The challenge of drug resistant tuberculosis

By

Giovanni Satta

A thesis submitted to University College London in fulfilment of the requirement for the degree of Doctor of Philosophy

Centre for Clinical Microbiology, Department of Infection
University College London
Royal Free Campus
Rowland Hill Street
London NW3 2PF

Declaration

I, Giovanni Satta, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Abstract

Tuberculosis remains a major cause of morbidity and mortality globally with 10.4 million new cases worldwide in 2015 (WHO data). Drug resistance has also emerged as a public health concern in high-income countries as a result of immigration from endemic areas, the HIV pandemic and immunosuppressive treatments. The challenge this expansion poses to public health authorities is illustrated by an outbreak of isoniazid resistant tuberculosis in London (UK) in which over 400 cases have been diagnosed since 1995.

The main scope of this thesis is to address the challenges of drug resistant tuberculosis, with a particular focus on the applications of whole genome sequencing to the drug discovery process. A multi-strategy approach has been adopted, starting with the use of next generation sequencing to understand pathogenicity and discover new potential drug targets, to the laboratory testing of new compounds.

Using a sub-cluster of strains from the London outbreak, our findings have confirmed the utility of whole genome sequencing in understanding the pathogenesis of clinical strains, discovering specific genetic mutations that may have allowed the outbreak to persist for years. At the same time, we have also integrated sequencing technologies within the screening process of new compounds and we have demonstrated its potential role in elucidating the mechanism of action of old and new drugs, in particular para-aminosalicylic acid and its newly created analogues. The repurposing of current drugs (carbapenems) has been evaluated as a potential treatment in case of drug resistance whilst gold nanoparticles have been assessed as alternative compounds. Hence, this thesis has reflected the complexity of addressing the challenge of drug resistant tuberculosis and how a multi strategy approach is essential for the eradication of this global pathogen.

Impact statement

This thesis has demonstrated the full potential of whole genome sequencing, in particular when evaluating new antituberculous compounds. Its introduction in the drug development process should become routine use and resistant laboratory mutants should be created to elucidate the mechanism of action of any new antibiotic.

From our findings, whole genome sequencing has also played an important role in unveiling additional insights into the pathogenicity of *M. tuberculosis*. We would support wide-scale investment in whole genome sequencing and the creation of comparable databases of genetic variations of clinical strains, plus the addition of biological information, including fitness assays and mutation rate information for a comprehensive picture. This will allow the interrogation of *M. tuberculosis* gene function and the discovery of new drug targets. In this context, the clinical isolates from the London outbreak represented a valuable sample set and their genetic analysis have highlighted different virulence genes that need further consideration as potential drug targets.

Whilst the repurposing of ertapenem and faropenem did not produce encouraging results, a new compound, AD25a, was successfully tested against *M. tuberculosis*. It has shown very promising results in our laboratory experiments and it should be further tested in animal models to assess its potential use as new drug against tuberculosis.

Finally, gold nanoparticles have shown an interesting interaction with the mycobacterial cell wall despite not having a direct antibacterial effect and further research should also focus on the development of similarly vectorized drug delivery systems.

A conference poster, peer-reviewed papers published at the time of printing this thesis and confirmation of a conference talk are attached at the end of the printed version under appendices.

Acknowledgments

Warm thanks go to my supervisors, Prof Timothy McHugh and Dr Marc Lipman for their continued support, advice and encouragement throughout my PhD.

I am especially grateful to the Center for Clinical Microbiology at UCL, Royal Free campus, for funding reagents and fees during my studies, and to all the scientists and researchers working there (including the NHS diagnostic laboratory at the Royal Free Hospital) for making my studies at UCL very enjoyable. Additional funding was received from the British Infection Association and I am grateful to the board of trustees for their award in the early stages of my research.

Particular thanks go to Dr Adam Witney, (Institute for Infection and Immunity, St. George's University of London) for his invaluable support with the bioinformatics analysis, Dr Innes Clatworthy (UCL Electron Microscopy Unit, Royal Free campus) for the electron microscopy pictures and X-ray microanalysis and Dr Peter Hart (Wellcome Trust) for proofreading this thesis.

I am particularly grateful to Dr Catherine Arnold and the Genomic Services and Development Unit (GSDU, Public Health England) for performing all our next generation sequencing and I am additionally thankful to all our collaborators for providing the different compounds, in particular Dr Andrew Boa (University of Hull) and Dr Africa Barrientos (Midatech Pharma).

I would also like to thank a number of colleagues for their support and input to this work. Dr Amani Alnimr and Dr Isobella Honeyborne trained me on how to work in the CL3 laboratory, research assistants Neelu Begum and Magdalena Karlikowska were extremely helpful in the extraction of some genomic DNA, Dr Julio Ortiz Canseco taught me how to use the MIGIT/BACTEC machine and helped with some of the antibiotic testing, Professor Francis Drobniewski and Dr Ximena Gonzalo (Imperial College London) kindly provided some isolates tested.

Finally, I am indebted to my family and friends for their support, encouragement and understanding during these last seven years.

Contents

This chapter has been edited due to copyright reasons. Its use is not permitted under the quotation exception, and permission to make it available has not been obtained. For these

List of tables

List of figures

Abbreviations

ADC	Albumin Dextrose Catalase
AK	Amikacin
AST	Antimicrobial Susceptibility Testing
BCG	Bacillus Calmette-Guerin
CAS	Central Asian Strain
CBN	Conformal Bayesian Network
CDC	Centers for Disease Control and Prevention
CFU	Colony Forming Unit
CL3	Containment Level 3
COSHH	Control of Substances Hazardous to Health
CTAB	Cetyl Trimethylammonium Bromide
DNA	Deoxyribonucleic Acid
DOT	Directly Observed Therapy
EAI	East African-Indian
EBA	Early Bactericidal Activity
EMB	Ethambutol
EUCAST	European Committee on Antimicrobial Susceptibility Testing
FAS	Fatty Acid Synthases
FDA	Food and Drug Administration
Gb	Giga-base
GC	Guanine-Cytosine
GSDU	Genomic Services and Development Unit
HIV	Human Immunodeficiency Virus

INH	Isoniazid
IS	Insertion Sequence
LAM	Latin America-Mediterranean
Levo	Levofloxacin
Mb	Mega-byte
MGIT	Mycobacteria Growth Indicator Tube
MIC	Minimal Inhibitory Concentration
MIRU	Mycobacterial Interspersed Repetitive Units
Moxi	Moxifloxacin
NaCl	Sodium Chloride
NAD	Nicotinamide Adenine Dinucleotide
NCTC	National Collection of Type Cultures
NGS	Next Generation Sequencing
NHS	National Health Service
NP	Nanoparticle
OADC	Oleic Albumin Dextrose Catalase
Oflox	Ofloxacin
PAML	Phylogenetic Analysis by Maximum Likelihood
PAS	Para-Amino-salicylic acid
PCR	Polymerase Chain Reaction
PHE	Public Health England
PZA	Pyrazinamide
QC	Quality Control
RIF	Rifampicin
RNA	Ribonucleic Acid

rRNA	Ribosomal Ribonucleic Acid
SDS	Sodium Dodecyl Sulphate
SM	Streptomycin
SNP	Single Nucleotide Polymorphism
TB	Tuberculosis
UK	United Kingdom
US	United State (of America)
VCF	Variant Call Format
WGS	Whole Genome Sequencing
WHO	World Health Organization
ZN	Ziehl-Neelsen

Chapter 1: Introduction

1.1 The disease

Tuberculosis (TB) is an infectious disease caused by bacteria of the *Mycobacterium tuberculosis* complex (of which *M. tuberculosis* is the most common human pathogen). TB generally affects the lungs (pulmonary TB), but every organ can be infected, in particular the central nervous system, lymphatic system, kidneys, spleen and bones. The disease is airborne transmitted by droplets when people who have active TB in their lungs cough, spit, speak, or sneeze. Symptoms of active disease include cough, fever, night sweats and weight loss (hence the popular historical name of *consumption*) and they may be mild for many months, causing further risk of transmission and significant delay in treatment. It is estimated that people with active TB can infect 10–15 other people through close contact over the course of a year (Sepkowitz, 1996).

Without proper treatment, 45% of HIV-negative people with TB and nearly all HIV-positive people with TB will die (WHO, 2017). The diagnosis is made through radiological imaging and traditional microscopy and culture of clinical samples, although new molecular diagnostic methodologies such as Xpert MTB/RIF® have expanded substantially since 2010, when the World Health Organisation (WHO) first recommended its use (WHO, 2010b). In 2017, England became the first country in the world to introduce and pioneer the use of whole genome sequencing (WGS) on a national scale for the diagnosis, detection of drug resistance and typing of *M. tuberculosis* (PHE, 2017). TB is a treatable disease but it requires a prolonged course of at least 3 antimicrobial drugs, for a minimum of 6 months. This is in contrast with other bacterial respiratory infections, where a few days of a single drug are generally sufficient (Grossman, Rotschafer, & Tan, 2005).

Robert Koch first identified and described the causative agent of TB in 1882 (Sakula, 1982), after centuries of speculation as to its possible infectious nature. TB has been a widespread disease since ancient times and, with no environmental reservoir, it is believed to have co-evolved with humans over millennia: this is demonstrated by skeletal remains of prehistoric humans (typically from 7000 BC) (Hershkovitz et al., 2008), and by the presence of *M. tuberculosis* DNA in the lungs of Egyptian mummies (600 BC) (Donoghue et al., 2010). Around 400 BC, Hippocrates defined *phthisis* as

the most widespread disease of the times, and it was almost always fatal (Daniel & Iversen, 2015). TB has been also linked with numerous legends: the mythical vampires have been considered to be people affected (Sledzik & Bellantoni, 1994). The spread of TB caused widespread public concern in the 19th century and, at the time, one in four deaths in England and one in six deaths in France were caused by TB. Infected people were encouraged to enter sanatoria (McCarthy, 2001), where fresh air and rest were believed to have a curative effect.

The first genuine success against TB occurred in 1906 when Albert Calmette and Camille Guerin developed an efficacious vaccine from an attenuated bovine-strain (BCG – Bacillus Calmette-Guerin) (Bohan, 2017). The BCG vaccine has been the subject of numerous efficacy trials and epidemiological studies conducted over several decades (Roy et al., 2014). Results confirm the protective effect against miliary TB and tuberculous meningitis (Trunz, Fine, & Dye, 2006) but its efficacy against pulmonary diseases varies geographically (Colditz, Brewer, Berkey, & et al., 1994) (Mangtani et al., 2017). Since 1946, the discovery of streptomycin and many other anti-TB drugs led to the belief that TB could be finally and completely eradicated like smallpox. However, at the beginning of the 1980s, the emergence of multi-drug resistant strains and the spread of HIV epidemic resulted in such a marked increase in the number of new cases that in 1993 WHO declared TB a global health emergency (WHO, 1993).

WHO passed a resolution in May 2014 approving, with full support, its "*End TB Strategy"* (WHO, 2015), with the aim to end the global TB epidemic by 2030, to reduce TB deaths by 95% and to cut new cases by 90%. After more than a century since its original identification, TB still remains a major global health challenge all over the world.

1.2 The organism

M. tuberculosis is a small (2-4 μm in length and 0.2-0.5 μm in width), aerobic, non-motile, non-spore forming bacillus. Its cell wall has a high content of high molecular weight lipids, making the cell surface hydrophobic and resistant to many disinfectants and common laboratory stains (Haas, 2000). If a Gram stain is performed, *M. tuberculosis* does not retain any dye or will stain as weakly Gram-positive due to the rich lipid cell wall (Madison, 2001). *M. tuberculosis* is a human-restricted pathogen, having no environmental or zoonotic reservoirs.

1.2.1 Mycobacterial cell wall

The distinctive feature of *M. tuberculosis* is the structure of its cell wall, composed of a thick peptidoglycan layer, with no outer membrane but a complex structure of carbohydrates and lipids (60% of the cell wall weight), made from arabinogalactan (D-arabinose and D-galactose), mycolic acids (long chain fatty acids) and free lipids (Fig. 1). Anchored in the plasma membrane are porins, transport proteins and lipoarabinomannan (LAM) which is functionally related to the O-antigenic lipopolysaccharides present in other bacteria. This particular structure confers upon mycobacteria resistance to acidic/alkaline substances, detergents, common antibacterials, osmotic lysis and radicals. The same structure is responsible for the unique characteristic of acid-fastness, that is the basis of Ziehl-Neelsen stain (ZN) (Murray, Rosenthal, & Pfaller, 2005).

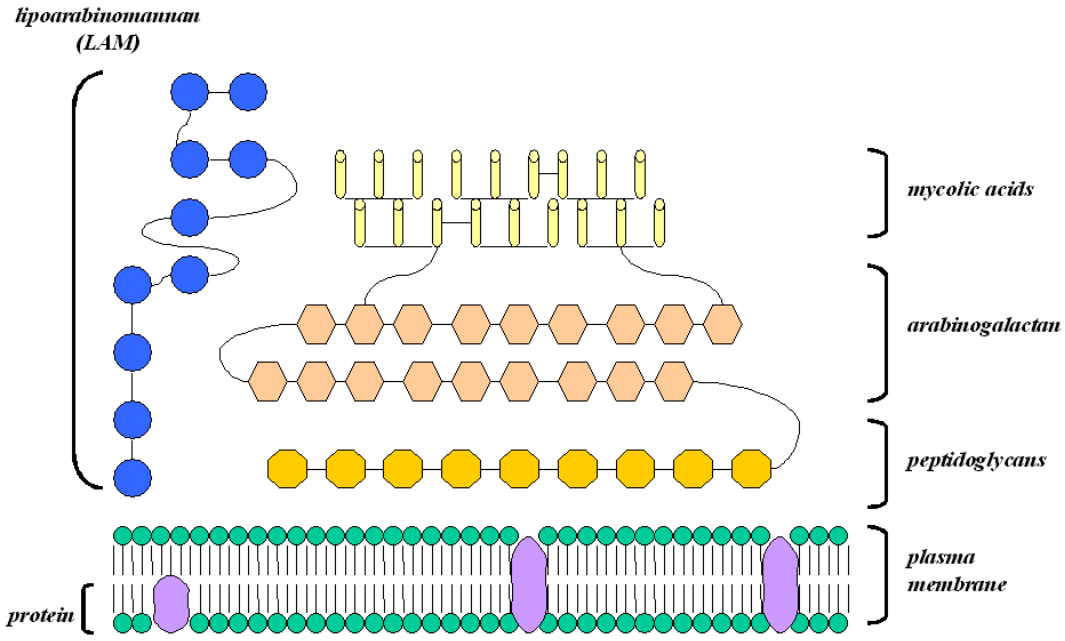


FIGURE 1: MYCOBACTERIAL CELL WALL.

Graphical representation of the mycobacterial cell wall, with a complex structure of peptidoglycans, arabinogalactan and mycolic acids (see text for further description) (original figure).

The term *acid-fast bacilli* is clinically synonymous with mycobacteria, although some other organisms (such as *Nocardia* and some *Corynebacteria*) can be acid-fast. In the ZN stain, a fixed smear is covered with carbol-fuchsin, heated, decolorized with alcohol and counterstained with methylene blue. Mycobacteria cannot be decolorized

with this acid solution and will retain the pink color of the fuchsin, contrasting with the blue background. The Kinyoun modification does not require any heating (Haas, 2000). It is calculated that at least 10,000 organisms/ml are necessary for the sputum to become smear positive. Many laboratories now use a fluorochrome stain with auramine and potassium permanganate as the counterstain and this is up to 10-fold more sensitive than carbol-fuchsin methods (Steingart et al., 2006).

Mycolic acids are a series of high molecular weight, long-chain hydroxy fatty acids, unique to mycobacteria and they play a critical role in the core structure of the cell wall (Brennan & Crick, 2007). They are hydrophobic and make the organism more resistant to chemical damage and dehydration, and limit the effectiveness of hydrophilic antibiotics and biocides (Lambert, 2002). Mycolic acids also allow *M. tuberculosis* to grow inside macrophages. They have been demonstrated to have inflammation-controlling properties (Korf et al., 2006). The mycolic acids found in *M. tuberculosis* can be divided into three main types: alpha-mycolic acids (forming 70% of the mycolic acids of the organism and containing several cyclopropane rings), methoxy- and keto-mycolic acids (forming the remaining 30%) (Takayama, Wang, & Besra, 2005).

Five distinct stages for the synthesis of mycolic acids have been described, with the involvement of various essential enzymes including type I and type II fatty acid synthases (*FAS-I* and *FAS-II*), *M. tuberculosis* beta-ketoacyl acyl carrier protein synthase III for initiation and condensation (*mtFabH*), ketoacyl synthases for elongation (*KasA* and *KasB*) and polyketide synthase for termination (*Pks13*), leading to full-length mycolates (Bhatt, Molle, Besra, Jacobs, & Kremer, 2007). Many of the genes and enzymes responsible for the synthesis of mycolic acid and the cell wall are unique to *M. tuberculosis* and have no known homologues in other bacteria, making them an attractive drug target. In fact, the integrity of the cell wall is essential for *M. tuberculosis* survival and its disruption will cause effective killing. Isoniazid (INH), a fundamental first-line antituberculous drug, inhibits mycolic acid synthesis and interferes with cell wall synthesis, thereby producing a bactericidal effect. INH also disrupts DNA, lipid, carbohydrate, and nicotinamide adenine dinucleotide (NAD) synthesis and metabolism (Timmins & Deretic, 2006).

1.2.2 Culture media and growth rate

M. tuberculosis grows slowly with an average generation time of 15 to 20 hours. This is an extremely slow rate compared with other bacteria, which usually divide in less

than an hour, and visible growth can take up to 6 weeks on solid media (compared to 12-24 hours for other common human pathogens) (Haas, 2000). The doubling time can go up to 33 hours inside human macrophages (Chanwong, Maneekarn, Makonkawkeyoon, & Makonkawkeyoon, 2007). This is probably due to the fact that *M. tuberculosis* has only one mechanism to translate mRNA into proteins because of the lack of the *rrn-B* operon. All slowly growing mycobacteria are thought to have either one or two rRNA operons per genome (Menendez et al., 2002).

Culture methods for Mycobacteria use either solid or liquid media. Solid media can be divided into two main types: agar-based (e.g. Middlebrook 7H10) and egg-based (e.g. Lowenstein-Jensen). Liquid media include various agar-based media (e.g. Middlebrook 7H10) and the automated Mycobacteria Growth Indicator Tube (MGIT) system, containing 7H9 broth base and OADC enrichment. The BACTEC MGIT 960 system is produced by Becton Dickinson (BD) (New Jersey, USA). The instrument scans the MGIT every 60 minutes for increased fluorescence. Analysis of the fluorescence is used to determine if viable organisms are growing and a positive tube contains approximately 10^5 to 10^6 colony-forming units per milliliter (CFU/mL). Culture tubes which remain negative for a minimum of 42 days (up to 56 days) and which show no visible signs of positivity are removed from the instrument as negatives and discarded (Tortoli et al., 1999). The system can also be used to perform susceptibility testing (Tortoli, Benedetti, Fontanelli, & Simonetti, 2002).

1.2.3 Cording and clumping

The aggregation of mycobacterial cells in a definite order, forming microscopic structures that resemble cords, is known as *cording* and is considered a virulence factor. In the 1950s, cording was shown to be due to trehalose dimycolate, a glycolipid molecule found in the *M. tuberculosis* cell wall, and this was consequently named *cord factor* (Retzinger, Meredith, Hunter, Takayama, & Kezdy, 1982) (Julian et al., 2010). This has been observed to influence immune responses, induce the formation of granulomas and inhibit tumor growth (Spargo, Crowe, Ioneda, Beaman, & Crowe, 1991). It is important not to confuse *cording* with *clumping*, which is a general property of mycobacteria due to the high lipid content in the wall and their hydrophobic surfaces. Clumping has been associated with more vigorous growth inside macrophages and it is a peculiar laboratory finding that can interfere with experiments using liquid culture (Hoal-van Helden, Hon, Lewis, Beyers, & van Helden, 2001).

Different measures can be applied to reduce clumping, including the use of detergents, sonication, glass beads and continuous agitation. Tween 80 is the detergent commonly used to minimize clumping of mycobacteria in the culture medium during exposure to test compounds. However, there is the potential concern that using this detergent may potentiate the activity of test compounds by artificially increasing the permeability and exposure of the mycobacterial cell. There is so far little evidence that Tween 80 can significantly affect susceptibility testing but it is important to test new compounds with two alternative methods. There is also a need for clarity regarding terminology used for Middlebrook-based media – It is very common for investigators to report the use of the medium Middlebrook 7H9 without indicating which supplements have been added (Franzblau et al., 2012).

1.2.4 Dormancy

In the immunocompetent person, the initial inflammatory process caused by *M. tuberculosis* is localized to a small area with the possible involvement of regional lymph nodes in what it is called a *primary complex*. Nevertheless, the primary complex is not always completely sterilized by the immune response, leading to a chronic infection which can develop due to the survival of some bacteria. Further immunosuppression (for example, HIV infection) can lead to reactivation of the primary complex and dissemination of the disease at pulmonary and extra-pulmonary level (Haas, 2000). Different factors can influence the course of the disease and are related to the immune response of the infected person and the virulence of the strain involved. One of the major mechanisms of virulence is the ability of mycobacteria to survive inside macrophages: this is not only connected with the particular structure of its cell wall, but also with a series of other mechanisms. Principally, the enormous success of *M. tuberculosis* in causing infection is based on three factors: first, its ability to control macrophages after primary infection/phagocytosis to prevent its own destruction; second, the formation of granulomas, creating a confined and acidic environment which causes immune system dysfunction; third, the capability to switch off its own metabolism and block replication, thus entering a stage of dormancy where it becomes extremely resistant to host defence and drug treatment (Gengenbacher & Kaufmann, 2012). This highlights the importance of studying the molecular biology of *M. tuberculosis* to better understand the pathogenic mechanisms and virulence factors and to contribute to the development of new vaccines and drugs. Indeed, the research in this field is extremely productive and different pathogenic mechanisms

have been described to date, though many remain unknown (Delogu, Sali, & Fadda, 2013).

The population theory – Initial studies (J. M. Dickinson & Mitchison, 1970) postulated the existence of four different bacterial subpopulations in the TB lesion. A first population is represented by actively replicating bacilli that, in the absence of drug resistance, are rapidly killed by INH and rifampicin (RIF). A second population is composed of slow growers still susceptible to RIF, whilst the third group is represented by intracellular bacteria in an acidic environment and killed by pyrazinamide (PZA), but tolerant to RIF. Lastly, the bacilli in the fourth population are quiescent cells, able to tolerate high doses of drugs and cause relapsing disease if not adequately killed. Another hypothesis has postulated the presence of three compartments (Gillespie, 2002): bronchioles and alveoli, cavitary lesion and a third population of bacilli, in acidic and intracellular medium and only susceptible to PZA. Understanding the natural biology of *M. tuberculosis* is essential for the development of active compounds and the reduction of its length of treatment.

Dormancy models – Various laboratory dormancy models for *M. tuberculosis* have been developed by simulating the latent TB infection, in which the bacteria persist in a non-replicative state and with reduced susceptibility to antimycobacterial agents (Cole et al., 1998). It is important to note that many authors tend to use the terms *dormant*, *latent* or *persistent* synonymously. *Latency* is a clinical condition defined as the presence of *M. tuberculosis* within the human host without causing any clinical disease. Conversely, *dormancy* is the metabolic condition where bacilli exhibit slow growth, with a downshift of metabolic activities, altered staining features, an inability to be cultivated on solid media and resistance to antimycobacterial agents (Alnimr, 2015).

Dormancy models can be roughly divided into two main groups: animal and *in vitro* models. The Cornell model, developed at the Cornell University in 1966 (McCune, Feldmann, & McDermott, 1966), was the first animal model described using infected mice treated with antimycobacterial drugs to reduce the bacterial burden to undetectable levels and subsequently waiting for spontaneous reactivation of infection (Alnimr, 2015). All the other various dormancy models have different advantages and disadvantages and it is difficult to choose which model is most likely to mimic the real *M. tuberculosis* dormancy in the human host. The 100-day culture model (Hu, Coates, & Mitchison, 2006) is based on the mentioned population theory

and it is useful to test the sterilizing activity of new drugs on the fourth population with reduced drug susceptibility, but it is technically challenging and the clumping of bacilli often interferes with experiments. Other models, such as the hypoxia-based (Wayne & Hayes, 1996) and the starvation model (Loebel, Shorr, & Richardson, 1933) (Betts, Lukey, Robb, McAdam, & Duncan, 2002), try to mimic the hostile environment of the granuloma/macrophage phagosome, hence forcing the bacilli to shut down their metabolic pathways. However, only in the starvation model do TB bacilli remain resistant to metronidazole, unlike the case for the Wayne model, and this is a condition believed to be reflective of an *in vivo* hypoxic granuloma (Klinkenberg, Sutherland, Bishai, & Karakousis, 2008).

Comparing various dormancy models is beyond the main scope of this thesis, but their importance has to be acknowledged. Assessment of novel compounds should include screening of their activity against dormant bacilli to identify actively sterilizing agents. However, it is also important to remember that the laboratory results from non-clinical models may not automatically translate to clinically relevant results. During laboratory studies, moxifloxacin had excellent sterilizing activity even against RIF-tolerant strains (Hu, Coates, & Mitchison, 2003). Unfortunately, this did not translate into shorter treatment, with the moxifloxacin regimens in the REMox trial not meeting the study's non-inferiority criteria (Gillespie, Crook, McHugh, Mendel, Meredith, & Murray, 2014). Despite the disappointing results, the trial highlighted the importance of developing a more accurate dormancy model with some authors reinforcing the predictive value of animal models for TB drug regimen development (S. Y. Li et al., 2015).

1.3 Genome of *M. tuberculosis*

The complete genome sequence of *M. tuberculosis* was deciphered in 1998 (Cole et al., 1998) and reannotated in 2002 (Camus, Pryor, Medigue, & Cole, 2002). It comprises more than 4 million base pairs (total of 4,411,529), with 4000 genes and a high content of guanine-cytosine (GC) (up to 65.6%). The genome possesses all the genes necessary for the synthesis of essential amino acids, vitamins and enzyme co-factors. A high proportion of genes encoding enzymes are involved in lipogenesis and lipolysis. Additionally, *M. tuberculosis* possesses the genes to synthesize glycolytic enzymes and enzymes necessary for the anabolic pentose phosphate pathway (production of NADPH and pentose sugars), the catabolic Krebs cycle and the glyoxylate cycle (for the synthesis of carbohydrates from lipids) (Cole et al., 1998).

The genome also encodes enzymes used in aerobic, microaerophilic and anoxic electron transfer and enables *M. tuberculosis* to survive in a number of different environments including the oxygen-rich lung, the macrophage and within an anoxic granuloma (Britton, Roche, & Winter, 1994). However, only 40% of genes have a known function and up to 50% of them are still labeled as unknown, uncharacterized or have a hypothetical function, based on their relatedness to corresponding genes in other bacteria (Mazandu & Mulder, 2012).

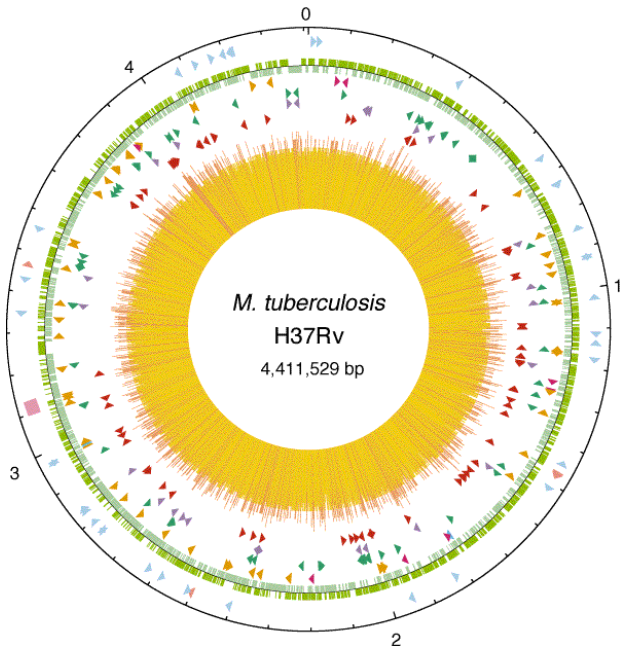


FIGURE 2: REPRESENTATION OF THE *M. TUBERCULOSIS* H37RV GENOME.

A diagrammatic representation of the *M. tuberculosis* H37Rv genome. The outer black circle shows the scale in megabase (mb), with 0 representing the origin of replication (total of 4,411,529 base pairs). The first ring from the exterior denotes the positions of stable RNA genes (tRNAs are blue, others are pink) and the direct repeat region (pink cube); the second ring inwards shows the coding sequence by strand (clockwise, dark green; anticlockwise, light green); the third ring depicts repetitive DNA (insertion sequences, orange; 13E12 REP family, dark pink; prophage, blue); the fourth ring shows the positions of the PPE family members (green); the fifth ring shows the PE family members (purple, excluding PGRS); and the sixth ring shows the positions of the PGRS sequences (dark red). The histogram (centre) represents G + C content, with <65% G + C in yellow, and >65% G + C in red. Open access article (Cole et al., 1998).

The introduction of next generation sequencing has revolutionized the field of TB research with new molecular biology techniques and genomic studies, such as transcriptomics, proteomics, functional and structural genomics. Better understanding of the complex biology of this pathogen will allow us to develop new drug targets to further populate the TB drug pipeline.

1.3.1 Classification

Mycobacterium spp belong to the family *Mycobacteriaceae* of the class *Actinobacteria*. The taxonomy of mycobacteria is regularly updated and the most recent classification was released in 2017 with over 170 recognized species (Forbes, 2017). Based on phenotypic and genetic differences, the genus can be classified into two main groups: slowly growing mycobacteria, including *M. tuberculosis*, and rapid growers, generally environmental organisms but able to cause disease in the immunocompromised patient (Wallace et al., 1983) (Table 1). *Mycobacterium fortuitum* and *Mycobacterium smegmatis* have been used as alternative laboratory models because of their rapid growth (2-4 days for visible colonies) and safety profile compared to *M. tuberculosis* (Gillespie, 2001) (Andries et al., 2005). However, it is debatable if they truly represent a reliable model as, from an evolutionary point of view, the different growth rates may reflect substantial variation between the species with up to 30% of *M. tuberculosis* proteins not present in *M. smegmatis* (Reyrat & Kahn, 2001) (Altaf, Miller, Bellows, & O'Toole, 2010). Alternatively, *Mycobacterium bovis* BCG has become an important tool for mycobacterial research, with only 3% of *M. tuberculosis* proteins missing and its safer profile (Altaf et al., 2010). It is an attenuated variant of *M. bovis* and it was developed by 230 passages over 13 years (Oettinger, Jorgensen, Ladefoged, Haslov, & Andersen, 1999). *M. bovis* is naturally resistant to pyrazinamide (PZA) but still susceptible to all other common antituberculous drugs (Marianelli et al., 2015). Finally, it is important to remember that *M. tuberculosis* belongs to a group of closely related organisms known as the *Mycobacterium tuberculosis* complex, including *M. bovis*, *Mycobacterium africanum, Mycobacterium orygis*, *Mycobacterium microti*, *Mycobacterium canetti*, *Mycobacterium caprae*, *Mycobacterium pinnipedii*, *Mycobacterium suricattae* and *Mycobacterium mungi*, with more than 99% match at nucleotide level and with an identical 16S rRNA sequence (McEvoy et al., 2007). In this thesis, we will always refer to *M. tuberculosis* unless stated otherwise.

Classification of mycobacteria	
Slow growers (growth in more than 7 days)	**Rapid growers (growth in less than 7 days)**
***Mycobacterium tuberculosis* complex**: - *Mycobacterium tuberculosis* - *Mycobacterium bovis* - *Mycobacterium africanum* Plus other less common species (see text for details)	- *Mycobacterium abscessus* - *Mycobacterium chelonae* - *Mycobacterium fortuitum* - *Mycobacterium smegmatis**
Other pathogenic mycobacteria: - *Mycobacterium avium/intracellulare* complex (MAC) - *Mycobacterium kansasii* - *Mycobacterium leprae* - *Mycobacterium malmoense* - *Mycobacterium marinum* - *Mycobacterium ulcerans*	
Additional rare pathogenic species were added in the most recent update in 2017 (Forbes, 2017).	**Mycobacterium smegmatis* is a non-pathogenic rapid grower commonly used for mycobacterial research

TABLE 1: CLASSIFICATION OF MYCOBACTERIA (SHINNICK & GOOD, 1994).

Classification of mycobacterial species into slow growers, rapid growers and other pathogenic mycobacteria (see text for further description).

1.3.2 Unique features

The *M. tuberculosis* genome exhibits some unique features that make *M. tuberculosis* radically different to other bacteria.

Lipid metabolism – A very large portion of its coding capacity is devoted to the production of enzymes involved in lipogenesis and lipolysis. The genome contains 250 genes (5.7%) involved in fatty acid metabolism and further 39 are involved in the polyketide metabolism generating the waxy coat, compared to only 50 genes in *Escherichia coli* (Cole, 1999). These conserved genes demonstrate the evolutionary importance of the waxy coat to the pathogen survival and experimental studies have validated the importance of a lipid metabolism for *M. tuberculosis*, with cholesterol uptake, catabolism and broader utilization contributing to virulence and pathogenesis (Wipperman, Sampson, & Thomas, 2014).

Repetitive regions – There are a number of repetitive DNA sequences including insertion sequences (IS), the direct repeat (DR) region, the major polymorphic tandem repeats (MPTR) and the polymorphic GC-rich repetitive sequence (PGRS) (Poulet & Cole, 1995a). The genome also contains two large protein families encoding acidic, asparagine- or glycine-rich proteins referred to PE (proline-glutamate) and PPE (proline- proline-glutamate), occupying about 10% of the coding capacity (Poulet & Cole, 1995b). Some of these genes are situated in the MPTR and the PGRS regions and are associated with antigenic variability and virulence (Karboul et al., 2008). These PE/PPE proteins have a conserved N-terminal motif, deletion of which impairs growth in macrophages and granulomas (Glickman & Jacobs, 2001).

Conserved genome – Comparison of the *M. tuberculosis* H37Rv genome sequence with that of a clinical isolate, CDC1551, revealed less heterogeneity when compared with other bacteria (Fleischmann et al., 2002). Various studies of clinical and laboratory strains have shown that the *M. tuberculosis* complex possesses a high degree of sequence homogeneity, demonstrated by a greater than 99.9% similarity in rDNA sequences (Frothingham, Hills, & Wilson, 1994) (Sreevatsan, Pan, & Stockbauer, 1997). Indeed, even when comparing *M. tuberculosis* genome with that of *Mycobacterium bovis*, there is less than 0.05% sequence divergence (Garnier et al., 2003). This is peculiar to the mycobacterial genus and profoundly differs from other species, with for example two strains of *Escherichia coli* showing a sequence diversity of 1.6% (Perna et al., 2001).

The lack of variation in *M. tuberculosis* may have resulted from limited selective pressure or from a very recent distribution in evolutionary terms. Indeed, some authors have hypothesised that the lack of genetic variation meant that *M. tuberculosis* is younger than 20,000 years old (Fleischmann et al., 2002) (Frothingham et al., 1994) (Sreevatsan, Pan, Stockbauer, et al., 1997). Additionally, with the exception of *Mycobacterium canetti* (Gutierrez et al., 2005), there is little evidence of horizontal gene transfer in *M. tuberculosis* and its evolution appears to have occurred in a clonal manner (Filliol et al., 2006) (Hirsh, Tsolaki, DeRiemer, Feldman, & Small, 2004).

1.3.3 Challenges when extracting mycobacterial DNA

Extracting adequate genomic DNA is an essential criterion prior to whole genome sequencing (WGS) as poor quality and quantity will impact on the genome coverage, ultimately influencing the final analysis. *M. tuberculosis* requires special consideration as various factors can influence its final DNA concentration and quality. Firstly, considering the slow growth rate of *M. tuberculosis*, the number of cells available on a LJ slope may be limited and only produce small amounts of DNA. Single colonies can be sub-cultured onto 7H10 solid media but this requires additional time and introduces a new contamination risk. Secondly, the mycobacterial cell wall is particularly robust and difficult to lyse due its unique composition. The complex structure of lipophilic molecules and branched polysaccharides within the cell wall is responsible for low cellular permeability and high resistance to extracellular stresses thereby resulting in a formidable protective barrier which is difficult to lyse. In general, cell wall disruption can be achieved with enzymatic or chemical lysis, mechanical disruption, or a combination of these (Amaro, Duarte, Amado, Ferronha, & Botelho, 2008). Thirdly, commercially available extraction kits are generally designed to provide enough material for simple PCR reactions with amplification of small and specific targets. However, the introduction of WGS has introduced new minimum requirements of DNA quantity and quality. As such, conventional DNA extraction methods for *M. tuberculosis* often provide low DNA yields of unsuitable quality for WGS (Kaser, Ruf, Hauser, & Pluschke, 2010).

Early protocols for the extraction of mycobacterial DNA have been described by various authors (Imai et al., 1994) (Gonzalez-y-Merchand, Estrada-Garcia, Colston, & Cox, 1996) (Belisle & Sonnenberg, 1998) (Thomson et al., 2005) but they do not generally focus on genomic DNA production for next generation sequencing (NGS)

applications. Effective protocols have been previously proposed for the isolation of high quality DNA from other difficult samples, such as plant and fungal tissues, including the original cetyltrimethyl ammonium bromide (CTAB) based method (Doyle & Doyle, 1987) (van Soolingen, Hermans, de Haas, Soll, & van Embden, 1991) (Ausubel et al., 2003) (Abu Almakarem, Heilman, Conger, Shtarkman, & Rogers, 2012). However, these protocols typically require higher reaction volumes, additional chemicals and are time-intensive. Some commercial extraction kits are also available on the market but they are often expensive, particularly when processing numerous samples. Loss of purified DNA through column washes or precipitations involved in these kits is also a concern, particularly when only small amount of sample is available (Healey, Furtado, Cooper, & Henry, 2014).

1.4 Evolution

Despite the limited genetic diversity between strains, the differences that have been observed have been useful in the investigation of the global evolution of the *M. tuberculosis* complex. With the lack of horizontal gene transfer, deletions and point mutations can be used to track the global evolution of this pathogen. It is widely believed that the *M. tuberculosis* complex emerged in Africa, most likely the Horn of Africa, and subsequent global spread occurred alongside human migrations (Blouin et al., 2012) (Comas et al., 2013). The recent application of WGS has allowed a better understanding of *M. tuberculosis* evolution and, in particular, the discovery of its origin in *M. canettii*, first described in 1969 and now considered the progenitor species from which *M. tuberculosis* emerged (Blouin et al., 2014).

The *M. tuberculosis* complex, containing the main human-infecting species, can be classified into seven major lineages (Kato-Maeda et al., 2011) (Blouin et al., 2012) (Galagan, 2014):

- Lineage 1 – Containing the East African-Indian (EAI), Manila family and some Manu (Indian) strains.
- Lineage 2 – Represented by the Beijing group.
- Lineage 3 – Consists of the Central Asian (CAS) strains.
- Lineage 4 – Formed by the Ghana and Haarlem and Latin America-Mediterranean (LAM) strains. It is also known as the Euro-American lineage. Subtypes within this type include Latin American Mediterranean, Uganda I, Uganda II, Haarlem, X, and Congo.

- Lineage 5 and 6 correspond to *M. africanum* and they are generally isolated in West Africa.
- Lineage 7 – Isolated from the Horn of Africa (Ethiopia).

It is important to clarify that major lineages can be predicted by two different methods: conformal Bayesian network (CBN) and RULES method. CBN can predict major lineages using spoligotyping (which is based on polymorphisms in the direct repeat locus), MIRU typing (based on the Mycobacterial Interspersed Repetitive Units), or using both at the same time (Aminian, Shabbeer, & Bennett, 2010). Alternatively, with the RULES method, a set of rules was developed to classify isolates into these strain groups, based on the presence or absence of spacer sequences in the spoligotype pattern and the values of the MIRU loci. An online tool (TB-lineage) is available to facilitate the automatic classification (Shabbeer et al., 2012).

A study using mycobacterial tandem repeat sequences as genetic markers has shown that the *M. tuberculosis* complex consists of two independent clades: one composed exclusively of *M. tuberculosis* lineages from humans and the other composed of both animal and human isolates. The latter also likely derived from a human pathogenic lineage, supporting the hypothesis of an original human host (Wirth et al., 2008). The same authors estimated the age of the *M. tuberculosis* complex at 40,000 years, coinciding with the expansion of human populations out of Africa. The most recent common ancestors of the EAI and LAM strains have been estimated to be 13,700 and 7,000 years ago, respectively, whilst the Beijing and CAS strains diverged about 17,100 years ago. All types of the *M. tuberculosis* complex began their current expansion about 5000 years ago. In particular, the Beijing strain appears to have been the most successful with around a 500-fold increase in effective population size since its expansion began, whilst those limited to Africa seem to be the least successful (Wirth et al., 2008).

1.5 Epidemiology of tuberculosis

WHO estimates that one third of the world population is latently infected with *M. tuberculosis* and there were an estimated 10.4 million new cases worldwide (5.9 million men, 3.5 million women and 1.0 million children) in 2015. People living with HIV accounted for 1.2 million (11%) of all new TB cases (WHO, 2017). TB has also re-emerged as a public health concern in high-income countries as a result of immigration from endemic areas, the HIV pandemic and immunosuppressive treatments. After decades of relative neglect, global efforts to control TB were

reinvigorated in 1993 when the WHO declared it as a global health emergency (WHO, 1993). Several challenges emerged over the years and to address these problems, a new plan entitled *Stop TB Partnership: The Global Plan to Stop TB 2006-2015* was developed by the WHO and its partners (Raviglione & Uplekar, 2006) (Maher et al., 2007). As already mentioned, a new resolution was passed in 2014 approving with full support its *End TB Strategy* (WHO, 2015), with the aim to end the global TB epidemic by 2030, to reduce TB deaths by 95% and to cut new cases by 90%. Despite significant improvements in controlling its spread and in optimizing diagnosis and treatment, TB still remains a major health challenge for the world.

1.5.1 Global and national epidemiology

The epidemiology of TB varies substantially around the globe, with the highest incidence (100 per 100,000 or higher) observed in sub-Saharan Africa, India, and the islands of Southeast Asia and Micronesia. Intermediate rates of TB (26 to 100 cases per 100,000) are reported in China, Central and South America, Eastern Europe, and northern Africa. Low rates (less than 25 cases per 100,000 inhabitants) occur in the United States, Western Europe, Canada, Japan, and Australia (Corbett, Marston, Churchyard, & De Cock, 2006) (Wright et al., 2009). In 2015, six countries accounted for 60% of the new cases: India, Indonesia, China, Nigeria, Pakistan and South Africa (WHO, 2017). Globally, mortality rates (excluding deaths among HIV-positive people) have decreased by 45% since 1990 and the current data suggests that the international target of a 50% reduction in TB mortality by 2015 have been achieved (Glaziou, Sismanidis, Floyd, & Raviglione, 2015). Worldwide, the rate of decline in TB incidence was only 1.5% from 2014 to 2015, and this needs to accelerate to a 4–5% annual decline by 2020 to reach the milestones of the *End TB Strategy*. In 2015, there were an estimated 480 000 new cases of multidrug-resistant TB (MDR-TB) and an additional 100 000 people with rifampicin-resistant TB (RR-TB). India, China and the Russian Federation accounted for 45% of these cases. Despite the progress made, TB remains one of the top 10 causes of death worldwide (WHO, 2017).

In England, there has been a steady decline in the incidence of TB over the past four years, down to 10.5 per 100,000 (5,758 cases) in 2015, a reduction of one-third since the peak of 15.6 per 100,000 (8,280 cases) in 2011 (PHE, 2015). The main burden of disease was concentrated in large urban areas, with the highest proportion of cases in London (rate of 26.2 cases per 100,000), followed by other hot spots in Leicester, Birmingham, Luton, Manchester and Coventry. The number of resistant case has also

decreased since 2011, with 54 cases (1.6%) with initial MDR/RR-TB in 2015. Ten of these cases had XDR-TB (PHE, 2015). NHS England, along with public health England (PHE), has announced a *Collaborative Tuberculosis Strategy for England 2015-2020,* a £11.5 million investment to decrease the number of TB cases and ultimately eliminate TB as a public health problem in England (NHS, 2015).

1.5.2 Outbreaks of tuberculosis

In industrialized countries TB is increasingly associated with specific population subgroups: prisoners, immigrants from countries with high endemicity, refugees, and homeless people (Burki, 2011). Outbreaks have been described in US and Western Europe (Mindra, Wortham, Haddad, & Powell, 2017) (Bergmire-Sweat et al., 1996) (Bedini et al., 2016) (Faccini et al., 2013) (Golesi, Brignatz, Bellenfant, Raoult, & Drancourt, 2013). A review by the CDC has highlighted 21 TB outbreaks in the US during the period 2009-2015: the majority of patients involved (83%) used alcohol excessively or illicit substances, another 45% had been homeless in the year before diagnosis and 26% patients had a mental illness (Mindra et al., 2017). The primary cause of an outbreak is thought to be the delay in identifying the index case, where the screening result interpretation might have contributed to the delay. Environmental factors, such as closed spaces with poor ventilation, people spending long periods living together and sharing rooms during cold winter months, all contributed to facilitate the spread of the infection (Kato & Kuwabara, 2014). Outbreaks in prisons, involving MDR strains, have been described (Bergmire-Sweat et al., 1996) (Bedini et al., 2016). However, shared public places are potentially sites of transmission, and school outbreaks have been reported in different countries (Faccini et al., 2013) (Golesi et al., 2013), including even outbreaks in small and remote communities, such as in Greenland (Lillebaek et al., 2013).

Of particular interest is the worldwide increase in isoniazid-resistant (INH-R) TB. This is important because INH resistance reduces the probability of treatment success, may facilitate the spread of multidrug resistance and may reduce the effectiveness of INH preventive therapy. It has been linked to poorer treatment outcomes, post-treatment relapse and death (Stagg, Lipman, McHugh, & Jenkins, 2017). The highest incidence of INH resistance is in the Eastern Europe area with 44.9% of new TB cases with INH-R whilst in all other regions combined, this is 13.9% (one in seven incident TB cases).

In the UK, the TB incidence is 10.5 active cases/100,000 and in London it has risen 30% in the last decade with a jump from 2,309 cases in 1999 to 3,426 in 2012 (PHE, 2015). The challenge this expansion poses to public health authorities is illustrated by the outbreak of INH-R TB in North London in which over 400 cases have been diagnosed since 1995 (Maguire, Ruddy, & Bothamley, 2006) (Ruddy, Davies, & Yates, 2004) (Casali et al., 2016). The original focus was identified in Enfield and Haringey with migration outwards across North London. Fifty per cent of cases were born in the UK; they were from white or black Caribbean ethnicity with a strong association with drug use and prison detention. Conventional epidemiological analysis revealed that adherence to treatment was poor in one third of these patients and several went on to acquire further resistance including MDR TB. A second clinically-relevant feature of this outbreak was the high transmission of infection to contacts (11%) compared with other documented outbreaks (0.7-2%), which could not be explained from the epidemiological data (Maguire et al., 2011). This leads us to the hypothesis that this INH-R strain has unique biological characteristics that serve to prolong the outbreak.

1.6 The challenges of drug resistance

Despite the introduction 40 years ago of an effective combination regimen (INH, RIF, PZA and EMB), TB remains a major cause of morbidity and mortality globally. New drugs and regimens are emerging and progressing through various clinical trials worldwide. However, treatment of drug resistance is still challenging with less than half of patients successfully completing that treatment, mainly due to high mortality and loss of follow up (Wallis et al., 2016). Other challenges include the need for accurate and early diagnosis, screening for drug resistance, administration of effective treatment regimens for at least 6 months and follow-up support (Zumla, Nahid, & Cole, 2013). Substantial efforts have been recently invested in drug development, particularly to create new regimens that can significantly shorten the duration of treatment, improve tolerability and reduce drug-drug interactions (Dooley et al., 2012). A multi strategy approach may be the solution to tackle the challenge of drug resistant TB. This will involve not only the discovery of new antibiotics (and revising the role of/repurposing old ones) but also a streamlined and rapid screening process in the laboratory phase to select the most promising compounds, research on alternatives to antibiotics (i.e. nanoparticles and bacteriophages) and the use of nanotechnologies with a potential role in increasing the uptake of drugs directly inside the cells (antibiotic carriers). Fundamentally, an all-inclusive approach to circumvent antibiotic

resistance in TB is required. This should involve public health bodies at the local, national and international levels with power to fast track the most promising drug and vaccine candidates and next generation diagnostic tests.

1.6.1 Drug resistance and fitness

It is commonly accepted that antibiotic resistance, whether caused by target alteration or by other mechanisms, generally confers a reduction in the fitness of bacteria. This fitness cost is expressed as reduced growth, virulence or transmission (Andersson, 2006). In *M. tuberculosis*, *rpoB* mutations have been found to cause a fitness disadvantage in some clinical isolates (O'Sullivan, McHugh, & Gillespie, 2010) but other authors have demonstrated that some mutations have only a minor cost, raising the possibility of reversion to susceptibility (Billington, McHugh, & Gillespie, 1999). Similar findings were demonstrated in INH-resistant strains due to *katG* Ser315Thr and the *rpsL* Lys43Arg mutations that were less likely to be found in multiple-case households (Salvatore et al., 2016). In contrast, the analysis of strains with *katG* or *inhA* promoter mutations from San Francisco over 9-year period showed that they were more likely to spread than strains with other mutations (Gagneux, Burgos, & DeRiemer, 2006). The severity of fitness cost also depends on the type of mutation conferring resistance and laboratory strains seems to suffer more from fitness disadvantage compared with clinical isolates. The latter could harbour compensatory mutations to balance the initial fitness cost (Gagneux, Long, et al., 2006) (Comas, Borrell, Roetzer, et al., 2011). The effect of resistance to other drugs and fitness cost is still unclear. Streptomycin-resistant strains do not seem to be affected by any fitness cost (Sander et al., 2002) and the application of WGS to MDR and XDR outbreak strains collected from human immunodeficiency virus (HIV)-infected patients has shown that these virulent isolates were able to evolve without any evident fitness cost or other XDR-specific mutation (Motiwala et al., 2010). Mathematical modelling has demonstrated that even when the average relative fitness of MDR strains is low, a small subpopulation of a relatively fit MDR strain may eventually outcompete both the drug-sensitive strains and the less fit MDR strains (Cohen & Murray, 2004). Thus, there is still uncertainty regarding the effect of drug resistance on fitness but studies are important to understand the spread of particular outbreak strains and the use of WGS may allow us to further elucidate the genetic mechanisms.

1.6.2 Mutation rate

The mechanisms by which drug resistance emerges are not wholly understood and WGS analysis suggests that the response to drug exposure might be much more complex than initially thought. There are indications that the acquisition of clinically significant resistance to certain drugs might be a stepwise process that often involves an initial low-level mutation that acts as a gateway for high-level resistance (Fonseca, Knight, & McHugh, 2015). A key question is why some strains of *M. tuberculosis* are preferentially associated with multi drug resistance. A recent study has demonstrated that different lineages have different potential of developing further resistance, with lineage 2 (East Asian and Beijing) acquiring drug resistance *in vitro* more rapidly than lineage 4 (Euro-American) (Ford et al., 2013). Higher resistance rates can be attributed to a higher mutation rate and *in vitro* models of mutation rate seem to correlate well with clinical strains as determined by WGS of clinical isolates (Ford et al., 2011). Additionally, WGS data also suggest that the mycobacterial mutation rate may be modulated within the host, raising the possibility that environmental factors might act as key mutagens during *M. tuberculosis* infection (McGrath, Gey van Pittius, van Helden, Warren, & Warner, 2014). Some authors have also demonstrated that sub-optimal concentrations of fluoroquinolones may significantly increase the mutation rate and this highlights the importance of adequate drug levels in preventing the occurrence of resistance (Gillespie, Basu, Dickens, O'Sullivan, & McHugh, 2005). Genome sequences of serial isolates from the same poorly compliant patient uncovered a dramatic turnover of competing lineages with rapid acquisition of antibiotic resistance. Resistance was mediated by individual mutations as well as a gradual increase in fitness. This was likely driven by stable gene expression reprogramming and it highlights the impressive adaptive potential of *M. tuberculosis* (Eldholm et al., 2014). Thus, the determination of mutation rate, combined with a fitness assay, is an essential step in understanding the potential of a specific strain to spread and further develop additional resistance.

1.7 Treatment of tuberculosis

TB is treated by taking different drugs for at least 6 months and prolonged courses are required in cases of extrapulmonary involvement. There are 10 drugs currently approved by the U.S. Food and Drug Administration (FDA) for treating TB, but further antibiotics (fluoroquinolones, aminoglycosides and others) are routinely used in the treatment of drug resistant cases (Nahid et al., 2016). Four main drugs are considered

essential by the WHO: Isoniazid (INH), Rifampicin (RIF), Pyrazinamide (PZA), Ethambutol (EMB) (WHO, 2009). Table 2 provides a list of common abbreviations of antituberculous drugs.

In general, TB treatment consists of two phases: an intensive phase (with at least four drugs for a period of two months) and a continuation phase (generally two drugs for at least four months). Patients should be managed with directly observed therapy (DOT) to ensure adherence and prevent the emergence of drug resistance. DOT involves providing antituberculous drugs directly to the patient and watching as the patient swallows the medication. Continuous treatment is particularly important in the intensive phase of therapy when the organism burden is highest and the chance of developing drug resistance is greatest. Different protocols are available for drug susceptible TB (Table 3) and various modifications can be made in case of drug side effects (in particular, hepatotoxicity) and if resistant strains are present (where second and third line antibiotics will be used). In instances of drug-resistance, the WHO treatment guidelines recommend extending the duration of the intensive phase to at least 8 months with a total duration (intensive and maintenance) of up to 24 months if there is history of previously failed treatment (Falzon et al., 2017). In the UK, guidelines from the National Institute for Health and Care Excellence (NICE) are recommended and updated regularly (https://www.nice.org.uk/guidance/ng33).

Common abbreviations of antituberculous drugs	
Isoniazid	INH or H
Rifampicin	RIF, R or RMP
Pyrazinamide	PZA or Z
Streptomycin	SM or S
Ethambutol	EMB or E
Ethionamide	ETH or ETA
Amikacin	AK OR AM
Kanamycin	KM
Capreomycin	CM
Cycloserine	CS
Para-amino-salicylic acid	PAS
Levofloxacin	Levo or LVX
Moxifloxacin	Moxi or MFX
Ofloxacin	Oflox or OFX

TABLE 2: COMMON ABBREVIATIONS OF ANTITUBERCULOUS DRUGS (HEIFETS, 2013).

Various abbreviations are in use for the different antituberculous drugs. One letter abbreviations are generally used to briefly summarized a combination protocol (i.e. HRZE for isoniazid, rifampicin, pyrazinamide, ethambutol).

Protocols for the treatment of susceptible pulmonary tuberculosis		
Intensive phase	**Continuation phase**	**Comments**
Protocol 1: INH, RIF, PZA, EMB - Daily for 8 weeks	Protocol 1: INH and RIF – Daily for 18 weeks	This is the preferred protocol for patients with newly diagnosed pulmonary TB
Protocol 2: INH, RIF, PZA, EMB - Daily for 8 weeks	Protocol 2: INH and RIF – Three times weekly for 18 weeks	Preferred alternative protocol when DOT is difficult to achieve during the continuation phase
Protocol 3: INH, RIF, PZA, EMB - Three times weekly for 8 weeks	Protocol 3: INH and RIF – Three times weekly for 18 weeks	Use caution with this protocol as missed doses can lead to treatment failure and selection of drug resistance
Protocol 4: INH, RIF, PZA, EMB – Daily for two weeks, the twice weekly for 6 weeks	Protocol 4: INH and RIF – Twice weekly for 18 weeks	Not to be used in HIV patients

TABLE 3: PROTOCOLS FOR THE TREATMENT OF PULMONARY TUBERCULOSIS CAUSED BY SUSCEPTIBLE STRAINS (NAHID ET AL., 2016).

The table above shows various protocols for the treatment of susceptible pulmonary tuberculosis as per WHO guidelines. Please note that protocols 3 and 4 have a higher failure rate, they should be only used in selected cases and they are not recommended in the British NICE guidelines.

1.7.1 Early bactericidal activity

The studies on the early bactericidal activity (EBA) of antituberculous drugs (during the first 2 days of treatment) are of particular interest. INH and RIF have a rapid killing effect on the actively replicating bacilli causing infection, whilst the EBA is much reduced for the other first line drugs, with EMB and PZA having a moderate and minimal EBA, respectively (Jindani, Dore, & Mitchison, 2003) (Botha et al., 1996). EBA studies of PZA have been confusing as it has almost no detectable activity during the first 2–4 days of treatment, but it does eventually kill at least 95% of the mycobacterial population by day 14 (Diacon & Donald, 2014). PZA differs from other anti-tuberculosis drugs in showing greater bactericidal activity the slower the bacillary metabolic activity is, with sterilizing activity against those slowly replicating bacilli. Streptomycin (SM), similar to other aminoglycosides, was initially identified as highly bactericidal (L. Dickinson, 1947), but this was not confirmed in EBA studies (Jindani, Aber, Edwards, & Mitchison, 1980) (Donald et al., 2002). This complexity highlights the importance of combination treatment, in particular to prevent the emergence of resistance, but also the synergistic action of antituberculous drugs. Thus, the introduction of any new compound also needs to be tested in combination with existing drugs to make sure it has a complementary action or, at least, not antagonistic action (Palomino, Ramos, & da Silva, 2009).

1.7.2 Development of drug resistance

Drug resistance in *M. tuberculosis* is not an unusual event as resistance to streptomycin, para-aminosalicylic acid (PAS) and INH developed shortly after their introduction as first line treatments. Genetic resistance to antituberculous drugs is due to spontaneous mutations at the low frequency of 10^{-6} to 10^{-8} mycobacterial replications (Y. Zhang & Yew, 2015). *M. tuberculosis* does not have mobile genetic elements such as plasmids and transposons, which are commonly present in other bacteria. Because such mutations conferring drug resistance are unlinked, the probability of simultaneously developing drug resistance to three drugs is very low, in the region of 10^{-18} to 10^{-20}. Thus, the chance of developing drug resistance during treatment with three active drugs is virtually non-existent and this highlights the importance of combination treatment (Canetti, 1965). However, once resistance has developed, the subsequent selection of resistant strains is due to suboptimal drug exposure for various reasons (poor compliance to treatment, wrong dose prescription, problems with the drug supply, unreliable drug formulations) and transmission due to

diagnostic delays, overcrowding and other inadequate infection control measures (Y. Zhang & Yew, 2015).

Resistance in *M. tuberculosis* is caused by spontaneous mutations in chromosomal genes due to sub-optimal drug concentrations (*acquired drug resistance*) (Kochi, Vareldzis, & Styblo, 1993) and it varies among antituberculous drugs, from a frequency of 10^{-8} organisms for INH to 10^{-9}-10^{-10} for RIF (David, 1970) (Karakousis, 2009) (McGrath et al., 2014). This differs to *intrinsic drug resistance* that refers to the ability of an organism to resist the activity of a particular antimicrobial compound due to its structural characteristics. In the case of *M. tuberculosis,* its unique cell wall represents a strong hydrophobic barrier against common antibiotics (Karakousis, Bishai, & Dorman, 2004).

1.7.3 Antituberculous drugs, mechanisms of action and resistance

The molecular mechanisms of drug resistance for first and second line drugs have been elucidated over decades of research. Recently, the introduction of whole genome sequencing (WGS) has also allowed the rapid application of this technology to identify the mechanism of new drugs before their clinical use by sequencing laboratory resistant mutants (Andries et al., 2005).

Isoniazid – INH is the most widely used antituberculous drug and it has been a core component for the treatment of TB since its discovery in 1952. It is only active against growing tubercle bacilli, with no activity under anaerobic conditions or against non-replicating bacilli (Y. Zhang & Yew, 2015). INH is a prodrug that is activated by the catalase peroxidase enzyme (encoded by the *katG* gene) into its active forms which is highly toxic for *M. tuberculosis*. These active forms include reactive oxygen species (superoxide, peroxide and hydroxyl radical) (Shoeb, Bowman, Ottolenghi, & Merola, 1985), nitric oxide (Timmins, Master, Rusnak, & Deretic, 2004), isonicotinic-acyl radical and anion (Rawat, Whitty, & Tonge, 2003), and certain electrophilic species (Johnsson, King, & Schultz, 1995). All of them are able to inhibit the inhA enzyme (enoyl-acyl carrier protein reductase), responsible for the elongation of fatty acids in mycolic acid synthesis (Banerjee et al., 1994).

Resistance to INH is a complex process. *KatG* mutations causing loss of the catalase and peroxidase activity are the most common cause of INH resistance (Y. Zhang, Heym, Allen, Young, & Cole, 1992). More than a hundred mutations have been described, conferring various levels of resistance with MICs from 0.2 to 256 mg/L, and

S315T is the most common mutation harboured by 50-95% of INH-resistant clinical isolates (Hazbon et al., 2006). Mutations in the *inhA* gene (encoding for the target enzyme) can also confer INH resistance (and cross-resistance to ETH) but they are generally low-level and less frequent compared to *katG* mutations (Banerjee et al., 1994). Other mutations can involve the *oxyR-aphC* region (*inhA* promoter) (Dalla Costa et al., 2009), *ndh* (NADH dehydrogenase) (Miesel, Weisbrod, Marcinkeviciene, Bittman, & Jacobs, 1998) and *furA-katG* intergenic region (causing downregulation of *katG*) (Ando et al., 2011). About 10-15% of low-level INH-resistant strains do not have mutations in the *katG* or *inhA* genes and they are likely to have alternative mechanisms for resistance (Hazbon et al., 2006).

Rifampicin – RIF was introduced in 1972 and is another important first-line drug with activity against both growing and stationary bacilli (Almeida Da Silva & Palomino, 2011). RIF interferes with RNA synthesis by binding to the beta-subunit of the RNA polymerase. Mutations in the *rpoB* gene are accountable for up to 96% of RIF-resistant isolates (Telenti et al., 1993). Various mutations have been described and some confer cross-resistance to other rifamycins (including rifabutin) (Cavusoglu, Karaca-Derici, & Bilgic, 2004). An important finding related to RIF resistance is that the majority of RIF-resistant strains also show resistance to other drugs, in particular INH, and RIF resistance has been proposed as a molecular surrogate to detect multi drug resistant strains (Traore, Fissette, Bastian, Devleeschouwer, & Portaels, 2000).

Pyrazinamide – PZA is a peculiar antituberculous drug as it has a high sterilizing activity at acidic pH (5.5) but no activity at all in normal culture conditions (McDermott & Tompsett, 1954). Its activity is also enhanced by low oxygen and anaerobic conditions (Wade & Zhang, 2004). Due to this ability to inhibit semi-dormant bacilli in the acidic granuloma environment, PZA was pivotal in shortening TB treatment to 6 months from what was previously 9-12 months (Mitchison, 1985). Pyrazinamide is a structural analogue of nicotinamide and is a pro-drug that requires activation to its active form (pyrazinoic acid) by the enzyme pyrazinamidase/nicotinamidase (PZase) (Konno, Feldmann, & McDermott, 1967). The mechanism of action is not completely understood. It has been postulated that pyrazinoic acid may accumulate inside the bacilli causing cellular damage (Y. Zhang & Mitchison, 2003) or it could inhibit the fatty acid synthase (Zimhony, Vilcheze, Arai, Welch, & Jacobs, 2007). Mutations in *pncA* (the gene encoding for PZase) are the main mechanisms of resistance but there is a high degree of diversity among PZA-resistant isolates and other genes have been implicated (Ramirez-Busby et al., 2017).

Ethambutol – EMB is another first line drug to use in combination with INH, RIF and PZA. It has a bacteriostatic action and is only active against growing bacilli, by interfering with the synthesis of cell wall arabinogalactan (Takayama & Kilburn, 1989). The enzyme arabinosyl transferase, encoded by *embB*, has been proposed as the drug target in *M. tuberculosis* (Belanger et al., 1996). Mutations in the *embB* gene, in particular codon 306, are responsible for up to 65% of resistant isolates (Sreevatsan, Stockbauer, et al., 1997). However, the remaining 35% of resistant cases do not have mutations in *embB* suggesting that there must be other mechanisms of resistance (Y. Zhang & Yew, 2015). Similar to PZA, EMB is another antibiotic that causes difficulties with laboratory testing and discordant results are frequently found (Laszlo, Rahman, Espinal, & Raviglione, 2002).

Streptomycin and aminoglycosides – Aminoglycosides are a class of antibiotics active against different bacteria, including *M. tuberculosis*. SM was the first antituberculous drug to be discovered, with rapid killing activity against growing bacilli but no effect on intracellular and non-growing bacilli. It inhibits the genome translation by binding to the 16S rRNA (Moazed & Noller, 1987). Resistance is mostly due to mutations in *rrs* or *rpsL* genes, which produce alterations of the SM binding site, but these have only been identified in around 50% of the resistant strains (Honore & Cole, 1994). Other mutations, in particular in the *gidB* gene, have been associated with low level resistance (Spies, da Silva, Ribeiro, Rossetti, & Zaha, 2008). KM and AMK have a similar mechanism of action and mutations in the *rss* gene also generally confer resistance to these drugs (Alangaden et al., 1998).

Fluoroquinolones – Moxi, Levo and Oflox are bactericidal antibiotics currently used as second line drugs in the treatment of TB and they target the type II topoisomerase (DNA gyrase) (Aubry, Fisher, Jarlier, & Cambau, 2006). This is an enzyme that catalyses the supercoiling of DNA and it is composed by two subunits (A and B) encoded by *gyrA* and *gyrB* (Berger & Wang, 1996). Moxi was originally proposed as a first-line drug with the aim of shortening duration of treatment (Nuermberger et al., 2004) but an international clinical trial failed to achieve this aim (Gillespie, Crook, McHugh, Mendel, Meredith, Murray, et al., 2014). Various mutations in *gyrA* have been associated with resistance (Cheng et al., 2004) (Sun et al., 2008), whilst *gyrB* mutations appear to be a much rarer occurrence (Pitaksajjakul et al., 2005). However, not all mutations are associated with cross-resistance to all the antibiotics of this class and mutations also confer different levels of MICs (J. Li et al.,

2014). Finally, efflux pumps against fluoroquinolones have also been described in *M. tuberculosis* (Pasca et al., 2004).

P-amino salicylic acid – PAS was one of the first antibiotics to be used against TB but its mechanism of action has been never clearly elucidated (Lehmann, 1946). It has been proposed that, being an analogue of para-amino benzoic acid (PABA), PAS competes with PABA for dihydropteroate synthase, interfering in the process of folate synthesis. A study using transposon mutagenesis identified mutations in the *thyA* gene associated with resistance to PAS that were also present in clinical isolates resistant to PAS (Mathys et al., 2009). A recent study has also identified various missense mutations in *folC* encoding dihydrofolate synthase that conferred resistance to PAS in laboratory isolates of *M. tuberculosis* (Zhao et al., 2014). Nevertheless, less than 40% of PAS-resistant strains had such mutations indicating that other mechanisms of resistance to the drug might exist and further studies are needed to finally elucidate mechanisms of action and resistance.

Beta-lactams – Even if beta-lactam antibiotics have been traditionally regarded as ineffective against TB, their role has been reconsidered due to the challenge of drug resistant TB (Diacon et al., 2016). Laboratory experiments suggest that resistance to beta-lactams is generally caused by the presence of beta-lactamase enzymes and it could be reversed with the addition of beta-lactamase inhibitors (Chambers et al., 1995). Early studies on the use of amoxicillin-clavulanate in patients with TB showed an early bactericidal activity comparable to ofloxacin (Chambers, Kocagoz, Sipit, Turner, & Hopewell, 1998). Carbapenems (a particular class of beta-lactams) are widely used for the treatment of many infections (Papp-Wallace, Endimiani, Taracila, & Bonomo, 2011) and their role against TB needs further evaluation (Jaganath, Lamichhane, & Shah, 2016).

Mechanisms of action of, and resistance to, other second and third line drugs (ethionamide, cycloserine, linezolid, bedaquiline and delamanid) have been described (Palomino et al., 2009) (Almeida Da Silva & Palomino, 2011) but they are beyond the remit of this thesis.

Mechanisms of action of common antituberculous drugs and genes associated with resistance		
Drug	**Mechanism of action**	**Genes associated with resistance**
Isoniazid	Inhibition of enoyl-acyl carrier protein reductase, involved in the synthesis of mycolic acid synthesis	*katG, inhA, ndh, oxyR-aphC, furA*
Rifampicin	Inhibition of RNA synthesis	*rpoB*
Pyrazinamide	Mechanism of action is not completely understood (postulated cellular damage or inhibition of fatty acids synthase)	*pncA*
Ethambutol	Inhibition of cell wall arabinogalactan	*embB*
Aminoglycosides	Inhibition of genome translation by binding to the 16S rRNA	*rrs, rpsL*
Fluoroquinolones	Inhibition of DNA gyrase	*gyrA, gyB*
P-amino salicylic acid	Unknown (proposed inhibition of folate synthesis)	*thyA, folC*

TABLE 4: MECHANISMS OF ACTION OF COMMON ANTITUBERCULOUS DRUGS AND GENES ASSOCIATED WITH RESISTANCE.

The main genes associated with drug resistance are shown above. Whilst the genes associated with isoniazid, rifampicin and fluoroquinolones resistance cover the majority of resistant strains, the mechanisms of resistance for ethambutol and para-amino salicylic acid are not completely understood and only 65% and 40% of resistant strains, respectively, have mutations involving those genes.

1.7.4 Multi drug resistant (MDR) and extensively drug resistant (XDR) tuberculosis

Since the discovery of streptomycin, numerous other drugs with antituberculous activity have been discovered and new forms of antibiotic resistance have been described (Johnson et al., 2006). Multi drug-resistant (MDR) strains are defined by resistance to the two most commonly used drugs in the current four-drug (or first line) regimen, INH and RIF. WHO treatment standards require at least four drugs in order to avoid the development of further resistance. Extensively drug-resistant (XDR) is defined as TB that is also resistant to any fluoroquinolone, and at least one of three injectable second-line drugs (capreomycin, kanamycin, and amikacin), in addition to INH and RIF. The emergence and spread of MDR and XDR TB have complicated how TB is treated. Successful XDR-TB treatment, particularly in resource-limited settings, may be almost impossible. In a 2006 XDR-TB outbreak in KwaZulu-Natal, South Africa, 52 of 53 people who contracted the disease died within months. It is estimated that 70% of XDR-TB patients die within a month of diagnosis (WHO, 2010a).

1.7.5 Susceptibility testing

Drug susceptibility testing for antituberculous drugs is a complex procedure and requires an understanding of many issues, including drug resistance mechanisms, potency and stability of drugs during laboratory manipulation and the antimycobacterial activity of compounds when incorporated into different media (CDC, 2012). The current methods have been developed over several decades and are restricted to specialised reference laboratories, as they are technically demanding, require appropriate isolation facilities and can be difficult to interpret. Three main methods are currently in use and recommended by international guidelines (Drobniewski, Rusch-Gerdes, & Hoffner, 2007):

1. *Absolute concentration*, where the drug is incorporated into solid agar using different dilutions (including Lowenstein–Jensen medium) or in a broth dilution method. Resistance is defined as the lowest concentration of the drug that inhibits growth (<20 colonies) (Drobniewski et al., 2007). Variation is generally due to erroneous drug concentration or inoculum size (Vareldzis et al., 1994).
2. *Resistance ratio*, where the minimum inhibitory concentration (MIC) for a given isolate is divided by the MIC for a standard susceptible strain (e.g., H37Rv). If the ratio is ≤2 or ≥8, the isolate is fully susceptible or highly resistant respectively.

Inoculum size still needs to be standardized to obtain reproducible results (Drobniewski et al., 2007).

3. *Proportion method*, where a strain is considered susceptible if it contains a proportion of resistant cells below a defined point. The proportion varies with different drugs (e.g., 1% for isoniazid and rifampicin). The introduction of the MGIT BACTEC systems (460 and 960) has allowed semi-automation of this method and lessens concerns regarding the initial inoculum size (Roberts et al., 1983) (Kruuner, Yates, & Drobniewski, 2006).

Other methods have been developed, mostly colorimetric methods that use redox indicators or nitrate reduction (Palomino, Martin, & Portaels, 2007) (Hall, Jude, Clark, & Wengenack, 2011). In particular, the resazurin assay (based on the Alamar Blue fluorescent dye) has been proposed as a simple and inexpensive method for detection of drug resistance in *M. tuberculosis* (Palomino et al., 2002) (A. Martin, Camacho, Portaels, & Palomino, 2003).

The first line drugs for resistance testing are isoniazid, rifampicin, ethambutol and pyrazinamide. This combination of agents reflects the aforementioned protocols for the treatment of drug susceptible *M. tuberculosis*. In case of resistance, secondary antituberculous drugs used for testing are capreomycin, kanamycin, amikacin, cycloserine, ethionamide, ofloxacin, moxifloxacin, para-aminosalicylic acid and streptomycin (Woods, 2000). The concentrations of these drugs are different, depending on the method or medium used (Heifets, 2013) (Table 5) and this highlights the importance of the critical knowledge and expertise necessary for performing such susceptibility testing. In addition, pyrazinamide susceptibility testing is particularly challenging as it requires an acidic pH. For this reason, it can only be performed on liquid cultures and no critical concentrations are available for solid media (Chang, Yew, & Zhang, 2011).

Molecular assays for detecting drug resistance have been developed and they offer several potential advantages, including lower turnaround times and minimal (or possibly no) initial culture period. The mutations associated with resistance are now well-known for some drugs, and potentially any laboratory could design its own in-house PCR (polymerase chain reaction). In particular, two molecular assays have clinical value: the Xpert® MTB/RIF assay (Cepheid, Sunnyvale, USA) and the Line probe assay (LPA) (Hain Lifescience, Nehren, Germany) (Steingart et al., 2013) (Tomasicchio et al., 2016). The Xpert assay is an automated real time-based system

with various advantages, including point of care testing, a closed tube system, ease of performance and result availability within hours. It is endorsed by WHO, but its high cost is a disadvantage. The LPA technology is suitable for reference laboratories, or laboratories where there is proven capacity to conduct molecular testing, as it requires appropriate laboratory equipment and adequately trained laboratory staff (WHO, 2017).

Finally, WGS has the potential to revolutionize *M. tuberculosis* susceptibility testing and it offers added value when compared with traditional susceptibility testing. Phenotypic susceptibility testing methods for TB are technically complex, time consuming and expensive (Koser et al., 2012). Extensive work on WGS and susceptibility testing led by the Wellcome Trust/University of Oxford has successfully demonstrated its implementation into routine diagnostic workflows (Pankhurst et al., 2016) (T M. Walker et al., 2015). This topic will be further discussed later on in this chapter.

Critical Concentrations (µg/ml) for *M. tuberculosis* resistance testing				
	Media/ Testing format			
Drug	**LJ**	**7H10 Agar**	**7H11 Agar**	**MGIT/BACTEC**
Isoniazid	0.2	0.2, 1.0	0.2, 1.0	0.1
Rifampin	40.0	1.0	1.0	1.0
Ethambutol	2.0	5.0, 10.0	7.5	5.0
Pyrazinamide	-	-	-	100
Streptomycin	4.0	2.0, 10.0	2.0, 4.0	1.0
Amikacin	20.0	4.0	6.0	1.0
Kanamycin	20.0	5.0	6.0	-
Capreomycin	20.0	10.0	10.0	1.25
Ethionamide	20.0	5.0	10.0	2.5, 5.0 Suggested (Kruuner et al., 2006)
Cycloserine	-	20.0	60.0	-
PAS	0.5	2.0	8.0	-
Levofloxacin	-	-	4.0	-
Moxifloxacin	-	-	4.0	0.5, 1.0

TABLE 5: CRITICAL CONCENTRATIONS (µg/ml) FOR TESTING RESISTANCE OF *M. TUBERCULOSIS* IN DIFFERENT SOLID AND LIQUID MEDIA (HEIFETS, 2013).

The critical concentration is defined as the lowest concentration of an antituberculous agent *in vitro* that will inhibit the growth of 99% of phenotypically wild type strains of *M. tuberculosis* complex.

1.7.6 Alternatives to antibiotics

The emergence of drug resistance has forced scientists to consider alternatives to antibiotics in the fight against TB. Alternative medicines range from traditional herbal remedies to drugs used for other medicinal purposes but with an unexpected antimicrobial effect (Semenya & Maroyi, 2013) (Amaral, Viveiros, & Kristiansen, 2006). Traditional herbal remedies against *M. tuberculosis* are mostly prepared from leaves (34%) and roots (21%) by the Bapedi traditional healers in South Africa and some literature claims that up to 71% of these remedies have antimicrobial properties and various medicinal uses in other countries (Nicoletti, Galeffi, Messana, & Marini-Bettolo, 1992) (Maroyi, 2011) (Semenya & Maroyi, 2013). They should not be completely discounted without further analysis as traditional herbal remedies have already allowed us to discover incredibly useful compounds: artemisin drugs, derived from the *Artemisia annua* (a herb employed in Chinese traditional medicine) are now a standard treatment worldwide for malaria (Dondorp, Nosten, Stepniewska, Day, & White, 2005). In terms of repurposing drugs, the antipsychotic drug Thioridazine has also shown activity against MDR-TB and other bacteria. Side effects would be a major limitation but it can still be considered where there are no alternative forms of therapy and the patient faces an unfavourable prognosis (compassionate use) (Amaral et al., 2006).

Other non-antibiotic approaches include phage therapy, phage lysins, bacteriocins, and predatory bacteria. Interestingly, several of these methods augment antibiotic efficacy by improving bacterial killing and decreasing antibiotic resistance selection (Allen, Trachsel, Looft, & Casey, 2014). In particular, bacteriophages against mycobacterial species (mycobacteriophages) have been described since 1954 (Froman, Will, & Bogen, 1954) and there is a periodical and growing interest as potential alternative to antibiotics against *M. tuberculosis* (Hatfull, 2014) (Basra, Anany, Brovko, Kropinski, & Griffiths, 2014). Immunotherapies and vaccines are beyond the scope of this thesis and would require extensive additional discussion, so will not be mentioned further.

1.7.7 Nanoparticles

The rapid emergence of antimicrobial resistance has caused new interest in nanoparticle technology as promising alternative compounds to treat various bacteria, from *E. coli* to *M. tuberculosis* (Yah & Simate, 2015). The antimicrobial actions of nanoparticles include cidal destruction of cell membranes, blockage of enzyme

pathways, alterations of microbial cell wall, and nucleic materials pathway (Galdiero et al., 2011). However, some mechanisms of the action of nanoparticles are yet to be fully elucidated since some of the NP drugs are still in their infancy (Yah & Simate, 2015).

There are various examples of nanotechnological research focussed on *M. tuberculosis*. Gold nanoparticles have also been developed for diagnostic purposes, as a direct and inexpensive alternative to DNA molecular amplification (Hussain, Samir, & Azzazy, 2013). Some nanoparticles may have a direct antibacterial activity against *M. tuberculosis* whilst others are promising intracellular delivery systems of antituberculous drugs. Various gold and silver nanoparticles, with different combinations of composition, types, and surface modifications, have shown a promising bactericidal effect, with reliable uptake into bacterial cells followed by lysis (Zhou, Kong, Kundu, Cirillo, & Liang, 2012). Nanoparticulate polymers in many forms (latex beads, PLGA, dendrimers, or nanoemulsions) can be used as synthetic carriers for first line and second line drugs. In this way, not only are the drugs sustainably released in organs and plasma, preventing suboptimal concentration (and related risk of developing resistance), but dose, toxicity and side effects can be reduced (Kataria, Sethi, Kaur, Punia, & Kumar, 2015). A variety of different nanoparticles have been tested both *in vitro* and *in vivo* as delivery platforms for antituberculous drugs (Gelperina, Kisich, Iseman, & Heifets, 2005). Mesoporous silica nanoparticle (MSNP) drug delivery systems for INH and RIF have been described which optimize drug loading and intracellular release (Clemens et al., 2012). Encapsulated moxifloxacin within biodegradable poly-butyl cyanoacrylate nanoparticles reduced the MIC for *M. tuberculosis* in macrophages by 10-fold versus that of free drug (Kisich et al., 2007). PLGA nanoparticles can deliver INH, RIF, and PZA and demonstrate prolonged release of drug, allowing less-frequent dosing to achieve a therapeutic effect in animal models (Sharma, Sharma, & Khuller, 2004). Some gold nanoparticles have shown antifungal but not antibacterial activity. A possibility is that bactericidal activity is actually the result of co-existing chemicals not completely removed from the nanoparticle preparation (i.e., gold ions, surface coating agents, and chemicals involved in the synthesis). Regardless of this, gold nanoparticles can also act as carriers or delivery vehicles of antibiotics, thus enhancing the bactericidal effect of the antibiotics (Y. Zhang, Shareena Dasari, Deng, & Yu, 2015). Whilst the nanoparticle delivery systems for TB drugs offer some potential advantages compared to normal drugs, they may also be associated with poor chemical stability, degradation in the

serum and potential toxicity (Foradada, Pujol, Bermudez, & Estelrich, 2000). There is also concern of macrophage cell damage, cytokine release, and inflammation (Semete, Booysen, Kalombo, et al., 2010), although their safety was demonstrated in previous experiments (Semete, Booysen, Lemmer, et al., 2010). More studies on nanoparticles and *M. tuberculosis* are needed to confirm their role as valid alternatives, thus addressing the challenge of drug resistant TB.

1.8 Development of new antituberculous drugs

There is a desperate need for new antituberculous drugs to tackle the problem of multi drug resistance. Currently there is no clear and rapid remedy for the global burden of TB. It is important to consider that traditional drug discovery approaches may not be sufficient. Any new anti-TB drug will be expected to fulfil at least one of the following objectives to be considered successful and worth the effort of years of laboratory research and clinical trials: potential to shorten and simplify the duration of treatment, effectiveness against MDR strains, ability to improve the treatment of latent infection, a reduced side effect profile, compatibility with other drugs (in particular HIV anti-retrovirals) and suitability for oral and paediatric formulations (Zumla et al., 2012). Simple activity against TB will not be considered sufficient if none of the above criteria are met, as the new drug will not add any benefit to the current treatment regimens. Only three drugs out of the 1,556 introduced to the market between 1975 and 2004 (pyrazinamide, rifabutin and rifapentine) are licensed for TB (Chirac & Torreele, 2006). More recently, combined and coordinated initiatives from the TB Alliance Partners (including University College London) have increased the number of compounds in the pipeline but more research and clinical trials are still needed (TBalliance, 2017). Moxifloxacin is one of the new drugs against TB and a clinical trial has been recently completed (REMoX TB). It was originally proposed as a first-line drug with the aim of shortening duration of treatment – laboratory experiments using a murine model showed that the combination of Moxi, RIF and PZA reduced the time needed to eradicate *M. tuberculosis* from the lungs by up to 2 months compared with the standard regimen (Nuermberger et al., 2004). However, the REMoX TB international clinical trial failed to achieve this aim and the addition of Moxi to the first line did not allow the treatment time to be shortened to four months (Gillespie, Crook, McHugh, Mendel, Meredith, Murray, et al., 2014). The European & Developing Countries Clinical Trials Partnership (EDCTP) spent approximately 28 million euro on this clinical trial (EDCTP, 2014) and its failure highlights the challenges when dealing with TB treatment.

Clofazimine (CFM), bedaquiline (BDQ) and pretomanid (PMD) are worth mentioning among other drugs in phase 2/3 clinical trials (TBalliance, 2017) (MSF, 2017). CFM is an interesting example of an old drug that was first introduced in 1969 for the treatment of leprosy and it has not been used much against TB due to its side effect profile. It is now back in the pipeline with new combinations and analogues, trying to optimize the compound and reduce its side effects (Sotgiu et al., 2017). BDQ is a member of the diarylquinoline class of drugs and laboratory tests have shown it to have strong bactericidal and sterilizing properties with activity against replicating and non-replicating bacilli (Riccardi, Del Puente, Magne, Taramasso, & Di Biagio, 2017). PMD (formerly PA-824) is a bicyclic nitroimidazole, causing both cell wall synthesis inhibition (like INH) and respiratory poisoning through nitric oxide release (Manjunatha, Boshoff, & Barry, 2009). In contrast to other antituberculous drugs, PMD exhibits bactericidal activity against both replicating and static bacilli, including strains resistant to first line drugs (Stover et al., 2000).

The challenge with new drugs is not only to develop a compound with high activity against *M. tuberculosis* but also to choose appropriate companion drugs to maximise the sterilizing effect thus reducing the length of treatment. Combinations of new and old drugs with high killing activity have been proposed (BDQ+PMD+MXF+PZA). From observations in the murine model, this combination could potentially reduce length of treatment in humans to only two months (S. Y. Li et al., 2017). However, the REMoX trial has shown us that excellent results in laboratory models may not necessarily translate into clinical practice and clinical trials are still essential to confirm whether regimens reduce the length of treatment. At the same time, there is a constant need to discover new drug targets to populate the pipeline with new potential compounds. Unfortunately, not all of these drugs will end up on the market and it will take at least a decade for some of them to become clinically available since the first discovery. Thus, there is also a need of a strict screening programme in the laboratory when the drug is still in its discovery phase to select and focus all attention and resources on the most promising candidates.

1.9 The potential of whole genome sequencing

Nearly two decades after the complete genome sequence of *M. tuberculosis* was deciphered (Cole, 1999), whole genome sequencing (WGS) has been applied to a vast range of clinical scenarios, with the potential to revolutionize the diagnosis, outbreak investigations, drug and vaccine development, and an overall better

understanding of *M. tuberculosis* evolution and its pathogenicity. During 2017, PHE will progressively introduce routine WGS to the clinical setting of the National Health Service. England will be the first country in the world to pioneer WGS on a national scale for the diagnosis, detection of drug resistance and typing of *M. tuberculosis* (PHE, 2017).

1.9.1 Outbreak investigation and epidemiology

WGS clearly has a role in public health interventions and in the detection of outbreaks and transmission events. Several studies (T. M. Walker, Ip, & Harrell, 2013) (Török et al., 2013) (Roetzer et al., 2013) (Stucki et al., 2014) confirm this role and the higher resolution compared to MIRU-VNTR typing, IS6110 RFLP typing and spoligotyping methods. The latest techniques can be inaccurate (Jagielski, van Ingen, et al., 2014) when tracing route of transmission or distinguishing isolates with minor differences, whilst WGS can identify Single Nucleotide Polymorphisms (SNPs) when comparing different strains. This can be particularly useful when trying to discriminate relapse from reinfection. In a recent paper after the completion of the RIFAQUIN trial (Witney et al., 2017), 36 patients had positive cultures before and after treatment and all strains were typed using 24-loci MIRU-VNTR, in silico spoligotyping and WGS. WGS and MIRU-VNTR both similarly differentiated relapses and reinfections, but WGS provided significant extra information. One pair of clinical samples had an intermediate number of SNP differences (more than 5) and was likely the result of a mixed infection with a pre-treatment minor genotype that was highly related to the post-treatment genotype; this was reclassified as a relapse, yet MIRU-VNTR typing was not powerful enough to detect this. From a global perspective, WGS has allowed the detection of genetic diversity in *M. tuberculosis* with unprecedented resolution and it has been used to understand its evolution, lineages, and variation at global, local and individual levels (Ford et al., 2012).

1.9.2 Drug susceptibility testing

Extensive work on WGS and susceptibility testing has been led by the Wellcome Trust/University of Oxford (and collaborators), with successful demonstration that this approach can be integrated into routine diagnostic workflows (Pankhurst et al., 2016) (T M. Walker et al., 2015). A recent systematic review on the use of WGS for the detection of drug resistance in *M. tuberculosis* has highlighted its role as a promising alternative to existing phenotypic and molecular drug susceptibility testing methods, with strong performance for first line drugs isoniazid and rifampicin (Papaventsis et

al., 2017). However, various practical limitations are still present. Direct WGS from clinical samples has been successfully demonstrated (Brown et al., 2015) but the extraction of genomic DNA from *M. tuberculosis* remains challenging. A powerful Information Technology (IT) infrastructure is needed for the analysis, storage and transfer of multiple samples and cloud-based services and reliable internet connections are necessary. There is also a discrepancy between phenotypic and genotypic susceptibility testing for all other drugs and the need of large database to clarify the role of different mutations. Despite these limitations, WGS has the potential to revolutionize the diagnosis of *M. tuberculosis* in both high- and low-income settings. A growing knowledge of the genetic mechanisms of resistance, combined with improved IT infrastructure, will facilitate the adoption of WGS and improve its clinical utility for drug testing.

1.9.3 Research applications

The complete genome of *M. tuberculosis* includes 4 million base pairs and 4000 genes. However, up to 50% of these genes are still labelled as unknown, uncharacterized or with hypothetical function (Mazandu & Mulder, 2012). WGS analysis should support us in unveiling the real role in the virulence and pathogenicity of these *M. tuberculosis* genes. It has already allowed us to understand TB evolution and its origins from *M. canettii* (Galagan, 2014). WGS also has the potential to revolutionize the process of drug target identification (Sala & Hartkoorn, 2011) and it has been successfully used for the identification of the target of bedaquiline. In this case, the authors selected and sequenced resistant *M. smegmatis* strains and identified mutations in the proton pump of adenosine triphosphate (ATP) synthase (Andries et al., 2005). Understanding bacterial pathogenesis is also expected to provide an instrumental contribution to vaccine development, particularly when targeting those pathogens (such as *M. tuberculosis*) for which the traditional approaches have thus far failed (Scarselli, Giuliani, Adu-Bobie, Pizza, & Rappuoli). Early studies on *M. tuberculosis* were only based on selected genes and led to the notion that strain variability was negligible (Musser, Amin, & Ramaswamy, 2000). The comparison by genome sequencing of clinical strain CDC1551 with the laboratory adapted H37Rv has demonstrated a more extensive variability than had been initially anticipated (Fleischmann et al., 2002), highlighting the importance of understanding genetic diversity in pathogenic strains. Additionally, although most antigens seem to be highly conserved, sequencing has revealed a small subset of antigens that do

show variation (Comas et al., 2010). Thus, further studies are needed to better understand their role in the pathogenesis of *M. tuberculosis*.

1.10 The complexity of whole genome sequencing analysis

The following paragraphs will focus on the analysis of WGS data, to provide the reader with an overview of its complexity through a literature review of the numerous software tools available (with a particular focus on TB). It is important to note that the analysis of WGS data is completely different from traditional PCR results. Electrophoretic gels, cycle threshold (CT) values and small fragment amplifications have been replaced by much more complex analyses requiring powerful computers, extensive storage and bio-informatics support.

1.10.1 Sanger sequencing versus next generation sequencing

Next-generation sequencing (NGS), also known as high-throughput sequencing, is the generic term used to describe modern sequencing technologies including:

- Illumina platform (San Diego, CA, USA).
- Roche 454 (Basel, Switzerland)
- Ion torrent by Thermo Fisher (Waltham, MA, USA)
- ABI SOLiD by Applied Biosystems (part of Thermo Fisher)

These platforms allow DNA and RNA to be sequenced much more quickly and cheaply than the previously used Sanger sequencing, transforming WGS research applications in genomics and molecular biology.

DNA sequencing has significantly evolved since the initial two-dimensional chromatography in the 1970s. The Sanger sequencing method, developed in 1977 by Frederick Sanger and colleagues (Sanger, Nicklen, & Coulson, 1977) overtook 2D chromatography to become the most widely used sequencing method for almost 40 years. A decade later, Applied Biosystems introduced the first automated sequencing platform, the AB370, followed by the AB3730xl in 1998. These platforms were based on capillary electrophoresis and using deoxyribonucleotide chain terminators labelled with fluorescent dyes. These were the instruments used in the Human Genome Project, completed in 2003 (Collins, Morgan, & Patrinos, 2003). The Genome Analyzer emerged in 2005 and it increased the sequencing capacity from 84 kilobase (kb) per run to 1 Gigabase (Gb) per run, revolutionizing sequencing capabilities and

leading the way into the “next-generation” of platforms in genomic medicine (Davies, 2013).

The NGS platforms mentioned are based on different sequencing methods (pyrosequencing, semiconductor sequencing, sequencing by ligation and reversible terminator sequencing) but the technical specifications of these instruments are beyond the scope of this thesis. In principle, the concept of NGS technology is similar to the automated Sanger sequencers: fluorescently labelled deoxyribonucleotides are incorporated in the DNA during multiple cycles of synthesis. The critical difference is that NGS extends this process to millions of DNA fragments in a massive way in parallel, instead of simply sequencing a single DNA fragment (Illumina, 2015). As the Illumina platform was used for all the NGS in this thesis, its workflow is further clarified. It includes four main steps:

1. Library preparation – The sequencing library is prepared by fragmenting the DNA and the ligation of specific adaptor oligonucleotides to fragments of the DNA to be sequenced.
2. Cluster generation – The library is loaded into a flow cell where the DNA fragments are captured on a surface where complementary oligos bind to the library adapters and each fragment is amplified through bridge amplification.
3. Sequencing – The Illumina platform uses sequencing by synthesis (SBS) chemistry, a proprietary method that detects single bases (fluorescence) as they are incorporated into the DNA template. Millions of clusters are sequenced in a parallel process. Sequenced fragments are then ordered by indices introduced during the library preparation and aligned to the reference strain.
4. Data analysis – The new sequence reads are converted into an electronic file ready for further analysis.

1.10.2 Main applications of NGS

NGS platforms enable a wide variety of applications, including both DNA and RNA samples. The most common applications can be divided into three main categories, each with its own sub-groups: genomics, transcriptomics and epigenomics (Illumina, 2015).

- *Genomics* – This includes whole genome sequencing (WGS), exome sequencing, *de novo* sequencing and targeted sequencing. The rapid drop in sequencing cost and the advent of new technologies can now allow the

sequencing of the entire genome of humans, plants and bacteria. However, exome sequencing is a cost-effective alternative to WGS as the exome represents only 1.5-2% of the human DNA but it encodes for the majority of proteins. It is also an example of targeted sequencing, forms of which can focus on specific genes allowing a higher sequencing coverage to identify rare variants. *De novo* sequencing refers to the identification of a new genome for which there is no reference sequence available for alignment.

- *Transcriptomics* – Transcriptome sequencing studies gene expression by sequencing the RNA (total RNA, mRNA and targeted RNA).

- *Epigenomics* – This is the study of changes in gene activity due to various mechanisms, including DNA methylation, small RNA-mediated regulation and DNA-protein interaction. In particular, ChIP Sequencing combines chromatin immunoprecipitation (ChIP) assays and NGS to analyse protein-DNA and protein-RNA interactions to better understand gene expression.

In this thesis, all *M. tuberculosis* samples (in particular, INH-R clinical strains and PAS-R mutants) have been whole genome sequenced unless stated otherwise.

1.10.3 Sequencing file formats

Different format files are used in the analysis of NGS data. FASTQ is generally the initial file format when receiving WGS results from the sequencing provider (or in-house analysis). All other formats have to be created using the initial FASTQ files with the numerous software tools that are described in this chapter. The most common files are:

- *FASTQ file*: Initially used for Sanger capillary sequencing, it has gradually become the most common format for sequencing reading data. It is constituted by four lines: the first line (starting with an @) is a read identifier, the second is the DNA sequence, the third another identifier (but starting with a +) and the fourth is the quality score symbol for each base in the read.

- *SAM/BAM files*: A SAM (Sequence Alignment Map) file represents a generic format for storing large nucleotide sequence alignments. It is created to obtain a well-defined interface between alignment and downstream analyses, including variant detection, genotyping and assembly. A BAM file is the binary version (or simply a lighter version) of a SAM file and it is more easily handled by common computers.

- *VCF file*: This Variant Call Format was developed by the 1000 Genomes project to encode single nucleotide polymorphisms (SNPs) and other structural genetic variants. Its flexibility and user extensibility allow representation of a wide variety of genomic variation with respect to a single reference sequence. It appears as a text file format (usually stored in a compressed manner) which contains meta-information lines, a header line, and then data lines, each containing information about a position in the genome.
- Many other files are available (i.e. *SRA*, *GFF*, *GTF*, *BED*) but they are less commonly used.

Once sequencing was completed at the Genomic Services and Development Unit (GSDU, Public Health England in London), all output results were downloadable as FASTQ files using a FTP (File Transfer Protocol) server.

1.10.4 Quality control

In the Illumina platform, the information containing the quality of sequencing is included in the FASTQ files and encoded as ASCII characters (American Standard Code for Information Interchange). Quality scores measure the probability that a base is called incorrectly. With NGS technology, each base in a read is assigned a quality score similar to that originally developed for Sanger sequencing experiments (Ewing, Hillier, Wendl, & Green, 1998) (Ewing & Green, 1998). The quality score of a given base (Q) is defined by the equation $\mathbf{Q = -10log_{10}(e)}$, where **e** is the estimated probability of the base call being wrong. Thus, a higher quality score implies that a base call is more reliable and less likely to be incorrect. During an Illumina sequencing run, a quality score is assigned to each base using various criteria (i.e. intensity profiles and signal-to-noise ratio) and a complex statistical calibration process where a quality table is derived from previous empirical data (Table 6). Another complex equation is also used to calculate the Shannon entropy, a simple quantitative measure of uncertainty to quantify sequence variability which is dependent on the number of variables (i.e., nucleotides in case of DNA) (Schmitt & Herzel, 1997). Read length, coverage and mapping coverage to the reference strain are other quality control measures. In general, it is important to have a good quality score for as many reads as possible. The read length is the number of bases sequenced. It is another measure of the coverage achieved as this represents the depth of coverage across the genome in comparison with H37Rv reference strain. Finally, base frequencies, comparing the content of Guanine (G) and Cytosine (C)

versus Adenine (A) and Thymine (T), are also available as a quality indicator as *M. tuberculosis* is known for having a higher GC content in its DNA.

Quality score	Probability of incorrect base call	Inferred base call accuracy
Q10	1:10	90%
Q20	1:100	99%
Q30	1:1000	99.9%

TABLE 6: RELATIONSHIP BETWEEN QUALITY SCORE AND BASE CALL ACCURACY, MODIFIED FROM ILLUMINA WEBSITE (ILLUMINA, 2014).

This quality table links Q-scores with error probabilities, showing the possible different quality scores. As an example, a quality score of 20 represents an error rate of 1 in 100, with a corresponding call accuracy of 99% (Illumina, 2014).

1.10.5 Analysis using Linux

Linux is a popular operating system used for the analysis of sequencing data as bioinformaticians have been able to create their own codes to compensate for the initial absence of commercially available software products. The common software packages that have been used for the sequencing analysis in this thesis are explained below.

- *Burrows-Wheeler Aligner* (BWA) (H. Li & Durbin, 2009) is a software package for mapping low-divergent sequences against a large reference genome. It consists of three algorithms: BWA-backtrack, BWA-SW and BWA-MEM. The first algorithm is designed for Illumina sequence reads up to 100bp, whilst the other two are used for longer sequences ranged from 70bp to 1Mbp.
- *SAM Tools* (Sequence Alignment/Map) (H. Li et al., 2009) provide various utilities for manipulating alignments in the SAM file format, such as indexing, variant caller and alignment viewer, and thus provides universal tools for processing read alignments.
- *Randomized Axelerated Maximum Likelihood* (RAxML) (Stamatakis, 2014b) is a popular program for phylogenetic analysis and it is able to accommodate large datasets.
- *Phylogenetic Analysis by Maximum Likelihood* (PAML) (Yang, 2007) is a package of programs for phylogenetic analyses and provide source codes for Linux. It is not good for tree making but it useful to estimate parameters and branch SNP counts.

In contrast with other operating systems (such as Microsoft Windows or MacOS) where a graphical interface and relatively intuitive programs are used, the Linux system requires significant knowledge of the command line. This represents its main limitation for people with little or no experience in bio-informatics but command-line interfaces are often preferred by more advanced computer users, as they provide a more powerful way to control a program.

1.10.6 Analysis using other software tools

An incredible amount of software tools for the analysis of next generation sequencing data are freely available on the internet. Most of them are Linux based, whilst others have a graphical interface or can be used directly online. The Sanger Institute has historically been a pioneer in genetic analysis and it offers a range of more than eighty

different online tools on its website (www.sanger.ac.uk). More recently, the Galaxy project (www.galaxyproject.org) is led by Pennsylvania State and John Hopkins universities and it is promising data intensive biology for everyone, with a similar number of various online programs. Additionally, the Broad Institute (www.broadinstitute.org), a collaboration between MIT and Harvard, is also offering different software products. These are only the major providers of such software and these sources are not inclusive of organism specific tools available elsewhere. Hundreds of different software products for the analysis of WGS data are freely available but their high number and various functionalities often leave the inexperienced user confused when faced with such diversity of choice. It is beyond the scope of this thesis to describe all the different tools available and only the most popular software packages used for sequencing analysis will be described below.

- *Artemis* – This is a free genome browser and annotation tool that allows visualisation of next generation sequencing data and the results of analyses within the context of the sequence. Artemis is part of the software tools offered by the Wellcome Trust, it is written in Java and available for UNIX, Macintosh and Windows systems. It can read various sequence database entries or raw format (Rutherford et al., 2000). A lightweight web-based version is also available. Artemis was extensively used in this thesis to analyse the position of SNPs and to identify the genes affected by SNPs and their potential function (Figure 3). The *M. tuberculosis* H37Rv complete genome (NCBI Reference Sequence: NC_000962.3) was used as the reference strain in all analyses.
- *TubercuList* – A freely available database and website (www.tuberculist.epfl.ch) created at the Ecole Polytechnique Fédérale de Lausanne and in collaboration with the Swiss Institute of Bioinformatics (Lew, Kapopoulou, Jones, & Cole, 2011). It presents genome-derived information about H37Rv, including gene function, expression profiles, drug development criteria and drug resistance information, in addition to direct access to PubMed articles where available. It has been used in this thesis to analyse and understand the potential significance of mutations and deletions and the role of the genes thus involved.
- *Circos* – A software package that allows the visualization of genomic data in a circular layout. It is particularly useful to demonstrate variations among different strains and its circular ideogram layout facilitates the display of such differences and/or relationships between genomes. An interactive online version is available

and it is licenced under GPL (General Public Licence, free software) (Krzywinski, Schein, & Birol, 2009).

It is worth mentioning for completeness that other commercial software packages are available on the market, such as CLC Genomics Workbench (by Qiagen Bioinformatics, Aarhus, Denmark) and BioBumerics (by AppliedMath, Sint-Martens-Latem, Belgium). They can potentially resolve the issue of using multiple tools, promising end-to-end NGS data analysis solutions, and they are compatible with different operating systems (including Microsoft, MacOS and Linux). However, due to their high commercial cost, they have not been used for any analyses within this thesis.

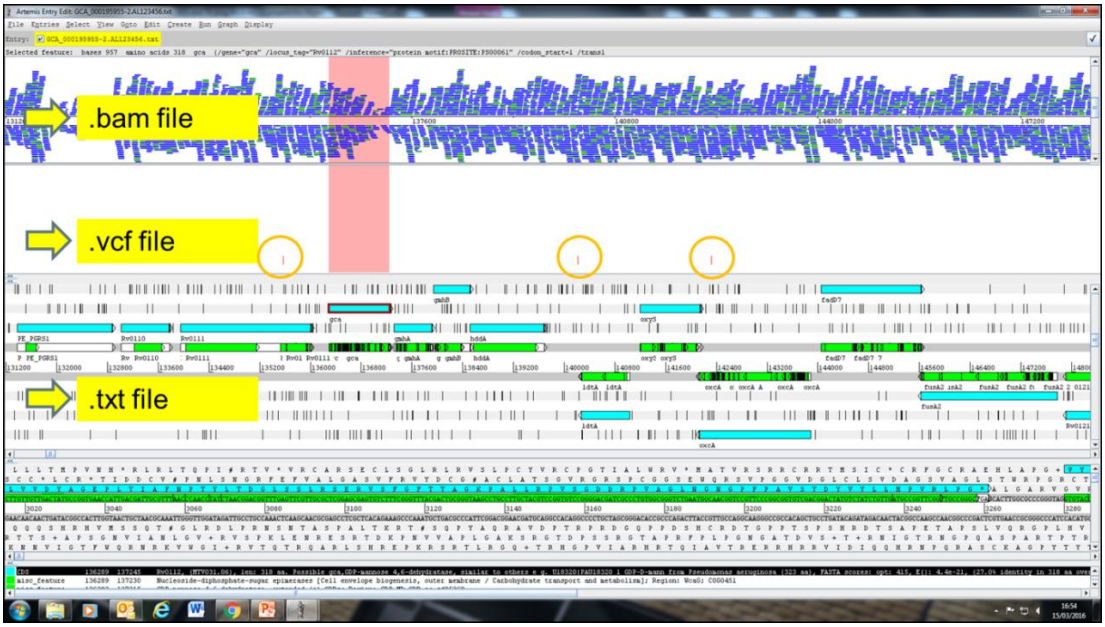


FIGURE 3: ARTEMIS SOFTWARE SCREENSHOT.

Artemis is used for the visualization of sequencing data and the analysis of SNPs. The reference strain is uploaded as a *.txt* file and it contains all the information regarding mycobacterial genes, their position and possible function. Functional information is visible at the bottom of the screenshot. The reference strain (NCBI Reference Sequence: NC_000962.3) is freely available from the National Center for Biotechnology Information website (https://www.ncbi.nlm.nih.gov/). The clinical strains for analysis were uploaded as a *.bam* file at the top of the screenshot, whilst a *.vcf* file was used for identifying SNPs, their position and genes involved.

1.10.7 Online software tools specific for *Mycobacterium tuberculosis*

The use of NGS, despite its great potential, may be hampered by the complexity of data and its analysis. TB is a major infection in low and middle-income countries. These areas typically have a chronic shortage of qualified healthcare workers and are resource limited. As such, there is a need for simpler, free of charge, software tools which are able to analyse data automatically from the initial FASTQ files and provide the basic functionalities as an all-in-one package. At the time of writing this thesis, four web-based software tools that satisfy the above criteria have been identified and their different functionalities are described below.

- *Mykrobe predictor* - A software package developed at the Wellcome Trust Centre for Human Genetics, University of Oxford. It is able to identify species and resistance profiles (www.mykrobe.com/products/predictor/#tb) (Bradley et al., 2015). This software presents a pleasant graphic interface in which raw sequence data in FASTQ format can be uploaded to generate a fast report (within 3 minutes), easily interpretable by clinicians. Files can be uploaded singularly (upload of paired end reads and multiple reads is not allowed) and resistance panels, virulence elements, phylogenetic lineages can be detected. The main limitation is that it does not provide information about sample quality, phylogenetic tree analysis and linkage networks (Table 7). Resistance to first line anti-TB drugs can be detected, apart from pyrazinamide. It can also detect resistance to some, but not all, second line drugs such as fluoroquinolones and aminoglycosides (Table 8). This software was used for the additional confirmation of identification at species level and isoniazid resistance for all FASTQ files of all clinical strains sequenced.
- *TB Profiler* – This was developed at the London School of Hygiene & Tropical Medicine and it processes raw sequence data identifying strain type and known drug resistance markers (http://tbdr.lshtm.ac.uk) (Coll et al., 2015). The web-based tool aligns data to a modified version of the H37Rv reference genome, which consists of the genes and flanking regional sequences containing 1,325 mutations from a curated whole genome drug resistance library. Sequence data can be submitted as single end or paired end reads (FASTQ file) and processing time is under 10 minutes per sample plus queuing time. The displayed output provides information about resistance to 11 TB drugs (Table 8) and lineages

specific mutations. However, phylogenetic tree and linkage networks are not available (Table 7). This software was used to double check discrepant results (and the associated mutations) between phenotypic and genotypic methods.

- *Phylo-Resistance Search Engine (PhyResSE)* – This is designed to enable nonspecialized users to extract phylogenetic and resistance information from WGS data (http://phyresse.org) (Feuerriegel et al., 2015). Single end or paired end reads can be freely submitted and multiple file selection for upload is supported up to a maximum upload of 2.1 GB. A validation process runs to reject improperly formatted data files. After processing, a quality report is produced (including mapping with the reference strain) and variants are called and provided in VCF format and as a HTML table carrying additional information about amino acid changes, genes involved and any association with genotype or resistance, including for the majority of first and second line drugs (Table 8). Unfortunately, the comparison is only versus the reference strain (NC_000962.3) and the software does not allow more complex comparisons among multiple samples. It was used to produce quality reports of the sequencing data.

- *Total Genotyping Solution for Mycobacterium tuberculosis (TGS-TB)* – An all-in-one web-based tool developed by Sekizuka and colleagues (https://gph.niid.go.jp/tgs-tb/) (Sekizuka et al., 2015). Information about the number of trimmed map reads and coverage region depth is provided. The resistance target list permits detection of genetic alterations for the majority of first and second line drugs (Table 8). In addition, up to 5 paired-ends FASTQ files can be uploaded simultaneously and the comparison of samples with a phylogenetic tree enable outbreak investigation. The phylogenetic tree can be downloaded as a nexus file and then processed using the software PopArt (http://popart.otago.ac.nz). The comparison between different strains can also be downloaded as a tabular format and this allow the identification of SNPs and their exact location (they can be filtered by searching under the *original*/*additional* SNPs column). Other software tools, such as Artemis, are then needed to identify the gene involved. TGS-TB was used in this thesis for identifying SNPs when comparing a limited number of samples.

These software tools hold great potential for future diagnostic use, in particular for the rapid analysis of susceptibility profiles by scientists with limited knowledge of the command line in Linux. However, they can only analyse few samples at any one time

and have limited value when investigating epidemiological links among different strains and in-depth analysis of SNPs.

1.10.8 Limitations of online tools

WGS holds great potential for rapid diagnosis, epidemiological analysis and research applications but its use may be hampered by the complexity of data and its analysis. This section has illustrated the multitude of software tools available, from Linux-based to online programs offering rapid identification of drug resistance using a FASTQ file. Online tools may be particularly appealing for diagnostic purposes but they have limited capabilities in terms of epidemiological analysis and number of strains uploaded at one time. Accreditation (ISO 15189 and others) is another possible problem. None of the mentioned software tools is licensed for clinical/diagnostic use and alternative programs and knowledge of the Linux command line are still essential for research purposes and for an in-depth analysis of mutations (deletions, insertions and SNPs).

A recent report from the European Committee on Antimicrobial Susceptibility Testing (EUCAST) subcommittee on the role of WGS in antimicrobial susceptibility testing (AST) of bacteria, concluded that available published evidence is currently insufficient and it does not support the use of WGS inferred susceptibility to guide clinical decision making (Ellington et al., 2017). The report has also highlighted two main issues: lack of international QC metrics and of a single public database. Minimum quality standards should exist and they should be comparable across different laboratories, with internationally agreed QC metrics to facilitate early harmonization of analytical processes and interpretive criteria. Only data sets that pass these QC metrics should be used in AST predictions and further analysis. This further explains and supports the focus on quality control that has been emphasised in this section.

The EUCAST report also highlighted some specific issues regarding *M. tuberculosis*, in particular the current discrepancies between genotype and phenotype and the need for large, openly accessible datasets. The latter is needed, to clarify the role of rare resistance mechanisms and the level of resistance conferred by different mutations. Ideally, there should be one single public database, regularly updated and strictly curated using agreed standards for the inclusion of resistance mutations. Additionally, from a clinical perspective WGS does not seem to resolve some of the current problems of traditional susceptibility testing, in particular with pyrazinamide (Chang et al., 2011). The proportion method measures resistance based on the critical

proportion of resistant colonies (which is set at 10% for pyrazinamide and at 1% for the other antituberculous drugs) (see Section 1.7.5). However, genotypic methods not only require the knowledge of all specific mutations conferring resistance but there is evidence that they can lead to systematic false-negative results for strains with low-level hetero-resistance (Folkvardsen, Svensson, et al., 2013) (Folkvardsen, Thomsen, et al., 2013).

WGS clearly has a role in public health interventions and in the detection of outbreaks and transmission events due to its higher resolution compared to MIRU-VNTR typing, IS6110 RFLP typing and spoligotyping methods (Gardy et al., 2011) (Török et al., 2013) and this will be further demonstrated in Chapter 4. Not all the online tools considered will be useful for this, with only TGS-TB and PhyResSE able to create phylogenetic trees and with TGS-TB performing more in-depth network analysis (including SNPs detection) but only for a limited number of samples. Several challenges remain before WGS can be routinely used in outbreak investigation and clinical practice, with the need of a robust bioinformatics infrastructure being the main limitation.

Unfortunately, many practical questions (how to deal with phenotype/genotype discrepant results, mutations conferring low level of resistance and hetero-resistance) remain unanswered and, based on these considerations, WGS is unlikely to completely replace phenotypic AST for *M. tuberculosis* in the near future. Microbiologists should be aware of these limitations but this should not prevent the use of WGS.

Functionality	Mykrobe predictor	TB Profiler	TGS-TB	PhyResSE
Resistance	✓	✓	✓	✓
Lineage	✓	✓	✓	✓
Phylogenetic tree	×	×	✓	✓
Network analysis	×	×	✓	×
Quality Control	×	×	✓	✓
Time: per sample of 100 Mb	3 min	5 min	10 min	15 min

TABLE 7: **COMPARISON OF THE FUNCTIONALITIES OFFERED BY WEB BASED SOFTWARE TOOLS FOR THE ANALYSIS OF WHOLE GENOME SEQUENCING DATA OF *MYCOBACTERIUM TUBERCULOSIS*.**

The various software tools offer different functionalities, with ✓ and × indicating available and not available, respectively. Mykrobe predictor and TB Profiler only allow the detection of resistance and lineage, whilst TGS-TB and PhyResSE have an integrated quality control system and they also allow the creation of phylogenetic trees. A Microsoft Windows computer with processor Intel CORE i7 and a broadband internet connection (11.9 Mbps download and 1.03 Mbps upload) were used to calculate the time per sample (100Mb).

Drug	Mykrobe Predictor	TB Profiler	TGS-TB	Phy ResSE
Isoniazid (INH)	✓	✓	✓	✓
Rifampicin (RIF)	✓	✓	✓	✓
Pyrazinamide (PZA)	×	✓	✓	✓
Ethambutol (EMB)	✓	✓	✓	✓
Streptomycin (SM)	✓	✓	✓	✓
Fluoroquinolones (FQs)	✓	✓	✓	✓
Kanamycin (KAN)	✓	✓	✓	✓
Amikacin (AMK)	✓	✓	✓	✓
Capreomycin (CAP)	✓	✓	✓	✓
Ethionamide (ETH)	×	✓	✓	✓
p-Amino salicylic acid (PAS)	×	✓	×	×
Cycloserine (CS)	×	×	×	×

TABLE 8: COMPARISON OF FIRST AND SECOND LINE ANTI-TUBERCULOUS DRUG RESISTANCE DETECTION OFFERED BY WEB BASED SOFTWARE TOOLS.

Detection of resistance to first and second line antituberculous drugs varies depending on the software tools used, with ✓ and × indicating available and not available, respectively. TB Profiler offers the widest choice of drugs whilst Mykrobe predictor is the most limited.

1.11 Aims of this thesis

The aim of this thesis is to address the challenge of drug resistant TB using a multi-strategy approach.

The primary objective is to demonstrate the utility of whole genome sequencing (WGS) in tackling this and next generation sequencing technologies will be adopted to better understand the pathogenicity of *M. tuberculosis* and the discovery of new potential drug targets.

1. In Chapter 3, we will try to create laboratory mutants from the reference strain H37Rv to match the same mutations of selected clinical strains with the aim to discover compensatory mutations and new potential drug targets.
2. In Chapter 4, we will analyse in more details those selected clinical strains from a known outbreak. We will test the hypothesis that this outbreak was the result of a particular *M. tuberculosis* strain with a selective advantage due to specific mutations in its genome. Fitness assays, mutation rate experiments and WGS will also be performed to complete the picture.

The project will then address the challenge of drug resistance from a different perspective, with the secondary objective of demonstrating the activity of various antituberculous agents. Different compounds will be tested to reflect the multi-strategy approach:

1. Para-aminosalicylic acid (PAS) analogues (Department of Chemistry, University of Hull) will be tested in Chapter 5. The hypothesis is that a modification of the original molecule could lead to an active compound against *M. tuberculosis*. We will also create laboratory mutants and apply next generation sequencing technologies to understand the mechanism of action.
2. Ertapenem and faropenem will be tested in Chapter 6, with the main aim to test their *in vitro* activity against different clinical strains of *M. tuberculosis* and to understand their potential clinical utility.
3. Gold nanoparticles (Midatech Pharma, UK) will be tested in Chapter 7 to understand if they have any inhibitory activity on *M. tuberculosis* cultures and/or if they are able to penetrate inside the bacterial cell.

The final discussion in Chapter 8 will critically analyse our findings and summarize the results of this thesis.

Chapter 2: Materials and methods

2.1 Health and safety considerations

M. tuberculosis is a Hazard Group 3 Pathogen and it needs to be handled in a Containment Level 3 (CL3) laboratory. All procedures were carried out following risk assessments and the completion of Control of Substances Hazardous to Health (COSHH) forms, in line with Health and Safety Executive regulations.

2.2 Culture methods

The reference strain H37Rv was used in all experiments unless stated otherwise. It was cultured from a frozen stock ordered directly from the supplier (PHE, National Collection of Type Cultures - NCTC) and it was sub-cultured from frozen aliquots a maximum of 5 times to avoid the natural selection of spontaneous mutations. Initial cultures were inoculated using 100 µL and gently moved to spread the inoculum all over the medium surface (or gently shaken in case of liquid media). 100 µL H37Rv stock was also used when sub-culturing into liquid media, whilst a single discrete colony (taken with a plastic disposable loop) was used when sub-cultures were prepared from solid media. For liquid cultures, the McFarland standards (Oxoid, Hampshire, UK) were used as a reference to adjust the turbidity of bacterial suspensions and ensure the appropriate amount of starting inoculum was used.

In general, all samples were incubated at 37°C for 7-14 days or until culture positive. Cultures of *M. tuberculosis* up to 2 weeks old are considered to be in the log phase of growth (Penuelas-Urquides et al., 2013) and in order to keep cultures growing optimally they should be sub-cultured at least once every 2 weeks.

Quality control measures were used throughout all the experiments described in this thesis. Plastic tips and loops were disposable and frequently changed to prevent cross-contamination and non-sterile solutions were autoclaved whenever possible or filtered through 0.2 µm pore size filters (Sigma-Aldrich, Dorset, UK) when higher temperatures were contraindicated. All inocula were also plated on Columbia blood agar (Oxoid, Hampshire, UK) and incubated for 48h at 37°C to exclude contamination with other rapidly growing bacteria. Additionally, positive and negative growth controls (H37Rv and a sterility control, respectively) were included in all experiments.

2.2.1 7H9

Middlebrook 7H9 (Becton-Dickinson, New Jersey, USA) was used as the main liquid media. A total of 4.7 g of 7H9 powder was dissolved in 900 ml of distilled water with the addition of 2 ml of Tween 80 (Sigma-Aldrich). Tween 80 is a nonionic surfactant and emulsifier derived from polyethoxylated sorbitan and oleic acid and it prevents the formation of clumps. This mixture was autoclaved at 121°C for 15 minutes, and, once cooled, 100 ml of Albumin Dextrose Catalase (ADC) supplement (Becton-Dickinson) was added. This complete liquid media was aliquoted into 10 ml tubes ready to be inoculated with H37Rv and/or other *M. tuberculosis* strains.

2.2.2 7H10

Middlebrook 7H10 (Becton-Dickinson) was used as the main solid media for the susceptibility testing of new compounds. A total of 3.8 g was dissolved in 180 ml of distilled water with the addition of 0.5 ml of glycerol (Sigma-Aldrich) and autoclaved at 121°C for 15 minutes. Once cooled, 20 ml of Oleic Albumin Dextrose Catalase (OADC) supplement (Becton-Dickinson) was added. OADC is heat sensitive and the broth needs to be cooled to 50 °C before it can be added. Complete 7H10 agar was poured into sterile Petri dishes. Six well culture plates were also used to optimize space and each well should contain 5 ml. Any addition of antibiotics should happen at this stage before the plates solidify.

2.2.3 Löwenstein–Jensen medium

Löwenstein–Jensen (LJ) medium was also used as a solid media, in particular for the subculture of clinical strains prior to genomic DNA extraction. It contains malachite green, glycerol, asparagine and various minerals and it is a common medium for mycobacterial growth (R. S. Martin, Sumarah, & Robart, 1975). It was also chosen for practical reasons as solid slopes containing 7 ml of medium with the addition of pyruvate were bought ready to be inoculated (E&O Laboratories, Bonnybridge, UK).

2.2.4 Other culture media

Kirchner's medium (E&O Laboratories, Scotland, UK) was sometimes used as an alternative liquid medium for sub-culturing H37Rv and other strains. This is a liquid medium for the enrichment and isolation of *Mycobacteria spp* from clinical specimens, particularly when the organisms may be present only in small numbers (Mitchison, Allen, & Manickavasagar, 1983). This was bought in 10 ml vials ready to be inoculated and this was its main advantage compared with the other media. However,

it does not contain Tween 80 and *M. tuberculosis* will grow forming clumps. This requires the additional step of passing the bacteria through an insulin needle (28 gauge) to separate the clumps and obtain single colonies before further experiments. Alternatively, Tween 80 could be added to each tube but the small amount needed (20 µL) makes it difficult and it increases the risk of contamination.

2.2.5 Freezing strains

Stocks of H37Rv, clinical strains and laboratory mutants were frozen to avoid sub-culturing them multiple times and to prevent the selection of spontaneous mutations. *M. tuberculosis* cultures were grown in liquid media (7H9 only) for 14 days. Cultures where then centrifuged and the bacteria were concentrated by decanting the supernatant and re-suspended in 2mL of fresh 7H9 broth. A mixture of 250 µL of this new bacterial broth and 250 µL of previously sterilized glycerol (Sigma-Aldrich) (500 µL in total) was transferred into 2 mL cryovials (Nunc, Roskilde, Denmark), gently inverted several times and placed in the -80°C freezer located in the CL3 research laboratory in 100-place boxes (Nalgene, Rochester, USA) for long term storage. When sub-culturing was required, culture methods as described in section 2.2 were followed after thawing the samples in a class 1 cabinet.

2.2.6 Miles and Misra colony count

The Miles and Misra method (Miles, Misra, & Irwin, 1938) was used to calculate the viable count of an inoculum by determining the number of colony forming units from a small drop of broth. This was done by counting the number of colonies which grew after 14 days on a 7H10 plate from the drop and estimating the number per mL. *M. tuberculosis* was cultured in 7H9 broth and ten-fold serial dilutions were made in 1mL volumes (i.e. adding 100µL of neat suspension into 900µL of sterile 7H9 media). Dilutions from 10^{-1} to 10^{-6} were used. Sterile 7H10 plates were also previously prepared. Each dilution was vortex-mixed and then 3 x 20µL plated onto one half of the 7H10 agar plate. Once dried in the class 1 cabinet, plates were sealed in plastic bags and incubated at 37°C for 14 days. After incubation, the viable count was calculated by using the following equation:

CFU/mL = mean no. colonies x 50 x dilution factor for the plate counted

E.g. if the counted colonies were 33/36/32 at dilution 10^{-5}, the CFU/mL would be $1.7x10^{7}$ CFU/mL (33.7 x 50 x 10^{5} or 168.5 x 10^{5} or 1.685 x 10^{7}, with approximation to 1.7).

2.3 TB susceptibility testing

2.3.1 Spot culture method

The spot culture method (Evangelopoulos & Bhakta, 2010) was used as a qualitative rapid screening method to test the susceptibility of new compounds. It is a modified version of the agar dilution assay but it required a lower volume of agar (six-well plates with 5 ml of agar each) and a smaller inoculum. It can also be used to determine the minimum inhibitory concentration (MIC) of antibiotics if combined with serial dilutions. In more detail:

- From the stock solution of a new compound, serial double dilutions were prepared by adding distilled water. If used as a rapid screening test, only the highest dilution of each new compound was used.
- 5 μL of each dilution was added to 5 ml agar in a six-well plate.
- 7H10 medium was added to each well, gently mixed and let to solidify.
- In order to obtain a standard inoculum of around 500-1000 cells, 1 McFarland ($3x10^8$ CFU/mL) broth was diluted 1:1000 and 10 μL of the final dilution was inoculated in each well containing the agar with antibiotics.
- Positive and negative controls were included.
- All plates were incubated for 2 weeks and the possible activity of a new compound was based on the growth/no growth of *M. tuberculosis* on the plate.
- In case of MIC determination, the MIC was determined as the lowest concentration that showed no visible growth.

2.3.2 Resazurin method

It is a simple and inexpensive method to test for drug resistance, using liquid media in 96-well microtiter assay plates (Palomino et al., 2002). It is based on the colour change of resazurin (from blue to pink) added to the liquid culture after one week. Resazurin (7-Hydroxy-3H-phenoxazin-3-one-10-oxide) is a blue dye and it is used as an oxidation-reduction indicator in cell viability assays for bacteria (presence of metabolic activity = growth). In case of no growth (active compound), there will be no change in colour. In more detail:

- 100 μL of broth 7H9 were added to a 96-well microtiter plate.

- 100 µL of the compound to be tested were added to well number 1 and then serial dilutions were made depending on the original concentration of the stock.
- Positive, negative and compound controls were always included.
- After dilution was completed, 100 µL of TB inoculum was added to each well and positive control.
- At the end each well will contained 200 µL in total with a TB inoculum of 1.5 x 10^5.
- After 1 week of incubation 10 µL of resazurin (original stock from 270 mg tablet in 40 mL of sterile distilled water) were added and the sample incubated for 24 hours.
- The next day, any change in colour from blue to pink indicated the presence of metabolic activity (non-active compound). If no colour change was present, this indicated the inhibition of mycobacterial growth (active compound).

2.3.3 BACTEC/MIGIT method

The BACTEC/MGIT method was used for the determination of the MIC of a test drug as previously described (Kruuner et al., 2006). Briefly, a range of concentrations in serial dilution of the test drug were added to MGIT tubes prior to inoculation with *M. tuberculosis* (H37Rv or other relevant clinical strains). A growth control (GC) tube with 1/100 the inoculum of the test isolate was included with each serial dilution to serve as a comparison for growth. In principle, there should be no growth in the MGIT tubes if the drug concentration used is active against *M. tuberculosis*. The GC tube and any other drug dilution at which *M. tuberculosis* is resistant should exhibit growth and the MGIT instrument should detect increasing fluorescence. The test was considered finished when the GC tube reached 400 growth units (GU). In the tubes containing test drugs, any growth above 100 GU was interpreted as resistance to the test compound, whilst any concentration associated with less than 100 GU was considered effective and the bacteria deemed sensitive. The MIC was defined as the lowest drug concentration where the GU was less than 100. The MGIT instrument is able to automatically report the results for common first and second line antituberculous drugs. However, the EpiceCentre software and the TB eXiST (TB eXtended Susceptibility Testing) module (Springer, Lucke, Calligaris-Maibach, Ritter, & Bottger, 2009) must be used for the interpretation of results for drugs outside the standard sets.

2.3.4 Drug supply and stock solutions

A summary of all compounds tested in this thesis is provided in Table 9. All drugs in the powder form where initially dissolved and diluted in sterile water. As different compounds were provided in different weights, 500 µL of sterile water was generally added to create the stock solution (to avoid unnecessary manipulation and risk of contamination) and stored at -80°C. The following formula was then used to calculate the concentration of the stock solution:

Weight (mg) = Volume (mL) x Concentration (µg/mL)/ assay potency (µg/mg)

The assay potency was generally indicated in the specification sheet when purchasing the drug and equal to 99%. Additionally, when performing dilutions, the following formula was used:

Volume Final x Concentration Final = Volume Initial x Concentration Initial

All stock solutions and dilutions were always performed in a class 2 cabinet and a drop of the new solution was plated on a Columbia blood agar plate to exclude contamination with rapidly growing bacteria.

List of compounds tested	Source	Notes
Para-aminosalicylic acid (PAS), DE97, DE98, DE99, DE471, DE160A, DE210A, DE267(2), DE540, DE536(2), DE541, 1H2NA, 2H1NA, 3H2NA, 3H2NH, PAMOIC ACID, AD25a	Research, University of Hull (UK)	PAS analogues Chapter 5
Ertapenem and Faropenem +/- Amoxicillin-clavulanate	Commercially available, Sigma-Aldrich	Chapter 6
Gold nanoparticles (NP-CML1 and NP-control)	Research, Midatech Pharma (UK)	Chapter 7

TABLE 9: LIST OF ALL COMPOUNDS TESTED IN THIS THESIS AND THEIR SOURCE (RESEARCH AND COMMERCIALLY AVAILABLE).

Para-aminosalicylic acid (PAS) and its analogues were kindly provided as dried powder by Dr Andrew Boa (Department of Chemistry, University of Hull, UK). Commercial antituberculous drugs (including ertapenem and faropenem) were purchased from Sigma-Aldrich, whilst the gold nanoparticles were provided in liquid solution by Midatech Pharma (UK).

2.4 Drug resistant mutants

Two main types of drug resistant mutants were selected:

- INH-resistant mutants (further discussed in Chapter 5)
- PAS-resistant mutants (further discussed in Chapter 7)

2.4.1 Isolate selection

In principle, *M. tuberculosis* H37Rv was grown for 14 days in culture broth to reach high concentration and log phase. Following the incubation period, the bacteria were concentrated and inoculated onto 7H10 plates containing different concentrations of antibiotics (2X, 4X and 8X minimum inhibitory of INH or PAS) and mutants were naturally selected. In more details:

- *M. tuberculosis* was sub-cultured into 10mL aliquots of sterile 7H9 broth and incubated for 14 days. These cultures were considered to be in the log phase and they generally reached a colony count of around 0.5-1 McFarland (150-300X10^6/mL). Kirchner's medium could be used as an alternative.
- To prevent the formation of clumps the bacteria needed to be passed through an insulin needle (28 gauges). Care was taken with this step. The needle must remain below the level of the liquid and the culture drawn up the needle slowly and the final 0.1µL left in the syringe so as to prevent generation of bubbles that may result in aerosols containing *M. tuberculosis*.
- 10mL were then transferred to 15mL plastic containers (Falcon tubes by Becton-Dickinson) using a plastic disposable pipette. These were centrifuged at 10,000 g for 10 minutes.
- One 10mL tube was used for each of the 7H10 plates already prepared (containing 2X, 4X and 8X MIC of INH). Double the inoculum (20 mL) was used for the creation of PAS-resistant mutants as PAS has a much higher genetic barrier compared to INH.
- The supernatant was discarded and the whole deposit was pipetted to inoculate one plate and a disposable spreader was used to ensure distribution on the plate surface.
- Plates without bacteria were included as a negative control for the experiment. A Columbia blood agar plate was used to check for any contamination.
- Plates were allowed to dry inside the safety cabinet prior to place them in sealable polyether bags and incubated for at least 14 days.

- After incubation, spontaneous mutants would grow on the plates and they were then selected for further testing.

2.4.2 DNA extraction of drug resistant mutants

A single colony was initially picked, sub-cultured to a LJ slope (with pyruvate) and incubated at 37°C. After 2-3 weeks of incubation (depending on the growth rate), a full loop (plastic and disposable) was transferred to a 2 mL microcentrifuge tube containing 400 µL of Tris-EDTA buffer (Sigma-Adrich) and heat-killed for 20 minutes at 100°C using an heating block. Different extraction methods were used depending on the type of genetic analysis needed. DNA from INH-R laboratory mutants was extracted using the commercial kit GenoLyse® from Hain Lifescience (Nehren, Germany) whilst the CTAB (cetyl trimethylammonium bromide) method was used for PAS-resistant mutants and INH-R clinical strains to obtain genomic DNA for next generation sequencing. Different DNA extraction methods are discussed in detail in Section 2.6.

2.4.3 Sequence analysis and confirmation of mutation type

One of the initial research plans was to create INH-R laboratory mutants from the reference strain H37Rv to match the same mutations of the INH-R clinical strains and compare them to discover compensatory mutations and new potential drug targets (Chapter 5). The mutations were chosen based on previous work from our group (R. Shorten, 2011). All INH-R clinical strains exhibited only two common mutations:

- *katG S315T* mutation
- *inhA C-T767* mutation

After 14 days of incubation, single colonies of possible INH-R laboratory mutants growing on the incubated plates were subcultured to LJ slope (with pyruvate) and DNA extraction performed as described. The line probe Genotype MTBDRplus from Hain Lifescience was used to confirmed INH resistance and type of mutation (please also refer to Chapter 3).

2.4.4 Fitness assay

Fitness assays was performed on both INH-R laboratory mutants and clinical strains as previously described (Pope, McHugh, & Gillespie, 2010). In more detail:

- Mutants and clinical strains were inoculated in Kirchner's medium (E&O Laboratories) and incubated at 37°C until 0.5 McFarland of growth was visible (average of 7-10 days).
- 0.5 mL of each broth was inoculated into a 7mL MGIT tube with growth supplement (800 µL in each MGIT tube). The tubes were incubated in the MGIT 960 system.
- Within 24 hours of these cultures signalling positive, 1:10 (dilution A) and 1: 10,000 (dilution B) dilutions were prepared.
- A 0.5 mL aliquot of each dilution was inoculated in triplicate into fresh MGIT tubes containing growth supplement and incubated in the MGIT 960 until they flag positive. The growth curve and time to positivity were printed out.
- A modified Youmans and Youmans method (Youmans & Youmans, 1949) was used to calculate the growth rate constant "K":

  K = log A (dilution A) – log B (dilution B)/T

  T is equal to the difference in the mean of three replicates of the time to positivity (TTP).
- The following equation was used to calculate the generation time "G":

  G = log 2/K.

The generation time of each strain is then compared to the generation time of the reference strain to calculate the final fitness assay value.

2.4.5 Mutation rate

The p_0 *method* (Luria & Delbruck, 1943) (Pope, O'Sullivan, & McHugh, 2008) was used to estimate the mutation rate in INH-R mutants and clinical strains. *M. tuberculosis* was used as a control and ciprofloxacin was used as antibiotic to calculate the mutation rates. In general, the antibiotics most suitable for mutation rate methods are those to which resistance arise as a result of point mutations in chromosomal genes, including the aminoglycosides, quinolones, rifampicin and pyrazinamide. In more details:

- 7H10 plates containing ciprofloxacin at a concentration equivalent to 4 X MIC (0.25 µL/mL x 4 = 1 µL/mL) were prepared in the containment level 2 laboratory and transferred into the containment level 3 laboratory for inoculation (20 plates with ciprofloxacin and 20 plates with rifampicin for each mutant and clinical strain to be tested).

- 10 mL aliquots of sterile 7H9 with ADC supplement were used to grow the INH-R mutants and clinical strains until they reached 0.5 McFarland (1.5×10^8 CFU/mL). Kirchner's medium could be used as an alternative.
- To prevent the formation of clumps the bacteria needed to be passed through an insulin needle (28 gauge) as previously described.
- The culture was diluted (1:100) by adding 0.1 mL to 9.9 mL broth giving 1.5×10^6 CFU/mL.
- Another 1:100 dilution was done by 0.1mL to 9.9mL broth giving 1.5×10^4 CFU/mL.
- 250 µL was distributed to microtubes to which 750 µL of 7H9 is added, giving a final number of 3750 bacterial cells in each tube.
- These microcultures were incubated for 14 days at 37°C prior to exposure to ciprofloxacin.
- Three sample cultures were used for colony count by the Miles and Misra method and the remaining cultures were then centrifuged at 10,000 g for 15 minutes and the supernatant discarded.
- The whole deposit was pipetted to inoculate the plates previously prepared. Plates were allowed to dry inside the safety cabinet prior to place them in sealable polyether bags and to incubate for 14 days at 37°C.
- As above, blank plates were included as a negative control for the experiment and a Columbia blood agar plate was used to check for any contamination.

Using the p_0 method, the proportion of cultures without mutants (p_0) is the zero term of the Poisson distribution given by the equation $\mathbf{p_0 = e^{-m}}$. This method should be used if the proportion of cultures without mutants is between 0.1 and 0.7, i.e., the number of mutational events per culture is between 0.3 and 2.3. The formula can be rearranged to give the number of mutational events as follows:

p_0: fraction of plates without mutants

m (number of mutants per culture) = $-\ln p_0$ (m = 0.3-2.3)

and

Mutation rate µ = m/mean CFU/mL

2.5 DNA requirements for next generation sequencing

Conventional protocols and commercially available extraction kits are generally designed for the isolation of small amounts of mycobacterial DNA to be used in polymerase chain reaction (PCR) applications for the amplification of a small fragment of the genome (De Almeida, Da Silva Carvalho, Rossetti, Costa, & De Miranda, 2013) (Kaser et al., 2010). The quality and the quantity of such extractions are not suitable for the amplification of the whole genome and new DNA requirements have emerged with the introduction of next generation sequencing (NGS).

Considering the initial prohibitive cost of NGS platforms, strict criteria were introduced to avoid unnecessary runs and only samples with confirmed high-quality DNA were processed. DNA spectrophotometric methods (e.g. NanoDrop™ by Thermo Fisher) were found to be inaccurate due to the presence of small nucleic acids and contaminating chemicals and their use is discouraged for the quantification of genomic DNA for WGS (Robin, Ludlow, LaRanger, Wright, & Shay, 2016). Quantification should be confirmed using a Qubit® fluorometer (Thermo Fisher) or other alternative methods to ensure that enough material is available for the library preparation. Concentration requirements are different depending on the type of analysis (from 1 ng/µl for viral amplicon-based NGS to a minimum of 10ng/µl for bacterial WGS), but a lower concentration may be acceptable depending on the library preparation used and sequencing provider. It is also important to assess the purity of the DNA and the absence of proteins and other organic compounds and the ratio of the absorbance at 260 and 280 nm (A260/280) is used to confirm this. Spectrophotometric analysis of DNA should show a 260/280 ratio between 1.8 and 2 indicating adequate sample purity (ThermoScientific, 2009). 1 µl of each DNA sample can also be run on an agarose gel alongside a quantitative DNA marker to confirm concentration and assess integrity.

In this study, WGS was performed using the Illumina HiSeq platform (Illumina, San Diego, USA) at the Genomic Services and Development Unit (GSDU, Public Health England, London). The minimum concentration of DNA required was 10 ng/µl and a minimum 260/280 ratio of 1.8. Running the DNA on an agarose gel was not required by the sequencing provider. The ratio 260/280 was checked using a NanoDrop™ 2000 Spectrophotometer (Thermo Fisher) and quantification performed using a Qubit® fluorometer by Thermo Fisher (dsDNA HS Assay kit – order code: Q32851) as per the manufacturer's instructions. Extracted DNA was transferred to bar-coded 96-

well plates provided by GSDU and returned for sequencing. Sequencing data (in the form of FASTQ files) were then retrieved via remote access using an FTP (File Transfer Protocol) server.

2.6 Extraction of DNA for next generation sequencing

Three different extraction methods were trialled in this thesis:

1. *GenoLyse*® (order code: 51610) from Hain Lifescience (Nehren, Germany) – This is a commercial kit and that allows the extraction of genomic bacterial DNA in only three quick steps. Its composition has not been disclosed by the manufacturer. This extraction method is generally sold in conjunction with other Hain Lifescience products, such as the Line Probe Genotype MTBDRplus described in Chapter 3. It is not marketed for the extraction of DNA for NGS but, as it was adopted in other experiments in this thesis and had better ease of use, it was decided to evaluate its potential value for WGS applications.
2. *DNeasy*® (order code: 69504) from Qiagen (Hilden, Germany) – This is a silica-based nucleic acid extraction kit with spin-columns and proteinase K, potentially eliminating the need for mechanical disruption and reducing labour costs. This was the method used by Brown et al. when extracting mycobacterial DNA directly from clinical samples (Brown et al., 2015). However, in this study the authors did use additional mechanical disruption with 0.1-mm-diameter glass beads (Becton Dickinson). In this thesis it was decided to evaluate the potential value of this kit but without the mechanical disruption step.
3. *CTAB based extraction* – This method was used as previously described (van Soolingen et al., 1991) and it contained lysozyme, proteinase K and CTAB. In the original protocol, a minimum incubation period with lysozyme for 4 hours is suggested, but it is also mentioned that a prolonged incubation may be necessary to achieve higher yields. It was decided to evaluate two variants of this protocol: one with 4 hours incubation after the addition of lysozyme (CTAB-4h) and another with 24 hours incubation (CTAB-24h).

2.6.1 GenoLyse®

All reagents were kept at 2-8°C and the protocol included three simple steps as per the manufacturer's instructions (IFU-51610-09):

1. Bacteria grown on LJ slopes were collected with a disposable loop and suspended in 100 µl of Lysis Buffer (A-LYS), vortexed and incubated for 5 minutes at 95°C.
2. After incubation, they were briefly centrifuged at 1,000 g for 1 minute and 100 µl of Neutralization Buffer (A-NB) was added prior to vortexing for 5 seconds.
3. All samples were centrifuged for 5 minutes at 10,000 g and the supernatant was transferred to a new tube. The extracted genomic DNA was then stored at -20°C until use.

A purification step with phenol:chloroform:isoamyl alcohol was also performed as an additional experiment in the attempt to increase the DNA purity and reduce protein contamination. This additional protocol is described in Section 3.3.3 (steps 7 to 15 only).

2.6.2 DNeasy®

After growth on solid media, a 10 µL loop of culture was transferred to a 2-mL microcentrifuge tube containing 400 µL of Tris-EDTA buffer (Sigma-Aldrich, order code: 93283). The tubes were then centrifuged for 10 minutes at 10,000 g and the supernatant discarded. Manufacturer's instructions were followed as per the DNeasy Blood and Tissue Handbook. Please note that all buffers are already included in the extraction kit, apart from the enzymatic lysis buffer in step 1.

1. Bacterial pellets were re-suspended in 180 µL of enzymatic lysis buffer and incubated for at least 1 hour at 37°C. The buffer was composed of 20 mM Tris-Cl (pH 8.0), 2 mM sodium EDTA, 1.2% Triton™ X100 and lysozyme at a concentration of 20mg/ml (see also Section 3.3.4 – Buffers and solutions).
2. After incubation, 25 µL of proteinase K (pre-made and included in the extraction kit) and 200 µL of Buffer AL (without ethanol – order code: 19075) were added and mixed by vortexing.
3. Samples were incubated at 95°C for 20 minutes to inactivate pathogens.
4. Ethanol (200 µL) was then added to the samples and each sample mixed by vortexing until a white and precipitate formed.
5. The precipitate was pipetted into the DNeasy Mini spin column placed in a 2-ml collection tube. Samples were then centrifuged at 8,000 g for 1 minute and the flow-though and the collection tubes were discarded.

6. Each DNeasy Mini spin column was placed into a new 2-ml collection tube and Buffer AW1 (500 µL – order code: 19081) was added. Tubes were centrifuged at 8,000 g for 1 minute and the flow-though and the collection tubes were discarded.
7. Each DNeasy Mini spin column was then placed into a new 2-ml collection tube and Buffer AW2 (500 µL – order code: 19072) was added. Tubes were centrifuged at 10,000 g for 1 minute and the flow-though and the collection tubes were discarded again. It was important to dry the membrane of the DNeasy Mini spin columns as ethanol residues could interfere with subsequent reactions.
8. The columns were placed into a clean 2-mL microcentrifuge tube and Buffer AE (200 µL – order code: 19077) was added directly onto the DNeasy membrane. Samples were incubated at room temperature for 1 minute and then centrifuged for 1 minute at 8000 g to elute. Elution with 100 µL (instead of 200 µL) would increase the final DNA concentration.

The extracted genomic DNA was then stored at -20°C until use.

2.6.3 CTAB based extraction

After growth on LJ slopes, a full loop of culture was transferred to a 2-mL microcentrifuge tube containing 400 µL of Tris-EDTA buffer (Sigma-Aldrich). Tubes were centrifuged in the microcentrifuge for 5 seconds to ensure that all organisms were at the bottom of the tube. Vortexing was not performed at any stage of the extraction as this can cause DNA shearing.

1. All tubes were placed in a head block at 95°C for 20 minutes to heat-kill *M. tuberculosis*.
2. Tubes were pulsed down as above, the supernatant removed and 300 µL of lysis buffer for mycobacteria (see also Section 3.3.4 – Buffers and solutions) were added to each tube.
3. Lysozyme (Sigma-Aldrich, order code: L3790) at a concentration of 10 mg/mL was also added to each tube (50 µL per tube), mixed gently with the pipette and incubated at 37°C in the incubator. Two different times of incubation were evaluated, 4 hours and 24 hours.
4. After incubation, 70 µL of 10% Sodium Dodecyl Sulphate (SDS) (Sigma-Aldrich, order code: 436143) and 5 µL of proteinase K at a concentration of 10 mg/mL (Promega, UK – order code: V3021) were added, gently mixed with the pipette and incubated at 60-65°C for 10 minutes. At the same time, the CTAB/NaCl was pre-warmed at 60-65°C.

5. 100 µL of 5M NaCl was added to each tube.
6. 100 µL of of pre-warmed CTAB/NaCl was added to each tube, mixed gently with the pipette and incubated at 60-65°C for 10 minutes.
7. After incubation, 500 µL of phenol:chloroform:isoamyl alcohol (25:24:1, Tris-saturated) (Sigma-Aldrich, order code: P2069) was added, mixed with the pipette and centrifuged at 10,000 g for 5 minutes at room temperature.
8. The supernatant aqueous phase was then transferred to a new 2-mL microcentrifuge tube. A small amount of the aqueous phase was left in the original tube to reduce the risk of contamination with proteins.
9. Chloroform:isoamyl alcohol (24:1) (Sigma-Aldrich, order code: C0549) was added (500 µL), mixed with the pipette and centrifuged at 10,000 g for 5 minutes at room temperature.
10. The supernatant aqueous phase was transferred to sterile DNAase free 1.5-mL Eppendorf tubes, 500 µL of 100% ethanol were added and the samples frozen at -20°C for at least 30 minutes.
11. After at least 30 minutes, the tubes were centrifuged at 10,000 g for 15 minutes at 4°C.
12. The supernatants were removed; the pellets were washed with 1 mL of ice-cold 70% ethanol and inverted gently.
13. Tubes were centrifuged again at 10,000 g for 5 minutes at 4°C.
14. Ethanol was removed as much as possible and the tubes left with open lids to allow the pellets to air-dry (at least 15 minutes).
15. Finally, the pellets were rehydrated with 100 µL of Tris-EDTA buffer, overnight at 4°C (or 1 hour at 65°C).

The extracted genomic DNA was then ready to be used for further applications and it could be stored at -20°C.

2.6.4 Buffers and solutions

1 M Tris-Cl (Promega, order code: H5123) – To prepare a 1 M stock solution of Tris-Cl, 121.1 g of Tris base were dissolved in 800 mL of distilled water. The pH was adjusted to the desired value by adding concentrated HCl (pH 7.4, add 70 mL – pH 7.6, add 60 mL – pH 8.0, add 42 mL). The final volume was adjusted to 1 L, dispensed into aliquots and sterilized by autoclaving.

0.5 M EDTA (Sigma-Aldrich, order code: D2900000) – To prepare a stock solution of EDTA at 0.5 M, 186.1 g of disodium EDTA were added into 800 ml of distilled water.

The solution was stirred vigorously to dissolve the salt, final volume adjusted to 1 L, dispensed into aliquots and sterilized by autoclaving.

Lysozyme (Sigma-Aldrich, order code: L6876) – To prepare a concentration of 20mg/mL, 100 mg of lysozyme were added to 5mL of enzymatic lysis buffer for the DNeasy extraction method (see below). Lysozyme used in the CTAB extraction method was bought already made at a concentration of 10mg/mL.

Triton™ X100 (Thermo-Fisher, order code 85111) – This was bought in a 50-mL container ready to be diluted.

Enzymatic lysis buffer for DNeasy extraction – This buffer was composed of 20 mM Tris-Cl (pH 8.0), 2 mM sodium EDTA, 1.2% Triton™ X100 and lysozyme at a concentration of 20 mg/ml. One litre of buffer was prepared by adding the following ingredients:

- 20ml of 1M Tris-Cl (final concentration of 20mM Tris-Cl in buffer)
- 2ml of 0.5M EDTA (final concentration of 1mM EDTA in buffer)
- 12ml Triton™ X-100 (final concentration 1.2% Triton® X-100 in buffer)

Lysozyme was added to 5 mL of the above buffer to reach a final concentration of 20 mg/mL.

Tris-EDTA buffer (Sigma-Aldrich, order code: 93283) – This was bought in a 100-mL container and ready to use.

10% SDS (Sigma-Aldrich, order code: 436143) – 10 g of SDS were dissolved in 100 mL of distilled water.

Proteinase K (Promega, order code: V3021) – 10 mg/mL stock solution was prepared by dissolving 100 mg in 10 mL of distilled water and 1 ml aliquots were stored at -20°C.

Sodium Chloride/NaCl (Sigma-Aldrich, order code: S7653) – 5 M solution of NaCl was prepared by dissolving 292 g of NaCl in 800 mL of distilled water. The final volume was adjusted to 1 L, dispensed into aliquots and sterilized by autoclaving.

CTAB/NaCl – 4.1 g of NaCl and 10 g of CTAB (Sigma-Aldrich, order code: H9151) were mixed in 100 mL of distilled water and heated to 65°C to completely dissolve.

Lysis buffer for CTAB method – The lysis buffer (100 ml) was composed of:

- 15% sucrose (Sigma-Aldrich, order codeLS7903), by adding 15 g into 100 mL of distilled water
- 0.05 M Tris-Cl (pH 8.5), by adding 5 mL of 1 M Tris-Cl stock solution
- 0.05 M EDTA, by adding 10 mL of 0.5 EDTA stock solution

Once prepared, the lysis buffer for the CTAB extraction was stored in the fridge at 4°C.

2.7 Statistical methods and analysis

The *t-test* was used to compare the means of different samples where necessary and a *P-value* of ≤0.05 was considered to be statistically significant. GrapPad Prism software version 7.05 was used for the statistical analysis and complex mathematical calculations.

Chapter 3: In vitro evolution of isoniazid resistance

3.1 Introduction

Isoniazid (INH) is one of the most commonly used antituberculous drugs and the backbone of all first line anti-TB treatment combinations. Data on the prevalence of resistance suggest that INH mono-resistance is much more frequent than rifampicin (RIF) mono-resistance. INH-resistance has an estimated global prevalence of 13.3% of all TB cases compared with 5.3% for MDR and 0.6% for RIF mono-resistance (WHO, 2017). This discrepancy in resistance rates was originally attributed to the higher mutation rate of INH when compared to RIF, 2.56×10^{-8} versus 2.25×10^{-10} (~100 times higher) (David, 1970), but a more recent paper has challenged this explanation as the *in vivo* mechanism of isoniazid resistance is not reflected by *in vitro* experiments (Bergval, Schuitema, Klatser, & Anthony, 2009).

INH resistant strains were clinically described soon after the drug's introduction in 1952 (Middlebrook, 1954). One of the earliest INH resistance mechanisms discovered was the complete deletion of the *katG* gene. This conferred resistance by reducing the catalase activity which is needed to metabolise INH into its active forms. These strains also exhibited a relative lack of virulence in mice and guinea-pigs, as functional KatG activity is required for virulence in these species (Y. Zhang et al., 1992). Other authors have also demonstrated that mutants with low levels of KatG enzymatic activity failed to persist in guinea pig spleen, showing less virulence (Z. Li, Kelley, Collins, Rouse, & Morris, 1998). However, it was later discovered that INH-R clinical isolates generally exhibit two main point mutations, *katG*-S315T (found in ≤75% INH-R strains) and *inhA*-C15T (in ≤25% strains) and not a complete deletion (M. Zhang et al., 2005). Additional studies in both clinical and laboratory strains also demonstrated that these strains retain some catalase activity despite the point mutation in *katG* and this is sufficient to confer intermediate to high levels of resistance to INH whilst allowing the strain to retain virulence in humans. In comparison, mutations in the *inhA* gene (involved in mycolic acid biosynthesis) do not affect the catalase activity and only confer low-level resistance to INH and also cross-resistance to ethionamide (Guo, Seet, Denkin, Parsons, & Zhang, 2006). Interestingly, the complete deletion of the *katG* is rare in clinical isolates and it is estimated to occur in only ~ 5% of INH-resistant clinical strains (Hazbon et al., 2006). This may be due to these strains having a reduced transmission rate (i.e. fitness cost of this mutation), whereas strains with only the point mutation, *katG*-S315T, have the selective

advantage of both INH-resistance whilst retaining transmissibility (van Soolingen et al., 2000).

The main aim of this chapter is to create laboratory mutants from the reference strain H37Rv with the same mutations of clinical isolates from the London INH-R outbreak. *KatG*-S315T and *inhA*-C15T were the most common mutations identified in these clinical isolates (Maguire et al., 2011) (Ruddy et al., 2004) (Casali et al., 2016) (R. J. Shorten, McGregor, Platt, & Jenkins, 2013). The driving hypothesis is that we will be able to understand the *in vitro* evolution of drug resistance and the development of compensatory mutations by comparing laboratory and clinical strains. Such mutations may have allowed the outbreak strains to be so successful over the years and they could represent new potential drug targets. It is worth mentioning for completion that the reference strain H37Rv used to create laboratory mutants belongs to Lineage 4 (Euro-American), the same as the clinical isolates from the London outbreak.

3.2 Materials and methods

3.2.1 Selection of mutants

INH-R laboratory mutants were selected using the reference strain H37Rv as previously described in Section 2.4. Considering that the critical concentration for INH is known to be between 0.2 and 1 µg/ml (see Section 1.7.5), six different drug concentrations were used to induce low to high level resistance: 1 µg/ml, 2 µg/ml, 4 µg/ml, 8 µg/ml, 16 µg/ml and 32 µg/ml. Additionally, these concentrations were used in three main experiments with different conditions for the stimulation of mutagenesis:

- **INH concentration:** In a first experiment, mutants were selected as described above using 7H10 plates with different concentrations of INH.
- **INH concentration, low pH and hypoxia:** In a second experiment, the 7H10 plates containing INH also included HCl to reduce the pH to 5.5 and incubated in hypoxic conditions (O2 reduced to 5%). This was done to simulate the acidic environment and hypoxic conditions found inside the lung granuloma during active human infection.
- **INH concentration with ciprofloxacin** In a third experiment, the 7H10 plates containing INH also contained a sub-inhibitory concentration of ciprofloxacin (0.125 µg/ml) in an attempt to increase the mutation rate and augment the selection of the required mutants.

3.2.2 Extraction of DNA and confirmation of mutations

Considering the high number of selected mutants and the need for a rapid screening of specific mutations (*katG*-S315T and *inhA*-C15T), the GenoLyse® commercial kit (Section 2.6.1) was used as a simple extraction method. No whole genome sequencing (WGS) was necessary at this stage.

The Line Probe Assay (LPA) *Genotype MTBDRplus* (Hain Lifescience, Nehren, Germany) was used to confirm the presence of required mutations and it was performed as per manufacturer's instructions. The assay has also an in-built quality control to confirm the identification of *M. tuberculosis*. The Line Probe requires the separate amplification of DNA extracts by PCR using biotin labelled primers for the *katG*, *inhA* and *rpoB* genes. The biotin-labelled amplicons are then hybridised with specific oligonucleotide probes immobilised as parallel lines on the membrane-based strips (Figure 4). On each strip there is a coloured marker line for orientation, one conjugate control, one amplification control, one probe to confirm the presence of *M. tuberculosis* and locus control probes for each gene (*rpoB*, *katG* and *inhA*). Additionally, there are eight wild type/sensitive probes and four mutant/resistant probes for *rpoB*, one wild type/sensitive and two mutant/resistant probes for *katG* and two wild type/sensitive and four mutant/resistant probes for *inhA*. If a specific mutation is present, the biotin-labelled amplicon binds to the corresponding mutation/resistant probe and this is detected by a staining reaction with alkaline phosphatase, which generates a purple/blue coloration (Figure 4). The various quality controls inbuilt in the assay also allow the detection of test failure and any extraction/amplification errors.

3.2.3 Fitness assay and mutation rate

Fitness and mutation rate assays were performed as described (Section 2.4.4 and Section 2.4.5.). They were performed on 18 randomly selected mutants that were generated (3 mutants for each of the 6 INH concentrations tested) and on the INH-R isolates from the London outbreak (16 samples plus 2 additional controls, for a total of 18 clinical isolates) (see also Chapter 4 for further information on the London outbreak).

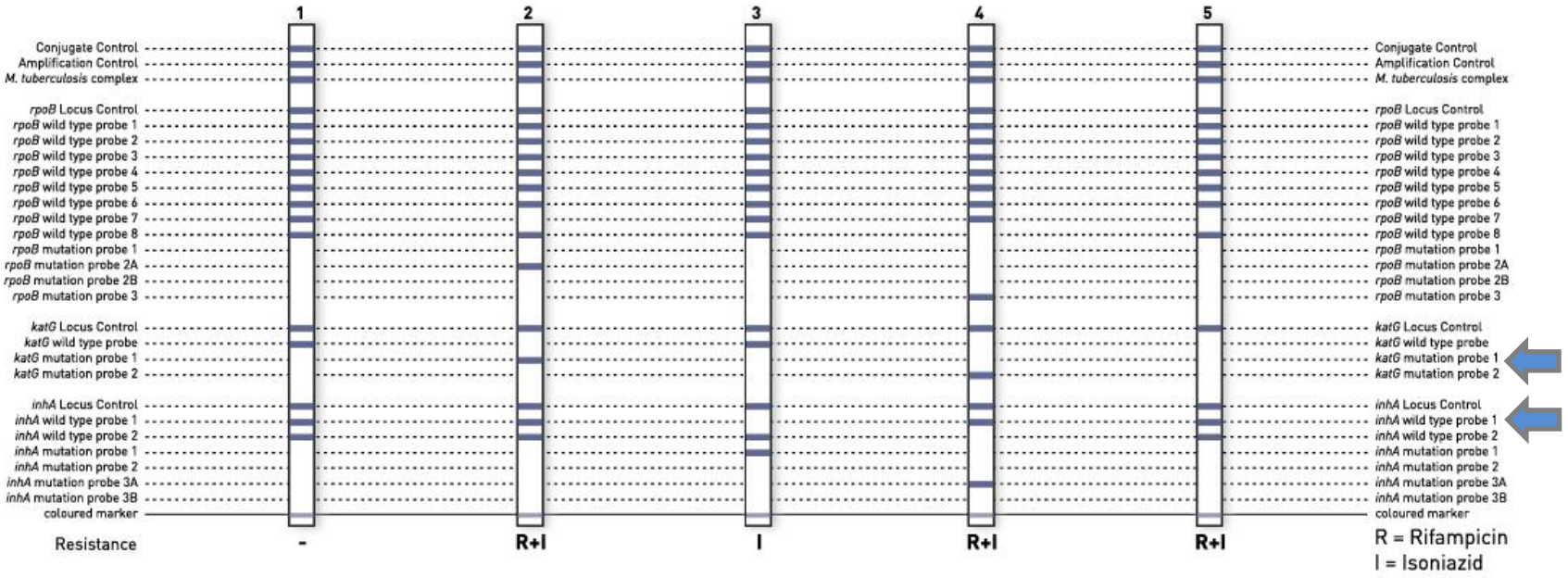


FIGURE 4: THE LINE PROBE ASSAY GENOTYPE *MTBDRPLUS* BY HAIN LIFESCIENCE.

The Line Probe Assay (LPA) Genotype *MTBDRplus* allows the identification of both INH and RIF resistance. In particular, the probe can detect both high and low level INH resistances, by analysing common mutations in the *katG* and *inhA* genes. The *katG* mutation probe 1 and 2 detect the mutation S315T, whilst the *inhA* mutation probe 1 detects the mutation C15T (indicated by arrows in the figure above). *InhA* mutation probes 2, 3A and 3B detect additional mutations A16G, T8C and T8A, respectively. The assay also contains additional probes to confirm the presence of both genes (*katG* and *inhA* locus controls). Picture and content from *Genotype MTBDRplus* Instruction for use (Version 2.0 – IFU-304A-02).

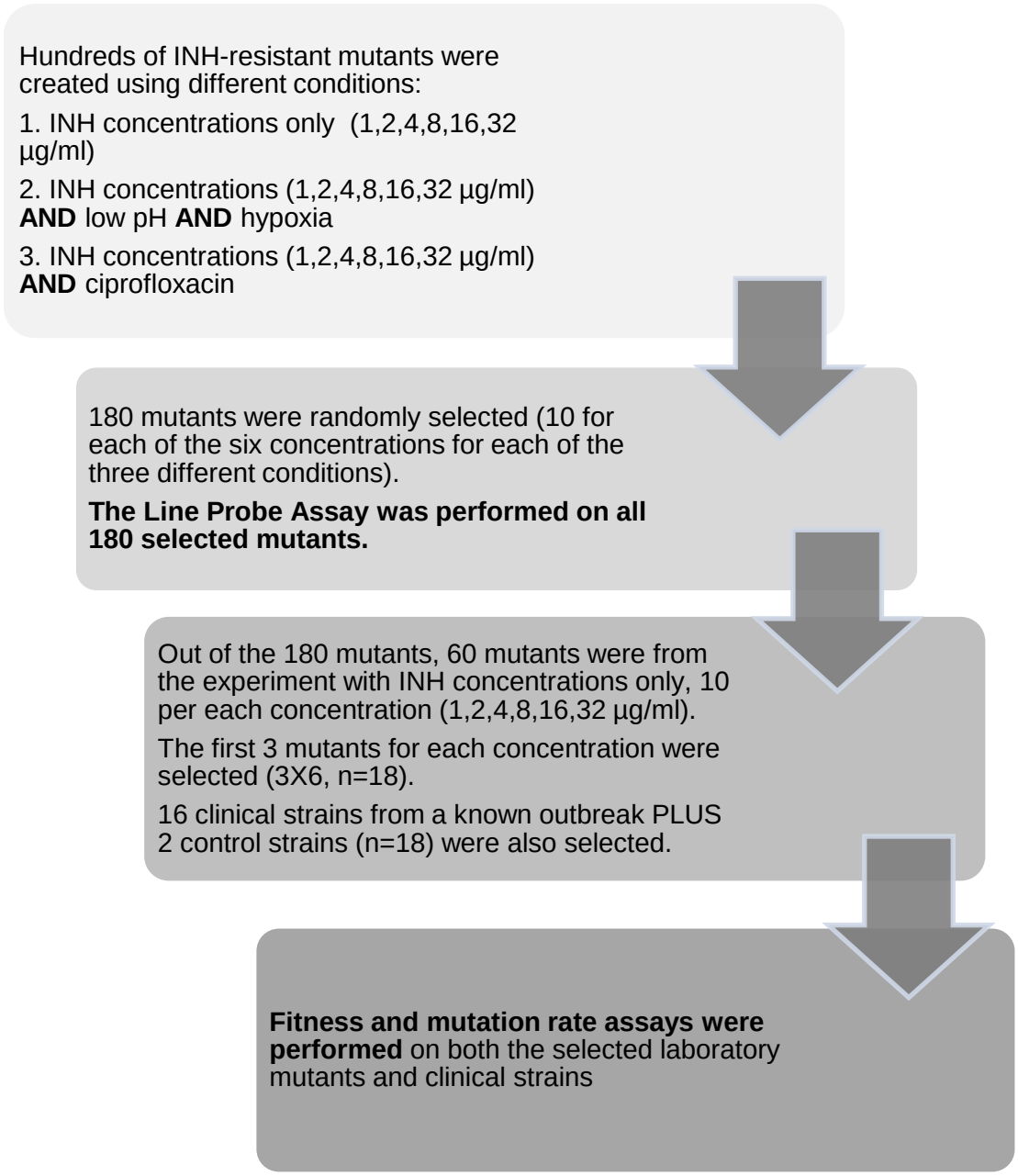


FIGURE 5: DIAGRAM EXPLAINING THE SELECTION OF RESISTANT MUTANTS FOR FURTHER ANALYSIS

The diagram above explains how resistant mutants were selected for further analysis. Originally, hundreds of different mutants were created using different conditions and concentrations. The Line Probe Assay was performed on 180 randomly selected mutants. Due to the complexity of fitness and mutation rate assays, these assays were only performed on clinical strains and selected INH-resistant laboratory mutants.

3.3 Results

In total, 180 mutants were randomly selected for further analysis (10 for each of the six INH concentrations for each of the three experiments) (Figure 5).

3.3.1 Line probe results

The results of the line probes assay are shown in Table 10. None of the 180 selected laboratory mutants exhibited the *katG*-S315T and *inhA*-C15T mutations that were required to match them with the clinical isolates of the London INH-R outbreak.

The results of the Genotype *MTBDRplus* probe assay also show that that the *katG* and *inhA* locus controls (inbuilt within the assay to confirm the presence of the genes) are largely unaffected by the lowest INH concentration (1 µg/ml) and the majority of laboratory mutants retain both genes and the wild type (WT) sequence (Table 10 and Figure 6). However, this is not the case for higher concentrations: as the concentration increases, more mutants exhibit a total deletion of the *katG* gene. At a concentration of 32 µg/ml, the majority of laboratory mutants seem to have lost the complete *katG* gene. The *inhA* gene is unaffected and it still retains the wild type variant (WT) in all mutants tested (Table 10 and Figure 6).

INH concentration (# mutants selected)	Additional conditions (if any)	Mutants with total deletion of *katG*
INH1 (n=10)	None	1
INH2 (n=10)	None	4
INH4 (n=10)	None	3
INH8 (n=10)	None	5
INH16 (n=10)	None	7
INH32 (n=10)	None	9
INH1 (n=10)	pH 5.5 and hypoxic conditions	1
INH2 (n=10)	pH 5.5 and hypoxic conditions	1
INH4 (n=10)	pH 5.5 and hypoxic conditions	4
INH8 (n=10)	pH 5.5 and hypoxic conditions	6
INH16 (n=10)	pH 5.5 and hypoxic conditions	7
INH32 (n=10)	pH 5.5 and hypoxic conditions	8
INH1 (n=10)	Addition of ciprofloxacin	3
INH2 (n=10)	Addition of ciprofloxacin	6
INH4 (n=10)	Addition of ciprofloxacin	6
INH8 (n=10)	Addition of ciprofloxacin	6
INH16 (n=10)	Addition of ciprofloxacin	7
INH32 (n=10)	Addition of ciprofloxacin	8

TABLE 10: **GENOTYPE *MTBDRPLUS* RESULTS FOR LABORATORY MUTANTS (N=180).**

The Line Probe Assay was performed on 180 randomly selected mutants. None of the laboratory mutants showed the *katG*-S315T and *inhA*-C15T mutations. The INH concentration used and the number of selected mutants are shown in the first column (for example, INH1 refers to a INH concentration of 1 µg/ml). The column on the far right indicates the number of mutants with a total deletion of the *katG* gene.

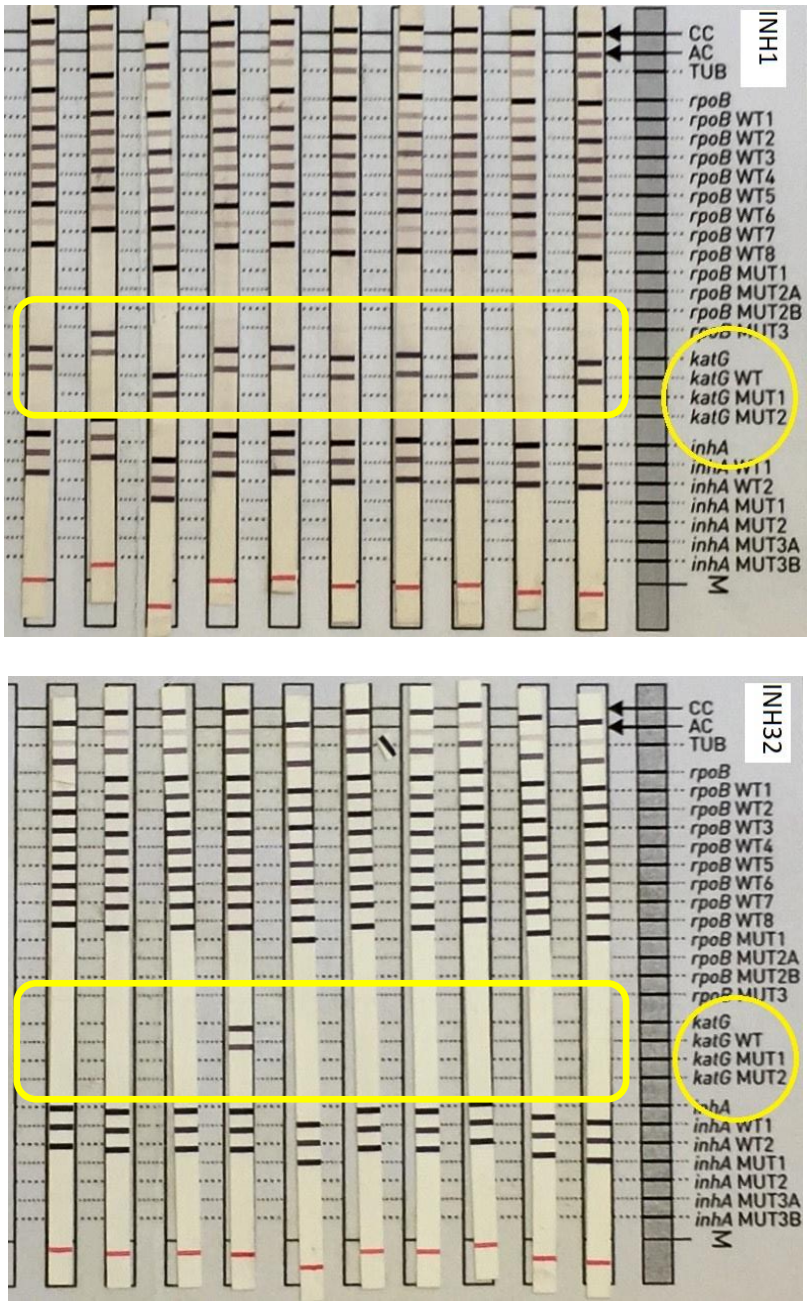


FIGURE 6: **LINE PROBE ASSAY GENOTYPE *MTBDRPLUS* RESULTS FOR INH-R MUTANTS AT LOW (1 µG/ML, TOP) AND HIGH INH CONCENTRATION (32 µG/ML, BOTTOM).**

At a low INH concentration (1 µg/ml, top figure), the majority of laboratory mutants still retain the *katG* gene in its wild type (WT) form (presence of two black/grey bands). Each line probe strip represents the result for one laboratory mutant. Only one isolate (second from the right) has lost the complete *katG* gene. However, at the highest INH concentration (32 µg/ml, bottom figure), the majority of laboratory mutants have lost the complete *katG* gene (apart from the forth sample from the left) (absence of bands).

3.3.2 Fitness and mutation rate assays

Fitness and mutation rate assays were performed on the selected 18 laboratory mutants (3 per each of the 6 INH concentrations) and on the INH-R isolates from the London outbreak (16 samples plus 2 additional controls, for a total of 18 clinical isolates). As the mutants were already randomly selected from the plate, the first three mutants for each INH concentration were chosen. There was a significant difference in the fitness value between clinical isolates and laboratory mutants but no difference in the mutation rate. Results are shown in Tables 11, 12, 13 and Figure 7.

INH-R clinical isolates	Median generation time G	Relative Fitness to H37Rv (G = 17.1)
02:113	17.8	0.96
02:292	19.4	0.88
02:302	16.3	1.05
03:013	19.2	0.89
03:039	20.8	0.82
03:303	18.6	0.92
03:313	17.4	0.98
04:018	22.8	0.75
04:194	17.8	0.96
04:198	17.4	0.98
04:211	18.4	0.93
04:493	16.9	1.01
04:503	18.2	0.94
05:046	16.6	1.03
07:116	17.2	0.99
07:118	20.8	0.82
05:177 (INH-R control)	18.6	0.92
05:094 (INH sensitive control)	19.4	0.88
Mutation rate assay (for ciprofloxacin)		
1.3×10^{-8} x cell division		

TABLE 11: FITNESS AND MUTATION RATE VALUES OF THE INH-R CLINICAL ISOLATES

The clinical isolates have fitness values ranging from 0.75 to 1.05 (n=16 clinical strains plus two control strains). The fitness value is expressed as a relative fitness to H37Rv by comparing their generation times G. The mutation rate was the same for both clinical isolates and the laboratory mutants. The names of the clinical strains were assigned by the Royal Free laboratory: the first two digits in the isolate number indicate the year of isolation (for example, 02:113 was isolated in 2002), whilst the second three digits is the original sample number.

Selected INH-R laboratory mutants	Median generation time G	Relative Fitness to H37Rv (G = 17.1)
INH1a	28.5	0.60
INH1b*	23.7	0.72
INH1c	21.9	0.78
INH2a*	19	0.90
INH2b*	19	0.90
INH2c	23.1	0.74
INH4a	20.8	0.82
INH4b*	21.9	0.78
INH4c	20.3	0.84
INH8a*	19.4	0.88
INH8b*	22.8	0.75
INH8c	25.1	0.68
INH16a*	19	0.90
INH16b	21.4	0.80
INH16c*	20.1	0.85
INH32a*	24.4	0.70
INH32b*	19.2	0.89
INH32c*	20.3	0.84
Mutation rate assay (for ciprofloxacin)		
1.3×10^{-8} x cell division		

TABLE 12: FITNESS AND MUTATION RATE VALUES OF THE INH-R LABORATORY MUTANTS

The laboratory mutants have fitness values ranging from 0.60 to 0.90 (n=18 laboratory mutants). The fitness value is expressed as a relative fitness to H37Rv by comparing their generation times G. The mutation rate was the same for both clinical isolates and the laboratory mutants. The name of the mutant refers to the original INH concentration used in the plate (for example, INH1 refers to an INH concentration of 1 µg/ml) and mutants were also numbered using alphabetical letters (*a, b, c* etc). The asterisk (*) indicate that the laboratory mutant has a full deletion of the *katG* gene.

Mean fitness value of clinical isolates	Mean fitness value of INH-R laboratory mutants
0.93 (SD 0.078)	0.80 (SD 0.087)
P <0.0001	

TABLE 13: **MEAN FITNESS VALUES FOR CLINICAL ISOLATES AND LABORATORY MUTANTS**

Clinical isolates have a mean fitness value of 0.93, whereas the laboratory mutants have a mean value of 0.80. SD = Standard Deviation. The mean value for the laboratory mutants was significantly lower compared to the mean value of the clinical isolates.

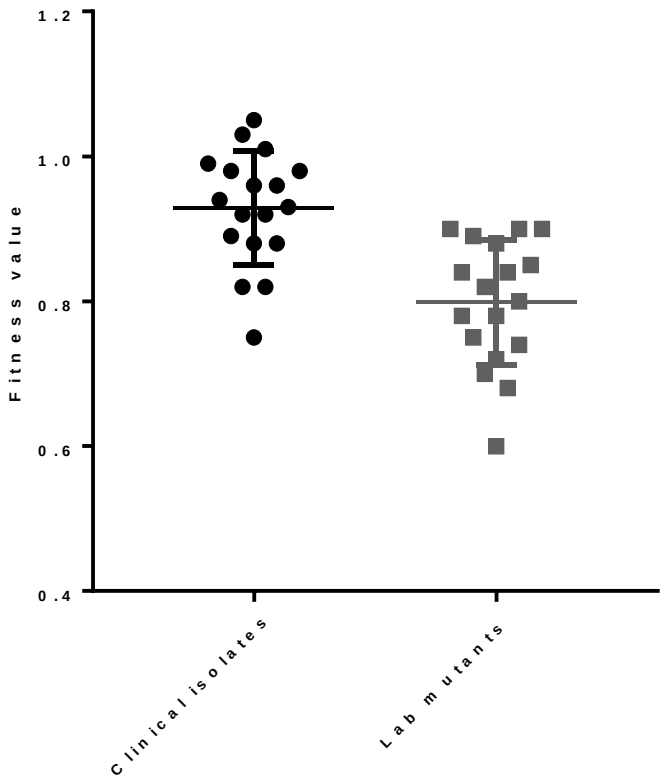


FIGURE 7: GRAPHICAL REPRESENTATION OF THE DIFFERENT FITNESS VALUES OF THE INH-R LABORATORY MUTANTS VERSUS CLINICAL ISOLATES.

The different fitness values of clinical isolates (n=18) and laboratory mutants (n=18) are represented on the *y*-axis. There was a statistically significant difference (P value <0.0001) between the mean fitness value of the clinical isolates versus the laboratory mutants.

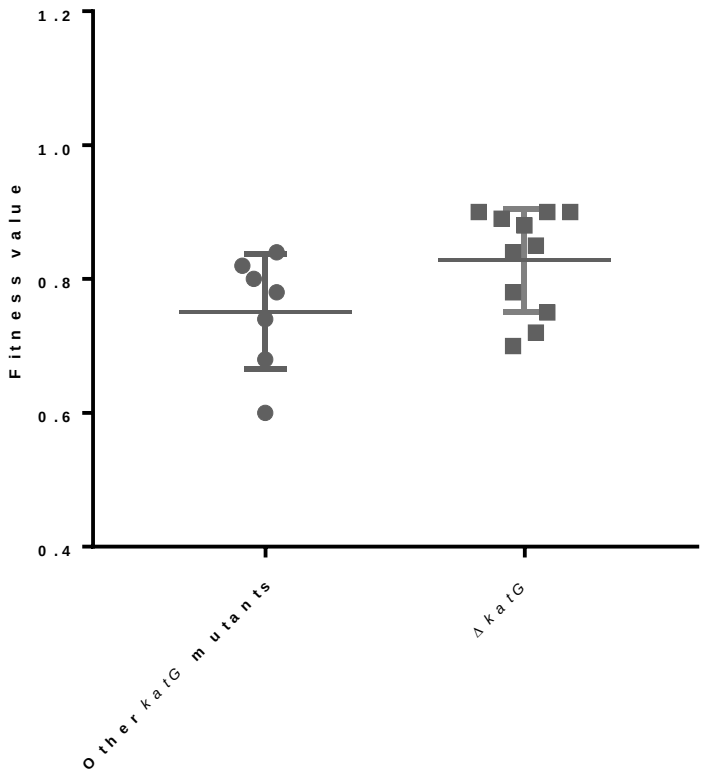


FIGURE 8: GRAPHICAL REPRESENTATION OF THE DIFFERENT FITNESS VALUES OF THE INH-R LABORATORY MUTANTS (*ΔkatG* VERSUS OTHER *katG* MUTANTS)

This graph represents a subgroup analysis of the different fitness values among laboratory mutants by comparing the mutants with a complete deletion of the *katG* gene (*ΔkatG*) (n=11) versus the other *katG* mutants (n=7). There was no significant difference between the two groups (P value 0.67).

3.4 Discussion

The original aim of this chapter was to create laboratory mutants from the reference strain H37Rv with the same mutations (*katG*-S315T and *inhA*-C15T) as the clinical isolates from the London INH-R outbreak. The secondary aim would have been to perform WGS on both laboratory and clinical isolates to identify any compensatory mutations that have allowed the outbreak strains to be so successful, thus identifying new potential drug targets. However, due to the failure in selecting the desired laboratory resistant mutants, WGS sequencing was considered unnecessary on these and it was only performed on the clinical isolates (Chapter 6).

INH-R laboratory mutants of the H37Rv reference strain were selected under different conditions (using different INH concentrations and including low pH and hypoxia or with the addition of sub-inhibitory concentrations of ciprofloxacin). None of the selected mutants carried the mutations that are predominant in the clinical isolates (*katG*-S315T and *inhA*-C15T). Additionally, a large number of mutants acquired INH-resistance by complete deletion of the *katG* gene, whilst the *inhA* gene was generally unaffected. As mentioned in Section 5.1, the complete deletion is rare in clinical isolates and it is estimated to occur in only 5% of them (Hazbon et al., 2006). The deletion also seemed to be associated with the INH concentration used in the plates, with higher concentrations selecting more Δ*katG* mutants, thus with a higher level of resistance. The acquisition of INH resistance was also associated with a significant reduction of the fitness value in the laboratory mutants (mean value: 0.80). Conversely, the selected clinical isolates from the London outbreak had a significantly higher fitness value (mean value: 0.93) ($p < 0.0001$) but still maintained the same mutation rate, supporting the hypothesis of some selective adaptation and advantage that has allowed them to be so successful in causing infection.

The *katG* gene encodes for the catalase peroxidase enzyme, whose main function is to protect the mycobacterial cell from the oxidative stress encountered inside macrophages or granuloma (Manca, Paul, Barry, Freedman, & Kaplan, 1999). It can be hypothesised that *in vitro* laboratory conditions are less challenging compared to the *in vivo* immune response. Supporting this, some authors (Bergval et al., 2009) have speculated that the protective conditions of the culture media do not reflect the animal infection model (even with the addition of H_2O_2) and they are responsible for the high prevalence of *katG* deletions, instead of point mutations observed in clinical

samples. In our experiment, lowering the pH and incubation in a hypoxic atmosphere did not prevent the selection of Δ*katG* mutants either. To further confirm this, there is only one study that has reported the creation of a *katG*-S315T mutant (Guo et al., 2006), but this was selected using a mouse model, infected with *M. tuberculosis* and treated with INH. Moreover, previous reports have suggested both a higher mutation rate (Riska, Jacobs, & Alland, 2000) and genetic instability (Y. Zhang et al., 1992) of the *katG* region, and this could represent another explanation for the high prevalence of Δ*katG* mutants observed in our and other laboratory studies (Bergval et al., 2009).

In this chapter, we did not investigate the presence of partial deletions or alternative mutations in the *katG* gene as it was beyond its original aim. Furthermore, such mutations would not have been clinically relevant and they would not have been detected by the Genotype *MTBDRplus* probe assay. The failure to create laboratory mutants with *katG*-S315T and *inhA*-C15T mutations strongly suggests that the *in vitro* model does not reflect the *in vivo* evolution of clinical isolates and an alternative approach is needed to better understand the genetic mechanisms that can allow persistent infection. WGS may help in addressing this challenge and this will be the focus of the next chapter.

Chapter 4: Genetic variation in *Mycobacterium tuberculosis* isolates from a London outbreak associated with isoniazid resistance

4.1 Introduction

An outbreak of isoniazid-resistant tuberculosis (INH-R TB) was first identified in London in 1995 and it has now been ongoing for over 20 years, making it the largest drug-resistant outbreak of TB documented to date worldwide (with over 500 cases) (Ruddy et al., 2004) (C. M. Smith et al., 2017). Conventional epidemiological analysis indicates that 50% of cases were born in the UK, were of white or black-Caribbean ethnicity and had a strong link to drug use and prison detention. Adherence to treatment was poor in one third of total patients and several went on to acquire further drug resistant strains including MDR TB (Maguire et al., 2006). A second clinically-relevant feature of this outbreak was the high transmission of infection to contacts (11%) compared with other documented outbreaks (0.7-2%) (Maguire et al., 2011). This could only be partially explained by the epidemiological data, suggesting that other factors must have contributed to the extent of this outbreak.

Shorten et al, previously reported that this outbreak is atypical, lacking a point source and serial transmission (R. J. Shorten et al., 2013). In fact, it is composed of multiple clusters (defined as samples related at ≥80% similarity by IS6110 fingerprinting and starting from a minimum of two cases) and spans a period of several years. One of the biggest clusters is the RFL15, from Royal Free London and initially called Lineage 15 (R. Shorten, 2011). Demographic characteristics of this cluster are generally comparable to the wider outbreak (Maguire, Dale, & McHugh, 2002): the majority of patients were male, UK born but with a high proportion of black Caribbean background, based in North London, associated with alcohol or drug abuse and prison and/or homelessness (C. M. Smith et al., 2017). It is also worth acknowledging that the RFL15 only represents a small sample of the ongoing outbreak but some of its features make it clinically interesting and unusual (Figure 9 and Table 14). The cluster is characterized by persistent transmission over the years (from 2002 to 2007) and it is composed of a variety of isolates with different antibiograms (including drug-susceptible, isoniazid- and streptomycin-mono-resistant as well as MDR isolates). Epidemiological factors may potentially explain some of the features: many patients in this outbreak were prisoners and drug users residing in the North London Boroughs.

Even if direct transmission could not be demonstrated, this does however indicate these isolates may have been circulating within specific communities.

Several reports suggest that the acquisition of resistance leads to a reduction in the fitness of the affected strain, but the full picture is far from clear (Gagneux, Long, et al., 2006) (Billington et al., 1999) (O'Sullivan et al., 2010) (O'Sullivan, McHugh, & Gillespie, 2005). A spontaneous mutation that confers drug resistance should provide an advantage in an appropriately selective environment (i.e. patient on treatment). If the mutation affects an essential gene/function causing a metabolic cost, then we could reasonably hypothesise that the mutant will be less "fit" than its sensitive precursor. However, this is not always the case, and the use of whole genome sequencing (WGS) has confirmed the presence of compensatory mutations that maintain a high competitive fitness (Comas, Borrell, & Roetzer, 2011).

Additionally, clinical strains of *M. tuberculosis* show a genomic diversity which varies from few single nucleotide polymorphisms (SNPs) (Musser et al., 2000) to large scale genomic rearrangements (Tsolaki, 2004). The majority of deletions are considered to be present in genes encoding for proteins not essential for the pathogenesis of the disease as in these analyses all strains were obtained from clinical cases with active TB. However, some deletions could conceivably result in a selective advantage at particular stages of infection or transmission; or even enable escape from the host immune response. Other deletions could confer a strong advantage, such as antibiotic resistance. An example of this is the deletion of the *katG* gene, resulting in isoniazid resistance (Heym, Alzari, & Honore, 1995).

This chapter focuses on the genetic variation in a subset of the London INH-R outbreak clinical samples with the hypothesis that these isolates have unique biological characteristics that serve to prolong the outbreak. Fitness assay, mutation rate and WGS were performed to test the hypothesis of selective advantage and compensatory mutations.

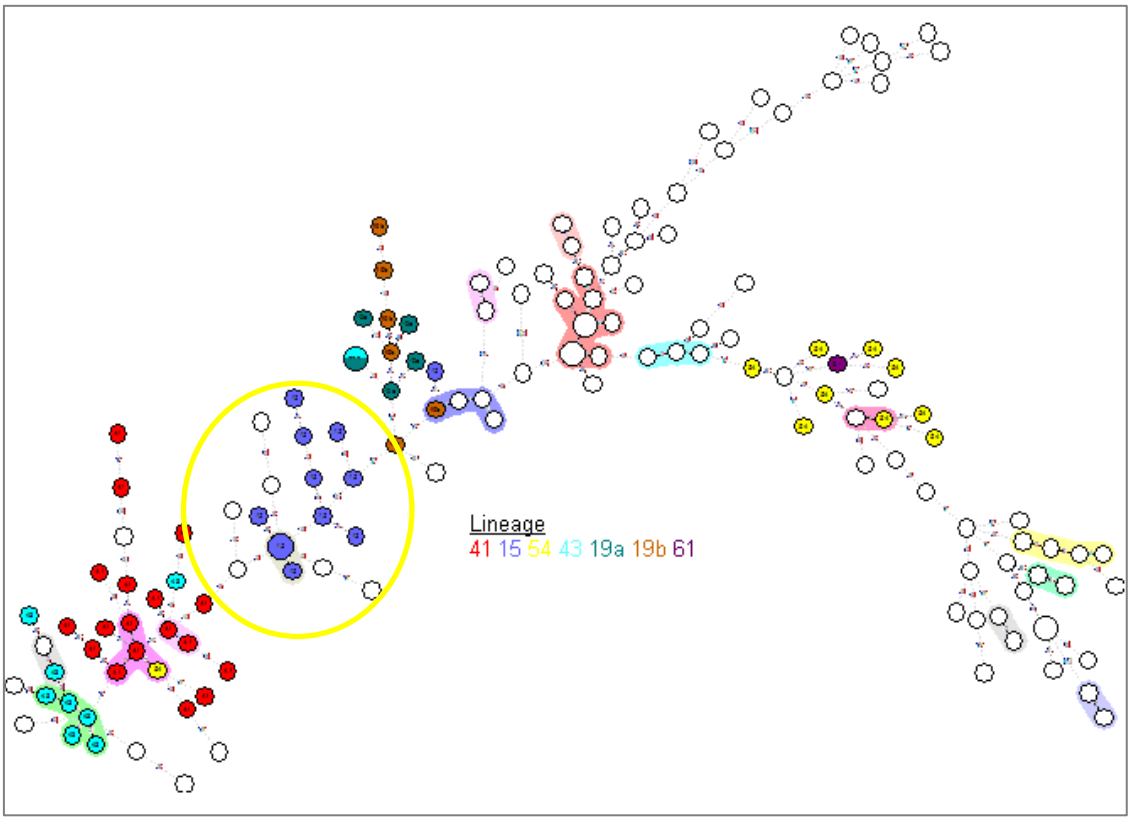


FIGURE 9: MINIMUM SPANNING TREE OF LONDON OUTBREAK ISOLATES TYPED BY MIRU-VNTR

The London outbreak isolates have been previously typed using MIRU-VNTR as part of the routine clinical service at Public Health England (PHE). Each node represents a single isolate, except larger nodes, which represent two identical isolates. The distance between nodes (isolates) are proportional to the number of differences on loci copy numbers between isolates. The outbreak isolates can be divided in various lineages (numerically numbered by PHE) and lineages 15, 19, 41, 43, 54 and 61 are shown above. In particular, lineage 15 (RFL15) is circled in yellow. The RFL15 only represents a small sample of the ongoing outbreak but some of its features make it clinically interesting and unusual. The cluster is characterized by persistent transmission over the years (from 2002 to 2007) and it is composed of a variety of isolates with different antibiograms (Courtesy of Dr Steven Platt, PHE).

Type	Isolate	Lineage	Resistant to	Sensitive to	Polymorphism for INH
Cluster RFL15	02.113	Euro American	-	HRZES	None detected
	02.292	Euro American	H	RZES	*inhA C→T*
	02:302*	Euro American	S	HRZE	None detected
	03.013	Euro American	S	HRZE	None detected
	03.039	Euro American	H	RZES	*inhA C→T*
	03.303*	Euro American	H	RZES	*inhA C→T*
	03.313	Euro American	S	HRZE	None detected
	04.018	Euro American	H,R,clari, ethi	ZES	*inhA C→T*
	04.194	Euro American	H	RZES	*kat*G S315T
	04:198*	Euro American	H	RZES	*inhA C→T*
	04.211	Euro American	H	RZES	*inhA C→T*
	04.493	Euro American	H	RZES	*inhA C→T*
	04.503	Euro American	H	RZES	*inhA C→T*
	05.046	Euro American	-	HRZES	None detected
	07.116	Euro American	H, ethi	RZES	*inhA C→T*
	07:118	Euro American	H	RZES	*kat*G S315T
Isoniazid resistant control	05.177 (Control sample 1)	Euro American	H	RZES	*inhA C→T*
Isoniazid sensitive control	05.094 (Control sample 2)	Euro American		HRZES	None detected
Isoniazid sensitive control	04.011 (Control sample 3)	East African Indian		HRZES	*kat*G R463L (but no resistance)
*no WGS available as unable to extract enough DNA					

TABLE 14: LIST OF CLINICAL ISOLATES FROM THE SELECTED CLUSTER RFL15.

The first two digits in the isolate number indicate the year of isolation, for example, 02:113 was isolated in 2002 (n=16 clinical strains plus three control strains). Sensitivities (H=Isoniazid, R=Rifampicin, S=Streptomycin, Clari=Clarithromycin, Ethi=Ethionamide) were based on phenotypical testing. Only the genetic mutations conferring isoniazid resistance are reported for simplicity. Samples in grey shade are closely related on further phylogenetic analysis (see phylogenetic tree, Figure 12).

4.2 Materials and Methods

4.2.1 Sample selection

The RFL15 isolates previously described (R. Shorten, 2011) were used for this investigation and were evaluated by fitness assay, mutation rate and WGS. Clinical isolates were originally frozen at -80°C and all experiments were performed directly from the original stock with only one passage. The reference strain H37Rv (from Public Health England, National Collection of Type Cultures), one unrelated INH-R isolate and two unrelated susceptible isolates were included as controls. Drug sensitivities were performed at the National Mycobacterium Reference Laboratory (Public Health England, London) as part of the routine clinical service. All isolates were isoniazid resistant, except 02:113 and 05:046, which were fully sensitive, and 02:302, 03:013 and 03:313, which were streptomycin mono-resistant. Samples 04.018 and 07.116 had additional resistance (Table 14). Despite the different sensitivity profiles, all isolates were included as part of RFL15 (Figure 9) and for a wider comparative genetic analysis.

4.2.2 Bioinformatics analysis

WGS data were downloaded from the PHE server as FASTQ files (Section 1.10.3). A quality control (QC) was run for all the sequenced clinical isolates and it included a Q-score 30 (Q30), read length, coverage, Shannon entropy and mapping coverage to H37Rv as explained in section 1.10.4. In addition to sending all the clinical isolates to the National Mycobacterium Reference Laboratory, all FASTQ files were also analysed using the Mykrobe predictor software (Section 1.10.7) for confirmation of identification at species level (*M. tuberculosis*) and resistance profile. Sequence data was then aligned to the H37Rv reference genome (RefSeq: NC_000962.3) using BWA-MEM 0.7.12 and sorted using SAMtools v0.1.19 (Section 1.10.5). All genome sites were called using SAMtools as described previously (Witney, Gould, & Arnold, 2015). The variant sites were filtered based on the following criteria: mapping quality (MQ) of >30, site quality score (QUAL) of >30, ≥4 reads covering each site with ≥2 reads mapping to each strand. Phylogenetic reconstruction was performed using RAxML v8.2.3 (Stamatakis, 2014a) with a GTR model of nucleotide substitution and a GAMMA model of rate heterogeneity; branch support values were determined using 1000 bootstrap replicates. Branch SNP counts were estimated by ancestral sequence reconstruction performed with PAML v4 (Yang, 2007). Circular plots were generated using Circos (Section 1.10.6). Dr Adam Witney, bioinformatics scientist

(Institute for Infection and Immunity, St. George's University of London, UK) kindly performed the alignment and the phylogenetic reconstruction as described above and he provided the *vcf* file (Section 1.10.3). This was used within Artemis for identifying SNPs, their position and genes involved.

4.3 Results

Sixteen clinical isolates and 3 unrelated control samples were available, but 13 isolates and the controls (16 samples in total) were included in the genetic analysis due to DNA extraction failures (Table 14). The fitness and the mutation rate results of the INH-R clinical isolates have already been shown in Chapter 3 (Table 11).

4.3.1 DNA extraction results

The 16 clinical strains were used to compare different extraction methods, as described in Section 2.6. The reference strain H37Rv was used as a positive control and sterile distilled water as the negative control. Purified DNA concentration was measured using a Qubit® and the 260/280 ratio calculated using a NanoDrop™ 2000 Spectrophotometer as described in Section 2.5.

- *Genolyse*® – The extraction was performed once and the DNA concentration (ng/µl) obtained for each clinical strain using GenoLyse kits is shown in Table 15.
- *DNeasy*® – The extraction was performed once and the DNA concentration (ng/µl) obtained using the DNeasy® kit for each clinical strain is shown in Table 16.
- *CTAB based extraction* – The concentrations (ng/µl) obtained for each clinical strain are shown in Table 17.

The CTAB based extraction method proved to be the optimal method for purifying mycobacterial DNA. This method provided high quality DNA at a viable concentration when prolonged incubation with lysozyme (24 hours) was used.

Only three samples (02:302, 03:303 and 04:198) failed to produce enough DNA despite multiple extraction attempts (Table 17). An additional experiment was carried out adding a mechanical disruption step as described by Kaser (Kaser et al., 2010) using 0.1mm silica beads (Lysing Matrix B) and a FastPrep®-24 Instrument (both from MP Biomedicals, Santa Ana, USA) before phenol:chloroform extraction and

ethanol precipitation. The addition of a mechanical disruption step to these three samples only marginally increased the DNA yield (2.31 ng/µl for 02:302 – 4.32 ng/µl for 03:303 and 2.54 ng/µl for 04:198) but it significantly decreased DNA purity with 260/280 ratios of 1.43, 1.54 and 1.49, respectively.

INH-R strains (n=16)	DNA yield (ng/µl) with GenoLyse®	DNA yield (ng/µl) with GenoLyse® + post purification
02:113	9.00	2.96
02:292	8.90	3.46
02:302	7.52	2.08
03:013	9.45	4.68
03:039	7.59	0.95
03:303	7.37	1.24
03:313	7.52	Too low to be measured
04:018	8.04	1.32
04:194	9.05	1.91
04:198	8.22	1.71
04:211	7.68	0.75
04:493	8.67	1.55
04:503	8.32	3.10
05:046	7.21	1.34
07:116	9.11	3.27
07:118	7.83	2.52
Positive control (H37Rv)	9.52	3.56
Negative control (H_20)	0	0
Mean +/- SD	**8.3 +/- 0.7**	**2.1 +/- 1.2**
Average 260/280 ratio:	**0.8 +/- 0.2**	**1.4 +/- 0.1**

TABLE 15: DNA YIELD (NG/µL) OBTAINED WITH GENOLYSE® EXTRACTION KIT.

The extraction with GenoLyse® did not achieve the required concentration. The average 260/280 ratio was also below the minimum required ratio of 1.8, indicating a high level of protein or solvent contamination. The addition of a purification step with phenol:chloroform:isoamyl alcohol increased the purity of the DNA but it also significantly reduced the final DNA yield. The extraction was performed once when considering all samples. However, the extraction for samples 02.113 and 02.292 (first two samples in numerical order) was repeated twice but it still failed to achieve the concentration required (only the highest concentration is shown above). This was done to check for reproducibility but minimizing workload and costs.

INH-R strains (n=16)	DNA yield (ng/µl) with DNeasy®
02:113	3.89
02:292	Too low to be measured
02:302	6.64
03:013	1.21
03:039	Too low to be measured
03:303	2
03:313	Too low to be measured
04:018	2.8
04:194	1.54
04:198	0.58
04:211	13
04:493	3.28
04:503	Too low to be measured
05:046	0.6
07:116	3.17
07:118	2.32
Positive control (H37Rv)	5.81
Negative control (H_20)	0
Mean +/- SD	**2.7 +/- 3.3**
Average 260/280 ratio	**1.5 +/- 0.2**

TABLE 16: DNA YIELD (NG/µL) OBTAINED WITH DNEASY® EXTRACTION KIT.

The extraction with DNeasy® did not achieve the required concentration. The average 260/280 was 1.5, and below the minimum level required of 1.8, indicating protein or solvent contamination. As with the previous method, the extraction was performed once when considering all samples. However, the extraction for samples 02.113 and 02.292 (first two samples in numerical order) was repeated twice but it still failed to achieve the concentration required. This was done to check for reproducibility but minimizing workload and costs.

INH-R strains (n=16)	DNA yield (ng/µl) with CTAB method and 4 hours lysozyme incubation	DNA yield (ng/µl) with CTAB method and 24 hours lysozyme incubation
02:113	0.75	11.9
02:292	0.20	22.3
02:302*	0.039	1.06
03:013	0.13	52.8
03:039	0.32	11.5
03:303*	0.69	2.82
03:313	7.36	57.7
04:018	1.41	12.3
04:194	1.40	12.1
04:198*	0.34	1.99
04:211	1.14	10.1
04:493	2.16	12.9
04:503	0.57	62.2
05:046	0.99	16.7
07:116	0.28	16.2
07:118	0.83	28.1
Positive control (H37Rv)	10.4	49.2
Negative control (H_20)	0	0
Mean +/- SD	**1.7 +/- 2.8**	**22.4 +/- 20.1**
Average ratio 260/280:	**1.7 +/- 0.1**	**1.8 +/- 0.3**

TABLE 17: DNA YIELD (NG/µL) OBTAINED WITH THE CTAB EXTRACTION METHOD, COMPARING 4H AND 24H LYSOZYME INCUBATION.

The extraction with the CTAB method and 4h lysozyme incubation did not achieve the required concentration. The average 260/280 was 1.75, still below the minimum level required of 1.8, again indicative of potential protein or solvent contamination. When extending the incubation period with lysozyme to 24 hours, the DNA yield dramatically increased to reach satisfactory levels. The average 260/280 ratio was also within the acceptable level required by PHE. The extraction with 24h incubation was performed once for the majority of samples as they passed the quality check and further extraction was unnecessary. The samples marked with an asterisk failed to reach adequate DNA yield despite multiple extractions (only the result with the highest yield is shown).

4.3.2 Quality control and initial analysis

Once the desired DNA yield was achieved, the samples were sent for sequencing and returned as FASTQ files (Section 1.10.3). All FASTQ files passed the quality control stage, confirming good genome coverage. Mycrobe predictor software reported some isolates (02.292, 03.039, 04.018, 04.211, 04.493, 04.503 and 07.116) as isoniazid and quinolones resistant (Figure 10), in contrast with the phenotypic information already available from the reference laboratory. These clinical isolates with discrepant results were also run through the TB profiler software for additional confirmation and to understand the mutations involved (Figure 11). A non-synonymous mutation affecting the gyrase gene (*gyrA*) and replacing the amino acid alanine with glycine in position 90 (*Ala90Gly*) was responsible for the alleged quinolone resistance. However, such mutation is known for involving the *gyrA* gene without conferring resistance; it can actually increase susceptibility to some fluoroquinolones (Malik, Willby, Sikes, Tsodikov, & Posey, 2012). The interpretation of this mutation as conferring resistance was a clear mistake due to the database not being updated.

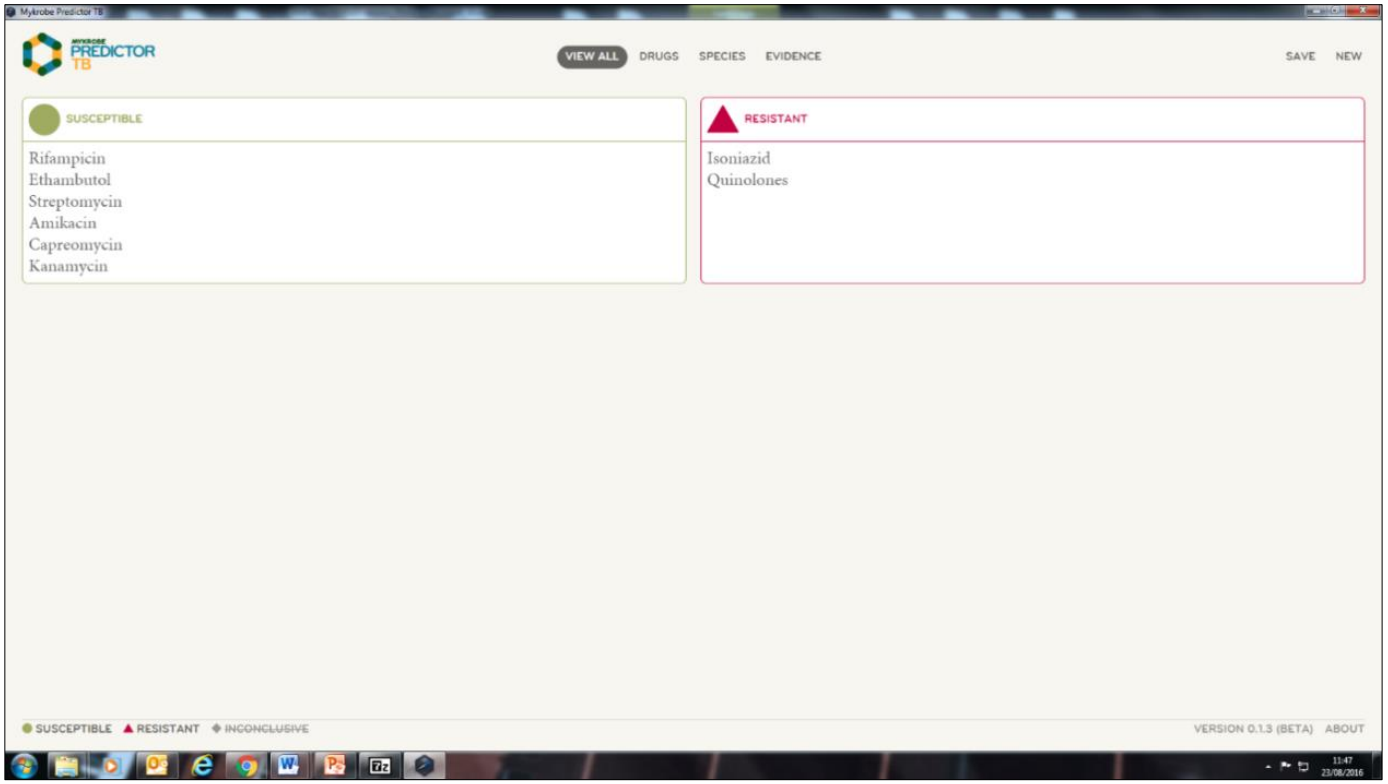
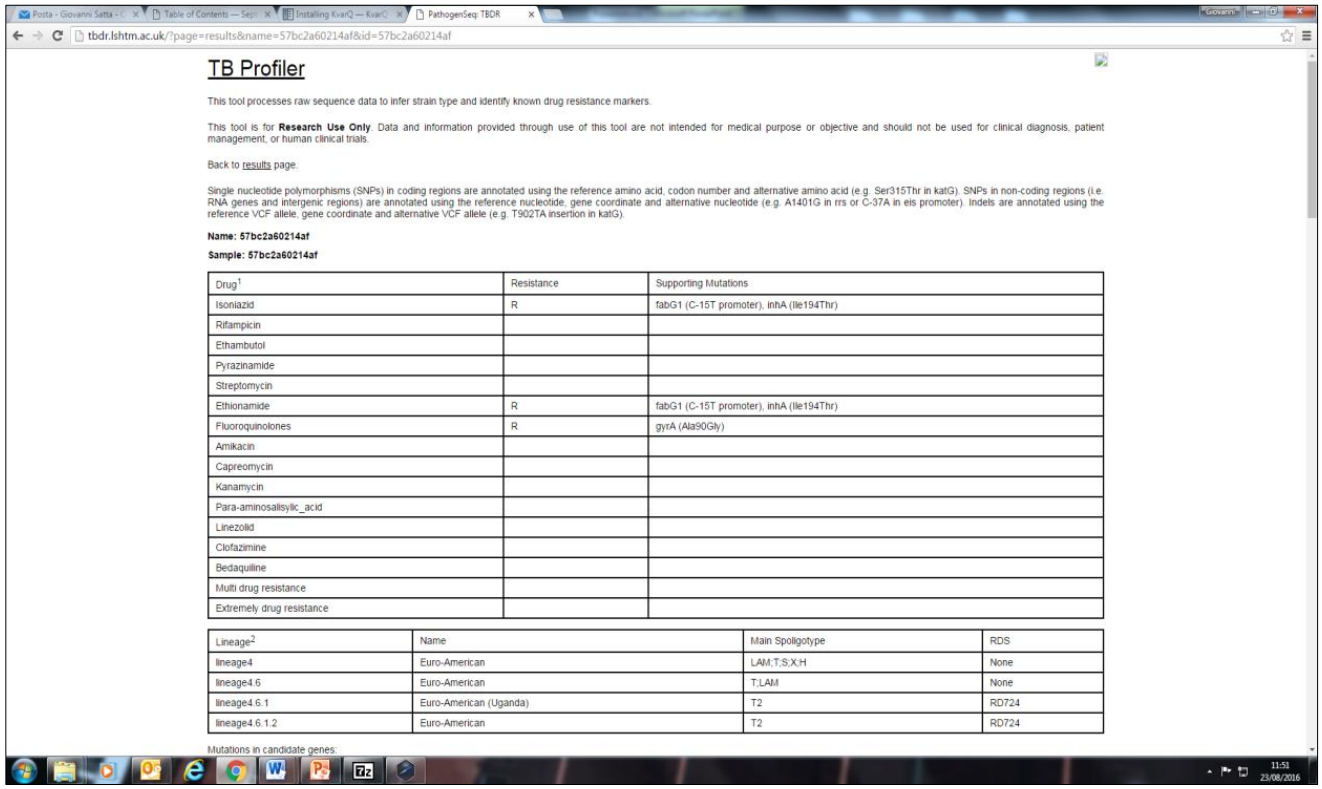


FIGURE 10: EXAMPLE OF REPORT USING MYKROBE PREDICTOR.

Mykrobe predictor report for sample 07.116 as an example (known isoniazid and ethionamide resistant). The software presents a user-friendly interface where susceptible and resistant drugs are reported based on genotypic information. The resistance to quinolones contrasts with the phenotypic information available from the reference laboratory. This software does not provide information about ethionamide resistance.



TB Profiler

This tool processes raw sequence data to infer strain type and identify known drug resistance markers.

This tool is for **Research Use Only**. Data and information provided through use of this tool are not intended for medical purpose or objective and should not be used for clinical diagnosis, patient management, or human clinical trials.

Back to results page.

Single nucleotide polymorphisms (SNPs) in coding regions are annotated using the reference amino acid, codon number and alternative amino acid (e.g. Ser315Thr in katG). SNPs in non-coding regions (i.e. RNA genes and intergenic regions) are annotated using the reference nucleotide, gene coordinate and alternative nucleotide (e.g. A1401G in rrs or C-37A in eis promoter). Indels are annotated using the reference VCF allele, gene coordinate and alternative VCF allele (e.g. T902TA insertion in katG).

Name: 57bc2a60214af

Sample: 57bc2a60214af

Drug[1]	Resistance	Supporting Mutations
Isoniazid	R	fabG1 (C-15T promoter), inhA (Ile194Thr)
Rifampicin		
Ethambutol		
Pyrazinamide		
Streptomycin		
Ethionamide	R	fabG1 (C-15T promoter), inhA (Ile194Thr)
Fluoroquinolones	R	gyrA (Ala90Gly)
Amikacin		
Capreomycin		
Kanamycin		
Para-aminosalisylic_acid		
Linezolid		
Clofazimine		
Bedaquiline		
Multi drug resistance		
Extremely drug resistance		

Lineage[2]	Name	Main Spoligotype	RDS
lineage4	Euro-American	LAM;T;S;X;H	None
lineage4.6	Euro-American	T;LAM	None
lineage4.6.1	Euro-American (Uganda)	T2	RD724
lineage4.6.1.2	Euro-American	T2	RD724

Mutations in candidate genes:

FIGURE 11: EXAMPLE OF REPORT USING TB PROFILER.

Example TB profiler report (sample 07.116, known isoniazid and ethionamide resistant). Isoniazid and ethionamide resistance are correctly reported but the isolate is erroneously classed as having fluoroquinolones resistance based on genotypic information (mutation in the *gyrA – Ala90Gly*).

4.3.3 Genetic analysis

Phylogenetic reconstruction (Figure 12) showed that all the clinical isolates, with the exception of 04.194, cluster together as part of RFL15. In particular, outbreak samples 02.292, 03.039, 04.018, 04.211, 04.493, 04.503 and 07.116, isolated over 6 years. Samples 04.018 and 07.116 also have additional resistances. Other isolates, including 05.046, 02.113 (both drug sensitive), 03.013 and 03.313 (both streptomycin mono-resistant only), diverge from the main group. The control samples (05.177, 05:094 and 04.011) are separate and independent strains.

Deletions – Based on the phylogenetic tree results, comparative analysis for the detection of deletions was initially performed between selected outbreak isolates (02.292, 03.039, 04.018, 04.211, 04.493, 04.503 and 07.116), the control sample 05.177 (with the same *inhA* mutation) and the outbreak isolate 03.313 (streptomycin mono-resistant). The isolates were originally chosen as being closely related and to prevent further genetic variation due to strain diversity. INH-R clinical samples demonstrated extensive deletions in sixteen genes compared with the control isolate used (05.177, still INH-R) (Table 18). Inclusion of sample 03.313 reduces the deleted gene set to 13 genes (the list of genes and their functional relevance is explained in Table 18 – BLAST ring is showed in Figure 13).

SNPs – Comparative analysis was performed between the same outbreak isolates (02.292, 03.039, 04.018, 04.211, 04.493, 04.503 and 07.116) and the control sample 05.177 for the detection of SNPs. A total of 563 SNPs were identified. These were compared to a classification of *M. tuberculosis* virulence factors (Forrellad, Klepp, & Gioffré, 2013) and 33 virulence genes have been identified as affected by at least one SNP (Tables 19, 20 and 21).

Insertions – Comparison between the outbreak isolate 04.211 and the reference strain H37Rv did not reveal the presence of any insertion (data not shown). The isolate 04.211 was chosen as the most representative for the cluster based on the phylogenetic tree results.

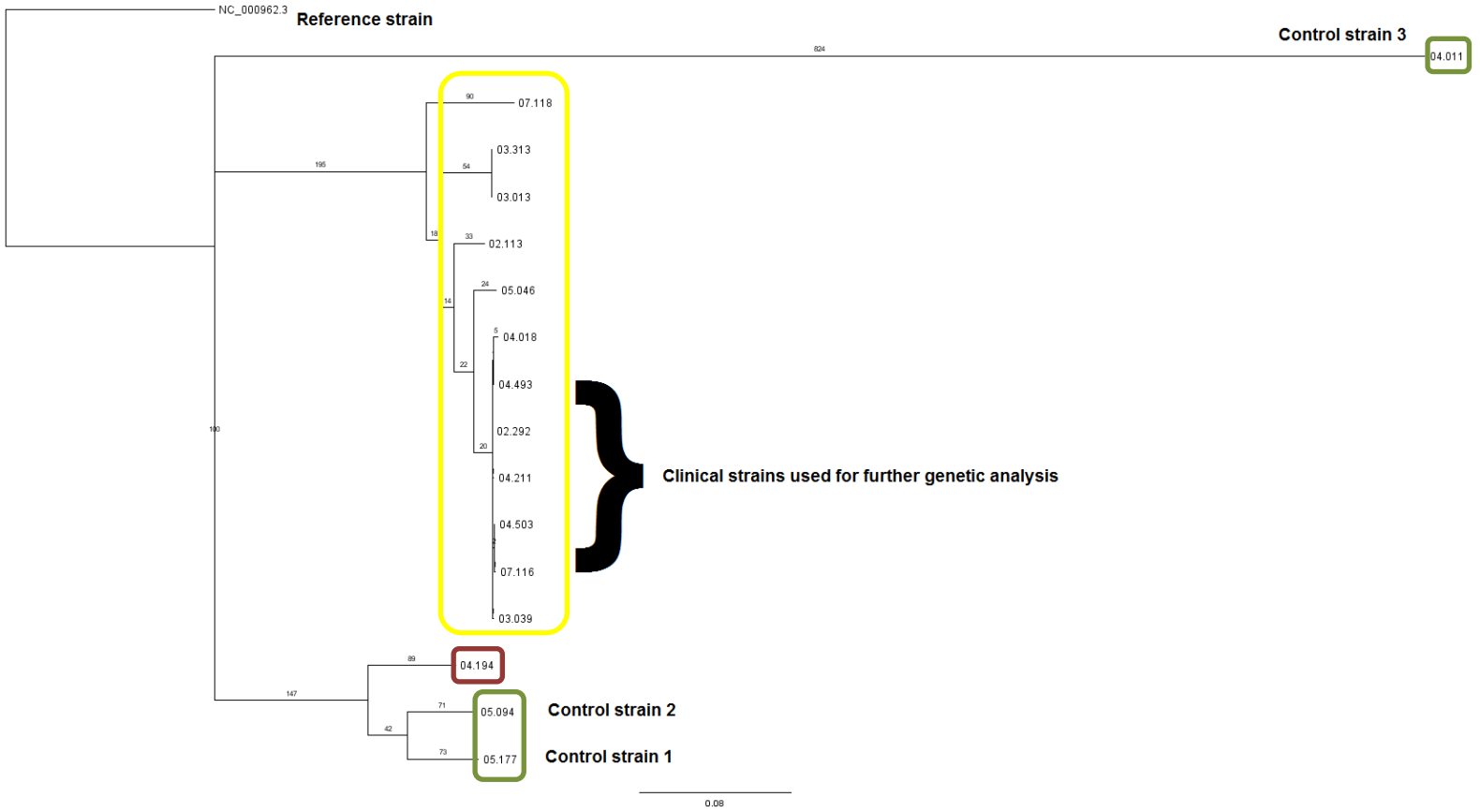


FIGURE 12: PHYLOGENETIC RECONSTRUCTION OF THE OUTBREAK AND CONTROL ISOLATES.

Phylogenetic reconstruction as described in Section 4.2.4. All clinical outbreak samples (circled in yellow), with the exception of 04.194 (circled in red), grouped together as part of Cluster RFL15. The control samples (05.177, 05:094 and 04.011 – circled in in green) diverge as separate and independent strains. All isolates are Euro American lineage except for 04.011 (Control strain 3) which is East African Indian.

Gene name and function	Isolates: 02.292, 03.039, 04.018, 04.211, 04.493, 04.503, 07.116	03.313	05.177 Control
Phirv1	✘	✘	✓
Rv1673c (hypothetical protein)	✘	✘	✓
Rv1675c (cAMP and macrophage regulator)	✘	✘	✓
plcD (phosphoesterase)	✘	✘	✓
Wag22 (antigen member, PE family)	✘	✘	✓
Rv1760 (triacylglycerol synthase)	✘	✘	✓
Rv1761c (unknown protein)	✘	✘	✓
Rv1762c (unknown protein)	✘	✘	✓
plcB (phosphoesterase)	✘	✓	✓
plcA (membrane associated phospholipase)	✘	✓	✓
PPE54 (part of PPE family)	✘	✓	✓
PE_PGRS50 (antigen member, PE family)	✘	✘	✓
PPE55 (part of PPE family)	✘	✘	✓
Rv3349c (probable transposase)	✘	✘	✓
Rv3371 (possible triacylglycerol synthase)	✘	✘	✓
Rv3486 (membrane protein, function unknown)	✘	✘	✓

TABLE 18: LIST OF GENES WITH COMPLETE DELETION IN THE INH-R OUTBREAK CLUSTER.

The symbol ✗ indicates that the gene contains extensive deletions, whilst the symbol ✓indicates that the gene is still present. Deletions were found in 16 genes in the INH-R outbreak isolates compared with the control isolate 05.177. Comparison is also made with the phylogenetically related isolate 03.313 (only streptomycin mono-resistant), reducing the total common deletions to 13 genes (highlighted in grey).

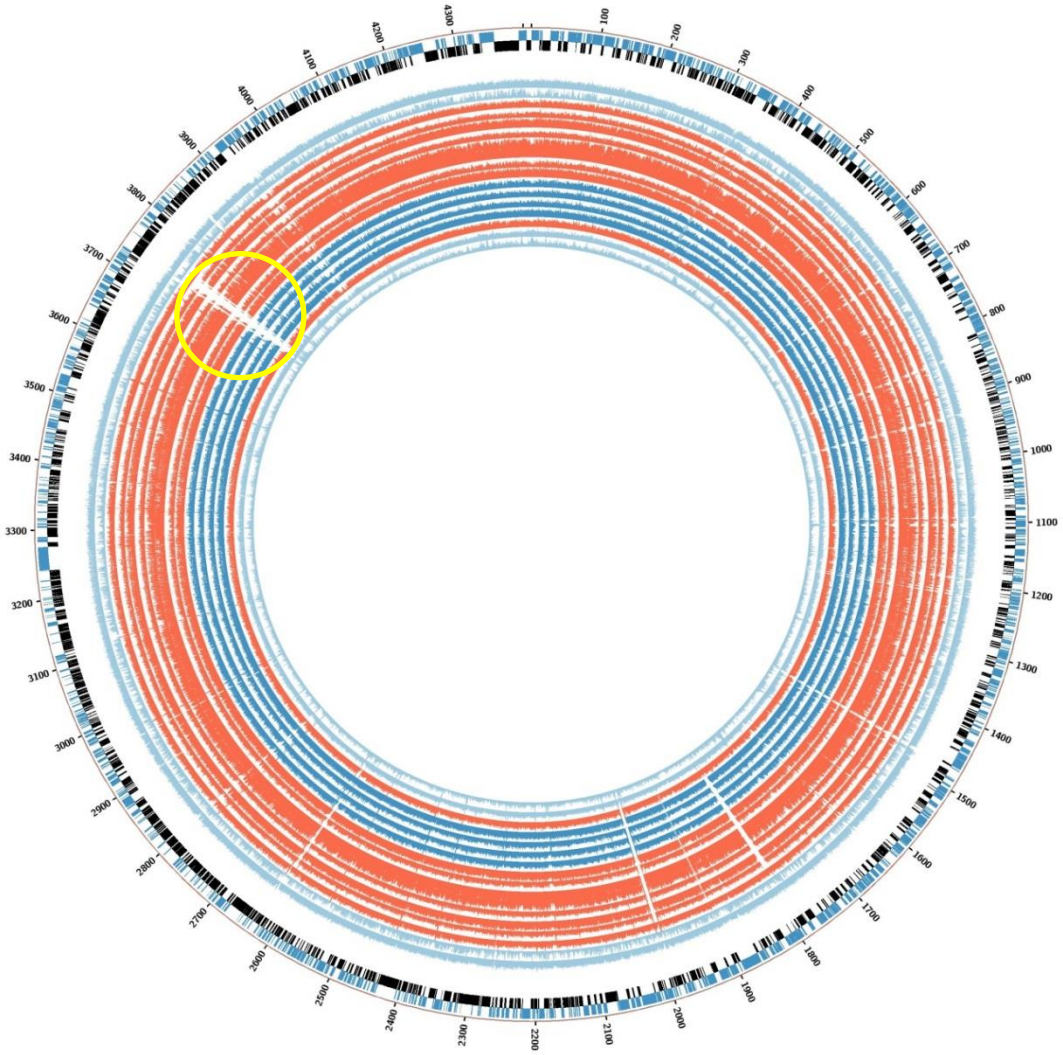


FIGURE 13: BLAST RING FOR THE GRAPHICAL REPRESENTATION OF DELETED GENES IN INH-R OUTBREAK.

Deleted regions can also be represented using a BLAST ring (white empty spaces). The most evident deletion in the *plc* genes is circled in yellow. The order is the same as for the phylogenetic tree (from outside in). Samples 02.292, 03.039, 04.018, 04.211, 04.494, 04.503, 07.116 (all closely related) are in red. Samples 05.046, 02.113, 03.013 and 03.313 are coloured deep blue whilst controls samples (05.177, 05.094 and 04.011) are light blue as *H37Rv*. Samples 04.194 is also shown in red.

Difference in SNPs between isolate 03.039 and other outbreak isolates									
Isolate 03.039									
SNPs difference	7	41	79	118	160	501	563	1096	629
Other outbreak isolates	**04.018**	**05.046**	**02.113**	**03.313**	**07.118**	**04.194**	**05.177 Control strain 1**	**04.011 Control strain 3**	**H37Rv**

TABLE 19: **DIFFERENCE IN SNPS BETWEEN ISOLATE 03.039 AND OTHER OUTBREAK ISOLATES**

The table above shows the difference in SNPs between isolate 03.039 and other outbreak isolates/control strains. From the phylogenetic tree (Figure 12), isolates 02.292, 03.039, 04.018, 04.211, 04.493, 04.503 and 07.116, all cluster together. The maximum difference within this sub-cluster is between isolates 03.039 and 04.018 as shown above (7 SNPs). The seven SNPs are due to different susceptibility profiles as 03.039 is only isoniazid-resistant, whilst 04.018 is isoniazid, rifampicin, clarithromycin and ethionamide resistant. The number of SNPs increases progressively to a maximum of 1096 when comparing 03.039 to 04.011. The choice was made of limiting the SNPs comparison to the control strain 05.177 (in grey) as all other additional SNPs (versus 04.011 and H37Rv) are simply due to lineage and strain variation and not relevant to the outbreak strains.

Gene name	SNP Position	Function	Result
ctpA	100895	Cation transporting P-type ATPase	Reduced CFUs*
mce1F	206453	Mce protein	Reduced tissue pathology and increased survival
Rv0176	208299	Mce protein	Reduced CFUs, reduced tissue pathology and increased survival
pckA	252105	Iron-regulated phosphoenolpyruvate carboxykinase	Reduced CFUs
pcaA	561021	Mycolic acid cyclopropane synthase	Reduced CFUs in lung
pstS1	1042991	phosphate-binding lipoprotein component of inorganic phasphate transport system	Reduced multiplication
mycP5	2033748	Pro-rich membrane-anchored serine protease (mycosin)	Reduced CFUs
Rv3083	3448567	Probable monooxygenase	Reduced CFUs
ponA2	4122882	Penicillin-binding protein	Moderate reduction in CFUs
eccE2	4367659	Type VII secretion system protein	Reduced CFUs
*CFUs = Colony Forming Units			

TABLE 20: LIST OF GENES ENCODING FOR KNOWN VIRULENCE FACTORS AND AFFECTED BY NON-SYNONYMOUS SNPS.

The table above shows the genes encoding for virulence factors and affected by a non-synonymous SNPs. The known gene function, SNP position and the presumed result caused by the mutation are also shown. These SNPs seem to reduce the CFU and increase the mycobacterial survival within the human host.

Gene name	Position	Function	Result
ctpI	130449	Cation transporting P-type ATPase	Reduced CFUs
mprB	1098698	Two component sensor kinase	Reduced CFUs in lung latent stage
oppA	1433114	Oligopeptide-binding lipoprotein of peptide transport system	Reduced CFUs and increased survival
oppD	1434648	Oligopeptide-transport ATP binding protein	Reduced CFUs in organs in the chronic infection, increased survival
irtA	1514732	Iron-regulated transporter	Reduced CFUs in macrophages and lung
irtB	1515922	Iron-regulated transporter	Reduced CFUs in macrophages and lung
Rv1410c	1586360	Aminoglycoside/tetracycline-transport membrane protein	Reduced CFUs
pks5	1726541	Polyketide synthase	Reduced CFUs in organs
eccB5	2017898	EsX-5 type VII secretion system protein	Reduced CFUs
eccD5	2032701	EsX-5 type VII secretion system protein	Reduced CFUs
katG	2155389	Catalase peroxidase peroxynitritase T	Reduced CFUs, resistance to INH
Rv1931c	2183054	Probable transcriptional regulatory protein	Reduced CFUs
fadE18	2184781	Acyl-CoA dehydrogenase	Reduced CFUs
mce3A	2210055	Mce protein, possible virulence factor	Reduced tissue pathology and increased survival
dosT	2273627	Histidine kinase response regulator	Moderate reduction CFUs
pks12	2294876	Polyketide synthase	Reduced CFUs
pafA	2355511	Proteasome accessory factor A	Reduced CFUs in organs and less tissue pathology
mpa	2375883	Proteasone ATPase	Reduced CFUs in organs and less tissue pathology
nuoG	3518089	NADH dehydrogenase	Increased animal survival
Rv3236c	3612571	Probable conserved integral membrane transport protein	Reduced phagosome ROS production
mmpL8	4291134	Predicted drug exporter of the RND superfamily	Reduced CFUs in organs
esxA	4352875	Early secretory antigenic target	Reduced CFUs
	*CFUs = Colony Forming Units		

TABLE 21: LIST OF GENES ENCODING FOR KNOWN VIRULENCE FACTORS AND AFFECTED BY SYNONYMOUS SNPS.

The table above shows the genes encoding for virulence factors and affected by a synonymous SNPs. It is difficult to assess the significance of synonymous SNPs. Even if they may not cause an amino-acid/protein alteration, they may still have an impact on messenger RNA splicing, stability, and structure as well as protein folding (Hunt, Sauna, Ambudkar, Gottesman, & Kimchi-Sarfaty, 2009).

4.4 Discussion

4.4.1 Extraction of genomic DNA

Efficient DNA extraction methods are essential to obtain high-quality DNA for next generation sequencing applications. Current methods employ different techniques to liberate DNA, each with inherent advantages and disadvantages. To generate high yield DNA of sufficient quality for WGS in this thesis, three different methods of DNA extraction were tested to find the optimal procedure and they were assessed based on the quantity of genomic DNA obtained and the ratio 260/280. Running the DNA on an agarose gel was not required by the sequencing provider.

GenoLyse® – This commercial kit is rapid (approximately 20 minutes) and easy to use. However, for WGS to be successful, PHE guidelines (GSDU user manual/PHE publications gateway number: 2015620) recommend a minimum DNA concentration of 10 ng/mL and a minimum 260/280 ratio of 1.8. The GenoLyse kit was unable to provide a DNA yield of sufficient quality and quantity for whole genome sequencing which met PHE's requirements. In order to augment yield and quality, an additional phenol:chloroform:isoamyl alcohol purification step was performed on the samples. Despite resulting in an increase in quality of the DNA, as measured by 260/280 ratios, yields were substantially lower when using this additional purification step. As such, the GenoLyse kit, even with additional purification steps, did not meet our requirements. This extraction method would be more suitable for other PCR-based applications, where a rapid extraction is needed and lower quality DNA can be tolerated. Nevertheless, it was worth evaluating its potential application to mycobacterial DNA extraction and next generation sequencing as there is no literature on this subject.

DNeasy® – This extraction kit produced very low DNA concentrations with yields for some isolates being beyond the lower limit of detection for the Qubit. This finding raises the concern, already expressed by other authors (Healey et al., 2014), that bigger fragments of DNA can potentially get lost through subsequent column washes or precipitations. However, these results were in contrast with Brown et al. (Brown et al., 2015) who successfully used this commercial kit in their extraction directly from clinical samples. However, on closer inspection in *Brown et al.* the kit still failed to provide any DNA in 17% of their samples. Moreover, they did not perform WGS on the purified DNA directly, instead pre-enriching the DNA with RNA baits prior to sequencing. This would suggest that the DNeasy kit yielded poorly concentrated DNA

for their sequencing and corroborates our findings on the use of this kit in the current study.

CTAB based extraction – This method proved to be the optimal method for purifying mycobacterial DNA and it provided high quality DNA at a viable concentration when prolonged incubation with lysozyme (24 hours) was used. Lysozyme is known to have antibacterial effects against mycobacteria (Kanetsuna, 1980) and is widely used in cellular permeabilization protocols. As such, it was not surprising that we observed an increase in DNA yield after prolonged lysozyme exposure. The presence of two detergents in this protocol is likely to further augment the DNA yield. CTAB binds to polysaccharides, polyphenols and other secondary metabolites so they can be separated from the DNA. In particular, the CTAB binds to the polysaccharides when the salt concentration is high and this explains the inclusion of NaCl in the procedure (Rogers & Bendich, 1985). Conversely, proteinase K and SDS are able to liberate and digest proteins, thereby preventing proteolytic DNA degradation by any nucleases. DNA is further purified with phenol:chloroform:isoamyl alcohol. The phenol/chloroform mixture further helps to denature proteins and facilitates the separation of the aqueous and organic phases, whilst the isoamyl alcohol reduces foaming during the extraction process. Finally, ethanol allows the DNA to precipitate and be separated from the rest of the solution. Isopropanol can also be used instead of ethanol but this may result in a yellowish and slimy pellet, difficult to re-suspend (Kaser et al., 2010) and thus it was avoided in this thesis. The cumulative action of all these reactions likely explain the excellent results obtained with this method but they also contribute to its extended duration of two days (compared with commercial kits).

Alternative extraction methods have been proposed by other authors. In particular, some authors used mechanical disruption as an adjunct step to cause cell lysis (Warren et al., 2006) (Belisle & Sonnenberg, 1998). However, they did not publish any data on the use of such extracted DNA for next generation sequencing applications. In this thesis, a mechanical disruption step was used in a small number of samples and it did not provide satisfactory results (more contaminated DNA compared with the traditional CTAB based method). More recently, a new extraction protocol including a combination of mechanical disruption, ethanol precipitation and DNA purification with AMPure XP solid-phase reversible immobilization beads (Beckman Coulter, UK) has been proposed as an alternative to the CTAB based method and directly from positive MGIT cultures (Votintseva et al., 2015). With increased availability and uptake of

WGS, it is likely that extraction protocols for the production of high yield and quality of DNA will follow suit.

4.4.2 Genetic analysis

Previous epidemiological studies have described the evolution of INH-R strains to MDR-TB, via the development of extra resistance to rifampicin and other drugs. In particular, it was observed that only strains with the *KatG* S315T substitution were associated with successful transmission and development of further phenotypic resistance (Hu, Hoffner, & Jiang, 2010) (Gagneux, Burgos, et al., 2006). However, in the RFL15 cluster, *inhA C-T767* is the predominant mutation. The fitness and mutation rate of the resistant clinical isolates were not affected when compared to the reference strain H37Rv and indeed the fitness was significantly higher when compared to INH-R laboratory mutants (Chapter 3). This indicates that if there were any fitness cost initially associated with the acquisition of resistance-conferring mutations then this was either very small, or these clinical isolates have compensated for it.

The application of WGS has allowed an in depth genetic analysis of the outbreak isolates. At the phylogenetic level, it is interesting to note that sample 04.194 does not seem to belong to RFL15 as previously reported based on MIRU and RFLP typing data although it was considered an outlier and partially divergent as it carries the *katG* mutation instead of the *inhA*. This supports the view that whole-genome sequencing offers a more precise means to delineate outbreaks (T. M. Walker, Ip, Harrell, et al., 2013).

The outbreak isolates show genetic variations with unique deletions and SNPs without additional insertions. Of the 13 genes identified as deleted, eight are conserved hypothetical proteins whose function is still unknown. These can be considered to involve non-essential genes. Nevertheless, they could potentially offer the advantage of escape from the host immune response and explain the reason why these strains remain fixed in the community prolonging the outbreak for years.

Some deletions are worth further attention and may confirm the hypothesis of escaping/reducing the host immune response. Of note are:

Cmr (*Rv1675c*) – This is a transcription factor known to be responsive to cAMP levels, and implicated in the biology of persistent TB infection (Ranganathan, Bai, & Lyubetskaya, 2016). It is a regulator of four different proteins (*mdh*, *groEL2*, *Rv1265*

and and *PE_PGRS6a*) during macrophage infection by *M. tuberculosis* and they are likely to play a role in host interactions (Gazdik, Bai, Wu, & McDonough, 2009). More recently (L. J. Smith et al., 2017), it has been demonstrated that the *cmr* gene can directly repress the DosS/T-DosR system. This is a two-component system able to respond to changes in redox state, hypoxia, nitric oxide and carbon monoxide (Kumar, Toledo, Patel, Lancaster, & Steyn, 2007). The DosR regulon is a critical component of tuberculosis pathogenesis and it is required for survival during anaerobic conditions of dormancy. For this reason, it is thought to be involved in establishing latent TB infections (Voskuil et al., 2003). In laboratory experiments, a mutant with a reduced *cmr* gene expression was able to better survive a nitrosative stress challenge (compared to H37Rv and complemented mutants). The *cmr* mutant was also transiently attenuated in a mouse aerosol infection model but it was unaffected in *in-vitro* infections of mouse bone marrow-derived macrophages (L. J. Smith et al., 2017). It can be postulated that the deletion affecting the *cmr* gene in the INH-R clinical isolates could have contributed to the overexpression of the DosR-regulon, thus allowing these strains to better establish latent infection.

Plc genes – The *M. tuberculosis* genome contains four different genes from this family (*plcA, plcB, plcC, plcD*) which encode for the enzyme *phospholipase C* (PLC). This region frequently contains deletions (Raynaud, Guilhot, & Rauzier, 2002) and it is considered to be involved in different functions related to virulence. Firstly, it may provide mycobacteria with nutrients by using fatty acids as a major energy source. Secondly, PLC may degrade the phagosomal membrane by changing its permeability or interrupting its function (Wheeler & Ratledge, 1992). Thirdly, PLC might release arachidonic acid metabolites from macrophage membranes and in this way interfering with signal transmission and the host immune response (Titball, 1993). This is could explain the role of PLC in chronic infections and the reduced growth (10 fold less) of *plc* deficient mutants in the late phase of infection in the mice model (Raynaud et al., 2002). PLC is associated with many different pathways and its mutations may allow *M. tuberculosis* to survive inside macrophages and sustain chronic infection (Viana-Niero, Haas, Soolingen, & Leão, 2004) (Yang, Yang, & Kong, 2005).

It is difficult to interpret the real role of all the 563 SNPs. 40% of the genes in the *M. tuberculosis* genome have had their function characterised, whilst another 44% have been proposed to have possible functional relevance. Previous authors (Fraser, Eisen, & Fleishmann, 2000) originally claimed a frequency of polymorphism of about 1/3000 bp (more than 1000 SNPs for the whole genome) thanks to detailed

comparative studies between H37Rv and CDC1551 strains. However, subsequent authors (Musser, 2001) studied 24 different genes encoding target proteins for the immune response of 16 different isolates of *M. tuberculosis*: among these, 19 genes were unvaried and just six nucleotide polymorphism sites were identified in the five genes where variation occurred. They estimated therefore an overall frequency of SNPs of about 1/10,000 bp (around 400 SNPs for the whole genome) and this rate was also confirmed by other earlier experiments (Ramaswanny, Amin, & Goksel, 2000). Additionally, some of the SNPs in our clinical isolates caused mutations in thirty-two virulence genes and their function is summarized in Table 17. However, only 10 genes (out of the thirty-two) have a nonsynonymous SNP (when the amino acid is altered), whilst the remaining genes were involved by a synonymous SNP (not causing a change in the amino acid). Overall, these are reported to cause a decreased capacity to growth (reduction of the colony forming unit – CFU), phagosome formation and increase survival - thus allowing persistence in the human host. This further confirms the hypothesis that these isolates have unique biological characteristics that serve to prolong this outbreak with the ability to persist in the host, potentially evading the immune response and allowing transmission to contacts (as confirmed by epidemiological data).

Additionally, it can be hypothesised that these virulence genes are somehow essential for *M. tuberculosis* (as they have been affected by single mutations rather than deletions). *CtpA* gene is essential for mycobacterial intracellular survival and the transport of heavy-metal ions (Leon-Torres, Novoa-Aponte, & Soto, 2015), whilst other genes (*mce1F*, *mycP5* and *pcaA*) have been associated with important regulatory and secretion functions or the prevention of phagosome maturation (Shimono et al., 2003) (Chen, 2016) (Corrales et al., 2012). This would make them an attractive target for the development of new drugs and further work (e.g. knock-out mutants) is necessary to confirm such hypothesis. Supporting this, one of those virulence genes (*pckA*) has already been proposed as potential target for new compounds (Marrero, Rhee, Schnappinger, Pethe, & Ehrt, 2010).

In conclusion, analysis of the genetic variations of INH-R clinical samples from the London outbreak suggests that this outbreak consists of successful, closely related, circulating strains with heterogeneous resistance profiles and mutations and little or no associated fitness cost or impact on their mutation rate. Deletions and SNPs may be a peculiar feature of these isolates and can potentially explain the persistence of this lineage in the community and the prolongation of the outbreak for years. The

genetic variation in clinical strains may provide potential drug targets for the development of new compounds.

Chapter 5: The role of whole genome sequencing in characterising the mechanism of action of TB drugs: demonstrated with PAS and analogue

5.1 Introduction

Para-aminosalicylic acid (PAS), also known as 4-aminosalicylic acid, was one of the first antibiotics to be used against tuberculosis (TB) and it is still one of the last remaining drugs available to treat extensively drug-resistant (XDR) disease. PAS was first synthesised in 1902 by Seidel and Bittner and then re-discovered as an antituberculous agent by Lehmann in 1943 (Lehmann, 1946). At the time of its introduction as treatment against TB, more than sixty preparations of PAS were available in USA (Bogen, Loomis, & Will, 1950), including aminosalicylic acid and various salts (sodium, potassium and calcium PAS). After ethambutol, rifampicin and pyrazinamide were introduced, the use of PAS fell so dramatically that PAS was almost unobtainable when the HIV and multi drug resistant (MDR) TB epidemics started in the 1980s. A new gastro-resistant formulation was released in 1994 (Jacobus Pharmaceutical Company, Princeton, NJ, USA) and it is the formulation currently used globally (Kibleur, Brochart, Schaaf, Diacon, & Donald, 2014). The PAS story highlights the importance of revisiting old drugs in the fight against drug resistant TB (Donald & Diacon, 2015).

PAS is generally regarded as a bacteriostatic agent (Lehmann, 1946), but some authors have demonstrated its bactericidal effects on metabolically active populations of bacilli (Xie, Siddiqi, & Rubin, 2005). This early bactericidal effect is likely to be dose-related and is particularly evident in the first two days of treatment (Jindani et al., 1980). Several clinical studies performed in the early days of clinical use showed the presence of a dose-related response to PAS and prevention of resistance with a single dose (Marsden, 1954) (Singh & Mitchison, 1955). A paper from the Medical Research Council (MRC) summarizing various clinical trials concluded that "*the height of peak serum concentrations may be a more important determining factor in preventing drug resistance during combined therapy than the maintenance of a high average concentration"* (MRC, 1955). This dose-dependent effect would be similar to the one achieved by aminoglycosides in contrast with a time-dependent effect of many other antibiotics (Conly, Gold, & Shafran, 1994).

Despite being on the market for decades, the mechanism of action of PAS is not completely understood yet. It was originally proposed that PAS had an impact on oxygen uptake by mycobacterial cells (Bernheim, 1940) but recent studies seem to highlight that the principal antitubercular action occurs through the alteration of bacterial folate metabolism. It has been proposed that, being an analogue of para-amino benzoic acid (PABA), PAS competes with PABA for dihydropteroate synthase, interfering in the process of folate synthesis (Minato et al., 2015). However, PAS may have other biological impacts and this could explain the uncertainty around its mechanism of action. There is additional evidence that PAS can also interfere with other metabolic pathways and in particular with the mycobacterial iron acquisition, causing disruption to cell replication (Ratledge & Brown, 1972) (Nagachar & Ratledge, 2010). Thus, the mechanism of action of PAS remains unclear.

Resistance to PAS was described shortly after its introduction into clinical use (Turnbull, Wallace, Stewart, & Crofton, 1953). Recent studies have demonstrated that PAS resistance in *M. tuberculosis* is probably caused by multiple mechanisms. A study using transposon mutagenesis identified mutations in the *thyA* gene that were also present in clinical isolates resistant to PAS (Mathys et al., 2009). The gene *thyA* encodes for a thymidylate synthase enzyme (essential for DNA replication and repair) and its deletion has been demonstrated to confer resistant to PAS (Fivian-Hughes, Houghton, & Davis, 2012). Other studies have also identified various missense mutations in *folC* (encoding a dihydrofolate synthase) and *ribD* (alternative dihydrofolate reductase) that conferred resistance to PAS in laboratory and clinical isolates of *M. tuberculosis* (Zhao et al., 2014) (X. Zhang et al., 2015). Nevertheless, mutations in *folC* were detected in only 34.8% of clinical isolates, whilst mutations of *thyA* and *ribD* were detected in 26.0% and 5.8%, respectively (X. Zhang et al., 2015). Hence, other mechanisms of resistance to the drug might exist. Efflux pumps have also been described conferring cross-resistant to PAS and other antibiotics including streptomycin (Ramon-Garcia et al., 2012). Whole genome sequencing (WGS) could potentially help in elucidating the mechanism of action/resistance through the whole genome analysis of PAS-resistant laboratory mutants.

Our collaborators at University of Hull (Dr Andrew Boa, Department of Chemistry) provided us with sixteen new compounds to test against *M. tuberculosis*. These compounds were created in the laboratory as PAS analogues (based on their chemical structures – Figure 14) and their antimycobacterial activity has never been tested before.

The main aim of this chapter was to test these new PAS analogues against *M. tuberculosis* and to understand their mechanism of action. The primary hypothesis that these new compounds have antimycobacterial activity was tested by using two susceptibility methods. The most promising compound, AD25a, was then selected for further analysis. The secondary hypothesis is that this new compound has the same mechanism of action of PAS. This was tested by creating laboratory resistant mutants against PAS and AD25a and by performing whole genome sequencing on all of them. The comparative analysis of the mutations provides further insight into the mechanism of action/resistance of PAS and/or AD25a.

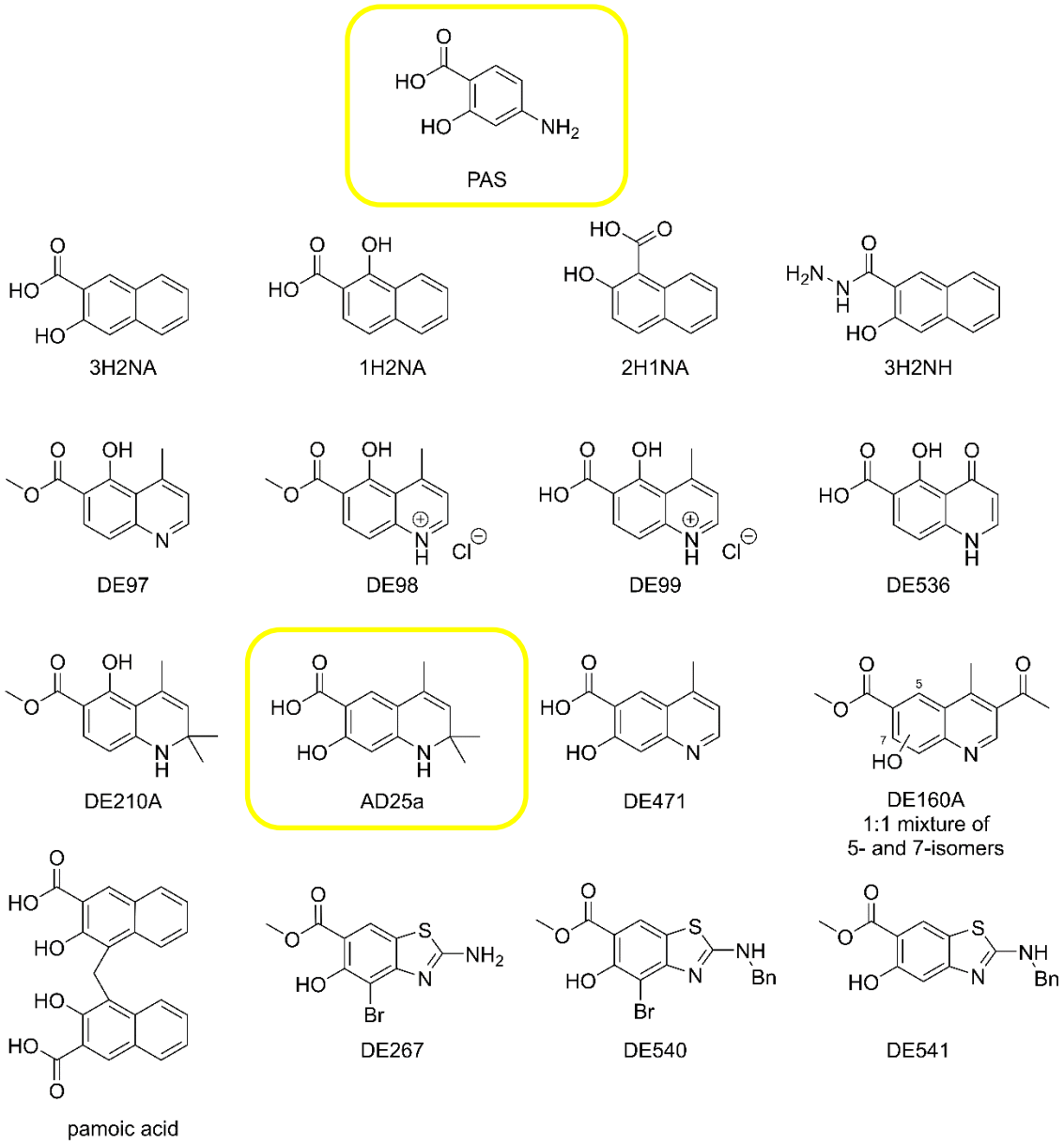


FIGURE 14: STRUCTURE OF PAS AND PAS-ANALOGUES

Chemical structures of PAS and its analogues. PAS ($C_7H_7NO_3$) is an amino derivative of salicylic acid and its main structure is easily resembled in the analogues shown above. The most promising compound (AD25a) was selected for further analysis. PAS and AD25a are highlighted in yellow.

5.2 Materials and methods

Sixteen different compounds were provided by our collaborators at the University of Hull (Dr Andrew Boa, Department of Chemistry). They also provided PAS to be included as a control in all experiments and for the creation of PAS-resistant mutants. Compounds were received as dry powder and were diluted in 0.5 ml of sterile water. Non-water-soluble compounds were initially dissolved in a small amount of DMSO (dimethyl sulfoxide) and then diluted in water. The solutions were sterilized through membrane filters (0.2 µm pore size) to prevent bacterial or fungal contamination.

5.2.1 Spot culture screening

The spot culture method was selected as a rapid screening test and the initial evaluation was performed using the maximum possible concentration of each compound (neat). The procedure has been described in Section 2.3.1 and the experiments were performed in duplicates.

5.2.2 Resazurin testing

Additional resazurin testing was performed on those compounds that exhibited a preliminary activity on the spot culture method and to confirm the critical concentration (µg/ml) still able to inhibit mycobacterial growth. The method has been described in Section 2.3.2 and all experiments were performed in duplicates. The most promising compound based on the critical concentration results was then selected for further analysis.

5.2.3 BACTEC/MGIT testing

The critical concentration of the most promising compound was also tested using the BACTEC/MGIT method. This method has been described in Section 2.3.3.

5.2.4 Selection of resistant mutants

PAS resistant laboratory mutants were selected in the laboratory using the reference strain H37Rv and 7H10 medium with the addition of PAS at a concentration of 2 and 4 µg/ml (critical concentration for PAS is 2 µg/ml) and the new compound AD25a (at a concentration of 35 µg/ml). The protocol has been described in Section 2.4.

5.2.5 Whole genome sequencing analysis

All selected mutants were sub-cultured on Lowenstein-Jensen (LJ) slopes, DNA extracted with the CTAB method WGS analysis performed as previously described in Chapter 4 and Section 6.2.4.

5.3 Results

5.3.1 Spot culture screening

The initial spot culture results are shown in Table 18. Five PAS analogues (DE471, 1H2NA, 2H1NA, 3H1NA and AD25a) inhibited the growth of *M. tuberculosis*, whereas the remaining 11 compounds did not show any growth inhibition.

5.3.2 Resazurin testing

The five active compounds on the spot culture were further tested with the resazurin method for the determination of their critical concentration. The results are shown in Table 19. One compound in particular, AD25a, showed the lowest critical concentration (0.04 µg/ml) among the analogues.

5.3.3 BACTEC/MGIT testing

The compound AD25a was selected for further testing using the BACTEC/MGIT method. This additional method confirmed its antituberculous activity and its lowest active concentration was 0.39 µg/ml. The results are show in Table 20.

Compound	Maximum concentration used (µg/ml)	Growth/No growth of *Mycobacterium tuberculosis*
DE97	27	Growth
DE98	27.3	Growth
DE99	24	Growth
DE471	28.6	No growth
DE160A	32.6	Growth
DE210A	41	Growth
DE267(2)	39.6	Growth
DE540	26.6	Growth
DE536(2)	35.3	Growth
DE541	25	Growth
1H2NA	74	No growth
2H1NA	53.3	No growth
3H2NA	53	No growth
3H2NH	36	Growth
PAMOIC ACID	37.6	Growth
AD25a	10.8	No growth
PAS control	26.9	No Growth

TABLE 22: **SPOT CULTURE RESULTS FOR PAS ANALOGUES**

Spot culture screening showed five compounds inhibiting the growth of *M. tuberculosis* H37Rv. The active compounds (DE471, 1H2NA, 2H1NA, 3H1NA and AD25a) are highlighted in grey. The experiment was repeated twice with identical results.

Compound	Critical concentration
AD25a	0.04 µg/ml
DE471	1.79 µg/ml
3H2NA	26.5 µg/ml
1H2NA	74 µg/ml
2H1NA	53.3 µg/ml

TABLE 23: RESAZURIN RESULTS FOR PAS ANALOGUES

The resazurin method was used to calculate the critical concentration of the five active compounds on the spot culture. AD25a (highlighted in grey) showed the lowest critical concentration of the compounds tested. The experiment was repeated twice with identical results.

Growth Units and different concentrations of AD25a							
Time (hours)	**GC**	**1.54 µg/ml**	**0.77 µg/ml**	**0.39 µg/ml**	**0.19 µg/ml**	**0.1 µg/ml**	**0.05 µg/ml**
0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0
40	0	0	0	0	0	0	1
47	0	0	0	0	0	1	5
136	8	0	0	0	>100	>100	>100
142	29	0	0	0	>100	>100	>100
159	156	0	0	0	>100	>100	>100
168	251	0	0	0	>100	>100	>100
174	**400**	**0**	**0**	**1**	**>100**	**>100**	**>100**

TABLE 24: BACTEC/MGIT RESULTS FOR COMPOUND AD25A

The table shows the growth units of H37Rv over time (in hours) at different concentrations of AD25a (from 1.54 µg/ml to 0.05 µg/ml). When the growth control (GC) reaches 400 growth units, any concentrations with a growth unit less than 100 is considered as active against *M. tuberculosis* (susceptible). If the growth units are more than 100, the concentration tested is considered as non-active (resistant). The lowest active concentration was 0.39 µg/ml (circled in yellow).

5.3.4 Resistant mutants and WGS analysis

The experiment to create the PAS-resistant mutants had to be repeated four times. The first two times (with the standard inoculum of 10 mL) there was no growth of mutants. After doubling the inoculum to 20 mL, it was possible to select only one PAS mutant for each critical concentration (2 and 4 µg/ml). The experiment with 20 mL inoculum was repeated twice; hence it was possible to select four mutants in total. These resistant mutants were designated PAS2 (1^{st} and 2^{nd}) and PAS4 (1^{st} and 2^{nd}). Conversely, creating resistant mutants against AD25a was relatively easy. This experiment was repeated twice and three mutants were selected each time, hence six AD25a mutants in total. These AD25a resistant mutants were designated: AD25a-1, AD25a-2, AD25a-3, AD25a-4, AD25a-5 and AD25a-6.

The WGS analysis identified a total of 28 single nucleotide polymorphisms (SNPs) in the AD25a resistant mutants (Table 25) and 40 SNPs in the PAS resistant mutants (Table 26). The SNPs involving the AD25a resistant mutants are completely different from the SNPs involving the PAS resistant mutants.

Gene	Function	AD25a 1	AD25a 2	AD25a 3	AD25a 4	AD25a 5	AD25a 6
		SNP Position					
Rv0197	Possible oxido-reductase	234494					
Rv0279c	PE family		338167				
Rv0388c	PPE family	467508					
Rv0532	PE family		623425				
Rv0532	PE family		623428				
Rv0532	PE family	623472	623472	623472	623472	623472	623472
Rv0578c	PE family	672491	672491				
Rv0746	PE family	836426	836426				
Rv0746	PE family	836454	836454	836454	836454		
Rv0746	PE family	836538	836538	836538	836538	836538	836538
Rv0747	PE family		839123	839123			
Rv0747	PE family		839129	839129			839129
Rv0747	PE family		840338				
Rv0747	PE family		840340				
Rv0833	PE family		845275				
Rv0834c	PE family		929943				
PE_PGRS25	PE family	1572865					
PE_PGRS25	PE family		1573326				
PE_PGRS25	PE family		1573335				
PE_PGRS31	PE family		2001220				
Rv2015c	Conserved hypothetical protein		2262896				
PE_PGRS43	PE family		2805256				
Rv2540	Chorismate synthase		2863654				
Rv3347	PPE family			3745738			
Rv3515	PE family					3948347	
PE_PGRS61	PE family	4094140					
Rv3655c	Hypothetical protein	4095000					

TABLE 25: LIST OF SNPS IN THE AD25A RESISTANT MUTANTS.

The table shows the genes involved with the respective function and SNP position in the genome. Only non-synonymous SNPs were considered. The gene *Rv2540* is highlighted in grey and its function is further discussed in the text. All these SNPs are different from the SNPs involving the PAS resistant mutants.

Gene	Function	PAS2 1st	PAS2 2nd	PAS4 1st	PAS4 2nd
		SNP Position			
Rv0486	Glycosyltransferase			576252	
Rv1072	Conserved membrane protein		1196380		
Rv1204c	Conserved protein				1347266
rrs	16s RNA gene		1472240		
rrs	16s RNA gene		1472242		1472242
rrs	16s RNA gene		1472251		1472251
rrs	16s RNA gene		1472252		1472252
rrs	16s RNA gene		1472253		1472253
rrs	16s RNA gene		1472256		1472256
rrs	16s RNA gene		1472259		1472259
rrs	16s RNA gene			1472616	
rrs	16s RNA gene	1473062			
rrs	16s RNA gene	1473068			
rrl	23s RNA gene		1475816		
rrl	23s RNA gene		1475817		
rrl	23s RNA gene		1476126		
rrl	23s RNA gene		1476131		
rrl	23s RNA gene		1476141		
rrl	23s RNA gene		1476153		
rrl	23s RNA gene		1476154		
rrl	23s RNA gene		1476164		
rrl	23s RNA gene		1476165		
rrl	23s RNA gene			1476455	
rrl	23s RNA gene			1476456	
Rv1431	Conserved membrane protein		1609581		
nadB	L-aspartate oxidase		1796003		
Rv2215	E2 component of pyruvate	2482556			
ahpD	Alkyl hydroperoxide reductase		2727108	2727108	
FolC	Folylpolyglutamate synthase				2747141
Rv2695	Conserved hypothetical protein		3012106		
cysA3	Thiosulfate sulfurtransferase		3484226		
Rv3202c	Possible ATP-dependent DNA helicase			3578723	
Rv3218	Conserved protein		3594639	3594639	
Rv3218	Conserved protein	3594791			
Rv3232c	Polyphosphate kinase		3609025		
RV3505	Acetyl-CoA dehydrogenase	3924194			
Rv3785	Hyphotetical protein				4232224
Rv3894c	Type VII secretion system		4378811		

TABLE 26: LIST OF SNPS IN THE PAS RESISTANT MUTANTS.

The table in the previous page shows the genes involved in the PAS-resistant mutants, with the respective function and SNP position in the genome. Only non-synonymous SNPs were considered. The genes *rrs*, *rrl* and *folC* are highlighted in grey and their function is further discussed in the text. All these SNPs are different from the SNPs involving the AD25a resistant mutants.

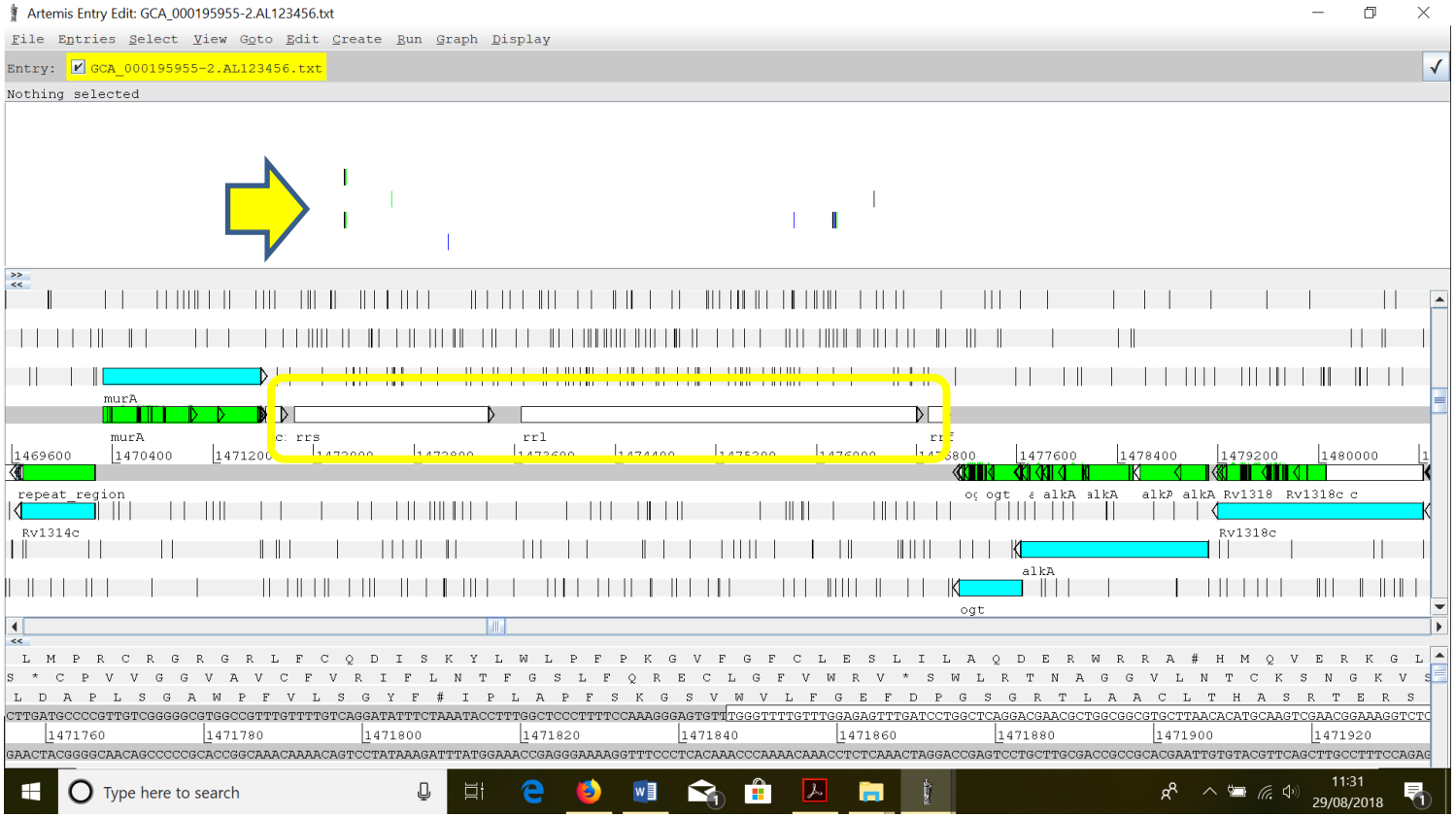


FIGURE 15: ALIGNMENT OF *rrs/rrl* SNPS IN THE PAS-RESISTANT MUTANTS

The figure above shows the SNPs positions in the *rrs/rrl* genes (16s and 23s RNA) of the PAS-resistant mutants. The SNPs are indicated by the top yellow arrow, whilst the *rrs/rrl* genes are circled in yellow in the middle of the screen. All four mutants have SNPs in the above genes. The figure has been created using Artemis and *.vcf* files (Section 1.10.6).

5.4 Discussion

PAS maintained its position as first line TB treatment for 25 years but after the introduction of ethambutol, rifampicin and pyrazinamide its use fell out of favour. A *Tubercle* editorial in 1973 stated that "*PAS may not quite be obsolete—but it is clearly obsolescent*", referring to the development of resistance and significant side effects (Editorial, 1973). Up to 75% of patients reported gastrointestinal symptoms (in particular, diarrhoea, nausea and vomiting) but this was likely associated with the type of formulation and the doses used at the time (Lehmann, 1949). More recent data recorded PAS being withdrawn in only 7% of patients with XDR TB due to the severity of side effects (Shean et al., 2013). In modern times, PAS has gained new life thanks to the increasing rates of MDR and XDR TB and it is often used to strengthen regimens (Donald & Diacon, 2015). Hence, fully understanding its mechanism of action and the development of PAS analogues is important to provide additional tools in the fight against drug resistant TB.

The activity of different PAS analogues was tested in this chapter. The primary hypothesis that these new compounds are active against *M. tuberculosis* was confirmed by using spot culture and resazurin methods. One compound in particular, AD25a, was selected for further analysis out of the original sixteen, on the basis of its activity and lower critical concentration. Resistant mutants were created against AD25a and PAS and WGS was performed on all the mutants in the attempt to understand and compare their mechanisms of action. However, WGS analysis was unable to confirm the secondary hypothesis that this new compound has the same mechanism of action as PAS. The complete difference in the mutation profiles suggests that AD25a has a mechanism of action different to that of PAS, despite AD25a being synthesised as a PAS analogue. Other authors have previously described the creation of analogues that had a different mechanism of action compared with the original drug (Al-Balas et al., 2009). It is difficult to predict the consequences of a change in the chemical structure and the impact that this will have on the mechanism of action. Hence, this seems a reasonable explanation in the case of AD25a.

The antituberculous activity of AD25a has been confirmed using three different methods (spot culture, resazurin and BACTEC/MGIT). There seems to be a discrepancy between the critical concentrations of AD25a using the resazurin method (0.04 µg/ml) and the BACTEC/MGIT method (0.39 µg/ml). However, this discrepancy

is unsurprising as variability among different susceptibility methods has already been described by other authors (Banu et al., 2014). In addition, it is already known (Section 1.7.5 and Table 5) that the same antibiotic can have different critical concentrations depending on the media used and the testing format. Considering PAS as an example, its critical concentration ranges from 0.5 µg/ml to 8 µg/ml depending on LJ slopes or 7H11, respectively. The critical concentration of AD25a at 0.39 µg/ml is still much lower than the critical concentration of PAS and many other antituberculous drugs (Heifets, 2013), making AD25a a very promising compound.

WGS analysis of AD25a mutants has highlighted different SNPs, the significance of which is unknown. The majority of them involve genes implicated in the synthesis of Pro-Glu (PE) and Pro-Pro-Glu (PPE) proteins (Table 25). The PE and PPE families include around 170 different genes in *M. tuberculosis* and they represent 10% of its genome (Mukhopadhyay & Balaji, 2011). Recent studies indicate that PE/PPE proteins could act as virulence factors by helping the bacteria to establish a successful infection inside the host through antigenic variation and other possible unconfirmed mechanisms (Mukhopadhyay & Balaji, 2011). However, it is difficult to explain their role as potential targets of antibiotics and other authors have suggested excluding them from the SNPs analysis due to their repetitive nature (Casali et al., 2016). Only one mutant out the six randomly selected showed an interesting mutation in the chorismate synthase gene (*Rv2540*). *Rv2540* is involved in the shikimate pathway synthesis of aromatic amino acids. It is also implicated in various functions such as UV protection, electron transport, signalling, communication and plant defence. Moreover, the enzymes in the shikimate pathway, and chorismate synthase specifically, are considered to be potential targets for new antimicrobial treatment for *M. tuberculosis* as studies have shown that the shikimate pathway is essential for mycobacterial viability whilst it is not present in human cells (Reichau et al., 2011). Unfortunately, this is a preliminary observation and has not been confirmed in other AD25a mutants. An alternative hypothesis is that the mechanism of resistance against AD25a is mediated via efflux pumps or other tolerance mechanisms. Efflux pumps against PAS and other antibiotics have already been described (Ramon-Garcia et al., 2012). Additionally, the selection of drug-tolerant *M. tuberculosis*, not drug-resistant, has been previously observed with isoniazid, confirming the presence of phenotypically but not genetically resistant bacilli (Ahmad et al., 2009).

In contrast with AD25a mutants and the inability to elucidate their mechanism of action, WGS analysis of PAS resistant mutants has provided some interesting results.

However, it is important to revisit the folate metabolic pathway to fully understand our data. As shown in Figure 16, folates are essential cofactors in the transfer of one-carbon groups in pathways for the synthesis of methionine, *N*-formylmethionyl-tRNA, glycine, serine, pantothenate, purines, and thymidine (Minato et al., 2015). The folate biosynthetic pathway starts when the aromatic precursor chorismate is converted to *p*-aminobenzoic acid (PABA) and coupled with the pteridine portion (from GTP) to generate dihydropteroate (H_2Pte) (*folP1* gene involved). The protein encoded by *folC*, dihydrofolate synthetase, adds the initial glutamate molecule to form dihydrofolate (H_2PteGlu) (Green & Matthews, 2007). We have already mentioned that mutations in the *folC* gene have been associated with PAS resistance but in less than 35% of the cases, whilst the same authors could not find any mutation in the *folP1* gene (Zhao et al., 2014). Only one PAS resistant mutant (out of four) showed a SNP affecting the *folC* gene (Table 26). This means that additional resistant mechanisms must be present and they could potentially involve a target further downstream.

All our PAS mutants showed mutations in the *rrs* and *rrl* genes (16S and 23S RNA gene, respectively) (Table 26). These mutations should affect the ribosomes, including the conversion from $tRNA^{Met}$ to $tRNA^{fMet}$ (Figure 16) and the overall synthesis of proteins. This highlights a new potential mechanism of resistance for PAS that has never been observed before. It also resembles the mechanism of resistance of streptomycin, albeit differently. In general, streptomycin resistance is frequently associated with mutations in the gene *rpsL*, encoding the ribosomal protein S12. The presence of several rRNA operons (*rrn*) and a single *rpsL* gene in most bacterial genomes prevents the isolation of streptomycin-resistant mutants in which resistance is mediated by mutations in the 16S rRNA gene (*rrs*) (Springer et al., 2001). In *M. tuberculosis*, whilst mutations in the *rrs* gene have been described, they are much less frequent compared to the mutations in the *rpsL* gene (Honore & Cole, 1994) (Sreevatsan et al., 1996) (Cooksey, Morlock, McQueen, Glickman, & Crawford, 1996) (Jagielski, Ignatowska, et al., 2014). Additionally, mutations in the *rrl* gene have not been associated with streptomycin resistance. Ours findings seem to indicate a new and completely different mechanism of action for PAS, affecting the ribosomes and the synthesis of proteins. As PAS and streptomycin have been companion drugs for decades and at a time when next generation sequencing technologies were not available, this mechanism of action/resistance may have been missed. We have discussed how resistance in PAS is probably caused by different mechanisms (as the

current evidence can only justify approximately half of the cases) and these mutations in the *rrs* and *rrl* genes could represent the missing link.

It is also important to highlight the difficulty in naturally selecting PAS-resistant mutants in the laboratory, in sharp contract with AD25a and other antibiotics (i.e. INH-resistant mutants in Chapter 3) and the early resistant clinical isolates described after PAS introduction (Turnbull et al., 1953). This could be due to the essentiality of genes involved and the negative impact that a deletion and/or an altered gene function would have on the cell survival. A recent paper (DeJesus et al., 2017) has provided a comprehensive essentiality analysis of the *M. tuberculosis* genome via saturating transposon mutagenesis and both *rrs* and *folC* are included in the list of essential genes.

In conclusion, the antituberculous activity of new PAS analogues has been confirmed suggesting a potential role in the fight against drug resistant TB. After nearly 70 years of clinical use of PAS, WGS analysis may help in elucidating its mechanism of action, previously not completely understood. At the same time, there are still limitations that need to be addressed: the presence of mutations in the genome needs additional confirmation that such mutations encode for significant metabolic changes and the possible overlap of PAS and streptomycin resistance requires further investigation.

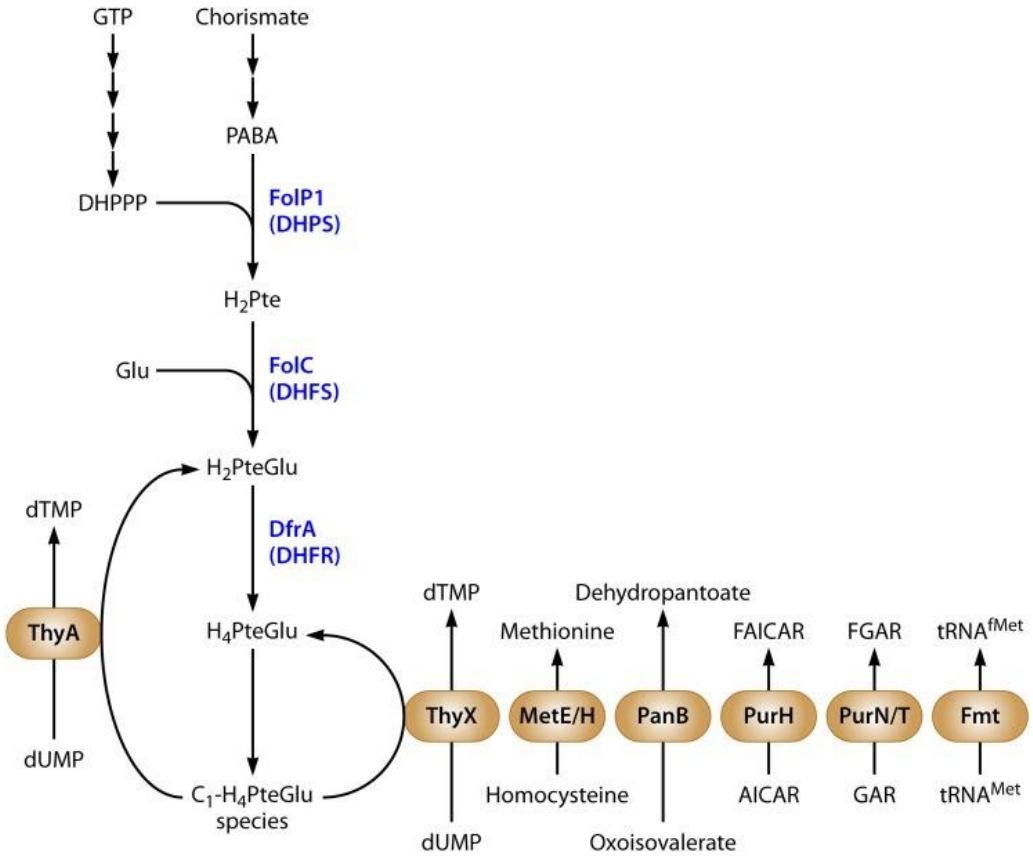


FIGURE 16: **FOLATE METABOLISM IN *M. TUBERCULOSIS*.**

Abbreviations: GTP, guanosine-5′-triphosphate; DHPPP, 7,8-dihydropterin pyrophosphate; PABA, para-aminobenzoic acid; Glu, glutamate; H_2Pte, dihydropteroate; $H_2PteGlu$, dihydrofolate; $H_4PteGlu$, tetrahydrofolate; C1-$H_4PteGlu$, various single-carbon-modified species of $H_4PteGlu$; DHPS, H_2Pte synthase; DHFS, dihydrofolate synthase; DHFR, dihydrofolate reductase; dTMP, deoxythymidine monophosphate; FAICAR, 5-formamidoimidazole-4-carboxamide ribotide; AICAR, 5-aminoimidazole-4-carboxamide ribonucleotide; FGAR, 5′-phosphoribosyl-N-formylglycinamide; GAR, 5′-phosphoribosylglycinamide; $tRNA^{Met}$, methionyl-tRNA; $tRNA^{fMet}$, N-formylmethionyl-tRNA (Minato et al., 2015). Authorization from the American Society for Microbiology (ASM) to republish the above figure in a doctoral thesis only.

Chapter 6: Ertapenem and faropenem for the treatment of drug resistant tuberculosis

6.1 Introduction

Carbapenems are a class of beta-lactam antibiotics which include imipenem, meropenem and ertapenem. More recently, two new carbapenems (faropenem and doripenem) have been marketed in a limited number of countries (in particular, India and Japan). Emerging evidence demonstrates that they target the mycobacterial cell wall, providing an alternative treatment for multidrug-resistant (MDR) and extensively drug-resistant (XDR) tuberculosis (TB), where options are limited (Jaganath et al., 2016). *M. tuberculosis* uses L,D-transpeptidases for the formation of its peptidoglycan (Gupta et al., 2010) and *in vitro* studies have demonstrated that imipenem is the most efficient carbapenem to inhibit these transpeptidases (Cordillot et al., 2013). Moreover, the production of a class C beta-lactamase by *M. tuberculosis* is encoded by the *BlaC* gene. *BlaC*, confers resistance to the majority of other, non-carbapenem, beta-lactams by opening and deactivating the beta-lactam ring. Carbapenems are not readily neutralized by class C beta-lactamases in this manner and thus have potential as therapeutic agents against *M. tuberculosis* (W. J. Li et al., 2013).

Though imipenem is the most effective transpeptidase inhibitor, meropenem seems to be more potent than imipenem and other carbapenems against whole bacilli in terms of minimum inhibitory concentration (MIC). However, laboratory data on drug efficacy and specificity are variable, and a wide range of MICs for these drugs are reported (Jaganath et al., 2016). Additionally, *in vitro* experiments have shown limited or no activity of meropenem as a single agent against *M. tuberculosis* and they have highlighted the necessary synergy with amoxicillin/clavulanate to obtain a significant reduction in its MIC (Gonzalo & Drobniewski, 2013). Clinical studies, mostly involving the treatment of MDR and XDR TB cases, have demonstrated an overall positive outcome when using the meropenem/clavulanate combination (Dauby, Muylle, Mouchet, Sergysels, & Payen, 2011) (Payen et al., 2012) (De Lorenzo et al., 2013). However, studies to date have suffered from small sample sizes (or even individual case reports) and the inclusion of linezolid, making it difficult to discriminate/distinguish the contribution of meropenem and assess its clinical efficacy. The only observational study comparing imipenem/clavulanate- versus meropenem/clavulanate-containing regimens in the treatment of MDR- and XDR-TB

has suggested a meropenem superiority and, currently, it is the carbapenem of choice when managing multi drug resistance (Tiberi, Sotgiu, et al., 2016).

Compared to imipenem and meropenem (both only available as intravenous formulations), ertapenem is relatively new and its long half-life enables once-daily dosing (versus the general administration of three and four times a day with meropenem and imipenem, respectively). As such, ertapenem is a very attractive alternative for ambulatory or homecare treatment as MDR and XDR TB patients often require prolonged and costly hospital admissions (Arbex, Siqueira, D'Ambrosio, & Migliori, 2015). However, there is a paucity of data on the efficacy of ertapenem against *M. tuberculosis*. Only two laboratory studies have been published on the subject (Veziris, Truffot, Mainardi, & Jarlier, 2011) (Kaushik et al., 2015). Both only evaluated the activity of Ertapenem against the reference strain H37Rv and they did not include any clinical isolates. Clinical *in vivo* studies found favourable outcomes in the majority of patients treated with ertapenem but the number of patients treated was again very limited and treatment was with a combination of different drugs (Tiberi, D'Ambrosio, et al., 2016) (van Rijn et al., 2016).

Faropenem is the only orally administered carbapenem, structurally similar to the other antibiotics of this class but with a sulphur atom substitution that gives stability and efficacy (Dalhoff, Nasu, & Okamoto, 2003). It is only licenced in India and Japan for the treatment of lower respiratory tract, ear, and throat infections (Gurumurthy et al., 2017). Laboratory experiments have shown rapid cytolysis of *M. tuberculosis* after exposure to faropenem, due to more efficient inactivation of the target enzyme (L,D-transpeptidases) (6- to 22-fold greater than same dose meropenem) (Dhar et al., 2015). However, not only did the authors not test faropenem's activity against clinical isolates (only the laboratory strains Erdman and H37Rv), but subsequent experiments, using whole-blood bactericidal activity (WBA) as the read out, showed no activity - possibly due to inadequate plasma levels of the drug (Gurumurthy et al., 2017).

The aim of this chapter was to test the *in vitro* activity of ertapenem and faropenem (with and without the addition of amoxicillin/clavulanate) against different clinical isolates of *M. tuberculosis* and the reference strain H37RV, to better understand their potential role as additional antibiotics in the management of drug resistant TB.

6.2 Materials and methods

6.2.1 Selection of isolates

Twenty samples in total (19 clinical isolates plus H37Rv) were tested against different concentrations of ertapenem and faropenem. The reference strain H37Rv was used as a control strain. Nine clinical isolates were collected from the Department of Microbiology at the Royal Free London, including some of the clinical isolates from the already described INH-R outbreak in London (Chapter 4). The other ten clinical isolates (including MDR and XDR samples) were kindly provided by Professor Francis Drobniewski and Dr Ximena Gonzalo from Imperial College London. Dr Gonzalo also tested all samples against Meropenem/amoxicillin/clavulanate. The samples were selected to offer a range of strains from fully susceptible to XDR, and a summary is provided in Table 27.

6.2.2 Susceptibility testing

Susceptibility testing was performed using the BACTEC960 MGIT system (Section 2.3.3). Four different concentrations of ertapenem and faropenem were tested with and without the addition of amoxicillin/clavulanate. These concentrations were selected based on previous pharmacokinetic/pharmacodynamic (PK/PD) data (Nix, Majumdar, & DiNubile, 2004) (Schurek et al., 2007) (Mushtaq, Hope, Warner, & Livermore, 2007). All drugs were from Sigma-Aldrich (Dorset, UK) as either sodium (ertapenem/faropenem) or potassium (amoxicillin/clavulanate) salts and diluted in sterile water. The amoxicillin/clavulanate potassium powder comes in a ratio of 2:1 and it was diluted to reach a final concentration of clavulanate of 2.5 µg/ml. A summary of the different concentrations used is provided in Table 28.

Strain	Phenotypical resistance profile	Meropenem clavulanate MIC (µg/ml)	Notes
03:013	S	32	
03:039	H	16	
04:018	H,R,clari,ethi	Failed	MDR
05:094	Fully susceptible	8	
07:116	H, ethi	4	
11:136	S,H,R	>32	MDR
11:156	S,H,R	4	MDR
11:191	H	16	
11:368	S,H,R	>32	MDR
324	Fully susceptible	8	
333	S,H,R	2	MDR
346	S,H,R	2	MDR
347	Fully susceptible	>32	
401	H,R	>32	MDR
408	S,H,R	>32	MDR
443	Fully susceptible	>32	
548	N/A*	>32	XDR
421	S,H,R,EMB,CAP,Moxi	>32	XDR
433	S,H,R,EMB,PYR,CAP,Moxi	8	XDR
H37Rv	Fully susceptible	2	Control reference strain

TABLE 27: **LIST OF *MYCOBACTERIUM TUBERCULOSIS* ISOLATES TESTED AGAINST ERTAPENEM AND FAROPENEM.**

The resistance profile is shown in the second column, with a variety of fully susceptible, MDR and XDR samples (S=Streptomycin, H=Isoniazid, R=Rifampicin, Clari=Clarithromycin, Ethi=Ethionamide, EMB=Ethambutol, CAP=Capreomycin, Moxi=Moxifloxacin, PYR=Pyrazinamide). The MIC against Meropenem/clavulanate is also shown. *This isolate was confirmed as XDR by WGS but phenotypical susceptibilities were not performed.

Ertapenem	Ertapenem Clavulanate	Faropenem	Faropenem Clavulanate
µg/ml	µg/ml	µg/ml	µg/ml
16-8-4-2	16-8-4-2 (+2.5 Clavulanate each)	8-4-2-1	8-4-2-1 (+2.5 Clavulanate each)

TABLE 28: **CONCENTRATIONS OF ERTAPENEM AND FAROPENEM TESTED.**

Four different concentrations of ertapenem (16, 8, 4 and 2 µg/ml) and faropenem (8, 4, 2 and 1µg/ml) were tested with and without the addition of amoxicillin/clavulanate (2.5 µg/ml).

6.3 Results

The results of ertapenem and faropenem testing (with and without amoxicillin/clavulanate) are shown in Table 29. Eighteen out of twenty samples were resistant to the highest concentration of ertapenem tested (including the addition of amoxicillin/clavulanate). Half of the samples tested have some degree of susceptibility to faropenem and the addition of amoxicillin/clavulanate further reduces the MIC level in seven isolates.

	MIC				
Strain	**ERT**	**ERT+C**	**FAR**	**FAR+C**	**MER+C**
03:013	>16	>16	8	4	32
03:039	>16	>16	8	4	16
04:018*	>16	>16	4	4	Failed
05:094	16	16	4	4	8
07:116	16	16	4	1	4
11:136*	>16	>16	8	4	>32
11:156*	>16	>16	8	2	4
11:191	>16	>16	>8	>8	16
11:368	>16	>16	>8	>8	>32
324	>16	>16	>8	>8	8
333*	>16	>16	>8	>8	2
346*	>16	>16	8	8	2
347	>16	>16	>8	8	>32
401*	>16	>16	>8	>8	>32
408*	>16	>16	8	8	>32
443	>16	>16	>8	>8	>32
548*	>16	>16	>8	>8	>32
421*	>16	>16	>8	8	>32
433*	>16	>16	>8	>8	8
H37Rv	>16	>16	>8	>8	2

TABLE 29: **RESULTS OF ERTAPENEM AND FAROPENEM TESTING AGAINST CLINICAL ISOLATES OF *M. TUBERCULOSIS*.**

An asterisk indicates that the strain is MDR or XDR (n=20). Susceptible concentrations are highlighted in grey (ERT=ertapenem, ERT+C=ertapenem plus amoxicillin/clavulanate, FAR=faropenem, FAR+C=faropenem plus amoxicillin/clavulanate, MER+C=meropenem plus amoxicillin/clavulanate). The MIC against meropenem/clavulanate is shown again to facilitate the comparison. The experiment was repeated once due to its costs. However, an alternative susceptibility testing with a dilution method was independently performed by Dr Gonzalo (Imperial College London), supporting the same results (data from Dr Gonzalo not shown).

6.4 Discussion

The results from this chapter have highlighted various levels of resistance of the clinical isolates and H37Rv to ertapenem *in vitro*, with no difference following the addition of amoxicillin/clavulanate. Conversely, 10 clinical isolates (out of 20) have shown different degrees of susceptibility to faropenem and the addition of amoxicillin/clavulanate further reduced the MIC in 7 clinical isolates. The concentrations tested were chosen based on previously published PK/PD data and the level of antibiotic achievable in the blood with the currently licensed dosing regimens (Nix et al., 2004) (Schurek et al., 2007) (Mushtaq et al., 2007). Previous laboratory studies reported the ertapenem MIC at 4 µg/ml (Veziris et al., 2011) whilst there are no breakpoints available for faropenem and *M. tuberculosis*. The method using the BACTEC960 MGIT system was preferred due to its reliability, automation and previous validation in numerous studies for the susceptibility testing of traditional and newer antimicrobials against *M. tuberculosis* (Bemer, Palicova, Rusch-Gerdes, Drugeon, & Pfyffer, 2002) (Lin, Desmond, Bonato, Gross, & Siddiqi, 2009) (Rusch-Gerdes, Pfyffer, Casal, Chadwick, & Siddiqi, 2006).

It is important to further discuss the PK/PD of carbapenems when evaluating their efficacy. Penicillins have a time-dependent mechanism of killing and carbapenems exhibit a bacteriostatic effect when at least 20% of the exposure time is above the MIC. The bactericidal effect is achieved when the exposure time above the MIC is at least 40% (Zhanel et al., 2007). Serum maximum concentration (C_{max}) depends on the type of carbapenem and comorbidities (i.e. renal failure, altered volume of distribution in case of sepsis etc), but in an healthy adult male the highest concentration for ertapenem is 154.9 µg/ml, followed by imipenem (60–70 µg/ml) and meropenem (50–60 µg/ml) (Zhanel et al., 2007). A clinical study evaluating ertapenem pharmacokinetics in 12 patients with MDR or XDR TB found that the C_{max} was 127.5 µg/ml and the half-life was 2.4 hours (van Rijn et al., 2016). Considering that ertapenem requires only one daily dose, the blood concentration should remain above the theoretical MIC of 4 µg/ml for at least 8 hours and additional administrations may be required if clinical isolates have a higher MIC. Similar data on faropenem are still lacking, including studies on the early bactericidal activity, but some clinical trials are currently ongoing at the time of this thesis (Diacon et al., 2016).

The significant level of resistance to ertapenem found in our experiments (despite the clinical success described by some authors) can be explained by two main theories:

real carbapenem resistance versus antibiotic degradation. Resistance to carbapenems can be conferred by various mechanisms, including transmissible beta-lactamases (Meletis, 2016). However, resistance in *M. tuberculosis* is likely to emerge due to spontaneous mutations or the presence of efflux pumps (Nguyen, 2016). A previous report showed that a single amino-acid substitution in *blaC* conferred the ability of this beta-lactamase to resist inactivation (Soroka et al., 2015). Whole genome sequencing was performed on some (n=5) of the clinical isolates tested (see Chapter 6) but no mutation was found on the *blaC* gene. Efflux pumps able to confer resistance to carbapenems have also been described (Dinesh, Sharma, & Balganesh, 2013) but further work is needed to better understand such mechanisms. This leaves the degradation of antibiotic as a plausible theory. There have been reported issues with the stability of ertapenem at room temperature (Keel, Sutherland, Crandon, & Nicolau, 2011) (Jain, Sutherland, Nicolau, & Kuti, 2014) but also degradation issues with imipenem at different temperatures and neutral pH ranges (G. B. Smith, Dezeny, & Douglas, 1990) (Keel et al., 2011). Given the slow replication of *M. tuberculosis*, this leads to a challenging situation in testing where the antibiotic is possibly degraded before killing or inhibiting bacterial growth. Some authors have also suggested the daily addition of antibiotics to this experimental set up (Watt, Edwards, Rayner, Grindey, & Harris, 1992) but this would hamper the evaluation of the dose tested and it would increase the risk of contamination. In addition, stability is not a problem for faropenem as it is thermo-stable at 37 degrees (Viaene, Chanteux, Servais, Mingeot-Leclercq, & Tulkens, 2002).

The results for faropenem are interesting yet disappointing. Only half of the isolates exhibited some susceptibility and only at the highest concentrations tested (apart from samples 07:116). The addition of amoxicillin/clavulanate did improve the MICs and this is in line with previous experiments with other carbapenems (in particular meropenem) (Gonzalo & Drobniewski, 2013). Faropenem did show some limited activity (MIC of 4 or higher) in strains completely resistant to meropenem (MIC of 32 or higher). However, the five isolates fully susceptible to first line antituberculous drugs were completely resistant at the highest concentration of faropenem used. This highlights a major issue of unpredictability and explains its limited use in clinical practice. In addition, further studies are still needed to assess which antibiotic level is actually achievable in the blood (and in the lung parenchyma) after the oral administration of faropenem as reaching the concentrations tested in our experiments may be challenging.

It is also worth mentioning that the use of carbapenems for a prolonged period of time (treatment of MDR TB is at least 12 months) may cause the emergence of resistance among other bacteria, in particular gram-negative. Studies on the topic are still missing and caution should be used when using this class of antibiotics for other unlicensed indications, including the treatment of tuberculosis.

The increasing global incidence of drug resistant TB demands additional therapeutic options. Although carbapenems are promising agents, this chapter has highlighted some of the limitations of their use. These data also raise questions regarding the mechanisms of action of newer carbapenem variants and further work is needed prior to their clinical use. Susceptibility testing of *M. tuberculosis* with ertapenem and faropenem showed a significant inherent level of resistance. Until further research is conducted, carbapenems will remain a third line choice to be used only in cases of MDR and XDR TB where other options are not available.

Chapter 7: An alternative approach: evaluation of gold nanoparticles

This chapter has been edited due to copyright reasons. Its use is not permitted under the quotation exception, and permission to make it available has not been obtained. For these reason, two versions, "edited" with no Chapter 7 and "complete" with Chapter 7, have been deposited.

Chapter 8: Discussion

8.1 The challenge of drug resistant tuberculosis

Despite the introduction of an effective combination treatment in the 1960's, tuberculosis remains a major cause of morbidity and mortality globally. In 2015, there were an estimated 480 000 new cases of multidrug-resistant TB (MDR-TB), with India, China and the Russian Federation accounting for 45% of these cases (WHO, 2017). This will also have a major financial impact on healthcare systems all over the world, with direct costs, mostly covered by the public sector, of $134,000 per MDR TB and $430,000 per XDR TB patient (Marks et al., 2014). New drugs have been discovered and they are progressing through various clinical trials worldwide, but treatment of drug resistance is still challenging, with less than half of patients with drug resistant TB being cured, mainly due to high mortality and loss of follow up (Wallis et al., 2016).

In 1962, Georges Canetti, a prominent French microbiologist and the discoverer of *Mycobacterium canettii*, stated that "*if we examine the situation not on a narrow national scale, but on a world-wide scale, one conclusion emerges: among the efforts to be made for progress on the road to the eradication of tuberculosis, there is one absolute priority: the improvement of chemotherapy*" (Canetti, 1962). Optimization of treatment regimens still remains one of the main challenges and substantial efforts have been recently invested in drug development, particularly to create new regimens that can significantly shorten the duration of treatment, improve tolerability and reduce drug-drug interactions (Dooley et al., 2012). Other challenges include the need for accurate and early diagnosis, screening for drug resistance, administration of treatments in developing countries and follow-up support to ensure treatment completion (Zumla et al., 2013). A multi-pronged approach must be the solution to tackle the challenge of drug resistant TB. This will involve not only the discovery of new antibiotics (and revising the role of/repurposing old ones) but also a streamlined and rapid screening process in the laboratory phase to select the most promising compounds, research on alternatives to antibiotics (i.e. nanoparticles and bacteriophages) and the use of nanotechnologies with a potential role in increasing the uptake of drugs directly inside the cells (antibiotic carriers). However, other authors have warned that the Stop TB Partnership eradication plan by 2050 will be unreachable without a new vaccine (R. Smith, 2009). Hence, it is also essential to fully understand TB pathogenesis and the role of mycobacterial genes to develop

better vaccines in addition to new drugs. Fundamentally, an *all-inclusive* approach to circumvent the challenges of TB is required.

The rest of this discussion will concentrate on the following main topics: the role of WGS in drug discovery, the current challenges in the drug development process and the repurposing of existing drug/evaluation of alternative approaches and their role in addressing the challenge of drug resistant TB (Table 31).

Role of WGS in drug discovery	Screening of novel antituberculous compounds	Repurposing exisiting antibiotics and alternative compounds
•**Chapter 4** - WGS to understand the genetic variations in a subset of clinical isolates from the London INH-R outbreak and new potential drug targets •**Chapter 5** - WGS to elucidate the mechanism of action of old and new drugs	•**Chapter 5** - Screening of novel compounds (PAS analogues). AD25a showed promising laboratory results	•**Chapter 6** - Testing of ertapenem and faropenem to understand their potential role as antituberculous drugs •**Chapter 7** - Evaluation of gold nanoparticles as direct antibacterial agents and/or antibiotic delivery carriers

TABLE 30: THE CHALLENGE OF DRUG RESISTANT TB AND THE NEED FOR A MULTI-STRATEGY APPROACH

The table above provide an overview of the thesis chapters and their relation to the challenge of drug resistant TB.

8.2 The role of whole genome sequencing in drug discovery

Despite decades of research on TB pathogenesis, up to 50% of mycobacterial genes are still labelled as unknown, uncharacterized or with a hypothetical function (Mazandu & Mulder, 2012). WGS can help in elucidating such mechanisms but there are still significant limitations we need to be aware of. As illustrated in Chapter 4 exploring the genetic variation of a sub-cluster of the London INH-R TB outbreak, WGS was able to demonstrate specific deletions and SNPs affecting different genes

that were unique to those clinical isolates and which could potentially explain the persistence of the outbreak over years. Other authors have also created SNP-calling programs combined with WGS that has allowed the creation of a catalogue of virulence genes and new potentially virulent sub-lineages (Mikheecheva, Zaychikova, Melerzanov, & Danilenko, 2017). However, the simple presence of such mutations does not necessary confirm their function. Further studies (transcriptomics, knock out models, recombineering) are still needed to confirm the impact of those genes in the pathogenesis of TB and if any of the virulence genes involved in these SNPs can be used as drug targets for the development of new compounds. Caution also needs to be used with this approach. Many of these genes may not be essential, accompanied by genetic redundancy or metabolic scavenging (overcoming nutrient insufficiency by using alternative pathways or substrates), leading some authors to consider the traditional pharmacological target-validation a far more reliable and a better predictor of success (Sassetti, Boyd, & Rubin, 2003) (Lechartier, Rybniker, Zumla, & Cole, 2014). Additionally, compensatory mutations in other part of the genome may balance the initial fitness cost and further complicate the overall picture (Comas, Borrell, & Roetzer, 2011).

WGS has the potential to revolutionize the process of drug target identification (Sala & Hartkoorn, 2011). As demonstrated in Chapter 5, WGS should be integrated into the drug screening process. It has been successfully used to determine the target of bedaquiline (BDQ) (Andries et al., 2005). Here, the authors selected and sequenced BDQ resistant *Mycobacterium smegmatis* strains and identified mutations in the proton pump of adenosine triphosphate (ATP) synthase associated with resistance. Others (Gonzalo, Casali, Broda, Pardieu, & Drobniewski, 2015) have selected mutants directly using *M. tuberculosis* to evaluate combination treatment against drug resistant strains. A scalable platform for the discovery of drug targets, based on combining high-throughput screening (HTS) with WGS of resistant isolates, has been proposed (T. Ioerger et al., 2013). Using this approach, the authors have identified promising new drug targets. Surprisingly, some structurally different compounds caused the selections of same genes mutations (conferring resistance) and this is the same confounding issue we have found in our PAS-resistant mutants and their *rrs* gene mutations. When sequencing the genome of resistant mutants, researchers should remember that the mutations conferring resistance may involve different stages of drug metabolism: mutations can affect enzymes that convert the compound into an active form (and then inhibit an unknown target), the target gene, regions

responsible for up- or down-regulation of targets and activating enzymes, efflux pumps or detoxifying enzymes. It is also possible that the same drug may be inactivated by different mechanisms of resistance (as in the case of PAS) and some of them may be common to other compounds, as already described for other antituberculous drugs such as isoniazid and ethionamide (Almeida Da Silva & Palomino, 2011). WGS could therefore contribute to elucidating global mechanisms of action/resistance.

Understanding bacterial pathogenesis is expected to provide an instrumental contribution to vaccine development, particularly to target those pathogens (such as *M. tuberculosis*) for which the traditional approaches have not been completely successful (Scarselli, Giuliani, Adu-Bobie, Pizza, & Rappuoli, 2005). The comparison, by genome sequencing, of *M. tuberculosis* clinical strains CDC1551 with the laboratory-adapted H37Rv has demonstrated a more extensive variability than had been anticipated (Fleischmann et al., 2002). This has highlighted the importance of understanding genetic diversity in pathogenic strains and of validating candidate targets with other genomic techniques. A functional genomic approach has been developed and inhibition techniques, such as signature-tagged mutagenesis (STM) and transposon site hybridization (TraSH), have been successfully applied to *M. tuberculosis* (Camacho et al., 2017) (Sassetti, Boyd, & Rubin, 2001) with the potential for better-attenuated vaccines.

The issue of genetic diversity has become particularly relevant since the introduction of WGS. The reference strain H37Rv is the most studied strain of *M. tuberculosis* in research laboratories (Cole, 1999) (Camus et al., 2002) and it was used in all our experiments unless stated otherwise. It was originally isolated in 1905 from a patient at the Trudeau Sanatorium (New York State) and it was maintained for several years by serial passages of culture. It was originally named H37 but it was later found to be dissociated into two main variants: the H37R (resistant to environment but avirulent) and the H37S (sensitive to environment but virulent) variants (Steenken, Oatway, & Petroff, 1934). A more virulent strain was renamed H37Rv in 1972 (with R standing for the rough morphology and V standing for virulent) and, having been used in many subsequent laboratory studies, it became the new reference for the species (Kubica, Kim, & Dunbar, 1972). The recent application of whole genome sequencing has allowed a better understanding of the genetic variations among the different H37Rv strains that were in stock and maintained in different research laboratories. The *M. tuberculosis* genome has continued to evolve even in a controlled laboratory

environment (T. R. Ioerger et al., 2010). Based on these findings, some authors are challenging its utility in the analysis of virulence-related loci and it is becoming apparent that use of H37Rv as a sole reference genome in analysing clinical isolates presents some limitations (O'Toole & Gautam, 2017). The same authors suggest the use of clinical strains for the investigation of virulence genes, further supporting our approach in Chapter 4.

From our findings, WGS has played an important role in unveiling additional insights into the pathogenicity of *M. tuberculosis* and the clinical strains from the London outbreak represented a valuable sample set that was appropriate for further analysis. In this context, we would support wide-scale investment in WGS and the creation of comparable databases of genetic variations of clinical strains, plus the addition of biological information, including fitness assays and mutation rate information for a comprehensive picture. This will allow the interrogation of *M. tuberculosis* gene function and the discovery of new drug targets. Additionally, WGS should be integrated as an essential component of the drug development process in early stage of pre-clinical laboratory studies.

8.3 Screening of novel antituberculous compounds

Enormous costs and time are involved in new drug development and this dictates a strict laboratory screening process in the preclinical phase to select only the most promising candidates and to increase their chance of reaching the market (Hughes & Karlen, 2014). The process of developing any new drug generally takes from 8 to 12 years and it is divided into different phases (Ciociola, Cohen, & Kulkarni, 2014). The laboratory and preclinical phases take around 4 years. If successful, the new drug will require approval from governing bodies (Food and Drug Administration in USA or European Medicine Agency in Europe) and it will be able to proceed to clinical trials (phase 1, 2 and 3, lasting a minimum of 1, 2 and 3 years, respectively). Efficacy and safety will be monitored throughout and comparison with other drugs may be evaluated. A new drug application will follow all these trials and, following acceptance, additional post-marketing studies can be requested (Ciociola et al., 2014). It is evident how this process can be difficult and expensive, with a calculated cost for each new drug now exceeding $1,000 million (DiMasi, Grabowski, & Hansen, 2016), which has more than doubled in the last decade (DiMasi, Hansen, & Grabowski, 2003). It is not surprising that from the laboratory phase to the market most compounds face a tough battle to become an approved drug. For approximately of every 5,000 compounds that

enter preclinical testing, only one is approved for marketing (Dahlin, Nelson, Haynes, & Sargeant, 2016). The creation of a streamlined laboratory screening process of new compounds thus becomes essential to accelerate and support the most promising candidates.

In drug development, the pre-clinical studies are completed on any new compound before it can be trialled in humans. There is a plethora of *in vitro* assays followed by validation in *in vivo* animal models to assess the activity of any new compound against *M. tuberculosis*. However, such variety of different tests makes it difficult to compare data from one laboratory to another. A review on the topic has confirmed the difference in screening protocols among different laboratories and the limitations of the various methods used (Franzblau et al., 2012). Recommendations included the use of a simple and sensitive screening test and the selection of more complex assays to confirm the activity of the most promising compounds. This supports the approach we used in Chapter 5. The early laboratory testing of new compounds (PAS analogues) was performed using the spot culture method as a rapid screening test. Further testing (resazurin assay and MGIT/BACTEC testing) confirmed the activity of the most promising candidate in a streamlined approach. The introduction of WGS in the early development of a drug represents an innovative addition and we have proposed its routine use when evaluating new compounds in Chapter 5.

The reference strain H37Rv was used when testing novel agents (Chapter 5 and Chapter 7), whilst known resistant clinical strains were also added for ertapenem and faropenem in Chapter 6. The choice of an appropriate bacterial strain is essential to obtain robust results. Authors have observed differences in susceptibility profiles even between H37Ra and H37Rv (Franzblau et al., 2012). *Mycobacterium smegmatis* and *Bacillus Calmette–Guerin* (BCG) are often proposed as alternative surrogates for mycobacterial drug discovery due to the relative low risk they pose to laboratory workers and their easy manipulation in the laboratory. However, a comparative study has shown that up to 58% of active compounds against *M. tuberculosis* would have been missed if using a *M. smegmatis* model. If directly comparing with the model of *M. bovis BCG*, the percentage of missing compounds would decrease to 34% instead (Altaf et al., 2010). Conversely, one could argue that the use of these surrogates has actually been successful so far and it has allowed the discovery of some current and most promising new agents, bedaquiline and delamanid (Andries et al., 2005) (Matsumoto et al., 2006).

The next stage in the process is to evaluate the anti-mycobacterial activity in animal models. These allow testing the safety of the new drugs and additional confirmation of their activity, including the effective dose to be used in further clinical studies. The guinea pig model is generally considered a better replica of human pathology but the mouse model is also used to test the activity of new compounds (Coxon, Cooper, Gillespie, & McHugh, 2012). The predictive value of mouse models for clinical trials has been questioned multiple times (Franzblau et al., 2012). Mitchison referred to two different occasions where mouse studies have potentially misdirected human clinical trials: the premature use of rifampicin and pyrazinamide for the treatment of latent TB infection in non-HIV patients and the substitution of moxifloxacin with another first line agent. In particular, the use of moxifloxacin instead of ethambutol produced little improvement in the mouse but definite improvement in patients, whereas the substitution of moxifloxacin for isoniazid produced marked improvement in the mouse but no improvement in patients (Mitchison & Chang, 2009) (Gillespie, Crook, McHugh, Mendel, Meredith, & Murray, 2014). Despite these concerns, it is important to bear in mind that mice are only “models” and they are cheap and readily affordable indicators to assess *in vivo* activity (Coxon et al., 2012). Other animal models have been described but they are much more expensive and limited in supply (Franzblau et al., 2012).

One of the main challenges of all laboratory models, both *in vitro* and *in vivo*, is that they do not necessarily reflect the real human infection and the complexity of the human immune response, including necrosis, the formation of granulomas and fibrosis. This issue was particularly evident in Chapter 5, where we were unable to select resistant laboratory mutants with desired mutations present in clinical strains. Other authors have also shown how the resistant mutants selected *in vitro* do not reflect the *in vivo* mechanisms (Bergval, Schuitema, Klatser, & Anthony, 2009). This lack of a reliable model is also encountered in the case of dormancy assays. The need for a prolonged duration of treatment is due to the propensity of *M. tuberculosis* to enter a dormant state, with altered metabolic activity and the ability to evade the immune system. A recent review of all available dormancy models has highlighted the lack of understanding of the mechanisms by which the tubercle bacilli survive in a dormant state and how none of the *in vitro* dormancy models reflect this reality (Alnimr, 2015).

The final hurdle in the development process of a new drug is the clinical applicability. One of the traditional approaches involves an initial monotherapy study (up to 2

weeks) with the frequent assessment of sputum viable counts to confirm if the novel agent is clinically active against *M. tuberculosis* (Gillespie, Gosling, & Charalambous, 2002). The problem is further complicated by the co-administration of other drugs, in particular antiretroviral drugs for HIV. Long and expensive clinical trials will be then needed to confirm the non-inferiority of any new regimen, but some authors are now proposing to accelerate such process by combining the phase 2 and 3 trial design (Phillips et al., 2016).

This thesis has addressed the early stages in the lengthy development of a new antituberculous drug. Various susceptibility methods have been used and WGS has been applied to the process. This approach has identified a very promising candidate, AD25a, and the next step would be to proceed towards *in vivo* models where AD25a's activity and safety profile can be further tested.

8.4 Repurposing existing antibiotics and alternative approaches

Due to the costly and lengthy process of drug discovery and development, revising/repurposing old drugs has emerged as an attractive strategy to increase the pool of available antituberculous medications, in particular for the treatment of patients infected with drug-resistant strains. Several old drugs have shown an inhibitory effect on the growth of *M. tuberculosis*. This is the case for both antibiotics and other drugs used for treating other infections and diseases (Maitra et al., 2015). A compound originally developed to treat a certain condition may interact with an additional target conferring a secondary biological effect. The most notable example of a successfully repurposed drug is sildenafil; it was originally developed for the treatment of hypertension but it turned out to be a selective inhibitor of the human phosphodiesterase, becoming one of the first line drugs for erectile dysfunction (Moreland, Goldstein, & Traish, 1998). Sildenafil has also been considered as an adjuvant treatment against TB in the mouse model (Maiga et al., 2012). One of the benefits of using old drugs is that they have already gone through toxicity profiling, *in vitro* and *in vivo* metabolic studies, thus their clinical alternative use can be greatly accelerated. There are several other examples of already available compounds with potential antituberculous activity. These include trimethoprim-sulfamethoxazole (Palomino & Martin, 2016), non-steroidal anti-inflammatory drugs (Maitra et al., 2016) and even antipsychotics (Alsaad et al., 2014). Chapter 6 focused on the antituberculous activity of two carbapenems - ertapenem and faropenem. Despite

some promising clinical reports with meropenem (De Lorenzo et al., 2013), our laboratory results were not very encouraging as we have found high rates of resistance to ertapenem and faropenem among the clinical isolates tested.

The emergence of drug resistance has forced scientists to consider alternative approaches in the fight against TB. These include herbal remedies (Semenya & Maroyi, 2013), bacteriophages (Hatfull, 2014) and the use of nanotechnologies (Yah & Simate, 2015), to mention a few. In Chapter 7 we assessed the activity of two compounds containing gold nanoparticles against *M. tuberculosis*. Both compounds did not inhibit *M. tuberculosis* growth but NP-CML1 did bind to the mycobacterial cell wall. Other authors have successfully tested gold nanoparticles able to cause mycobacterial cell lysis (Zhou et al., 2012) and further studies on the topic should be encouraged. Of particular interest is the role of nanoparticles (not just gold) as potential drug carriers. Researchers from India have encapsulated first line antituberculous drugs with alginate nanoparticles. The concentrations (in the lungs of guinea pigs) of three doses of drug-loaded alginate nanoparticles nebulised 15 days apart were comparable with 45 doses of oral free drugs, making these carriers a very attractive alternative in terms of tissue penetration and compliance (Ahmad, Sharma, & Khuller, 2005).

Final conclusion

Tuberculosis remains a global concern and the emergence of drug resistance has become a major challenge. The main scope of this thesis was to address this challenge, with a particular focus on the applications of WGS to the drug development process and the laboratory testing of new compounds which have potential value in the management of drug resistant TB. Our findings have confirmed the utility of WGS in understanding the pathogenesis of clinical strains. At the same time, we have also integrated WGS within the screening process of new compounds and we have demonstrated its potential role in elucidating the mechanism of action of new and old drugs. Whilst the repurposing of carbapenems did not produce encouraging results, a new promising compound, AD25a, was successfully tested against *M. tuberculosis*. Finally, gold nanoparticles have shown an interesting interaction with the mycobacterial cell wall despite not having a direct antibacterial effect. Hence, this thesis has reflected the complexity of addressing the challenge of drug resistant TB and how a multi strategy approach is essential for the eradication of this global pathogen.

Bibliography

Abu Almakarem, A. S., Heilman, K. L., Conger, H. L., Shtarkman, Y. M., & Rogers, S. O. (2012). Extraction of DNA from plant and fungus tissues in situ. *BMC Res Notes, 5*, 266. doi:10.1186/1756-0500-5-266

Ahmad, Z., Klinkenberg, L. G., Pinn, M. L., Fraig, M. M., Peloquin, C. A., Bishai, W. R., . . . Karakousis, P. C. (2009). Biphasic kill curve of isoniazid reveals the presence of drug-tolerant, not drug-resistant, *Mycobacterium tuberculosis* in the guinea pig. *J Infect Dis, 200*(7), 1136-1143. doi:10.1086/605605

Ahmad, Z., Sharma, S., & Khuller, G. K. (2005). Inhalable alginate nanoparticles as antitubercular drug carriers against experimental tuberculosis. *Int J Antimicrob Agents, 26*(4), 298-303. doi:10.1016/j.ijantimicag.2005.07.012

Al-Balas, Q., Anthony, N. G., Al-Jaidi, B., Alnimr, A., Abbott, G., Brown, A. K., . . . Coxon, G. D. (2009). Identification of 2-aminothiazole-4-carboxylate derivatives active against *Mycobacterium tuberculosis* H37Rv and the beta-ketoacyl-ACP synthase mtFabH. *PLoS One, 4*(5), e5617. doi:10.1371/journal.pone.0005617

Alangaden, G. J., Kreiswirth, B. N., Aouad, A., Khetarpal, M., Igno, F. R., Moghazeh, S. L., . . . Lerner, S. A. (1998). Mechanism of resistance to amikacin and kanamycin in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 42*(5), 1295-1297.

Allen, H. K., Trachsel, J., Looft, T., & Casey, T. A. (2014). Finding alternatives to antibiotics. *Ann N Y Acad Sci, 1323*, 91-100. doi:10.1111/nyas.12468

Almeida Da Silva, P. E., & Palomino, J. C. (2011). Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs. *J Antimicrob Chemother, 66*(7), 1417-1430. doi:10.1093/jac/dkr173

Alnimr, A. M. (2015). Dormancy models for *Mycobacterium tuberculosis*: A minireview. *Braz J Microbiol, 46*(3), 641-647. doi:10.1590/s1517-838246320140507

Alsaad, N., Wilffert, B., van Altena, R., de Lange, W. C., van der Werf, T. S., Kosterink, J. G., & Alffenaar, J. W. (2014). Potential antimicrobial agents for the treatment of multidrug-resistant tuberculosis. *Eur Respir J, 43*(3), 884-897. doi:10.1183/09031936.00113713

Altaf, M., Miller, C. H., Bellows, D. S., & O'Toole, R. (2010). Evaluation of the *Mycobacterium smegmatis* and BCG models for the discovery of *Mycobacterium*

tuberculosis inhibitors. *Tuberculosis (Edinb), 90*(6), 333-337. doi:10.1016/j.tube.2010.09.002

Amaral, L., Viveiros, M., & Kristiansen, J. E. (2006). "Non-Antibiotics": alternative therapy for the management of MDRTB and MRSA in economically disadvantaged countries. *Curr Drug Targets, 7*(7), 887-891.

Amaro, A., Duarte, E., Amado, A., Ferronha, H., & Botelho, A. (2008). Comparison of three DNA extraction methods for *Mycobacterium bovis, Mycobacterium tuberculosis and Mycobacterium avium subsp. avium*. *Lett Appl Microbiol, 47*(1), 8-11. doi:10.1111/j.1472-765X.2008.02372.x

Aminian, M., Shabbeer, A., & Bennett, K. P. (2010). A conformal Bayesian network for classification of *Mycobacterium tuberculosis* complex lineages. *BMC Bioinformatics, 11 Suppl 3*, S4. doi:10.1186/1471-2105-11-s3-s4

Andersson, D. I. (2006). The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr Opin Microbiol, 9*(5), 461-465. doi:10.1016/j.mib.2006.07.002

Ando, H., Kitao, T., Miyoshi-Akiyama, T., Kato, S., Mori, T., & Kirikae, T. (2011). Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol Microbiol, 79*(6), 1615-1628. doi:10.1111/j.1365-2958.2011.07547.x

Andries, K., Verhasselt, P., Guillemont, J., Göhlmann, H. W. H., Neefs, J.-M., Winkler, H., . . . Jarlier, V. (2005). A Diarylquinoline Drug Active on the ATP Synthase of *Mycobacterium tuberculosis*. *Science, 307*(5707), 223-227. doi:10.1126/science.1106753

Arbex, M. A., Siqueira, H. R., D'Ambrosio, L., & Migliori, G. B. (2015). The challenge of managing extensively drug-resistant tuberculosis at a referral hospital in the state of Sao Paulo, Brazil: a report of three cases. *J Bras Pneumol, 41*(6), 554-559. doi:10.1590/s1806-37562015000000299

Aubry, A., Fisher, L. M., Jarlier, V., & Cambau, E. (2006). First functional characterization of a singly expressed bacterial type II topoisomerase: the enzyme from *Mycobacterium tuberculosis*. *Biochem Biophys Res Commun, 348*(1), 158-165. doi:10.1016/j.bbrc.2006.07.017

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., & Struhl, K. (2003). *Current Protocols in Molecular Biology*. New Jersey, USA: John Wiley & Sons.

Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., . . . Jacobs, W. R., Jr. (1994). *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science, 263*(5144), 227-230.

Banu, S., Rahman, S. M., Khan, M. S., Ferdous, S. S., Ahmed, S., Gratz, J., . . . Houpt, E. R. (2014). Discordance across several methods for drug susceptibility testing of drug-resistant *Mycobacterium tuberculosis* isolates in a single laboratory. *J Clin Microbiol, 52*(1), 156-163. doi:10.1128/jcm.02378-13

Basra, S., Anany, H., Brovko, L., Kropinski, A. M., & Griffiths, M. W. (2014). Isolation and characterization of a novel bacteriophage against *Mycobacterium avium subspecies paratuberculosis*. *Arch Virol, 159*(10), 2659-2674. doi:10.1007/s00705-014-2122-3

Bedini, A., Garlassi, E., Stentarelli, C., Petrella, S., Meacci, M., Meccugni, B., . . . Mussini, C. (2016). Multidrug-resistant tuberculosis outbreak in an Italian prison: tolerance of pyrazinamide plus levofloxacin prophylaxis and serial interferon gamma release assays. *New Microbes New Infect, 12*, 45-51. doi:10.1016/j.nmni.2016.03.010

Belanger, A. E., Besra, G. S., Ford, M. E., Mikusova, K., Belisle, J. T., Brennan, P. J., & Inamine, J. M. (1996). The *embAB* genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A, 93*(21), 11919-11924.

Belisle, J. T., & Sonnenberg, M. G. (1998). Isolation of genomic DNA from mycobacteria. *Methods Mol Biol, 101*, 31-44. doi:10.1385/0-89603-471-2:31

Bemer, P., Palicova, F., Rusch-Gerdes, S., Drugeon, H. B., & Pfyffer, G. E. (2002). Multicenter evaluation of fully automated BACTEC Mycobacteria Growth Indicator Tube 960 system for susceptibility testing of *Mycobacterium tuberculosis*. *J Clin Microbiol, 40*(1), 150-154.

Berger, J. M., & Wang, J. C. (1996). Recent developments in DNA topoisomerase II structure and mechanism. *Curr Opin Struct Biol, 6*(1), 84-90.

Bergmire-Sweat, D., Barnett, B. J., Harris, S. L., Taylor, J. P., Mazurek, G. H., & Reddy, V. (1996). Tuberculosis outbreak in a Texas prison, 1994. *Epidemiol Infect, 117*(3), 485-492.

Bergval, I. L., Schuitema, A. R., Klatser, P. R., & Anthony, R. M. (2009). Resistant mutants of *Mycobacterium tuberculosis* selected in vitro do not reflect the in vivo

mechanism of isoniazid resistance. *J Antimicrob Chemother, 64*(3), 515-523. doi:10.1093/jac/dkp237

Bernheim, F. (1940). The effect of salicylate on the oxygen uptake of the tubercle bacillus. *Science, 92*(2383), 204. doi:10.1126/science.92.2383.204

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., & Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol, 43*(3), 717-731.

Bhatt, A., Molle, V., Besra, G. S., Jacobs, W. R., Jr., & Kremer, L. (2007). The *Mycobacterium tuberculosis* FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol Microbiol, 64*(6), 1442-1454. doi:10.1111/j.1365-2958.2007.05761.x

Billington, O. J., McHugh, T. D., & Gillespie, S. H. (1999). Physiological cost of rifampin resistance induced in vitro in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 43*(8), 1866-1869.

Blouin, Y., Cazajous, G., Dehan, C., Soler, C., Vong, R., Hassan, M. O., . . . Vergnaud, G. (2014). Progenitor "*Mycobacterium canettii*" clone responsible for lymph node tuberculosis epidemic, Djibouti. *Emerg Infect Dis, 20*(1), 21-28. doi:10.3201/eid2001.130652

Blouin, Y., Hauck, Y., Soler, C., Fabre, M., Vong, R., Dehan, C., . . . Vergnaud, G. (2012). Significance of the identification in the Horn of Africa of an exceptionally deep branching *Mycobacterium tuberculosis* clade. *PLoS One, 7*(12), e52841. doi:10.1371/journal.pone.0052841

Bogen, E., Loomis, R. N., & Will, D. W. (1950). Para-aminosalicylic acid treatment of tuberculosis; a review. *Am Rev Tuberc, 61*(2), 226-246.

Bohan, C. (2017). The experimental stable of the BCG vaccine: safety, efficacy, proof, and standards, 1921-1933. *Studies in History and Philosophy of Science, 36*(4), 696-721.

Botha, F. J., Sirgel, F. A., Parkin, D. P., van de Wal, B. W., Donald, P. R., & Mitchison, D. A. (1996). Early bactericidal activity of ethambutol, pyrazinamide and the fixed combination of isoniazid, rifampicin and pyrazinamide (Rifater) in patients with pulmonary tuberculosis. *S Afr Med J, 86*(2), 155-158.

Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., . . . Iqbal, Z. (2015). Rapid antibiotic-resistance predictions from genome sequence data for

Staphylococcus aureus and *Mycobacterium tuberculosis*. *Nat Commun, 6*, 10063. doi:10.1038/ncomms10063

Brennan, P. J., & Crick, D. C. (2007). The cell-wall core of *Mycobacterium tuberculosis* in the context of drug discovery. *Curr Top Med Chem, 7*(5), 475-488.

Britton, W. J., Roche, P. W., & Winter, N. (1994). Mechanisms of persistence of mycobacteria. *Trends Microbiol, 2*(8), 284-288.

Brown, A. C., Bryant, J. M., Einer-Jensen, K., Holdstock, J., Houniet, D. T., Chan, J. Z. M., . . . Breuer, J. (2015). Rapid Whole-Genome Sequencing of *Mycobacterium tuberculosis* Isolates Directly from Clinical Samples. *Journal of Clinical Microbiology, 53*(7), 2230-2237. doi:10.1128/jcm.00486-15

Burki, T. (2011). Surveillance, co-infection, resistance: tuberculosis in Europe. *Lancet Infect Dis, 11*(5), 350-351.

Camacho, L. R., (2017). Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Molecular Microbiology, 34*(2), 257-267. doi:10.1046/j.1365-2958.1999.01593.x

Camus, J. C., Pryor, M. J., Medigue, C., & Cole, S. T. (2002). Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology, 148*(Pt 10), 2967-2973. doi:10.1099/00221287-148-10-2967

Canetti, G. (1962). The eradication of tuberculosis: theoretical problems and practical solutions. *Tubercle, 43*, 301-321.

Canetti, G. (1965). Present aspects of bacterial resistance in tuberculosis. *Am Rev Respir Dis, 92*(5), 687-703. doi:10.1164/arrd.1965.92.5.687

Casali, N., Broda, A., Harris, S. R., Parkhill, J., Brown, T., & Drobniewski, F. (2016). Whole Genome Sequence Analysis of a Large Isoniazid-Resistant Tuberculosis Outbreak in London: A Retrospective Observational Study. *PLoS Med, 13*(10), e1002137. doi:10.1371/journal.pmed.1002137

Cavusoglu, C., Karaca-Derici, Y., & Bilgic, A. (2004). In-vitro activity of rifabutin against rifampicin-resistant *Mycobacterium tuberculosis* isolates with known *rpoB* mutations. *Clin Microbiol Infect, 10*(7), 662-665. doi:10.1111/j.1469-0691.2004.00917.x

CDC. (2012). Drug Susceptibility Testing. Retrieved from https://www.cdc.gov/tb/topic/laboratory/drug_testing.htm

Chambers, H. F., Kocagoz, T., Sipit, T., Turner, J., & Hopewell, P. C. (1998). Activity of amoxicillin/clavulanate in patients with tuberculosis. *Clin Infect Dis, 26*(4), 874-877.

Chambers, H. F., Moreau, D., Yajko, D., Miick, C., Wagner, C., Hackbarth, C., . . . Nikaido, H. (1995). Can penicillins and other beta-lactam antibiotics be used to treat tuberculosis? *Antimicrob Agents Chemother, 39*(12), 2620-2624.

Chang, K. C., Yew, W. W., & Zhang, Y. (2011). Pyrazinamide susceptibility testing in *Mycobacterium tuberculosis*: a systematic review with meta-analyses. *Antimicrob Agents Chemother, 55*(10), 4499-4505. doi:10.1128/aac.00630-11

Chanwong, S., Maneekarn, N., Makonkawkeyoon, L., & Makonkawkeyoon, S. (2007). Intracellular growth and drug susceptibility of *Mycobacterium tuberculosis* in macrophages. *Tuberculosis (Edinb), 87*(2), 130-133. doi:10.1016/j.tube.2006.06.001

Chen, J. M. (2016). Mycosins of the Mycobacterial Type VII ESX Secretion System: the Glue That Holds the Party Together. *MBio, 7*(6). doi:10.1128/mBio.02062-16

Cheng, A. F., Yew, W. W., Chan, E. W., Chin, M. L., Hui, M. M., & Chan, R. C. (2004). Multiplex PCR amplimer conformation analysis for rapid detection of *gyrA* mutations in fluoroquinolone-resistant *Mycobacterium tuberculosis* clinical isolates. *Antimicrob Agents Chemother, 48*(2), 596-601.

Chirac, P., & Torreele, E. (2006). Global framework on essential health R&D. *Lancet, 367*(9522), 1560-1561. doi:10.1016/s0140-6736(06)68672-8

Ciociola, A. A., Cohen, L. B., & Kulkarni, P. (2014). How drugs are developed and approved by the FDA: current process and future directions. *Am J Gastroenterol, 109*(5), 620-623. doi:10.1038/ajg.2013.407

Clemens, D. L., Lee, B. Y., Xue, M., Thomas, C. R., Meng, H., Ferris, D., . . . Horwitz, M. A. (2012). Targeted intracellular delivery of antituberculosis drugs to *Mycobacterium tuberculosis*-infected macrophages via functionalized mesoporous silica nanoparticles. *Antimicrob Agents Chemother, 56*(5), 2535-2545. doi:10.1128/aac.06049-11

Cohen, T., & Murray, M. (2004). Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. *Nat Med, 10*(10), 1117-1121. doi:10.1038/nm1110

Colditz, G. A., Brewer, T. F., Berkey, C. S., & et al. (1994). Efficacy of bcg vaccine in the prevention of tuberculosis: Meta-analysis of the published literature. *JAMA, 271*(9), 698-702. doi:10.1001/jama.1994.03510330076038

Cole, S. T. (1999). Learning from the genome sequence of *Mycobacterium tuberculosis* H37Rv. *FEBS Lett, 452*(1-2), 7-10.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., . . . Barrell, B. G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature, 393*(6685), 537-544.

Coll, F., McNerney, R., Preston, M. D., Guerra-Assuncao, J. A., Warry, A., Hill-Cawthorne, G., . . . Clark, T. G. (2015). Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med, 7*(1), 51. doi:10.1186/s13073-015-0164-0

Collins, F. S., Morgan, M., & Patrinos, A. (2003). The Human Genome Project: lessons from large-scale biology. *Science, 300*(5617), 286-290. doi:10.1126/science.1084564

Comas, I., Borrell, S., & Roetzer. (2011). Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet, 44*. doi:10.1038/ng.1038

Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., Kato-Maeda, M., . . . Gagneux, S. (2011). Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet, 44*(1), 106-110. doi:10.1038/ng.1038

Comas, I., Chakravartti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., . . . Gagneux, S. (2010). Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet, 42*(6), 498-503. doi:10.1038/ng.590

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., . . . Gagneux, S. (2013). Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet, 45*(10), 1176-1182. doi:10.1038/ng.2744

Conly, J., Gold, W., & Shafran, S. (1994). Once-daily aminoglycoside dosing: A new look at an old drug. *Can J Infect Dis, 5*(5), 205-206.

Cooksey, R. C., Morlock, G. P., McQueen, A., Glickman, S. E., & Crawford, J. T. (1996). Characterization of streptomycin resistance mechanisms among *Mycobacterium tuberculosis* isolates from patients in New York City. *Antimicrob Agents Chemother, 40*(5), 1186-1188.

Corbett, E. L., Marston, B., Churchyard, G. J., & De Cock, K. M. (2006). Tuberculosis in sub-Saharan Africa: opportunities, challenges, and change in the era of antiretroviral treatment. *Lancet, 367*(9514), 926-937. doi:10.1016/s0140-6736(06)68383-9

Cordillot, M., Dubee, V., Triboulet, S., Dubost, L., Marie, A., Hugonnet, J. E., . . . Mainardi, J. L. (2013). In vitro cross-linking of *Mycobacterium tuberculosis* peptidoglycan by L,D-transpeptidases and inactivation of these enzymes by carbapenems. *Antimicrob Agents Chemother, 57*(12), 5940-5945. doi:10.1128/aac.01663-13

Corrales, R. M., Molle, V., Leiba, J., Mourey, L., de Chastellier, C., & Kremer, L. (2012). Phosphorylation of mycobacterial PcaA inhibits mycolic acid cyclopropanation: consequences for intracellular survival and for phagosome maturation block. *J Biol Chem, 287*(31), 26187-26199. doi:10.1074/jbc.M112.373209

Coxon, G. D., Cooper, C. B., Gillespie, S. H., & McHugh, T. D. (2012). Strategies and challenges involved in the discovery of new chemical entities during early-stage tuberculosis drug discovery. *J Infect Dis, 205 Suppl 2*, S258-264. doi:10.1093/infdis/jis191

Dahlin, E., Nelson, G. M., Haynes, M., & Sargeant, F. (2016). Success rates for product development strategies in new drug development. *J Clin Pharm Ther, 41*(2), 198-202. doi:10.1111/jcpt.12362

Dalhoff, A., Nasu, T., & Okamoto, K. (2003). Beta-lactamase stability of faropenem. *Chemotherapy, 49*(5), 229-236. doi:10.1159/000072446

Dalla Costa, E. R., Ribeiro, M. O., Silva, M. S., Arnold, L. S., Rostirolla, D. C., Cafrune, P. I., . . . Rossetti, M. L. (2009). Correlations of mutations in *katG*, *oxyR-ahpC* and *inhA* genes and in vitro susceptibility in *Mycobacterium tuberculosis* clinical strains segregated by spoligotype families from tuberculosis prevalent countries in South America. *BMC Microbiol, 9*, 39. doi:10.1186/1471-2180-9-39

Daniel, T. M., & Iversen, P. A. (2015). Hippocrates and tuberculosis. *The International Journal of Tuberculosis and Lung Disease, 19*(4), 373-374(372). doi:info:doi/10.5588/ijtld.14.0736

Dauby, N., Muylle, I., Mouchet, F., Sergysels, R., & Payen, M. C. (2011). Meropenem/clavulanate and linezolid treatment for extensively drug-resistant tuberculosis. *Pediatr Infect Dis J, 30*(9), 812-813. doi:10.1097/INF.0b013e3182154b05

David, H. L. (1970). Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*. *Appl Microbiol, 20*(5), 810-814.

Davies, K. (2013). The era of genomic medicine. *Clin Med (Lond), 13*(6), 594-601. doi:10.7861/clinmedicine.13-6-594

De Almeida, I. N., Da Silva Carvalho, W., Rossetti, M. L., Costa, E. R., & De Miranda, S. S. (2013). Evaluation of six different DNA extraction methods for detection of *Mycobacterium tuberculosis* by means of PCR-IS6110: preliminary study. *BMC Res Notes, 6*, 561. doi:10.1186/1756-0500-6-561

De Lorenzo, S., Alffenaar, J. W., Sotgiu, G., Centis, R., D'Ambrosio, L., Tiberi, S., . . . Migliori, G. B. (2013). Efficacy and safety of meropenem-clavulanate added to linezolid-containing regimens in the treatment of MDR-/XDR-TB. *Eur Respir J, 41*(6), 1386-1392. doi:10.1183/09031936.00124312

DeJesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., . . . Ioerger, T. R. (2017). Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome via Saturating Transposon Mutagenesis. *MBio, 8*(1). doi:10.1128/mBio.02133-16

Delogu, G., Sali, M., & Fadda, G. (2013). The biology of *Mycobacterium tuberculosis* infection. *Mediterr J Hematol Infect Dis, 5*(1), e2013070. doi:10.4084/mjhid.2013.070

Dhar, N., Dubee, V., Ballell, L., Cuinet, G., Hugonnet, J. E., Signorino-Gelo, F., . . . McKinney, J. D. (2015). Rapid cytolysis of *Mycobacterium tuberculosis* by faropenem, an orally bioavailable beta-lactam antibiotic. *Antimicrob Agents Chemother, 59*(2), 1308-1319. doi:10.1128/aac.03461-14

Diacon, A. H., & Donald, P. R. (2014). The early bactericidal activity of antituberculosis drugs. *Expert Rev Anti Infect Ther, 12*(2), 223-237. doi:10.1586/14787210.2014.870884

Diacon, A. H., van der Merwe, L., Barnard, M., von Groote-Bidlingmaier, F., Lange, C., Garcia-Basteiro, A. L., . . . Barros-Aguirre, D. (2016). beta-Lactams against Tuberculosis--New Trick for an Old Dog? *N Engl J Med, 375*(4), 393-394. doi:10.1056/NEJMc1513236

Dickinson, J. M., & Mitchison, D. A. (1970). Observations in vitro on the suitability of pyrazinamide for intermittent chemotherapy of tuberculosis. *Tubercle, 51*(4), 389-396.

Dickinson, L. (1947). Effect of streptomycin on experimental tuberculosis in guinea-pigs. *Br J Pharmacol Chemother, 2*(1), 23-26.

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ, 47*, 20-33. doi:10.1016/j.jhealeco.2016.01.012

DiMasi, J. A., Hansen, R. W., & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. *J Health Econ, 22*(2), 151-185. doi:10.1016/s0167-6296(02)00126-1

Dinesh, N., Sharma, S., & Balganesh, M. (2013). Involvement of efflux pumps in the resistance to peptidoglycan synthesis inhibitors in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 57*(4), 1941-1943. doi:10.1128/aac.01957-12

Donald, P. R., & Diacon, A. H. (2015). Para-aminosalicylic acid: the return of an old friend. *Lancet Infect Dis, 15*(9), 1091-1099. doi:10.1016/s1473-3099(15)00263-7

Donald, P. R., Sirgel, F. A., Venter, A., Smit, E., Parkin, D. P., Van de Wal, B. W., . . . Mitchison, D. A. (2002). The early bactericidal activity of streptomycin. *Int J Tuberc Lung Dis, 6*(8), 693-698.

Dondorp, A., Nosten, F., Stepniewska, K., Day, N., & White, N. (2005). Artesunate versus quinine for treatment of severe falciparum malaria: a randomised trial. *Lancet, 366*(9487), 717-725. doi:10.1016/s0140-6736(05)67176-0

Donoghue, H. D., Lee, O. Y.-C., Minnikin, D. E., Besra, G. S., Taylor, J. H., & Spigelman, M. (2010). Tuberculosis in Dr Granville's mummy: a molecular re-examination of the earliest known Egyptian mummy to be scientifically examined and given a medical diagnosis. *Proceedings of the Royal Society B: Biological Sciences, 277*(1678), 51-56. doi:10.1098/rspb.2009.1484

Dooley, K. E., Mitnick, C. D., Ann DeGroote, M., Obuku, E., Belitsky, V., Hamilton, C. D., . . . Nuermberger, E. (2012). Old drugs, new purpose: retooling existing drugs for optimized treatment of resistant tuberculosis. *Clin Infect Dis, 55*(4), 572-581. doi:10.1093/cid/cis487

Doyle, J., & Doyle, J. (1987). A rapid procedure for DNA purification from small quantities of fresh leaf tissue. *Phytochem Bull., 19*, 11-15.

Drobniewski, F., Rusch-Gerdes, S., & Hoffner, S. (2007). Antimicrobial susceptibility testing of *Mycobacterium tuberculosi*s (EUCAST document E.DEF 8.1)--report of the Subcommittee on Antimicrobial Susceptibility Testing of *Mycobacterium tuberculosis* of the European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID). *Clin Microbiol Infect, 13*(12), 1144-1156. doi:10.1111/j.1469-0691.2007.01813.x

EDCTP. (2014). REMoxTB trial results confirm moxifloxacin safety but not shortened tuberculosis treatment - EDCTP. Retrieved from

http://www.edctp.org/news/remoxtb-trial-results-confirm-moxifloxacin-safety-shortened-tuberculosis-treatment/

Editorial. (1973). PAS. *Tubercle, 54*(2), 165-167.

Eldholm, V., Norheim, G., von der Lippe, B., Kinander, W., Dahle, U. R., Caugant, D. A., . . . Balloux, F. (2014). Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. *Genome Biol, 15*(11), 490. doi:10.1186/s13059-014-0490-3

Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., . . . Woodford, N. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clin Microbiol Infect, 23*(1), 2-22. doi:10.1016/j.cmi.2016.11.012

Evangelopoulos, D., & Bhakta, S. (2010). Rapid methods for testing inhibitors of mycobacterial growth. *Methods Mol Biol, 642*, 193-201. doi:10.1007/978-1-60327-279-7_15

Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res, 8*(3), 186-194.

Ewing, B., Hillier, L., Wendl, M. C., & Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res, 8*(3), 175-185.

Faccini, M., Codecasa, L. R., Ciconali, G., Cammarata, S., Borriello, C. R., De Gioia, C., . . . Castaldi, S. (2013). Tuberculosis outbreak in a primary school, Milan, Italy. *Emerg Infect Dis, 19*(3), 485-487. doi:10.3201/eid1902.120527

Falzon, D., Schunemann, H. J., Harausz, E., Gonzalez-Angulo, L., Lienhardt, C., Jaramillo, E., & Weyer, K. (2017). World Health Organization treatment guidelines for drug-resistant tuberculosis, 2016 update. *Eur Respir J, 49*(3). doi:10.1183/13993003.02308-2016

Feuerriegel, S., Schleusener, V., Beckert, P., Kohl, T. A., Miotto, P., Cirillo, D. M., . . . Fellenberg, K. (2015). PhyResSE: web tool delineating *Mycobacterium tuberculosis* antibiotic resistance and lineage from whole-genome sequencing data. *J Clin Microbiol, 53*. doi:10.1128/jcm.00025-15

Filliol, I., Motiwala, A. S., Cavatore, M., Qi, W., Hazbon, M. H., Bobadilla del Valle, M., . . . Alland, D. (2006). Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and

recommendations for a minimal standard SNP set. *J Bacteriol, 188*(2), 759-772. doi:10.1128/jb.188.2.759-772.2006

Fivian-Hughes, A. S., Houghton, J., & Davis, E. O. (2012). *Mycobacterium tuberculosis* thymidylate synthase gene thyX is essential and potentially bifunctional, while *thyA* deletion confers resistance to p-aminosalicylic acid. *Microbiology, 158*(Pt 2), 308-318. doi:10.1099/mic.0.053983-0

Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., Peterson, J., . . . Fraser, C. M. (2002). Whole-Genome Comparison of *Mycobacterium tuberculosis* Clinical and Laboratory Strains. *Journal of Bacteriology, 184*(19), 5479-5490. doi:10.1128/JB.184.19.5479-5490.2002

Folkvardsen, D. B., Svensson, E., Thomsen, V. O., Rasmussen, E. M., Bang, D., Werngren, J., . . . Rigouts, L. (2013). Can molecular methods detect 1% isoniazid resistance in *Mycobacterium tuberculosis*? *J Clin Microbiol, 51*(5), 1596-1599. doi:10.1128/jcm.00472-13

Folkvardsen, D. B., Thomsen, V. O., Rigouts, L., Rasmussen, E. M., Bang, D., Bernaerts, G., . . . Svensson, E. (2013). Rifampin heteroresistance in *Mycobacterium tuberculosis* cultures as detected by phenotypic and genotypic drug susceptibility test methods. *J Clin Microbiol, 51*(12), 4220-4222. doi:10.1128/jcm.01602-13

Fonseca, J. D., Knight, G. M., & McHugh, T. D. (2015). The complex evolution of antibiotic resistance in *Mycobacterium tuberculosis*. *Int J Infect Dis, 32*, 94-100. doi:10.1016/j.ijid.2015.01.014

Foradada, M., Pujol, M. D., Bermudez, J., & Estelrich, J. (2000). Chemical degradation of liposomes by serum components detected by NMR. *Chem Phys Lipids, 104*(2), 133-148.

Forbes, B. A. (2017). Mycobacterial Taxonomy. *J Clin Microbiol, 55*(2), 380-383. doi:10.1128/jcm.01287-16

Ford, C. B., Lin, P. L., Chase, M. R., Shah, R. R., Iartchouk, O., Galagan, J., . . . Fortune, S. M. (2011). Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet, 43*(5), 482-486. doi:10.1038/ng.811

Ford, C. B., Shah, R. R., Maeda, M. K., Gagneux, S., Murray, M. B., Cohen, T., . . . Fortune, S. M. (2013). *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat Genet, 45*(7), 784-790. doi:10.1038/ng.2656

Ford, C. B., Yusim, K., Ioerger, T., Feng, S., Chase, M., Greene, M., . . . Fortune, S. (2012). *Mycobacterium tuberculosis* heterogeneity revealed through whole genome sequencing. *Tuberculosis (Edinb), 92*(3), 194-201. doi:10.1016/j.tube.2011.11.003

Forrellad, M. A., Klepp, L. I., & Gioffré, A. (2013). Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence, 4*. doi:10.4161/viru.22329

Franzblau, S. G., DeGroote, M. A., Cho, S. H., Andries, K., Nuermberger, E., Orme, I. M., . . . Lenaerts, A. J. (2012). Comprehensive analysis of methods used for the evaluation of compounds against *Mycobacterium tuberculosis*. *Tuberculosis (Edinb), 92*(6), 453-488. doi:10.1016/j.tube.2012.07.003

Fraser, C. M., Eisen, J., & Fleishmann. (2000). Comparative genomics and understanding of microbial biology. *Emerg Infect Dis, 6*. doi:10.3201/eid0605.000510

Froman, S., Will, D. W., & Bogen, E. (1954). Bacteriophage active against virulent *Mycobacterium tuberculosis*. I. Isolation and activity. *Am J Public Health Nations Health, 44*(10), 1326-1333.

Frothingham, R., Hills, H. G., & Wilson, K. H. (1994). Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex. *J Clin Microbiol, 32*(7), 1639-1643.

Gagneux, S., Burgos, M. V., & DeRiemer, K. (2006). Impact of bacterial genetics on the transmission of isoniazid resistance in *Mycobacterium tuberculosis*. *PloS Pathol., 2*. doi:10.1371/journal.ppat.0020061

Gagneux, S., Long, C. D., Small, P. M., Van, T., Schoolnik, G. K., & Bohannan, B. J. (2006). The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science, 312*(5782), 1944-1946. doi:10.1126/science.1124410

Gajendiran, G., Balashanmugam, P., Kalaichelvan, P., & Balasubramanian, S. (2016). Multi-drug delivery of tuberculosis drugs by π-back bonded gold nanoparticles with multiblock copolyesters. *Mater. Res. Express, 3 065401*. doi:doi:10.1088/2053-1591/3/6/065401

Galagan, J. E. (2014). Genomic insights into tuberculosis. *Nat Rev Genet, 15*(5), 307-320. doi:10.1038/nrg3664

Galdiero, S., Falanga, A., Vitiello, M., Cantisani, M., Marra, V., & Galdiero, M. (2011). Silver nanoparticles as potential antiviral agents. *Molecules, 16*(10), 8894-8918. doi:10.3390/molecules16108894

Gardy, J. L., Johnston, J. C., Ho Sui, S. J., Cook, V. J., Shah, L., Brodkin, E., . . . Tang, P. (2011). Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med, 364*(8), 730-739. doi:10.1056/NEJMoa1003176

Garnier, T., Eiglmeier, K., Camus, J. C., Medina, N., Mansoor, H., Pryor, M., . . . Hewinson, R. G. (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A, 100*(13), 7877-7882. doi:10.1073/pnas.1130426100

Gazdik, M. A., Bai, G., Wu, Y., & McDonough, K. A. (2009). Rv1675c (cmr) regulates intramacrophage and cyclic AMP-induced gene expression in *Mycobacterium tuberculosis*-complex mycobacteria. *Mol Microbiol, 71*. doi:10.1111/j.1365-2958.2008.06541.x

Gelperina, S., Kisich, K., Iseman, M. D., & Heifets, L. (2005). The potential advantages of nanoparticle drug delivery systems in chemotherapy of tuberculosis. *Am J Respir Crit Care Med, 172*(12), 1487-1490. doi:10.1164/rccm.200504-613PP

Gengenbacher, M., & Kaufmann, S. H. (2012). *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiol Rev, 36*(3), 514-532. doi:10.1111/j.1574-6976.2012.00331.x

Gillespie, S. H. (2001). Antibiotic resistance in the absence of selective pressure. *Int J Antimicrob Agents, 17*(3), 171-176.

Gillespie, S. H. (2002). Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrob Agents Chemother, 46*(2), 267-274.

Gillespie, S. H., Basu, S., Dickens, A. L., O'Sullivan, D. M., & McHugh, T. D. (2005). Effect of subinhibitory concentrations of ciprofloxacin on *Mycobacterium fortuitum* mutation rates. *J Antimicrob Chemother, 56*(2), 344-348. doi:10.1093/jac/dki191

Gillespie, S. H., Crook, A. M., McHugh, T. D., Mendel, C. M., Meredith, S. K., & Murray, S. R. (2014). Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. *N Engl J Med, 371*. doi:10.1056/NEJMoa1407426

Gillespie, S. H., Crook, A. M., McHugh, T. D., Mendel, C. M., Meredith, S. K., Murray, S. R., . . . Nunn, A. J. (2014). Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. *N Engl J Med, 371*(17), 1577-1587. doi:10.1056/NEJMoa1407426

Gillespie, S. H., Gosling, R. D., & Charalambous, B. M. (2002). A reiterative method for calculating the early bactericidal activity of antituberculosis drugs. *Am J Respir Crit Care Med, 166*(1), 31-35. doi:10.1164/rccm.2112077

Glaziou, P., Sismanidis, C., Floyd, K., & Raviglione, M. (2015). Global Epidemiology of Tuberculosis. In *Cold Spring Harb Perspect Med* (Vol. 5).

Glickman, M. S., & Jacobs, W. R., Jr. (2001). Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell, 104*(4), 477-485.

Gobin, J., & Horwitz, M. A. (1996). Exochelins of *Mycobacterium tuberculosis* remove iron from human iron-binding proteins and donate iron to mycobactins in the *M. tuberculosis* cell wall. *J Exp Med, 183*(4), 1527-1532.

Golesi, F., Brignatz, J., Bellenfant, M., Raoult, D., & Drancourt, M. (2013). *Mycobacterium tuberculosis* Beijing outbreak in a school in Marseille, France, 2012. *Euro Surveill, 18*(2).

Gonzalez-y-Merchand, J. A., Estrada-Garcia, I., Colston, M. J., & Cox, R. A. (1996). A novel method for the isolation of mycobacterial DNA. *FEMS Microbiol Lett, 135*(1), 71-77.

Gonzalo, X., Casali, N., Broda, A., Pardieu, C., & Drobniewski, F. (2015). Combination of amikacin and doxycycline against multidrug-resistant and extensively drug-resistant tuberculosis. *Int J Antimicrob Agents, 45*(4), 406-412. doi:10.1016/j.ijantimicag.2014.11.017

Gonzalo, X., & Drobniewski, F. (2013). Is there a place for beta-lactams in the treatment of multidrug-resistant/extensively drug-resistant tuberculosis? Synergy between meropenem and amoxicillin/clavulanate. *J Antimicrob Chemother, 68*(2), 366-369. doi:10.1093/jac/dks395

Green, J. M., & Matthews, R. G. (2007). Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives. *EcoSal Plus, 2*(2). doi:10.1128/ecosalplus.3.6.3.6

Grossman, R., Rotschafer, J., & Tan, J. (2005). Antimicrobial treatment of lower respiratory tract infections in the hospital setting. *118*(7), 29-38.

Guo, H., Seet, Q., Denkin, S., Parsons, L., & Zhang, Y. (2006). Molecular characterization of isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis* from the USA. *J Med Microbiol, 55*(Pt 11), 1527-1531. doi:10.1099/jmm.0.46718-0

Gupta, R., Lavollay, M., Mainardi, J. L., Arthur, M., Bishai, W. R., & Lamichhane, G. (2010). The *Mycobacterium tuberculosis* protein *LdtMt2* is a nonclassical

transpeptidase required for virulence and resistance to amoxicillin. *Nat Med, 16*(4), 466-469. doi:10.1038/nm.2120

Gurumurthy, M., Verma, R., Naftalin, C. M., Hee, K. H., Lu, Q., Tan, K. H., . . . Paton, N. I. (2017). Activity of faropenem with and without rifampicin against *Mycobacterium tuberculosis*: evaluation in a whole-blood bactericidal activity trial. *J Antimicrob Chemother, 72*(7), 2012-2019. doi:10.1093/jac/dkx081

Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omais, B., Marmiesse, M., . . . Vincent, V. (2005). Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog, 1*(1), e5. doi:10.1371/journal.ppat.0010005

Haas, D. (2000). *Mycobacterium tuberculosis*. In M. GL, J. Bennett, & R. Dolin (Eds.), *Mandell, Douglas, and Bennett's principles and practice of infectious diseases* (5th ed., Vol. 2, pp. 2576-2578). Philadelphia, PA: Churchill Livingstone/Elsevier.

Hall, L., Jude, K. P., Clark, S. L., & Wengenack, N. L. (2011). Antimicrobial susceptibility testing of *Mycobacterium tuberculosis* complex for first and second line drugs by broth dilution in a microtiter plate format. *J Vis Exp*(52). doi:10.3791/3094

Hatfull, G. F. (2014). Mycobacteriophages: windows into tuberculosis. *PLoS Pathog, 10*(3), e1003953. doi:10.1371/journal.ppat.1003953

Hazbon, M. H., Brimacombe, M., Bobadilla del Valle, M., Cavatore, M., Guerrero, M. I., Varma-Basil, M., . . . Alland, D. (2006). Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 50*(8), 2640-2649. doi:10.1128/aac.00112-06

Healey, A., Furtado, A., Cooper, T., & Henry, R. J. (2014). Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods, 10*, 21. doi:10.1186/1746-4811-10-21

Heifets, L. (2013). Role of phenotypic methods for drug susceptibility testing of *M. tuberculosis* isolates in the era of MDR and XDR epidemics. In T. D. McHugh (Ed.), *Tuberculosis: laboratory diagnosis and treatment strategies.* London, UK: CABI.

Hershkovitz, I., Donoghue, H. D., Minnikin, D. E., Besra, G. S., Lee, O. Y. C., Gernaey, A. M., . . . Spigelman, M. (2008). Detection and Molecular Characterization of 9000-Year-Old *Mycobacterium tuberculosis* from a Neolithic

Settlement in the Eastern Mediterranean. *PLoS One, 3*(10), e3426. doi:10.1371/journal.pone.0003426

Heym, B., Alzari, P. M., & Honore, N. (1995). Missense mutations in the catalase-peroxidase gene, *katG*, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol Microbiol., 15*. doi:10.1111/j.1365-2958.1995.tb02238.x

Hirsh, A. E., Tsolaki, A. G., DeRiemer, K., Feldman, M. W., & Small, P. M. (2004). Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci U S A, 101*(14), 4871-4876. doi:10.1073/pnas.0305627101

Hoal-van Helden, E. G., Hon, D., Lewis, L. A., Beyers, N., & van Helden, P. D. (2001). Mycobacterial growth in human macrophages: variation according to donor, inoculum and bacterial strain. *Cell Biol Int, 25*(1), 71-81. doi:10.1006/cbir.2000.0679

Honore, N., & Cole, S. T. (1994). Streptomycin resistance in mycobacteria. *Antimicrob Agents Chemother, 38*(2), 238-242.

Hu, Y., Coates, A. R., & Mitchison, D. A. (2003). Sterilizing activities of fluoroquinolones against rifampin-tolerant populations of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 47*(2), 653-657.

Hu, Y., Coates, A. R., & Mitchison, D. A. (2006). Sterilising action of pyrazinamide in models of dormant and rifampicin-tolerant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis, 10*(3), 317-322.

Hu, Y., Hoffner, S., & Jiang, W. (2010). Extensive transmission of isoniazid resistant *M. tuberculosis* and its association with increased multidrug-resistant TB in two rural counties of eastern China: a molecular epidemiological study. *BMC Infect Dis., 10*. doi:10.1186/1471-2334-10-43

Hughes, D., & Karlen, A. (2014). Discovery and preclinical development of new antibiotics. *Ups J Med Sci, 119*(2), 162-169. doi:10.3109/03009734.2014.896437

Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M., & Kimchi-Sarfaty, C. (2009). Silent (synonymous) SNPs: should we care about them? *Methods Mol Biol, 578*, 23-39. doi:10.1007/978-1-60327-411-1_2

Hussain, M. M., Samir, T. M., & Azzazy, H. M. (2013). Unmodified gold nanoparticles for direct and rapid detection of *Mycobacterium tuberculosis* complex. *Clin Biochem, 46*(7-8), 633-637. doi:10.1016/j.clinbiochem.2012.12.020

Illumina. (2014). Understanding Illumina Quality Scores. Retrieved from https://www.illumina.com/content/dam/illumina-

marketing/documents/products/technotes/technote_understanding_quality_scores.pdf

Illumina. (2015). An introduction to next-generation sequencing technology. Retrieved from https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

Imai, T., Ohta, K., Kigawa, H., Kanoh, H., Taniguchi, T., & Tobari, J. (1994). Preparation of high-molecular-weight DNA: application to mycobacterial cells. *Anal Biochem, 222*(2), 479-482. doi:10.1006/abio.1994.1520

Ioerger, T., O'Malley, T., Liao, R., Guinn, K., Hickey, M., Mohaideen, N., . . . Sacchettini, J. C. (2013). Identification of New Drug Targets and Resistance Mechanisms in *Mycobacterium tuberculosis*. *PLOS ONE, 8*(9), e75245. doi:10.1371/journal.pone.0075245

Ioerger, T. R., Feng, Y., Ganesula, K., Chen, X., Dobos, K. M., Fortune, S., . . . Sacchettini, J. C. (2010). Variation among Genome Sequences of H37Rv Strains of *Mycobacterium tuberculosis* from Multiple Laboratories. *J Bacteriol, 192*(14), 3645-3653. doi:10.1128/jb.00166-10

Jaganath, D., Lamichhane, G., & Shah, M. (2016). Carbapenems against *Mycobacterium tuberculosis*: a review of the evidence. *Int J Tuberc Lung Dis, 20*(11), 1436-1447. doi:10.5588/ijtld.16.0498

Jagielski, T., Ignatowska, H., Bakula, Z., Dziewit, L., Napiorkowska, A., Augustynowicz-Kopec, E., . . . Bielecki, J. (2014). Screening for streptomycin resistance-conferring mutations in *Mycobacterium tuberculosis* clinical isolates from Poland. *PLoS One, 9*(6), e100078. doi:10.1371/journal.pone.0100078

Jagielski, T., van Ingen, J., Rastogi, N., Dziadek, J., Mazur, P., & Bielecki, J. (2014). Current Methods in the Molecular Typing of *Mycobacterium tuberculosis* and Other Mycobacteria. *BioMed Research International, 2014*, 21. doi:10.1155/2014/645802

Jain, J. G., Sutherland, C., Nicolau, D. P., & Kuti, J. L. (2014). Stability of ertapenem 100 mg/mL in polypropylene syringes stored at 25, 4, and -20 degrees C. *Am J Health Syst Pharm, 71*(17), 1480-1484. doi:10.2146/ajhp130664

Jindani, A., Aber, V. R., Edwards, E. A., & Mitchison, D. A. (1980). The early bactericidal activity of drugs in patients with pulmonary tuberculosis. *Am Rev Respir Dis, 121*(6), 939-949. doi:10.1164/arrd.1980.121.6.939

Jindani, A., Dore, C. J., & Mitchison, D. A. (2003). Bactericidal and sterilizing activities of antituberculosis drugs during the first 14 days. *Am J Respir Crit Care Med, 167*(10), 1348-1354. doi:10.1164/rccm.200210-1125OC

Johnson, R., Streicher, E. M., Louw, G. E., Warren, R. M., van Helden, P. D., & Victor, T. C. (2006). Drug resistance in *Mycobacterium tuberculosis*. *Curr Issues Mol Biol, 8*(2), 97-111.

Johnsson, K., King, D. S., & Schultz, P. G. (1995). Studies on the Mechanism of Action of Isoniazid and Ethionamide in the Chemotherapy of Tuberculosis. *Journal of the American Chemical Society, 117*(17), 5009-5010. doi:10.1021/ja00122a038

Juarez-Hernandez, R. E., Franzblau, S. G., & Miller, M. J. (2012). Syntheses of mycobactin analogs as potent and selective inhibitors of *Mycobacterium tuberculosis*. *Org Biomol Chem, 10*(37), 7584-7593. doi:10.1039/c2ob26077h

Julian, E., Roldan, M., Sanchez-Chardi, A., Astola, O., Agusti, G., & Luquin, M. (2010). Microscopic cords, a virulence-related characteristic of *Mycobacterium tuberculosis*, are also present in nonpathogenic mycobacteria. *J Bacteriol, 192*(7), 1751-1760. doi:10.1128/jb.01485-09

Kanetsuna, F. (1980). Effect of lysozyme on mycobacteria. *Microbiol Immunol, 24*(12), 1151-1162.

Karakousis, P. (2009). Mechanisms of Action and Resistance of the Antimycobacterial Agents. In D. Mayers (Ed.), *Antimicrobial Drug Resistance* (pp. 271-291). New York: Humana Press.

Karakousis, P., Bishai, W., & Dorman, S. (2004). *Mycobacterium tuberculosis* cell envelope lipids and the host immune response. *Cell Microbiol, 6*(2), 105-116.

Karboul, A., Mazza, A., Gey van Pittius, N. C., Ho, J. L., Brousseau, R., & Mardassi, H. (2008). Frequent homologous recombination events in Mycobacterium tuberculosis PE/PPE multigene families: potential role in antigenic variability. *J Bacteriol, 190*(23), 7838-7846. doi:10.1128/jb.00827-08

Kaser, M., Ruf, M. T., Hauser, J., & Pluschke, G. (2010). Optimized DNA preparation from mycobacteria. *Cold Spring Harb Protoc, 2010*(4), pdb.prot5408. doi:10.1101/pdb.prot5408

Kataria, M., Sethi, M., Kaur, J., Punia, S., & Kumar, K. (2015). Formulation of Nanoparticles Against TB--A Review. *Recent Pat Inflamm Allergy Drug Discov, 9*(2), 120-127.

Kato, S., & Kuwabara, K. (2014). Lessons learned from tuberculosis outbreak cases. *Kekkaku, 89*(2), 77-88.

Kato-Maeda, M., Gagneux, S., Flores, L. L., Kim, E. Y., Small, P. M., Desmond, E. P., & Hopewell, P. C. (2011). Strain classification of *Mycobacterium tuberculosis*: congruence between large sequence polymorphisms and spoligotypes. *Int J Tuberc Lung Dis, 15*(1), 131-133.

Kaushik, A., Makkar, N., Pandey, P., Parrish, N., Singh, U., & Lamichhane, G. (2015). Carbapenems and Rifampin Exhibit Synergy against *Mycobacterium tuberculosis* and *Mycobacterium abscessus*. *Antimicrob Agents Chemother, 59*(10), 6561-6567. doi:10.1128/aac.01158-15

Keel, R. A., Sutherland, C. A., Crandon, J. L., & Nicolau, D. P. (2011). Stability of doripenem, imipenem and meropenem at elevated room temperatures. In *Int J Antimicrob Agents* (Vol. 37, pp. 184-185). Netherlands.

Kibleur, Y., Brochart, H., Schaaf, H. S., Diacon, A. H., & Donald, P. R. (2014). Dose regimen of para-aminosalicylic acid gastro-resistant formulation (PAS-GR) in multidrug-resistant tuberculosis. *Clin Drug Investig, 34*(4), 269-276. doi:10.1007/s40261-014-0172-7

Kisich, K. O., Gelperina, S., Higgins, M. P., Wilson, S., Shipulo, E., Oganesyan, E., & Heifets, L. (2007). Encapsulation of moxifloxacin within poly(butyl cyanoacrylate) nanoparticles enhances efficacy against intracellular *Mycobacterium tuberculosis*. *Int J Pharm, 345*(1-2), 154-162. doi:10.1016/j.ijpharm.2007.05.062

Klinkenberg, L. G., Sutherland, L. A., Bishai, W. R., & Karakousis, P. C. (2008). Metronidazole lacks activity against *Mycobacterium tuberculosis* in an in vivo hypoxic granuloma model of latency. *J Infect Dis, 198*(2), 275-283. doi:10.1086/589515

Kochi, A., Vareldzis, B., & Styblo, K. (1993). Multidrug-resistant tuberculosis and its control. *Res Microbiol, 144*(2), 104-110.

Konno, K., Feldmann, F. M., & McDermott, W. (1967). Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am Rev Respir Dis, 95*(3), 461-469. doi:10.1164/arrd.1967.95.3.461

Korf, J. E., Pynaert, G., Tournoy, K., Boonefaes, T., Van Oosterhout, A., Ginneberge, D., . . . Grooten, J. (2006). Macrophage reprogramming by mycolic acid promotes a tolerogenic response in experimental asthma. *Am J Respir Crit Care Med, 174*(2), 152-160. doi:10.1164/rccm.200507-1175OC

Koser, C. U., Ellington, M. J., Cartwright, E. J., Gillespie, S. H., Brown, N. M., Farrington, M., . . . Peacock, S. J. (2012). Routine use of microbial whole genome

sequencing in diagnostic and public health microbiology. *PLoS Pathog, 8*(8), e1002824. doi:10.1371/journal.ppat.1002824

Kraemer, S. M. (2004). Iron oxide dissolution and solubility in the presence of siderophores. *Aquat. Sci.*(66; 3). doi:10.1007/s00027-003-0690-5

Kruuner, A., Yates, M. D., & Drobniewski, F. A. (2006). Evaluation of MGIT 960-based antimicrobial testing and determination of critical concentrations of first- and second-line antimicrobial drugs with drug-resistant clinical strains of *Mycobacterium tuberculosis*. *J Clin Microbiol, 44*(3), 811-818. doi:10.1128/jcm.44.3.811-818.2006

Krzywinski, M., Schein, J., & Birol, I. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res, 19*. doi:10.1101/gr.092759.109

Kubica, J., Kim, T., & Dunbar, F. (1972). Designation of Strain H37Rv as the Neotype of *Mycobacterium tuberculosis*. doi:doi:10.1099/00207713-22-2-99

Kumar, A., Toledo, J. C., Patel, R. P., Lancaster, J. R., Jr., & Steyn, A. J. (2007). *Mycobacterium tuberculosis* DosS is a redox sensor and DosT is a hypoxia sensor. *Proc Natl Acad Sci U S A, 104*(28), 11568-11573. doi:10.1073/pnas.0705054104

Lambert, P. A. (2002). Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria. *J Appl Microbiol, 92 Suppl*, 46s-54s.

Laszlo, A., Rahman, M., Espinal, M., & Raviglione, M. (2002). Quality assurance programme for drug susceptibility testing of *Mycobacterium tuberculosis* in the WHO/IUATLD Supranational Reference Laboratory Network: five rounds of proficiency testing, 1994-1998. *Int J Tuberc Lung Dis, 6*(9), 748-756.

Lechartier, B., Rybniker, J., Zumla, A., & Cole, S. T. (2014). Tuberculosis drug discovery in the post-post-genomic era. *EMBO Molecular Medicine, 6*(2), 158-168. doi:10.1002/emmm.201201772

Lehmann, J. (1946). Para-aminosalicylic acid in the treatment of tuberculosis. *Lancet, 1*(6384), 15.

Lehmann, J. (1949). The treatment of tuberculosis in Sweden with para-aminosalicylic acid; a review. *Dis Chest, 16*(6), 684-703, illust.

Leon-Torres, A., Novoa-Aponte, L., & Soto, C. Y. (2015). CtpA, a putative Mycobacterium tuberculosis P-type ATPase, is stimulated by copper (I) in the mycobacterial plasma membrane. *Biometals, 28*(4), 713-724. doi:10.1007/s10534-015-9860-x

Lew, J. M., Kapopoulou, A., Jones, L. M., & Cole, S. T. (2011). TubercuList--10 years after. *Tuberculosis (Edinb), 91*(1), 1-7. doi:10.1016/j.tube.2010.09.008

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics., 25*. doi:10.1093/bioinformatics/btp352

Li, J., Gao, X., Luo, T., Wu, J., Sun, G., Liu, Q., . . . Gao, Q. (2014). Association of gyrA/B mutations and resistance levels to fluoroquinolones in clinical isolates of *Mycobacterium tuberculosis*. *Emerg Microbes Infect, 3*(3), e19. doi:10.1038/emi.2014.21

Li, S. Y., Irwin, S. M., Converse, P. J., Mdluli, K. E., Lenaerts, A. J., & Nuermberger, E. L. (2015). Evaluation of moxifloxacin-containing regimens in pathologically distinct murine tuberculosis models. *Antimicrob Agents Chemother, 59*(7), 4026-4030. doi:10.1128/aac.00105-15

Li, S. Y., Tasneen, R., Tyagi, S., Soni, H., Converse, P. J., Mdluli, K., & Nuermberger, E. L. (2017). Bactericidal and sterilizing activity of a novel regimen with bedaquiline, pretomanid, moxifloxacin and pyrazinamide in a murine model of tuberculosis. *Antimicrob Agents Chemother*. doi:10.1128/aac.00913-17

Li, W. J., Li, D. F., Hu, Y. L., Zhang, X. E., Bi, L. J., & Wang, D. C. (2013). Crystal structure of L,D-transpeptidase LdtMt2 in complex with meropenem reveals the mechanism of carbapenem against *Mycobacterium tuberculosis*. In *Cell Res* (Vol. 23, pp. 728-731). England.

Li, Z., Kelley, C., Collins, F., Rouse, D., & Morris, S. (1998). Expression of *katG* in *Mycobacterium tuberculosis* is associated with its growth and persistence in mice and guinea pigs. *J Infect Dis, 177*(4), 1030-1035.

Lillebaek, T., Andersen, A. B., Rasmussen, E. M., Kamper-Jorgensen, Z., Pedersen, M. K., Bjorn-Mortensen, K., . . . Thomsen, V. O. (2013). *Mycobacterium tuberculosis* outbreak strain of Danish origin spreading at worrying rates among greenland-born persons in Denmark and Greenland. *J Clin Microbiol, 51*(12), 4040-4044. doi:10.1128/jcm.01916-13

Lin, S. Y., Desmond, E., Bonato, D., Gross, W., & Siddiqi, S. (2009). Multicenter evaluation of Bactec MGIT 960 system for second-line drug susceptibility testing

of *Mycobacterium tuberculosis* complex. *J Clin Microbiol, 47*(11), 3630-3634. doi:10.1128/jcm.00803-09

Loebel, R. O., Shorr, E., & Richardson, H. B. (1933). The Influence of Foodstuffs upon the Respiratory Metabolism and Growth of Human Tubercle Bacilli. *J Bacteriol, 26*(2), 139-166.

Luria, S. E., & Delbruck, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics, 28*(6), 491-511.

Madison, B. M. (2001). Application of stains in clinical microbiology. *Biotech Histochem, 76*(3), 119-125.

Maguire, H., Brailsford, S., Carless, J., Yates, M., Altass, L., Yates, S., . . . Bothamley, G. (2011). Large outbreak of isoniazid-monoresistant tuberculosis in London, 1995 to 2006: case-control study and recommendations. *Euro Surveill, 16*(13).

Maguire, H., Dale, J. W., & McHugh, T. D. (2002). Molecular epidemiology of tuberculosis in London 1995-7 showing low rate of active transmission. *Thorax, 57*. doi:10.1136/thorax.57.7.617

Maguire, H., Ruddy, M., & Bothamley, G. (2006). Multidrug resistance emerging in North London outbreak. *Thorax, 61*. doi:10.1136/thx.2005.052423

Maher, D., Dye, C., Floyd, K., Pantoja, A., Lonnroth, K., Reid, A., . . . Espinal, M. (2007). Planning to improve global health: the next decade of tuberculosis control. *Bull World Health Organ, 85*(5), 341-347.

Maiga, M., Agarwal, N., Ammerman, N. C., Gupta, R., Guo, H., Maiga, M. C., . . . Bishai, W. R. (2012). Successful shortening of tuberculosis treatment using adjuvant host-directed therapy with FDA-approved phosphodiesterase inhibitors in the mouse model. *PLoS One, 7*(2), e30749. doi:10.1371/journal.pone.0030749

Maitra, A., Bates, S., Kolvekar, T., Devarajan, P. V., Guzman, J. D., & Bhakta, S. (2015). Repurposing-a ray of hope in tackling extensively drug resistance in tuberculosis. *Int J Infect Dis, 32*, 50-55. doi:10.1016/j.ijid.2014.12.031

Maitra, A., Bates, S., Shaik, M., Evangelopoulos, D., Abubakar, I., McHugh, T. D., . . . Bhakta, S. (2016). Repurposing drugs for treatment of tuberculosis: a role for non-steroidal anti-inflammatory drugs. In *Br Med Bull* (Vol. 118, pp. 138-148).

Malik, S., Willby, M., Sikes, D., Tsodikov, O. V., & Posey, J. E. (2012). New insights into fluoroquinolone resistance in *Mycobacterium tuberculosis*: functional genetic analysis of *gyrA* and *gyrB* mutations. *PLoS One, 7*(6), e39754. doi:10.1371/journal.pone.0039754

Manca, C., Paul, S., Barry, C. E., 3rd, Freedman, V. H., & Kaplan, G. (1999). *Mycobacterium tuberculosis* catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro. *Infect Immun, 67*(1), 74-79.

Mangtani, P., Nguipdop-Djomo, P., Keogh, R. H., Trinder, L., Smith, P. G., Fine, P. E., . . . Rodrigues, L. C. (2017). Observational study to estimate the changes in the effectiveness of bacillus Calmette-Guerin (BCG) vaccination with time since vaccination for preventing tuberculosis in the UK. *Health Technol Assess, 21*(39), 1-54. doi:10.3310/hta21390

Manjunatha, U., Boshoff, H. I., & Barry, C. E. (2009). The mechanism of action of PA-824: Novel insights from transcriptional profiling. *Commun Integr Biol, 2*(3), 215-218.

Marianelli, C., Armas, F., Boniotti, M. B., Mazzone, P., Pacciarini, M. L., & Di Marco Lo Presti, V. (2015). Multiple drug-susceptibility screening in *Mycobacterium bovis*: new nucleotide polymorphisms in the embB gene among ethambutol susceptible strains. *Int J Infect Dis, 33*, 39-44. doi:10.1016/j.ijid.2014.12.043

Marks, S. M., Flood, J., Seaworth, B., Hirsch-Moverman, Y., Armstrong, L., Mase, S., . . . Sheeran, K. (2014). Treatment practices, outcomes, and costs of multidrug-resistant and extensively drug-resistant tuberculosis, United States, 2005-2007. *Emerg Infect Dis, 20*(5), 812-821. doi:10.3201/eid2005.131037

Maroyi, A. (2011). An ethnobotanical survey of medicinal plants used by the people in Nhema communal area, Zimbabwe. *J Ethnopharmacol, 136*(2), 347-354. doi:10.1016/j.jep.2011.05.003

Marrero, J., Rhee, K. Y., Schnappinger, D., Pethe, K., & Ehrt, S. (2010). Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection. *Proc Natl Acad Sci U S A, 107*(21), 9819-9824. doi:10.1073/pnas.1000715107

Marsden, J. A. (1954). The single daily dose of sodium para-aminosalicylic acid used in conjunction with streptomycin in tuberculosis. *Med J Aust, 2*(24), 934-935.

Martin, A., Camacho, M., Portaels, F., & Palomino, J. C. (2003). Resazurin microtiter assay plate testing of *Mycobacterium tuberculosis* susceptibilities to second-line drugs: rapid, simple, and inexpensive method. *Antimicrob Agents Chemother, 47*(11), 3616-3619.

Martin, R. S., Sumarah, R. K., & Robart, E. M. (1975). Comparison of four culture media for the isolation of *Mycobacterium tuberculosis*: a 2-year study. *J Clin Microbiol, 2*(5), 438-440.

Mathys, V., Wintjens, R., Lefevre, P., Bertout, J., Singhal, A., Kiass, M., . . . Bifani, P. (2009). Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 53*(5), 2100-2109. doi:10.1128/aac.01197-08

Matsumoto, M., Hashizume, H., Tomishige, T., Kawasaki, M., Tsubouchi, H., Sasaki, H., . . . Komatsu, M. (2006). OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice. *PLoS Med, 3*(11), e466. doi:10.1371/journal.pmed.0030466

Mazandu, G. K., & Mulder, N. J. (2012). Function Prediction and Analysis of *Mycobacterium tuberculosis* Hypothetical Proteins. *International Journal of Molecular Sciences, 13*(6), 7283-7302. doi:10.3390/ijms13067283

McCarthy, O. R. (2001). The key to the sanatoria. *Journal of the Royal Society of Medicine, 94*(8), 413-417.

McCune, R. M., Feldmann, F. M., & McDermott, W. (1966). Microbial persistence. II. Characteristics of the sterile state of tubercle bacilli. *J Exp Med, 123*(3), 469-486.

McDermott, W., & Tompsett, R. (1954). Activation of pyrazinamide and nicotinamide in acidic environments in vitro. *Am Rev Tuberc, 70*(4), 748-754.

McEvoy, C. R., Falmer, A. A., Gey van Pittius, N. C., Victor, T. C., van Helden, P. D., & Warren, R. M. (2007). The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb), 87*(5), 393-404. doi:10.1016/j.tube.2007.05.010

McGrath, M., Gey van Pittius, N. C., van Helden, P. D., Warren, R. M., & Warner, D. F. (2014). Mutation rate and the emergence of drug resistance in *Mycobacterium tuberculosis*. *J Antimicrob Chemother, 69*(2), 292-302. doi:10.1093/jac/dkt364

McMahon, M. D., Rush, J. S., & Thomas, M. G. (2012). Analyses of *MbtB*, *MbtE*, and *MbtF* suggest revisions to the mycobactin biosynthesis pathway in *Mycobacterium tuberculosis*. *J Bacteriol, 194*(11), 2809-2818. doi:10.1128/jb.00088-12

Meletis, G. (2016). Carbapenem resistance: overview of the problem and future perspectives. *Ther Adv Infect Dis, 3*(1), 15-21. doi:10.1177/2049936115621709

Menendez, M. C., Garcia, M. J., Navarro, M. C., Gonzalez-y-Merchand, J. A., Rivera-Gutierrez, S., Garcia-Sanchez, L., & Cox, R. A. (2002). Characterization of an rRNA operon (rrnB) of *Mycobacterium fortuitum* and other mycobacterial species: implications for the classification of mycobacteria. *J Bacteriol, 184*(4), 1078-1088.

Meniche, X., Otten, R., Siegrist, M. S., Baer, C. E., Murphy, K. C., Bertozzi, C. R., & Sassetti, C. M. (2014). Subpolar addition of new cell wall is directed by DivIVA in

mycobacteria. *Proc Natl Acad Sci U S A, 111*(31), E3243-3251. doi:10.1073/pnas.1402158111

Middlebrook, G. (1954). Isoniazid-resistance and catalase activity of tubercle bacilli; a preliminary report. *Am Rev Tuberc, 69*(3), 471-472.

Miesel, L., Weisbrod, T. R., Marcinkeviciene, J. A., Bittman, R., & Jacobs, W. R., Jr. (1998). NADH dehydrogenase defects confer isoniazid resistance and conditional lethality in *Mycobacterium smegmatis*. *J Bacteriol, 180*(9), 2459-2467.

Miethke, M., & Marahiel, M. A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev, 71*(3), 413-451. doi:10.1128/mmbr.00012-07

Mikheecheva, N. E., Zaychikova, M. V., Melerzanov, A. V., & Danilenko, V. N. (2017). A Nonsynonymous SNP Catalog of Mycobacterium tuberculosis Virulence Genes and Its Use for Detecting New Potentially Virulent Sublineages. *Genome Biol Evol, 9*(4), 887-899. doi:10.1093/gbe/evx053

Miles, A. A., Misra, S. S., & Irwin, J. O. (1938). The estimation of the bactericidal power of the blood. *J Hyg (Lond), 38*(6), 732-749.

Minato, Y., Thiede, J. M., Kordus, S. L., McKlveen, E. J., Turman, B. J., & Baughn, A. D. (2015). *Mycobacterium tuberculosis* folate metabolism and the mechanistic basis for para-aminosalicylic acid susceptibility and resistance. *Antimicrob Agents Chemother, 59*(9), 5097-5106. doi:10.1128/aac.00647-15

Mindra, G., Wortham, J. M., Haddad, M. B., & Powell, K. M. (2017). Tuberculosis Outbreaks in the United States, 2009-2015. *Public Health Rep, 132*(2), 157-163. doi:10.1177/0033354916688270

Mitchison, D. A. (1985). The action of antituberculosis drugs in short-course chemotherapy. *Tubercle, 66*(3), 219-225.

Mitchison, D. A., Allen, B. W., & Manickavasagar, D. (1983). Selective Kirchner medium in the culture of specimens other than sputum for mycobacteria. *J Clin Pathol, 36*(12), 1357-1361.

Mitchison, D. A., & Chang, K. C. (2009). Experimental models of tuberculosis: can we trust the mouse? *Am J Respir Crit Care Med, 180*(3), 201-202. doi:10.1164/rccm.200905-0708ED

Moazed, D., & Noller, H. F. (1987). Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature, 327*(6121), 389-394. doi:10.1038/327389a0

Moreland, R. B., Goldstein, I., & Traish, A. (1998). Sildenafil, a novel inhibitor of phosphodiesterase type 5 in human corpus cavernosum smooth muscle cells. *Life Sci, 62*(20), Pl 309-318.

Motiwala, A. S., Dai, Y., Jones-Lopez, E. C., Hwang, S. H., Lee, J. S., Cho, S. N., . . . Alland, D. (2010). Mutations in extensively drug-resistant *Mycobacterium tuberculosis* that do not code for known drug-resistance mechanisms. *J Infect Dis, 201*(6), 881-888. doi:10.1086/650999

MRC. (1955). Various combinations of isoniazid with streptomycin or with PAS in the treatment of pulmonary tuberculosis; seventh report to the Medical Research Council by their Tuberculosis Chemotherapy Trials Committee. *Br Med J, 1*(4911), 435-445.

MSF. (2017, 2017-01-20). TB PRACTECAL: Pioneering new MSF-run clinical trial for drug-resistant TB treatment starts. Retrieved from https://www.msf.org.uk/article/tb-practecal-pioneering-new-msf-run-clinical-trial-drug-resistant-tb-treatment-starts

Mukhopadhyay, S., & Balaji, K. N. (2011). The PE and PPE proteins of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb), 91*(5), 441-447. doi:10.1016/j.tube.2011.04.004

Murray, P., Rosenthal, K., & Pfaller, M. (2005). Mycobacterium. In *Medical Microbiology* (5th ed., pp. 297-298.). Philadelphia, PA: Elsevier-Mosby.

Mushtaq, S., Hope, R., Warner, M., & Livermore, D. M. (2007). Activity of faropenem against cephalosporin-resistant Enterobacteriaceae. *J Antimicrob Chemother, 59*(5), 1025-1030. doi:10.1093/jac/dkm063

Musser, J. M. (2001). Single nucleotide polymorphisms in *Mycobacterium tuberculosis* structural genes. *Emerg Infect Dis., 7*. doi:10.3201/eid0703.017334

Musser, J. M., Amin, A., & Ramaswamy, S. (2000). Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. *Genetics., 155*.

Nagachar, N., & Ratledge, C. (2010). Knocking out salicylate biosynthesis genes in *Mycobacterium smegmatis* induces hypersensitivity to p-aminosalicylate (PAS). *FEMS Microbiol Lett, 311*(2), 193-199. doi:10.1111/j.1574-6968.2010.02091.x

Nahid, P., Dorman, S. E., Alipanah, N., Barry, P. M., Brozek, J. L., Cattamanchi, A., . . . Vernon, A. (2016). Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice

Guidelines: Treatment of Drug-Susceptible Tuberculosis. *Clin Infect Dis, 63*(7), e147-e195. doi:10.1093/cid/ciw376

Neilands, J. B. (1995). Siderophores: structure and function of microbial iron transport compounds. *J Biol Chem, 270*(45), 26723-26726.

Nguyen, L. (2016). Antibiotic resistance mechanisms in M. tuberculosis: an update. *Arch Toxicol, 90*(7), 1585-1604. doi:10.1007/s00204-016-1727-6

NHS, E. (2015). NHS England launches £11.5M strategy to wipe out tuberculosis in the UK. Retrieved from https://www.england.nhs.uk/2015/01/tuberculosis/

Nicoletti, M., Galeffi, C., Messana, I., & Marini-Bettolo, G. B. (1992). Hypoxidaceae. Medicinal uses and the norlignan constituents. *J Ethnopharmacol, 36*(2), 95-101.

Nix, D. E., Majumdar, A. K., & DiNubile, M. J. (2004). Pharmacokinetics and pharmacodynamics of ertapenem: an overview for clinicians. *J Antimicrob Chemother, 53 Suppl 2*, ii23-28. doi:10.1093/jac/dkh205

Nuermberger, E. L., Yoshimatsu, T., Tyagi, S., O'Brien, R. J., Vernon, A. N., Chaisson, R. E., . . . Grosset, J. H. (2004). Moxifloxacin-containing regimen greatly reduces time to culture conversion in murine tuberculosis. *Am J Respir Crit Care Med, 169*(3), 421-426. doi:10.1164/rccm.200310-1380OC

O'Sullivan, D. M., McHugh, T. D., & Gillespie, S. H. (2005). Analysis of rpoB and pncA mutations in the published literature: an insight into the role of oxidative stress in *Mycobacterium tuberculosis* evolution? *J Antimicrob Chemother, 55*(5), 674-679. doi:10.1093/jac/dki069

O'Sullivan, D. M., McHugh, T. D., & Gillespie, S. H. (2010). Mapping the fitness of *Mycobacterium tuberculosis* strains: a complex picture. *Journal of Medical Microbiology, 59*(12), 1533-1535. doi:doi:10.1099/jmm.0.019091-0

O'Toole, R. F., & Gautam, S. S. (2017). Limitations of the *Mycobacterium tuberculosis* reference genome H37Rv in the detection of virulence-related loci. *Genomics*. doi:10.1016/j.ygeno.2017.07.004

Oettinger, T., Jorgensen, M., Ladefoged, A., Haslov, K., & Andersen, P. (1999). Development of the *Mycobacterium bovis* BCG vaccine: review of the historical and biochemical evidence for a genealogical tree. *Tuber Lung Dis, 79*(4), 243-250. doi:10.1054/tuld.1999.0206

Palomino, J. C., & Martin, A. (2016). The potential role of trimethoprim-sulfamethoxazole in the treatment of drug-resistant tuberculosis. *Future Microbiol, 11*(4), 539-547. doi:10.2217/fmb.16.2

Palomino, J. C., Martin, A., Camacho, M., Guerra, H., Swings, J., & Portaels, F. (2002). Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother, 46*(8), 2720-2722.

Palomino, J. C., Martin, A., & Portaels, F. (2007). Rapid drug resistance detection in *Mycobacterium tuberculosis*: a review of colourimetric methods. *Clin Microbiol Infect, 13*(8), 754-762. doi:10.1111/j.1469-0691.2007.01698.x

Palomino, J. C., Ramos, D. F., & da Silva, P. A. (2009). New anti-tuberculosis drugs: strategies, sources and new molecules. *Curr Med Chem, 16*(15), 1898-1904.

Pandey, R., & Khuller, G. K. (2006). Nanotechnology based drug delivery system(s) for the management of tuberculosis. *Indian J Exp Biol, 44*(5), 357-366.

Pandey, R., Zahoor, A., Sharma, S., & Khuller, G. K. (2003). Nanoparticle encapsulated antitubercular drugs as a potential oral drug delivery system against murine tuberculosis. *Tuberculosis (Edinb), 83*(6), 373-378.

Pankhurst, L. J., del Ojo Elias, C., Votintseva, A. A., Walker, T. M., Cole, K., Davies, J., . . . Crook, D. W. (2016). Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *The Lancet Respiratory Medicine, 4*(1), 49-58. doi:10.1016/s2213-2600(15)00466-x

Papaventsis, D., Casali, N., Kontsevaya, I., Drobniewski, F., Cirillo, D. M., & Nikolayevskyy, V. (2017). Whole genome sequencing of *Mycobacterium tuberculosis* for detection of drug resistance: a systematic review. *Clin Microbiol Infect, 23*(2), 61-68. doi:10.1016/j.cmi.2016.09.008

Papp-Wallace, K. M., Endimiani, A., Taracila, M. A., & Bonomo, R. A. (2011). Carbapenems: past, present, and future. *Antimicrob Agents Chemother, 55*(11), 4943-4960. doi:10.1128/aac.00296-11

Pasca, M. R., Guglierame, P., Arcesi, F., Bellinzoni, M., De Rossi, E., & Riccardi, G. (2004). Rv2686c-Rv2687c-Rv2688c, an ABC fluoroquinolone efflux pump in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 48*(8), 3175-3178. doi:10.1128/aac.48.8.3175-3178.2004

Payen, M. C., De Wit, S., Martin, C., Sergysels, R., Muylle, I., Van Laethem, Y., & Clumeck, N. (2012). Clinical use of the meropenem-clavulanate combination for extensively drug-resistant tuberculosis [Case study]. *Int J Tuberc Lung Dis, 16*(4), 558-560. doi:10.5588/ijtld.11.0414

Penuelas-Urquides, K., Villarreal-Trevino, L., Silva-Ramirez, B., Rivadeneyra-Espinoza, L., Said-Fernandez, S., & de Leon, M. B. (2013). Measuring of

Mycobacterium tuberculosis growth. A correlation of the optical measurements with colony forming units. *Braz J Microbiol, 44*(1), 287-289. doi:10.1590/s1517-83822013000100042

Perna, N. T., Plunkett, G., 3rd, Burland, V., Mau, B., Glasner, J. D., Rose, D. J., . . . Blattner, F. R. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature, 409*(6819), 529-533. doi:10.1038/35054089

PHE. (2015). Tuberculosis in England: annual report. Retrieved from https://www.gov.uk/government/publications/tuberculosis-in-england-annual-report

PHE. (2017). England world leaders in the use of whole genome sequencing to diagnose TB. Retrieved from https://www.gov.uk/government/news/england-world-leaders-in-the-use-of-whole-genome-sequencing-to-diagnose-tb

Phillips, P. P., Dooley, K. E., Gillespie, S. H., Heinrich, N., Stout, J. E., Nahid, P., . . . Hoelscher, M. (2016). A new trial design to accelerate tuberculosis drug development: the Phase IIC Selection Trial with Extended Post-treatment follow-up (STEP). *BMC Med, 14*, 51. doi:10.1186/s12916-016-0597-3

Pitaksajjakul, P., Wongwit, W., Punprasit, W., Eampokalap, B., Peacock, S., & Ramasoota, P. (2005). Mutations in the *gyrA* and *gyrB* genes of fluoroquinolone-resistant *Mycobacterium tuberculosis* from TB patients in Thailand. *Southeast Asian J Trop Med Public Health, 36 Suppl 4*, 228-237.

Pope, C. F., McHugh, T. D., & Gillespie, S. H. (2010). Methods to determine fitness in bacteria. *Methods Mol Biol, 642*, 113-121. doi:10.1007/978-1-60327-279-7_9

Pope, C. F., O'Sullivan, D. M., & McHugh, T. D. (2008). A practical guide to measuring mutation rates in antibiotic resistance. *Antimicrob Agents Chemother., 52*. doi:10.1128/aac.01152-07

Poulet, S., & Cole, S. T. (1995a). Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in Mycobacterium tuberculosis. *Arch Microbiol, 163*(2), 87-95.

Poulet, S., & Cole, S. T. (1995b). Repeated DNA sequences in mycobacteria. *Arch Microbiol, 163*(2), 79-86.

Ramaswanny, S. W., Amin, A. G., & Goksel, S. (2000). Molecular genetic analysis of nucleotide polymorphisms associated with ethambutol resistance in human isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother., 44*. doi:10.1128/aac.44.2.326-336.2000

Ramirez-Busby, S. M., Rodwell, T. C., Fink, L., Catanzaro, D., Jackson, R. L., Pettigrove, M., . . . Valafar, F. (2017). A Multinational Analysis of Mutations and

Heterogeneity in PZase, RpsA, and PanD Associated with Pyrazinamide Resistance in M/XDR *Mycobacterium tuberculosis*. *Sci Rep, 7*(1), 3790. doi:10.1038/s41598-017-03452-y

Ramon-Garcia, S., Mick, V., Dainese, E., Martin, C., Thompson, C. J., De Rossi, E., . . . Ainsa, J. A. (2012). Functional and genetic characterization of the tap efflux pump in *Mycobacterium bovis* BCG. *Antimicrob Agents Chemother, 56*(4), 2074-2083. doi:10.1128/aac.05946-11

Ranganathan, S., Bai, G., & Lyubetskaya, A. (2016). Characterization of a cAMP responsive transcription factor, Cmr (Rv1675c), in TB complex mycobacteria reveals overlap with the DosR (DevR) dormancy regulon. *Nucleic Acids Res, 44*. doi:10.1093/nar/gkv889

Ratledge, C., & Brown, K. A. (1972). Inhibition of mycobactin formation in Mycobacterium smegmatis by p-aminosalicylate. A new proposal for the mode of action of p-aminosalicylate. *Am Rev Respir Dis, 106*(5), 774-776. doi:10.1164/arrd.1972.106.5.774

Ratledge, C., & Ewing, M. (1996). The occurrence of carboxymycobactin, the siderophore of pathogenic mycobacteria, as a second extracellular siderophore in *Mycobacterium smegmatis. Microbiology, 142 (Pt 8)*, 2207-2212. doi:10.1099/13500872-142-8-2207

Raviglione, M. C., & Uplekar, M. W. (2006). WHO's new Stop TB Strategy. *Lancet, 367*(9514), 952-955. doi:10.1016/s0140-6736(06)68392-x

Rawat, R., Whitty, A., & Tonge, P. J. (2003). The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: adduct affinity and drug resistance. *Proc Natl Acad Sci U S A, 100*(24), 13881-13886. doi:10.1073/pnas.2235848100

Raynaud, C., Guilhot, C., & Rauzier, J. (2002). Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol., 45*. doi:10.1046/j.1365-2958.2002.03009.x

Reichau, S., Jiao, W., Walker, S. R., Hutton, R. D., Baker, E. N., & Parker, E. J. (2011). Potent inhibitors of a shikimate pathway enzyme from *Mycobacterium tuberculosis*: combining mechanism- and modeling-based design. *J Biol Chem, 286*(18), 16197-16207. doi:10.1074/jbc.M110.211649

Retzinger, G. S., Meredith, S. C., Hunter, R. L., Takayama, K., & Kezdy, F. J. (1982). Identification of the physiologically active state of the mycobacterial glycolipid

trehalose 6,6'-dimycolate and the role of fibrinogen in the biologic activities of trehalose 6,6'-dimycolate monolayers. *J Immunol, 129*(2), 735-744.

Reyrat, J. M., & Kahn, D. (2001). *Mycobacterium smegmatis*: an absurd model for tuberculosis? In *Trends Microbiol* (Vol. 9, pp. 472-474). England.

Riccardi, N., Del Puente, F., Magne, F., Taramasso, L., & Di Biagio, A. (2017). Bedaquiline: a new hope for shorter and better anti-tuberculosis regimens. *Recent Pat Antiinfect Drug Discov*. doi:10.2174/1574891x12666170619101904

Riska, P. F., Jacobs, W. R., Jr., & Alland, D. (2000). Molecular determinants of drug resistance in tuberculosis. *Int J Tuberc Lung Dis, 4*(2 Suppl 1), S4-10.

Roberts, G. D., Goodman, N. L., Heifets, L., Larsh, H. W., Lindner, T. H., McClatchy, J. K., . . . Wright, P. (1983). Evaluation of the BACTEC radiometric method for recovery of mycobacteria and drug susceptibility testing of *Mycobacterium tuberculosis* from acid-fast smear-positive specimens. *J Clin Microbiol, 18*(3), 689-696.

Robin, J. D., Ludlow, A. T., LaRanger, R., Wright, W. E., & Shay, J. W. (2016). Comparison of DNA Quantification Methods for Next Generation Sequencing. *Sci Rep, 6*, 24067. doi:10.1038/srep24067

Roetzer, A., Diel, R., Kohl, T. A., Rückert, C., Nübel, U., Blom, J., . . . Niemann, S. (2013). Whole Genome Sequencing versus Traditional Genotyping for Investigation of a *Mycobacterium tuberculosis* Outbreak: A Longitudinal Molecular Epidemiological Study. *PLOS Medicine, 10*(2), e1001387. doi:10.1371/journal.pmed.1001387

Rogers, S. O., & Bendich, A. J. (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol Biol, 5*(2), 69-76. doi:10.1007/bf00020088

Roy, A., Eisenhut, M., Harris, R. J., Rodrigues, L. C., Sridhar, S., Habermann, S., . . . Abubakar, I. (2014). Effect of BCG vaccination against *Mycobacterium tuberculosis* infection in children: systematic review and meta-analysis. *BMJ, 349*. doi:10.1136/bmj.g4643

Ruddy, M. C., Davies, A. P., & Yates, M. D. (2004). Outbreak of isoniazid resistant tuberculosis in north London. *Thorax., 59*. doi:10.1136/thx.2003.010405

Rusch-Gerdes, S., Pfyffer, G. E., Casal, M., Chadwick, M., & Siddiqi, S. (2006). Multicenter laboratory validation of the BACTEC MGIT 960 technique for testing susceptibilities of *Mycobacterium tuberculosis* to classical second-line drugs and

newer antimicrobials. *J Clin Microbiol, 44*(3), 688-692. doi:10.1128/jcm.44.3.688-692.2006

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., & Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics, 16*(10), 944-945.

Sakula, A. (1982). Robert Koch: centenary of the discovery of the tubercle bacillus, 1882. *Thorax, 37*(4), 246.

Sala, C., & Hartkoorn, R. C. (2011). Tuberculosis drugs: new candidates and how to find more. *Future Microbiology, 6*(6), 617-633. doi:10.2217/fmb.11.46

Salvatore, P. P., Becerra, M. C., Abel zur Wiesch, P., Hinkley, T., Kaur, D., Sloutsky, A., & Cohen, T. (2016). Fitness Costs of Drug Resistance Mutations in Multidrug-Resistant *Mycobacterium tuberculosis*: A Household-Based Case-Control Study. *J Infect Dis, 213*(1), 149-155. doi:10.1093/infdis/jiv347

Sander, P., Springer, B., Prammananan, T., Sturmfels, A., Kappler, M., Pletschette, M., & Bottger, E. C. (2002). Fitness cost of chromosomal drug resistance-conferring mutations. *Antimicrob Agents Chemother, 46*(5), 1204-1211.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A, 74*(12), 5463-5467.

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2001). Comprehensive identification of conditionally essential genes in mycobacteria. doi:10.1073/pnas.231275498

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology, 48*(1), 77-84. doi:10.1046/j.1365-2958.2003.03425.x

Scarselli, M., Giuliani, M. M., Adu-Bobie, J., Pizza, M., & Rappuoli, R. The impact of genomics on vaccine design. *Trends in Biotechnology, 23*(2), 84-91. doi:10.1016/j.tibtech.2004.12.008

Scarselli, M., Giuliani, M. M., Adu-Bobie, J., Pizza, M., & Rappuoli, R. (2005). The impact of genomics on vaccine design. *Trends in Biotechnology, 23*(2), 84-91. doi:10.1016/j.tibtech.2004.12.008

Schmitt, A. O., & Herzel, H. (1997). Estimating the entropy of DNA sequences. *J Theor Biol, 188*(3), 369-377. doi:10.1006/jtbi.1997.0493

Schurek, K. N., Wiebe, R., Karlowsky, J. A., Rubinstein, E., Hoban, D. J., & Zhanel, G. G. (2007). Faropenem: review of a new oral penem. *Expert Rev Anti Infect Ther, 5*(2), 185-198. doi:10.1586/14787210.5.2.185

Sekizuka, T., Yamashita, A., Murase, Y., Iwamoto, T., Mitarai, S., Kato, S., & Kuroda, M. (2015). TGS-TB: Total Genotyping Solution for *Mycobacterium tuberculosis* Using Short-Read Whole-Genome Sequencing. *PLoS One, 10*(11), e0142951. doi:10.1371/journal.pone.0142951

Semenya, S. S., & Maroyi, A. (2013). Medicinal plants used for the treatment of tuberculosis by Bapedi traditional healers in three districts of the Limpopo Province, South Africa. *Afr J Tradit Complement Altern Med, 10*(2), 316-323.

Semete, B., Booysen, L., Lemmer, Y., Kalombo, L., Katata, L., Verschoor, J., & Swai, H. S. (2010). In vivo evaluation of the biodistribution and safety of PLGA nanoparticles as drug delivery systems. *Nanomedicine, 6*(5), 662-671. doi:10.1016/j.nano.2010.02.002

Semete, B., Booysen, L. I., Kalombo, L., Venter, J. D., Katata, L., Ramalapa, B., . . . Swai, H. (2010). In vivo uptake and acute immune response to orally administered chitosan and PEG coated PLGA nanoparticles. *Toxicol Appl Pharmacol, 249*(2), 158-165. doi:10.1016/j.taap.2010.09.002

Sepkowitz, K. A. (1996). How contagious is tuberculosis? *Clin Infect Dis, 23*(5), 954-962.

Shabbeer, A., Cowan, L. S., Ozcaglar, C., Rastogi, N., Vandenberg, S. L., Yener, B., & Bennett, K. P. (2012). TB-Lineage: an online tool for classification and analysis of strains of *Mycobacterium tuberculosis* complex. *Infect Genet Evol, 12*(4), 789-797. doi:10.1016/j.meegid.2012.02.010

Sharma, A., Sharma, S., & Khuller, G. K. (2004). Lectin-functionalized poly (lactide-co-glycolide) nanoparticles as oral/aerosolized antitubercular drug carriers for treatment of tuberculosis. *J Antimicrob Chemother, 54*(4), 761-766. doi:10.1093/jac/dkh411

Shean, K., Streicher, E., Pieterson, E., Symons, G., van Zyl Smit, R., Theron, G., . . . Dheda, K. (2013). Drug-associated adverse events and their relationship with outcomes in patients receiving treatment for extensively drug-resistant tuberculosis in South Africa. *PLoS One, 8*(5), e63057. doi:10.1371/journal.pone.0063057

Shimono, N., Morici, L., Casali, N., Cantrell, S., Sidders, B., Ehrt, S., & Riley, L. W. (2003). Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the mce1 operon. *Proc Natl Acad Sci U S A, 100*(26), 15918-15923. doi:10.1073/pnas.2433882100

Shinnick, T. M., & Good, R. C. (1994). Mycobacterial taxonomy. *Eur J Clin Microbiol Infect Dis, 13*(11), 884-901.

Shoeb, H. A., Bowman, B. U., Jr., Ottolenghi, A. C., & Merola, A. J. (1985). Peroxidase-mediated oxidation of isoniazid. *Antimicrob Agents Chemother, 27*(3), 399-403.

Shorten, R. (2011). *The Molecular Epidemiology of Mycobacterium tuberculosis in North London.* (PhD), University College London, Retrieved from http://discovery.ucl.ac.uk/1310448/1/1310448.pdf

Shorten, R. J., McGregor, A. C., Platt, S., & Jenkins, C. (2013). When is an outbreak not an outbreak? Fit, divergent strains of *Mycobacterium tuberculosis* display independent evolution of drug resistance in a large London outbreak. *J Antimicrob Chemother, 68*. doi:10.1093/jac/dks430

Singh, B., & Mitchison, D. A. (1955). Bactericidal activity of streptomycin and isoniazid in combination with p-aminosalicylic acid against *Mycobacterium tuberculosis*. *J Gen Microbiol, 12*(1), 76-84. doi:10.1099/00221287-12-1-76

Sledzik, P. S., & Bellantoni, N. (1994). Bioarcheological and biocultural evidence for the New England vampire folk belief. *American Journal of Physical Anthropology, 94*(2), 269-274. doi:10.1002/ajpa.1330940210

Smith, C. M., Trienekens, S. C., Anderson, C., Lalor, M. K., Brown, T., Story, A., . . . Maguire, H. (2017). Twenty years and counting: epidemiology of an outbreak of isoniazid-resistant tuberculosis in England and Wales, 1995 to 2014. *Euro Surveill, 22*(8). doi:10.2807/1560-7917.es.2017.22.8.30467

Smith, G. B., Dezeny, G. C., & Douglas, A. W. (1990). Stability and kinetics of degradation of imipenem in aqueous solution. *J Pharm Sci, 79*(8), 732-740.

Smith, L. J., Bochkareva, A., Rolfe, M. D., Hunt, D. M., Kahramanoglou, C., Braun, Y., . . . Green, J. (2017). Cmr is a redox-responsive regulator of DosR that contributes to *M. tuberculosis* virulence. *Nucleic Acids Res, 45*(11), 6600-6612. doi:10.1093/nar/gkx406

Smith, R. (2009). Eradication of tuberculosis by 2050 impossible without new vaccine. *BMJ, 338*, b1291. doi:10.1136/bmj.b1291

Snow, G. A. (1970). Mycobactins: iron-chelating growth factors from mycobacteria. *Bacteriol Rev, 34*(2), 99-125.

Soroka, D., Li de la Sierra-Gallay, I., Dubee, V., Triboulet, S., van Tilbeurgh, H., Compain, F., . . . Arthur, M. (2015). Hydrolysis of clavulanate by *Mycobacterium*

tuberculosis beta-lactamase BlaC harboring a canonical SDN motif. *Antimicrob Agents Chemother, 59*(9), 5714-5720. doi:10.1128/aac.00598-15

Sotgiu, G., Tiberi, S., Centis, R., D'Ambrosio, L., Fuentes, Z., Zumla, A., & Migliori, G. B. (2017). Applicability of the shorter 'Bangladesh regimen' in high multidrug-resistant tuberculosis settings. *Int J Infect Dis, 56*, 190-193. doi:10.1016/j.ijid.2016.10.021

Spargo, B. J., Crowe, L. M., Ioneda, T., Beaman, B. L., & Crowe, J. H. (1991). Cord factor (alpha,alpha-trehalose 6,6'-dimycolate) inhibits fusion between phospholipid vesicles. *Proc Natl Acad Sci U S A, 88*(3), 737-740.

Spies, F. S., da Silva, P. E., Ribeiro, M. O., Rossetti, M. L., & Zaha, A. (2008). Identification of mutations related to streptomycin resistance in clinical isolates of *Mycobacterium tuberculosis* and possible involvement of efflux mechanism. *Antimicrob Agents Chemother, 52*(8), 2947-2949. doi:10.1128/aac.01570-07

Springer, B., Kidan, Y. G., Prammananan, T., Ellrott, K., Bottger, E. C., & Sander, P. (2001). Mechanisms of streptomycin resistance: selection of mutations in the *16S* rRNA gene conferring resistance. *Antimicrob Agents Chemother, 45*(10), 2877-2884. doi:10.1128/aac.45.10.2877-2884.2001

Springer, B., Lucke, K., Calligaris-Maibach, R., Ritter, C., & Bottger, E. C. (2009). Quantitative drug susceptibility testing of *Mycobacterium tuberculosis* by use of MGIT 960 and EpiCenter instrumentation. *J Clin Microbiol, 47*(6), 1773-1780. doi:10.1128/jcm.02501-08

Sreevatsan, S., Pan, X., & Stockbauer, K. E. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionary recent global dissemination. *Proc Natl Acad Sci U S A., 94*. doi:10.1073/pnas.94.18.9869

Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S., & Musser, J. M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A, 94*(18), 9869-9874.

Sreevatsan, S., Pan, X., Stockbauer, K. E., Williams, D. L., Kreiswirth, B. N., & Musser, J. M. (1996). Characterization of *rpsL* and *rrs* mutations in streptomycin-resistant *Mycobacterium tuberculosis* isolates from diverse geographic localities. *Antimicrob Agents Chemother, 40*(4), 1024-1026.

Sreevatsan, S., Stockbauer, K. E., Pan, X., Kreiswirth, B. N., Moghazeh, S. L., Jacobs, W. R., Jr., . . . Musser, J. M. (1997). Ethambutol resistance in

Mycobacterium tuberculosis: critical role of *embB* mutations. *Antimicrob Agents Chemother, 41*(8), 1677-1681.

Stagg, H. R., Lipman, M. C., McHugh, T. D., & Jenkins, H. E. (2017). Isoniazid-resistant tuberculosis: a cause for concern? *Int J Tuberc Lung Dis, 21*(2), 129-139. doi:10.5588/ijtld.16.0716

Stamatakis, A. (2014a). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*. doi:10.1093/bioinformatics/btu033

Stamatakis, A. (2014b). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. In *Bioinformatics* (Vol. 30, pp. 1312-1313).

Steenken, W., Oatway, W. H., & Petroff, S. A. (1934). Biological studies of the tubercle bacillus: Dissociation and pathogenicity of the R and S variants of the human tubercle bacillus (H37). *J Exp Med, 60*(4), 515-540.

Steingart, K. R., Henry, M., Ng, V., Hopewell, P. C., Ramsay, A., Cunningham, J., . . . Pai, M. (2006). Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis, 6*(9), 570-581. doi:10.1016/s1473-3099(06)70578-3

Steingart, K. R., Sohn, H., Schiller, I., Kloda, L. A., Boehme, C. C., Pai, M., & Dendukuri, N. (2013). Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev*(1), Cd009593. doi:10.1002/14651858.CD009593.pub2

Stover, C. K., Warrener, P., VanDevanter, D. R., Sherman, D. R., Arain, T. M., Langhorne, M. H., . . . Baker, W. R. (2000). A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature, 405*(6789), 962-966. doi:10.1038/35016103

Stucki, D., Ballif, M., Bodmer, T., Coscolla, M., Maurer, A. M., Droz, S., . . . Fenner, L. (2014). Tracking a tuberculosis outbreak over 21 years: strain-specific single-nucleotide polymorphism typing combined with targeted whole-genome sequencing. *J Infect Dis, 211*(8), 1306-1316. doi:10.1093/infdis/jiu601

Sun, Z., Zhang, J., Zhang, X., Wang, S., Zhang, Y., & Li, C. (2008). Comparison of gyrA gene mutations between laboratory-selected ofloxacin-resistant *Mycobacterium tuberculosis* strains and clinical isolates. *Int J Antimicrob Agents, 31*(2), 115-121. doi:10.1016/j.ijantimicag.2007.10.014

Takayama, K., & Kilburn, J. O. (1989). Inhibition of synthesis of arabinogalactan by ethambutol in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother, 33*(9), 1493-1499.

Takayama, K., Wang, C., & Besra, G. S. (2005). Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev, 18*(1), 81-101. doi:10.1128/cmr.18.1.81-101.2005

TBalliance. (2017). Clinical Portfolio. Retrieved from https://www.tballiance.org/portfolio

Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., . . . Bodmer, T. (1993). Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet, 341*(8846), 647-650.

ThermoScientific. (2009). Technical Bulletin - T042. Retrieved from http://www.nhm.ac.uk/content/dam/nhmwww/our-science/dpts-facilities-staff/Coreresearchlabs/nanodrop.pdf

Thomson, L. M., Traore, H., Yesilkaya, H., Doig, C., Steingrimsdottir, H., Garcia, L., & Forbes, K. J. (2005). An extremely rapid and simple DNA-release method for detection of *M. tuberculosis* from clinical specimens. *J Microbiol Methods, 63*(1), 95-98. doi:10.1016/j.mimet.2005.03.005

Tiberi, S., D'Ambrosio, L., De Lorenzo, S., Viggiani, P., Centis, R., Sotgiu, G., . . . Migliori, G. B. (2016). Ertapenem in the treatment of multidrug-resistant tuberculosis: first clinical experience. In *Eur Respir J* (Vol. 47, pp. 333-336). England.

Tiberi, S., Sotgiu, G., D'Ambrosio, L., Centis, R., Abdo Arbex, M., Alarcon Arrascue, E., . . . Migliori, G. B. (2016). Comparison of effectiveness and safety of imipenem/clavulanate- versus meropenem/clavulanate-containing regimens in the treatment of MDR- and XDR-TB. *Eur Respir J, 47*(6), 1758-1766. doi:10.1183/13993003.00214-2016

Timmins, G. S., & Deretic, V. (2006). Mechanisms of action of isoniazid. *Mol Microbiol, 62*(5), 1220-1227. doi:10.1111/j.1365-2958.2006.05467.x

Timmins, G. S., Master, S., Rusnak, F., & Deretic, V. (2004). Nitric oxide generated from isoniazid activation by *KatG*: source of nitric oxide and activity against *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 48*(8), 3006-3009. doi:10.1128/aac.48.8.3006-3009.2004

Titball, R. W. (1993). Bacterial phospholipases C. *Microbiol Rev., 57*.

Tomasicchio, M., Theron, G., Pietersen, E., Streicher, E., Stanley-Josephs, D., van Helden, P., . . . Dheda, K. (2016). The diagnostic accuracy of the MTBDRplus and MTBDRsl assays for drug-resistant TB detection when performed on sputum and culture isolates. *6*, 17850. doi:10.1038/srep17850

Tortoli, E., Benedetti, M., Fontanelli, A., & Simonetti, M. T. (2002). Evaluation of automated BACTEC MGIT 960 system for testing susceptibility of *Mycobacterium tuberculosis* to four major antituberculous drugs: comparison with the radiometric BACTEC 460TB method and the agar plate method of proportion. *J Clin Microbiol, 40*(2), 607-610.

Tortoli, E., Cichero, P., Piersimoni, C., Simonetti, M. T., Gesu, G., & Nista, D. (1999). Use of BACTEC MGIT 960 for recovery of mycobacteria from clinical specimens: multicenter study. *J Clin Microbiol, 37*(11), 3578-3582.

Traore, H., Fissette, K., Bastian, I., Devleeschouwer, M., & Portaels, F. (2000). Detection of rifampicin resistance in *Mycobacterium tuberculosis* isolates from diverse countries by a commercial line probe assay as an initial indicator of multidrug resistance. *Int J Tuberc Lung Dis, 4*(5), 481-484.

Trunz, B. B., Fine, P. E. M., & Dye, C. (2006). Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *The Lancet, 367*(9517), 1173-1180. doi:http://dx.doi.org/10.1016/S0140-6736(06)68507-3

Tsolaki, A. G. (2004). Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc Natl Acad Sci U S A., 101*. doi:10.1073/pnas.0305634101

Turnbull, F. W., Wallace, A. T., Stewart, S., & Crofton, J. W. (1953). Streptomycin resistance in patients with pulmonary tuberculosis previously treated with PAS alone. *Br Med J, 1*(4822), 1244-1246.

Török, M. E., Reuter, S., Bryant, J., Köser, C. U., Stinchcombe, S. V., Nazareth, B., . . . Peacock, S. J. (2013). Rapid Whole-Genome Sequencing for Investigation of a Suspected Tuberculosis Outbreak. *Journal of Clinical Microbiology, 51*(2), 611-614. doi:10.1128/jcm.02279-12

van Rijn, S. P., van Altena, R., Akkerman, O. W., van Soolingen, D., van der Laan, T., de Lange, W. C., . . . Alffenaar, J. W. (2016). Pharmacokinetics of ertapenem in patients with multidrug-resistant tuberculosis. *Eur Respir J, 47*(4), 1229-1234. doi:10.1183/13993003.01654-2015

van Soolingen, D., de Haas, P. E., van Doorn, H. R., Kuijper, E., Rinder, H., & Borgdorff, M. W. (2000). Mutations at amino acid position 315 of the *katG* gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of *Mycobacterium tuberculosis* in the Netherlands. *J Infect Dis, 182*(6), 1788-1790. doi:10.1086/317598

van Soolingen, D., Hermans, P. W., de Haas, P. E., Soll, D. R., & van Embden, J. D. (1991). Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol, 29*(11), 2578-2586.

Vareldzis, B. P., Grosset, J., de Kantor, I., Crofton, J., Laszlo, A., Felten, M., . . . Kochi, A. (1994). Drug-resistant tuberculosis: laboratory issues. World Health Organization recommendations. *Tuber Lung Dis, 75*(1), 1-7. doi:10.1016/0962-8479(94)90096-5

Veziris, N., Truffot, C., Mainardi, J. L., & Jarlier, V. (2011). Activity of carbapenems combined with clavulanate against murine tuberculosis. *Antimicrob Agents Chemother, 55*(6), 2597-2600. doi:10.1128/aac.01824-10

Viaene, E., Chanteux, H., Servais, H., Mingeot-Leclercq, M. P., & Tulkens, P. M. (2002). Comparative stability studies of antipseudomonal beta-lactams for potential administration through portable elastomeric pumps (home therapy for cystic fibrosis patients) and motor-operated syringes (intensive care units). *Antimicrob Agents Chemother, 46*(8), 2327-2332.

Viana-Niero, C., Haas, P. E., Soolingen, D., & Leão, S. C. (2004). Analysis of genetic polymorphisms affecting the four phospholipase C (plc) genes in *Mycobacterium tuberculosis* complex clinical isolates. *Microbiology., 150*. doi:10.1099/mic.0.26778-0

Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., & Schoolnik, G. K. (2003). Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J Exp Med, 198*(5), 705-713. doi:10.1084/jem.20030205

Votintseva, A. A., Pankhurst, L. J., Anson, L. W., Morgan, M. R., Gascoyne-Binzi, D., Walker, T. M., . . . Crook, D. W. (2015). Mycobacterial DNA extraction for whole-genome sequencing from early positive liquid (MGIT) cultures. *J Clin Microbiol, 53*(4), 1137-1143. doi:10.1128/jcm.03073-14

Wade, M. M., & Zhang, Y. (2004). Anaerobic incubation conditions enhance pyrazinamide activity against *Mycobacterium tuberculosis*. *J Med Microbiol, 53*(Pt 8), 769-773. doi:10.1099/jmm.0.45639-0

Walker, T. M., Ip, C. L., & Harrell, R. H. (2013). Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis, 13*. doi:10.1016/s1473-3099(12)70277-3

Walker, T. M., Ip, C. L., Harrell, R. H., Evans, J. T., Kapatai, G., & Dedicoat, M. J. (2013). Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis, 13*. doi:10.1016/s1473-3099(12)70277-3

Walker, T. M., Kohl, T. A., Omar, S. V., Hedge, J., Del Ojo Elias, C., Bradley, P., . . . Peto, T. E. A. (2015). Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *The Lancet Infectious Diseases, 15*(10), 1193-1202. doi:10.1016/s1473-3099(15)00062-6

Wallace, R. J., Jr., Swenson, J. M., Silcox, V. A., Good, R. C., Tschen, J. A., & Stone, M. S. (1983). Spectrum of disease due to rapidly growing mycobacteria. *Rev Infect Dis, 5*(4), 657-679.

Wallis, R. S., Maeurer, M., Mwaba, P., Chakaya, J., Rustomjee, R., Migliori, G. B., . . . Zumla, A. (2016). Tuberculosis--advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers. *Lancet Infect Dis, 16*(4), e34-46. doi:10.1016/s1473-3099(16)00070-0

Walz, A. J., Mollmann, U., & Miller, M. J. (2007). Synthesis and studies of catechol-containing mycobactin S and T analogs. *Org Biomol Chem, 5*(10), 1621-1628. doi:10.1039/b703116e

Warren, R., de Kock, M., Engelke, E., Myburgh, R., Gey van Pittius, N., Victor, T., & van Helden, P. (2006). Safe *Mycobacterium tuberculosis* DNA extraction method that does not compromise integrity. *J Clin Microbiol, 44*(1), 254-256. doi:10.1128/jcm.44.1.254-256.2006

Watt, B., Edwards, J. R., Rayner, A., Grindey, A. J., & Harris, G. (1992). In vitro activity of meropenem and imipenem against mycobacteria: development of a daily antibiotic dosing schedule. *Tuber Lung Dis, 73*(3), 134-136. doi:10.1016/0962-8479(92)90145-a

Wayne, L. G., & Hayes, L. G. (1996). An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun, 64*(6), 2062-2069.

Wheeler, P. R., & Ratledge, C. (1992). Control and location of acyl-hydrolysing phospholipase activity in pathogenic mycobacteria. *J Gen Microbiol, 138*(4), 825-830. doi:10.1099/00221287-138-4-825

WHO. (1993). Tuberculosis: a global emergency. *World Health Forum, 14*(4), 438.

WHO. (2009). Treatment of tuberculosis: guidelines. 4th. Retrieved from http://apps.who.int/iris/bitstream/10665/44165/1/9789241547833_eng.pdf

WHO. (2010a, 2010-12-07 20:49:04). Emergence of XDR-TB. Retrieved from http://www.who.int/mediacentre/news/notes/2006/np23/en/

WHO. (2010b, 2016-10-05 15:48:58). WHO, Xpert MTB/RIF: WHO Policy update and Implementation manual. *WHO.* Retrieved from http://www.who.int/tb/laboratory/xpert_launchupdate/en/

WHO. (2015, 2015-09-08 13:13:45). WHO, End TB Strategy. *WHO.* Retrieved from http://www.who.int/tb/post2015_strategy/en/

WHO. (2017, 2017-03-21 11:55:32). WHO, Tuberculosis. *Key Facts.* Retrieved from http://www.who.int/mediacentre/factsheets/fs104/en/

Wipperman, M. F., Sampson, N. S., & Thomas, S. T. (2014). Pathogen roid rage: cholesterol utilization by *Mycobacterium tuberculosis*. *Crit Rev Biochem Mol Biol, 49*(4), 269-293. doi:10.3109/10409238.2014.895700

Wirth, T., Hildebrand, F., Allix-Beguec, C., Wolbeling, F., Kubica, T., Kremer, K., . . . Niemann, S. (2008). Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS Pathog, 4*(9), e1000160. doi:10.1371/journal.ppat.1000160

Witney, A. A., Bateson, A. L., Jindani, A., Phillips, P. P., Coleman, D., Stoker, N. G., . . . McHugh, T. D. (2017). Use of whole-genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. *BMC Med, 15*(1), 71. doi:10.1186/s12916-017-0834-4

Witney, A. A., Gould, K. A., & Arnold, A. (2015). Clinical application of whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases. *J Clin Microbiol, 53*. doi:10.1128/jcm.02993-14

Woods, G. L. (2000). Susceptibility testing for mycobacteria. *Clin Infect Dis, 31*(5), 1209-1215. doi:10.1086/317441

Wright, A., Zignol, M., Van Deun, A., Falzon, D., Gerdes, S. R., Feldman, K., . . . Raviglione, M. (2009). Epidemiology of antituberculosis drug resistance 2002-07: an updated analysis of the Global Project on Anti-Tuberculosis Drug Resistance Surveillance. *Lancet, 373*(9678), 1861-1873. doi:10.1016/s0140-6736(09)60331-7

Xie, Z., Siddiqi, N., & Rubin, E. J. (2005). Differential antibiotic susceptibilities of starved *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother, 49*(11), 4778-4780. doi:10.1128/aac.49.11.4778-4780.2005

Yah, C. S., & Simate, G. S. (2015). Nanoparticles as potential new generation broad spectrum antimicrobial agents. *Daru, 23*, 43. doi:10.1186/s40199-015-0125-6

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol, 24*. doi:10.1093/molbev/msm088

Yang, Z., Yang, D., & Kong, Y. (2005). Clinical relevance of *Mycobacterium tuberculosis plcD* gene mutations. *Am J Respir Crit Care Med, 171*. doi:10.1164/rccm.200408-1147OC

Youmans, G. P., & Youmans, A. S. (1949). A METHOD FOR THE DETERMINATION OF THE RATE OF GROWTH OF TUBERCLE BACILLI BY THE USE OF SMALL INOCULA. *J Bacteriol, 58*(2), 247-255.

Zhanel, G. G., Wiebe, R., Dilay, L., Thomson, K., Rubinstein, E., Hoban, D. J., . . . Karlowsky, J. A. (2007). Comparative review of the carbapenems. *Drugs, 67*(7), 1027-1052.

Zhang, M., Yue, J., Yang, Y., Zhang, H., Lei, J., Jin, R., . . . Wang, H. (2005). Detection of mutations associated with isoniazid resistance in *Mycobacterium tuberculosis* isolates from China. *J Clin Microbiol, 43*(11), 5477-5482. doi:10.1128/jcm.43.11.5477-5482.2005

Zhang, X., Liu, L., Zhang, Y., Dai, G., Huang, H., & Jin, Q. (2015). Genetic determinants involved in p-aminosalicylic acid resistance in clinical isolates from tuberculosis patients in northern China from 2006 to 2012. *Antimicrob Agents Chemother, 59*(2), 1320-1324. doi:10.1128/aac.03695-14

Zhang, Y., Heym, B., Allen, B., Young, D., & Cole, S. (1992). The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature, 358*(6387), 591-593. doi:10.1038/358591a0

Zhang, Y., & Mitchison, D. (2003). The curious characteristics of pyrazinamide: a review. *Int J Tuberc Lung Dis, 7*(1), 6-21.

Zhang, Y., Shareena Dasari, T. P., Deng, H., & Yu, H. (2015). Antimicrobial Activity of Gold Nanoparticles and Ionic Gold. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev, 33*(3), 286-327. doi:10.1080/10590501.2015.1055161

Zhang, Y., & Yew, W. W. (2015). Mechanisms of drug resistance in *Mycobacterium tuberculosis*: update 2015. *Int J Tuberc Lung Dis, 19*(11), 1276-1289. doi:10.5588/ijtld.15.0389

Zhao, F., Wang, X. D., Erber, L. N., Luo, M., Guo, A. Z., Yang, S. S., . . . Deng, J. Y. (2014). Binding pocket alterations in dihydrofolate synthase confer resistance to para-aminosalicylic acid in clinical isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother, 58*(3), 1479-1487. doi:10.1128/aac.01775-13

Zhou, Y., Kong, Y., Kundu, S., Cirillo, J. D., & Liang, H. (2012). Antibacterial activities of gold and silver nanoparticles against *Escherichia coli* and *Bacillus Calmette-Guerin*. *J Nanobiotechnology, 10*, 19. doi:10.1186/1477-3155-10-19

Zimhony, O., Vilcheze, C., Arai, M., Welch, J. T., & Jacobs, W. R., Jr. (2007). Pyrazinoic acid and its n-propyl ester inhibit fatty acid synthase type I in replicating tubercle bacilli. *Antimicrob Agents Chemother, 51*(2), 752-754. doi:10.1128/aac.01369-06

Zumla, A., Abubakar, I., Raviglione, M., Hoelscher, M., Ditiu, L., McHugh, T. D., . . . Atun, R. (2012). Drug-resistant tuberculosis--current dilemmas, unanswered questions, challenges, and priority needs. *J Infect Dis, 205 Suppl 2*, S228-240. doi:10.1093/infdis/jir858

Zumla, A., Nahid, P., & Cole, S. T. (2013). Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Discov, 12*(5), 388-404. doi:10.1038/nrd4001

Appendices

A conference poster, peer-reviewed papers published at the time of printing this thesis and a conference talk are listed below:

- Satta G, Witney AA, Shorten RJ, Karlikowska M, Lipman M, McHugh TD. Genetic variation in *Mycobacterium tuberculosis* isolates from a London outbreak associated with isoniazid resistance. Poster 4857, Federation of Infection Societies/10th Healthcare Infection Society International Conference (Edinburgh, 6th-8th November 2016).
- Satta G, Witney AA, Shorten RJ, Karlikowska M, Lipman M, McHugh TD. Genetic variation in *Mycobacterium tuberculosis* isolates from a London outbreak associated with isoniazid resistance. BMC Med. 2016 Aug 16;14(1):117.
- Satta G, Atzeni A, McHugh TD. *Mycobacterium tuberculosis* and whole genome sequencing: a practical guide and online tools available for the clinical microbiologist. Clin Microbiol Infect. 2017 Feb;23(2):69-72.
- Satta G, Lipman M, Smith GP, Arnold C, Kon OM, McHugh TD. *Mycobacterium tuberculosis* and whole-genome sequencing: how close are we to unleash its full potential? Clin Microbiol Infect. 2018 Jun;24(6):604-609.
- Conference speaker at the Infectious Diseases Diagnostics Conference (London, 21st-22nd Feb 2018).