2274	MAV_3142	MAV_4769
MAV_0363	MAV_1505	MAV_2278	MAV_3144	MAV_4771
MAV_0364	MAV_1512	MAV_2302	MAV_3148	MAV_4779
MAV_0365	MAV_1513	MAV_2326	MAV_3149	MAV_4794
MAV_0367	MAV_1535	MAV_2339	MAV_3159	MAV_4796
MAV_0370	MAV_1562	MAV_2346	MAV_3197	MAV_4798
MAV_0374	MAV_1564	MAV_2348	MAV_3213	MAV_4799
MAV_0396	MAV_1565	MAV_2368	MAV_3216	MAV_4800
MAV_0397	MAV_1588	MAV_2369	MAV_3219	MAV_4801
MAV_0409	MAV_1589	MAV_2371	MAV_3223	MAV_4803
MAV_0428	MAV_1599	MAV_2373	MAV_3241	MAV_4804
MAV_0441	MAV_1616	MAV_2375	MAV_3256	MAV_4809
MAV_0463	MAV_1617	MAV_2412	MAV_3257	MAV_4810
MAV_0472	MAV_1621	MAV_2415	MAV_3259	MAV_4814
MAV_0474	MAV_1622	MAV_2417	MAV_3267	MAV_4815
MAV_0476	MAV_1623	MAV_2419	MAV_3268	MAV_4825
MAV_0477	MAV_1624	MAV_2428	MAV_3281	MAV_4827
MAV_0478	MAV_1633	MAV_2429	MAV_3284	MAV_4833
MAV_0479	MAV_1637	MAV_2436	MAV_3286	MAV_4835
MAV_0480	MAV_1645	MAV_2441	MAV_3287	MAV_4836
MAV_0484	MAV_1646	MAV_2449	MAV_3311	MAV_4837
MAV_0485	MAV_1647	MAV_2457	MAV_3312	MAV_4840
MAV_0486	MAV_1649	MAV_2462	MAV_3333	MAV_4843
MAV_0488	MAV_1656	MAV_2464	MAV_3334	MAV_4844
MAV_0489	MAV_1662	MAV_2474	MAV_3335	MAV_4868
MAV_0490	MAV_1679	MAV_2481	MAV_3344	MAV_4872
MAV_0491	MAV_1681	MAV_2490	MAV_3345	MAV_4883

MAV_0492	MAV_1683	MAV_2491	MAV_3348	MAV_4889
MAV_0493	MAV_1693	MAV_2493	MAV_3353	MAV_4890
MAV_0494	MAV_1720	MAV_2494	MAV_3355	MAV_4898
MAV_0497	MAV_1721	MAV_2508	MAV_3356	MAV_4905
MAV_0498	MAV_1739	MAV_2512	MAV_3361	MAV_4921
MAV_0499	MAV_1740	MAV_2514	MAV_3379	MAV_4925
MAV_0501	MAV_1742	MAV_2520	MAV_3380	MAV_4928
MAV_0502	MAV_1775	MAV_2522	MAV_3403	MAV_4929
MAV_0508	MAV_1788	MAV_2525	MAV_3404	MAV_4933
MAV_0514	MAV_1799	MAV_2527	MAV_3407	MAV_4947
MAV_0536	MAV_1809	MAV_2530	MAV_3418	MAV_4959
MAV_0541	MAV_1810	MAV_2531	MAV_3419	MAV_4962
MAV_0547	MAV_1815	MAV_2540	MAV_3420	MAV_4972
MAV_0605	MAV_1818	MAV_2541	MAV_3421	MAV_4979
MAV_0606	MAV_1819	MAV_2544	MAV_3423	MAV_4991
MAV_0622	MAV_1821	MAV_2545	MAV_3424	MAV_4993
MAV_0628	MAV_1822	MAV_2546	MAV_3425	MAV_4999
MAV_0649	MAV_1823	MAV_2549	MAV_3426	MAV_5004
MAV_0675	MAV_1824	MAV_2550	MAV_3427	MAV_5023
MAV_0679	MAV_1826	MAV_2555	MAV_3434	MAV_5024
MAV_0680	MAV_1827	MAV_2558	MAV_3443	MAV_5026
MAV_0682	MAV_1828	MAV_2560	MAV_3444	MAV_5028
MAV_0684	MAV_1829	MAV_2561	MAV_3445	MAV_5031
MAV_0685	MAV_1830	MAV_2562	MAV_3446	MAV_5032
MAV_0686	MAV_1831	MAV_2563	MAV_3458	MAV_5033
MAV_0687	MAV_1832	MAV_2565	MAV_3463	MAV_5034
MAV_0688	MAV_1833	MAV_2566	MAV_3508	MAV_5036
MAV_0689	MAV_1834	MAV_2567	MAV_3524	MAV_5038
MAV_0690	MAV_1836	MAV_2569	MAV_3531	MAV_5041
MAV_0691	MAV_1837	MAV_2571	MAV_3534	MAV_5044
MAV_0693	MAV_1839	MAV_2572	MAV_3537	MAV_5045
MAV_0696	MAV_1840	MAV_2573	MAV_3538	MAV_5046
MAV_0697	MAV_1841	MAV_2575	MAV_3539	MAV_5047
MAV_0712	MAV_1842	MAV_2578	MAV_3544	MAV_5049
MAV_0762	MAV_1846	MAV_2579	MAV_3545	MAV_5051
MAV_0763	MAV_1847	MAV_2580	MAV_3546	MAV_5052
MAV_0770	MAV_1848	MAV_2581	MAV_3547	MAV_5055
MAV_0775	MAV_1850	MAV_2582	MAV_3549	MAV_5056
MAV_0779	MAV_1852	MAV_2583	MAV_3557	MAV_5057
MAV_0780	MAV_1853	MAV_2584	MAV_3567	MAV_5058
MAV_0781	MAV_1854	MAV_2585	MAV_3569	MAV_5059

MAV_0782	MAV_1857	MAV_2586	MAV_3596	MAV_5060
MAV_0784	MAV_1859	MAV_2587	MAV_3630	MAV_5061
MAV_0785	MAV_1860	MAV_2588	MAV_3648	MAV_5062
MAV_0786	MAV_1862	MAV_2589	MAV_3653	MAV_5063
MAV_0787	MAV_1863	MAV_2591	MAV_3656	MAV_5064
MAV_0788	MAV_1864	MAV_2593	MAV_3663	MAV_5065
MAV_0789	MAV_1865	MAV_2595	MAV_3667	MAV_5066
MAV_0790	MAV_1867	MAV_2596	MAV_3697	MAV_5067
MAV_0791	MAV_1870	MAV_2598	MAV_3713	MAV_5070
MAV_0792	MAV_1872	MAV_2599	MAV_3715	MAV_5071
MAV_0793	MAV_1873	MAV_2601	MAV_3734	MAV_5072
MAV_0794	MAV_1874	MAV_2602	MAV_3738	MAV_5073
MAV_0795	MAV_1875	MAV_2605	MAV_3750	MAV_5074
MAV_0796	MAV_1877	MAV_2606	MAV_3778	MAV_5075
MAV_0797	MAV_1879	MAV_2607	MAV_3781	MAV_5076
MAV_0799	MAV_1880	MAV_2609	MAV_3788	MAV_5078
MAV_0800	MAV_1881	MAV_2610	MAV_3789	MAV_5080
MAV_0801	MAV_1882	MAV_2611	MAV_3791	MAV_5081
MAV_0802	MAV_1884	MAV_2612	MAV_3792	MAV_5082
MAV_0803	MAV_1888	MAV_2613	MAV_3793	MAV_5083
MAV_0804	MAV_1890	MAV_2615	MAV_3800	MAV_5084
MAV_0805	MAV_1891	MAV_2616	MAV_3807	MAV_5086
MAV_0806	MAV_1893	MAV_2617	MAV_3808	MAV_5087
MAV_0807	MAV_1895	MAV_2618	MAV_3815	MAV_5088
MAV_0808	MAV_1896	MAV_2619	MAV_3821	MAV_5089
MAV_0809	MAV_1898	MAV_2620	MAV_3823	MAV_5090
MAV_0810	MAV_1899	MAV_2621	MAV_3836	MAV_5091
MAV_0811	MAV_1900	MAV_2622	MAV_3862	MAV_5092
MAV_0812	MAV_1901	MAV_2623	MAV_3865	MAV_5094
MAV_0813	MAV_1902	MAV_2624	MAV_3881	MAV_5097
MAV_0814	MAV_1904	MAV_2625	MAV_3883	MAV_5103
MAV_0816	MAV_1905	MAV_2626	MAV_3892	MAV_5107
MAV_0817	MAV_1906	MAV_2628	MAV_3900	MAV_5117
MAV_0818	MAV_1907	MAV_2630	MAV_3909	MAV_5176
MAV_0820	MAV_1908	MAV_2631	MAV_3918	MAV_5180
MAV_0821	MAV_1909	MAV_2632	MAV_3929	MAV_5181
MAV_0822	MAV_1910	MAV_2633	MAV_3930	MAV_5182
MAV_0823	MAV_1911	MAV_2634	MAV_3941	MAV_5195
MAV_0824	MAV_1912	MAV_2636	MAV_3942	MAV_5199
MAV_0825	MAV_1913	MAV_2637	MAV_3944	MAV_5209
MAV_0826	MAV_1914	MAV_2639	MAV_3945	MAV_5214

MAV_0827	MAV_1915	MAV_2642	MAV_3946	MAV_5216
MAV_0828	MAV_1916	MAV_2643	MAV_3947	MAV_5218
MAV_0829	MAV_1918	MAV_2644	MAV_3948	MAV_5236
MAV_0830	MAV_1919	MAV_2645	MAV_3950	MAV_5244
MAV_0831	MAV_1920	MAV_2646	MAV_3951	MAV_5255
MAV_0832	MAV_1921	MAV_2647	MAV_3955	MAV_5256
MAV_0833	MAV_1923	MAV_2650	MAV_3956	MAV_5259
MAV_0834	MAV_1924	MAV_2653	MAV_3957	MAV_5260
MAV_0835	MAV_1926	MAV_2654	MAV_3968	MAV_5261
MAV_0836	MAV_1927	MAV_2655	MAV_3969	MAV_5262
MAV_0838	MAV_1929	MAV_2657	MAV_3970	MAV_5263
MAV_0840	MAV_1930	MAV_2658	MAV_3972	MAV_5273
MAV_0841	MAV_1931	MAV_2660	MAV_3994	MAV_5286
MAV_0848	MAV_1932	MAV_2661	MAV_3995	MAV_5287
MAV_0861	MAV_1936	MAV_2662	MAV_3997	MAV_5289
MAV_0866	MAV_1937	MAV_2663	MAV_4004	MAV_5295
MAV_0869	MAV_1940	MAV_2664	MAV_4011	MAV_5305
MAV_0899	MAV_1941	MAV_2665	MAV_4012	

Table 10 Genes unique to *Mycobacterium leprae* when the genome was compared to the genomes of 7 other mycobacterial species

ML0008	ML0464	ML0946	ML1575	ML2178
ML0009	ML0470	ML0950	ML1601	ML2201
ML0023	ML0472	ML0953	ML1602	ML2244
ML0024	ML0473	ML0957	ML1603	ML2249
ML0025	ML0527	ML0958	ML1604	ML2252
ML0056	ML0568	ML0959	ML1605	ML2253
ML0067	ML0573	ML0963	ML1717	ML2283
ML0070	ML0574	ML0964	ML1788	ML2346
ML0124	ML0575	ML1010	ML1821	ML2347
ML0141	ML0576	ML1011	ML1829	ML2407
ML0152	ML0588	ML1018	ML1915	ML2428A
ML0162	ML0638	ML1057	ML1928	ML2468
ML0217	ML0656	ML1148	ML1932	ML2562
ML0218	ML0659	ML1186	ML1949	ML2567
ML0291	ML0664	ML1189	ML1976	ML2630
ML0292	ML0678	ML1243	ML1979	ML2651
ML0293	ML0679	ML1275	ML1989	ML2666
ML0308	ML0777	ML1292	ML1990	
ML0369	ML0796	ML1344	ML2035	

ML0394	ML0837	ML1384	ML2044	
ML0397	ML0863	ML1419	ML2129	
ML0398	ML0927	ML1420	ML2170	
ML0411	ML0928	ML1505	ML2172	
ML0447	ML0938	ML1523	ML2176	
ML0448	ML0939	ML1553	ML2177	

Table 11 Genes unique to *Mycobacterium marinum* when the genome was compared to the genomes of 7 other mycobacterial species

MMAR_0008	MMAR_1386	MMAR_2550	MMAR_3553	MMAR_4439
MMAR_0023	MMAR_1391	MMAR_2553	MMAR_3564	MMAR_4440
MMAR_0024	MMAR_1401	MMAR_2555	MMAR_3565	MMAR_4441
MMAR_0025	MMAR_1402	MMAR_2557	MMAR_3568	MMAR_4442
MMAR_0026	MMAR_1404	MMAR_2565	MMAR_3569	MMAR_4444
MMAR_0027	MMAR_1405	MMAR_2566	MMAR_3570	MMAR_4445
MMAR_0028	MMAR_1406	MMAR_2567	MMAR_3571	MMAR_4446
MMAR_0029	MMAR_1410	MMAR_2575	MMAR_3572	MMAR_4447
MMAR_0030	MMAR_1412	MMAR_2579	MMAR_3573	MMAR_4449
MMAR_0036	MMAR_1413	MMAR_2588	MMAR_3574	MMAR_4450
MMAR_0045	MMAR_1415	MMAR_2589	MMAR_3575	MMAR_4474
MMAR_0050	MMAR_1420	MMAR_2591	MMAR_3577	MMAR_4475
MMAR_0077	MMAR_1421	MMAR_2596	MMAR_3578	MMAR_4482
MMAR_0079	MMAR_1422	MMAR_2600	MMAR_3579	MMAR_4483
MMAR_0080	MMAR_1423	MMAR_2601	MMAR_3580	MMAR_4484
MMAR_0081	MMAR_1424	MMAR_2602	MMAR_3581	MMAR_4489
MMAR_0084	MMAR_1425	MMAR_2603	MMAR_3582	MMAR_4500
MMAR_0086	MMAR_1426	MMAR_2604	MMAR_3583	MMAR_4501
MMAR_0087	MMAR_1428	MMAR_2607	MMAR_3584	MMAR_4502
MMAR_0089	MMAR_1429	MMAR_2613	MMAR_3585	MMAR_4504
MMAR_0090	MMAR_1438	MMAR_2623	MMAR_3586	MMAR_4505
MMAR_0091	MMAR_1439	MMAR_2626	MMAR_3587	MMAR_4506
MMAR_0092	MMAR_1440	MMAR_2645	MMAR_3588	MMAR_4507
MMAR_0096	MMAR_1441	MMAR_2650	MMAR_3591	MMAR_4510
MMAR_0097	MMAR_1442	MMAR_2655	MMAR_3592	MMAR_4511
MMAR_0100	MMAR_1447	MMAR_2656	MMAR_3593	MMAR_4512
MMAR_0105	MMAR_1449	MMAR_2662	MMAR_3596	MMAR_4513
MMAR_0108	MMAR_1450	MMAR_2668	MMAR_3597	MMAR_4514
MMAR_0109	MMAR_1452	MMAR_2689	MMAR_3598	MMAR_4537
MMAR_0111	MMAR_1453	MMAR_2690	MMAR_3599	MMAR_4538
MMAR_0113	MMAR_1454	MMAR_2695	MMAR_3600	MMAR_4548

MMAR_0115	MMAR_1457	MMAR_2709	MMAR_3601	MMAR_4554
MMAR_0117	MMAR_1458	MMAR_2734	MMAR_3602	MMAR_4556
MMAR_0118	MMAR_1459	MMAR_2735	MMAR_3603	MMAR_4560
MMAR_0119	MMAR_1462	MMAR_2744	MMAR_3604	MMAR_4561
MMAR_0125	MMAR_1464	MMAR_2748	MMAR_3609	MMAR_4562
MMAR_0128	MMAR_1465	MMAR_2774	MMAR_3610	MMAR_4571
MMAR_0129	MMAR_1469	MMAR_2775	MMAR_3613	MMAR_4585
MMAR_0130	MMAR_1484	MMAR_2776	MMAR_3628	MMAR_4588
MMAR_0131	MMAR_1497	MMAR_2778	MMAR_3629	MMAR_4589
MMAR_0132	MMAR_1507	MMAR_2788	MMAR_3632	MMAR_4595
MMAR_0133	MMAR_1508	MMAR_2795	MMAR_3633	MMAR_4596
MMAR_0134	MMAR_1513	MMAR_2803	MMAR_3634	MMAR_4597
MMAR_0136	MMAR_1524	MMAR_2804	MMAR_3636	MMAR_4598
MMAR_0138	MMAR_1529	MMAR_2805	MMAR_3640	MMAR_4599
MMAR_0140	MMAR_1536	MMAR_2806	MMAR_3641	MMAR_4600
MMAR_0141	MMAR_1539	MMAR_2809	MMAR_3644	MMAR_4601
MMAR_0142	MMAR_1540	MMAR_2817	MMAR_3648	MMAR_4602
MMAR_0143	MMAR_1543	MMAR_2818	MMAR_3658	MMAR_4603
MMAR_0157	MMAR_1544	MMAR_2823	MMAR_3659	MMAR_4606
MMAR_0175	MMAR_1545	MMAR_2830	MMAR_3662	MMAR_4607
MMAR_0183	MMAR_1546	MMAR_2831	MMAR_3666	MMAR_4611
MMAR_0185	MMAR_1548	MMAR_2832	MMAR_3680	MMAR_4612
MMAR_0189	MMAR_1549	MMAR_2833	MMAR_3681	MMAR_4618
MMAR_0190	MMAR_1550	MMAR_2841	MMAR_3692	MMAR_4627
MMAR_0193	MMAR_1552	MMAR_2842	MMAR_3728	MMAR_4628
MMAR_0194	MMAR_1553	MMAR_2846	MMAR_3729	MMAR_4630
MMAR_0195	MMAR_1558	MMAR_2847	MMAR_3742	MMAR_4635
MMAR_0196	MMAR_1560	MMAR_2851	MMAR_3758	MMAR_4672
MMAR_0197	MMAR_1561	MMAR_2853	MMAR_3763	MMAR_4684
MMAR_0198	MMAR_1562	MMAR_2858	MMAR_3778	MMAR_4686
MMAR_0199	MMAR_1563	MMAR_2859	MMAR_3787	MMAR_4687
MMAR_0200	MMAR_1564	MMAR_2874	MMAR_3790	MMAR_4700
MMAR_0201	MMAR_1565	MMAR_2893	MMAR_3793	MMAR_4702
MMAR_0202	MMAR_1576	MMAR_2894	MMAR_3800	MMAR_4703
MMAR_0203	MMAR_1579	MMAR_2895	MMAR_3811	MMAR_4723
MMAR_0209	MMAR_1581	MMAR_2897	MMAR_3816	MMAR_4735
MMAR_0210	MMAR_1592	MMAR_2898	MMAR_3824	MMAR_4783
MMAR_0225	MMAR_1594	MMAR_2900	MMAR_3825	MMAR_4786
MMAR_0227	MMAR_1595	MMAR_2901	MMAR_3832	MMAR_4801
MMAR_0229	MMAR_1596	MMAR_2903	MMAR_3857	MMAR_4802
MMAR_0230	MMAR_1597	MMAR_2905	MMAR_3870	MMAR_4803

MMAR_0232	MMAR_1598	MMAR_2907	MMAR_3873	MMAR_4804
MMAR_0237	MMAR_1599	MMAR_2911	MMAR_3874	MMAR_4813
MMAR_0238	MMAR_1600	MMAR_2912	MMAR_3875	MMAR_4814
MMAR_0239	MMAR_1601	MMAR_2935	MMAR_3876	MMAR_4816
MMAR_0242	MMAR_1602	MMAR_2940	MMAR_3878	MMAR_4817
MMAR_0245	MMAR_1604	MMAR_2941	MMAR_3880	MMAR_4822
MMAR_0254	MMAR_1605	MMAR_2942	MMAR_3881	MMAR_4826
MMAR_0266	MMAR_1606	MMAR_2943	MMAR_3882	MMAR_4828
MMAR_0272	MMAR_1609	MMAR_2946	MMAR_3883	MMAR_4829
MMAR_0278	MMAR_1610	MMAR_2948	MMAR_3884	MMAR_4831
MMAR_0279	MMAR_1618	MMAR_2952	MMAR_3885	MMAR_4835
MMAR_0280	MMAR_1619	MMAR_2955	MMAR_3886	MMAR_4840
MMAR_0282	MMAR_1629	MMAR_2958	MMAR_3887	MMAR_4841
MMAR_0299	MMAR_1632	MMAR_2959	MMAR_3888	MMAR_4843
MMAR_0308	MMAR_1666	MMAR_2960	MMAR_3889	MMAR_4845
MMAR_0333	MMAR_1669	MMAR_2969	MMAR_3890	MMAR_4846
MMAR_0338	MMAR_1674	MMAR_2970	MMAR_3891	MMAR_4847
MMAR_0358	MMAR_1680	MMAR_2973	MMAR_3892	MMAR_4853
MMAR_0362	MMAR_1698	MMAR_2980	MMAR_3893	MMAR_4872
MMAR_0363	MMAR_1699	MMAR_2983	MMAR_3894	MMAR_4875
MMAR_0364	MMAR_1709	MMAR_2988	MMAR_3895	MMAR_4884
MMAR_0365	MMAR_1724	MMAR_2989	MMAR_3896	MMAR_4899
MMAR_0366	MMAR_1745	MMAR_2992	MMAR_3897	MMAR_4903
MMAR_0367	MMAR_1752	MMAR_3002	MMAR_3898	MMAR_4914
MMAR_0368	MMAR_1766	MMAR_3003	MMAR_3899	MMAR_4922
MMAR_0369	MMAR_1779	MMAR_3005	MMAR_3903	MMAR_4923
MMAR_0371	MMAR_1829	MMAR_3006	MMAR_3905	MMAR_4936
MMAR_0382	MMAR_1833	MMAR_3022	MMAR_3906	MMAR_4938
MMAR_0431	MMAR_1845	MMAR_3023	MMAR_3908	MMAR_4951
MMAR_0454	MMAR_1846	MMAR_3033	MMAR_3909	MMAR_4952
MMAR_0469	MMAR_1847	MMAR_3038	MMAR_3910	MMAR_4956
MMAR_0478	MMAR_1848	MMAR_3040	MMAR_3911	MMAR_4958
MMAR_0479	MMAR_1849	MMAR_3042	MMAR_3912	MMAR_4969
MMAR_0480	MMAR_1850	MMAR_3043	MMAR_3913	MMAR_4970
MMAR_0483	MMAR_1851	MMAR_3064	MMAR_3914	MMAR_4995
MMAR_0484	MMAR_1852	MMAR_3072	MMAR_3915	MMAR_4999
MMAR_0491	MMAR_1853	MMAR_3073	MMAR_3916	MMAR_5009
MMAR_0492	MMAR_1858	MMAR_3074	MMAR_3917	MMAR_5013
MMAR_0495	MMAR_1861	MMAR_3088	MMAR_3919	MMAR_5020
MMAR_0498	MMAR_1874	MMAR_3105	MMAR_3923	MMAR_5038
MMAR_0507	MMAR_1878	MMAR_3108	MMAR_3924	MMAR_5044

MMAR_0540	MMAR_1879	MMAR_3146	MMAR_3925	MMAR_5047
MMAR_0544	MMAR_1908	MMAR_3148	MMAR_3926	MMAR_5056
MMAR_0572	MMAR_1909	MMAR_3149	MMAR_3927	MMAR_5059
MMAR_0573	MMAR_1924	MMAR_3162	MMAR_3928	MMAR_5061
MMAR_0574	MMAR_1940	MMAR_3165	MMAR_3929	MMAR_5075
MMAR_0577	MMAR_1942	MMAR_3177	MMAR_3930	MMAR_5093
MMAR_0578	MMAR_1943	MMAR_3183	MMAR_3931	MMAR_5099
MMAR_0579	MMAR_1944	MMAR_3184	MMAR_3932	MMAR_5117
MMAR_0580	MMAR_1946	MMAR_3199	MMAR_3933	MMAR_5135
MMAR_0581	MMAR_1954	MMAR_3212	MMAR_3934	MMAR_5143
MMAR_0582	MMAR_1982	MMAR_3213	MMAR_3935	MMAR_5146
MMAR_0583	MMAR_2033	MMAR_3214	MMAR_3936	MMAR_5176
MMAR_0584	MMAR_2052	MMAR_3215	MMAR_3937	MMAR_5177
MMAR_0588	MMAR_2053	MMAR_3216	MMAR_3938	MMAR_5178
MMAR_0590	MMAR_2054	MMAR_3217	MMAR_3939	MMAR_5193
MMAR_0593	MMAR_2066	MMAR_3219	MMAR_3940	MMAR_5207
MMAR_0594	MMAR_2067	MMAR_3220	MMAR_3941	MMAR_5218
MMAR_0597	MMAR_2097	MMAR_3221	MMAR_3942	MMAR_5219
MMAR_0601	MMAR_2099	MMAR_3265	MMAR_3944	MMAR_5227
MMAR_0607	MMAR_2100	MMAR_3267	MMAR_3945	MMAR_5237
MMAR_0608	MMAR_2102	MMAR_3272	MMAR_3947	MMAR_5243
MMAR_0609	MMAR_2112	MMAR_3274	MMAR_3948	MMAR_5247
MMAR_0614	MMAR_2113	MMAR_3279	MMAR_3949	MMAR_5255
MMAR_0618	MMAR_2116	MMAR_3280	MMAR_3951	MMAR_5258
MMAR_0619	MMAR_2128	MMAR_3282	MMAR_3972	MMAR_5259
MMAR_0625	MMAR_2130	MMAR_3290	MMAR_3977	MMAR_5260
MMAR_0666	MMAR_2132	MMAR_3300	MMAR_3984	MMAR_5261
MMAR_0674	MMAR_2133	MMAR_3302	MMAR_3989	MMAR_5262
MMAR_0675	MMAR_2134	MMAR_3303	MMAR_3990	MMAR_5263
MMAR_0676	MMAR_2141	MMAR_3316	MMAR_4001	MMAR_5267
MMAR_0690	MMAR_2142	MMAR_3321	MMAR_4002	MMAR_5277
MMAR_0691	MMAR_2143	MMAR_3322	MMAR_4005	MMAR_5279
MMAR_0692	MMAR_2144	MMAR_3323	MMAR_4017	MMAR_5291
MMAR_0693	MMAR_2146	MMAR_3325	MMAR_4018	MMAR_5294
MMAR_0694	MMAR_2185	MMAR_3326	MMAR_4041	MMAR_5306
MMAR_0723	MMAR_2238	MMAR_3327	MMAR_4044	MMAR_5307
MMAR_0730	MMAR_2245	MMAR_3329	MMAR_4046	MMAR_5308
MMAR_0737	MMAR_2248	MMAR_3330	MMAR_4060	MMAR_5311
MMAR_0741	MMAR_2256	MMAR_3363	MMAR_4062	MMAR_5321
MMAR_0755	MMAR_2272	MMAR_3364	MMAR_4073	MMAR_5322
MMAR_0784	MMAR_2274	MMAR_3367	MMAR_4107	MMAR_5339

MMAR_0786	MMAR_2307	MMAR_3372	MMAR_4109	MMAR_5340
MMAR_0787	MMAR_2308	MMAR_3373	MMAR_4112	MMAR_5341
MMAR_0795	MMAR_2309	MMAR_3374	MMAR_4113	MMAR_5344
MMAR_0806	MMAR_2310	MMAR_3376	MMAR_4116	MMAR_5345
MMAR_0828	MMAR_2311	MMAR_3377	MMAR_4119	MMAR_5348
MMAR_0837	MMAR_2312	MMAR_3378	MMAR_4127	MMAR_5349
MMAR_0838	MMAR_2315	MMAR_3379	MMAR_4149	MMAR_5362
MMAR_0847	MMAR_2316	MMAR_3380	MMAR_4158	MMAR_5401
MMAR_0904	MMAR_2319	MMAR_3381	MMAR_4159	MMAR_5402
MMAR_0905	MMAR_2323	MMAR_3383	MMAR_4169	MMAR_5414
MMAR_0906	MMAR_2324	MMAR_3385	MMAR_4175	MMAR_5415
MMAR_0916	MMAR_2326	MMAR_3391	MMAR_4186	MMAR_5416
MMAR_0918	MMAR_2328	MMAR_3395	MMAR_4187	MMAR_5417
MMAR_0919	MMAR_2329	MMAR_3398	MMAR_4201	MMAR_5418
MMAR_0921	MMAR_2330	MMAR_3400	MMAR_4208	MMAR_5419
MMAR_0924	MMAR_2331	MMAR_3405	MMAR_4218	MMAR_5420
MMAR_0925	MMAR_2332	MMAR_3411	MMAR_4223	MMAR_5421
MMAR_0926	MMAR_2333	MMAR_3412	MMAR_4224	MMAR_5422
MMAR_0927	MMAR_2335	MMAR_3414	MMAR_4237	MMAR_5423
MMAR_0930	MMAR_2336	MMAR_3422	MMAR_4239	MMAR_5424
MMAR_0931	MMAR_2337	MMAR_3424	MMAR_4241	MMAR_5425
MMAR_0933	MMAR_2338	MMAR_3425	MMAR_4247	MMAR_5426
MMAR_0935	MMAR_2339	MMAR_3426	MMAR_4248	MMAR_5427
MMAR_0937	MMAR_2347	MMAR_3428	MMAR_4256	MMAR_5428
MMAR_0943	MMAR_2350	MMAR_3430	MMAR_4258	MMAR_5429
MMAR_0947	MMAR_2351	MMAR_3432	MMAR_4260	MMAR_5430
MMAR_0951	MMAR_2372	MMAR_3436	MMAR_4261	MMAR_5431
MMAR_0954	MMAR_2403	MMAR_3439	MMAR_4262	MMAR_5432
MMAR_0988	MMAR_2404	MMAR_3442	MMAR_4266	MMAR_5433
MMAR_0989	MMAR_2405	MMAR_3446	MMAR_4267	MMAR_5434
MMAR_0998	MMAR_2408	MMAR_3447	MMAR_4269	MMAR_5435
MMAR_1080	MMAR_2435	MMAR_3448	MMAR_4270	MMAR_5447
MMAR_1095	MMAR_2441	MMAR_3451	MMAR_4287	MMAR_5454
MMAR_1127	MMAR_2444	MMAR_3464	MMAR_4293	MMAR_5462
MMAR_1128	MMAR_2458	MMAR_3466	MMAR_4316	MMAR_5463
MMAR_1129	MMAR_2459	MMAR_3467	MMAR_4319	MMAR_5486
MMAR_1130	MMAR_2461	MMAR_3468	MMAR_4321	MMAR_5521
MMAR_1139	MMAR_2477	MMAR_3473	MMAR_4341	MMAR_5543
MMAR_1161	MMAR_2480	MMAR_3474	MMAR_4346	MMAR_5544
MMAR_1162	MMAR_2481	MMAR_3475	MMAR_4347	MMAR_5545
MMAR_1170	MMAR_2483	MMAR_3478	MMAR_4348	MMAR_5546

MMAR_1171	MMAR_2490	MMAR_3494	MMAR_4349	MMAR_5547
MMAR_1172	MMAR_2492	MMAR_3495	MMAR_4350	MMAR_5548
MMAR_1181	MMAR_2508	MMAR_3496	MMAR_4352	MMAR_5549
MMAR_1193	MMAR_2509	MMAR_3499	MMAR_4354	MMAR_5550
MMAR_1195	MMAR_2510	MMAR_3500	MMAR_4362	MMAR_5552
MMAR_1199	MMAR_2511	MMAR_3506	MMAR_4373	MMAR_5553
MMAR_1207	MMAR_2512	MMAR_3508	MMAR_4377	MMAR_5555
MMAR_1208	MMAR_2513	MMAR_3517	MMAR_4398	MMAR_5556
MMAR_1213	MMAR_2514	MMAR_3519	MMAR_4399	MMAR_5557
MMAR_1234	MMAR_2516	MMAR_3522	MMAR_4407	MMAR_5558
MMAR_1246	MMAR_2519	MMAR_3524	MMAR_4416	MMAR_5561
MMAR_1250	MMAR_2520	MMAR_3525	MMAR_4421	MMAR_5562
MMAR_1268	MMAR_2521	MMAR_3526	MMAR_4422	MMAR_5563
MMAR_1297	MMAR_2522	MMAR_3528	MMAR_4425	MMAR_5565
MMAR_1324	MMAR_2537	MMAR_3536	MMAR_4426	MMAR_5567
MMAR_1325	MMAR_2538	MMAR_3540	MMAR_4428	MMAR_5570
MMAR_1326	MMAR_2539	MMAR_3544	MMAR_4431	MMAR_5571
MMAR_1327	MMAR_2541	MMAR_3545	MMAR_4432	MMAR_5573
MMAR_1328	MMAR_2542	MMAR_3546	MMAR_4433	MMAR_5580
MMAR_1329	MMAR_2543	MMAR_3547	MMAR_4434	MMAR_5582
MMAR_1331	MMAR_2545	MMAR_3548	MMAR_4436	
MMAR_1348	MMAR_2546	MMAR_3550	MMAR_4437	
MMAR_1377	MMAR_2549	MMAR_3551	MMAR_4438	

Table 12 Genes unique to *Mycobacterium vanbaalenii* when the genome was compared to the genomes of 7 other mycobacterial species

Mvan_0009	Mvan_1075	Mvan_2323	Mvan_3692	Mvan_5023
Mvan_0010	Mvan_1076	Mvan_2331	Mvan_3693	Mvan_5026
Mvan_0011	Mvan_1077	Mvan_2332	Mvan_3697	Mvan_5028
Mvan_0012	Mvan_1078	Mvan_2333	Mvan_3698	Mvan_5029
Mvan_0021	Mvan_1080	Mvan_2334	Mvan_3700	Mvan_5042
Mvan_0030	Mvan_1082	Mvan_2335	Mvan_3702	Mvan_5057
Mvan_0031	Mvan_1083	Mvan_2336	Mvan_3703	Mvan_5058
Mvan_0032	Mvan_1084	Mvan_2344	Mvan_3705	Mvan_5059
Mvan_0033	Mvan_1085	Mvan_2354	Mvan_3708	Mvan_5060
Mvan_0035	Mvan_1086	Mvan_2359	Mvan_3711	Mvan_5061
Mvan_0036	Mvan_1087	Mvan_2360	Mvan_3712	Mvan_5064
Mvan_0038	Mvan_1088	Mvan_2363	Mvan_3713	Mvan_5065
Mvan_0039	Mvan_1089	Mvan_2364	Mvan_3714	Mvan_5074
Mvan_0041	Mvan_1090	Mvan_2365	Mvan_3715	Mvan_5075

Mvan_0042	Mvan_1091	Mvan_2367	Mvan_3719	Mvan_5079
Mvan_0044	Mvan_1092	Mvan_2368	Mvan_3720	Mvan_5082
Mvan_0045	Mvan_1096	Mvan_2369	Mvan_3722	Mvan_5095
Mvan_0046	Mvan_1097	Mvan_2370	Mvan_3725	Mvan_5104
Mvan_0047	Mvan_1098	Mvan_2371	Mvan_3728	Mvan_5106
Mvan_0050	Mvan_1099	Mvan_2372	Mvan_3731	Mvan_5107
Mvan_0052	Mvan_1100	Mvan_2373	Mvan_3732	Mvan_5112
Mvan_0055	Mvan_1101	Mvan_2385	Mvan_3733	Mvan_5124
Mvan_0058	Mvan_1102	Mvan_2389	Mvan_3735	Mvan_5125
Mvan_0063	Mvan_1103	Mvan_2396	Mvan_3737	Mvan_5126
Mvan_0065	Mvan_1104	Mvan_2406	Mvan_3738	Mvan_5127
Mvan_0076	Mvan_1105	Mvan_2417	Mvan_3739	Mvan_5140
Mvan_0082	Mvan_1106	Mvan_2418	Mvan_3740	Mvan_5141
Mvan_0083	Mvan_1107	Mvan_2423	Mvan_3741	Mvan_5142
Mvan_0084	Mvan_1108	Mvan_2424	Mvan_3742	Mvan_5143
Mvan_0086	Mvan_1109	Mvan_2459	Mvan_3743	Mvan_5147
Mvan_0087	Mvan_1112	Mvan_2478	Mvan_3744	Mvan_5151
Mvan_0088	Mvan_1114	Mvan_2484	Mvan_3745	Mvan_5152
Mvan_0090	Mvan_1115	Mvan_2486	Mvan_3746	Mvan_5171
Mvan_0093	Mvan_1117	Mvan_2495	Mvan_3748	Mvan_5177
Mvan_0094	Mvan_1118	Mvan_2496	Mvan_3766	Mvan_5178
Mvan_0095	Mvan_1121	Mvan_2499	Mvan_3779	Mvan_5179
Mvan_0096	Mvan_1124	Mvan_2505	Mvan_3782	Mvan_5194
Mvan_0097	Mvan_1125	Mvan_2507	Mvan_3786	Mvan_5212
Mvan_0098	Mvan_1131	Mvan_2510	Mvan_3792	Mvan_5231
Mvan_0100	Mvan_1132	Mvan_2511	Mvan_3794	Mvan_5232
Mvan_0101	Mvan_1133	Mvan_2512	Mvan_3799	Mvan_5233
Mvan_0111	Mvan_1134	Mvan_2513	Mvan_3800	Mvan_5239
Mvan_0113	Mvan_1135	Mvan_2514	Mvan_3801	Mvan_5240
Mvan_0114	Mvan_1136	Mvan_2515	Mvan_3810	Mvan_5243
Mvan_0124	Mvan_1137	Mvan_2516	Mvan_3811	Mvan_5245
Mvan_0128	Mvan_1138	Mvan_2517	Mvan_3813	Mvan_5246
Mvan_0145	Mvan_1139	Mvan_2518	Mvan_3826	Mvan_5247
Mvan_0146	Mvan_1142	Mvan_2519	Mvan_3827	Mvan_5248
Mvan_0148	Mvan_1146	Mvan_2520	Mvan_3839	Mvan_5249
Mvan_0151	Mvan_1147	Mvan_2521	Mvan_3853	Mvan_5250
Mvan_0159	Mvan_1148	Mvan_2522	Mvan_3856	Mvan_5251
Mvan_0161	Mvan_1149	Mvan_2523	Mvan_3860	Mvan_5252
Mvan_0162	Mvan_1150	Mvan_2524	Mvan_3861	Mvan_5263
Mvan_0163	Mvan_1151	Mvan_2526	Mvan_3862	Mvan_5270
Mvan_0164	Mvan_1152	Mvan_2528	Mvan_3863	Mvan_5291

Mvan_0165	Mvan_1153	Mvan_2530	Mvan_3864	Mvan_5292
Mvan_0168	Mvan_1154	Mvan_2531	Mvan_3865	Mvan_5293
Mvan_0169	Mvan_1155	Mvan_2532	Mvan_3866	Mvan_5294
Mvan_0170	Mvan_1158	Mvan_2533	Mvan_3867	Mvan_5295
Mvan_0171	Mvan_1160	Mvan_2534	Mvan_3868	Mvan_5299
Mvan_0177	Mvan_1164	Mvan_2535	Mvan_3869	Mvan_5301
Mvan_0178	Mvan_1167	Mvan_2540	Mvan_3870	Mvan_5302
Mvan_0182	Mvan_1169	Mvan_2541	Mvan_3871	Mvan_5303
Mvan_0190	Mvan_1175	Mvan_2545	Mvan_3878	Mvan_5304
Mvan_0202	Mvan_1177	Mvan_2548	Mvan_3880	Mvan_5313
Mvan_0203	Mvan_1179	Mvan_2552	Mvan_3890	Mvan_5314
Mvan_0204	Mvan_1180	Mvan_2557	Mvan_3891	Mvan_5315
Mvan_0205	Mvan_1182	Mvan_2571	Mvan_3892	Mvan_5316
Mvan_0206	Mvan_1183	Mvan_2572	Mvan_3893	Mvan_5323
Mvan_0207	Mvan_1185	Mvan_2591	Mvan_3896	Mvan_5324
Mvan_0208	Mvan_1190	Mvan_2592	Mvan_3897	Mvan_5325
Mvan_0209	Mvan_1191	Mvan_2593	Mvan_3898	Mvan_5326
Mvan_0210	Mvan_1193	Mvan_2594	Mvan_3900	Mvan_5328
Mvan_0212	Mvan_1198	Mvan_2595	Mvan_3901	Mvan_5329
Mvan_0213	Mvan_1199	Mvan_2596	Mvan_3902	Mvan_5332
Mvan_0214	Mvan_1201	Mvan_2606	Mvan_3903	Mvan_5334
Mvan_0217	Mvan_1202	Mvan_2607	Mvan_3904	Mvan_5350
Mvan_0220	Mvan_1203	Mvan_2612	Mvan_3916	Mvan_5357
Mvan_0229	Mvan_1205	Mvan_2613	Mvan_3928	Mvan_5361
Mvan_0230	Mvan_1206	Mvan_2619	Mvan_3929	Mvan_5366
Mvan_0236	Mvan_1207	Mvan_2620	Mvan_3930	Mvan_5374
Mvan_0239	Mvan_1211	Mvan_2621	Mvan_3931	Mvan_5389
Mvan_0246	Mvan_1216	Mvan_2622	Mvan_3941	Mvan_5401
Mvan_0247	Mvan_1218	Mvan_2623	Mvan_3954	Mvan_5402
Mvan_0255	Mvan_1220	Mvan_2624	Mvan_3955	Mvan_5403
Mvan_0262	Mvan_1221	Mvan_2626	Mvan_3956	Mvan_5408
Mvan_0263	Mvan_1222	Mvan_2632	Mvan_3957	Mvan_5409
Mvan_0264	Mvan_1223	Mvan_2638	Mvan_3958	Mvan_5414
Mvan_0266	Mvan_1225	Mvan_2648	Mvan_3959	Mvan_5415
Mvan_0273	Mvan_1237	Mvan_2653	Mvan_3966	Mvan_5426
Mvan_0274	Mvan_1238	Mvan_2654	Mvan_3967	Mvan_5446
Mvan_0275	Mvan_1242	Mvan_2656	Mvan_3968	Mvan_5447
Mvan_0278	Mvan_1247	Mvan_2673	Mvan_3969	Mvan_5449
Mvan_0282	Mvan_1248	Mvan_2674	Mvan_3971	Mvan_5451
Mvan_0292	Mvan_1255	Mvan_2675	Mvan_3972	Mvan_5460
Mvan_0293	Mvan_1256	Mvan_2676	Mvan_3985	Mvan_5464

Mvan_0294	Mvan_1260	Mvan_2678	Mvan_3986	Mvan_5465
Mvan_0305	Mvan_1262	Mvan_2679	Mvan_3989	Mvan_5466
Mvan_0309	Mvan_1264	Mvan_2680	Mvan_3990	Mvan_5468
Mvan_0321	Mvan_1266	Mvan_2681	Mvan_3991	Mvan_5473
Mvan_0324	Mvan_1268	Mvan_2682	Mvan_3995	Mvan_5474
Mvan_0329	Mvan_1273	Mvan_2685	Mvan_3997	Mvan_5475
Mvan_0331	Mvan_1283	Mvan_2701	Mvan_3998	Mvan_5476
Mvan_0332	Mvan_1292	Mvan_2702	Mvan_4009	Mvan_5477
Mvan_0333	Mvan_1316	Mvan_2709	Mvan_4010	Mvan_5478
Mvan_0334	Mvan_1320	Mvan_2710	Mvan_4011	Mvan_5479
Mvan_0335	Mvan_1321	Mvan_2711	Mvan_4012	Mvan_5494
Mvan_0338	Mvan_1322	Mvan_2712	Mvan_4013	Mvan_5497
Mvan_0340	Mvan_1323	Mvan_2713	Mvan_4014	Mvan_5498
Mvan_0341	Mvan_1324	Mvan_2723	Mvan_4016	Mvan_5500
Mvan_0347	Mvan_1325	Mvan_2728	Mvan_4017	Mvan_5509
Mvan_0348	Mvan_1326	Mvan_2739	Mvan_4018	Mvan_5510
Mvan_0349	Mvan_1328	Mvan_2756	Mvan_4019	Mvan_5511
Mvan_0351	Mvan_1331	Mvan_2768	Mvan_4020	Mvan_5512
Mvan_0353	Mvan_1342	Mvan_2770	Mvan_4022	Mvan_5513
Mvan_0354	Mvan_1357	Mvan_2776	Mvan_4023	Mvan_5514
Mvan_0355	Mvan_1358	Mvan_2779	Mvan_4024	Mvan_5533
Mvan_0356	Mvan_1359	Mvan_2780	Mvan_4025	Mvan_5536
Mvan_0357	Mvan_1363	Mvan_2781	Mvan_4026	Mvan_5539
Mvan_0358	Mvan_1366	Mvan_2782	Mvan_4028	Mvan_5540
Mvan_0359	Mvan_1369	Mvan_2801	Mvan_4029	Mvan_5541
Mvan_0360	Mvan_1370	Mvan_2812	Mvan_4034	Mvan_5543
Mvan_0370	Mvan_1371	Mvan_2823	Mvan_4035	Mvan_5544
Mvan_0380	Mvan_1373	Mvan_2826	Mvan_4037	Mvan_5548
Mvan_0385	Mvan_1378	Mvan_2852	Mvan_4042	Mvan_5553
Mvan_0386	Mvan_1380	Mvan_2853	Mvan_4043	Mvan_5554
Mvan_0387	Mvan_1389	Mvan_2855	Mvan_4045	Mvan_5555
Mvan_0388	Mvan_1390	Mvan_2860	Mvan_4055	Mvan_5559
Mvan_0402	Mvan_1391	Mvan_2861	Mvan_4059	Mvan_5561
Mvan_0426	Mvan_1392	Mvan_2864	Mvan_4060	Mvan_5562
Mvan_0427	Mvan_1394	Mvan_2866	Mvan_4071	Mvan_5563
Mvan_0433	Mvan_1396	Mvan_2878	Mvan_4077	Mvan_5565
Mvan_0434	Mvan_1400	Mvan_2879	Mvan_4078	Mvan_5566
Mvan_0435	Mvan_1402	Mvan_2880	Mvan_4079	Mvan_5567
Mvan_0436	Mvan_1403	Mvan_2881	Mvan_4080	Mvan_5568
Mvan_0437	Mvan_1415	Mvan_2882	Mvan_4081	Mvan_5569
Mvan_0439	Mvan_1416	Mvan_2887	Mvan_4095	Mvan_5570

Mvan_0447	Mvan_1417	Mvan_2893	Mvan_4096	Mvan_5571
Mvan_0453	Mvan_1419	Mvan_2904	Mvan_4111	Mvan_5577
Mvan_0455	Mvan_1421	Mvan_2905	Mvan_4112	Mvan_5578
Mvan_0456	Mvan_1422	Mvan_2906	Mvan_4115	Mvan_5584
Mvan_0459	Mvan_1423	Mvan_2907	Mvan_4116	Mvan_5585
Mvan_0460	Mvan_1424	Mvan_2908	Mvan_4118	Mvan_5588
Mvan_0461	Mvan_1426	Mvan_2912	Mvan_4121	Mvan_5589
Mvan_0467	Mvan_1446	Mvan_2914	Mvan_4123	Mvan_5592
Mvan_0468	Mvan_1452	Mvan_2915	Mvan_4126	Mvan_5593
Mvan_0469	Mvan_1453	Mvan_2918	Mvan_4127	Mvan_5607
Mvan_0470	Mvan_1454	Mvan_2919	Mvan_4129	Mvan_5612
Mvan_0471	Mvan_1455	Mvan_2922	Mvan_4137	Mvan_5622
Mvan_0472	Mvan_1456	Mvan_2923	Mvan_4142	Mvan_5623
Mvan_0474	Mvan_1457	Mvan_2926	Mvan_4143	Mvan_5634
Mvan_0475	Mvan_1458	Mvan_2930	Mvan_4144	Mvan_5638
Mvan_0476	Mvan_1459	Mvan_2933	Mvan_4147	Mvan_5643
Mvan_0478	Mvan_1460	Mvan_2934	Mvan_4152	Mvan_5653
Mvan_0479	Mvan_1461	Mvan_2935	Mvan_4158	Mvan_5670
Mvan_0480	Mvan_1463	Mvan_2937	Mvan_4161	Mvan_5671
Mvan_0483	Mvan_1471	Mvan_2939	Mvan_4191	Mvan_5676
Mvan_0484	Mvan_1474	Mvan_2940	Mvan_4192	Mvan_5678
Mvan_0485	Mvan_1475	Mvan_2941	Mvan_4197	Mvan_5679
Mvan_0486	Mvan_1477	Mvan_2944	Mvan_4199	Mvan_5680
Mvan_0487	Mvan_1481	Mvan_2945	Mvan_4205	Mvan_5681
Mvan_0488	Mvan_1492	Mvan_2946	Mvan_4209	Mvan_5686
Mvan_0489	Mvan_1496	Mvan_2949	Mvan_4219	Mvan_5688
Mvan_0491	Mvan_1497	Mvan_2951	Mvan_4222	Mvan_5689
Mvan_0492	Mvan_1498	Mvan_2953	Mvan_4239	Mvan_5690
Mvan_0493	Mvan_1500	Mvan_2954	Mvan_4252	Mvan_5691
Mvan_0494	Mvan_1503	Mvan_2955	Mvan_4266	Mvan_5708
Mvan_0495	Mvan_1504	Mvan_2959	Mvan_4278	Mvan_5710
Mvan_0498	Mvan_1505	Mvan_2960	Mvan_4280	Mvan_5711
Mvan_0499	Mvan_1515	Mvan_2962	Mvan_4282	Mvan_5714
Mvan_0501	Mvan_1517	Mvan_2963	Mvan_4285	Mvan_5719
Mvan_0502	Mvan_1535	Mvan_2967	Mvan_4310	Mvan_5720
Mvan_0503	Mvan_1543	Mvan_2968	Mvan_4312	Mvan_5723
Mvan_0508	Mvan_1544	Mvan_2969	Mvan_4313	Mvan_5730
Mvan_0509	Mvan_1547	Mvan_2973	Mvan_4315	Mvan_5731
Mvan_0515	Mvan_1557	Mvan_2974	Mvan_4316	Mvan_5732
Mvan_0517	Mvan_1564	Mvan_2975	Mvan_4318	Mvan_5733
Mvan_0519	Mvan_1565	Mvan_2976	Mvan_4319	Mvan_5734

Mvan_0520	Mvan_1566	Mvan_2978	Mvan_4320	Mvan_5736
Mvan_0522	Mvan_1569	Mvan_2979	Mvan_4343	Mvan_5739
Mvan_0525	Mvan_1584	Mvan_2980	Mvan_4351	Mvan_5740
Mvan_0526	Mvan_1595	Mvan_2982	Mvan_4352	Mvan_5741
Mvan_0529	Mvan_1596	Mvan_2994	Mvan_4353	Mvan_5742
Mvan_0532	Mvan_1597	Mvan_2995	Mvan_4354	Mvan_5744
Mvan_0533	Mvan_1599	Mvan_2998	Mvan_4357	Mvan_5745
Mvan_0534	Mvan_1602	Mvan_3002	Mvan_4358	Mvan_5747
Mvan_0535	Mvan_1603	Mvan_3006	Mvan_4359	Mvan_5748
Mvan_0536	Mvan_1604	Mvan_3011	Mvan_4360	Mvan_5754
Mvan_0538	Mvan_1605	Mvan_3014	Mvan_4362	Mvan_5756
Mvan_0539	Mvan_1609	Mvan_3020	Mvan_4363	Mvan_5759
Mvan_0540	Mvan_1615	Mvan_3022	Mvan_4365	Mvan_5761
Mvan_0541	Mvan_1617	Mvan_3024	Mvan_4366	Mvan_5763
Mvan_0542	Mvan_1623	Mvan_3026	Mvan_4368	Mvan_5764
Mvan_0545	Mvan_1625	Mvan_3028	Mvan_4372	Mvan_5765
Mvan_0546	Mvan_1629	Mvan_3031	Mvan_4373	Mvan_5767
Mvan_0547	Mvan_1633	Mvan_3032	Mvan_4374	Mvan_5771
Mvan_0548	Mvan_1634	Mvan_3048	Mvan_4375	Mvan_5774
Mvan_0549	Mvan_1635	Mvan_3052	Mvan_4381	Mvan_5775
Mvan_0552	Mvan_1636	Mvan_3054	Mvan_4382	Mvan_5782
Mvan_0553	Mvan_1637	Mvan_3055	Mvan_4385	Mvan_5785
Mvan_0558	Mvan_1646	Mvan_3056	Mvan_4386	Mvan_5786
Mvan_0559	Mvan_1647	Mvan_3062	Mvan_4387	Mvan_5788
Mvan_0563	Mvan_1649	Mvan_3072	Mvan_4395	Mvan_5794
Mvan_0564	Mvan_1650	Mvan_3091	Mvan_4396	Mvan_5797
Mvan_0565	Mvan_1657	Mvan_3092	Mvan_4397	Mvan_5798
Mvan_0568	Mvan_1658	Mvan_3093	Mvan_4398	Mvan_5799
Mvan_0574	Mvan_1666	Mvan_3096	Mvan_4400	Mvan_5800
Mvan_0575	Mvan_1667	Mvan_3099	Mvan_4401	Mvan_5803
Mvan_0576	Mvan_1670	Mvan_3100	Mvan_4402	Mvan_5805
Mvan_0579	Mvan_1671	Mvan_3102	Mvan_4403	Mvan_5806
Mvan_0580	Mvan_1672	Mvan_3103	Mvan_4404	Mvan_5807
Mvan_0582	Mvan_1673	Mvan_3106	Mvan_4405	Mvan_5809
Mvan_0583	Mvan_1675	Mvan_3110	Mvan_4407	Mvan_5811
Mvan_0586	Mvan_1679	Mvan_3111	Mvan_4408	Mvan_5813
Mvan_0587	Mvan_1680	Mvan_3114	Mvan_4410	Mvan_5815
Mvan_0588	Mvan_1683	Mvan_3117	Mvan_4411	Mvan_5816
Mvan_0595	Mvan_1684	Mvan_3141	Mvan_4413	Mvan_5817
Mvan_0599	Mvan_1685	Mvan_3142	Mvan_4414	Mvan_5818
Mvan_0605	Mvan_1686	Mvan_3143	Mvan_4417	Mvan_5820

Mvan_0606	Mvan_1690	Mvan_3144	Mvan_4418	Mvan_5821
Mvan_0607	Mvan_1691	Mvan_3145	Mvan_4420	Mvan_5823
Mvan_0608	Mvan_1692	Mvan_3146	Mvan_4425	Mvan_5824
Mvan_0609	Mvan_1693	Mvan_3148	Mvan_4442	Mvan_5825
Mvan_0610	Mvan_1696	Mvan_3149	Mvan_4447	Mvan_5826
Mvan_0611	Mvan_1697	Mvan_3151	Mvan_4450	Mvan_5827
Mvan_0612	Mvan_1698	Mvan_3152	Mvan_4455	Mvan_5828
Mvan_0613	Mvan_1699	Mvan_3153	Mvan_4462	Mvan_5830
Mvan_0614	Mvan_1700	Mvan_3154	Mvan_4463	Mvan_5833
Mvan_0615	Mvan_1702	Mvan_3155	Mvan_4486	Mvan_5834
Mvan_0616	Mvan_1705	Mvan_3156	Mvan_4502	Mvan_5835
Mvan_0622	Mvan_1706	Mvan_3157	Mvan_4513	Mvan_5842
Mvan_0623	Mvan_1709	Mvan_3158	Mvan_4516	Mvan_5845
Mvan_0626	Mvan_1710	Mvan_3160	Mvan_4517	Mvan_5850
Mvan_0627	Mvan_1711	Mvan_3161	Mvan_4518	Mvan_5858
Mvan_0628	Mvan_1718	Mvan_3177	Mvan_4530	Mvan_5862
Mvan_0632	Mvan_1723	Mvan_3178	Mvan_4532	Mvan_5863
Mvan_0633	Mvan_1724	Mvan_3192	Mvan_4536	Mvan_5864
Mvan_0640	Mvan_1741	Mvan_3193	Mvan_4541	Mvan_5867
Mvan_0645	Mvan_1749	Mvan_3195	Mvan_4542	Mvan_5868
Mvan_0647	Mvan_1751	Mvan_3199	Mvan_4551	Mvan_5869
Mvan_0648	Mvan_1755	Mvan_3204	Mvan_4555	Mvan_5870
Mvan_0649	Mvan_1756	Mvan_3209	Mvan_4557	Mvan_5871
Mvan_0650	Mvan_1757	Mvan_3211	Mvan_4558	Mvan_5872
Mvan_0651	Mvan_1758	Mvan_3213	Mvan_4559	Mvan_5873
Mvan_0652	Mvan_1773	Mvan_3226	Mvan_4560	Mvan_5875
Mvan_0653	Mvan_1774	Mvan_3227	Mvan_4561	Mvan_5876
Mvan_0655	Mvan_1793	Mvan_3228	Mvan_4562	Mvan_5877
Mvan_0656	Mvan_1808	Mvan_3230	Mvan_4563	Mvan_5878
Mvan_0657	Mvan_1815	Mvan_3232	Mvan_4564	Mvan_5879
Mvan_0658	Mvan_1818	Mvan_3233	Mvan_4568	Mvan_5880
Mvan_0667	Mvan_1820	Mvan_3236	Mvan_4574	Mvan_5881
Mvan_0679	Mvan_1821	Mvan_3237	Mvan_4576	Mvan_5882
Mvan_0681	Mvan_1822	Mvan_3238	Mvan_4581	Mvan_5887
Mvan_0682	Mvan_1828	Mvan_3239	Mvan_4584	Mvan_5888
Mvan_0689	Mvan_1829	Mvan_3240	Mvan_4585	Mvan_5889
Mvan_0691	Mvan_1830	Mvan_3241	Mvan_4586	Mvan_5890
Mvan_0694	Mvan_1831	Mvan_3244	Mvan_4611	Mvan_5893
Mvan_0696	Mvan_1838	Mvan_3245	Mvan_4615	Mvan_5897
Mvan_0702	Mvan_1839	Mvan_3246	Mvan_4616	Mvan_5898
Mvan_0717	Mvan_1840	Mvan_3247	Mvan_4617	Mvan_5899

Mvan_0720	Mvan_1843	Mvan_3248	Mvan_4618	Mvan_5900
Mvan_0727	Mvan_1846	Mvan_3253	Mvan_4619	Mvan_5903
Mvan_0729	Mvan_1847	Mvan_3255	Mvan_4620	Mvan_5904
Mvan_0738	Mvan_1848	Mvan_3258	Mvan_4621	Mvan_5905
Mvan_0747	Mvan_1849	Mvan_3262	Mvan_4622	Mvan_5906
Mvan_0748	Mvan_1850	Mvan_3263	Mvan_4623	Mvan_5907
Mvan_0753	Mvan_1851	Mvan_3265	Mvan_4624	Mvan_5908
Mvan_0770	Mvan_1852	Mvan_3270	Mvan_4626	Mvan_5909
Mvan_0778	Mvan_1853	Mvan_3271	Mvan_4627	Mvan_5910
Mvan_0779	Mvan_1860	Mvan_3273	Mvan_4628	Mvan_5913
Mvan_0784	Mvan_1863	Mvan_3286	Mvan_4630	Mvan_5914
Mvan_0786	Mvan_1871	Mvan_3287	Mvan_4636	Mvan_5916
Mvan_0787	Mvan_1889	Mvan_3300	Mvan_4637	Mvan_5918
Mvan_0789	Mvan_1905	Mvan_3315	Mvan_4641	Mvan_5919
Mvan_0790	Mvan_1913	Mvan_3318	Mvan_4643	Mvan_5920
Mvan_0793	Mvan_1918	Mvan_3319	Mvan_4650	Mvan_5921
Mvan_0814	Mvan_1920	Mvan_3324	Mvan_4685	Mvan_5924
Mvan_0815	Mvan_1921	Mvan_3338	Mvan_4690	Mvan_5925
Mvan_0820	Mvan_1923	Mvan_3343	Mvan_4695	Mvan_5926
Mvan_0839	Mvan_1929	Mvan_3345	Mvan_4697	Mvan_5930
Mvan_0849	Mvan_1930	Mvan_3346	Mvan_4698	Mvan_5931
Mvan_0851	Mvan_1933	Mvan_3347	Mvan_4700	Mvan_5933
Mvan_0853	Mvan_1934	Mvan_3348	Mvan_4701	Mvan_5934
Mvan_0856	Mvan_1935	Mvan_3363	Mvan_4703	Mvan_5935
Mvan_0857	Mvan_1938	Mvan_3364	Mvan_4704	Mvan_5937
Mvan_0860	Mvan_1940	Mvan_3367	Mvan_4705	Mvan_5938
Mvan_0861	Mvan_1941	Mvan_3369	Mvan_4706	Mvan_5939
Mvan_0862	Mvan_1954	Mvan_3370	Mvan_4707	Mvan_5941
Mvan_0873	Mvan_1955	Mvan_3371	Mvan_4711	Mvan_5942
Mvan_0884	Mvan_1956	Mvan_3380	Mvan_4719	Mvan_5943
Mvan_0885	Mvan_1957	Mvan_3381	Mvan_4721	Mvan_5946
Mvan_0890	Mvan_1966	Mvan_3383	Mvan_4723	Mvan_5947
Mvan_0892	Mvan_1979	Mvan_3384	Mvan_4724	Mvan_5948
Mvan_0893	Mvan_2000	Mvan_3385	Mvan_4727	Mvan_5949
Mvan_0895	Mvan_2009	Mvan_3387	Mvan_4733	Mvan_5950
Mvan_0896	Mvan_2015	Mvan_3388	Mvan_4734	Mvan_5951
Mvan_0897	Mvan_2016	Mvan_3389	Mvan_4735	Mvan_5952
Mvan_0899	Mvan_2029	Mvan_3390	Mvan_4736	Mvan_5953
Mvan_0900	Mvan_2036	Mvan_3414	Mvan_4753	Mvan_5954
Mvan_0901	Mvan_2043	Mvan_3421	Mvan_4755	Mvan_5955
Mvan_0902	Mvan_2046	Mvan_3425	Mvan_4756	Mvan_5957

Mvan_0903	Mvan_2047	Mvan_3428	Mvan_4757	Mvan_5961
Mvan_0904	Mvan_2048	Mvan_3432	Mvan_4758	Mvan_5962
Mvan_0907	Mvan_2053	Mvan_3439	Mvan_4765	Mvan_5963
Mvan_0909	Mvan_2054	Mvan_3446	Mvan_4766	Mvan_5964
Mvan_0910	Mvan_2057	Mvan_3454	Mvan_4780	Mvan_5965
Mvan_0913	Mvan_2061	Mvan_3456	Mvan_4789	Mvan_5966
Mvan_0917	Mvan_2062	Mvan_3465	Mvan_4790	Mvan_5967
Mvan_0919	Mvan_2063	Mvan_3493	Mvan_4793	Mvan_5969
Mvan_0920	Mvan_2064	Mvan_3494	Mvan_4794	Mvan_5970
Mvan_0921	Mvan_2068	Mvan_3495	Mvan_4817	Mvan_5971
Mvan_0924	Mvan_2069	Mvan_3499	Mvan_4819	Mvan_5976
Mvan_0925	Mvan_2070	Mvan_3500	Mvan_4820	Mvan_5977
Mvan_0928	Mvan_2072	Mvan_3501	Mvan_4821	Mvan_5978
Mvan_0929	Mvan_2074	Mvan_3502	Mvan_4828	Mvan_5979
Mvan_0930	Mvan_2078	Mvan_3503	Mvan_4837	Mvan_5980
Mvan_0941	Mvan_2086	Mvan_3504	Mvan_4859	Mvan_5981
Mvan_0943	Mvan_2093	Mvan_3506	Mvan_4861	Mvan_5982
Mvan_0945	Mvan_2102	Mvan_3507	Mvan_4862	Mvan_5983
Mvan_0946	Mvan_2109	Mvan_3509	Mvan_4863	Mvan_5984
Mvan_0956	Mvan_2125	Mvan_3511	Mvan_4865	Mvan_5985
Mvan_0981	Mvan_2130	Mvan_3562	Mvan_4867	Mvan_5986
Mvan_0984	Mvan_2145	Mvan_3588	Mvan_4868	Mvan_5987
Mvan_0985	Mvan_2146	Mvan_3589	Mvan_4878	Mvan_5988
Mvan_0987	Mvan_2147	Mvan_3591	Mvan_4879	Mvan_5989
Mvan_0990	Mvan_2150	Mvan_3592	Mvan_4886	Mvan_5990
Mvan_0996	Mvan_2160	Mvan_3606	Mvan_4888	Mvan_5991
Mvan_1001	Mvan_2183	Mvan_3607	Mvan_4889	Mvan_5992
Mvan_1002	Mvan_2196	Mvan_3609	Mvan_4891	Mvan_5993
Mvan_1007	Mvan_2198	Mvan_3610	Mvan_4895	Mvan_5994
Mvan_1008	Mvan_2199	Mvan_3612	Mvan_4897	Mvan_5996
Mvan_1009	Mvan_2200	Mvan_3613	Mvan_4899	Mvan_5997
Mvan_1010	Mvan_2211	Mvan_3626	Mvan_4900	Mvan_5998
Mvan_1012	Mvan_2214	Mvan_3627	Mvan_4901	Mvan_6000
Mvan_1015	Mvan_2215	Mvan_3629	Mvan_4902	Mvan_6001
Mvan_1016	Mvan_2217	Mvan_3630	Mvan_4903	Mvan_6002
Mvan_1018	Mvan_2218	Mvan_3632	Mvan_4904	Mvan_6003
Mvan_1019	Mvan_2224	Mvan_3633	Mvan_4905	Mvan_6004
Mvan_1020	Mvan_2225	Mvan_3634	Mvan_4914	Mvan_6005
Mvan_1021	Mvan_2226	Mvan_3637	Mvan_4918	Mvan_6006
Mvan_1022	Mvan_2230	Mvan_3639	Mvan_4919	Mvan_6007
Mvan_1023	Mvan_2231	Mvan_3642	Mvan_4920	Mvan_6008

Mvan_1030	Mvan_2232	Mvan_3653	Mvan_4923	Mvan_6009
Mvan_1031	Mvan_2233	Mvan_3654	Mvan_4924	Mvan_6010
Mvan_1034	Mvan_2234	Mvan_3655	Mvan_4934	Mvan_6011
Mvan_1035	Mvan_2235	Mvan_3656	Mvan_4942	Mvan_6012
Mvan_1036	Mvan_2247	Mvan_3657	Mvan_4944	Mvan_6013
Mvan_1041	Mvan_2250	Mvan_3658	Mvan_4945	Mvan_6014
Mvan_1042	Mvan_2253	Mvan_3659	Mvan_4948	Mvan_6015
Mvan_1044	Mvan_2254	Mvan_3660	Mvan_4949	Mvan_6016
Mvan_1045	Mvan_2256	Mvan_3662	Mvan_4955	Mvan_6017
Mvan_1046	Mvan_2271	Mvan_3663	Mvan_4968	Mvan_6019
Mvan_1047	Mvan_2277	Mvan_3664	Mvan_4977	Mvan_6021
Mvan_1048	Mvan_2282	Mvan_3665	Mvan_4979	Mvan_6022
Mvan_1049	Mvan_2283	Mvan_3666	Mvan_4980	Mvan_6023
Mvan_1050	Mvan_2285	Mvan_3667	Mvan_4981	Mvan_6024
Mvan_1052	Mvan_2286	Mvan_3668	Mvan_4982	Mvan_6025
Mvan_1053	Mvan_2289	Mvan_3673	Mvan_4985	Mvan_6026
Mvan_1054	Mvan_2291	Mvan_3676	Mvan_4994	Mvan_6027
Mvan_1055	Mvan_2292	Mvan_3679	Mvan_4997	Mvan_6028
Mvan_1063	Mvan_2293	Mvan_3680	Mvan_4999	Mvan_6043
Mvan_1064	Mvan_2294	Mvan_3682	Mvan_5000	Mvan_6044
Mvan_1067	Mvan_2297	Mvan_3687	Mvan_5003	Mvan_6052
Mvan_1069	Mvan_2307	Mvan_3689	Mvan_5004	Mvan_6056
Mvan_1072	Mvan_2310	Mvan_3690	Mvan_5007	
Mvan_1074	Mvan_2319	Mvan_3691	Mvan_5014	

For this project focus will be on the genes unique to *M. tuberculosis* (Table 5), which might give further insight into the virulence of this organism causing fatal disease and to investigate the possibility of one or more of these unique genes being a novel drug target.

To further investigate the uniqueness of 200 genes of *M. tuberculosis* H37Rv that were identified comparing to 7 selected genomes, the genome of the 96 species belonging to the Mycobacterium tuberculosis complex were compared following the protocol previously published by our laboratory ().

4.1.1 Analysis of gene homologs

To further investigate the uniqueness of the 200 genes identified for *M. tuberculosis* H37Rv when compared to 8 selected genomes (Table 5, Section 4.1), the genes were compared to the genomes of 96 mycobacterial species following the protocol previously published by our laboratory (Van wyk *et al*, 2019).

The output for the analysis of homology between the unique genes of *M. tuberculosis* and the 96 mycobacterial species is presented in Table 13. Only the top hits for each species is presented in the table with the gene code, name and percentage identity and homology.

Based on these results, a separate table was populated to present possible homologs and is presented in Table 14.

Due to the nature of the data, these tables are provided as excel sheets and are presented as a standalone file with the thesis.

Table 13 Comparison of unique genes of *M. tuberculosis* to 96 species in the Mycobacterium tuberculosis complex

Sheet 1: Blast results tabulated presenting only the top hits against each of the 96 species in the Mycobacterium complex.

Sheet 2: The amino acid sequences of each of the M. tuberculosis genes used for the comparison.

Sheet 3: Definition of the Mycobacterium codes used for the 96 species in the Mycobacterium complex used for the comparison.

Table 14 Presentation of possible homologs from the comparison of unique genes of *M. tuberculosis* to 96 species in the Mycobacterium tuberculosis complex

Green indicates likely homologs with identity match ≥60%.

Yellow indicates possible homologs with identity match ≥50% and homology ≥60%

Red indicate unlikely homologs.

4.2. Genome mapping

Gene mapping is used to identify the location of a gene and the distance between the genes on the chromosome. Genetic mapping can help us to identify new genes and to understand their function (Genetic Mapping Fact Sheets, 2015). Gene mapping is usually the starting point for any downstream studies.

Synteny analysis is a term used to describe the physical location of a gene on the chromosome (Syed *et a*l, 2014). Synteny analysis can provide excellent information on genome duplication of the unique genes identified during the genomic comparison, where localization of these genes belonging to the same family adjacent to one another is a direct indication that these genes are possibly duplicated during the evolution (Qhanya et al., 2015). As part of this chapter, the physical localization (Synteny) of protein identified in Chapter 2 on the chromosome of *M. tuberculosis* was mapped along with up-stream and down-stream proteins.

Genome mapping of proteins refer to the identification of neighbouring genes with respect to the proteins. Genome mapping reveals information on the possible role of the proteins in a physiological function. For example, CYP128A1 of *M. tuberculosis* function is predicted based on its physical localization with genes involved in the biosynthesis of menaquinone-like molecule (Holsclaw *et al.*, 2008).

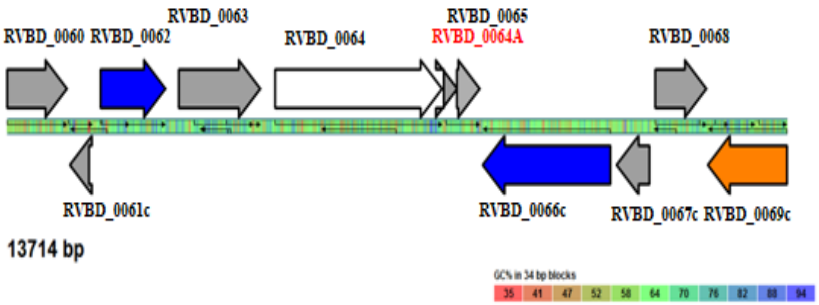
4.2.1 Pathway categories identified on mapping of the genes

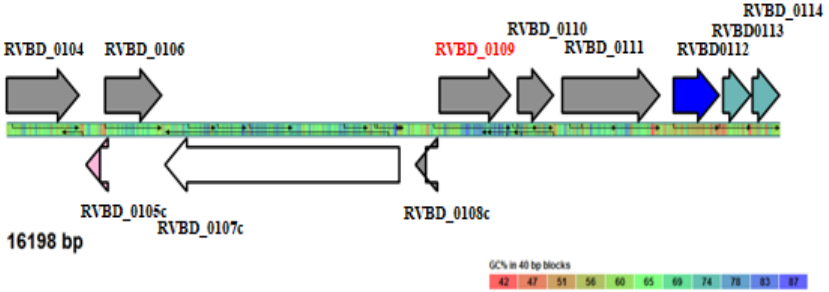
The pathway categories are represented by the colours of the proteins, indicated below:

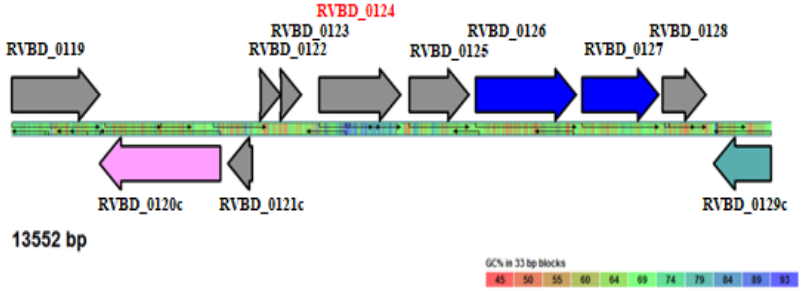
- Carbohydrate metabolism
- Energy metabolism
- Lipid metabolism
- Nucleotide metabolism
- Amino acid metabolism
- Metabolism of other amino acids
- Glycan biosynthesis and metabolism
- Metabolism of cofactors and vitamins
- Metabolism of terpenoids and polyketides
- Biosynthesis of other secondary metabolites
- Xenobiotics biodegradation and metabolism
- Enzyme families
- Genetic information processing
- Environmental information processing
- Cellular processes
- Organismal systems
- Human diseases
- Unclassified

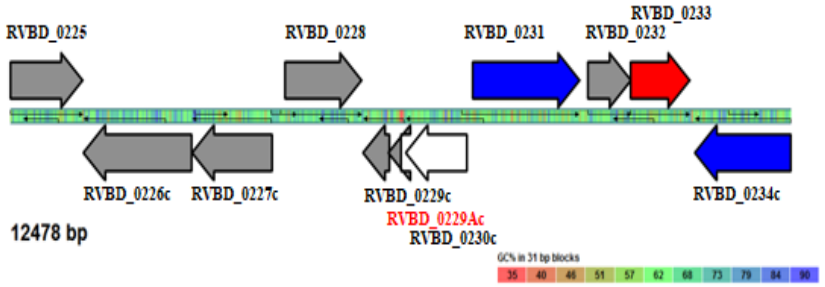
4.2.2. Representation of the gene maps

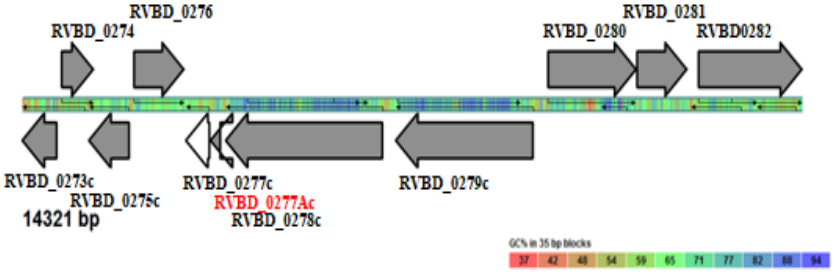
The unique proteins with 5 upstream and 5 downstream proteins were mapped and is presented below:

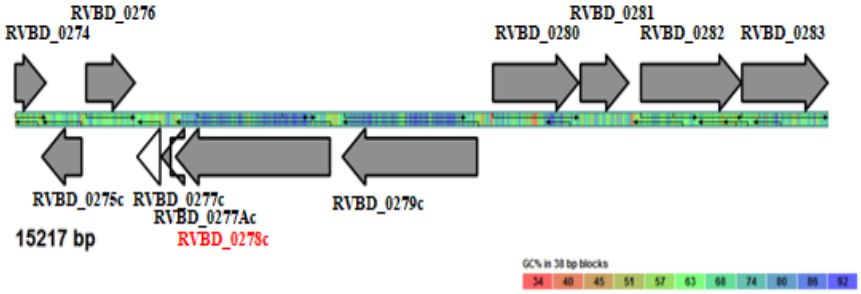


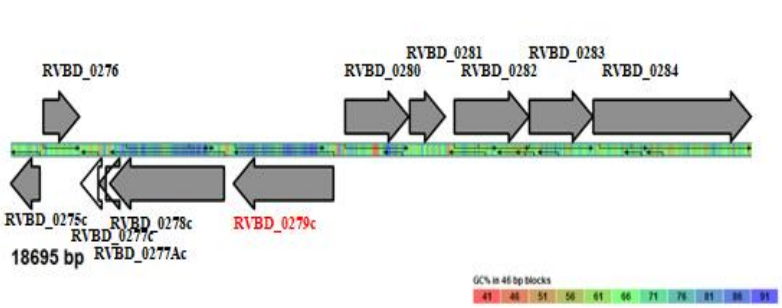


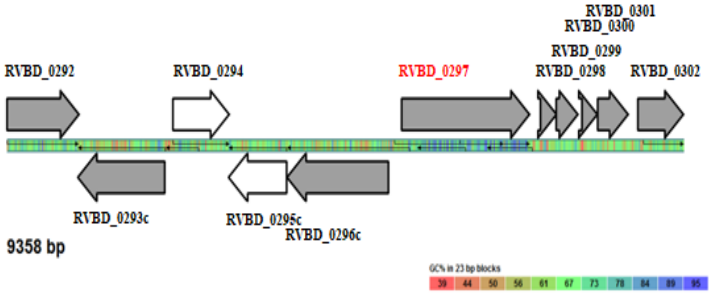


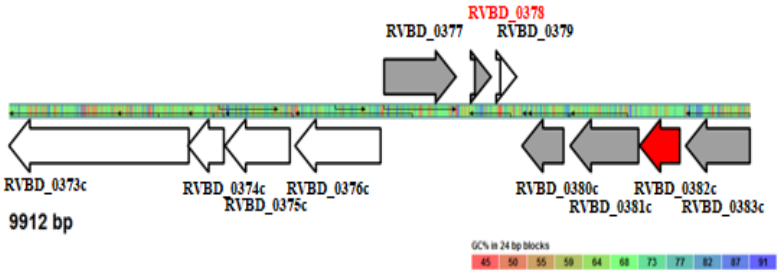


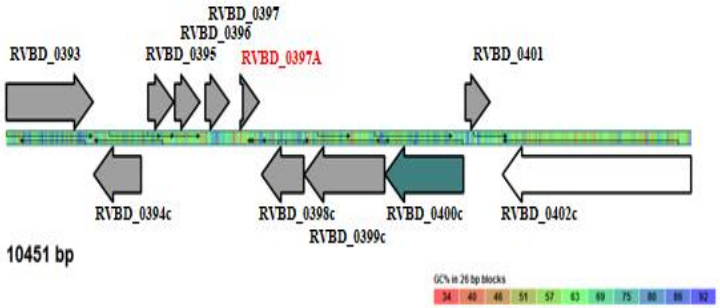


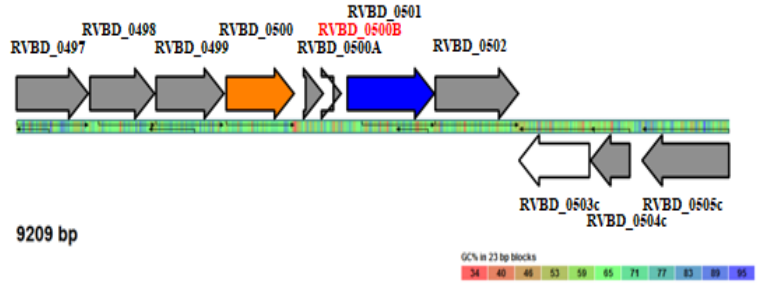


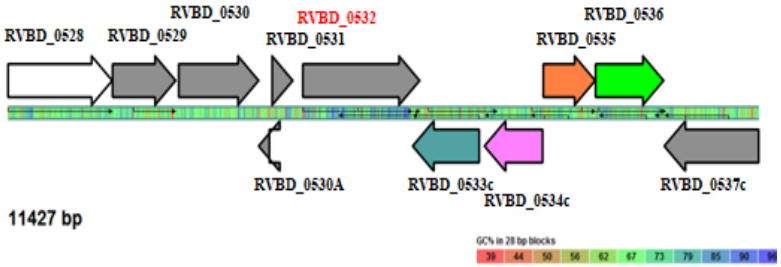


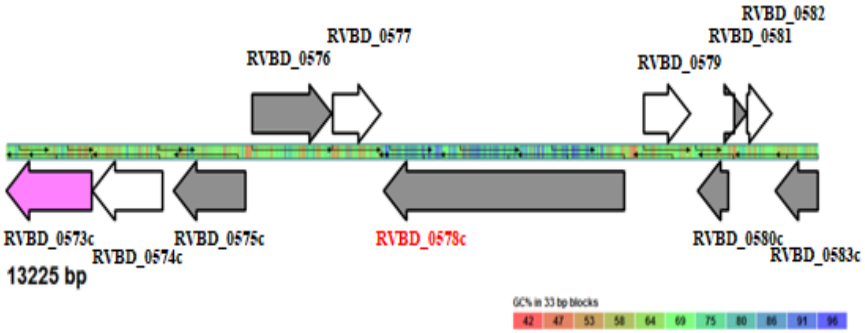


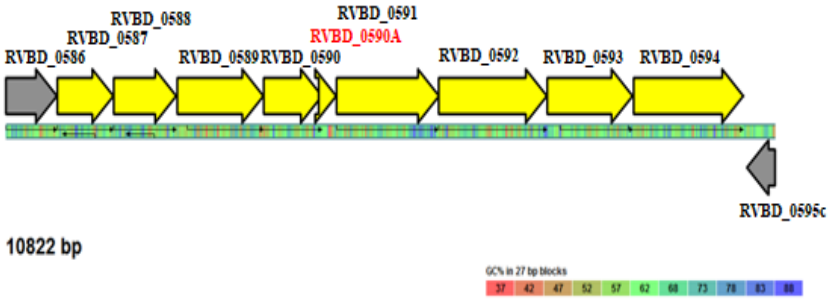


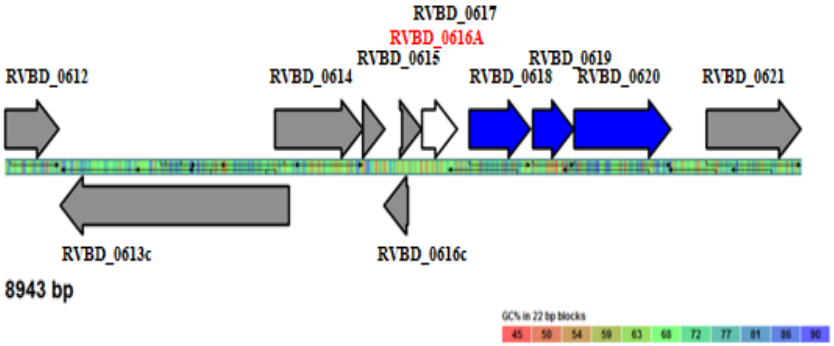


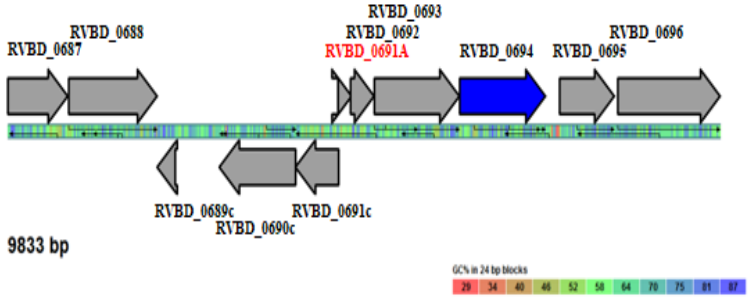


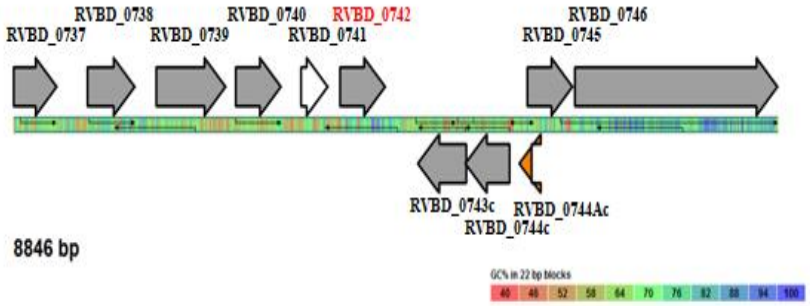


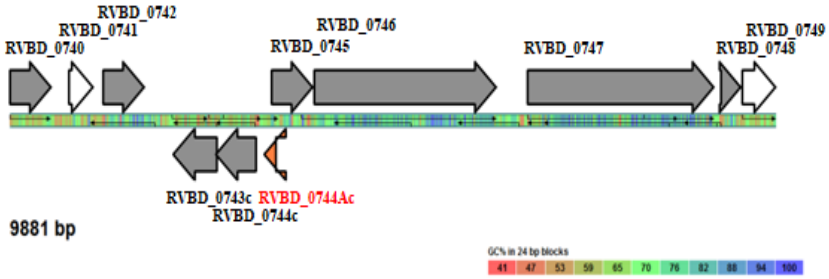


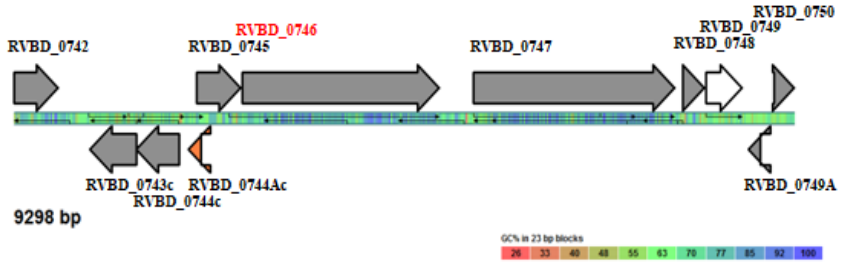


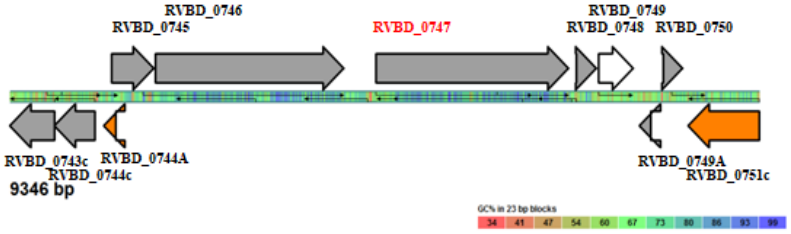


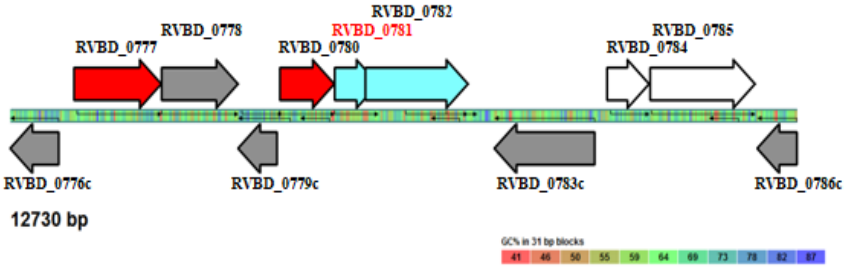


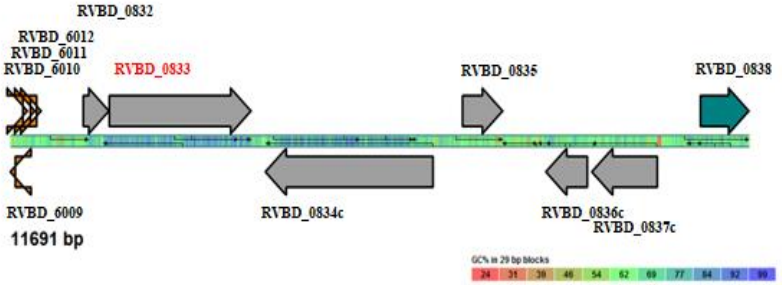


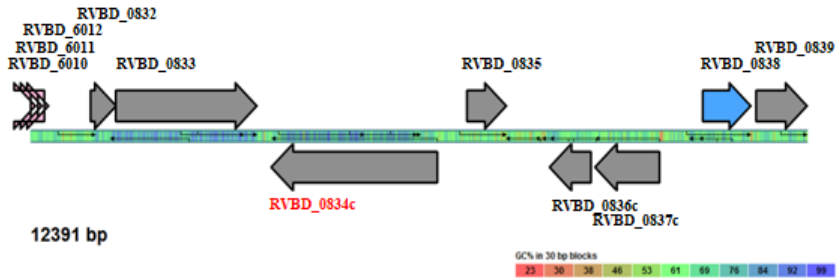


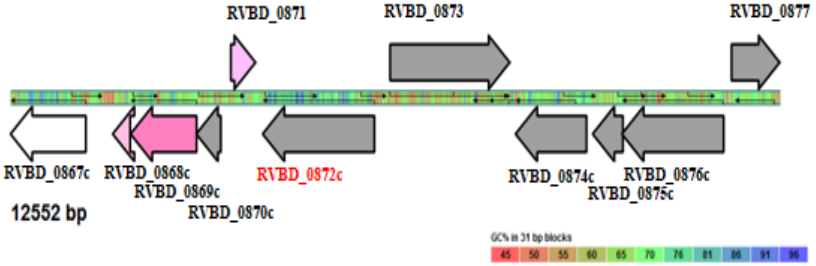


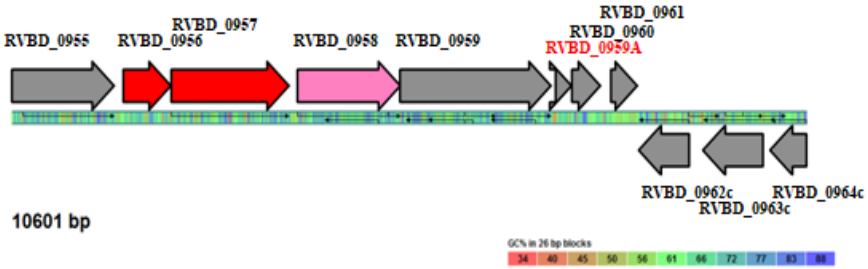


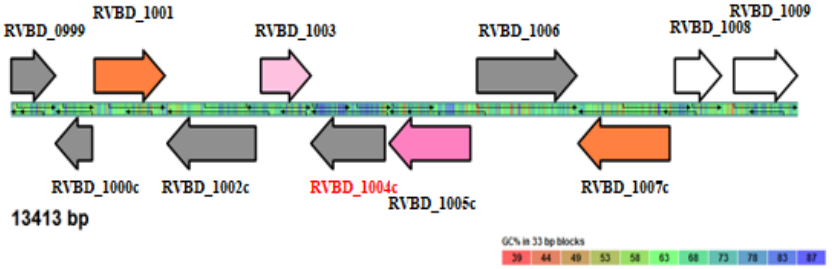


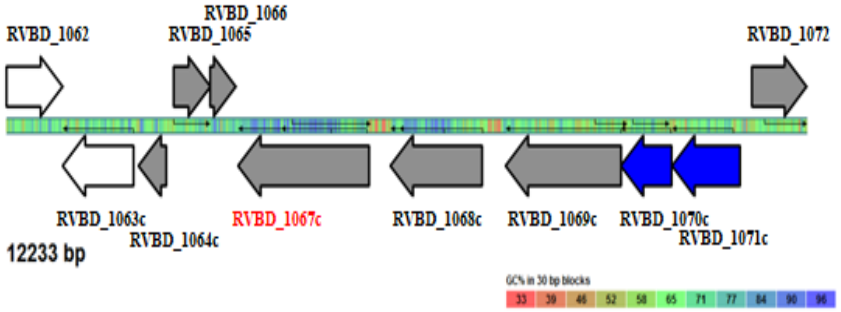


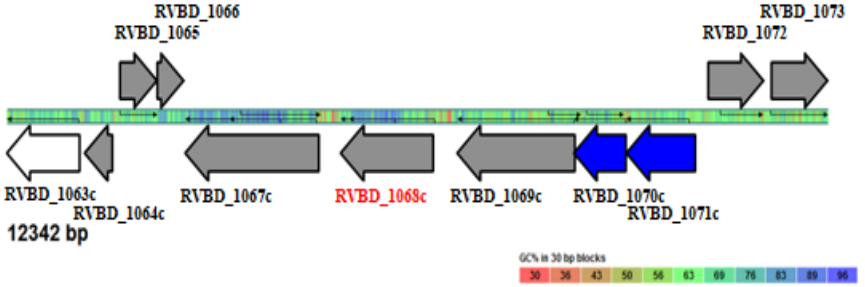


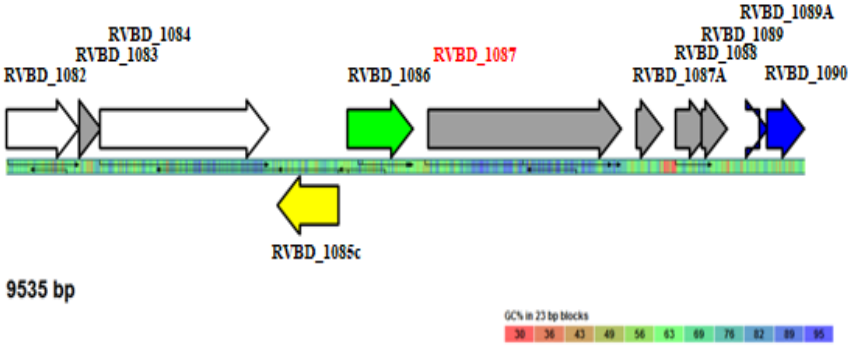


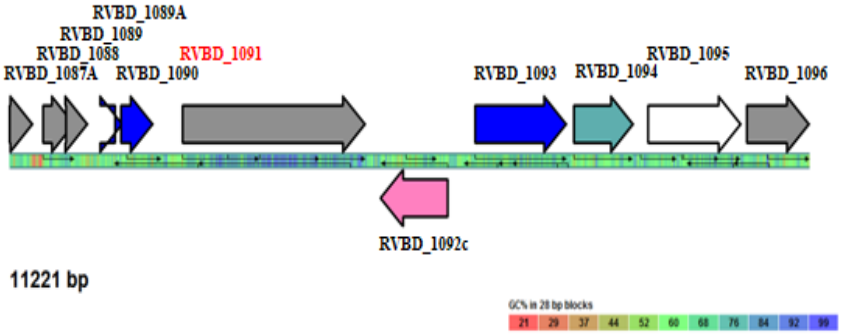


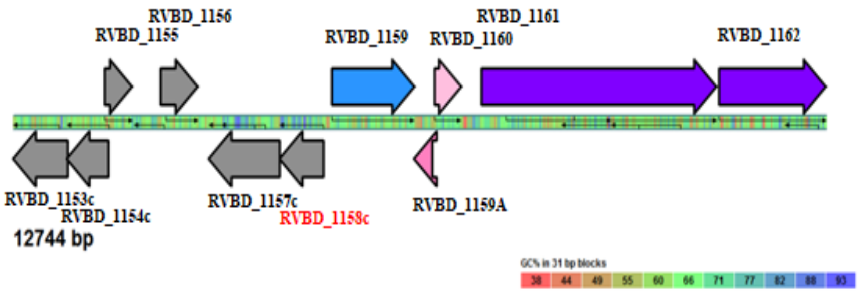


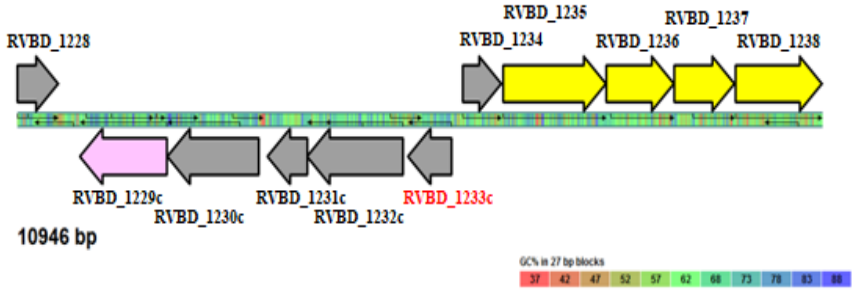


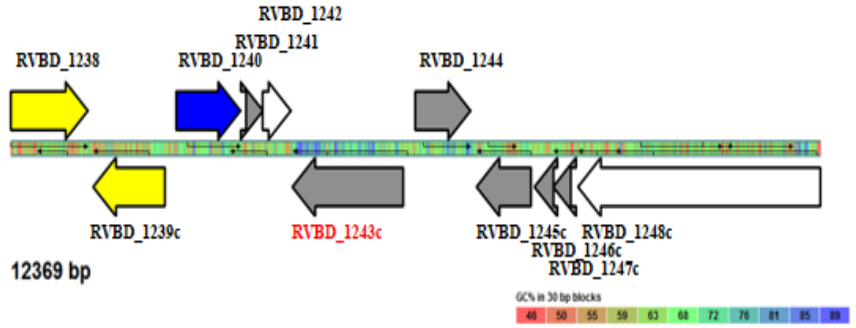


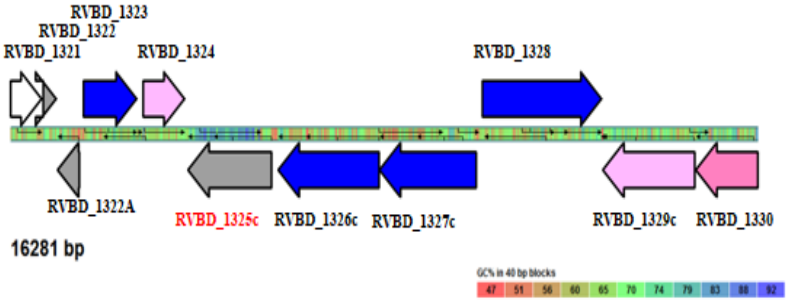


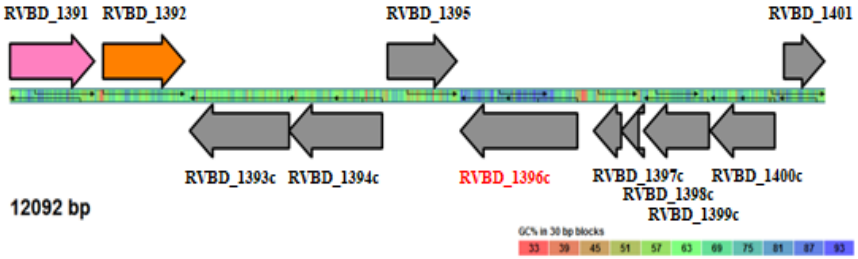


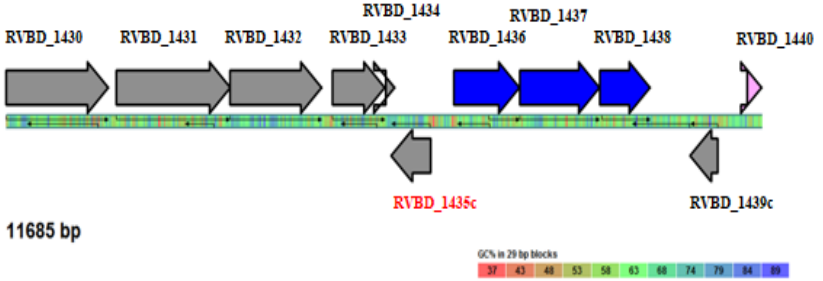


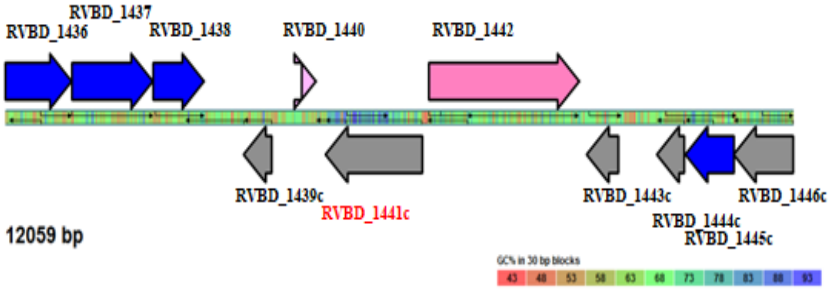


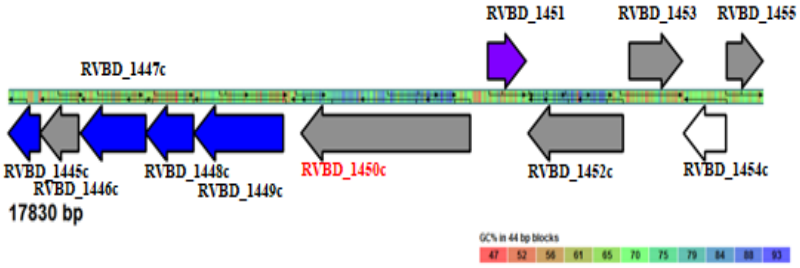


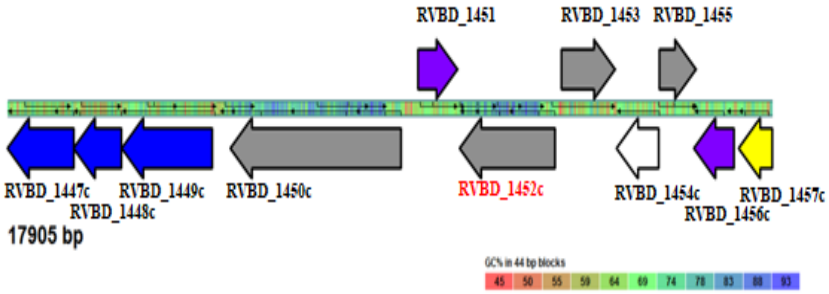


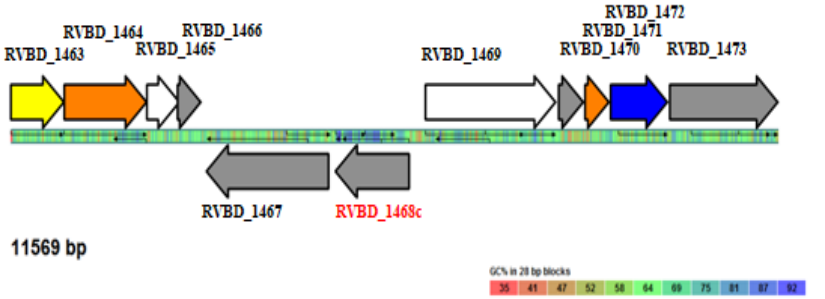


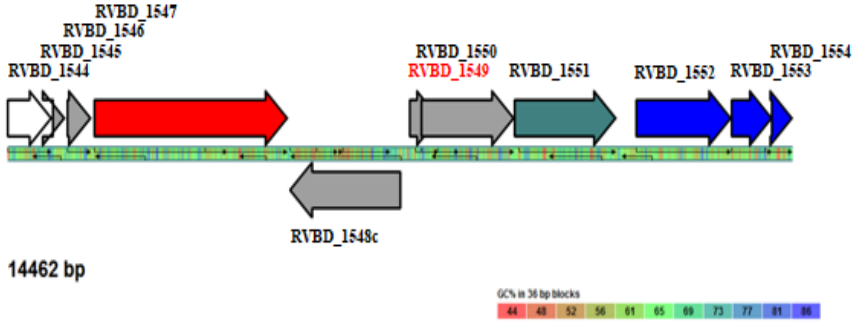


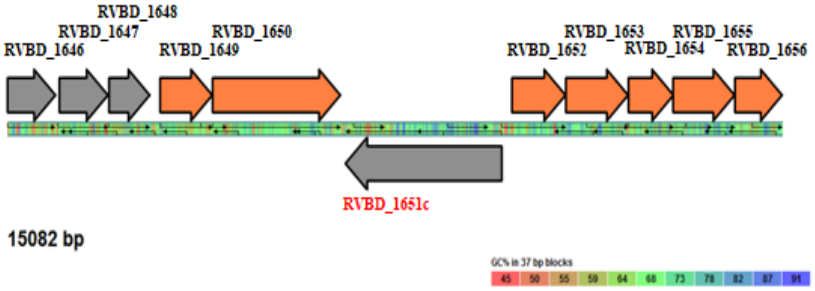


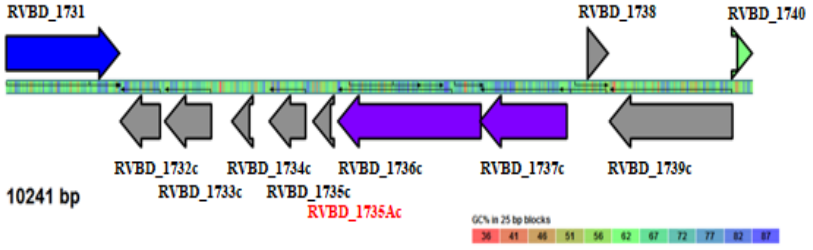


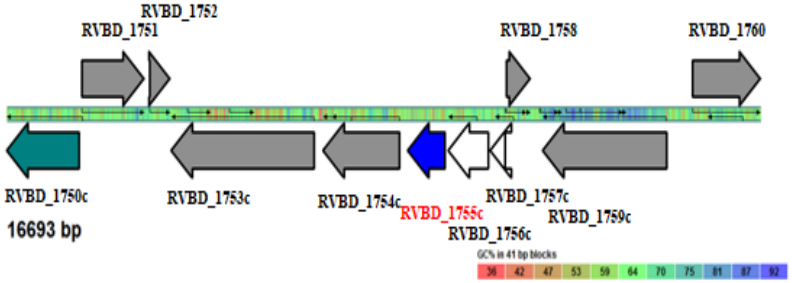


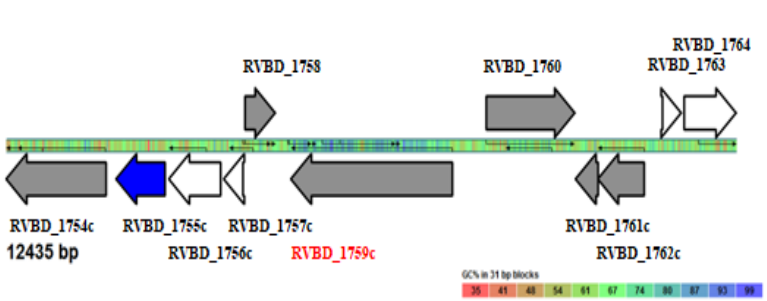


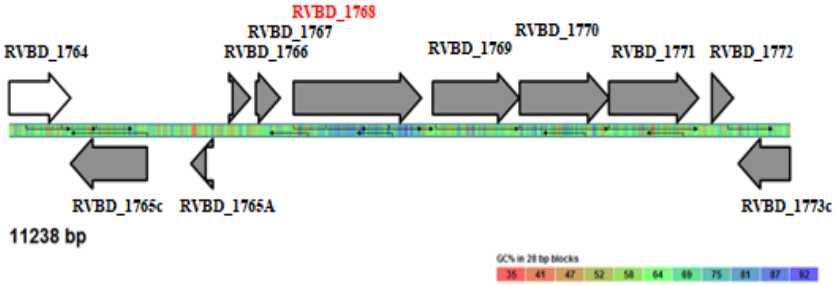


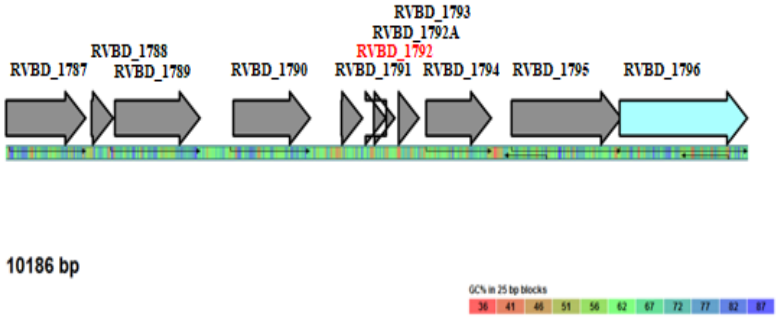


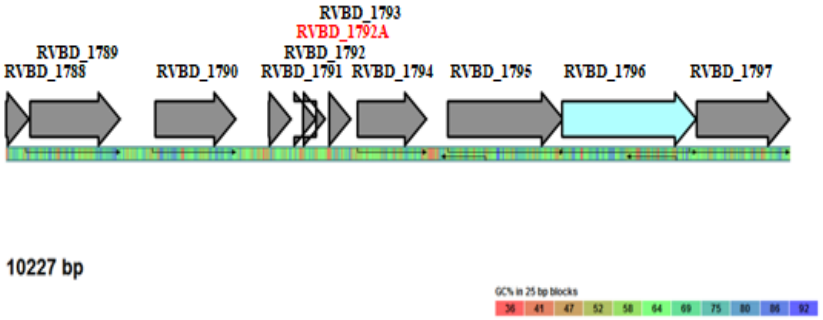


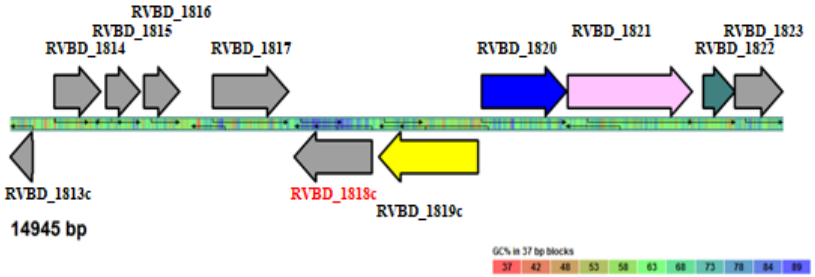


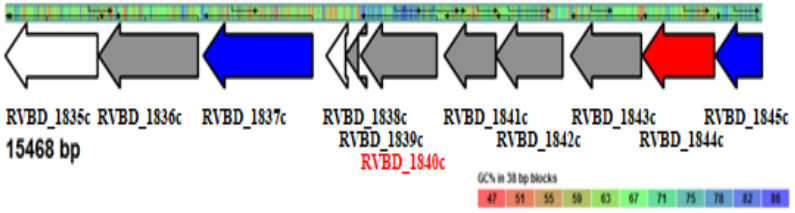


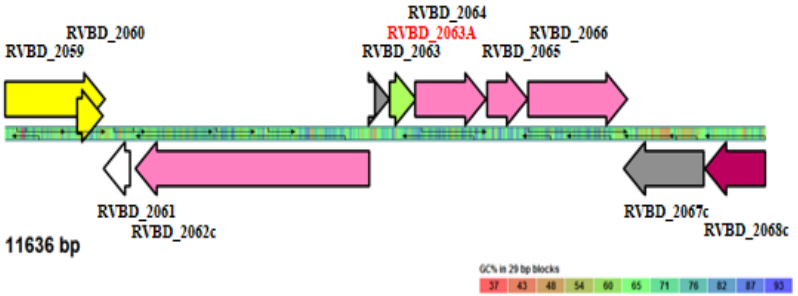


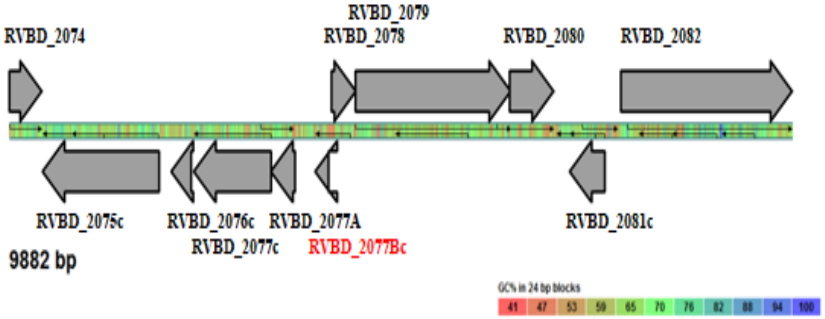


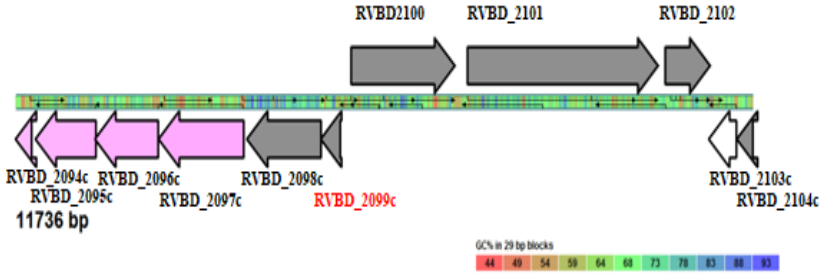


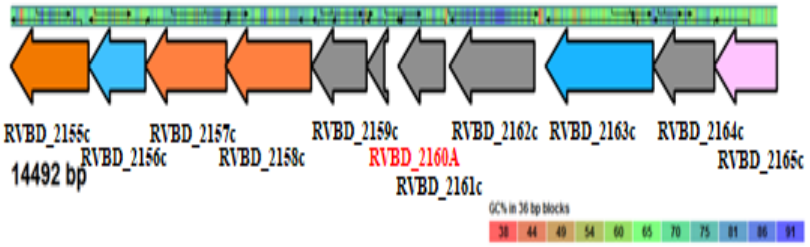


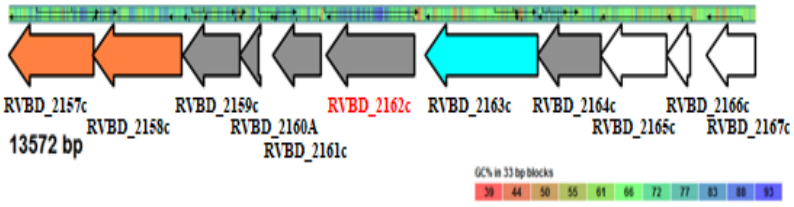


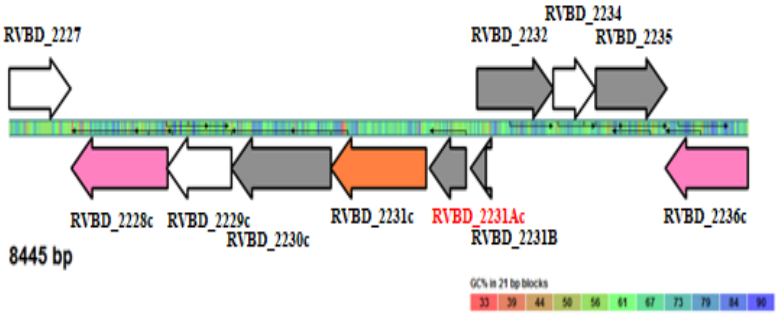


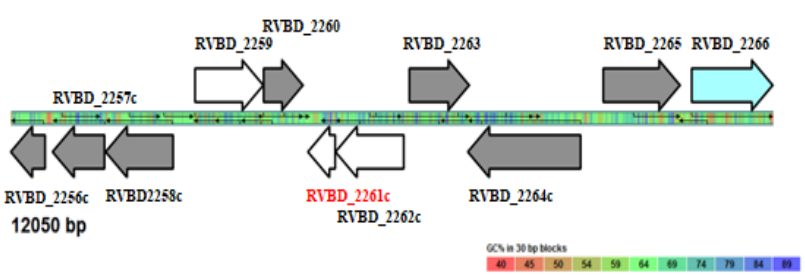


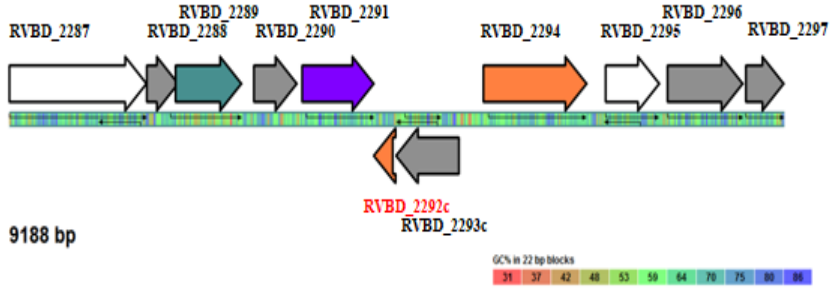


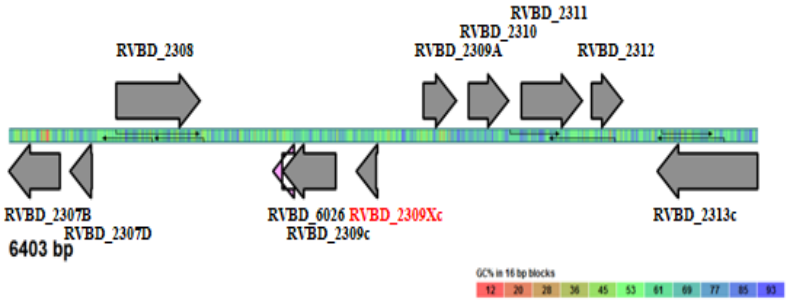


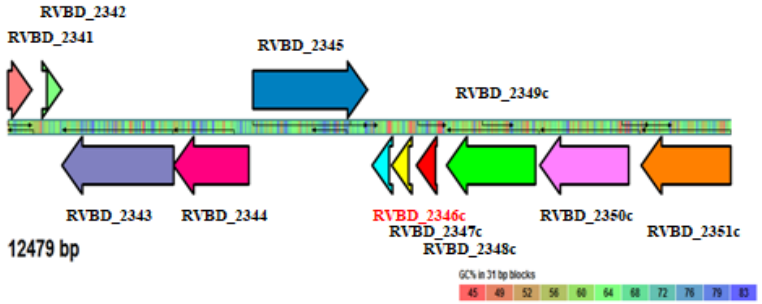


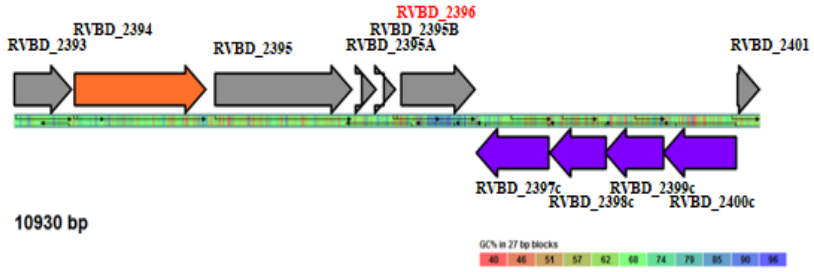


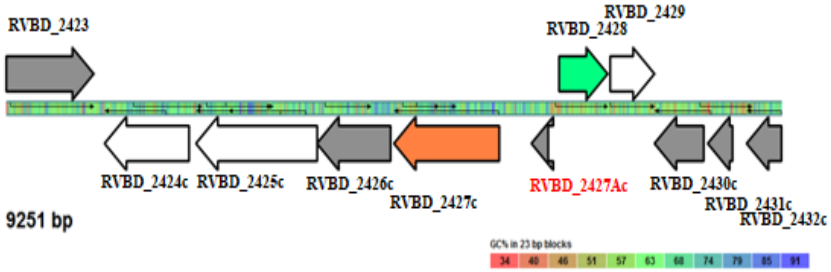


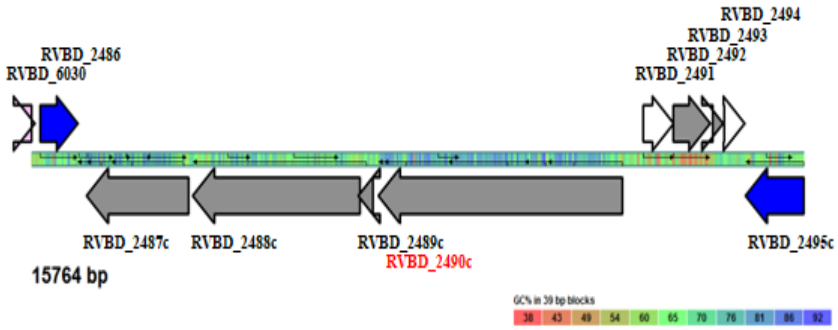


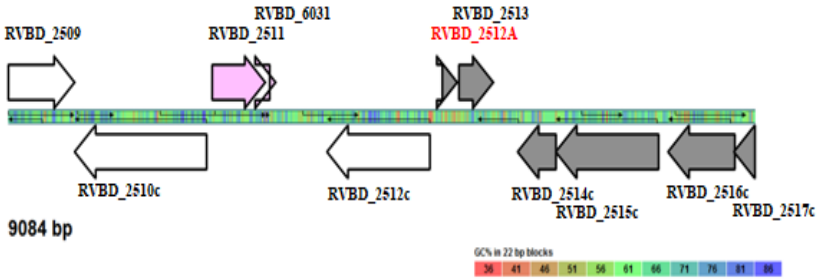


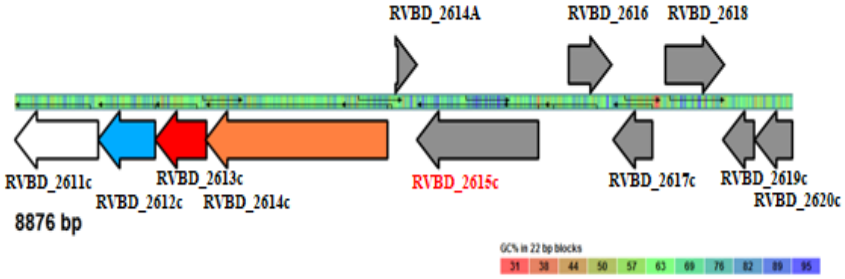


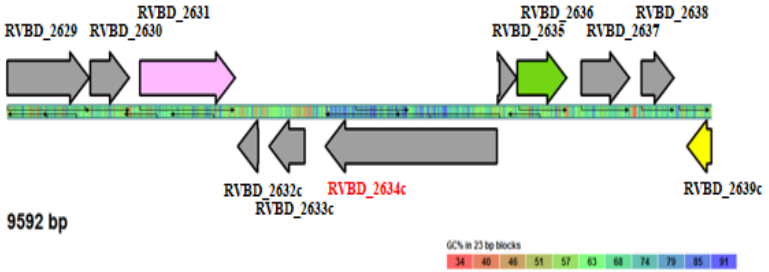


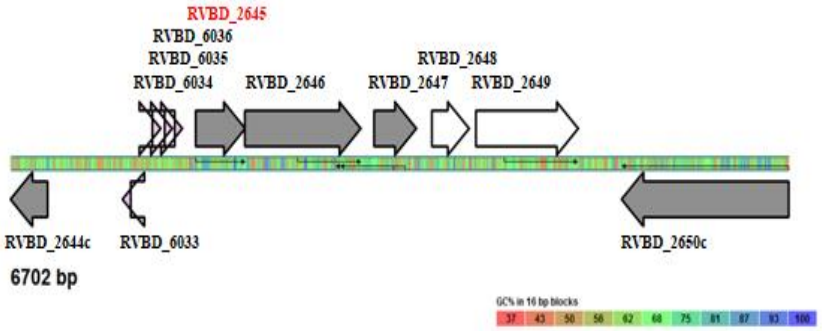


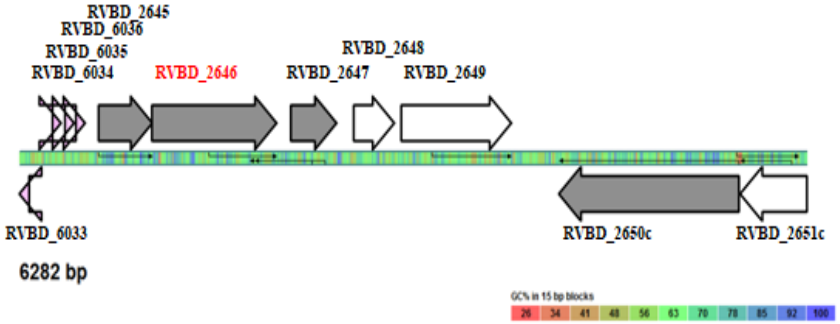


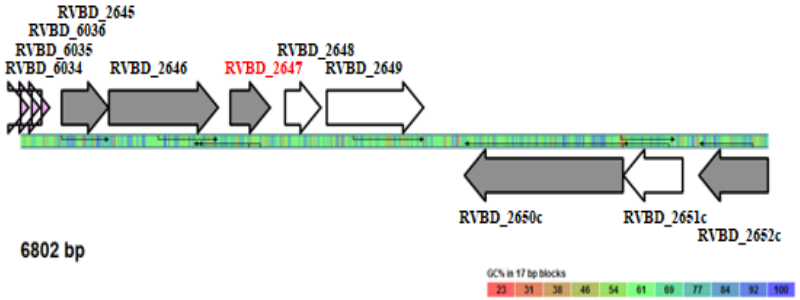


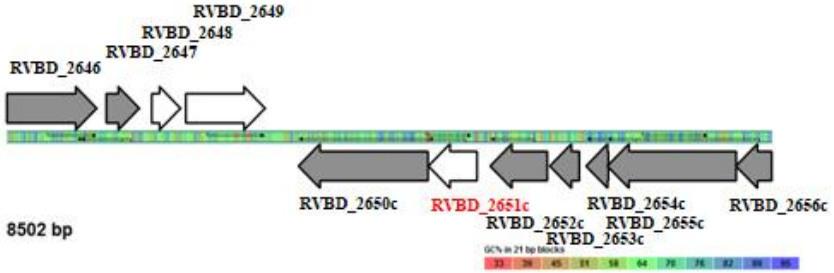


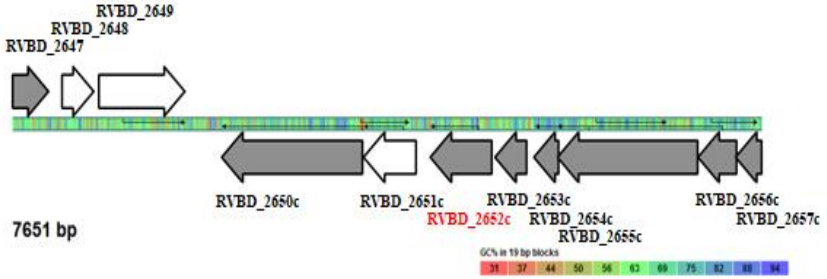


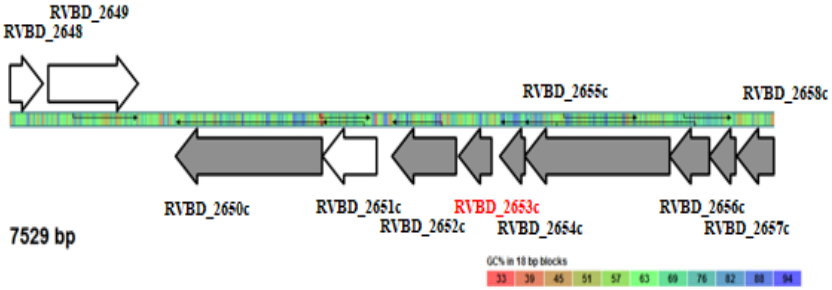


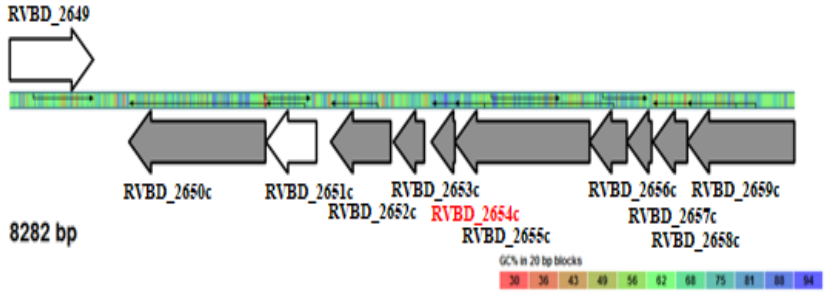


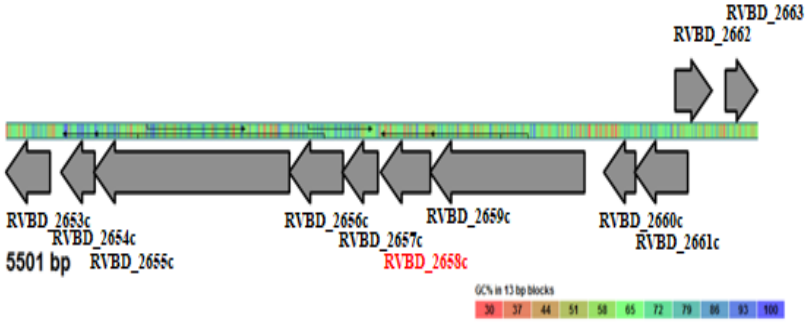


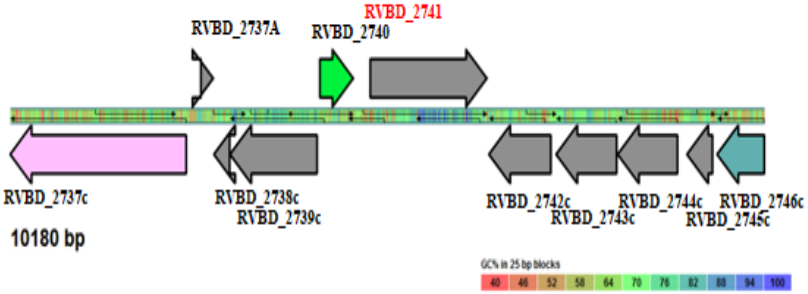


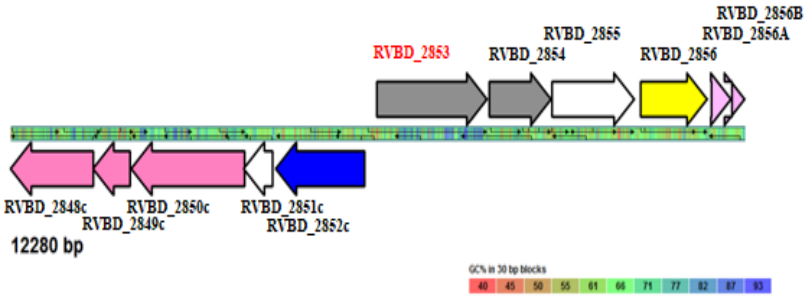


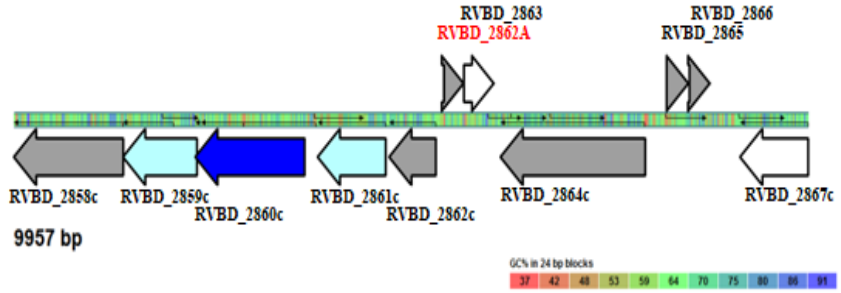


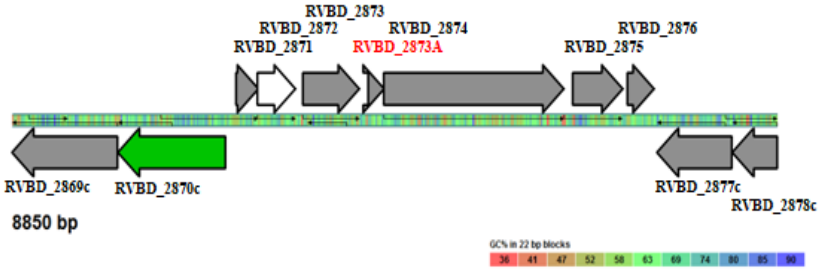


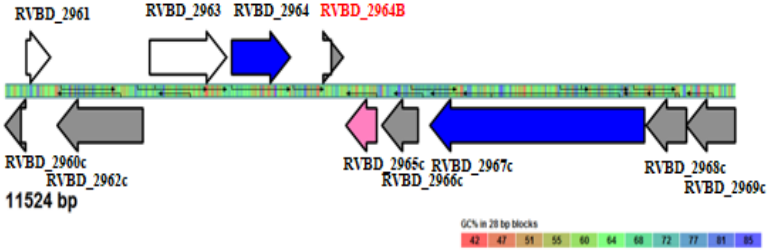


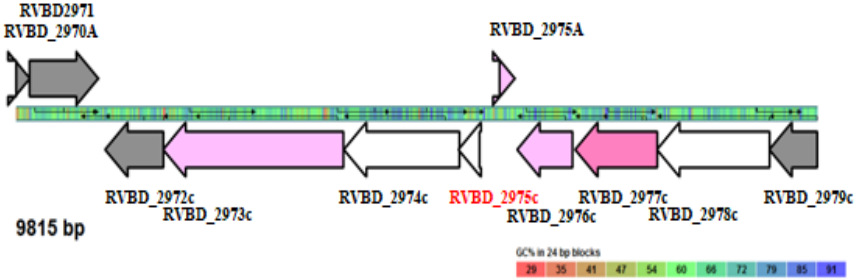


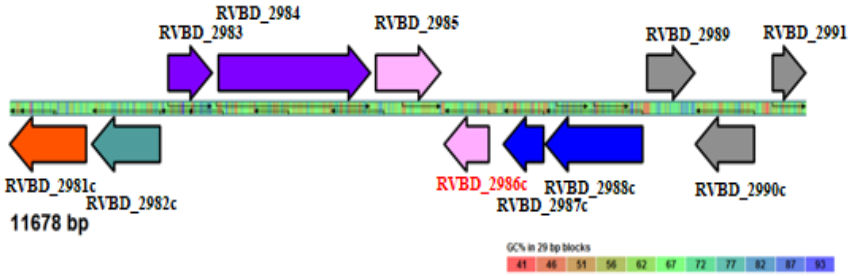


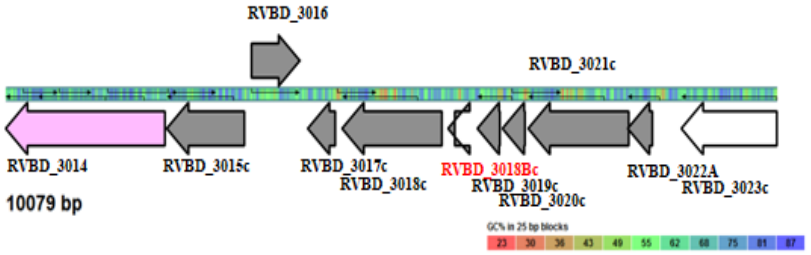


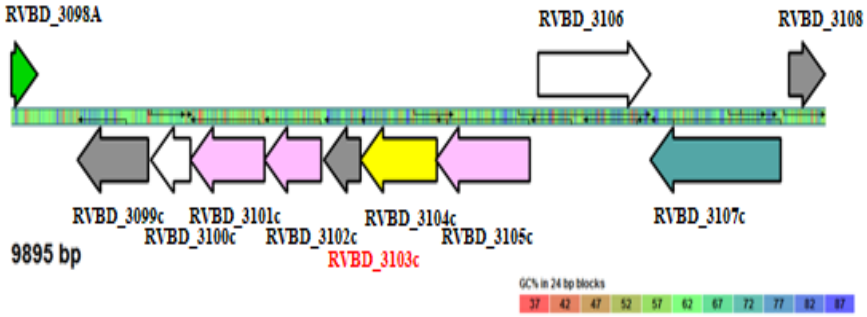


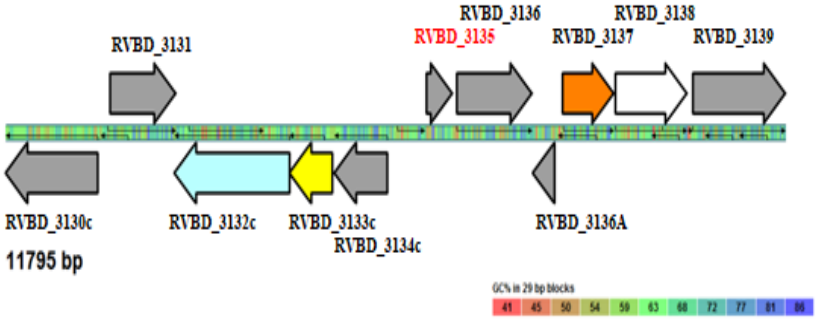


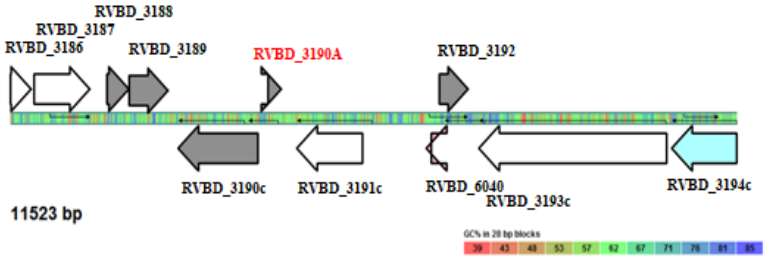


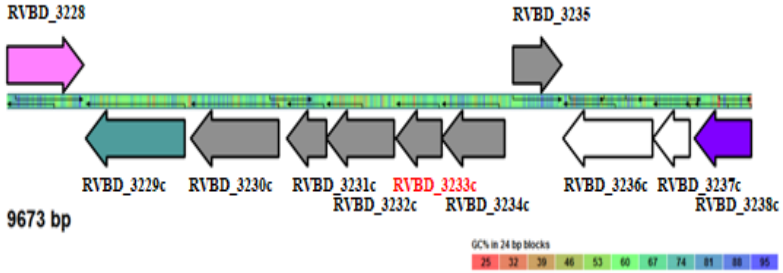


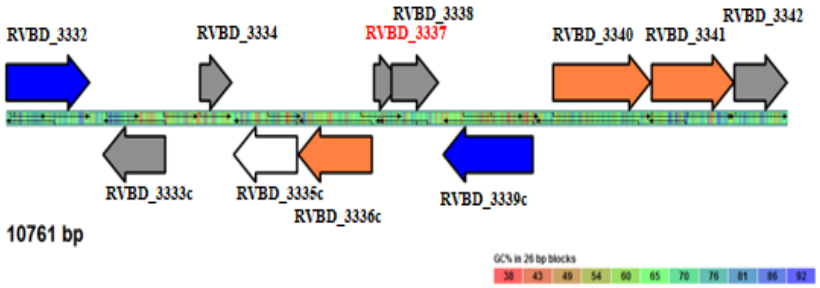


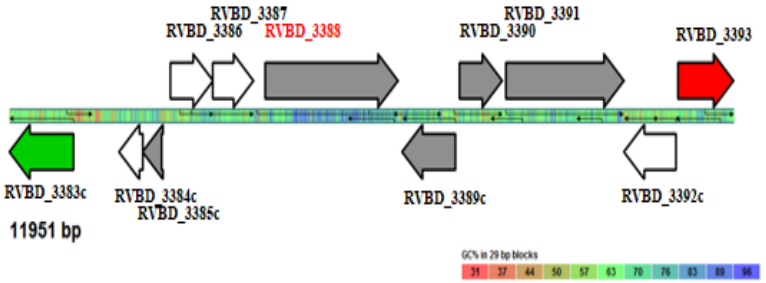


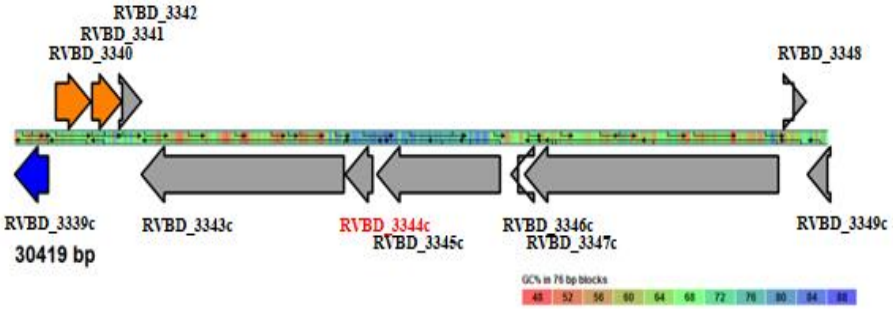


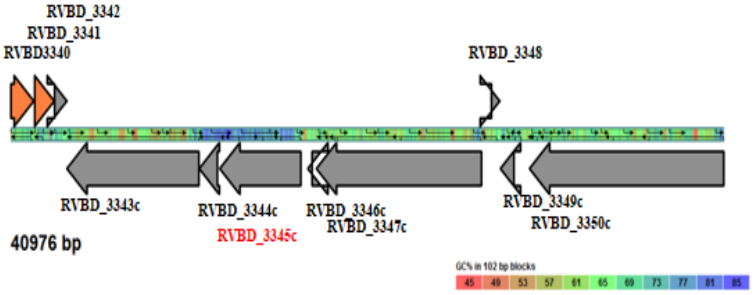


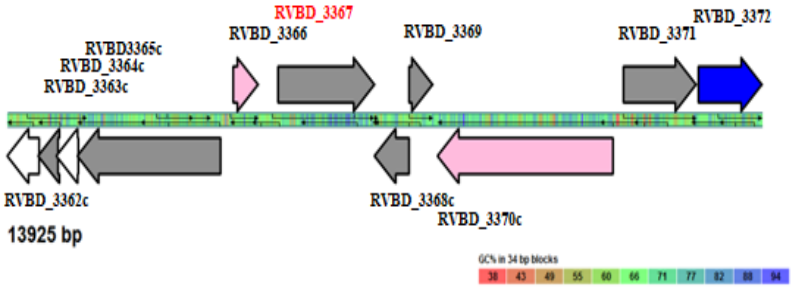


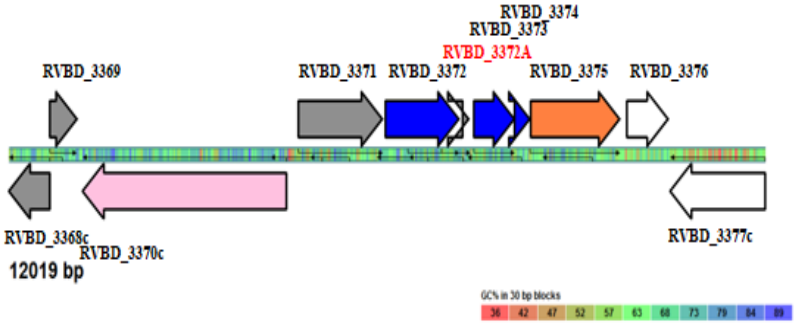


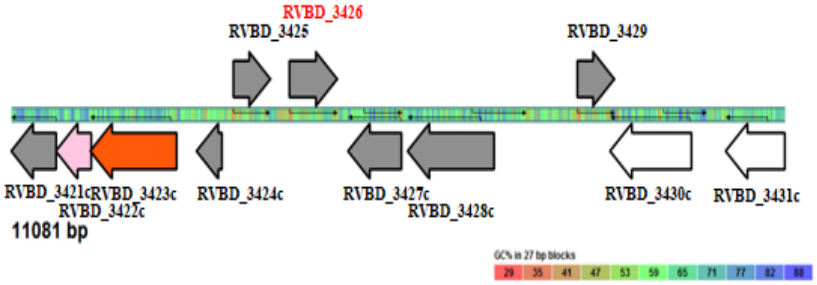


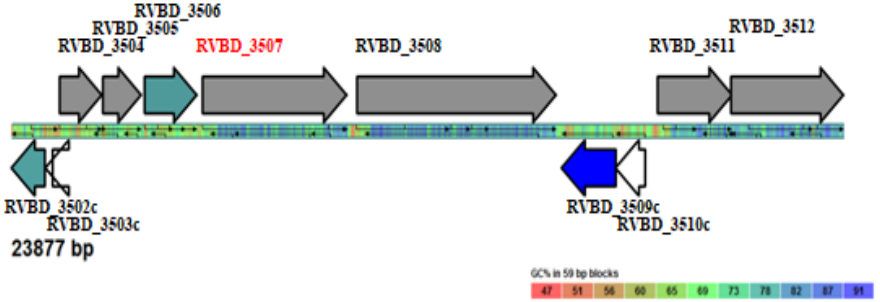


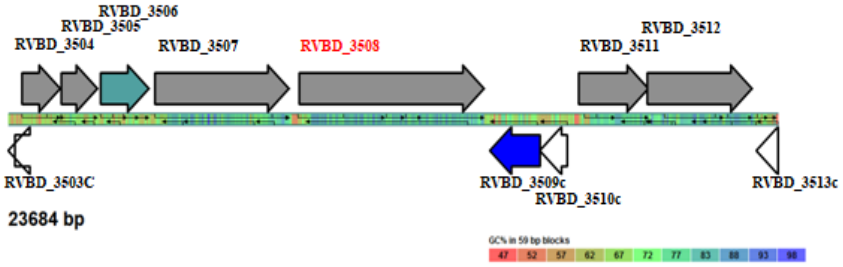


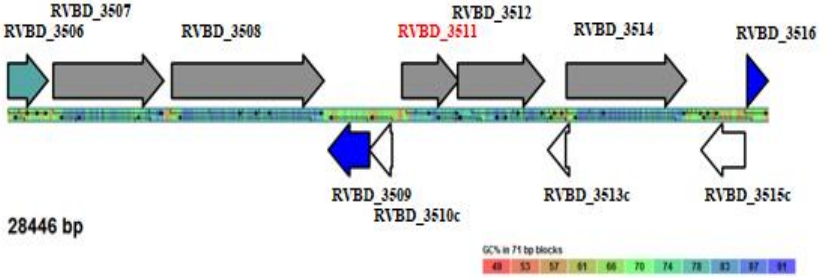


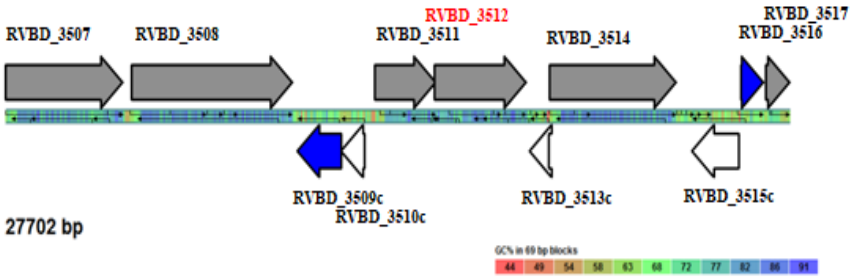


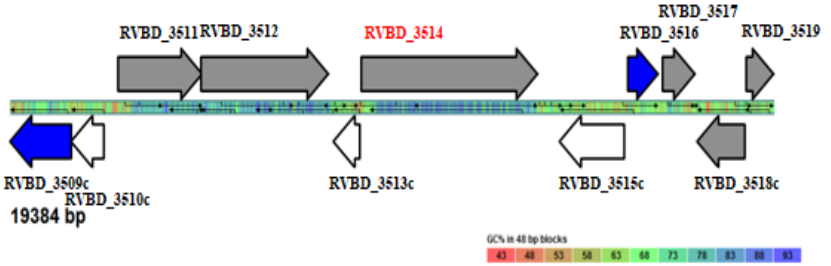


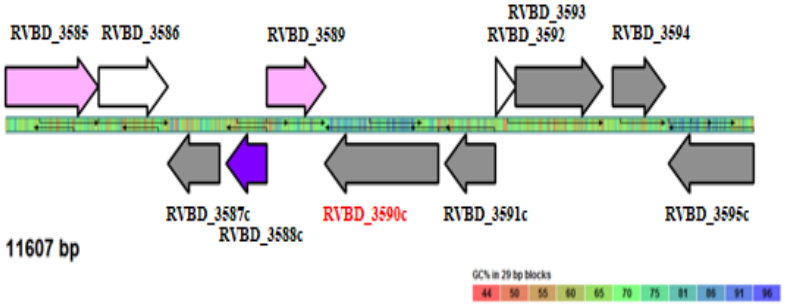


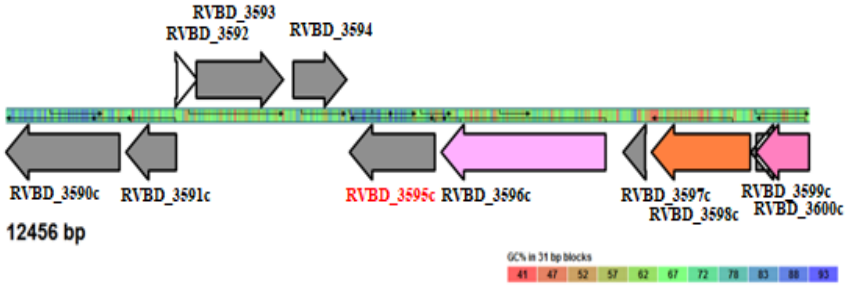


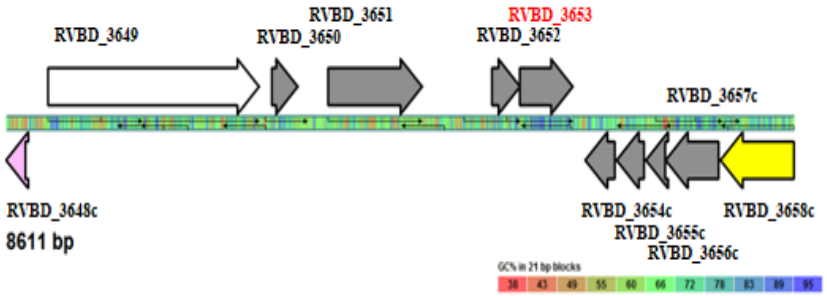


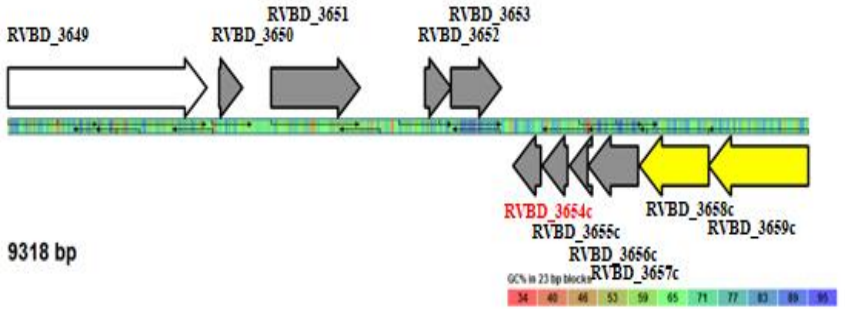


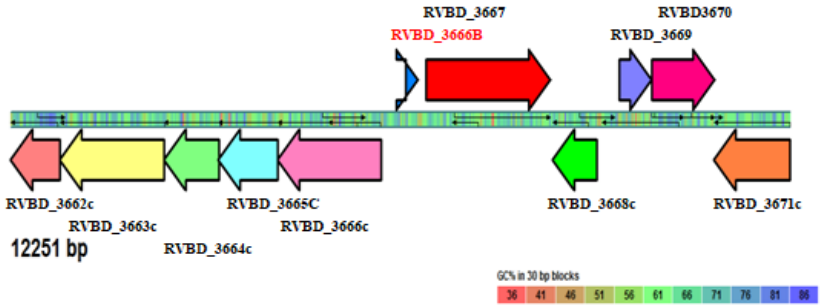


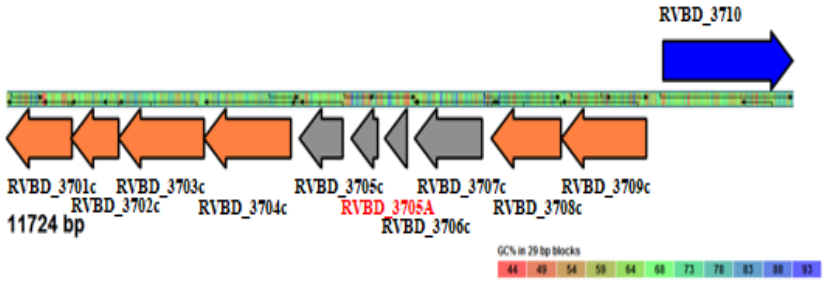


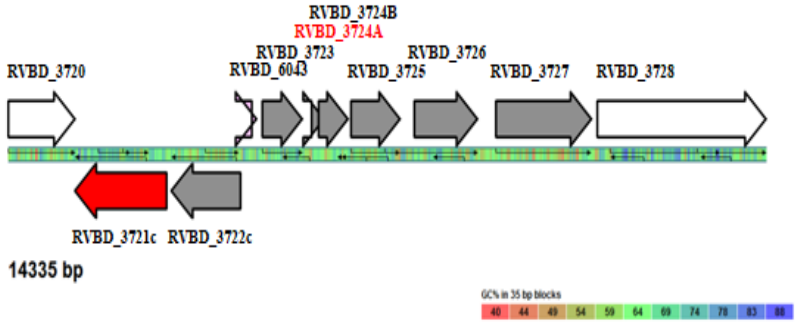


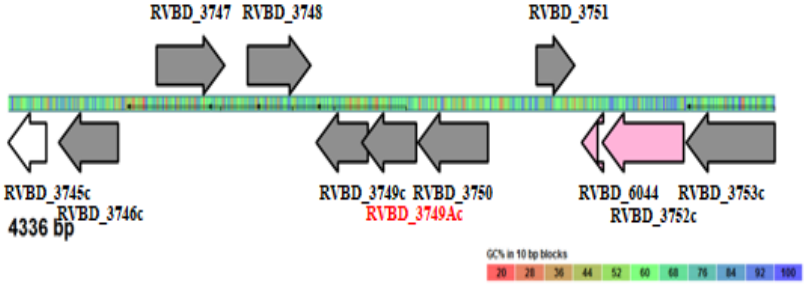


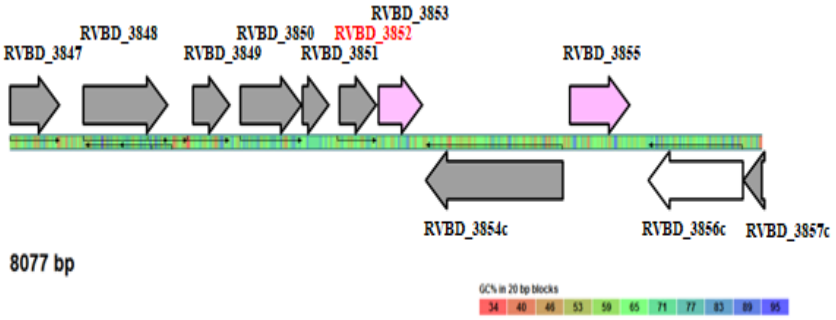












The unique proteins mapped in the above figures, belonged to pathway categories as summarised in Table 15.

Table 15 Summary of pathway unique proteins fall in as revealed by genome mapping

Function	Proteins in each category (RVBD_)
Carbohydrate metabolism	1755c, 2666B
Amino acid metabolism	0744A, 2292c
Xenobiotics biodegradation and metabolism	0064A, 0109, 0124, 0229Ac, 0277Ac, 0278c, 0279c, 0297, 0378, 0297A, 0500B, 0532, 0578c, 0616A, 0691A, 0742, 0746, 0747, 0833, 0834c, 0872c, 0959, 1004c, 1067c, 1068c, 1087,1091, 1158c, 1233c, 1243c, 1325c, 1396c, 1435c, 1441c, 1450c, 1452c, 1468c, 1549, 1651c, 1735Ac, 1759c, 1768, 1792, 1792A, 1818c, 1840c, 2077Bc, 2099c, 2160A, 2162c, 2231Ac, 2231Ac, 2309Xc, 2396, 2427Ac, 2490c, 2512A, 2615c, 2634c, 2645, 2646, 2647, 2652c, 2653c, 2654c, 2658c, 2741, 2853, 2862A, 2873A, 2964B, 3018Bc, 3103c, 3135, 3190A, 3233c, 3337, 3388, 3344c, 3345c, 3367, 3372A, 3426, 3507, 3508, 3511, 3512, 3514, 3590c, 3595c, 3653, 3654c, 2705A, 3724A, 3749Ac, 3852
Enzyme families	0781, 2346c
Genetic information processing	2986c
Environmental information processing	0590A
Cellular processes, Organismal system, Human diseases	2063A
Unclassified	2261c, 2651c, 2975c

The majority of the proteins are xenobiotics degradation proteins.

Unclassified proteins included:

- RVBD_2261c
- RVBD_2651c
- RVBD_2975c

4.3. References

Genetic Mapping Fact Sheets (2015) National Human Genome Research Institute.

https://www.genome.gov/10000715/genetic-mapping-fact-sheet/

Holsclaw, CM, Sogi, KM, Gilmore, SA, Schelle, MW, Leavell, MD, Bertozzi, CR, and Leary, JA (2008) Structural characterization of a novel sulphated menaquinone produced by stf3 from *Mycobacterium tuberculosis*. **ACS Chemical Bioloy**. 3(10): 619 - 624

Syed, K, Shale, K, Nazir, KNH, Krasevec, N, Mashele, SS and Pagadala, NS (2014) Genome-wide identification, annotation and characterization of novel thermostable cytochrome P450 monooxygenases from the thermophilic biomass-degrading fungi *Thielavia terrestris* and *Myceliophthora thermophila*. **Genes & Genomics**. 36(3): 321 - 333

Qhanya, LB, Matowane, G, Chen, W, Sun, Y. Letsimo, EM, Parvez, M, Yu, JH, Mashele, SS and Syed, K (2015) Genome-wide annotation and comparative analysis of cytochrome P450 monooxygenases in Basidiomycete Biotrophic plant pathogens. **PloS one**. 10(11): p.e 0142100.

Van Wyk R, Van Wyk M, Mashele SM, Nelson DR, Syed K. Comprehensive comparative analysis of cholesterol catabolic genes/proteins in mycobacterial species. International Journal of Molecular Science; Vol 20, Issue 5, March 2019.

CHAPTER 5

CONCLUSION

This study is the first of its kind, utilizing comparative genomics to identify unique genes/protein in *M. tuberculosis via* comparative analysis of mycobacterial species belonging to six different ecological niches.

In total 200 unique genes were identified for *M. tuberculosis* H37Rv with genome comparison. The genome annotation available for 120 of these 200 genes on TB database, allowing us to further analyse the uniqueness of these genes *via* gene comparison and to attempt functional classification of these unique genes *via* genome mapping. Further research and work are needed to identify the sequences of all the mycobacterial genes and make it available on a database.

In conclusion, comparative analysis of genes among mycobacterial species can be utilized to identify common and unique proteins in mycobacterial species belonging to each mycobacterial category. Identifying the location of a gene on the chromosome *via* genome mapping can play a role in functional classification of unique genes.

More in depth research is needed to test the uniqueness of the identified genes and to investigate whether these genes could serve as possible novel drug targets.