Deciphering the Mechanisms of Isoniazid Resistance in *Mycobacterium tuberculosis*

By Simoné Nagel

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

Supervisor: Dr. Elizabeth Streicher

Co-supervisor: Dr. Marisa Klopper

December 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature:

Date: December 2018

Abstract

Isoniazid (INH) serves as the backbone of combined anti-tuberculosis therapy. However, the effectivity of this drug has been compromised due to increasing resistance to it in *Mycobacterium tuberculosis* (*M. tuberculosis*). This resistance arises through spontaneous mutations in certain genes/genomic regions of *M. tuberculosis*. The current drug resistance testing algorithm for INH consists of a genotypic MTBDR*plus* line-probe assay (LPA), which reports resistance to INH based on mutations in *katG* gene and *inhA* promoter, and a phenotypic drug susceptibility test (DST), used to confirm INH susceptibility in rifampicin resistant cases on LPA. However, the algorithm is problematic as there are discrepancies between mutations causing INH resistance detected on LPA and phenotypic DST. This may result in incorrect diagnoses and the prescription of incorrect drug treatment regimens. The current study investigates 398 clinical isolates obtained from NHLS Port Elizabeth with discrepant status between the LPA and the phenotypic DST. The isolates were investigated by means of a series of standardized methods such as Sanger sequencing, Whole Genome Sequencing and spoligotyping. This was done to determine if the reasons for the discrepancies between the LPA and phenotypic DST were due to novel mechanisms of resistance to INH. Genotyping via spoligotyping indicated a prevalence of the Beijing genotype in this cohort, as well as mixed infections and novel spoligotype patterns that don't resemble any known lineage or family markers. Sanger sequencing revealed canonical mutations that were missed by the LPA as well as mutations occurring elsewhere in the *katG* gene, indicating there are blind spots in the LPA that may have detrimental effects on patient response to treatment. Per the whole genome sequencing results, novel mechanisms were associated with spontaneous non-canonical single nucleotide polymorphisms in *M. tuberculosis*. These polymorphisms occurred in genetic regions outside the LPA region of hybridisation and conferred varying levels of resistance to INH. The clinical isolates in this dataset consisted of 29.2% low-level resisistant (minimum inhibitory concentration between 0.1-1.0 µg/L to INH) and 61% high-level resistant (minimum inhibitory concentration of > 1µg/ml to INH) isolates.The degree of resistance to INH is relevant, as it informs prescription of treatment regimens. It is recommended that cases of low-level resistance to INH be treated with the short course Bangladesh regimen consisting of high dose INH, whereas high-level resistance to INH requires the removal of this antibiotic from the regimen.Homologous recombineering was also done to functionally confirm the role these novel mechanisms have on INH resistance, and the subsequent level of resistance caused. The results obtained in this study will ultimately be used to improve current diagnostic methods for detection of INH resistance.

Opsomming

Isoniazid (INH) dien as die ruggraat van gekombineerde anti-tuberkulose terapie. Tans is die effektiewiteit van hierdie antibioika in gedrang as gevolg van toenemende weerstand daarteen in *Mycobacterium tuberculosis* (*M. tuberculosis*). Hierdie weerstand ontstaan deur spontane mutasies in sekere gene/genotipiese streke van *M. tuberculosis*. Die huidige middel weerstand toets algoritme vir INH bestaan uit 'n genotipiese MTBDR*plus* (LPA) toets, wat INH weerstandigheid rapporteer gebaseer op mutasies in *katG* gene en *inhA* promotor, en 'n fenotipiese middel vatbaarheid toets (DST), gebruik om INH vatbaarheid te bevestig in rifampisien weerstandige gevalle op LPA. Die algoritme is egter problematies omdat daar teenstrydighede is tussen mutasies wat INH weerstand veroorsaak op LPA en fenotipiese DST. Dit kan lei tot verkeerde diagnoses en die voorskrif van verkeerde behandeling vir tuberkulose. Die huidige studie ondersoek 398 kliniese isolate van NHLS Port Elizabeth met teenstrydige weerstandigheids status tussen die LPA en die fenotipiese DST. Die isolate is ondersoek deur middel van gestandardiseerde metodes soos Sanger volgordebepaling, hele genoom volgordebepaling (WGS) en spoligotipering. Dit was gedoen om die redes vir die teenstrydighede tussen die LPA en fenotipiese DST te bepaal. Spoligotipering het aangedui dat die meeste isolate in die studie aan die Beijing genotipe behoort, sowel as gevalle van gemengde infeksies en onbekende patrone. Sanger volgordebepaling het klassieke mutasies wat deur die LPA gemis was asook mutasies wat elders in die *katG* geen,voorkom-gediagnoseer, wat aandui dat daar blinde kolle is in die LPA. Die blinde kolle mag ‘n nadelige effek hê op die behandeling van die pasient. Volgens die hele genoom volgordebepaling resultate, het die unieke meganismes van weerstand gepaardgegaan met punt mutasies in *M. tuberculosis*. Hierdie mutasies het voorgekom in genetiese streke buite die wat ondersoek word deur die LPA en verleen wisselende vlakke van weerstand teen INH. Die kliniese isolate in hierdie studie bestaan uit 29.2% lae-vlak weerstandige (minimum inhiberende konsentrasie tussen 0.1-1.0 µg/L tot INH) en 61% hoëvlak weerstandige (minimum inhiberende konsentrasie van > 1µg/ml tot INH) isolate. Die vlak van weerstand teen INH is belangrik, aangesien dit die voorskrif van antibiotika vir behandeling beinvloed. Ons beveel aan dat gevalle van lae-vlak weerstand tot INH behandel word met die kort kursus Bangladesh antibiotika kursus wat bestaan uit hoë dosis INH, terwyl hoëvlak weerstand teen INH vereis dat INH uitgesluit word vir behandeling. Homologosiese rekombinasie is gedoen om funksioneel te bevestig dat die rare mutasies ‘n rol het in INH weerstand. Die resultate verkry in hierdie studie sal uiteindelik gebruik word om huidige diagnostiese metodes vir die opsporing van INH weerstand te verbeter.

Acknowledgements

I would like to acknowledge the invaluable support and expertise of my supervisor Dr. Elizabeth Streicher and co-supervisor Dr. Marisa Klopper- without your patience this thesis would not have been possible. Additionally, I would like to thank Prof. Robin Warren, Dr. Margaretha de Vos, Dr. Anzaan Dippenaar and Prof. Paul van Helden for all the questions answered and vital insights. I would also like to thank Christian Otum for his assistance and encouragement.

I am grateful for the financial assistance by the National Research Foundation (NRF) towards my work. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF. I would also like to acknowledge the financial support of the DST-NRF Centre of Excellence for Biomedical TB Research (CBTBR) - the assistance was vital to this work.

To my fantastic parents, thank you for your prayers and endless support. You were always a phone call away, ready with advice and words of encouragement. I appreciate the sacrifices you have made for my education, more than you will ever know.

Dedications

This work is dedicated to Rachelle and Etienne Nagel, for your endless belief in my abilities.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

INH: Isoniazid

TB: Tuberculosis

M. tuberculosis: *Mycobacterium tuberculosis*

DST: Phenotypic drug susceptibility test

WHO: World Health Organization

FASII: type 2 fatty acid synthase system

MDR: multi-drug resistant

LPA: Line Probe Assay

DNA: Deoxyribonucleic acid

RIF: Rifampicin

PCR: Polymerase chain reaction

WGS: Whole genome sequencing

MIC: minimum inhibitory concentration

XDR: extremely-drug resistant -

BMI: Body mass index

NAD: Nicotinamide adenine dinucleotide

INH-NAD Adduct: Isoniazid nicotinamide adenine dinucleotide Adduct

InhA: enoyl-acyl carrier protein reductase

NAT: arylamine *N*-acetyltransferase

M.smegmatis: *Mycobacterium smegmatis*

ACP: acyl carrier protein

KAS: β-Ketoacyl synthase enzyme consisting of genes *kasA* and *kasB*

ORF: open reading frame

NdhII: type II NADH dehydrogenase

BCG: Bacillus Calmette-Guérin

ahpC: alkyl hydroperoxide reductase

MGIT: Mycobacterium Growth Indicator Tube

MODS: microscopic observation of drug susceptibility

CFU: colony forming units

ETH: ethionamide

PAS: para-aminosalicylate, p-aminosalicylate

CYP2E1: cytochrome P450IIE1

AUC: area under the curve

NRF: National Research Foundation

NHLS: National Health Laboratory Service

UV: ultra violet

GC: growth control

OADC: Oleic Albumin Dextrose Catalase

ADC: Albumin Dextrose Catalase

DR: direct repeat

PCR: Polymerase Chain Reaction

SSPE: saline sodium phosphate

EDTA: ethylene diamine tetra-acetic acid

SDS: sodium dodecyl sulphate

Tm: temperature

HRM: High-Resolution Melt

CAF: Central Analytical Facility

$MgCl_2$: magnesium chloride

dNTP: Deoxyribonucleotide triphosphate

MSG: Mono Sodium Glutamic Acid

PCI: Phenol, chloroform, isoamylalcohol

CI: chloroform & isoamylalcohol

Na-Ac: sodium-acetate

CDC: Center for Disease Control

USAP: Universal Sequence Analysis Pipeline

NGS: Next Generation Sequencing

QC: quality control

SNP: single nucleotide polymorphism

BWA: Burrows-Wheeler Aligner

SAM: Sequence Alignment Map

BAM: Binary Alignment Map

GATK: Genome Analysis Toolkit

BLAST: Basic Local Alignment Search Tool

InDels: insertions and deletions

VCF: variant calling files

PROVEAN: Protein Variation Effect Analyzer

PAGE: Polyacrylamide Gel Electrophoresis

E.coli: *Escherichia coli*

LB: Luria Broth

OD: optical density

SITVIT: Spoligo international type, VNTR international type (International Spoligotype database)

ng: nanogram

NTM: non-tuberculous mycobacteria

bp: base pair

kbp: kilo base pair

cm: centimetre

#: number

Chapter 1 - General Introduction

1.1 Background

Aerosolised *Mycobacterium tuberculosis* (*M. tuberculosis*) bacilli are causative of the infectious disease tuberculosis. Its infectious nature has ensured that the TB epidemic is a health concern worldwide. Per statistics supplied by the World Health Organization (WHO), tuberculosis is the third leading cause of mortality due to infectious disease[1]. This disease is particularly dangerous when it presents in persons that are immuno-compromised.

To combat TB, a strict regimen of combined drug therapy is prescribed. This involves taking multiple antibiotics simultaneously to maximize their effects and protect the various drugs. The first-line drugs that are generally prescribed according to the standard regimen recommended by WHO consist of rifampicin, isoniazid, ethambutol and pyrazinamide[2].

The first line drug isoniazid (INH) serves as a central component of combined anti-tuberculosis drug therapy. This drug is bactericidal against actively replicating bacilli and bacteriostatic against latent infections[3]. The mechanism of action of activated INH is to target the type 2 fatty acid synthase (FAS II) system responsible for mycolic acid synthesis within *M. tuberculosis*[4].

Prior to its entry into *M. tuberculosis* cells, INH remains inactive and is not harmful to *M. tuberculosis*. However, upon diffusion into the mycobacterial cells it becomes activated via peroxidation by the KatG catalase peroxidase enzyme (produced by the mycobacterial *katG* gene)[5].

Under ideal conditions *M. tuberculosis* is extremely vulnerable to INH. However, this bacterium is notoriously tenacious and cases of resistance to INH emerged soon after its implementation as an anti-tuberculosis drug in 1952[6]. This resistance to INH arises through spontaneous mutations in the genetic make-up of *M. tuberculosis*, changing its genome and interfering with the mechanism of the anti-tuberculosis drugs. Of these mutations that confer INH resistance, the most well-known are mutations in *katG* and the *inhA* promoter. However, many examples of *M. tuberculosis* strains that display resistance despite lacking mutations in *katG* and the *inhA* promoter are described in the literature[7]–[10]. Some other genes that have been implicated in INH resistance include mutations in *ahpC-oxyR* intergenic region[11], *kasA* gene[12], *inhA* gene[12] and *fabG-inhA* regulatory

region[13] to name but a few. Knowledge of how these mutations contribute to resistance may be used in future to combat the increase in isoniazid mono-resistant and multi-drug resistant (MDR) tuberculosis, and to improve current diagnostic methods.

Traditionally, phenotypic tests have been employed in the South African drug resistance testing algorithm for isoniazid. However, while phenotypic methods such as MIC testing are popular in resource poor settings due to their low cost, they are time consuming because of the slow growth rate *of M. tuberculosis*. Thus, the implementation of molecular methods has become necessary to diagnose drug resistance in *M. tuberculosis* and speed up diagnosis, as well as to ensure a patient is on the correct treatment regimen. These methods function on a genotypic level to detect mutations of the *M. tuberculosis* genome that are associated with drug resistance[14].

As mutations in *katG*, *inhA* promoter, *kasA*[12]and the *oxyR-ahpC* intergenic region[11] are associated with INH resistance, these must be screened for first to obtain a faster diagnosis of INH resistance. The Line Probe Assay (LPA) MTBDR*plus* is a molecular method that can partially achieve this, as it screens for mutations in regions of *katG* and *inhA* promoter which are known to be associated with INH resistance[14]. The LPA tests specifically for mutations in *katG* 315 (S315T1 and S315T2) as these mutations are most commonly associated with resistance to INH (Zhang and Yew, 2009), as well as -15 (C15T), -16 (A16G) and -8 (T8C/ T8A) mutations in the *inhA* promoter[15]. Thus, any mutations that cause INH resistance, but occur outside of these specific regions will be missed by the LPA, thereby misdiagnosing the patient as INH susceptible.

The LPA method involves isolating bacterial deoxyribonucleic acid (DNA) from the *M. tuberculosis* isolate, amplifying the DNA via polymerase chain reaction (PCR) and then observing the hybridization to a probe with a biotin tag. If the probe is bound, a band will be visible. Probes on the GenoType MTBDR*plus*[16] are designed to indicate the presence or absence of *M. tuberculosis*, INH resistance by the markers described above, and rifampicin resistance[15],[17].

As this test has a better positive predictive value for rifampicin (RIF) resistance compared to INH resistance, a secondary phenotypic drug susceptibility test (DST) must be done to confirm INH susceptibility in cases reported as RIF mono-resistant on the LPA. However,

there may be discrepancies between mutations causing INH resistance detected on molecular testing and phenotypic DST[14].

The discrepancies between the two tests for INH resistance may result in incorrect diagnoses and thus the prescription of incorrect drug treatment regimens to patients. This project is based on the hypothesis that the discrepancies between LPA and DST are due to unknown mechanisms of resistance. Thus, we endeavor to verify the accuracy of routine DST (LPA and phenotypic) through Sanger sequencing and diagnose mixed infections or heteroresistance via spoligotyping. Additionally, we identify novel mechanisms of resistance to the first line anti-tuberculosis drug isoniazid through Whole Genome Sequencing (WGS) analysis, followed by validation of selected candidates through recombineering and minimum inhibitory concentration determination. Ultimately, these findings will aid to improve current diagnostic methods.

1.2 Problem Statement

There are discrepancies between the genotypic and phenotypic tests for INH resistance, which may result in incorrect diagnoses and the subsequent prescription of an incorrect drug treatment regimen to patients.

1.3 Hypothesis

The discrepancies between LPA and DST are due to unknown mechanisms of resistance, outside the region of hybridization of the LPA.

1.4 Aims

We aim to identify the reason for discrepancies between the LPA and DST and if no conventional reason can be elucidated, to investigate novel mechanisms of resistance.

1.5 Objectives

1. To identify the reason for discrepancies between genotypic and phenotypic INH resistance testing via MIC, Spoligotyping and Sanger sequencing.
2. To identify unknown mechanisms of resistance to INH via whole genome sequencing.
3. To verify the identified mechanism of resistance via homologous recombineering.

REFERENCES

[1] WHO | Global tuberculosis report 2018. WHO 2018.

[2] WHO. Treatment of tuberculosis: guidelines for national programmes http://www.who.int/tb/features_archive/new_treatment_guidelines_may2010/en/ (accessed Jun 8, 2016).

[3] Zhang, Y.; Yew, W. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* **2009**, *13* (11), 1320–1330 DOI: 10.2741/1289.

[4] Slayden, R. A.; Lee, R. E.; Barry, C. E. Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2000**, *38* (3), 514–525 DOI: 10.1046/j.1365-2958.2000.02145.x.

[5] Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S. The catalase—peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **1992**, *358* (6387), 591–593 DOI: 10.1038/358591a0.

[6] Middlebrook, G. Isoniazid-resistance and catalase activity of tubercle bacilli. *Am. Rev. Tuberc.* **1954**, *69* (3), 471–472.

[7] Lee, A. S. G.; Lim, I. H. K.; Tang, L. L. H.; Telenti, A.; Wong, S. Y. Contribution of *kasA* Analysis to Detection of Isoniazid-Resistant *Mycobacterium tuberculosis* in Singapore. *Antimicrob Agents Chemother.* **1999**, *43* (8), 12–15.

[8] Ramaswamy, S. V; Reich, R.; Dou, S.-J.; Jasperse, L.; Pan, X.; Wanger, A.; Quitugua, T.; Graviss, E. a. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **2003**, *47* (4), 1241–1250 DOI: 10.1128/AAC.47.4.1241.

[9] Viswanathan, G.; Yadav, S.; Raghunand, T. R. Identification of novel loci associated with mycobacterial isoniazid resistance. *Tuberculosis* **2016**, *96*, 21–26 DOI: 10.1016/j.tube.2015.09.008.

[10] Zhang, M.; Yue, J.; Yang, Y.; Zhang, H.; Lei, J.; Jin, R.; Zhang, X.; Wang, H. Detection of Mutations Associated with Isoniazid Resistance in *Mycobacterium tuberculosis* Isolates from China. *J. Clin. Microbiol.* **2005**, *43* (11), 1–7 DOI: 10.1128/JCM.43.11.5477.

[11] Cardoso, R. F.; Cooksey, R. C.; Morlock, G. P.; Barco, P.; Cecon, L.; Forestiero, F.; Leite, C. Q. F.; Sato, D. N.; Shikama, M. D. L.; Mamizuka, E. M.; et al. Screening and characterization of mutations in isoniazid-resistant *Mycobacterium tuberculosis* isolates obtained in Brazil. *Antimicrob. Agents Chemother.* **2004**, *48* (9), 3373–3381

DOI: 10.1128/AAC.48.9.3373-3381.2004.

[12] Hernando Hazbon, M.; Brimacombe, M.; Del Valle, M. B.; Cavatore, M.; Guerrero, M. I.; Varma-Basil, M.; Billman-Jacobe, H.; Lavender, C.; Fyfe, J.; Garcia-Garcia, L.; et al. Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **2006**, *50* (8), 2640–2649 DOI: 10.1128/AAC.00112-06.

[13] Ramaswamy, S. V; Reich, R.; Dou, S.-J.; Jasperse, L.; Pan, X.; Wanger, A.; Quitugua, T.; Graviss, E. a. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **2003**, *47* (4), 1241–1250 DOI: 10.1128/AAC.47.4.1241.

[14] Hughes, J.; Osman, M. Diagnosis and management of drug-resistant tuberculosis in South African adults. *South African Med. J.* **2014**, *104* (12), 894 DOI: 10.7196/samj.9097.

[15] Albert, H.; Bwanga, F.; Mukkada, S.; Nyesiga, B.; Ademun, J. P.; Lukyamuzi, G.; Haile, M.; Hoffner, S.; Joloba, M.; O'Brien, R. Rapid screening of MDR-TB using molecular Line Probe Assay is feasible in Uganda. *BMC Infect. Dis.* **2010**, *10* (41), 1–7 DOI: 10.1186/1471-2334-10-41.

[16] Hain. GenoType MTBDR*plus* | Detection resistance to rifampicin and/or isoniazid of MTBC complex http://www.hain-lifescience.de/en/products/microbiology/mycobacteria/tuberculosis/genotype-mtbdrplus.html (accessed May 2, 2017).

[17] Marttila, H. J.; Soini, H.; Vyshnevskiy, B. I.; Otten, T. F.; Vasilyef, A. V; Huovinen, P.; Viljanen, M. K. Rapid detection of rifampin-resistant *Mycobacterium tuberculosis* by sequencing and line probe assay. *Scand J Infect Dis* **1998**, *30* (2), 129–132.

Chapter 2 - Literature Review (published)

Simoné Nagel, Elizabeth M. Streicher, Marisa Klopper, Robin M. Warren and Paul D. van Helden. Isoniazid Resistance and Dosage as Treatment for Patients with Tuberculosis. *Curr. Drug Metab.* 2017, *18*, 1–11 DOI: 10.2174/1389200218666171031121905.

Isoniazid resistance and dosage as treatment for patients with tuberculosis.

Simoné Nagel[a], Elizabeth M. Streicher*[a], Marisa Klopper[a], Robin M. Warren[a] and Paul D. van Helden[a]

[a]*SA MRC Centre for Tuberculosis Research, DST/NRF Centre of Excellence for Biomedical Tuberculosis research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.*

*Corresponding author: * Elizabeth M. Streicher: DST/NRF Centre of Excellence for Biomedical Tuberculosis research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, P.O box 241, Cape Town, 8000, South Africa. Fax number: + 27 21 938 9863, Tel number: + 27 21 938 9073, Email address: lizma@sun.ac.za*

Abstract: ***Background:*** The first-line TB antibiotic isoniazid (INH) serves as a central component of combined first-line anti-tuberculosis drug therapy. However, resistance to INH has hindered the functioning of this drug. Resistance is caused by several known and unknown mutations in genes/regions in *Mycobacterium tuberculosis* (*M. tuberculosis*), followed by selection of these mutants in the presence of the drug. INH resistance can be categorised as either "high-level" (minimum inhibitory concentration (MIC) of > 1µg/ml to INH) or "low-level" (MIC between 0.1-1.0 µg/L) resistance and is dependent on the specific mutation acquired. The level of resistance is relevant, as INH resistance is often considered to be the first step in development of multi-drug resistant (MDR) and extremely Resistant (XDR) TB. Isoniazid is a pro-drug in which first pass metabolism happens via N-acetyltransferase and is fast, intermediate or slow, depending on the genetics of the host. Thus, low-level INH resistance, particularly in the presence of fast metabolism, could allow additional mutations, development of high-level resistance and progression to multi-drug resistance.

Methods: A structured search of bibliographic databases for peer-reviewed research literature was performed. Set parameters and specific inclusion criteria where used to filter the literature, based on our specified review questions. The quality and relevance of included papers was deduced using standard tools. The relevant content of cited papers was described, and an inferential qualitative content analysis methodology was utilised to analyse the inferences and findings of included studies using a conceptual framework.

Results: Seventy-eight papers were included in the review, of which a sub-set (36) of the papers describe how different genetic mutations result in low or high-level resistance to INH. These papers were also used to set up a diagram detailing how each mutation affects

INH functionality in order to visualise the interactome of INH and *M.tuberculosis* A further twenty-eight out of the seventy-eight papers detail the methods for testing for INH resistance, current treatment regimens and factors that influence treatment outcome in order to better understand the role of INH within the current anti-tuberculosis treatment therapy and how its use can be optimised.

Conclusion: The findings of this review suggest that low-level INH resistance, in the presence of fast-acetylation, is an underrated component of the global TB epidemic worldwide, and may be a significant problem in terms of treatment outcome and progression to antibiotic resistance. Thus, more research must be done to test whether personalised diagnostics and targeted high dose treatment with INH will reduce the incidence of isoniazid mono-resistant and multi-drug resistant (MDR) tuberculosis.

Keywords: Low-level resistance, dosage, tuberculosis, isoniazid, resistance, metabolism and acetylation.

INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (henceforth referred to as *M. tuberculosis*). Its infectious nature, together with its propensity for antibiotic resistance, has ensured that the TB epidemic is a global health concern.

To treat TB, a strict regimen of extended combined drug therapy is prescribed. The first-line drugs that are recommended consist of rifampicin, isoniazid (INH), ethambutol and pyrazinamide[2]. This current drug treatment regimen and dosage was developed after extensive trials, and although it is currently considered best practice, it is not perfect[3],[4]. For example, this multi-drug regimen for TB treatment can lead to hepatotoxicity in some individuals[5]. Anti-TB drug-induced hepatotoxicity ranges from 1% to 36%, and mortality within this group is not uncommon[6],[7]. Additionally, treatment failures can also occur under certain conditions, however it is important to note that the current drug regimen generates successful outcomes for most, provided there is adherence and no antibiotic

resistance[6]. Out of all the drugs in the regimen, INH has been found to be the greatest contributor to drug-induced hepatotoxicity[5],[7], and a better understanding of isoniazid metabolism and treatment modification may provide a solution.

INH specifically, serves as an integral component of combined first-line drug therapy. This first-line drug serves as the backbone of treatment for drug sensitive tuberculosis infections, as it is bactericidal against actively replicating bacilli[8], and is also responsible for much of the early bactericidal activity of the multi-drug regimen[9]. INH is also bacteriostatic against latent infections[8], and is recommended for preventative therapy[10].

For overburdened health care systems to function optimally, cost effective standardised therapies that serve to manufacture and deliver drugs in the most effective way may be desirable. This is especially true for therapies for diseases that have reached epidemic status, such as tuberculosis. Thus, it is for TB treatment, where a standardised dose for INH is provided. However, despite the advantages of standardised therapy, it may not be the most effective treatment in some cases[4]. Therefore, it behoves us to think more carefully about INH usage. While dosage per weight is a factor that we can easily control for, factors involved in metabolism of isoniazid are more complicated and need careful consideration. For example, high concentrations of drug, or certain metabolic products may result in adverse effects, limiting dosage. In turn, a limited dosage either administered in attempt to avoid adverse events or from rapid degradative metabolism may result in relatively low levels of INH, which may predispose the target organism to develop antibiotic resistance. In the case of already existing low-level *M. tuberculosis* isoniazid resistance due to genomic mutations, the host is placed at potential risk for the development of high-level INH resistance or resistance to additional drugs[11].

In the context of the information provided thus far, it is not surprising that INH mono-resistance develops commonly and is widespread. In fact, 4.9% of tuberculosis cases in South Africa were found to be INH mono-resistant per a survey between 2012 and 2014 by the National Institute for Communicable Diseases[12]. In agreement with this finding, the worldwide prevalence of mono-resistance to INH is between 4–13% for all TB cases with a global average of 8.1% for new TB cases[13],[14]. It is also suspected to be the stepping stone for development of further resistance[15], progressing to MDR, XDR and what might be regarded as totally resistant TB cases.

Some important considerations regarding the use of isoniazid are (not in priority order); 1) absorption, 2) host body mass index (BMI), 3) host tolerance, 4) host metabolism, 5) concomitant host health problems, such as HIV infection or diabetes, 6) microbial metabolism and 7) microbial resistance. If we understand the factors involved in INH metabolism for example (particularly fast metabolism), and low-level resistance in *M. tuberculosis*, we might justify modifying the fixed dose regimen. A higher INH dose may pose extra risk for adverse events of course, but the benefit gained may outweigh such risk if patients are monitored carefully. This review proposes to justify this hypothesis.

1. Isoniazid: Mechanism of Action against *M. tuberculosis*

Isoniazid is a pro-drug that, in its activated form, is effective against actively replicating *M. tuberculosis*. Upon diffusion into the mycobacterial cells, INH becomes activated via peroxidation by the KatG catalase peroxidase enzyme (produced by the mycobacterial *katG* gene)[16]. The target of activated INH within the bacillus is the type 2 fatty acid synthase (FAS II) system responsible for mycolic acid synthesis in *M. tuberculosis*[17].

Mycolic acids are essential components of the unique mycobacterial cell wall, along with peptidoglycan and arabinogalactan[18]. Mycolic acids form part of the waxy layer of mycobacteria and also play a role in cording[19]. Importantly, mycolic acids assist in *M. tuberculosis* survival in sub-optimal conditions for extended periods of time[20]. Thus, by preventing the synthesis of mycolic acids, INH is directly affecting the ability of *M. tuberculosis* to survive under most conditions within the host and renders it more susceptible to killing by other antibiotics.

The current model for the INH mechanism of action in the *M. tuberculosis* bacillus suggests that the *katG*-mediated activation of INH results in reactive organic species known as isonicotinic-acyl radical or anion, as well as reactive oxygen species such as superoxide and electrophilic species[8]. The isonicotinic-acyl radical or anion reacts covalently with NAD(H) to form an INH-NAD Adduct. It is suspected that the NAD(H) co-factor binding involved in the Adduct may bring about a conformational change that renders InhA more accessible to INH[21]. It is the modified INH that targets InhA (enoyl-acyl carrier protein reductase), which is thought to be the primary target of INH inhibition as well as being involved in elongation of fatty acids in mycolic acid synthesis[8],[22]. The

INH-NAD Adduct binds to InhA, thus blocking its activity[21]. The elongation of C_{16} and C_{24} fatty acids are prevented within the FAS II (fatty acid elongation system) and thus mycolic acid synthesis is blocked, leading to cell death[23],[24]. Thus, the current understanding is that activated INH targets the active site of enoyl acyl carrier protein (ACP) reductase, named InhA, within the FAS II system[25]. However, an alternative metabolic pathway for INH exists in the bacillus. Specifically, the *nat* gene of *M. tuberculosis* (known as "*tbnat*") has also been postulated to play a role in INH susceptibility in combination with other factors[26]. This gene encodes a microbial arylamine *N*-acetyltransferase (NAT) enzyme. It was first documented in *M. tuberculosis* in 1999[27], and it has been shown to be expressed only in active *M. tuberculosis* cells[28]. Additionally, *tbnat* occurs in a gene cluster that is vital for the survival of *M. tuberculosis* in macrophages[28]. The level of expression may vary per polymorphic forms and exposure to inducing agents such as INH. Payton et al. showed in 1999 that overexpression of the *M. tuberculosis nat* gene in *Mycobacterium smegmatis*, resulted in a threefold increase in resistance[27]. In addition, they showed that *tbnat* can *N*-acetylate INH under in vitro conditions[27].The activation of INH in *M. tuberculosis* via catalase-peroxidase activity cannot occur if INH is *N*-acetylated[26],[29]. There is therefore some evidence to suggest that some variants or expression levels of this enzyme, although rare, may contribute to INH resistance in conjunction with various other factors[26].

2a. *M. tuberculosis* genes associated with isoniazid resistance and the level of resistance

Unfortunately, resistance to INH emerged soon after its introduction as an anti-tuberculosis drug[30]. In *M. tuberculosis*, resistance is not plasmid encoded, but arises through mutations in genes/regions in the genome which may alter the processing of antibiotics and lead to resistance. Of the mutations that confer INH resistance, the most well-known are in the *katG* gene and the *inhA* promoter[31],[32]. Some other genes that have been implicated in INH resistance include mutations in *ahpC-oxyR* intergenic region[33], *kasA* gene[34] and *inhA* gene[34]. However, other genes and gene mutations have been implicated in INH resistance or susceptibility but their mechanisms are not yet well understood[32]. Genes associated with INH resistance have been summarised in **Table 1.**

Table 1: Genes and gene mutations implicated in INH resistance per literature

Locus associated with resistance	Function	Level of resistance
ahpC-oxyR Intergenic Region	Regulation of *katG* expression [41].	Resistant*1 [48]
ahpC	Involved in oxidative stress response [45]	Low-level*2 [48]
oxyR	Compensatory under conditions of peroxide-induced stress [17]	Resistant [48]
furA	Assists in catalase activity and INH susceptibility [49]	Resistant [48]
furA-katG Intergenic region	Regulates transcription of *katG* [49]	Resistant [49]
fabG-inhA Regulatory region	Alternative promoter for *inhA* [38]	Resistant [50]
inhA	Involved in mycolic acid biosynthesis and resistance against isoniazid and ethionamide [24]	Low-level [34]
kasA	Involved in fatty acid biosynthesis (mycolic acids synthesis) [17]	Inconclusive [51]
srmR homolog	Unknown	Inconclusive [44],[52]
Ndh	Transfer of electrons from NADH to the respiratory chain [50]	Resistant [48]
iniB	Unknown	Inconclusive [48],[53]
iniA	Unknown	Inconclusive [53],[54]
iniC	Unknown	Inconclusive [53],[54]
Rv1592c	Unknown	Inconclusive [48]
fadE24	Unknown	Inconclusive [54]
Rv1772	Unknown	Inconclusive [54]
efpA	Translocation of the substrate across the membrane (efflux pump for antibiotic) [35]	Inconclusive [52]
fabD	Malonyl CoA:ACP transacylase [17]	Inconclusive [52]
accD6	Involved in fatty acid biosynthesis (mycolic acids synthesis) [38]	Inconclusive [48]
mabA	Fatty acid biosynthesis pathway [34]	Low-level [36]
katG	Catalase-peroxidase activity that converts INH to its active form [50]	Resistant [37]
nhoA / *nat*	Acetylation of INH (thereby inactivating it) [6]	Inconclusive [48]
embB	Involved in the biosynthesis of the mycobacterial cell wall (synthesis of arabinogalactan) [8]	Resistant [55]
Rv0340	Unknown	Inconclusive [53],[54]
fbpC	Mycolyltransferase activity required for the biogenesis of trehalose dimycolate (cord factor) [35]	Inconclusive [48]

*1 **High-level resistance**: MIC of >1 µg/ml INH
*2 **Low-level resistance**: MIC between 0.1-1.0 µg/ml INH

2b. Genes associated with high-level resistance

i. *katG* gene

Mutations in the *katG* gene are the most common mechanisms associated with INH resistance in *M. tuberculosis*[35]. In addition, *katG* mutations confer medium to high-level resistance (minimum inhibitory concentration of > 1mg/ml) to INH[36],[37]. Taken together, these factors make *katG* mutations the most important mechanism of INH resistance.

Mutations or deletions within the *katG* gene prevent the activation of INH upon its diffusion into *M. tuberculosis*, rendering INH ineffective against *M. tuberculosis* bacilli and resulting in high INH resistance levels[8],[36]. The catalase or peroxidase enzyme encoded for by the *katG* gene is often lost in highly resistant strains, however the catalase ability may be sustained in cases of low-level resistance caused by some *katG* mutations [8],[34],[38].

The majority of INH resistant isolates can be attributed to a *katG* mutation at codon position 315 where serine is substituted for a threonine [8],[31]. This information is important for diagnostic and treatment purposes because it has been shown that initial mono-resistance to INH, specifically *katG* S315T mutations, often precede multi-drug resistance in *M. tuberculosis* [39]. Pym et al., 2002 showed in a mouse model that a resistant *katG* S315T mutant produces active catalase-peroxidase and is virulent, revealing that a loss of bacterial fitness does not result from this frequent mutation[40]. Since *M. tuberculosis* is an aerobic organism, the catalase and peroxidase activity of *katG* is important for the breakdown of radicals that may cause cellular damage[34],[41].

However, a whole genome study done by Torres et al. 2015, observed 14 alternative mutations in *katG* which were implicated in INH resistance[38]. Eight of these polymorphisms were novel (had not previously been associated with resistance). Thus, mutations at *katG*315 are common, but there are other *katG* mutations that may encode INH resistance and complicate resistance diagnoses. However, more work must be done to establish that resistance is truly owing to all the *katG* mutations in these isolates, and are not due to other INH resistance associated mechanisms.

ii. *kasA* gene

The role of *kasA* in mycolic acid synthesis is to elongate long chain fatty acids, which are essential precursor components of mycolic acid chains as discussed previously. Extension of these fatty acids via the FASII system occurs when an acyl carrier protein (ACP), alternatively referred to as InhA, binds to the active site of β-Ketoacyl synthase enzyme known as KAS (consisting of genes *kasA* and *kasB*)[42].

At this stage, we know that some polymorphisms in the *kasA* gene are present in certain *M. tuberculosis* sub-lineages and have no role in INH resistance, whilst others such as G961A, G805A, and G269S are found solely in INH resistant strains and are hypothesised to have a role in INH resistance[43]. However, it should be noted that the Jagielski study was not focused solely on *kasA* mutations and more work must be done on *kasA* before the tenuous link between the mutations and INH resistance can be confirmed or refuted.

2c. Genes associated with low-level resistance

i. *inhA* (*mabA*) promoter

As InhA is the primary target for activated INH (the INH-NAD Adduct), mutations in the gene itself, or its promoter region (*mabA-inhA* operon) may disrupt this interaction and result in isoniazid resistant isolates[24],[34]. Extensive data shows that mutations or polymorphisms in the *inhA* promoter are the second most common explanation for INH resistance. More specifically, investigation of INH resistant isolates showed that a -15 mutation in the *inhA* promoter was on average present in 19.5% of isolates from 49 different countries[44]. Moreover, this mutation was seen in as many as 43% of INH resistant cases in specific geographical locations[44].

It has been shown that *M. tuberculosis* strains with these mutations spread more rapidly than strains with other mutations[45],[46]. Specifically, Müller et al., 2011, showed that these isolates were more likely to acquire additional mutations, leading to XDR bacilli[46]. However, a key difference in INH resistance encoded by this gene in comparison to *katG* mutations, is that resistance is low-level (MIC of less than 1 mg/ml)[36]. This is due to the mechanism of resistance to INH i.e. upregulation of the target causes resistance by titration (the drug concentration is too low to block all the active sites available). Mutations in the *inhA* promoter have also been shown to confer cross-resistance to another pro-drug,

ethionamide (ETH)[47]. ETH is activated by EthA to form an ETH-NAD Adduct that inactivates InhA, similar to activated INH[37]. However, in at least one documented case, the -15 C - T *inhA* promoter mutation was found to be present in isolates with high-level resistance to INH and ETH[37].

Notably, while low-level resistance is conferred when each of these mutations occurs alone, this high-level resistance was present in all strains where simultaneous mutations in the *inhA* promoter region and *inhA* structural gene occurred, implying a synergistic mechanism[37].

Additionally, a study done by Ando et al. in 2014 discovered a non-synonymous mutation in the *mabA* gene at position 609 (glycine to alanine)[36]. This mutation was found in isolates that previously had no identifiable mechanism of resistance. Per that study, the mechanism of INH resistance associated with a *mabA* G609A mutation is to cause up regulation of *inhA* and to change the region adjacent to the mutation into an alternative promoter[36]. Thus this mutation is thought to confer INH resistance by causing overexpression of *inhA* in the same way that the canonical *inhA* promoter mutations do[36].

ii. *inhA* gene coding sequence

Mutations in the *inhA* open reading frame of *M. tuberculosis* are not as common as in *katG* or *inhA* promoter but still occur often enough to be of interest[34]. If these mutations are present it is usually indicative of an *M. tuberculosis* strain with low-level resistance (MIC between 0.1-1.0 µg/L)[34],[55]. In particular, the S94A mutation in *inhA* decreases INH-NAD Adduct affinity for the NADH-binding site of InhA, therefore resulting in increased INH as well as ethionamide resistance[24].

Other mutations in the *inhA* open ORF that have been implicated in INH resistance include mutations at position I21T/V, position A190S, position I194T and at position I258T/V [24],[44].

Due to the many different mutations in this region that have been associated with INH resistance, as well as the prevalence of mutations in this region being associated with low-level INH resistance, it can be assumed that the *inhA* gene coding region is significant to INH action and resistance.

iii. *ndh* gene

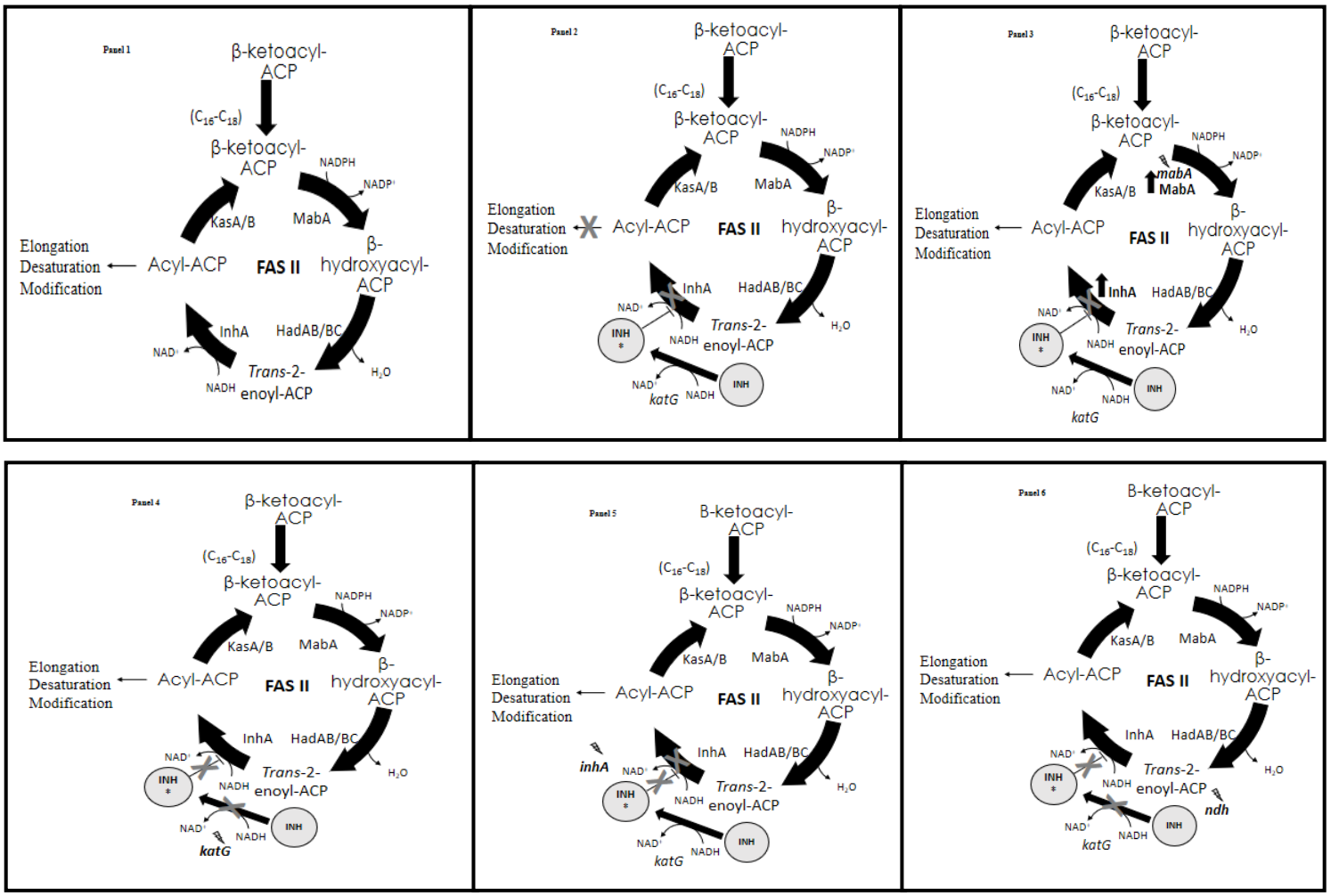
The *ndh* gene encodes for a type II NADH dehydrogenase (NdhII). It is vital for the activation of INH upon its entry into *M. tuberculosis* cells and without it the INH-NAD Adduct cannot be formed [47]. A study by Vilchèze et al., showed for the first time that *ndh* gene mutations conferred a novel resistance pathway, and that these mutations conferred co-resistance to both isoniazid and ethionamide in *Mycobacterium smegmatis* (a fast grower) and *Mycobacterium bovis* BCG (a slow growing strain like *M. tuberculosis*)[47]. It was shown that the mechanism of this co-resistance in *Mycobacterium bovis* BCG was not due to overexpression of *inhA*, thus indicating that its mechanism was novel. All strains with a *ndh* mutation were found to have defective NdhII activity[47]. Due to the defective NdhII activity, the NADH/NAD+ ratios were increased intracellularly. It was hypothesised that this increased NADH/NAD+ ratio competitively inhibited the binding of the INH-NAD Adduct to InhA thereby preventing its activation and effectively blocking the INH mechanism of action[47].

A study done in Singapore in 2001 by Lee et al., discovered *ndh* mutations in INH resistant strains that were absent in all INH susceptible strains. Specifically, 8 out of 84 INH resistant isolates had *ndh* mutations, with 7 of these containing no additional mutations that have been implicated in INH resistance. Only one of the 8 had another mutation in addition to the *ndh* mutation and this was in the *ahpC* gene (which is not thought to confer resistance on its own)[56]. A further study by Viswanathan et al. 2016, confirmed the presence of *ndhII* mutations in INH resistant isolates from the above mentioned Singapore cohort[32]. The mutations in this study, as well as their association with INH resistance were validated by MIC determination and complementation studies with the corresponding *M. smegmatis* and *M. tuberculosis* genes[32]. Thus, this study confirmed that *ndhII* mutations confer INH resistance[32].

To add some perspective and more diverse epidemiological data, a 2006 study by Hernando Hazbon et al. worked with 1011 *M. tuberculosis* clinical isolates from Australia, Colombia, India, Mexico, New York, Spain, and Texas[34]. This study found mutations in the *ndh* gene (or its upstream regions) in 30 of these isolates. The most common *ndh* mutation, V18A, was detected in 28 isolates, but 16 were in INH susceptible strains and 12 in resistant isolates. Alternatively, two resistant isolates had an R268H *ndh* mutation, which was not present in susceptible isolates. Thus, only two *ndh* mutations were present

solely in INH resistant strains. This suggests that not all *ndh* polymorphisms are causative of INH resistance, and that epidemiologically, the role of those *ndh* mutations that do cause INH resistance, is minimal[34].

The interactome of the *M. tuberculosis* cell and INH is represented in **Figure 1**. It includes INH mechanism of action and diagrammatically shows how mutations within certain genes prevent INH from functioning effectively. **Panel 1** shows the normal functioning of the FASII system (i.e. when INH is not present in the system). Alternatively, **Panel 2** diagrammatically presents the functioning of the FASII system in the presence of activated INH and thus displays the INH mechanism of action. Additionally, the INH point of entry into the cell is represented, to better visualise the formation of the INH-NAD Adduct. This modified INH targets and blocks InhA activity, thus blocking elongation of fatty acids within the FASII system of *M. tuberculosis*. The mutations represented **Panel 3** are mutations in *ndh*, *kasA*, *inhA* and *mabA* promoter. In each case **Panel 3** shows the effects of the mutations in these genes, i.e. which step of the FASII system is blocked. Of interest in the context of this review is how mutations in the *mabA* promoter region, *inhA* gene and *ndh* affect resistance. These genes are known to result in low-level INH resistance in *M. tuberculosis*. Low-level INH resistance, particularly in the presence of fast metabolism, could allow additional mutations, development of high-level resistance and progression to multi-drug resistance. A better understanding of the factors involved in INH metabolism is thus vital.



Key for **Figure 1:**

X : Step in pathway blocked, : Mutation, : Upregulation, * : Activated compound.

Figure 1: Interactome of *M. tuberculosis* cell and INH.

Panel 1: Normal functioning of FAS II system
Panel 2: Interruption of FAS II system by activated INH
Panel 3: Effect of mutations:
- Mutation in ***mabA*/promoter**: upregulation of MabA and InhA → resistance by titration

Panel 4: Effect of mutation:
- Mutation in ***katG***: INH not activated

Panel 5: Effect of mutation:
- Mutation in ***inhA*** gene: decreases INH-NAD adduct affinity for the NADH-binding site of InhA → increased INH + ETH resistance

Panel 6: Effect of mutation;
- Mutation in ***ndh***: INH-NAD adduct cannot be formed → INH not activated

2d. Gene associated with compensatory mechanism

i. *ahpC-oxyR* Intergenic Region

The *ahpC-oxyR* intergenic region is thought to be responsible for the regulation of *katG* expression[41]. Both the *ahpC* and *oxyR* genes are upregulated in *M. tuberculosis* under conditions of oxidative stress. The promoter region of *ahpC* tends to acquire mutations when *katG* is absent or mutated and there is no catalase or peroxidase activity[8]. Often in INH resistant strains that have a *katG* mutation, overexpression of the *aphC* gene is induced to compensate for the lack of catalase/peroxidase activity by providing an alternative mechanism for clearing reactive oxygen species. However, overexpression of AphC did not confer significant INH resistance, or at most low-level INH resistance[8].

The alkylhydroperoxide reductase enzyme, encoded by *ahpC,* performs the same function as the *ahpC-oxyR* intergenic region and can compensate for the absence of catalase and peroxidase activity should there be *katG* mutations or a gene deletion [8]. The *oxyR* gene activates *katG* expression when the latter gene is intact, while *oxyR* is upregulated in the absence of functional *katG*, indicating a negative feedback mechanism[41],[57]. Apart from the catalase-peroxidase enzyme encoded for by *katG*, the *oxyR* gene is considered the primary regulator of the peroxide stress responses in *M. tuberculosis.* This gene is activated by exposure to low dosages of hydrogen peroxide and expresses proteins that protect the cell against these oxidative stress conditions[41].

Mutations in the alkyl hydroperoxide reductase (*ahpC*) promoter have been associated with strains of *M. tuberculosis* that are resistant to INH[34],[45]. These mutations are hypothesised to be compensatory as they are believed to assist strains with adjusting to oxidative stress caused by gene mutations. However, mutations in the *ahpC* promoter region are fairly rare and are usually found in conjunction with *katG* mutations or deletions[45].

3. Methods of testing for Isoniazid resistance and current treatment regimens

Traditional methods to test for susceptibility to anti-tuberculosis drugs, such as INH, involve phenotypic tests. Examples of these currently in use include minimum inhibitory concentration (MIC) testing, or microscopic observation of drug susceptibility (MODS)[32].

Measuring the MIC of an antibiotic against *M. tuberculosis* involves culturing *M. tuberculosis* strains on media containing serial dilutions of the antibiotic. The level of resistance is then determined by comparing the colony forming units (CFU's) of a control plate without antibiotic, to the CFU's of the plate with the lowest possible concentration where growth was completely inhibited[58]. However, phenotype testing using the MGIT 960 instrument coupled with TB Exist software, allows for more accurate and considerably faster output times[11]. MIC testing is based on the simple principle that if *M. tuberculosis* is able to grow in the presence of an antibiotic, it must be resistant to the antibiotic at that concentration and all concentrations below[59].

Lakshmi et al., noted in 2014 that *M. tuberculosis* isolates that were low-level resistant to INH were also more likely to be cross-resistant to ETH[60],[61]. The MIC testing method was used in that study to test for cross-resistance and the level of resistance. The authors concluded that this method (viz. MIC) was effective enough to classify the level of resistance of standard cases, but was less effective in diagnosing borderline INH resistant isolates. As such, it was concluded that a modified MIC method would be required to more accurately determine cases of low-level resistance and cross-resistance to ETH[60].

Thus, while phenotypic methods such as MIC testing and MODS are particularly useful in resource poor settings due to their low cost, they are labour intensive and time consuming because of the slow growth rate of *M. tuberculosis*. To reduce time-to-diagnosis, molecular methods have been developed and implemented. These methods function on a genotypic level to pinpoint mutations of the *M. tuberculosis* genome that are associated with drug resistance[32]. There is clear justification for this, since rapid identification of antibiotic resistance in *M. tuberculosis* is vital to ensure an effective and competent therapeutic regimen to limit the spread of multi-drug resistant strains[53].

As outlined earlier, isoniazid resistance occurs via genetic mutations primarily within specific regions of the *M tuberculosis* genome, such as *katG* or the *inhA* operon. According to Seifert et al., 2015, mutations in *katG*315 and *inhA* promoter -15, in addition to ten of the most commonly occurring mutations in the *inhA* promoter and the *ahpC-oxyR* intergenic region explain 84% of global phenotypic isoniazid resistance[44]. In order for molecular–based drug susceptibility testing (DST) methods to achieve 100% sensitivity and specificity, all of the genes that confer INH resistance must be identified, classified, and characterised in terms of severity (level of resistance)[62].

Since mutations in *katG*, *inhA*, *kasA* and the *oxyR-ahpC* intergenic region are most often associated with INH resistance, these are obvious targets for diagnostics. The *katG* 315 codon is a prominent target, since it encodes high- level resistance and mutations often precede multi-drug resistance in *M. tuberculosis.*

The Line Probe Assay (LPA) MTBDR*plus* is a molecular method that can partially achieve this, as it screens for mutations in the *katG* and *inhA* promoter which are known to be associated with INH resistance[32],[63],[64] The LPA method involves isolating bacterial DNA from the *M. tuberculosis* isolate, amplifying the DNA via PCR and then observing the hybridisation to a probe with a biotin-labelled tag. If the probe is bound, a band, part of a "barcode", will be visible. Probes on the MTBDR*plus* are designed to indicate the presence or absence of *M. tuberculosis*, INH resistance and rifampicin (RIF) resistance[65],[66].

The LPA tests specifically for mutations *in katG* 315 (S315T1 and S315T2) as these mutations are most commonly associated with resistance to INH[8], as well as -15 (C15T), -16 (A16G) and -8 (T8C/ T8A) mutations in the *inhA* promoter[66]. Thus, any mutations that cause INH resistance, but occur outside of these specific regions will be missed by the LPA, thereby misclassifying the patient as INH susceptible.

In routine laboratories in South Africa, INH resistance in *M. tuberculosis* isolates is determined by means of Hain GenoType MTBDR*plus.* However, if this method predicted INH susceptibility and the isolate is RIF resistant, a second specimen will be tested by culture-based DST for INH. It is important to note that for INH resistance, the sensitivity of the MTBDR*plus* assay, compared to conventional drug susceptibility testing, was found to be 90%, although the LPA has a better positive predictive value for RIF resistance compared to INH resistance[53],[67]. This disparity is likely due to the restricted number of RIF resistance causing mutations, which are mostly detected by the LPA, compared to the diversity of INH resistance causing mutations.

Furthermore, it has been observed[53], that the LPA is not 100% effective in detecting mutations in the limited number of targets included on the LPA. However, it should be noted that this may be due to the test being incorrectly used[63].

The LPA is also not an effective means of diagnosing resistance when heteroresistance or mixed infections occur. The resultant outcome will be dependent on the proportion of the

mix of susceptible and resistant strains, favouring the more common strain. For example, the patient could be incorrectly diagnosed as having an INH susceptible strain of *M. tuberculosis*, should the resistant strain be a minor component of bacterial burden at time of diagnosis, and the incorrect drug treatment regimen would be prescribed[67].

A comprehensive study by Brossier et al., from 2006 suggests that the current systems in place in routine laboratories enables the detection of approximately 89% of the INH-resistant strains with high-level resistance but only 17% of the strains characterized by low-level INH resistance[53].

4. Host based INH metabolism

Quite apart from any metabolic pathways in *M. tuberculosis*, INH is also metabolised by the human host by arylamine *N*-acetyltransferase (NAT). This enzyme is a polymorphic drug-metabolising enzyme, of which there are two isoenzymes, NAT1 and NAT2 in humans[28]. NAT1 acetylates p-aminosalicylate (PAS), however of interest in the context of this review is arylamine *N*-acetyltransferase 2 (NAT2), which occurs mainly in the liver and gut of humans. In the context of the human host, the effectivity of the INH drug (i.e. treatment outcome) partly depends on the metabolic rate of *N*-acetyltransferase[6]. As stated earlier, once acetylated, INH cannot be activated by the microbial catalase-peroxidase complex to the active form. This highly polymorphic gene and its effects have been extensively studied and phenotypes and their reaction to INH therapy will be discussed more in depth later.

Xenobiotic molecules are metabolised by living organisms via a variety of enzymes. Metabolic rates may be dependent on several factors, such as ethnicity, body weight, HIV status, etc.[6],[9]. However, in humans, arguably the most important metabolic activity pertaining to INH is acetylation, mediated by *N*-acetyltransferase[6]. Once acetylated, INH is rendered ineffective as an antibiotic against *M. tuberculosis*. The genes for NAT2 are highly polymorphic[68], but essentially the alleles encode either a fast or slow-acting form of the enzyme. Thus, depending on homozygous or heterozygous state, an individual can be classified as a fast, intermediate or slow phenotype for INH acetylation, as alleles are co-dominant[68]. Fast-acetylation of INH may result in low serum concentrations of the anti-tuberculosis drug, consequently resulting in a higher risk for treatment failure[6],[43].

Conversely, slow-acetylators are at higher risk for INH-induced hepatotoxicity[6],[43]. This INH hepatotoxicity is due to oxidative stress mediated by hepatic CYP2E1[69]. INH is metabolized by NAT2, largely in the liver of the human host, to acetyl-INH via the primary metabolic pathway of INH, which is then hydrolysed to acetyl hydrazine. Acetyl hydrazine can be oxidized by cytochrome P450IIE1 (CYP2E1) to form *N*-hydroxy acetyl hydrazine, which is known to yield reactive hepatotoxins in slow-acetylators of INH. This includes acetyldiazene (or its products such as the reactive acetyl onium ion), acetyl radicals and ketene, which can bind covalently with hepatic macromolecules and result in liver damage[69].

Thus, a single standard dosage of INH poses a risk of adverse events for slow-acetylators and risk of treatment failure for fast-acetylators.

5. Factors that influence treatment outcome

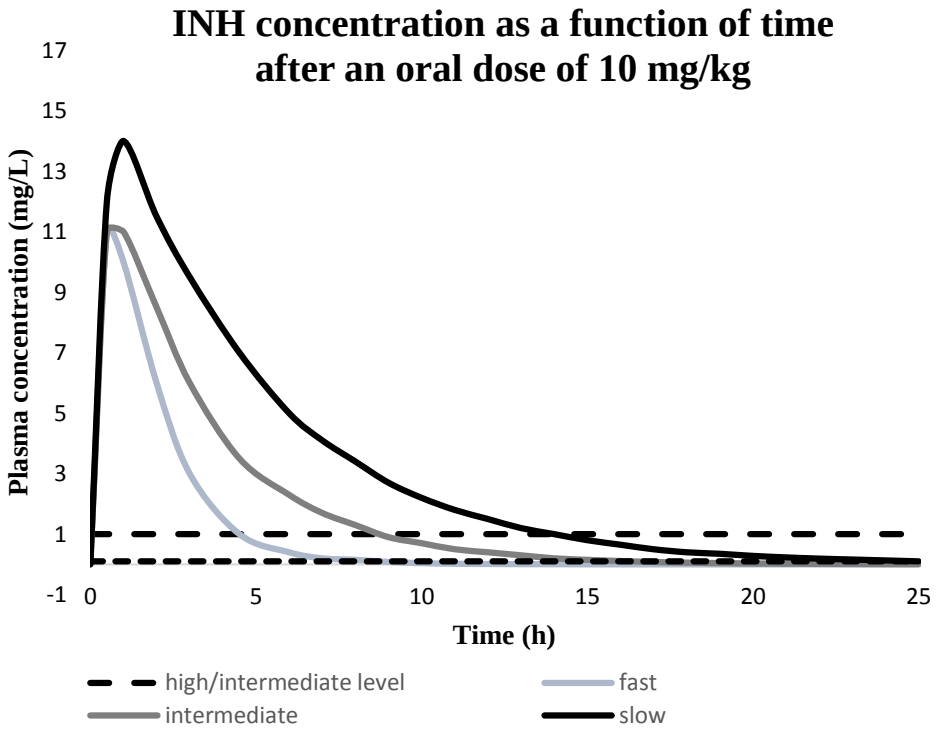
Low-level resistance to INH has been associated with failure of tuberculosis treatment therapy or relapse[67], as antibiotics will only act effectively if concentrations are above the MIC for *M. tuberculosis* and are adequately maintained (adequate area under the curve or "AUC").

For TB treatment, resistance to INH is commonly defined by an MIC of >1ug/L. However, this is a somewhat arbitrary definition, since there is no binary outcome for resistance, such that the bacilli are either fully susceptible to INH or resistant at MIC>1ug/L. In South Africa, for example, it was reported that the MICs of INH for approximately 50% of the INH resistant organisms studied were in the low-level resistance category, where low-level resistance to INH is characterised by an MIC between 0.1-1.0 µg/L[68]. As a result of serum levels of INH using current therapeutic dosage exceeding this level of resistance, particularly shortly after dosing[36], we may initially predict that low-level resistance is irrelevant. However, in the context of fast metabolism of INH, re-evaluation may be required.

With respect to antibiotic therapy, the AUC and minimum inhibitory concentration (MIC) represent the parameters with the best correlation to possible clinical outcomes[69]. However, to assume that all patients will react the same is not in the best interest of patients. INH elimination in humans is trimodal (slow, intermediate and fast-acetylators) and substantially different half-lives of INH are experienced[9],[66]. In the fast-acetylator

(see **Figure 2**), INH concentration rises after administration and then declines to reach the critical concentration of 1ug/L after only 4,5 hours, in contrast to the slow-acetylator where decline to serum concentration of 1ug/L is reached only after 14,5 hours.

From this we may conclude that in the case of patients receiving daily treatment and infected by a low-level INH resistant strain, effective INH concentration is present for only 17% of the time in fast-acetylators, whereas slow-acetylators experience effective dosage



for approximately 60% of the time, or more than 4-fold longer.

Figure 2: Concentration time curve of INH after an oral dose of 10mg/kg. The black line represents a slow-acetylator, dark grey represents an intermediate-acetylator and the light grey represents a fast-acetylator. High/ intermediate resistance (MIC of > 1µg/ml to INH) is represented by everything above the coarse-dashed line and low-level resistance (MIC 0.1-1.0 µg/L) is represented by the space between the dashed lines. Adapted from Parkin et al; 1997[69].

INH dosing based on the NAT2 genotype (slow, intermediate or fast-acetylators) may help to maintain patients within the therapeutic range required for INH to combat *M. tuberculosis* effectively irrespective of whether they have low-level resistant strains of *M.*

tuberculosis. This approach would minimize treatment failures in fast- acetylators, particularly those with low-level resistant strains and reduce adverse events such as drug induced hepatoxicity in slow-acetylators.

Moodley et al., (2014) did a study in mice to assess whether higher doses of isoniazid may be active against drug resistant TB, particularly in cases of low-level resistance to INH[70] . Note that this study considered the effectivity of INH based on the MIC only, and not on the rate of acetylation[70]. Approximately 50% of the INH resistant organisms were low-level resistant in this study. The data suggests that high doses of INH may be a viable treatment option for cases of low-level INH resistance in patients[70].

Although Burhan et al. (2013) reported that most TB patients with low serum level INH and fast-acetylation had good treatment outcomes on a first-line antibiotic multi-drug regimen, this study excluded patients with any resistant isolates and did not report on longer term follow up to assess whether mono-resistance had developed or relapse occurred[71]. Of these patients, 1% failed to be cured and patients with a "poor" response had very low plasma antibiotic levels (in this case "poor response" is defined as at least one positive culture at week 4, 8, or 24/32). Others[72], have shown that in cases where low-level resistance exists, a portion of the bacillary population may survive the initial exposure to INH during treatment. Curiously, Burhan et al (2013) reported that several patients with low serum antibiotic concentration that had a negative culture at week 4 or 8, presented with a positive culture at weeks 8, 24 or 32[71]. This study had no longer term follow up, so it is not evident whether these were forerunners of resistance development or relapse. However, it is important to note, follow-up would not necessarily show acquisition or relapse. More intensive molecular typing would be required for that, and it may even be possible to tell from the existing time-points. It was shown years ago,[73] that slow-acetylators responded well to the conventional doses of INH (namely, 3-5 mg. INH/Kg. /day), but that fast-acetylators would potentially require higher doses of INH (up to 32mg./Kg. of body weight/day).

Van Deun et al. 2010 justify high-dosage INH as a supplement to 9 months of treatment with gatifloxacin, clofazimine, ethambutol, and pyrazinamide[74]. This intensive phase supplement of high-dose INH (in addition to prothionamide and kanamycin) for a minimum of 4 months resulted in a relapse free cure in 87,9% of patients in Bangladesh. Alternatively, a regimen lacking INH was shown to result in much lower effectiveness

(57,1% cure). Thus, the results of this study suggest that high-dose INH plays a vital role in treatment of MDR-TB[74]. In this study, MIC determination was not done and thus treatment was not individualised. We suggest that targeted treatment would yield even better outcomes. A similar conclusion was reached by an independent study[75], i.e. that high dose INH can be effective in the presence of the *inhA* gene mutations.

The failure to take a personalised dosage and diagnosis of low-level resistance into use has at least two negative consequences: firstly, there is the risk that patients will not respond to treatment optimally, particularly during the 4-month extension phase when only 2 antibiotics (viz INH and RIF) are given. Secondly, and perhaps more importantly from a programmatic point of view, there is the risk of development of antibiotic resistance. Our own studies have shown that low-level INH resistance, as defined by mutations in the *inhA* promoter, are more common than mutations in *katG* in certain South African settings[46],[76]. This data suggests that such mutations conferring low-level resistance (and simultaneously including resistance to ETH) are a gateway to the development of further resistance, including high-level INH resistance, encoded by "classical" *katG* mutations that confer high-level resistance to INH[72].

Unfortunately, our knowledge of INH resistant strains and whether tailored INH treatment regimens will provide improved outcomes and reduce generation of antibiotic resistance, is limited. We do not advocate substitution of INH currently, due to its low cost and its general effectiveness. However, we believe we have the ability to make the use of INH more effectively with our current understanding of the mechanisms of action and metabolism of INH[10] . This is particularly important in view of the ease with which *M. tuberculosis* can acquire first low-level, and subsequently high-level resistance to antibiotics. This is thought to be driving the increasing epidemic of MDR and XDR-TB [76].

CONCLUSION

We suggest that low-level INH resistance, in the presence of fast-acetylation, is an underrated component of the global TB epidemic worldwide, and may be a significant problem in terms of treatment outcome and progression to antibiotic resistance. In combination, these two factors can be a generator of mono-antibiotic (isoniazid) resistance, eventually leading to progression to multi-drug resistance [46],[76]. We need to investigate this carefully and test whether personalised diagnostics and targeted high dose treatment with INH will reduce the incidence of isoniazid mono-resistant and multi-drug resistant (MDR) tuberculosis. Since we have shown that serum concentrations of INH can drop well below MIC, particularly in the case of fast-acetylators, we should arguably also rethink the intermittent (5 days a week) therapy widely employed in first-line TB treatment. Although this program saves on costs, it has been associated with relapse and acquired drug resistance [77]. The World Health Organization (WHO) recommended in 2008 that daily regimens be employed as the preferred dosing schedule for TB [77], however intermittent regimens are still an option in several national TB guidelines [78], which is not ideal.

CONFLICT OF INTEREST

The authors of this review declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors acknowledge an affiliation with the South African Medical Research Council. The content is the solely the responsibility of the authors and does not necessarily represent the official views of the South African Medical Research Council.

Additionally, the support of the DST-NRF Centre of Excellence for Biomedical TB research towards this review is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and not necessarily to be attributed to the CBTBR.

Dr E. M. Streicher is supported by the National Research Foundation (NRF) Research Career Advancement Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NRF.

REFERENCES

Journal References

[3] Fox, W.; Ellard, G. A.; Mitchison, D. A. Studies on the treatment of tuberculosis undertaken by the British Medical Research Council tuberculosis units, 1946-1986, with relevant subsequent publications. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **1999**, *3* (10), S231–S279.

[4] Zumla, A.; Nahid, P.; Cole, S. T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nature* **2013**, *12* (May), 388–405 DOI: 10.1038/nrd4001.

[5] Nolan, C. M.; Goldberg, S. V.; Buskin, S. E. Hepatotoxicity Associated With Isoniazid Preventive Therapy. *Jama* **1999**, *281* (11), 1014 DOI: 10.1001/jama.281.11.1014.

[6] Jung, J. A.; Kim, T.; Lee, H.; Jeong, B.; Jeon, K.; Koh, W. A proposal for an individualized pharmacogenetic- guided isoniazid dosage regimen for patients with tuberculosis. *Drug Design, Development and Therapy* **2015**, *2015* (9), 5433–5438.

[7] Du, H.; Chen, X.; Fang, Y.; Yan, O.; Xu, H.; Li, L.; Li, W.; Huang, W. Slow *N*-acetyltransferase 2 genotype contributes to anti-tuberculosis drug-induced hepatotoxicity: a meta-analysis. *Molecular biology reports* **2013**, *40* (5), 3591–3596 DOI: 10.1007/s11033-012-2433-y.

[8] Zhang, Y.; Yew, W. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* **2009**, *13* (11), 1320–1330 DOI: 10.2741/1289.

[9] Gumbo, T.; Louie, A.; Liu, W.; Brown, D.; Ambrose, P. G.; Bhavnani, S. M.; Drusano, G. L. Isoniazid Bactericidal Activity and Resistance Emergence : Integrating Pharmacodynamics and Pharmacogenomics To Predict Efficacy in Different Ethnic Populations. *J Infect Dis.* **2007**, *51* (7), 2329–2336 DOI: 10.1128/AAC.00185-07.

[10] Stagg, H. R.; Lipman, M. C.; Mchugh, T. D.; Jenkins, H. E. Isoniazid-resistant tuberculosis : a cause for concern? *INT J Tuberc Lung Dis* **2017**, *21* (2), 129–139.

[11] Cambau, E.; Viveiros, M.; Machado, D.; Raskine, L.; Ritter, C.; Tortoli, E.; Matthys, V.; Hoffner, S.; Richter, E.; Molino, M. L. P. Del; Cirillo, D. M.; Van Soolingen, D.; Bottger, E. C. Revisiting susceptibility testing in MDR-TB by a standardized quantitative phenotypic assessment in a European multicentre study. *Journal of Antimicrobial Chemotherapy* **2015**, *70* (November 2014), 686–696 DOI: 10.1093/jac/dku438.

[12] National Institute for Communicable Diseases. South African Tuberculosis Drug Resistance Survey 2012–14. **2016**.

[13] Villegas, L.; Otero, L.; Sterling, T. R.; Huaman, M. A. Prevalence , Risk Factors , and Treatment Outcomes of Isoniazid- and Rifampicin- Mono-Resistant Pulmonary Tuberculosis in Lima, Peru. *PLoS ONE* **2016**, *85*, 1–11 DOI: 10.1371/journal.pone.0152933.

[14] Chien, J.; Chen, Y.; Wu, S.; Lee, J.; Wang, J.; Yu, C. Treatment outcome of patients with isoniazid mono-resistant tuberculosis. *Clinical Microbiology and Infection* **2015**, *21*, 59–68 DOI: 10.1016/j.cmi.2014.08.008.

[15] Dye, C.; Espinal, M. A. Will tuberculosis become resistant to all antibiotics? *The Royal Society* **2001**, *268* (1462), 45–52 DOI: 10.1098/rspb.2000.1328.

[16] Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S. The catalase—peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **1992**, *358* (6387), 591–593 DOI: 10.1038/358591a0.

[17] Slayden, R. A.; Lee, R. E.; Barry, C. E. Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*. *Molecular Microbiology* **2000**, *38* (3), 514–525 DOI: 10.1046/j.1365-2958.2000.02145.x.

[18] Augusto Basso, L.; Hildebrando Pereira da Silva, L.; Germano Fett-Neto, A.; Filgueira de Azevedo Junior, W.; de Souza Moreira, Í.; Sérgio Palma, M.; Batista Calixto, J.; Astolfi Filho, S.; Ribeiro dos Santos, R.; Botelho Pereira Soares, M.; Santiago Santos, D.; de Biociências, F.; Claro, R.; Gonçalo Moniz-Fiocruz, F. The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases. *Mem Inst Oswaldo Cruz, Rio de Janeiro* **2005**, *100* (1006), 475–506.

[19] Sandy, J.; Holton, S.; Fullam, E.; Sim, E.; Noble, M. Binding of the anti-tubercular drug isoniazid to the arylamine *N*-acetyltransferase protein from *Mycobacterium smegmatis*. *Protein Science* **2005**, *14* (3), 775–782 DOI: 10.1110/ps.041163505.

[20] Barry, C. E.; Lee, R. E.; Mdluli, K.; Sampson, A. E.; Schroeder, B. G.; Slayden, R. A.; Yuan, Y.; Anderson, R. J. Mycolic acids: structure, biosynthesis and physiological functions. *Prog. Lipid Res.* **1998**, *37* (2), 143–179.

[21] Johnsson, K.; King, D. S.; Schultz, P. G. Studies on the Mechanism of Action of Isoniazid and Ethionamide in the Chemotherapy of Tuberculosis. *J. Am. Chem. SOC. Lancet Nature J . Am. Chem. SOC. C.; Neal, R. N. Mol. Pharmacol* **1995**, *117* (78), 5009–5010.

[22] Rozwarski, D.; Grant, G.; Barton, D.; Jacobs, W. J. .; Sacchettini, J. Tuberculosis drug target. *Science* **1998**, *279* (5347), 98–102.

[23] Takayama, K.; Schnoes, H. K.; Armstrong, E. L.; Boyle, R. W. Site of inhibitory action of isoniazid in the synthesis of mycolic acids in *Mycobacterium tuberculosis*. *Journal of lipid research* **1975**, *16*, 308–317.

[24] Vilchèze, C.; Jacobs Jr, W. R. Resistance to Isoniazid and Ethionamide in *Mycobacterium tuberculosis*: Genes, Mutations, and Causalities. *Micrlogy Spectrum* **2014**, *2* (4), 1–21 DOI: 10.1128/microbiolspec.MGM2-0014-2013.

[25] Slayden, R. A.; Barry, C. E. The genetics and biochemistry of isoniazid resistance in Mycobacterium tuberculosis. *Microbes and Infection* **2000**, *2*, 659–669.

[26] Sholto-douglas-vernon, C.; Sandy, J.; Victor, T. C.; Sim, E.; Helden, P. D. Van. Mutational and expression analysis of tbnat and its response to isoniazid. *Journal of Medical Microbiology* **2005**, *54*, 1189–1197 DOI: 10.1099/jmm.0.46153-0.

[27] Payton, M.; Auty, R. O. Y.; Delgoda, R.; Everett, M. Cloning and Characterization of Arylamine *N* -Acetyltransferase Genes from *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* : Increased Expression Results in Isoniazid Resistance. *Journal of Bacteriology* **1999**, *181* (4), 1343–1347.

[28] Sim, E.; Abuhammad, A.; Ryan, A. Arylamine *N*-acetyltransferases : from drug metabolism and pharmacogenetics to drug. *British Journal of Pharmacology* **2014**, *171*, 2705–2725 DOI: 10.1111/bph.12598.

[29] Upton, A. M.; Mushtaq, A.; Victor, T. C.; Sampson, S. L.; Sandy, J.; Smith, D.; Helden, P. V. Van; Sim, E. Arylamine *N*- acetyltransferase of *Mycobacterium tuberculosis* is a polymorphic enzyme and a site of isoniazid metabolism. *Molecular Microbiology* **2001**, *42* (2), 309–317 DOI: 10.1046/j.1365-2958.2001.02648.x.

[30] Middlebrook, G. Isoniazid-resistance and catalase activity of tubercle bacilli. *American review of tuberculosis* **1954**, *69* (3), 471–472.

[31] Dobner, P., Rush-Gerdes, S., Bretzel, B.; Feldmann, K., Rifai, M. Loscher, T., Rinder, H. Usefulness of *Mycobacterium tuberculosis* genomic mutations in the *katG* and *inhA* for the prediction of isoniazid resistance. *INT J Tuberc Lung Dis* **1997**, *1* (4), 365–369.

[32] Viswanathan, G.; Yadav, S.; Raghunand, T. R. Identification of novel loci associated with mycobacterial isoniazid resistance. *Tuberculosis* **2016**, *96*, 21–26 DOI: 10.1016/j.tube.2015.09.008.

[33] Cardoso, R. F.; Cooksey, R. C.; Morlock, G. P.; Barco, P.; Cecon, L.; Forestiero, F.; Leite, C. Q. F.; Sato, D. N.; Shikama, M. D. L.; Mamizuka, E. M.; Hirata, R. D. C.; De Mello, F.F.; Hirata, M. H. Screening and characterization of mutations in isoniazid-resistant *Mycobacterium tuberculosis* isolates obtained in Brazil. *Antimicrobial Agents and Chemotherapy* **2004**, *48* (9), 3373–3381 DOI: 10.1128/AAC.48.9.3373-3381.2004.

[34] Hernando Hazbon, M.; Brimacombe, M.; Del Valle, M. B.; Cavatore, M.; Guerrero, M. I.; Varma-Basil, M.; Billman-Jacobe, H.; Lavender, C.; Fyfe, J.; Garcia-Garcia, L.; Leon, C.I.; Bose, M.; Chaves, F.; Murray, M.; Eisenach, K.D.; Sifuentes-Osornio, J.; Cave, M. D.; De Leon, A. P.; Alland, D. Population genetics study of isoniazid resistance mutations and evolution of multi-drug-resistant *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy* **2006**, *50* (8), 2640–2649 DOI: 10.1128/AAC.00112-06.

[35] Jeeves, R. E.; Marriott, A. A. N.; Pullan, S. T.; Hatch, K. A.; Allnutt, J. C.; Freire-Martin, I.; Hendon-Dunn, C. L.; Watson, R.; Witney, A. A.; Tyler, R. H.; Arnold, C.; Marsh, P. D.; McHugh, T. D.; Bacon, J. *Mycobacterium tuberculosis* is resistant to isoniazid at a slow growth rate by single nucleotide polymorphisms in *katG* codon ser315. *PLoS ONE* **2015**, *10* (9), 1–21 DOI: 10.1371/journal.pone.0138253.

[36] Ando, H.; Miyoshi-Akiyama, T.; Watanabe, S.; Kirikae, T. A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*. *Molecular Microbiology* **2014**, *91* (3), 538–547 DOI: 10.1111/mmi.12476.

[37] Machado, D.; Perdigão, J.; Ramos, J.; Couto, I.; Portugal, I.; Ritter, C.; Boettger, E. C.; Viveiros, M. High-level resistance to isoniazid and ethionamide in multi-drug-resistant *Mycobacterium tuberculosis* of the Lisboa family is associated with *inhA* double mutations. *Journal of Antimicrobial Chemotherapy* **2013**, *68* (8), 1728–1732 DOI: 10.1093/jac/dkt090.

[38] Torres, J. N.; Paul, L. V; Rodwell, T. C.; Victor, T. C.; Amallraja, A. M.; Elghraoui, A.; Goodmanson, A. P.; Ramirez-Busby, S. M.; Chawla, A.; Zadorozhny, V.; Streicher, E. M.; Sirgel, F. A.; Catanzaro, D.; Rodrigues, C.; Gler, M. T.; Crudu, V.; Catanzaro, A.; Valafar, F. Novel *katG* mutations causing isoniazid resistance in clinical *M. tuberculosis* isolates. *Emerging Microbes & Infections* **2015**, *4*, 1–8 DOI: 10.1038/emi.2015.42

[39] Manson, A. L.; Cohen, K. A.; Abeel, T.; Desjardins, C. A.; Armstrong, D. T.; Iii, C. E. B.; Brand, J.; Global, T.; Consortium, G.; Chapman, S. B.; Cho, S.; Gabrielian, A.; Gomez, J.; Jodals, A. M.; Joloba, M.; Jureen, P.; Lee, J. S.; Malinga, L.; Maiga, M.; Nordenberg, D.; Noroc, E.; Romancenco, E.; Salazar, A.; Ssengooba, W.; Velayati, A. A.; Winglee, K.; Zalutskaya, A.; Via, L. E.; Cassell, G. H.; Dorman, S. E.; Ellner, J.; Farnia, P.; Galagan, J. E.; Rosenthal, A.; Crudu, V.; Homorodean, D.; Hsueh, P.; Narayanan, S.; Pym, A. S.; Skrahina, A.; Swaminathan, S.; Van Der Walt, M.; Alland, D.; Bishai, W. R.; Cohen, T.; Hoffner, S.; Birren, B. W.; Earl, A. M.Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multi-drug resistance. *Nature Publishing Group* **2017**, *49* (3), 395–402 DOI: 10.1038/ng.3767.

[40] Pym, A. S.; Saint-Joanis, B.; Cole, S. T. Effect of *katG* mutations on the virulence of *Mycobacterium tuberculosis* and the implication for transmission in humans.

Infection and immunity **2002**, *70* (9), 4955–4960 DOI: 10.1128/IAI.70.9.4955-4960.2002.

[41] Loots, D. T. An Altered *Mycobacterium tuberculosis* Metabolome Induced by *katG* Mutations Resulting in Isoniazid Resistance. *Antimicrobial Agents and Chemotherapy* **2014**, *58* (4), 2144–2149 DOI: 10.1128/AAC.02344-13.

[42] Slayden, R. a; Barry, C. E. The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *Mycobacterium tuberculosis*. *Tuberculosis (Edinburgh, Scotland)* **2002**, *82* (4-5), 149–160 DOI: 10.1054/tube.2002.0333.

[43] Jagielski, T.; Bakula, Z.; Roeske, K.; Kaminski, M.; Napiorkowsa, A.; Augustynowicz-Kopec, E.; Zwolska, Z.; Bielecki, J. Mutation profiling for detection of isoniazid resistance in *Mycobacterium tuberculosis* clinical isolates. *Journal of Antimicrobial Chemotherapy* **2015**, No. 70, 3214–3221.

[44] Seifert, M.; Catanzaro, D.; Catanzaro, A.; Rodwell, T. C. Genetic mutations associated with isoniazid resistance in *Mycobacterium tuberculosis* : A systematic review. *PLoS ONE* **2015**, *10* (3), 1–13 DOI: 10.1371/journal.pone.0119628.

[45] Gagneux, S.; Burgos, M. V.; DeRiemer, K.; Enciso, A.; Munoz, S.; Hopewell, P. C.; Small, P. M.; Pym, A. S. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathogens* **2006**, *2* (6), 603–610 DOI: 10.1371/journal.ppat.0020061.

[46] Müller, B.; Streicher, E. M.; Hoek, K. G. P.; Tait, M.; Trollip, A.; Bosman, M. E.; Coetzee, G. J.; Hoosain, E.; Pittius, N. C. G. Van; Victor, T. C.; Helden, P. D. Van.; Warren, R. M. *inhA* promoter mutations : a gateway to extensively drug-resistant tuberculosis in South Africa ? *Int J Tuberc Lung Dis* **2011**, *15* (March 2010), 344–351.

[47] Vilchèze, C.; Weisbrod, T. R.; Chen, B.; Kremer, L.; Hazbo, M. H.; Wang, F.; Alland, D.; Sacchettini, J. C.; Jacobs, W. R. Altered NADH/NAD+ Ratio Mediates Coresistance to Isoniazid and Ethionamide in Mycobacteria. *Antimicrob Agents Chemother* **2005**, *49* (2), 708–720 DOI: 10.1128/AAC.49.2.708.

[48] Nusrath, A.; Subbian, S.; Elizabeth, L.; Selvakumar, N. Infection , Genetics and Evolution Overview on mechanisms of isoniazid action and resistance in *Mycobacterium tuberculosis*. *MEEGID* **2016**, *45* (1), 474–492 DOI: 10.1016/j.meegid.2016.09.004.

[49] Kit, G.; Siu, H.; Yam, W. C.; Zhang, Y.; Kao, R. Y. T. An Upstream Truncation of the *furA-katG* Operon Confers High- Level Isoniazid Resistance in a *Mycobacterium tuberculosis* Clinical Isolate with No Known Resistance-Associated Mutations. *Antimicrobial Agents and Chemotherapy* **2014**, *58* (10), 6094–6100 DOI: 10.1128/AAC.03277-14.

[50] Ando, H.; Kitao, T.; Miyoshi-akiyama, T.; Kato, S.; Mori, T.; Kirikae, T. Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Molecular Microbiology* **2011**, *79* (February), 1615–1628 DOI: 10.1111/j.1365-2958.2011.07547.x.

[51] Sun, Y. J.; Lee, A. S. G.; Wong, S. Y.; Paton, N. I. Analysis of the role of *Mycobacterium tuberculosis kasA* gene mutations in isoniazid resistance. *Clinical Microbiology and Infection* **2007**, *13* (8), 833–835 DOI: 10.1111/j.1469-0691.2007.01752.x.

[52] Ramaswamy, S. V; Reich, R.; Dou, S.-J.; Jasperse, L.; Pan, X.; Wanger, A.; Quitugua, T.; Graviss, E. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **2003**, *47* (4), 1241–1250 DOI: 10.1128/AAC.47.4.1241.

[53] Brossier, F.; Veziris, N.; Truffot-pernot, C.; Jarlier, V.; Sougakoff, W. Performance of the Genotype MTBDR Line Probe Assay for Detection of Resistance to Rifampin and Isoniazid in Strains of *Mycobacterium tuberculosis* with Low- and High-Level Resistance. *Journal of Clinical Microbiology* **2006**, *44* (10), 3659–3664 DOI: 10.1128/JCM.01054-06.

[54] Nusrath, A.; Subbian, S.; Elizabeth, L.; Selvakumar, N. Infection , Genetics and Evolution Overview on mechanisms of isoniazid action and resistance in *Mycobacterium tuberculosis*. *Infection, Genetics and Evolution* **2016**, *45* (1), 474–492 DOI: 10.1016/j.meegid.2016.09.004.

[55] Phelan, J.; Coll, F.; Mcnerney, R.; Ascher, D. B.; Pires, D. E. V; Furnham, N.; Coeck, N.; Hill-Cawthorne, G. A.; Nair, M. B.; Mallard, K.; Ramsay, A.; Campino, S.; Hibberd, M. L.; Pain, A.; Rigouts, L.; Clark, T. G. *Mycobacterium tuberculosis* whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance. *BMC Medicine* **2016**, *14* (31), 1–13 DOI: 10.1186/s12916-016-0575-9.

[56] Lee, A. S. G.; Teo, A. S. M.; Wong, S. Novel Mutations in *ndh* in *Mycobacterium tuberculosis* Isolates. *Antimicrobial agents and chemotherapy* **2001**, *45* (7), 2157–2159 DOI: 10.1128/AAC.45.7.2157.

[57] Farr, S. B.; Kogoma, T. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiological Reviews* **1991**, *55* (4), 561–585.

[58] Inderlied, C.; Salfinger, M. Antimicrobial agents and susceptibility tests: mycobacteria. *Manual of clinical microbiology* **1995**, *6*, 1385–1404.

[59] Sirgel, F. A.; Wiid, I. J. F.; Helden, P. D. Van. Measuring Minimum Inhibitory Concentrations in Mycobacteria. *Mycobacteria Protocols* **2009**; pp 173–186.

[60] Lakshmi, R.; Ramachandran, R.; Sundar, A. S. Stratifying low level Isoniazid resistance using additional intermediate drug concentration. *International Journal of Mycobacteriology* **2014**, *3* (2), 152–154 DOI: 10.1016/j.ijmyco.2014.02.001.

[61] Morlock, G. P.; Metchock, B.; Sikes, D.; Crawford, J. T.; Cooksey, R. C. *ethA*, *inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. *Antimicrobial agents and chemotherapy* **2003**, *47* (12), 3799–3805 DOI: 10.1128/aac.47.12.3799-3805.2003.

[62] Kiepiela, P.; Bishop, K. S.; Smith, A. N.; Roux, L.; York, D. F. Genomic mutations in the *katG*, *inhA* and *aphC* genes are useful for the prediction of isoniazid resistance in *Mycobacterium tuberculosis* isolates from Kwazulu Natal, South Africa. *Tubercle and lung disease* **2000**, *80* (1), 47–56 DOI: 10.1054/tuld.1999.0231.

[64] Johnson, R.; Streicher, E. .; Louw, G. .; Warren, R. .; Van Helden, P. D.; Victor, T. C. Drug resistance in *Mycobacterium tuberculosis*. *Curr. Issues Mol. Biol* **2006**, *8*, 97–112.

[65] Marttila, H. J.; Soini, H.; Vyshnevskiy, B. I.; Otten, T. F.; Vasilyef, A. V; Huovinen, P.; Viljanen, M. K. Rapid detection of rifampin-resistant *Mycobacterium tuberculosis* by sequencing and line probe assay. *Scand J Infect Dis* **1998**, *30* (2), 129–132.

[66] Albert, H.; Bwanga, F.; Mukkada, S.; Nyesiga, B.; Ademun, J. P.; Lukyamuzi, G.; Haile, M.; Hoffner, S.; Joloba, M.; O'Brien, R. Rapid screening of MDR-TB using molecular Line Probe Assay is feasible in Uganda. *BMC infectious diseases* **2010**, *10* (41), 1–7 DOI: 10.1186/1471-2334-10-41.

[67] Hughes, J.; Osman, M. Diagnosis and management of drug-resistant tuberculosis in South African adults. *South African Medical Journal* **2014**, *104* (12), 894 DOI: 10.7196/samj.9097.

[68] Parkin, D. P.; Vandenplas, S.; Botha, F. J. H.; Vandenplas, M. L.; Seifart, H.; Helden, P. D. Van; Walt, B. J. Van Der; Donald, P. R.; Jaarsveld, P. P. Van. Trimodality of Isoniazid Elimination Phenotype and Genotype in Patients with Tuberculosis. *Am J Respir Crit Care Med* **1997**, *155*, 1717–1722.

[69] Lian, Y.; Zhao, J.; Wang, Y.; Zhao, J.; Peng, S. Metallothionein protects against isoniazid-induced liver injury through the inhibition of CYP2E1-dependent oxidative and nitrosative impairment in mice. *Food and Chemical Toxicology* **2017**, *102*, 32–38 DOI: 10.1016/j.fct.2017.01.016.

[70] Moodley, C.; Ammerman, N.; Ncgobo, B.; Swanson, R.; Dorasamy, A.; Moodley, S.; Mgaga, Z.; Pillay, L.; Bester, L.; Singh, S.; Grosset, J.; Rodrigues, C.; Almeida, D. Are high doses of isoniazid beneficial in treatment of infection with isoniazid-resistant *M. tuberculosis*? *Kwazulu-Natal Research institute for tuberculosis and HIV* **2014**, 1–5.

[71] Burhan, E.; Ruesen, C.; Ruslami, R.; Ginanjar, A.; Mangunnegoro, H.; Ascobat, P.; Donders, R.; Crevel, R. Van; Aarnoutse, R. Isoniazid, Rifampin, and Pyrazinamide Plasma Concentrations in Relation to Treatment Response in Indonesian Pulmonary Tuberculosis Patients. *Antimicrob Agents And Chemotherapy* **2013**, *57* (8), 3614–3619 DOI: 10.1128/AAC.02468-12.

[72] Chen, P.; Bishai, W. R. Novel Selection for Isoniazid (INH) Resistance Genes Supports a Role for NAD 2 -Binding Proteins in Mycobacterial INH Resistance. *Infection and immunity* **1998**, *66* (11), 5099–5106.

[73] Mitchell, R. S.; Bell, J. C. Clinical implications of isoniazid, pas and streptomycin blood levels in pulmonary tuberculosis. *Trans Am Clin Climatol Assoc.* **1958**, *69*, 98–105.

[74] Van Deun, A.; Maug, A. K. J.; Salim, M. A. H.; Das, P. K.; Sarker, M. R.; Daru, P.; Rieder, H. L. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *American Journal of Respiratory and Critical Care Medicine* **2010**, *182* (5), 684–692 DOI: 10.1164/rccm.201001-0077OC.

[75] Sotgiu, G.; Tiberi, S.; Centis, R.; D'Ambrosio, L.; Fuentes, Z.; Zumla, A.; Migliori, G. B. Applicability of the shorter "Bangladesh regimen" in high multidrug-resistant tuberculosis settings. *International Journal of Infectious Diseases* **2017**, *56*, 190–193 DOI: 10.1016/j.ijid.2016.10.021.

[76] Warren, R. M.; Streicher, E. M.; Gey van Pittius, N. C.; Marais, B. J.; van der Spuy, G. D.; Victor, T. C.; Sirgel, F.; Donald, P. R.; van Helden, P. D. The clinical relevance of Mycobacterial pharmacogenetics. *Tuberculosis*. 2009, pp 199–202.

URL (Web Page):

[1] WHO. Tuberculosis: http://www.who.int/mediacentre/factsheets/fs104/en/ (accessed Apr 12, 2016).

[2] WHO. Treatment of tuberculosis: guidelines for national programmes http://www.who.int/tb/features_archive/new_treatment_guidelines_may2010/en/ (accessed Jun 8, 2016).

[63] WHO. Molecular Line Probe Assays for Rapid Screening of Patients at Risk of Miltidrug-resistant tuberculosis (MDR-TB) http://www.who.int/tb/laboratory/line_probe_assays/en/.

[77] WHO. Treatment of tuberculosis: guidelines. Geneva, Switzerland http://www.who.int/tb/publications/2010/9789241547833/en/.

[78] NICE. National Institute for Health and Care Excellence (NICE) Tuberculosis Guidelines. *United Kingdom* **2016**, No. January.

Chapter 3: Materials and Methods

3.1 Introduction and Summary

In this study we investigated the reasons for discrepant isoniazid (INH) resistance results in isolates that were sent to the National Health Laboratory Services (NHLS) Port Elizabeth for routine diagnosis. Standardized methods were used to elucidate the mechanisms of INH resistance for samples from NHLS with discrepant status. These procedures include Minimum Inhibitory Concentration (MIC), Sanger sequencing, Whole Genome Sequencing (WGS), spoligotyping and homologous recombineering. The experimental procedure is outlined below in **Figure 3**.

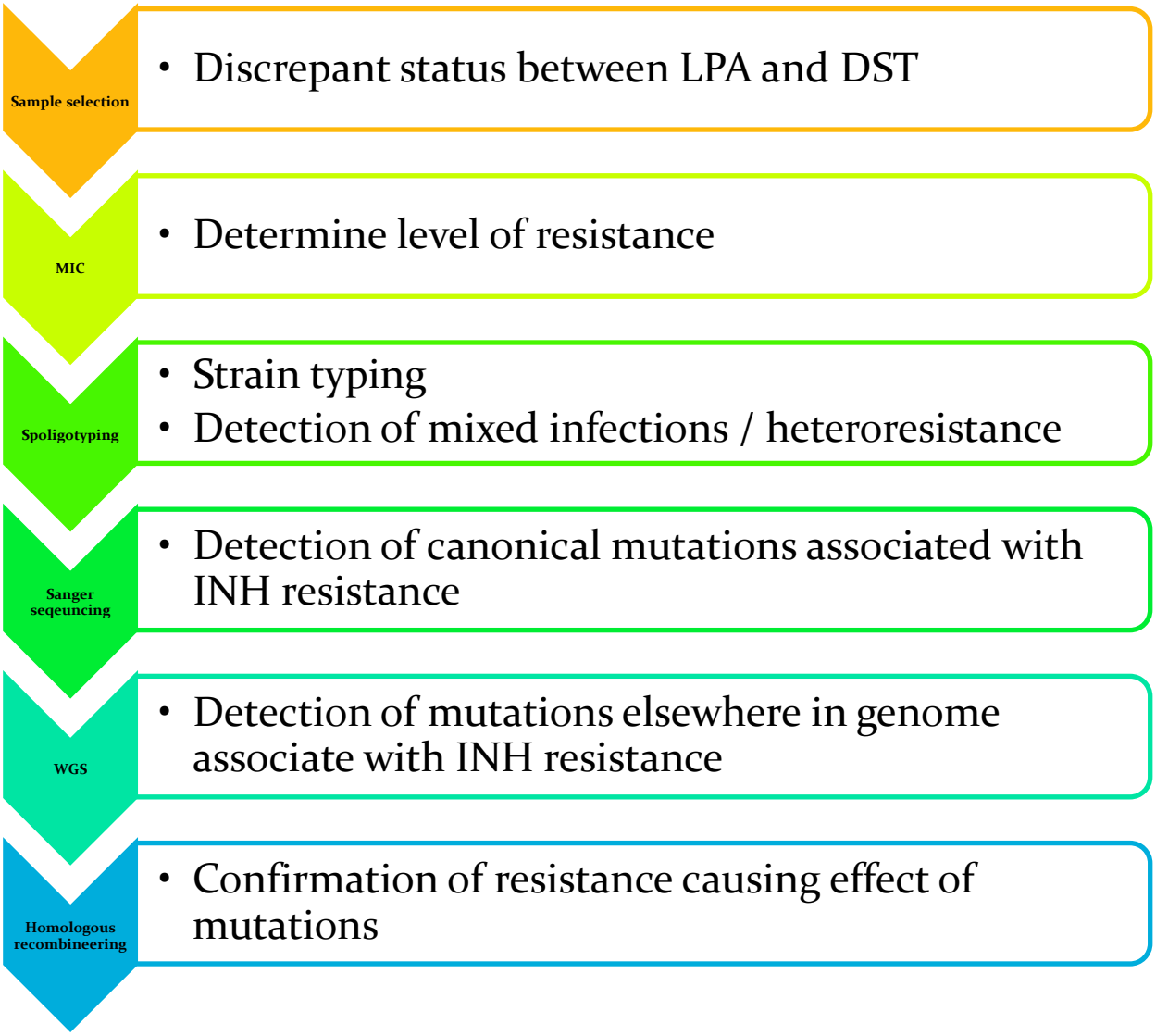


Figure 3: Summary of techniques used in this study.

3.1.1 Ethics approval

Ethical approval was obtained from the University of Stellenbosch Health Research Ethics Committee - reference number N09-11-296.

3.2 *M. tuberculosis* sample selection

The mechanisms of INH resistance were analyzed in 397 *M. tuberculosis* isolates obtained from the NHLS Port Elizabeth with discrepant status. These samples were collected between June 2014 and May 2016 from the Eastern Cape Province and were found to be susceptible to INH by a routine molecular diagnostic test, the Hain GenoType MTBDR*plus* LPA[1], but subsequent phenotypic drug-susceptibility testing (DST) on MGIT960 classified them as resistant to INH. The samples in this dataset were selected based on their discrepant status.

3.3 MIC determination

MIC determination was done according to the method described by Sirgel et al. in 2009[2]. The MIC of INH was determined using the BACTEC™ MGIT 960™ system[3] to confirm drug resistance.

The BACTEC™ MGIT 960 is used for rapid qualitative susceptibility testing of anti-tuberculosis drugs. MGIT tubes contain an oxygen-quenched fluorochrome, tris 4, 7-diphenyl-1, 10-phenonthroline ruthenium chloride pentahydrate, embedded in silicone at the bottom of the tube. The free oxygen in the medium is utilized by bacterial growth and replaced with carbon dioxide. Depletion of free oxygen results in fluorescence of the oxygen sensor in the MGIT tubes when exposed to UV light. The intensity of fluorescence is directly proportional to the extent of oxygen depletion and thus bacterial growth.

Typically, qualitative susceptibility testing is dependent on growth reaching the critical concentration as a standard, to differentiate between susceptible and resistant strains of *M. tuberculosis* to INH. However, due to the specifications of this study, the level of resistance to the drug was additionally determined quantitatively via extended individual drug susceptibility testing (exposure to different drug concentrations).

The BACTEC™ MIGIT method for DST is based on proportionality. The MGIT determines susceptibility by comparing growth of the test isolate at a standard inoculum size and in the presence of an antimicrobial to that in a drug-free control (growth control, GC) with a 1/100 diluted inoculum to represent 1% of the standard inoculum size. A resistant strain shows better growth in the presence of an antimicrobial as opposed to the GC. More growth in the GC as compared to the drug-containing test tube indicates that more than 99% of the bacterial population has been inhibited and the isolate is therefore considered susceptible. Thus, the bacterial population is considered to be resistant when 1% or more of the population is resistant to the test drug – in this case INH.

In the MGIT system, low-level resistance to INH is when there is growth between 0.1 and 1 µg/ml INH. High-level resistant status is designated to a sample if there is growth at concentrations higher than 1 µg/ml INH.

The reagents used for this method can be found in the **Supplementary data**. The storage of bacterial strains was done by centrifuging strains of *M. tuberculosis* grown in liquid Middlebrook 7H9 medium, at 3000 x g for 15 min. Next, the pellets were resuspended in glycerol peptone water medium and vortexed for 30 secs. Sterile glass beads (diameter of 4 mm) were covered with the suspension and preserved in Nunc vials at -70°C to -80°C.

To subculture from frozen cultures, the MGIT tubes were labelled with the specimen number and 0.8 ml of reconstituted OADC growth supplement was added aseptically to each MGIT tube. For each isolate, a single glass bead coated with *M. tuberculosis* and stored at -70°C was then removed from Nunc vials, under sterile conditions, and added to MGIT tubes labeled with corresponding specimen numbers. Thawing of remaining vials was prevented by keeping it on dry ice and by returning it to a -70°C freezer as soon as possible. The inoculated tubes were mixed well, scanned and placed into the BACTEC™ MGIT 960 instrument. The tubes were then incubated at 37°C until the instrument flagged them as positive.

Subsequently, 0.5 ml of sub-culture was transferred from tubes into the freshly prepared MGIT tube and incubated at 37°C. It is important to note that the day the MGIT tube became positive, is considered Day 0, after which the sub-culture was incubated for at least one more day (Day 1) before using it as inoculum for susceptibility testing. A positive tube was used for drug susceptibility testing up to and including the 5th day (Day 5) after it

became instrument positive. Cultures were thoroughly mixed through inversion or vortexing. The undiluted suspension was then used for inoculation of drug sets.

Subcultures of clinical isolates with discrepant status were grown at 4 different concentrations of INH to determine the level of resistance. These isolates were grown in the presence of INH at concentrations of 0.1, 1, 5 and 10 µg/ml and a no drug control was included. For each sample an aliquot of the no growth control, an aliquot of lowest concentration at which growth occurred (referred to in workflow as "drug tube") as well as a crude DNA aliquot of the original sample were made. The purpose of the no growth control is to serve as reference for a concentration of INH that *M. tuberculosis* is not resistant to and the drug tube is determinant of the level of resistance to INH. Thus, 3 tubes were obtained per clinical isolate via MIC, and these tubes were subsequently used for spoligotyping, Sanger sequencing and Whole Genome Sequencing (WGS). **Figure 4** below describes the workflow employed for each clinical isolate obtained from NHLS Port Elizabeth.

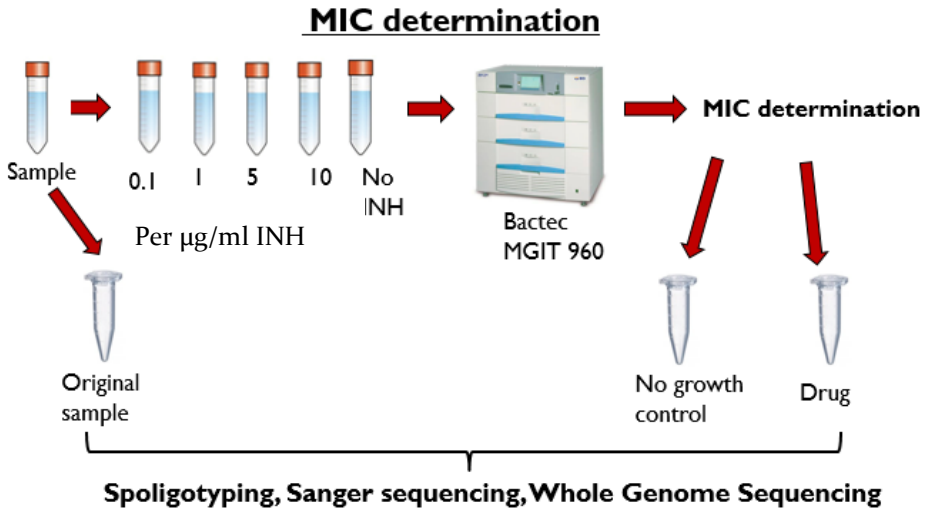


Figure 4: Schematic representation of method employed for MIC determination.

3.4 Strain typing

All 397 clinical isolates that formed part of this study were genotyped by spoligotyping according to the method described by Kamerbeek et al. in 1997[4]. It was done to classify the isolates in strain families and to detect the presence of mixed infections.

Spoligotyping is an internationally standardised reverse line probe assay based on the direct repeat (DR) region of *M. tuberculosis*, which consists of direct repeat sequences with unique spacer sequences that have been extensively studied in TB research[5]. This method is simple and has allowed the organization of the isolates of this study into strain families in accordance with the international spoligotype database[6].

Briefly, the 3 tubes per sample (original sample, growth control and drug tube) obtained in method a) were amplified via PCR. The PCR mix used for amplification is summarised in **Table 2** and the PCR cycles in **Table 3**

Table 2: PCR mix for amplification of DNA for genotyping

PCR mix	25 µl
ddH_2O	6.5 µl
Kapa Taq readymix[7]	12.5 µl
DRa[8]	2 µl
DRb[8]	2 µl
DNA template	2 µl

Table 3: PCR cycles for spoligotyping

Temperature	Time	Step	
95˚C	3 min	Activation	
94˚C	1 min	Denature	x30 (can vary from 28 to 35 cycles)
55˚C	1 min	Annealing	
72˚C	30 sec	Extension	
72˚C	10 min	Final Extension	

The solutions used can be found in **Supplementary data**. After amplification, 20 µl of PCR products was added to 150 µl 2x saline sodium phosphate ethylene diamine tetra-acetic acid EDTA/SSPE 0.1% sodium dodecyl sulphate (SDS). Next, the products were heat denatured at 99ºC for 10 minutes and immediately cooled on ice to prevent re-annealing of the single DNA strands. The membrane was washed at 60˚C in 2xSSPE/0.1% for 5 min and then placed on a support cushion on the miniblotter[9] with the slots perpendicular to the line pattern. Residual fluid was removed from the slots via aspiration and the slots were each filled with 150 µl diluted denatured PCR product and care was taken to avoid air bubbles. Hybridization occurred at 60˚C for 60 min on a horizontal surface. Next, the samples were removed from the miniblotter via aspiration. The membrane was washed twice in 2xSSPE/0.5%SDS for 5-10 min at 60˚C. Subsequently, 50 ml 2xSSPE/0.5%SDS (42˚C) with 12.5 µl Strepdavidine-peroxidase conjugate (500U/ml) was added and incubated at 42 ˚C for 45-60 min. After this, the membrane was washed twice with 2xSSPE/0.5%SDS at 42˚C˚C for 5-10 min. Next the membrane was rinsed with 2xSSPE for 5 min at room temperature and then incubated in 20 ml (10 ml solution 1 + solution 2) ECL mix[10] for 1 min 30 sec. The membrane was then exposed to film on X-ray after 5-20 min exposure time with a hyperprocessor[10]. Lastly the membrane was stripped for re-use by washing in 1%SDS at 80˚C for 30 min-1 hour and then storing at 4˚C in 20 mM EDTA pH8 overnight.

3.5 Sequencing of *inhA* promoter and *katG*

3.5.1 Primer design for Standard Hotstart PCR reactions

PCR primers for amplification of genes of interest were designed against the whole genome sequence of the *M. tuberculosis* H37Rv reference strain[11] using Primer 3 software, version 0.2[12]. The primers used in this study are summarized in **Table 4**.

Table 4: Primers used for PCR amplification

Gene	Primer	Primer sequence	Tm	Size of PCR product	Reference
katG	RTB 59	TGGCCGCGGCGGTCGACATT	62ºC	419 bp	[13]
	RTB 38	GGTCAGTGGCCAGCATCGTC			
inhA	inhAPfor	GCGTAACCCCAGTGCGAAAG	59ºC	282 bp	This study
promotor	inhAPrev	ACATTCGACGCCAAACAGCC			

New primers were designed for this study to amplify *inhA* promoter. This was done to increase the Tm of *inhA* promoter to that of *katG* so that PCR amplification can be done for both genes simultaneously.

3.5.2 Standard Hotstart PCR

DNA from the above mentioned samples (3 per isolate) were subjected to standard hotstart PCR reactions on the Applied Biosystems thermal cycler 2720 – to amplify the *katG* gene and *inhA* promoter region, followed by Sanger sequencing. The PCR mix[14] used for amplification is detailed in **Table 5** and the PCR cycles in **Table 6.**

Table 5: PCR mix for amplification of DNA for Standard Hotstart PCR

Total mix	25 µl
H_2O	9.375 µl
Q solution	5 µl
10x buffer	2.5 µl
$MgCl_2$	2 µl
dNTPs	4 µl
Primer mix (2 primers)	1 µl
Syto 9[15]	1 µl
Qiagen Hotstart Taq[15]	0.125 µl
DNA	1 µl

Table 6: PCR cycles for Hotstart

Temperature	Time	Step	
95˚C	15 min	Activation	
94˚C	1 min	Denature	x40
55˚C	1 min	Annealing	
72˚C	1 min	Extension	
72˚C	10 min	Final Extension	

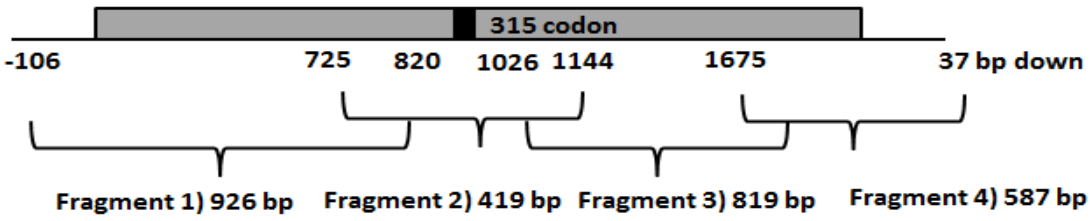
After PCR, confirmation of amplification was performed by High Resolution Melt (HRM) on the Rotor Gene 6000[15]. HRM analysis is molecular technique used for detection of genetic alterations related to drug resistance in *M. tuberculosis*. The DNA amplified via PCR is melted and the genetic alterations are detected according to the dissociation profile[16]. Amplification was judged to have occurred based on a melt curve. All PCR products that showed amplification were then Sanger sequenced on the ABI PRISM DNA sequencer model 377[17] at the Central Analytical Facility (CAF) of Stellenbosch University.

The Sanger sequences of the *katG* and *inhA* promoter of each clinical isolate were then analysed using Bioedit[18]. This was done with the purpose of looking for specific mutations in these regions that could potentially be associated with INH resistance. However, several isolates had no detectable mutations in these regions and were subsequently selected for Whole Genome Sequencing to detect mutations that could potentially exist elsewhere in the genome.

3.6 Sequencing *katG*

DNA from the samples with no detectable mutation in the region of *inhA* promoter or *katG* Sanger sequenced in section **3.5** had the rest of *katG* amplified via a PCR reaction – the *katG* gene region was split into 3 fragments (namely fragment 4, fragment 3 and fragment 1). These fragments exclude “fragment 2” around the *katG* 315 region that was amplified

in section **3.6**. The four different fragments of *katG* amplified in this study, as well as their



relevant positions within the gene, are presented diagrammatically in **Figure 5**.

Figure 5. Four fragments of *katG* amplified via Sanger sequencing in this study (1)-(4), where relevant fragment size is indicated prior to "bp" and numbers above square brackets indicate relative genetic position.

3.6.1 Primer design for Standard Hotstart PCR reactions (rest of *katG*)

PCR primers for amplification of three fractions of the gene of interest were designed against the whole genome sequence of the *M. tuberculosis* H37Rv reference strain [11] using Primer software 3 version 0.2[12]. The primers used in this study are detailed in **Table 7.**

Table 7: Primers used for PCR amplification of rest of *katG*

Gene	Primer	Primer sequence	Tm	Size of PCR product	Reference
***katG* fragment 1**	katG1F	GCGGAGGTCATCTACTGGGG	62ºC	926 bp	This study
	katG1R	TACCGAAAGTGTGACCGCCG			
***katG* fragment 3**	katG3F	GCTGACGAAGAGCCCTGC	62ºC	819 bp	This study
	katG3R	TTCGCCTTGTCGAGCAGCAT			

***katG* fragment 4**	katG4F	GTTGTGGTTGATCGGCGGG	62ºC	587 bp	This study
	katG4F	GCTGGCCACAACATCACGG			

Specialized primers were designed for the PCR reaction as there were no established primers in the literature.

3.6.2 Standard Hotstart PCR for rest of *katG*

DNA from the samples with no canonical mutation in *katG* or *inhA* promoter in section **3.5** were subjected to a hotstart PCR reaction on the Applied Biosystems thermal cycler 2720 – to amplify the rest of the *katG* gene. This was followed by Sanger sequencing. The PCR mix[14] used for amplification is summarised in **Table 5** and the PCR cycles in **Table 6**.

Qiagen Hotstart taq was used and 1 µl Syto9[15] was added in addition to the standard PCR mix as fluorescence for detection by the High-Resolution Melt (HRM) experiments on the Rotor Gene 6000[15], as described in section 3.5.All PCR products that showed amplification were then Sanger sequenced on the ABI PRISM DNA sequencer model 377[17] at the Central Analytical Facility (CAF) of Stellenbosch University.

The Sanger sequences of *katG* of each clinical isolate were then analysed using Bioedit [18]. This was done with the purpose of looking for specific mutations in these regions that could potentially be associated with INH resistance. However, several isolates had no detectable mutations in these regions and were subsequently selected for Whole Genome Sequencing to detect mutations that could potentially exist elsewhere in the genome.

3.7 DNA isolation and Whole Genome Sequencing

3.7.1 DNA isolation

Isolates of *M. tuberculosis* with no elucidated mechanism of resistance were cultured on 7H11 solid media for DNA extraction. Cultures on 7H11 plates were heat inactivated at 80˚C for 2 hours and lysozyme powder [19] was brought to room temperature. During the waiting period, the conical bottom of a 50 ml BD Falcon tube was filled with 5 mm glass beads and 6 ml extraction buffer (which consists of Mono Sodium Glutamic Acid (MSG), Tris, EDTA and hydrochloric acid) was added. After the cultures were heat inactivated, the confluent growth (usually from 2 plates) was scraped into tubes and the tubes were then placed in a vortex for 2 minutes until a milky consistency was achieved. Next, enough lysozyme for the plates (+ 1ml extra) was reconstituted by making use of a 100 mg/ml working concentration. A volume of 500 µl was added per tube and mixed by inversion in order to obtain a homogenous mixture. The tubes were then incubated at 37˚C for 2 hours. After the specified time, 600 µl proteinase K buffer[19] and 300 µl proteinase K[19] (10mg/ml) were added to the tubes and they were incubated at 45˚C overnight. The next day PCI (Phenol[19], chloroform[19], isoamylalcohol[19]; in proportion 25:24:1) was poured into a flask. After phase separation occurred, 5 ml PCI from the bottom layer was added to each tube of culture. Each tube was inverted every 30 min for 2 hours at room temperature in the fume hood. Next, the tubes were centrifuged at 3000 rpm for 20 min at room temperature (21˚C). From these tubes 5ml CI (chloroform/isoamylalcohol; 24:1) was put into new 50 ml tubes and the top layer PCI was aspirated with a 5ml pipette and added to CI. The new tubes were then inverted and centrifuged at 3000 rpm for 20 min at room temperature. We placed 600 µl sodium-acetate (Na-Ac; 3M) [19] in new 50 ml Falcon tubes and aspirated the top phase to the Na-Ac tubes. Following this, 2 rows of 1.5ml Eppendorf tubes were packed. The first row contained 1ml 70% ethanol in each tube and the second tube was empty. We added 7ml ice-cold isopropanol to the Falcon tubes and they were gently inverted until precipitated DNA could be observed. The DNA was then immediately fished with a glass rod. The rods covered in fished DNA were in placed in the ethanol Eppendorf tubes and the second row tubes of empty Eppendorf tubes with their glass rods were labeled. After 10 min the rods with DNA were placed in empty Eppendorf tubes and left to dry at room temperature. Once dry, 300 µl TE buffer (Sigma - Aldrich) was added to

the tubes, left for 5 min and the rods were removed while gently stirring to leave the DNA behind. The DNA was incubated at 4˚C overnight or alternatively, 65˚C for 2 hours. The isolated DNA was stored at -20˚C until use in further experiments.

3.8 Whole Genome Sequencing

Whole genome sequencing was done in collaboration with CDC in Atlanta, GA. Library preparation was done by making use of 1 µg of genomic DNA per sample. DNA was fragmented to an average length of 600 bp (focused ultrasonicator, Covaris LE220), and samples underwent size selection via 0.6X AMPure XP bead purification. Library preparation was completed as per the manufacturer's recommendation using the NEBNext Ultra DNA library prep kit for Illumina (New England BioLabs) with dual indices. Samples were pooled to a final loading concentration of 1 pM and sequenced on an Illumina NextSeq 550 platform using the 500/550 mid-output v2 (300 cycle) kit.

3.8.1 Bioinformatic analysis of WGS data

In order to analyse the WGS sequenced reads obtained, an in-house tool was used. The universal sequence analysis pipeline (USAP) - developed by the bioinformatics team at University Stellenbosch – consists of several software packages and python scripts in order to optimally extrapolate bioinformatics data. The reference sequence used was *M. tuberculosis* H37Rv (accession number: AL123456.3). The work flow is diagrammatically summarised in **Figure 6**.

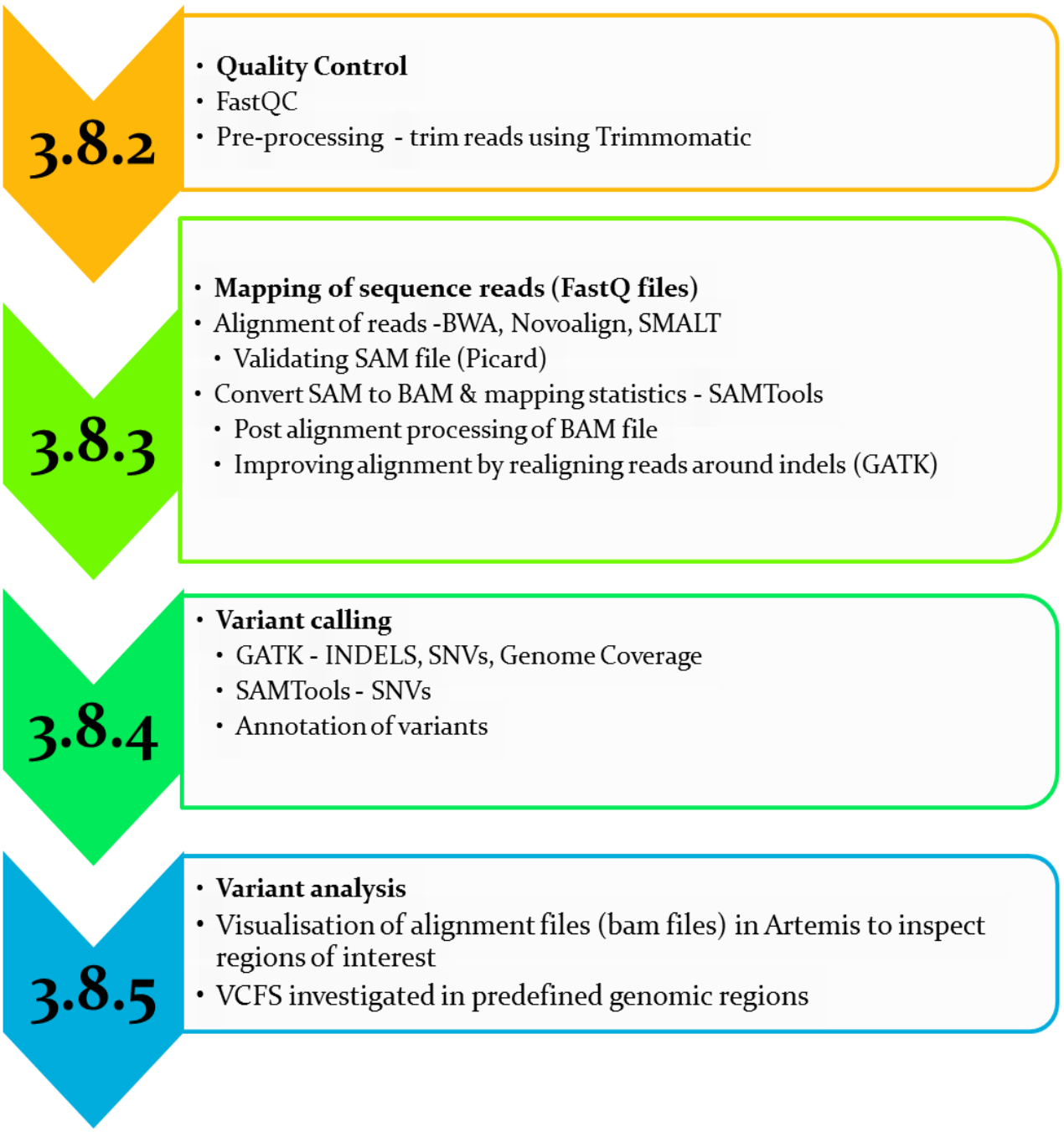


Figure 6: Work-flow employed for bioinformatic analysis of Illumina WGS reads to elucidate novel mechanisms of resistance to INH.

3.8.2 Quality Assessment

To assess the quality of the raw sequencing reads, each isolate was analysed using the FastQC program (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FastQC is a java script with a HTML output format[20]. The purpose of this program is to interpret sequencing quality based on the FastQ files that were sequenced on the Illumina platform. The FastQ files are analysed with several modules in order to extrapolate data relating to basic statistics, mean quality scores overall as well as reads per base position, distribution of read lengths, the distribution of uncalled bases, level of duplication in reads and identification of individual reads that may have a bias due to overrepresentation.

3.8.3 Computational Analysis

The FastQ files of isolates that passed the quality assessment were put through a WGS pipeline named USAP. USAP amalgamates a variety of software packages and utilises open source WGS analysis tools in conjunction with customised scripts to filter and annotate reads[21].

Trimming of reads

Trimmomatic[22] was used to remove low quality bases, duplicates and adapters with a Q value obtained via the Illumina platform, and a cut off value of 36 bp for the minimal read length[23]. Thus taking into consideration coverage, target length and error rate.

Alignment to a reference genome

The remaining high-quality reads were aligned to the *M. tuberculosis* H37Rv reference genome with three different alignment tools:

1) Burrows-Wheeler Aligner (BWA)[24] (http://bio-bwa.sourceforge.net/),
2) NovoAlign[25] (Novocraft - http://www.novocraft.com/products/novoalign/)
3) SMALT[26] (http://www.sanger.ac.uk/science/tools/smalt-0).

The Sequence Alignment/Map (SAM) file output format[27] , is a version of the alignment that improves the ability to work with it in the following steps.

SAM file conversion, validation and indexing

The "ValidateSamFile.jar" command of the Java based Picard Tools program (https://broadinstitute.github.io/picard/) was used to validate the obtained SAM files. The SAM files were converted into their binary (BAM) file format, using the "view" and "sort" command of SAMTools[27] (https://github.com/samtools/bcftools/releases/download/), followed by the "index" command, to simplify the processing steps for the next programs and create the indexed BAM (BAI) file.

Alignment statistics and realignment of BAM files

Specifics about the total number of input reads for the alignment, the number of duplicated, mapped and mate paired (associated at the same locus in the BAM file) reads were generated in a text file (txt) format with SAMTools "flagstat" command[28]. The java based Genome Analysis Toolkit (GATK) (https://software.broadinstitute.org/gatk/) was used to re-align the sequence reads around InDels, following the best practices recommendation. An interval file was generated from the BAM file format using the "RealignerTargetCreator" command, in order to be able to realign around small InDels, for which the GATK "IndelRealigner" command was used[28].

Removal of PCR duplicates

Thereafter, BAM files were re-sorted with the java based Picard Tools (http://broadinstitute.github.io/picard/) command "SortSam.jar", re-indexed with the "index" command from SAMTools, followed by the identification and removal of any PCR duplicates with the "MarkDuplicates.jar" command from Picard.

3.8.4 Validation of Variants

Variant calling

Single nucleotide polymorphisms (SNPs) and InDels were identified with the "mpileup" SAMTools command and the GATK[28] "UnifiedGenotyper" tool. Information about the variant positions, the phred (base calling program for DNA sequence traces to assess their quality) scaled probability and the alternative sequence are given in the obtained variant calling format (vcf) files[29] , which were used for further variant comparisons.

Annotation and functional classification of SNPs/InDels

High confidence SNPs and InDels were identified with the triplicate of mappers (BWA, Novoalign, SMALT) covered by more than 30% of the average depth of coverage of the reads and had a heterogeneity frequency of more than 0.8. They were annotated and if applicable, the consequent amino acid change, or effect on the reading frame of corresponding genes were calculated with a python script included in USAP.

Genome coverage

The genome coverage was calculated with the GATK "DepthOfCoverage" tool and areas that resulted with zero depth coverage, which implies bad mapping or a deletion, were subsequently identified with the Bash "grep" command in the obtained txt file.

Analysis of the effect of non-synonymous SNPs with PROVEAN

Amino acid changes as a result of non-synonymous SNPs were then categorised into neutral or deleterious changes, depending on the predicted impact on the biological function of a protein[30]. The Protein Variation Effect Analyzer (PROVEAN) tool for all species was used, which requires the protein sequence and affected amino acid, its position and the obtained variant, as an input. The standard cut-off score of -2.5 was used, which considers values below or equal to this threshold as deleterious and values above as neutral[31].

Investigation and visualisation

Finally, genome wide depth of coverage was determined with a specialized function in the Bedtools software package, for each mapping file. The depth of coverage files was searched for regions that had zero depth of coverage, i.e. no reads map to these regions. This step produces a file for each isolate with a list of positions that indicate possible deletions. However, visual inspection of every region was required to determine if the region indicated was a true deletion. This visual inspection was done using Artemis[32].

3.9 Homologous recombineering

For this study, the point mutagenesis via recombination technique adapted from van Kessel et al (2008)[33] was used. An in vitro point mutation was generated via a plasmid containing a recombineering gene, which required the selection of mutants from the selective media. The plasmid pJV75amber was used as a vector because it contained the gene, Che9c61 which is responsible for recombination. An oligonucleotide containing the mutation was recombined into a transformed culture. This oligonucleotide was recombined into the bacterial genome owing to the activity of the Che9c61 gene. The methods were as follows:

3.9.1 Oligonucleotide design for Homologous recombineering

Oligonucleotides containing the SNP of interest were designed on the lagging strand of the relevant gene sequence of *M. smegmatis* reference gene[11], [33]. Subsequently, a nucleotide BLAST[34] was run to determine sequence homology between the relevant gene in *M. smegmatis* and *M. tuberculosis* H37Rv[11]. The gene regions had 100% query cover and identity between *M. smegmatis* and *M. tuberculosis*. Thus, it was determined that homologous recombineering could be done in *M. smegmatis* and parallels could be drawn between the mechanism in the two organisms due to the sequence homology. The oligonucleotides were ordered from IDT[35] and it was specified that they should be PAGE purified. The oligonucleotides used in this study are summarized in **Table 8**.

Table 8. Oligonucleotides used for Homologous recombineering

Gene & codon	Oligonucleotide sequence	Reference genome
***katG* codon 347**	TACGGCAACGAGTGGGAGCTCACCAAGAGC[C**A**C]GCGGGCGCCAACCAGTGGAAGCCCAAGGAC	MSMEG_6384[36]
***katG* codon 420**	AAGGCGTGGTTCAAGCTGCTGCACCGCG[A**C**A]CGGGTCCGGTGACCCGCTACCTGGGCCCCGAG	MSMEG_6384[36]
***nat* codon 207**	TCAACACATCCGGGGTCACACTTCGTCACC[G**A**G]CTCACGGTCGCGGTGGTCACCGACGACG	MSMEG_0306[37]

Electrocompetent *Eschericia coli* cells

A culture of *Eschericia coli* (*E.coli*) XL1 was streaked out on a Luria Agar plate to serve as a multiplication vector for the plasmid due to its fast doubling time. Single colonies were picked and sub-cultured in 100 µl Luria Broth (LB) overnight at 37˚C. A volume of 20 µl culture was inoculated and grown in 100 ml LB to mid-log phase (OD_{600} = 0.8) at 37˚C. Next, the culture was transferred into 50 ml sterile tubes and centrifuged at 5000 rpm at 4˚C for 20 min. The supernatant was discarded and the pellet was resuspended in 1 ml 10% Glycerol. The wash step was repeated twice. Then the pellet was resuspended in 1 ml of 10% Glycerol again and 100 µl aliquots were stored at -80˚C for use in subsequent experiments.

E.coli transformation with pJV75amber

The pJV75amber plasmid (**Figure 7**) contains Kanamycin and Hygromycin resistance genes, thereby allowing for antibiotic selection.

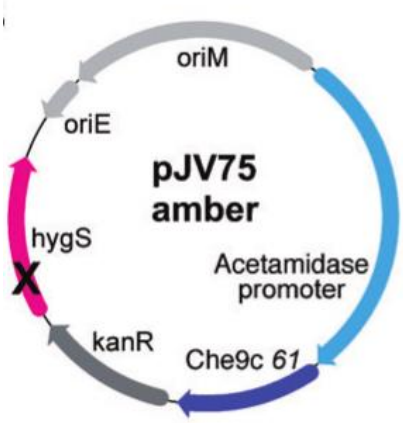


Figure 7. Plasmid pJV75amber[33] with origin of replication oriE confers expression of Che9c gp61 from the inducible acetamidase promoter. pJV75amber contains a hygromycin-resistant cassette (with two adjacent nonsense mutations which inactivate its function).

The plasmid was eluted on 20 µl H_20 and centrifuged at 13000 rpm for 5 min. Next, 1 µl pJV75amber was added to 100 µl *E. coli* competent cells and the cells were electroporated (Resistance: 200 ohms, Voltage: 2.5 kV, Capacitance: 125 µFD) in a Gene Pulser™[38]. Next, 500 µl SOC media was added and incubated at 37˚C for one hour. The transformed culture (25 µl, 50 µ, 100 µl) was then plated out onto Luria Agar containing Kanamycin (25 µg/ml) and incubated at 37˚C overnight. Single colonies were selected and cultured in 10-15 ml LB broth containing Kanamycin (25 µg/ml) in a 50ml flask tube at 37˚C overnight. Glycerol stock (10%) cultures were aliquoted and stored at -80˚C. Transformation was confirmed by DraI and EcoRV restriction enzyme digestion and gel electrophoresis (100 Volts for 1.5 hours) on 0.8% 1x SB buffer agarose gel after staining with Ethidium Bromide. Subsequently, the pJV75amber plasmid was isolated by the Wizard® Plus SV Minipreps DNA Purification system[39].

Electrocompetent *M. smegmatis* cells

M. smegmatis cells were cultured in Middlebrook 7H9 broth[3] supplemented with 0.05% Tween80 and 10% ADC at 37˚C until mid-log phase (OD_{600} = 0.5). The cells were harvested by centrifugation at 4000 rpm at 0˚C for 15 min. The supernatant was discarded and the pellet was resuspended in 50 ml Wash Solution (10% Glycerol containing 0.05% Tween80). The wash step was performed twice. The final pellet was resuspended in 1 ml Wash Solution and an aliquot of 200 µl culture was snap frozen in liquid nitrogen and stored at -80˚C.

M. smegmatis transformation with pJV75amber

The electrocompetent culture was centrifuged at 1500 rpm for 5 min at room temperature. The supernatant was discarded and the pellet resuspended in 100 µl Wash Solution. The wash step was repeated twice. A volume of 2 µl pJV75amber was added and the cells were electroporated (Resistance: 1000 ohms, Voltage: 2.5 kV, Capacitance: 125 µFD) in a Gene Pulser™[38]. Next, 900 µl of supplemented Middlebrook 7H9 broth was added and incubated for 3 hours at 37˚C. Then, the transformation culture (150 µl and 350 µl) was plated out on Middlebrook 7H10 plates[3] supplemented with 0.05% Tween80, 10% ADC and Kanamycin (25 µg/ml). Single colonies were picked and cultured in 10 ml of Middlebrook 7H9 broth[3] supplemented with 0.05% Tween80, 10% ADC and Kanamycin (25 µg/ml) at 37˚C for 2 days. The *M. smegmatis* culture was examined for contamination through Ziehl Neelsen staining.

Electrocompetent *M. smegmatis* containing pJV75amber

A volume of 100 µl *M. smegmatis* containing pJV75amber was cultured in 100 ml of Middlebrook 7H9 broth supplemented with 0.05% Tween80, 10% ADC, 20% succinate and kanamycin (25 µg/ml) at 37˚C until mid-log phase (OD_{600} = 0.4). Acetamide (1 ml) was added to the culture and incubated at 37˚C for 3 hours. The culture was put on ice for 2 hours and centrifuged at 400 rpm at 0˚C for 15 min. The supernatant was discarded and the pellet was resuspended in 25 ml 10% Glycerol (sterile). The wash step was performed thrice, effectively halving the volume of the 10% Glycerol per step. The final pellet was resuspended in 1 ml 10% sterile Glycerol and an aliquot of 100 µl culture was stored at -80˚C.

M. smegmatis cells containing pJV75amber with oligonucleotide transformation

The electrocompetent *M. smegmatis* culture was centrifuged at 1500 rpm for 5 min and the supernatant was discarded. The pellet was resuspended in 100 µl of 10% Glycerol with 0.05% Tween 80. A volume of 1 µl (500 ng) of oligonucleotide for *ndh* 18, *katG* 347 and 420 (see **Table 8**) was added in addition to Jcv198 oligonucleotide (5'-CCCTGTTACTTCTCGACCGTATTGATTCGG[ATG**]**ATTCCTACGCGAGCCTGCGG AACGACCAGG-3') (mutant codon in square brackets) and the cells were electroporated (Resistance: 1000 ohms, Voltage: 2.5 kV, Capacitance: 125 µFD) in a Gene Pulser™[38]. Next, 900 µl of supplemented Middlebrook 7H9 broth was added an incubated at 37˚C for 3 hours. Lastly, 110 µl transformation culture was plated onto Middlebrook 7H10 plates supplemented with 0.05% Tween80, 10% ADC and Hygromycin (100 µg/ml) at 37˚C for a duration of 4 days.

Selection of mutants

Single colonies were picked from the Middlebrook 7H10 plates supplemented with 0.05% Tween80, 10% ADC and Hygromycin (100 µg/ml) and cultured in 500 µl of Middlebrook 7H9 broth[3] supplemented with 0.05% Tween80, 10% ADC and Hygromycin (100 µg/ml) at 37˚C for 2 days before PCR amplification and DNA sequencing. A volume of 100 µl 50% Glycerol was added to 250 µl of mutant cultures and stored at -80˚C.

PCR amplification and DNA sequencing *ndh* codon 18, *katG* codon 347 and 420

Primers were designed to amplify individually codon 347 and 420 of *katG* as well as codon 18 of *ndh*. This was done by making use of the genome sequence of the *M. tuberculosis* orthologue for the *katG* (MSMEG_6384[36]) gene in *M. smegmatis*, as well as the orthologue for the *nat* gene (MSMEG_0306[37]) and Primer software 3 version 0.2 [12]. The primers used in this study are summarized in **Table 9**.

Table 9. Primers used for PCR amplification of homologous products

Gene	Primer	Primer sequence	Tm	Size of PCR product	Reference
katG	Smeg_katG for	AGGTCACCTGGACCCACAC	60ºC	340 bp	MSMEG_6384[36]
	Smeg_katG rev	CAGGTCACTCGACCGGAAG			
Nat	nat smeg for	CCAACACCGACCTCACCGC	61ºC	378 bp	MSMEG_0306[37]
	nat smeg rev	GCCACGTTGATCCCGAACCG			

The PCR mix [14] used per sample for amplification on the Applied Biosystems thermal cycler 2720 is summarised in **Table 10** and the PCR cycles in **Table 11**.

Table 10. PCR mix for amplification of DNA for Standard Hotstart PCR

Total mix	125 µl
H_2O	9.375 µl
Q solution	5 µl
10x buffer	2.5 µl
$MgCl_2$	2 µl
dNTPs	4 µl
Primer mix (2 primers)	50 µl
Syto9	50 µl
Qiagen Hotstart Taq[14]	0.125 µl
DNA (from unpurified transformed *M. smegmatis*)	2 µl

Table 11. PCR cycles for Hotstart

Temperature	Time	Step	
95˚C	15 min	Activation	
95˚C	45 sec	Denature	x45
62˚C	45 sec	Annealing	
72˚C	45 sec	Extension	
72˚C	10 min	Final Extension	

After PCR, mutant screening was conducted via High Resolution Melt (HRM) on the Rotor Gene 6000[15]. All PCR products that showed amplification were then Sanger sequenced on the ABI PRISM DNA sequencer model 377[17] at the Central Analytical Facility (CAF) of Stellenbosch University. This was done with in order to confirm the presence of the introduced mutations.

MIC determination of transformants

M. smegmatis colonies that were transformed and mutant confirmed, were plated onto Middlebrook 7H10 agar supplemented with 0.05% Tween80, 10% ADC at varying concentrations of INH (0, 2, 8, 16, 32 µg/ml)[41] and incubated for 3 days at 37˚C. A total of sixteen colonies were picked from the selective media and subsequently cultured in Middlebrook 7H9 broth supplemented with 0.05% Tween80, 10% ADC and INH (8 µg/ml) and incubated at 37˚C for 2 days. A total of 1 µl for each culture was spotted onto a Middlebrook 7H10 plate supplemented with 0.05% Tween80, 10% ADC and INH (16 µg/ml) using a grid and incubated at 37˚C for 3 days. The preceding two steps were repeated by increasing the concentration of INH in 8 µg/ml increments until a concentration of 40 µg/ml was achieved.

REFERENCES

[1] Hain. GenoType MTBDRplus | Detection resistance to rifampicin and/or isoniazid of MTBC complex http://www.hain-lifescience.de/en/products/microbiology/mycobacteria/tuberculosis/genotype-mtbdrplus.html (accessed May 2, 2017).

[2] Sirgel, F. A.; Wiid, I. J. F.; van Helden, P. D. Measuring Minimum Inhibitory Concentrations in Mycobacteria. In *Mycobacteria Protocols*; 2009; pp 173–186.

[3] Becton Dickinson Microbiology systems, Sparks, USA http://www.bd.com/en-us/offerings/capabilities/microbiology-solutions (accessed Apr 19, 2018).

[4] Kamerbeek, J.; Schouls, L.; Kolk, A.; Agterveld, M. Van; Kuijper, S.; Bunschoten, A.; Molhuizen, H.; Shaw, R.; Kamerbeek, J.; Schouls, L. E. O.; et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* **1997**, *35* (4), 907–914.

[5] Streicher, E. M.; Victor, T. C.; Van Der Spuy, G.; Sola, C.; Rastogi, N.; Van Helden, P. D.; Warren, R. M. Spoligotype signatures in the *Mycobacterium tuberculosis* complex. *J. Clin. Microbiol.* **2007**, *45* (1), 237–240 DOI: 10.1128/JCM.01429-06.

[6] Demay, C.; Liens, B.; Burguière, T.; Hill, V.; Couvin, D.; Millet, J.; Sola, C.; Zozio, T. SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect. Genet. Evol.* **2012**, *12* (4), 755–766 DOI: 10.1016/j.meegid.2012.02.004.

[7] Kapa Biosystems, Wilmington, Massachusetts, USA https://www.roche.com/media/store/releases/med-cor-2015-08-19.htm (accessed Apr 19, 2018).

[8] Isogen Life Science, IJsselstein, The Netherlands http://www.isogen-lifescience.com/ (accessed Apr 19, 2018).

[9] MN45; Immunetics, Cambridge, MA, USA http://immunetics.com/index.html (accessed Apr 19, 2018).

[10] Amersham, GE Healthcare Life Sciences, Chicago, IL, USA http://www.aecam.co.za/ (accessed Apr 19, 2018).

[11] Mycobrowser https://mycobrowser.epfl.ch/ (accessed Apr 19, 2018).

[12] Whitehead Scientific, South Africa http://whitesci.co.za/ (accessed Apr 19, 2018).

[13] Pretorius, G. S.; Van Helden, P. D.; Sirgel, F.; Eisenach, K. D.; Victor, T. C. Mutations in *katG* gene sequences in isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis* are rare. *Antimicrob. Agents Chemother.* **1995**, *39* (10), 2276–2281 DOI: 10.1128/AAC.39.10.2276.

[14] Qiagen HotStarTaq DNA polymerase Kit https://www.qiagen.com/us/shop/pcr/end-point-pcr-enzymes-and-kits/regular-pcr/hotstartaq-dna-polymerase/#orderinginformation (accessed Apr 19, 2018).

[15] QIAGEN, Venlo, The Netherlands https://www.qiagen.com/us/ (accessed Apr 19,

2018).

[16] Ramirez, M. V.; Cowart, K. C.; Campbell, P. J.; Morlock, G. P.; Sikes, D.; Winchell, J. M.; Posey, J. E. Rapid detection of multidrug-resistant *Mycobacterium tuberculosis* by use of real-time PCR and high-resolution melt analysis. *J. Clin. Microbiol.* **2010**, *48* (11), 4003–4009 DOI: 10.1128/JCM.00812-10.

[17] PerkinElmer | http://www.perkinelmer.com/ (accessed May 14, 2018).

[18] BioEdit Sequence Alignment Editor for Windows 95/98/NT/XP http://www.mbio.ncsu.edu/bioedit/bioedit.html (accessed Apr 19, 2018).

[19] South Africa | Sigma-Aldrich https://www.sigmaaldrich.com/south-africa.html (accessed Apr 19, 2018).

[20] Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed Apr 19, 2018).

[21] Dippenaar, A.; David, S.; Parsons, C.; Ann, M.; Hlokwe, T.; Claudius, N.; Pittius, G. Van; Abdu, S.; Musa, A.; Pain, A.; et al. Infection , Genetics and Evolution Progenitor strain introduction of *Mycobacterium bovis* at the wildlife-livestock interface can lead to clonal expansion of the disease in a single ecosystem. *Infect. Genet. Evol.* **2017**, *51*, 235–238 DOI: 10.1016/j.meegid.2017.04.012.

[22] Trimmomatic http://www.usadellab.org/cms/?page=trimmomatic (accessed Apr 19, 2018).

[23] Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30* (15), 2114–2120 DOI: 10.1093/bioinformatics/btu170.

[24] Albert, H.; Bwanga, F.; Mukkada, S.; Nyesiga, B.; Ademun, J. P.; Lukyamuzi, G.; Haile, M.; Hoffner, S.; Joloba, M.; O'Brien, R. Rapid screening of MDR-TB using molecular Line Probe Assay is feasible in Uganda. *BMC Infect. Dis.* **2010**, *10* (41), 1–7 DOI: 10.1186/1471-2334-10-41.

[25] NovoAlign | Novocraft http://www.novocraft.com/products/novoalign/ (accessed Apr 19, 2018).

[26] Ponstingl, H.; Ning, Z. SMALT-a new mapper for DNA sequencing reads. *F1000 Posters* **2010**.

[27] Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25* (14), 1754–1760 DOI: 10.1093/bioinformatics/btp324.

[28] McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20* (9), 1297–1303 DOI: 10.1101/gr.107524.110.

[29] Danecek, P.; Auton, A.; Abecasis, G.; Albers, C. A.; Banks, E.; DePristo, M. A.; Handsaker, R. E.; Lunter, G.; Marth, G. T.; Sherry, S. T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27* (15), 2156–2158 DOI: 10.1093/bioinformatics/btr330.

[30] Choi, Y.; Sims, G. E.; Murphy, S.; Miller, J. R.; Chan, A. P. Predicting the

Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* **2012**, *7* (10), e46688 DOI: 10.1371/journal.pone.0046688.

[31] Choi, Y.; Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31* (16), 2745–2747 DOI: 10.1093/bioinformatics/btv195.

[32] Rutherford, K.; Parkhill, J.; Crook, J.; Horsnell, T.; Rice, P.; Rajandream, M.-A.; Barrell, B. Artemis: sequence visualization and annotation. *Bioinformatics* **2000**, *16* (10), 944–945 DOI: 10.1093/bioinformatics/16.10.944.

[33] van Kessel, J. C. Van; Hatfull, G. F. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering : characterization of antimycobacterial drug targets. **2008**, *67* (January), 1094–1107 DOI: 10.1111/j.1365-2958.2008.06109.x.

[34] NCBI| Nucleotide BLAST https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (accessed May 2, 2018).

[35] OligoAnalyzer 3.1 | IDT https://eu.idtdna.com/calc/analyzer (accessed May 2, 2018).

[36] Mycobrowser https://mycobrowser.epfl.ch/genes/MSMEG_6384 (accessed Aug 8, 2018).

[37] Mycobrowser https://mycobrowser.epfl.ch/genes/MSMEG_0306 (accessed May 11, 2018).

[38] Bio-rad, Hercules, CA 94547, USA |Gene Pulser Xcell™ Electroporation Systems http://www.bio-rad.com/en-cn/product/gene-pulser-xcell-electroporation-systems?ID=b1a35eb3-d55c-47b3-aaf3-95e4d1d85848 (accessed May 2, 2018).

[39] Promega, Madison, WI 53711-5399, USA https://worldwide.promega.com/resources/protocols/technical-bulletins/0/wizard-plus-sv-minipreps-dna-purification-system-protocol/ (accessed May 2, 2018).

[40] Mycobrowser https://mycobrowser.epfl.ch/genes/MSMEG_3621 (accessed May 11, 2018).

[41] Xu, X.; Vilchèze, C.; Av-Gay, Y.; Gómez-Velasco, A.; Jacobs, W. R. Precise null deletion mutations of the mycothiol synthesis genes reveal their role in isoniazid and ethionamide resistance in *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* **2011**, *55* (7), 3133–3139 DOI: 10.1128/AAC.00020-11.

Chapter 4: Results

3.1 Introduction

The study collection period stretches from June 2014 to May 2016 and involved the identification of 397 *M. tuberculosis* isolates that were found to be susceptible to INH based on the genotypic test: GenoType MTBDR*plus*, but resistant with the phenotypic DST test using MGIT960^{TM}. Samples originated from across the Eastern Cape and samples with discrepant results between the LPA and DST were identified at the NHLS Port Elizabeth Laboratory.

3.2 MIC determination

The MIC determination of the 397 clinical isolates used in this study was necessary to determine the level of resistance to INH for each isolate. Low-level resistance to INH is defined as growth between 0.1 and 1 µg/ml INH. High-level resistant status is designated to a sample if there is growth at concentrations higher than 1 µg/ml INH. Isolates are INH susceptible if no growth is observed below concentrations of 0.1µg/ml INH.

As per these specifications, it was found that 42 (10.6%) of the clinical isolates were susceptible to INH, thus these isolates could be treated with INH effectively. Additionally, 116 (29.2%) were found to be low-level resistant to INH per the MIC data. The relevance of this high proportion of INH low-level resistant isolates will be interpreted more in depth in the discussion section. It is of interest as patients with low-level resistant strains of *M. tuberculosis* could be treated with higher dosages of INH, and their drug treatment regimen does not have to be adjusted severely.

A high proportion of the isolates (61%, n=242) from this study showed resistance to INH at varying degrees higher than 1 µg/ml. For 155 isolates, growth occurred at concentrations ≥10 µg/ml INH. This effectively rules out INH as a treatment option as effective MIC cannot be clinically attained and alternative treatment regimens must be considered to treat these patients. For 10 isolates (2.5%) it was not possible to obtain MIC, as these were contaminated or lost viability. **Table 12** below is a summary of the results of the resistance status obtained via MIC determination.

Table 12. Resistance status of clinical isolates by MGIT960^{TM} MIC determination

Resistance status	Concentration	Number of isolates
Susceptible	≤0.1	28
Low-level resistant	>0.1 ≤1	116
Resistant	≥10	155
	>5 ≤10	11
	>1 ≤5	76
Contaminated (not *M. tuberculosis*)	-	5
Lost viability	-	6

3.3 Strain typing via spoligotyping

The isolates included in this study were classified according to strain type by spoligotyping. Three spoligotype patterns were obtained for each isolate: the original culture and 2 samples from the MIC testing (the growth control and the tube with the highest concentration of INH where growth was observed). This was done to identify possible mixed infections which could have been the reason for the discrepancy of the INH testing methods.

Table 13 below summarizes the proportions of the genotypes observed by spoligotyping of the original cultures. Out of the 397 isolates, 216 (54.4%) were found to be part of the Beijing family. The results observed were largely in accordance with the expectation based on the prevalence of the Beijing strain family in South Africa[1],[2]. Other major strain families identified include: LAM (9.3%), S (8.3%), EAI (6.8%), T (6.3%), X (4.8%), CAS (3.3%) and H (1.0%). We found 5 different types for 7 different isolates that is not recoded in the International Spoligotype database (SITVIT[3]) and 1 of these patterns is of unknown family/lineage. The remaining 5.8% of isolates consisted of mixed infections (0.25%); unknown lineage (2.01%), one case of BCG (0.25%), 2 isolates with the MANU genotype (0.5%), non-tuberculosis mycobacteria (1.26%) and no spoligotype pattern could be obtained for 1.5% of isolates due to lost viability, contamination or not part of the *Mycobacterium tuberculosis* complex. This data is visually represented in **Figure 8.**

We identified one original culture with a mixed infection of LAM and Beijing. A further four isolates with mixed infections were observed based on the spoligotype patterns. This is of interest as the presence of more than one strain type in a single isolate could potentially justify the discrepancies between the culture based method and molecular technique in diagnosing INH resistance. The statement “not in SITVIT” in the **Table 13** indicates that the spoligotype pattern isn’t present on the SITVIT database. “Unknown” in the lineage or SIT family column means that the spoligotype pattern of this isolate does not resemble any known lineage or family markers.

Table 13. Genotypes observed in sample set via spoligotyping

Lineage	N	SIT family	n	SIT	Spoligotype pattern	n
Beijing	216	BEIJING	212	1		212
		BEIJING-LIKE	4	796		4
LAM	37	LAM3	34	33		20
				4		6
				719		4
				1293		1
				1841		1
				NOT IN SITVIT		1
				NOT IN SITVIT		1
		LAM9	1	42		1
		LAM11_ZWE	2	815		1
				2017		1
S	33			34		28
				71		3
				1113		1
				NOT IN SITVIT		1
EAI	27	EAI1_SOM	21	48		20
				806		1
		EAI5	6	947		3
				236		3
T	25	T1	23	53		18
				154		2
				1112		1
				2296		1
				803		1
		T3	2	37		2
X	19	X1	4	119		2
				2226		1
				336		1
		X2	4	137		3
				478		1
		X3	11	92		9
				2286		1
				2020		1
CAS	13	CAS1_KILI	9	21		9
		CAS1_DELHI	2	1092		1

				26		1
		CAS	2	NOT IN SITVIT		2
H	4	H1		62		1
				610		1
				NOT IN SITVIT		1
		H3		631		1
Mix	1	BEIJING/LAM3		Mix		1
Unknown	8	Unknown		NOT IN SITVIT		8
Bovis	1	BOVIS1_BCG		482		1
MANU	2	MANU1	1	337		1
			1	**100**		**1**
NTM	5	Possible NTM				5
no result	6					6
						397

*If “no result” in the lineage column indicates that no results could be obtained for these isolates due to loss of viability due to contamination or this isolate could possibly be a non-tuberculosis mycobacterium (NTM).

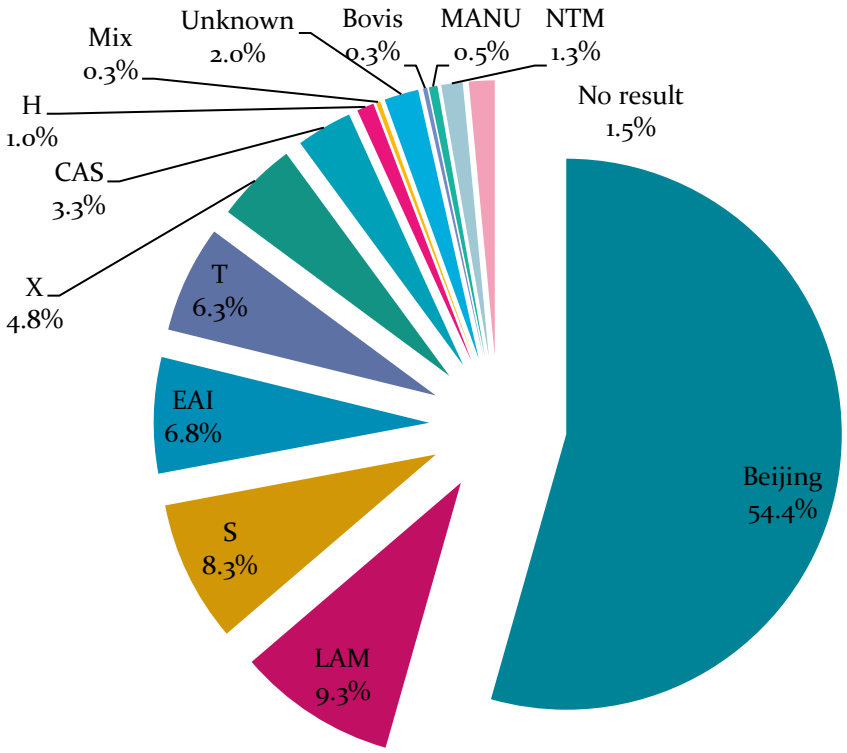


Figure 8. Proportion of genotypes observed via spoligotyping

3.4 Sequencing of *inhA* promoter and *katG*

We performed Sanger sequencing of *katG* gene and *inhA* promoter to identify mutations potentially associated with INH resistance. As in the case of the spoligotyping, we performed the PCR and sequencing on 3 samples per isolate: the original culture and 2 samples from the MIC testing (the growth control and the tube with the highest concentration of INH where growth was observed).

This was done to enable identification of possible mixed infections and to identify heteroresistance, when the *M. tuberculosis* strain is in the process of acquiring INH resistance, but would be missed by the LPA.

It was found that the 42 samples that were susceptible to INH by MGIT960 MIC determination had wild type *katG* and *inhA* promoter. These samples were susceptible to INH on the LPA and were thus correctly diagnosed as they were susceptible on the DST as well. Sanger sequencing detected 8 isolates with classical mutations conferring resistance that should have been detected by LPA through the absence of a wild type probe, and the

presence of a mutant probe. Additionally, 7 isolates had mutations identified at codon 315 of *katG*, which is different to the mutant probes of the LPA test, but should have been detected due to the absence of a wild type probe.

By doing Sanger sequencing on the original culture and samples obtained from the MIC determination, we could identify 10 isolates that were in the process of gaining resistance to INH. These isolates were therefore classified as heteroresistant (resistant and susceptible in the same isolate). An example of the sequencing results of one these identified are given in **Figure 9**.

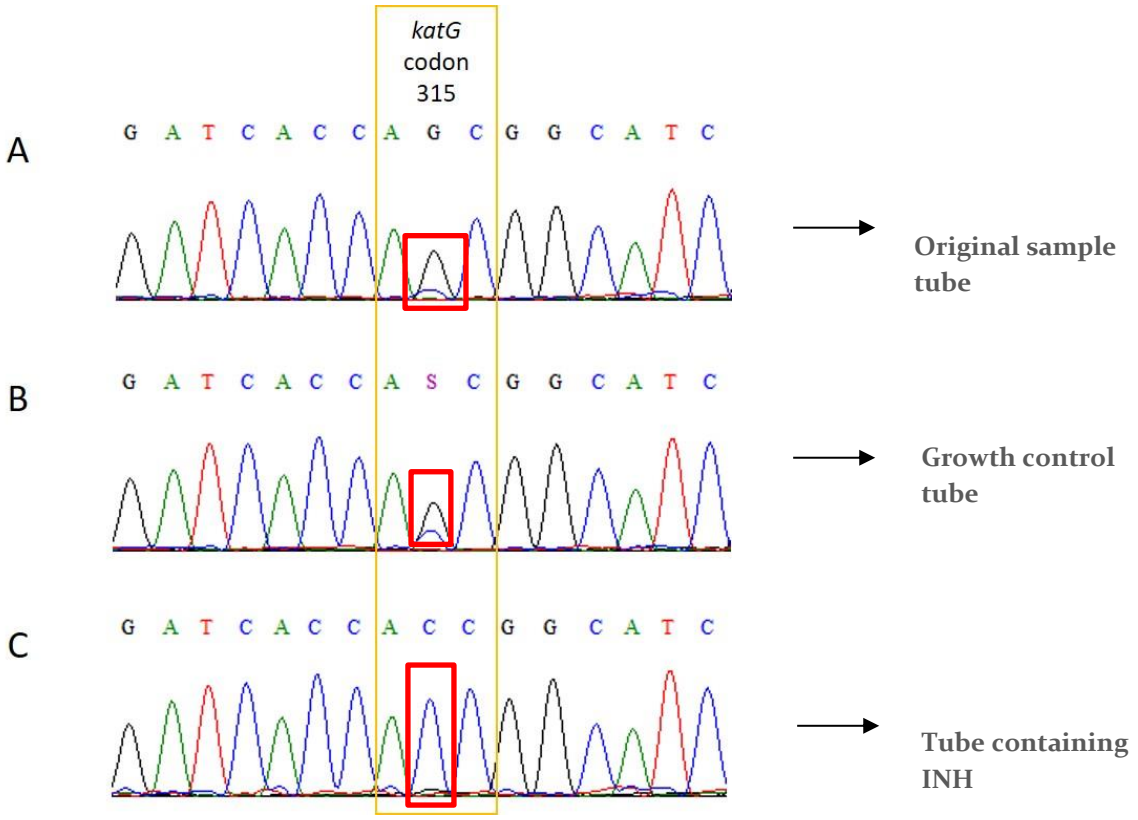


Figure 9. Image of Sanger sequencing results for as an example of heteroresistance.

A: The Sanger sequencing results of the original sample, indicating codon 315 as wild type (AGC), although a slight peak of the mutant (C) is underlying. B: Sample of the growth control tube indicates the presence of both the wild type and the mutant. C: The sample from the tube that contained 10µg/ml INH, indicates only the mutant (ACC) present. The wild type could not survive in the presence of INH, only the mutant remained.

Additionally, Sanger sequencing of the areas surrounding the *katG* codon 315 and *inhA* promoter enabled identification 17 different mutations in 89 isolates. These mutations may have been the cause of INH resistance, but would have been missed by LPA method. See **Table 14** and **15**.

For 200 (50.25%) of the isolates that were found to be resistant to INH, we could not identify a possible explanation for the discrepancy between genotypic and phenotypic methods by Sanger sequencing of *katG* or *inhA* promoter.

Table 14. Mutations identified via Sanger sequencing around *katG* codon 315

Mutation identified	# of high level resistant isolates	# of low level resistant isolates	Strain family
2 bp ins at codon 346 of *katG*	1		Beijing
katG 271 ACT-ATT	3		S
katG 272 TTC-TCC	1		Beijing
katG 273 GGT-GTT	2		Beijing
katG 360 GGT-TGT	1		Beijing
katG 279 GGC-GAC; *inhA* -77 G-A	2		Unknown
katG 285 GGC-GAC		18	Beijing
katG 291 GCT-GAT	2		Beijing
katG 295 CAG-CCG	3		Beijing
katG 300 TGG-CGG	2		T1
katG 300 TGG-GGG	2		T3
katG 325 CCG-CGG	1		Beijing
katG 325 CCG-TCG		1	Beijing
katG 333 CTC-TTC	1		H3
katG 337 TAC-TAG	1		Beijing
katG 341 TGG-CGG	1		Beijing
katG 347 CCT-CAT	1		T1

Table 15. Mutations identified via Sanger sequencing around *inhA* promoter region

Mutation identified	**# of high level resistant isolates**	**# of low level resistant isolates**	**Strain family**
inhA promoter -15 C-T		2	Beijing
inhA promoter mix wt/-15 C-T	2		BEIJING/LAM3
inhA promoter -17 G-T	2		Beijing
inhA promoter-77 G-A	2		NOT IN SITVIT
fabG1 12 CCA-CGA	1		X3

Note, *fabG1* is included in this table as it is an alternative promoter for *inhA* and increases its expression level[4]. Thus it is relevant for diagnostics.

3.5 Sequencing *katG*

Preliminary WGS results from a pilot investigation of 53 isolates with discrepant status indicated that a significant proportion had mutations elsewhere in the *katG* gene, thus outside the *katG* 315 region of *M. tuberculosis*. Mutations in this region serve as indicators of high-level resistance to INH per the LPA, thus mutations elsewhere in the *katG* gene would be missed by the molecular test but may be contributing to resistance. Thus, we performed Sanger sequencing on the entire *katG* gene for isolates with a mechanism of resistance that could not be elucidated via Sanger sequencing around the 315 region. Only SNPs confirmed to have a deleterious effect on the protein via PROVEAN[5] were included. The results are summarized in **Table 16**.

Table 16: *katG* mutations outside region of LPA detection

Mutation identified	# of high level resistant isolates	# of low level resistant isolates	Strain family
katG 346 TCC-TTC	1		Beijing
katG 394 ACG-ATG	1		Beijing
katG 412 TGG-TAG	1		F11/LAM3
katG 406 GAC-AAC	1		EAI1_SOM
katG 419 GAC-TAC	14		S
katG 420 ATG-ACG	1		CAS1_KILI
katG 421 GGT-CGT	1		Beijing
katG 438 TGG-CGG	4		Beijing
katG 439 CAG-CCG		1	Beijing
katG 457 AGC-ATC	1		Beijing
katG 466 GGA-CGA		1	Beijing
katG 487 GAC-GAG	2		Beijing
katG 501 CCA-TCA	1		S
katG 538 GTC-TCC	1		Beijing
katG 552 ATA-AAA	1		Beijing
katG 553 GAG-AAG	1		X3
katG 570 GGC-CGC	1		LAM3

katG 594 TTC-CTC		1	Beijing
katG 644 GGC-CGC	1		LAM3

3.6 Whole Genome Sequencing Analysis

We had the opportunity to perform WGS on a total of 130 of the first 300 isolates for which we were unable to find an explanation for INH discrepancies between DST techniques. This was done to search for other possible mechanisms of INH resistance in the genome of *M. tuberculosis*. The analysis of the results of the WGS revealed some interesting phenomena. In order to keep with the precedent set in the previous section, **Table 17** below describes *katG* mutations that were missed by the LPA and Sanger sequencing.

Table 17: *katG* mutations outside region of LPA detection observed via WGS

Mutation identified	**# of high level isolates with mutation**	**# of low level isolates with mutation**	**Strain family**
large deletion (+/-199226) including entire *katG*	1		X3
large deletion (+/-19252 bp) including entire *katG*	1		X3
large deletion (+/-14671 bp) including entire *katG*	1		LAM3 and S /convergent
katG deletion +/- 3268 bp (part of *katG* deleted)	1		Beijing

Mutation identified	# of high level isolates with mutation	# of low level isolates with mutation	Strain family
katG deletion +/- 3232 bp (part of *katG* deleted)	1		Beijing
katG deletion +/- 1252 bp (part of katG deleted)	2		LAM3 and S /convergent
katG deletion +/- 1045 bp (part of *katG* deleted)	1		T1
katG promoter -11 T-G	2		Beijing
katG 1 GTG-GCG	2		Beijing
katG 1 GTG-GTA	1		T1
katG 71 ATC-ACC + 91 TGG-CGG + 121 GGC-CGC + 140 AGC-GGC		1	Beijing
katG 91 TGG-CGG	5		Beijing (3) & S (2)
katG 93 GCC-ACC		2	Beijing
katG 94 GAC-GGC	4		Beijing
katG 104 CGG-CAG	1		LAM3 and S /convergent
katG 104 CGG-TGG	1		LAM3
katG 105 ATG- ATA	1		Beijing
katG 109 GCT-GTT		1	LAM11_ZWE
katG 113 TAC-TGC + *katG* promoter-	1	1	Beijing

Mutation identified	# of high level isolates with mutation	# of low level isolates with mutation	Strain family
9 A-C			
katG 121 GGC-GTC	1		Beijing
katG 121 GGC-CGC		1	Beijing
katG 127 CAG-CCG	1		LAM3
katG 138 AAC-AGC	1		Beijing
katG 140 AGC-GGC		1	Beijing
katG 142 GAC-GCC	5		Beijing
katG 144 GCG-GTG	1		Beijing
katG 172 GCG-ACG	2		Beijing
katG 185 TTC- TTG	1		Beijing
katG 189 GAC-CAC	1	1	Beijing
katG 189 GAC-GGC		1	Beijing
katG 203 CGG-CTG	1		Beijing
katG 232 CCG-TCG	1	2	X2
	1		S
katG 380 ACT-ATT		1	Beijing
katG 394 ACG-ATG		1	Beijing
katG 419 GAC-TAC	1		S

Mutation identified	# of high level isolates with mutation	# of low level isolates with mutation	Strain family
katG 420 ATG-ACG		1	CAS1_KILI
katG 421 GGT-AGT	1		T1
katG 463 CGG-CTG	1		Beijing
katG 479 GCG-CCG	1		Beijing
katG 483 TTC-TCC	1		LAM3
katG 496 CGC-GGC	1		Beijing
katG 466 GGA-CGA + 594 TTC-CTC		1	Beijing
katG 695 GAC-GGC	1		CAS1_KILI
katG 711 TAT-GAT	1		Beijing

Visual inspection via Artemis[6] revealed 1045 bp and 1252 bp deletions within the *katG* gene. Additionally, there were two isolates with deletions of +/- 3200 bp. The first was 3232 bp which included the first 430 bp of *katG* and the second was 3268 bp which included the first 478 bp of *katG*. Interestingly, in a further 2 isolates we identified a 19 kbp and a 14 kbp deletion that included the *katG* gene. These large deletions were missed by the initial screening as there is an underlying population with an intact wild type *katG* gene. **Figure 10** displays these interesting large deletions obtained via Artemis analysis.

A

B

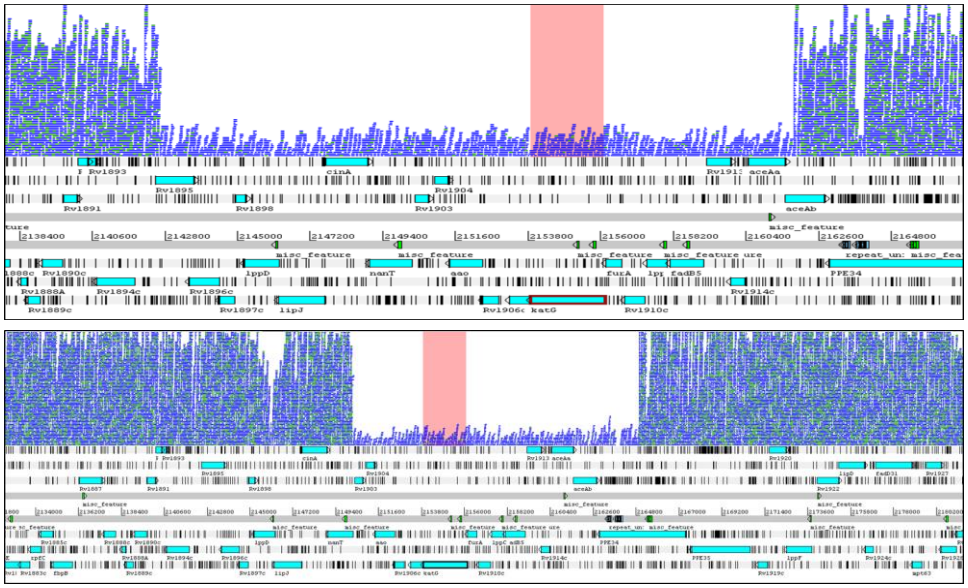


Figure 10. Visual inspection by Artemis to detect large deletions identified in the study

A: result for sample 61 with +/-19 kbp deletion in majority of population (*katG* included). B: result for sample 66 with +/-14 kbp deletion in majority of population (*katG* included). For both figures the parts with low depth of coverage represents the deletions. These samples have mixed infections and the reads in the region with low depth of coverage represents the wild type strain that does not have the deletion. The pink shaded area represents the *katG* gene.

Based on the results obtained via WGS and Sanger sequencing it is evident that a considerable proportion of isolates in this study had mutations in the *katG* gene that could be associated with INH resistance. Additionally, different mutations were associated with varying degrees of resistance. Based on this premise, **Figure 11** shows all *katG* mutations detected in this study and their associated level of resistance.

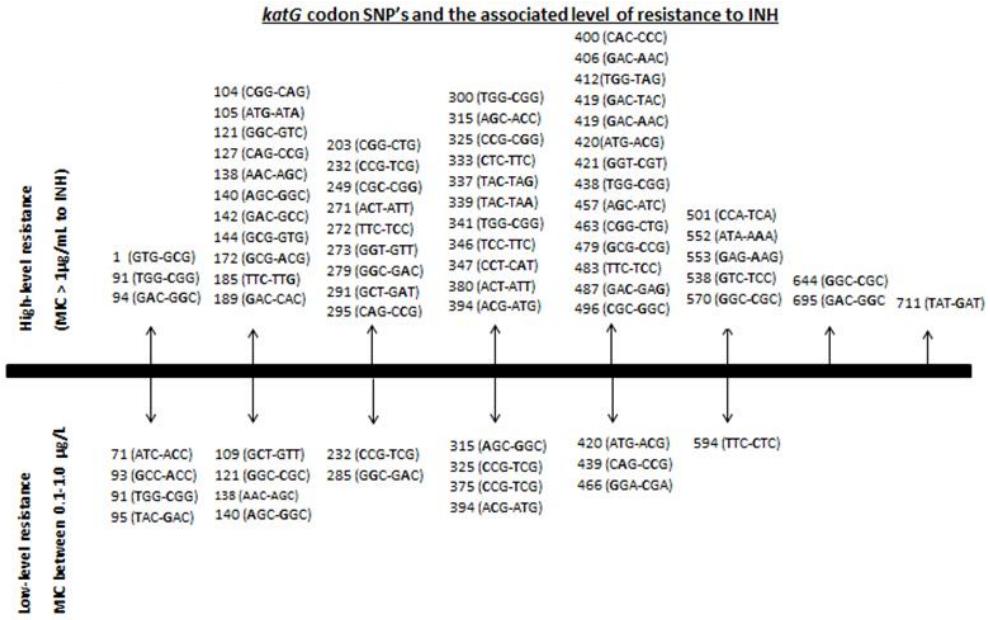


Figure 11. All *katG* SNP mutations identified in this study and their associated level of resistance.

Next, isolates were screened for mutations in genes (excluding *katG*) and promoters that have been linked to INH resistance per the literature[7],[8]. The genes include Rv0129c (*fbpC*)[9]; Rv0340[9]; Rv0341 (*iniB*)[9]; Rv0342 (*iniA*)[9]; Rv0343 (*iniC*)[9]; Rv1483 (*mabA*)[9]; Rv1484 (*inhA*)[9], Rv1592c[9]; Rv1772[9]; Rv2242 (*srmR*)[9]; Rv2243 (*fabD*)[9]; Rv2247 (*accD6*)[9]; Rv3139 (*fadE24*)[9]; Rv3566c (*nhoA*)[9], Rv2245 (*kasA*), Rv2428 (*aphC*) [10]; Rv1854c (*ndh*) [11], Rv1909c (*furA*) [12]; Rv2427a (*oxyR*) [13]; Rv2846c (*efpA*) [14]; Rv3795 (*embB*) [15]; Rv3199c (*nudC*); Rv3566c (*nat*); Rv0486 [16] and Rv3854C [17]. The promoter mutations included mutations at positions 1673356-1673439 (*inhA* promoter) [18]. The results for SNPs associated with low-level resistance to INH are summarised in **Table 18.** Conversely, **Table 19** categorises the mutations that are potentially causative of high-level resistance to INH.

Table 18. SNPs associated with low-level resistance to INH

Gene	Codon position	Amino acid change	Strain families of associated SNP
Rv1592c	322	I/V	Beijing, LAM3, H1, X2 & X3
fabD	270	P/L	LAM3
fabG1	230	L/L	LAM3
inhA	194	I/T	EAI1_SOM
inhA	94	S/A	Beijing, T1 & S
ndh	18	V/A	S

Table 19. SNPs associated with high-level resistance to INH

Gene	Codon position	Amino acid change	Strain families of associated SNP
Rv1592c	322	I/V	Beijing, EAI1_SOM, EAI5, X1& T1
	430	V/A	T1
accD6	229	D/G	Beijing, EAI1_SOM, T1, X1 & S
fabD	199	A/T	T1
fabG1	230	L/L	S & X1
fadE24	243	L/Q	T1
	430	M/L	CAS1_KILI & CAS1_DELHI
fbpC	158	G/S	CAS1_KILI & CAS1_DELHI

Gene	Codon position	Amino acid change	Strain families of associated SNP
	158 & 160	G/S; S/A	CAS1_DELHI
inhA	194	I/T	EAI5, EAI1_SOM & S
iniA	481	H/Q	EAI1_SOM & EAI5
iniC	476	R/S	EAI1_SOM & EAI5
nat	207	G/R	S
ndh	18	V/A	S
	268	R/H	EAI1_SOM & EAI5
nhoA	207	G/R	S
nudC	126	S/A	S
	237	Q/P	Beijing, EAI1_SOM & EAI5
	239	P/R	Beijing
	72	P/S	X1
oxyR	13	L/F	EAI1_SOM & EAI5
	131	S/T	EAI1_SOM & EAI5
srmR	323	M/T	CAS1_KILI & CAS1_DELHI

No mechanism of resistance could be elucidated for six of the 130 WGS isolates per the current literature pertaining to SNPs associated with INH resistance. These included five isolates from strain family LAM3, of which three where low-level resistant and two was high-level resistant to INH. The last isolate was of Beijing lineage and high-level resistant to INH.

When combining all the results obtained in this study, we were able to deduce the reason for INH testing discrepancies for 278 isolates (70%). However, for 119 isolates (30%) we are yet unable to elucidate the mechanism of INH resistance, and thus the reason for the discrepancies at the Routine laboratory. **Table 20** below summarises the combined results obtained in this study. Note that the isolates with no elucidated mechanism of resistance to INH include the six that couldn't be diagnosed via WGS, as well as isolates that were not whole genome sequenced and the mechanisms of resistance couldn't be deduced via Sanger sequencing. Future studies will whole genome sequence the rest of the dataset.

Table 20. Combined results of MIC, mutation analysis by Sanger sequencing and WGS

Interpretation based on combined results	**# of isolates**
INH sensitive - wild type *katG* & *inhA* promoter	28
INH resistant - mutations present (should have been detected by MTBDR*plus*)	9
INH resistant - mutations present - should have been detected by MTBDR*plus* (*katG* wt probe absent)	3
Low level resistance - inhA prom mutation identified (should have been detected with MTBDR*plus*)	1
Low level resistance - mutations present - should have been detected by MTBDR*plus* (*katG* wt probe absent)	5
INH resistant - mutation not part of MTBDR*plus* assay	148
Low level resistance - mutation not part of MTBDR*plus* assay	53
INH resistant - mutations not detected by MTBDR*plus* due to low proportion (heteroresistance)	7
Low level resistance - *inhA* prom not detected by MTBDR*plus* due to low proportion (heteroresistance)	4
INH resistant - mutations not detected by MTBDR*plus* due to low proportion (mixed infections)	6
Low level resistance - mutations not detected by MTBDR*plus* due to low proportion (mixed infections)	2
BCG; intrinsic resistance	1
Contaminated / NTM	5
INH resistant - no mutation identified	72
Low level resistance - no mutation identified	47
Lost viability	6
Total	**397**

3.7 Point Mutagenesis via Homologous recombineering

Point mutagenesis served as functional validation of the effect of certain mutations on INH resistance. This was done by introducing single point mutations into a mycobacterial genome via oligonucleotide derived single-stranded DNA recombineering and mycobacteriophage-encoded proteins in a *M. smegmatis* strain carrying the replicating plasmid pJV75amber (**Figure 7**) that expresses Che9c gp61. This Che9c gp61-mediated recombination was efficient enough to introduce single base changes without direct selection, where the mutant strains were identified by PCR. Two genes implicated in INH resistance in this study were modelled in *M. smegmatis* per the method described by van Kessel et al.[19] Thus, by replacing a base pair within the gene at the specified position the adaption of *M. tuberculosis* is simulated and MIC's were done to elucidate the effect on resistance. **Figure 12** is the result of the Ziehl Neelsen staining done to confirm that the mutants generated were in fact mycobacteria. Observation of the culture under the microscope revealed acid fast bacilli consisting of straight and curved rods thus confirming that it is *M. smegmatis* and not a contaminant.

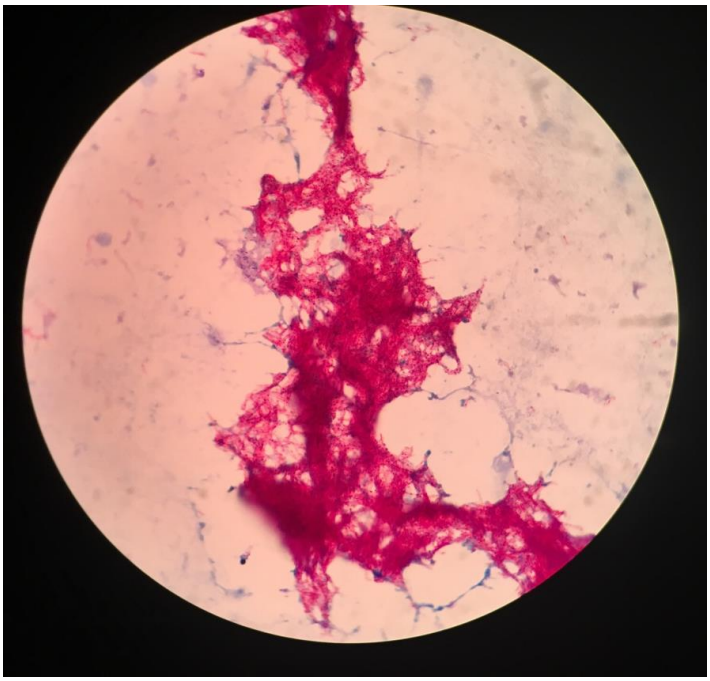


Figure 12. Ziehl Neelsen stain confirming mycobacteria

After the mutants were generated, Sanger sequencing was done in order to confirm that the SNP substitution had occurred. The ABI chromatograms represented by **Figure 13** illustrated that *ndh* colony 2 had a threonine substituted for a cytosine in codon 18 (position 53). Note, the SNP is shown as being at position 253 bp and this is due to 200 bp upstream of the position of interest being sequenced as the mutation occurs close to the start of the gene – thus ensuring the desired region would be amplified accurately. **Figure 14** & **15** present the ABI chromatograms of the respective *katG* point mutants, where colony 9 of *katG* 347 had a cytosine substituted for an alanine and *katG* 420 colony 3 contained the desired SNP of threonine substituted for a cysteine in the relevant codon. Note that **Figure 15** has the codon change represented by "Y" as the wild-type is underlying.

Colony 2

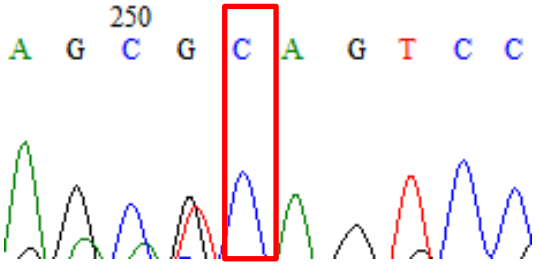


Figure 13. Colony of mutant *M. smegmatis ndh* with T-C substitution at position 53 (+200 bp upstream)

Colony 9

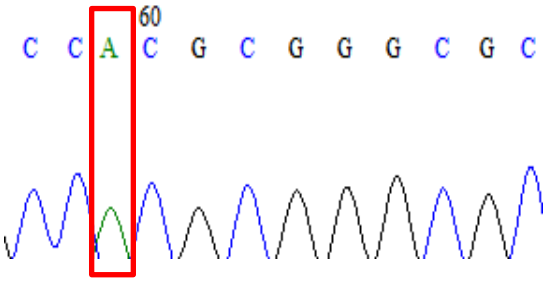


Figure 14. Colony of mutant *M. smegmatis katG* 347 with C-A substitution at position 1040

Colony 3

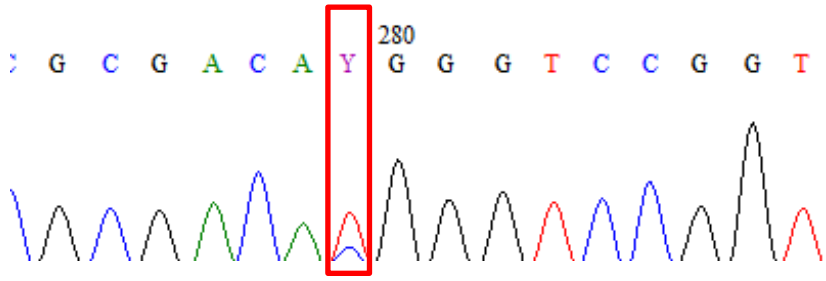


Figure 15. Colony of mutant *M. smegmatis katG* 420 with T-C substitution at position 1259

Following confirmation that the SNP was present using Sanger sequencing, MIC determination was done on the *ndh* and *katG* mutant colonies and wild type *M. smegmatis* at varying concentrations (0, 0.1, 0.5, 1, 2, 4, 8, 16, 32 & 40 µg/ml) of INH per the literature for resistance[20]. This was done in order to validate the level of resistance to INH conferred by the mutation. **Figure 16** shows growth of the wild-type at 0.1 µg/ml of INH. The radius of the colonies for MIC is depicted in centimetres (cm) and a double ended arrow (not to scale). Growth of the wild type in the presence of drug at a low concentration is within the parameters of expectation due to spontaneous mutation that occurs naturally within *M. smegmatis*. However, growth was inhibited at higher concentrations of INH.

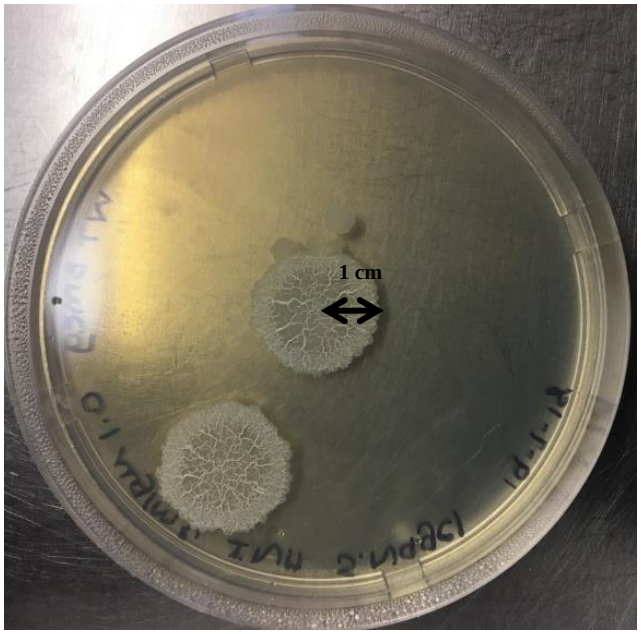


Figure 16. Wild type *M. smegmatis* colonies in the presence of 0.1 µg/ml INH.

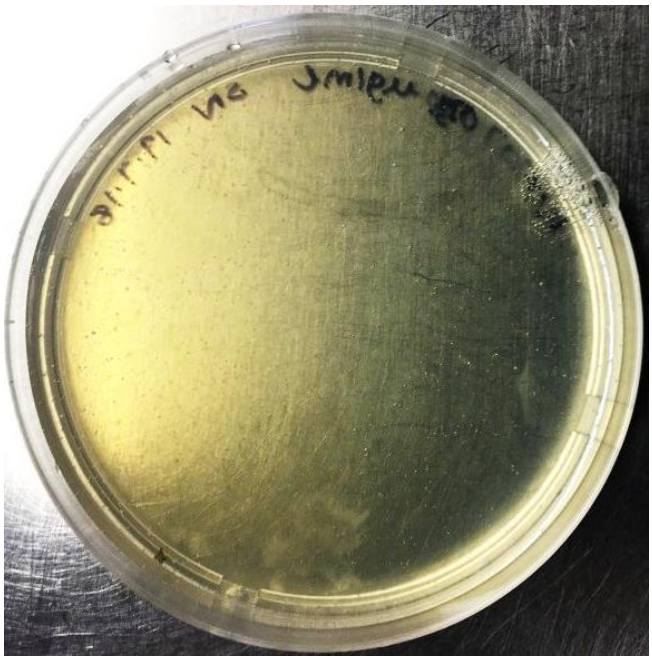


Figure 17. Wild type *M. smegmatis* in the presence of 0.5 µg/ml INH.

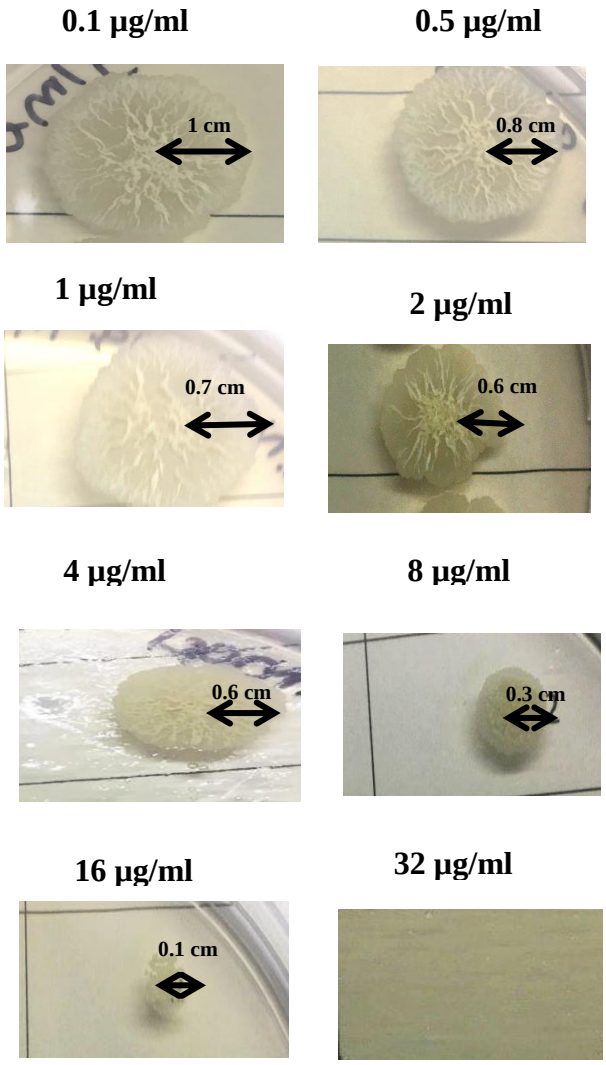


Figure 18. Minimum Inhibitory Concentration of *M. smegmatis* Colony 2 successfully transformed to have *katG* C347A mutation at varying concentrations of INH.

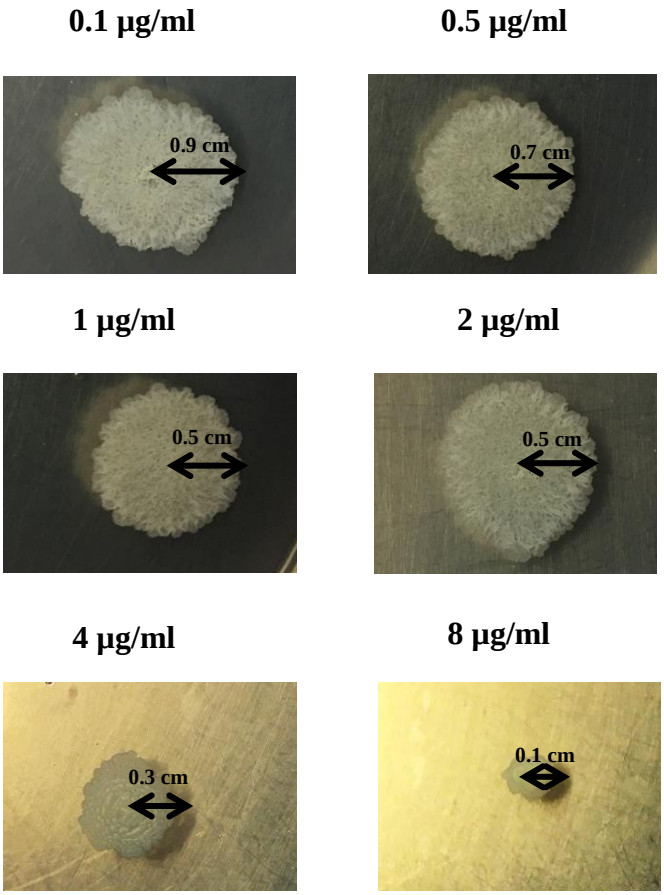


Figure 19. Minimum Inhibitory Concentration of *M. smegmatis* Colony 9 successfully transformed to have *katG* C347A mutation at varying concentrations of INH.

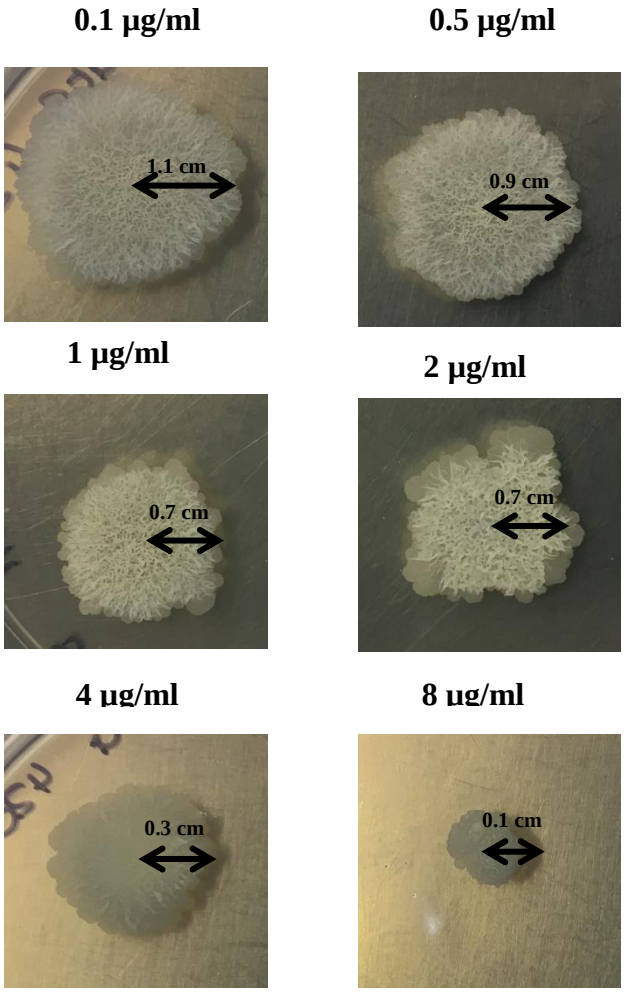


Figure 20. Minimum Inhibitory Concentration of *M. smegmatis* Colony 3 successfully transformed to have *katG* T420A mutation at varying concentrations of INH.

Low-level resistance to INH occurred in 25.6% of isolates per the MIC data suggesting that the Bangladesh regimen could be used effectively to treat these cases. Alternatively, 60.3% of isolates were diagnosed as high-level resistant, thus necessitating alternative treatment regimens that exclude INH. Genotyping via spoligotyping revealed a prevalence of the Beijing genotype in this cohort, as well as mixed infections and novel spoligotype patterns that don't resemble any known lineage or family markers. Sanger sequencing revealed canonical mutations that were missed by the LPA as well as mutations occurring elsewhere in the *katG* gene, indicating there are blind spots in the LPA that may have detrimental effects on patient response to treatment.

Further analysis of genes associated with INH resistance via WGS revealed that 50.6% of isolates had resistance causative mutations in regions outside the region of hybridisation by the LPA, as well as large deletions in *katG* that would not have been detected via the molecular method. This serves as preliminary proof of the value of WGS as a diagnostic tool. Finally, homologous recombineering showed the functional effect of the *ndh* V18A, *katG* C347A and *katG* T420A SNPs on INH resistance. MIC data generated indicated that the *ndh* V18A mutation conferred high-level resistance to INH as growth occurred up to a concentration of 16 µg/ml of INH. Furthermore, per the MIC's both the *katG* C347A and T420A SNPs conferred high-level resistance to INH, with mutant colonies showing growth on selective media up to a concentration of 8 µg/ml of INH.

REFERENCES

[1] Johnson, R.; Streicher, E. .; Louw, G. .; Warren, R. .; Van Helden, P. D.; Victor, T. C. Drug resistance in *Mycobacterium tuberculosis*. *Curr. Issues Mol. Biol* **2006**, *8*, 97–112.

[2] Muller, B.; Chihota, V. N.; Pillay, M.; Klopper, M.; Streicher, E. M.; Coetzee, G.; Trollip, A.; Hayes, C.; Bosman, M. E.; Pittius, N. C. G. Van; et al. Programmatically Selected Multidrug-Resistant Strains Drive the Emergence of Extensively Drug-Resistant Tuberculosis in South Africa. *PLoS One* **2013**, *8* (8), 1–9 DOI: 10.1371/journal.pone.0070919.

[3] Demay, C.; Liens, B.; Burguière, T.; Hill, V.; Couvin, D.; Millet, J.; Sola, C.; Zozio, T. SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect. Genet. Evol.* **2012**, *12* (4), 755–766 DOI: 10.1016/j.meegid.2012.02.004.

[4] Ando, H.; Miyoshi-Akiyama, T.; Watanabe, S.; Kirikae, T. A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2014**, *91* (3), 538–547 DOI: 10.1111/mmi.12476.

[5] Choi, Y.; Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31* (16), 2745–2747 DOI: 10.1093/bioinformatics/btv195.

[6] Rutherford, K.; Parkhill, J.; Crook, J.; Horsnell, T.; Rice, P.; Rajandream, M.-A.; Barrell, B. Artemis: sequence visualization and annotation. *Bioinformatics* **2000**, *16* (10), 944–945 DOI: 10.1093/bioinformatics/16.10.944.

[7] Coll, F. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* **2015**, *7* (1) DOI: http://dx.doi.org/10.1186/s13073-015-0164-0.

[8] Walker, T. M.; Kohl, T. A.; Omar, S. V; Hedge, J.; Del, C.; Elias, O.; Bradley, P.; Iqbal, Z.; Feuerriegel, S.; Niehaus, K. E.; et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance : a retrospective cohort study. No. appendix 1, 1193–1202 DOI: 10.1016/S1473-3099(15)00062-6.

[9] Ramaswamy, S. V; Reich, R.; Dou, S.-J.; Jasperse, L.; Pan, X.; Wanger, A.; Quitugua, T.; Graviss, E. a. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **2003**, *47* (4), 1241–1250 DOI: 10.1128/AAC.47.4.1241.

[10] Hernando Hazbon, M.; Brimacombe, M.; Del Valle, M. B.; Cavatore, M.; Guerrero, M. I.; Varma-Basil, M.; Billman-Jacobe, H.; Lavender, C.; Fyfe, J.; Garcia-Garcia, L.; et al. Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **2006**, *50* (8), 2640–2649 DOI: 10.1128/AAC.00112-06.

[11] Cardoso, R. F.; Cardoso, M. A.; Leite, C. Q. F.; Sato, D. N.; Mamizuka, E. M.; Hirata, R. D. C.; De Mello, F. F.; Hirata, M. H. Characterization of *ndh* gene of isoniazid resistant and susceptible *Mycobacterium tuberculosis* isolates from Brazil. *Mem. Inst. Oswaldo Cruz* **2007**, *102* (1), 59–61 DOI: 10.1590/S0074-02762007000100009.

[12] Ando, H.; Kitao, T.; Miyoshi-akiyama, T.; Kato, S.; Mori, T.; Kirikae, T. Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2011**, *79* (February), 1615–1628 DOI: 10.1111/j.1365-2958.2011.07547.x.

[13] Lee, A. S. G.; Lim, I. H. K.; Tang, L. L. H.; Telenti, A.; Wong, S. Y. Contribution of *kasA* Analysis to Detection of Isoniazid-Resistant *Mycobacterium tuberculosis* in Singapore. *Antimicrob Agents Chemother.* **1999**, *43* (8), 12–15.

[14] Zhang, M.; Yue, J.; Yang, Y.; Zhang, H.; Lei, J.; Jin, R.; Zhang, X.; Wang, H. Detection of Mutations Associated with Isoniazid Resistance in *Mycobacterium tuberculosis* Isolates from China. *J. Clin. Microbiol.* **2005**, *43* (11), 1–7 DOI: 10.1128/JCM.43.11.5477.

[15] Nebenzahl-Guimaraes, H.; Jacobson, K. R.; Farhat, M. R.; Murray, M. B. Systematic review of allelic exchange experiments aimed at identifying mutations that confer drug resistance in *Mycobacterium tuberculosis*. *J. Antimicrob. Chemother.* **2014**, *69* (2), 331–342 DOI: 10.1093/jac/dkt358.

[16] Projahn, M.; Köser, C. U.; Homolka, S.; Summers, D. K.; Archer, J. A. C.; Niemann, S. Polymorphisms in isoniazid and prothionamide resistance genes of the *Mycobacterium tuberculosis* complex. *Antimicrob. Agents Chemother.* **2011**, *55* (9), 4408–4411 DOI: 10.1128/AAC.00555-11.

[17] Ioerger, T. R.; O'Malley, T.; Liao, R.; Guinn, K. M.; Hickey, M. J.; Mohaideen, N.; Murphy, K. C.; Boshoff, H. I. M.; Mizrahi, V.; Rubin, E. J.; et al. Identification of New Drug Targets and Resistance Mechanisms in *Mycobacterium tuberculosis*. *PLoS One* **2013**, *8* (9), 1–13 DOI: 10.1371/journal.pone.0075245.

[18] Gagneux, S.; Burgos, M. V.; DeRiemer, K.; Enciso, A.; Munoz, S.; Hopewell, P. C.; Small, P. M.; Pym, A. S. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog.* **2006**, *2* (6), 603–610 DOI: 10.1371/journal.ppat.0020061.

[19] van Kessel, J. C. Van; Hatfull, G. F. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering : characterization of antimycobacterial drug targets. **2008**, *67* (January), 1094–1107 DOI: 10.1111/j.1365-2958.2008.06109.x.

[20] Xu, X.; Vilchèze, C.; Av-Gay, Y.; Gómez-Velasco, A.; Jacobs, W. R. Precise null deletion mutations of the mycothiol synthesis genes reveal their role in isoniazid and ethionamide resistance in *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* **2011**, *55* (7), 3133–3139 DOI: 10.1128/AAC.00020-11.

Chapter 5: Discussion

In Routine laboratories in South Africa, INH resistance in *M. tuberculosis* isolates are determined by means of the molecular method, Hain GenoType MTBDR*plus*. However, if this method predicted INH susceptibility and the isolate is RIF resistant, a second isolate will be cultured and conventional DST will be performed for INH. In this study, 397 isolates were identified over a 2 year period by the regional NHLS laboratory in Port Elizabeth, to be scored by the LPA as susceptible to INH, but found to be resistant by conventional DST on the MGIT 960 system. Thus, we aimed to elucidate the reasons for the discrepancies between the 2 methods that are standardised and routinely used. The reason for discrepancies between genotypic and phenotypic INH resistance were subsequently tested via MIC, spoligotyping and Sanger sequencing and unknown mechanisms of resistance to INH were identified via whole genome sequencing.

5.1 Rationalisation for discrepancy: level of drug resistance of non-canonical SNPs

Per our MIC determination, 10.6% of the clinical isolates were susceptible to INH. Thus these isolates were incorrectly diagnosed as INH resistant by the routine DST. Arguably, a potential reason for this could be the level of resistance to INH of the isolates. If the breakpoint is close to the critical concentration at which it is tested, false results may be generated. Alternatively, subculture of the original sample may have resulted in heteroresistance or mixed infection being selected out as this is the initial culturing step before INH is added, in favour of the susceptible form. Lastly, lab error may have accounted for incorrect recording of results. Thus, complexities of the technique may result in inaccurate results, which could potentiate incorrect diagnosis of drug-resistant tuberculosis in patients. The current regimen of extended combined drug therapy for INH susceptible consists of rifampicin, isoniazid, ethambutol and pyrazinamide [1]. However, diagnosis of INH resistance may lead to the exclusion of INH in the regimen. As INH serves as the backbone of treatment for drug sensitive tuberculosis infections due to its bactericidal effects against actively replicating bacilli [2], its exclusion may be detrimental to patient outcomes.

Additionally, 116 (29.2%) were found to be low-level resistant to INH per the MIC data. Low-level resistance to INH has been associated with failure of tuberculosis therapy or relapse [3], as antibiotics will only act effectively if concentrations are above the MIC for *M. tuberculosis* and are adequately maintained (adequate area under the curve or "AUC"). For TB treatment, resistance to INH is commonly defined by an MIC of >1 µg/ml. In the South African context, the MICs of INH for approximately 50% of the INH resistant organisms studied were in the low-level resistance category, where low-level resistance to INH is characterised by an MIC between 0.1-1.0 µg/ml [4]. A study done on animal models suggests that high doses of INH may be a viable treatment option for cases of low-level INH resistance in patients[5]. Van Deun et al. 2010 recommend high-dose INH as a supplement to 9 months of treatment with gatifloxacin, clofazimine, ethambutol, and pyrazinamide for treatment of low-level INH resistant cases [6]. It was shown that a regimen lacking INH results in much lower effectiveness. This suggests that high-dose INH plays a vital role in treatment of MDR-TB [6]. We suggest that targeted treatment would yield even better outcomes. A similar conclusion was reached by an independent study [7], i.e. that high dose INH can be effective in the presence of the *inhA* gene mutations.

Furthermore, we suggest that novel SNPs were observed in this study in low-level resistant isolates in genetic regions outside the hybridised regions of the LPA, viz.*inhA* promoter and *katG*. These polymorphisms, namely: *Rv1592c* I322V, *fabD* P270L, *fabG1* L230L *and ndh* V18A would subsequently be missed by the LPA due to the selection bias in the molecular test. This may incorrectly exclude patients who may have benefited from the high dose regimen and unfavourably affect outcomes.

The disregard for personalised dosage and diagnosis of low-level resistance may have severe consequences such as patients not responding to treatment optimally, particularly during the 4-month extension phase. Additionally, there is an increased risk for patient development of antibiotic resistance. Our own studies have shown that low-level INH resistance, as defined by mutations in the *inhA* promoter, are more common than mutations in *katG* in certain South African settings[8],[9]. This data suggests that such mutations conferring low-level resistance (and simultaneously including resistance to ETH) are a

gateway to the development of further resistance, including high-level INH resistance, encoded by "classical" *katG* mutations that confer high-level resistance to INH [10].

A high proportion of the isolates (61%, n=242) from this study showed resistance to INH at varying degrees higher than 1 µg/ml. Growth of 155 isolates occurred at concentrations ≥10 µg/ml INH. This effectively rules out INH as a treatment option as effective MIC cannot be clinically attained and alternative treatment regimens must be considered to treat these patients.

In summary, INH is likely a viable treatment option at high dosages for isolates that were resistant to INH at a low-level (<1ug/ml). A specific mutation, namely the *inhA* promotor, has been linked to lower levels of resistance in previous studies [11]. For 9.3% of the isolates, no mutations were identified in the classical loci conferring resistance to INH included in the LPA, and these isolates were all found to be susceptible to INH (MIC<0.1 µg/ml). Additionally, for 4.5% of the isolates we identified a mutation at the loci represented on the LPA and these were found to be resistant with the MIC testing (>0.1 µg/ml).

5.2 Rationalisation for discrepancy: mixed infections/heteroresistance

Another possible reason for the discrepancies between the two methods used by the routine laboratory could be due to INH resistant and susceptible bacteria being present in one isolate. This could be caused by two known ways. Firstly due to mixed strain infections, where the patient is infected by 2 different strains, a dominant susceptible strain and an underlying resistant strain[12]Secondly, when the bacteria are in the process of acquiring drug resistance, the population present in the isolates could have the wild-type population with the mutant population underlying, termed heteroresistance [13]. By the combination of Sanger sequencing and spoligotyping of the original cultures and the samples from the MIC determination, we were able to show the reason for the discrepancies were either mixed infection (n=8) or heteroresistance (n=11) in 4.8% of all the isolates. Sanger sequencing allowed for detection of underlying populations based on the ABI chromatograms. Simultaneously, Sanger sequencing was done to enable identification of

possible mixed infections and to identify heteroresistance, when the *M. tuberculosis* strain is in the process of acquiring INH resistance, but would be missed by the LPA. .

The LPA is not consequent in cases of mixed infection or heteroresistance, as the detection of mutations in the canonical regions is dependent on what population of *M. tuberculosis* (resistant or susceptible) is present in the greatest concentration. Thus, if the susceptible strain is dominant, the isolate will be misdiagnosed as susceptible to INH and the regimen prescribed will include INH to the detriment of the patient as it my perpetuate progression of resistance[14]. This hurdle may be overcome via culturing on selective media however, the DST is a secondary technique solely employed in cases of suspected drug resistance. Recommendation for culturing of all *M. tuberculosis* isolates (regardless of susceptibility or resistance) is logical but not necessarily viable due to the labour intensity and cost, which the molecular technique is intended to circumvent. Thus, cases of heteroresistance and mixed infection significantly complicate diagnoses as the primary technique employed for detection of INH resistance is not reliable.

5.3 Rationalisation for discrepancy: BCG/ *M. bovis*

We also identified one spoligotype pattern that indicates *M. bovis*/ BCG. A possible reason for it being missed by the LPA is that that the mutation that results in INH resistance is outside the region hybridised by the LPA. For instance, a study by Yang et al. in 2015 found that the *InbR* gene, which regulates the expression of multiple genes as well as the *iniBAC* operon, serves as a repressor. This hypothesis was proven when overexpression of InbR protein decreased bacterial susceptibility to INH in *M. bovis/* BCG[15]. Thus, an alternative mechanism of resistance (not caused by canonical mutations) would result in misdiagnosis of drug resistant mycobacteria as susceptible.

5.4 Mechanisms of INH resistance: *katG*

By Sanger sequencing of the areas surrounding the mutations that are part of the LPA, we were able to identify additional mutations that are associated with resistance to INH. The specific genes and regions of *M. tuberculosis* that were amplified via PCR and Sanger sequenced are the *katG* gene and *inhA* promoter.

5.4.1 Resistance affiliated with *katG* mutations

The *katG* gene was selected for investigation for a number of reasons. Mutations in the *katG* gene are the most common mechanisms associated with INH resistance in *M. tuberculosis* [16], categorically associated with conferring medium to high-level resistance (minimum inhibitory concentration of > 1mg/ml) to INH [11],[17]. Thus, *katG* mutations are significant when inferring the mechanism of INH resistance.

Mutations or deletions within the *katG* gene prevent the activation of INH upon its diffusion into *M. tuberculosis*, rendering INH ineffective against *M. tuberculosis* bacilli and resulting in high INH resistance levels [2],[11]. The catalase-peroxidase enzyme encoded for by the *katG* gene is often lost in highly resistant strains, however the catalase ability may be sustained in cases of low-level resistance caused by some *katG* mutations [2],[18],[19].

The *katG* S315T SNP has been strongly linked to high-level resistance to INH in the literature. According to Jeeves et al., 2015 this SNP is present in 50–90% of all isoniazid-resistant clinical isolates[16]. This information is important for diagnostic and treatment purposes because it has been shown that initial mono-resistance to INH, specifically *katG* S315T mutations, often precede multi-drug resistance in *M. tuberculosis* [20].

Thus, due to its significance, the S315T mutation is screened for in the LPA and should theoretically have been diagnosed as INH resistant if the SNP was present. However, 12 isolates in this study had the canonical S315T mutation and were misdiagnosed as INH susceptible by the MTBDR*plus* LPA. We theorise that this is due to human error (perhaps subjective interpretation of results obtained) or possibly mixed infection with an underlying population of susceptible *M. tuberculosis* of the same genotype.

Despite the significance of the canonical *katG* mutation, due to its prevalence and link to high-level resistance, mutations elsewhere in this gene may be causative of INH

resistance[21] by increased susceptibility to oxidative stress, consequently resulting in adaption by upregulating the synthesis of a number of compounds involved in (i) increased uptake and use of alkanes and fatty acids as a source of carbon and energy and (ii) the synthesis of a number of compounds directly involved in reducing oxidative stress (REF). As indicated in **Table 14** and **16,** various SNP's were identified elsewhere in this gene.

Of particular interest in **Table 14** is the *katG* G285D mutation, as it was identified in 18 isolates, all of the Beijing lineage, and all low-level resistant to INH. As this SNP was only present in low-level resistant isolates, it is a good candidate for the justification of the high dose INH regimen. Kandler et al., 2018 found a missense mutation at this position in the gene and suggested its association with INH resistance based on the current literature[20],[22],[23], as well as confirming the rare *katG* W300G (**Table 14**) mutation has been experimentally associated with INH resistance in *M. tuberculosis* via functional genetics[22],[24].

The presence of *katG* SNP's in low-level resistant isolates may be indicative of flaws in the premise of *katG* mutations being conferring medium to high level resistance to INH. There were other substitutions uncovered via Sanger sequencing in the *katG* gene in low-level resistant isolates, however at lower proportions than *katG* G285D . These include an isolate with a SNP in *katG* S325P (**Table 14**), as well as single Beijing isolates with *katG* Q439P, G406R and F594L respectively (**Table 16**). A proportion (n=14) of high-level resistant isolates had a SNP in *katG* D419Y and as they were all part of the S family, it may be specific to this lineage, indicating possible transmission of an MDR-TB strain with high level of INH resistance.

Furthermore, the results in **Table 14** show isolates with double mutations in genes that may have a role in the mechanism of resistance. Torres et al., 2015 reported novel *katG* mutations, however that occurred either in combination with *katG* S315T or with -15C-T *inhA* promoter mutation [19]. Thus, it is not uncommon for combination mutations to be associated with INH resistance in *M. tuberculosis.* However, two isolates of unknown lineage had *katG* G279D as well as *inhA* -77 G-A SNP's. These occurred in high-level resistant isolates and have not previously been reported in the current literature. Consequently, this may be a novel mechanism of resistance where adaption of the bacterium was causative of resistance at minimal fitness cost. Due to limited information

available, it is difficult to infer with certainty the mechanistic role played by these mutations, but it serves as evidence of the complexities associated with drug resistance diagnoses as it is uncertain which SNP is causative of resistance or whether they both contribute. Future studies will investigate the effect of these combination mutations functionally.

In summary, mutations detected in isolates with low-level resistance included *inhA* promoter -15 C-T and *katG* G285D, S325P, Q439P, G406R and F594L. Patients infected with a strain of *M. tuberculosis* containing these mutations serve as ideal candidates for inclusion of high dose INH within a treatment regimen, consequently limiting cases where first-line therapy is not prescribed.

5.4.2 Resistance affiliated with *katG* deletions

This body of work has established the role of mutations or deletions of the mycobacterial *katG* gene in INH resistance. Pretorius et al., 1995 stated that *katG* deletions are rare and not significant to resistance to INH[25], however our data suggests otherwise. Deletions of varying sizes in the *katG* genetic region are highlighted in **Table 17.** The large deletions identified are arguably the most significant finding of this study, were the two isolates with 19 kbp are both of the X3 strain family. Although we can conclude with reasonable certainty that these deletions occur on two separate occasions due to the variation in size, there is insufficient data to draw irrefutable parallels with the phenomenon of large deletions and lineage.

Three of the deletions span large areas (14 and 19 kbp) and various different genes are lost (see **Figure 9**). These deletions bring into question which compensatory mechanisms are at play, and future studies will confirm the presence of these deletions and functionally confirm their role in resistance and possibly fitness and virulence. The *furA* and *katG* promoters are deleted in both samples and thus further investigation must be done to find out how transcription of the rest of the genes downstream of the deletions are affected. It has been demonstrated in the literature that a deletion preceding the initiation codon of *katG* (thus including *furA*) was causative of high-level resistance to INH[26]. Per that publication, *furA* is essential for full catalase activity and INH susceptibility in *M. tuberculosis* and the deletion was responsible for the reduction of catalase-peroxidase

activity in the isolate, thereby resulting in resistance to INH[26]. Tsolaki et al., 2004 reported a similar case, where the entire *katG* gene was deleted, thereby conferring INH resistance[27]. The other four isolates with partially deleted *katG* genes are also significant, as codon 315 of *katG* is still intact and thus have resulted in a susceptible result with LPA.

The *katG* promoter -11 T-G mutation in **Table 17** is novel, as it has not been described in the literature to the best of our knowledge. We hypothesise that this SNP negatively affects transcription of the *katG* gene, thereby preventing InhA from binding and interrupting the INH mechanism of action and resulting in high-level resistance to INH in *M. tuberculosis*[28]. This is significant as the bacteria can survive and be virulent despite the defective promoter and thus suggesting compensatory mechanisms have evolved.

The mutations in the *katG* start codon as observed in two isolates are also interesting, because it affects the recognition site of the start of the gene. The *katG* V1A mutation has previously been reported in INH high-level resistant isolates. Kandler et al., 2018 confirmed the role of this SNP in INH resistance via functional genetics (point mutagenesis by homologous recombineering) and stated that resistance was expected given that the G**T**G-G**C**G mutation meant that codon one would no longer be recognized as a start codon[22].

WGS revealed multiple mutations in a single isolate. However, unlike in section **5.4.1,** the combination SNP's summarised in **Table 17** occurred in a single gene, namely the *katG* gene. Some combination mutations, such as *katG* Y113C + *katG* promoter -9 A-C, occurred in low and high-level INH resistant isolates. Alternatively, the combination mutations *katG* I71T + W91R + G121R + S140G and *katG* G466R + F594L occur solely in low-level resistant isolates. A potential reason for this may be that multiple mutations in a single gene could have a cascade effect on fitness cost, resulting in lower resistance. This reasoning is also applicable in cases of non-canonical mutations in alternative genes with known INH drug resistance associations.

Coll., 2015[29] and Walker et al., 2015[30] serve as the current reference standards for SNPs that have defined roles in INH resistance. Subsequently, mutations such as *katG* W91R, A93T and D94G identified via WGS are considered novel mechanisms of resistance, as they have not yet been described in the literature to our knowledge. The *katG* 104 SNP had two variations in this cohort, where R104Q occurred in LAM3/S convergent high-level

resistant isolate and R104W occurred in a LAM3 high-level resistant strain. This establishes a tenuous link between this mutation in this codon and lineage.

There were also two different versions of the *katG* 121 SNP in **Table 17**. The *katG* G121V mutation resulted in high-level resistance in an isolate of Beijing lineage, whereas *katG* G121R resulted in low-level resistance in an isolate belonging to the same lineage. Other novel SNP's conferring high-level resistance to INH include *katG* N138S, D142A, A144V, A172T, F185L, D189H, R203L, D419Y, G421S, R463L, A479P, F483S, R496G, D695G and Y711D. Potential candidates for the high dose INH regimen include low-level resistant isolates namely: *katG* S140G, D189G, T380I, T394M and M420T.

5.5 Mechanisms of INH resistance: *inhA* promoter

As stated in previous sections, mutations or polymorphisms in the *inhA* promoter are the second most common explanation for INH resistance. The mutations detected included the *fabG1* gene as serves as an alternative promoter for *inhA*, thereby increasing expression level and cementing its relevance for next generation molecular diagnostics[19],[22],[31]. More specifically, investigation of INH resistant isolates showed that a -15 mutation in the *inhA* promoter was on average present in 19.5% of isolates from 49 different countries [23]. Low-level resistant Beijing isolates with this mutation can be observed in **Table 15**. Moreover, this mutation was seen in as many as 43% of INH resistant cases in specific geographical locations, thus its link to resistance is well established [23] and strains with these mutations spread more rapidly than strains with other mutations [8],[32]. Resistance caused by *inhA* promoter mutations is typically low-level (MIC of less than 1 µg/ml) [11].

However, the data in **Table 15** categorises the non-canonical *inhA* promoter mutations detected in this study via Sanger sequencing as being present in high-level resistant isolates. A possible reason for the higher level of resistance is mixed infections, where the low-level resistant population may be present at lower proportions and thus not detected. Additionally, MIC would select for the high-level resistant population, as the lower resistant population would not grow at incrementally higher concentrations of INH.

5.4.2 Mechanisms of INH resistance: SNPs in other genes associated with INH resistance

Other genes (excluding *katG* and *inhA* promoter) also have also been reported to be associated with resistance to the first-line antibiotic INH (REFs). Some SNPs such as *Rv1592c* I322V, *fabG1* L230L, *inhA* I194T and *ndh* V18A identified by WGS, were present in low and high-level INH resistant isolates. Thus, the only candidates for the Bangladesh INH high-dose regimen in **Table 18** are the strains harbouring *fabD* P270L and *inhA* S94A.

The role of these alternative genes in INH resistance vary. For instance, SNP's in the *Rv1592c*, *fadE24, iniA, iniC* and *srmR* genes have been implicated in INH resistance, but the functional mechanism remains speculative/unknown and the level of resistance is inconclusive[23],[33]. However, the mechanism of action is well defined in *fabD* as it a malonyl CoA:ACP transacylase[34] as well as in the *fabG1-inhA* regulatory region, which serves as an alternative promoter for *inhA*[19].

Mutations in the *inhA* open reading frame of *M. tuberculosis* are not as common as in *katG* or *inhA* promoter, but still occur often enough to be of interest [18]. The S94A mutation in *inhA* decreases INH-NAD Adduct affinity for the NADH-binding site of InhA, subsequently resulting in increased INH as well as ethionamide resistance [35]. Other mutations in the *inhA* gene that have been associated with INH resistance encompass mutations at codons I21T/V, A190S, I194T and I258T/V [23],[35]. As a result of the different mutations in this region that have been associated with INH resistance, as well as the prevalence of mutations in this region being associated with low-level INH resistance, it can be assumed that the *inhA* gene coding region is significant to INH action and resistance.

The *accD6* D229G mutation in **Table 19** has been shown to play a role in the building blocks for de novo fatty acid biosynthesis in the literature[19]. Similarly, the *fbpC* gene has a role in mycolyltransferase activity that is vital for the formation of cord factors, and if disrupted with mutations affects the INH mechanism of action as INH targets the type 2 fatty acid synthase (FAS II) system in *M. tuberculosis*. Thus, by preventing the synthesis of mycolic acids, the *fbpC* SNP is directly affecting the target of INH.

Finally, the *ahpC-oxyR* intergenic region is thought to be responsible for the regulation of *katG* expression[21]. Both the *ahpC* and *oxyR* genes are up regulated in *M. tuberculosis* under conditions of oxidative stress. The promoter region of *ahpC* acquires mutations when *katG* is absent or mutated and there is no catalase or peroxidase activity[2]. These polymorphisms result in the production of an alkylhydroperoxide reductase enzyme that performs the same function to compensate for the absence of catalase peroxidase[2] The *oxyR* gene is also upregulated either to combat peroxidation in the absence of *katG* or to activate *katG* expression[36],[37].

Although we were able to identify the possible mechanism of INH resistance in most of the isolates through WGS, there were still 9 isolates were no mutations detected by USAP, Artemis or the variant files in any of the known genes that could have resulted in INH resistance. Thus, the cause of resistance for these isolates remains unknown.

It is concerning that highly resistant cases would have been missed entirely if the LPA was the sole method of resistance testing. These INH resistant cases were only discovered due to secondary DST resistance testing. These culturing methods are not a solution to the problem, as they are time consuming, which is problematic if the patient is on inadequate treatment and is potentially spreading the disease further in the community.

Furthermore, it has been observed in literature [13], as well as in the results of this study, that the LPA is not 100% effective in detecting mutations in the regions it does hybridize to. It is also not an effective means of resolving/diagnosing resistance when heteroresistance or mixed infections occur. Mixed infections require additional testing beyond just the LPA to be detected, as only the population present in the highest proportion would be detected by the LPA. The patient would be incorrectly diagnosed as having a drug susceptible strain of INH and the incorrect drug treatment regimen would be prescribed.

It is hypothesized that the usually uncommon mutations observed in this study are observed more frequently for one of two reasons (or a combination): The first reason is that they occur due to an outbreak in this region. However, this only applies to isolates in this study that have mutations that occur in multiple isolates from the same strain lineage. The second reason is that there are systems in place to test for and treat well-known mutations such as the *katG* 315 mutation that is readily detected by the LPA in most

instances. Thus, the common mutations are treated, giving the rarer mutations the opportunity to prevail.

Since a number of mutations in other regions of *katG* (not surrounding codon 315) were detected, sequencing the rest of the *katG* gene may be useful as a first line approach to identify uncommon mutations causing INH resistance.

Additionally, a significant amount of isolates have been identified to be resistant to INH but not have the classical mutations tested for by molecular test such as the LPA. Thus, the current drug resistance testing algorithm used is logical. The data suggests that it is not sufficient to employ the LPA as the sole resistance testing mechanism in all cases that are not diagnosed as rifampicin mono-resistant. Instead, it serves as evidence that the secondary phenotypic DST may have to be employed in all cases, whether the isolates are found to be susceptible or resistant to INH, to ensure that no incorrect diagnoses are made. Although, it should be noted, that this is time consuming and expensive and should thus not be the sole solution to the current problem of discrepancies in detection of resistance causing mutations for routine laboratories. Our recommendation is that WGS could be implemented in order to minimize the chances of missing non-canonical INH resistance mutations outside targeted hotspots and identify acquisition of resistance (heteroresistance) earlier. Future studies should obtain WGS data for all isolates in this study in order to ensure that no novel SNPs go undiagnosed, and to ensure there is no bias.

5.4.3 Confirmation of INH resistance mechanisms by Point Mutagenesis in *M. smegmatis* via homologous recombineering

Single point mutations were generated in suspected drug targets of a wild-type *M. smegmatis* model using oligonucleotides. *M. smegmatis* is well established as a model for *M. tuberculosis*[38] due to its non-virulence and fast-growing status. This technique has been adapted from van Kessel et al., 2008 and results in high recombination frequency in *M. tuberculosis* and *M. smegmatis*[38] when an oligonucleotide binds to the lagging strand of single stranded DNA. The purpose of this was to investigate the role of individual SNPs identified by WGS in this study role in INH resistance. The three single point mutations sucessfully generated were *ndh* V18A, *katG* C347A and *katG* T420C. The presence of the polymorphism was confirmed via Sanger sequencing in each and the effect was measured via MIC.

Cardoso et al., 2007, hypothesised that a mutation in *ndh* V18A could be associated with resistance as a mutation in this gene could result in an alteration of the NADH/NAD+ ratio[39]. This prevents the peroxidation reactions required for the activation of INH, as well as the displacement of the NADH-isonicotinic acyl complex from InhA enzyme binding site. This study found a V18A mutation in five INH high-level resistant isolates and one low-level, thus establishing a tenuous link to resistance. This mutation was also encountered in both resistant and susceptible strains by Hernando Hazbon et al., 2006. Thus, this SNP was selected, to confirm or refute its role in INH resistance. Additionally, a search using the program PROVEAN revealed that the influence of this SNP on the protein is deleterious and thus affects the functioning of the gene.

MSMEG_0306[40] was selected as the reference sequence due to its high genetic homology (93% query cover per NCBI BLAST[41]) to the *ndh* (Rv1854c) gene of *M. tuberculosis*. This was done with the with the purpose of inferring the effect of resistance from the model. The four colonies containing the SNP were cultured in increasing concentrations of INH and the MIC for INH in the *ndh* mutants was > 8 µg/ml after three days of incubation. As INH resistance is categorised as "high-level" at minimum inhibitory concentration (MIC) of > 1 µg/ml INH, the results obtained indicate that the *ndh* V18A polymorphism confers high-level resistance to INH in *M. smegmatis*.

Selecting an orthologue for the *katG* gene in *M. smegmatis* was more complex than for *ndh*, as there are three possible options. These are MSMEG_3729, MSMEG_3461 and MSMEG_6384. However, MSMEG_6384 was found to have the greatest sequence homology to *katG* (Rv1908c) of *M. tuberculosis* with a query cover of 95%[41]. The justification for the investigation of uncharacterised *katG* mutations is that the LPA MTBDR*plus* screens for mutations in regions of *katG* and *inhA* promoter which are known to be associated with INH resistance[3]. The LPA tests specifically for mutations in *katG* 315 (S315T1 and S315T2) as these mutations are most commonly associated with resistance to INH[2], as well as -15 (C15T), -16 (A16G) and -8 (T8C/T8A) mutations in the *inhA* promoter (Albert et al., 2010). Thus, any mutations that cause INH resistance, but occur outside these specific regions (such as the *katG* mutations in this study) will be missed by the LPA, thereby misdiagnosing the patient as INH susceptible. Therefore, it is vital to ensure that all mutations that are suspected to confer INH resistance are confirmed, with the purpose of improving current diagnostic methods.

The point mutation in *katG* C347A was selected for genetic functional validation as it has not been associated with INH resistance in the literature[29],[30], thus it may be a novel mechanism of resistance. A single high-level resistant isolate containing this SNP was identified in the subset via Sanger sequencing (see **Figure 10**). Functional genetics confirmed this finding as the MIC for the single mutant colony was < 8 µg/ml of INH as seen in **Figure 25**, indicating that the base pair change confers high-level resistance in the model. The *katG* T420C mutation was encountered in one low-level and one high-level resistant isolate in our dataset as seen in **Figure 10**. A clear link to resistance has not been established per the literature[29],[30]. The results in **Figure 26** indicate that mutant colony 3 has an MIC of > 8 µg/ml of INH and confers high-level resistance. The MIC's obtained for each point mutation is vital, as knowledge on SNPs and the level of resistance they cause is limited[42]. The variation in growth shows that different resistance mutations lead to specific MIC's, where those conferring low-level resistance may still be treated by increased dosing.

However, despite the results of this experiment showing promise, there are certain limitations to the homologous recombineering experiment. The pJV75amber plasmid used as a vector could potentially initiate other undesireable recombination events which may have an influence on INH resistance. Additionally, the use of *M. smegmatis* allows inference as to the potential effects of the point mutations in *M. tuberculosis* but are not entirely conclusive.

Future studies will aim to overcome these hurdles by performing homologous recombineering in H37Rv and sequencing the entire genome using WGS. Then the genome of wild-type could be compared to the recombineered H37Rv in order to confirm there are no alternative mechanisms that are causative of resistance. Alternatively, further possible recombineering events caused by the *Che9c 61* gene can be eliminated via curing of the pJV75amber plasmid by several passages on media that doesn't contain the selective drugs (Kanamycin and Hygromycin). If the plasmid is successfully removed and the isolate remains resistant to INH then it is likely that an alternative mechanism of resistance has occurred.

REFERENCES

[1] WHO. Treatment of tuberculosis: guidelines for national programmes http://www.who.int/tb/features_archive/new_treatment_guidelines_may2010/en/ (accessed Jun 8, 2016).

[2] Zhang, Y.; Yew, W. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* **2009**, *13* (11), 1320–1330 DOI: 10.2741/1289.

[3] Hughes, J.; Osman, M. Diagnosis and management of drug-resistant tuberculosis in South African adults. *South African Med. J.* **2014**, *104* (12), 894 DOI: 10.7196/samj.9097.

[4] Parkin, D. P.; Vandenplas, S.; Botha, F. J. H.; Vandenplas, M. L.; Seifart, H.; Helden, P. D. Van; Walt, B. J. Van Der; Donald, P. R.; Jaarsveld, P. P. Van. Trimodality of Isoniazid Elimination Phenotype and Genotype in Patients with Tuberculosis. *Am J Respir Crit Care Med* **1997**, *155*, 1717–1722.

[5] Moodley, C.; Ammerman, N.; Ncgobo, B.; Swanson, R.; Dorasamy, A.; Moodley, S.; Mgaga, Z.; Pillay, L.; Bester, L.; Singh, S.; et al. Are high doses of isoniazid beneficial in treatment of infection with isoniazid-resistant *M. tuberculosis*? *Kwazulu-Natal Res. Inst. Tuberc. HIV* **2014**, 1–5.

[6] Van Deun, A.; Maug, A. K. J.; Salim, M. A. H.; Das, P. K.; Sarker, M. R.; Daru, P.; Rieder, H. L. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *Am. J. Respir. Crit. Care Med.* **2010**, *182* (5), 684–692 DOI: 10.1164/rccm.201001-0077OC.

[7] Sotgiu, G.; Tiberi, S.; Centis, R.; D'Ambrosio, L.; Fuentes, Z.; Zumla, A.; Migliori, G. B. Applicability of the shorter "Bangladesh regimen" in high multidrug-resistant tuberculosis settings. *Int. J. Infect. Dis.* **2017**, *56*, 190–193 DOI: 10.1016/j.ijid.2016.10.021.

[8] Müller, B.; Streicher, E. M.; Hoek, K. G. P.; Tait, M.; Trollip, A.; Bosman, M. E.; Coetzee, G. J.; Hoosain, E.; Pittius, N. C. G. Van; Victor, T. C.; et al. *inhA* promoter mutations : a gateway to extensively drug-resistant tuberculosis in South Africa ? *Int J Tuberc Lung Dis* **2011**, *15* (March 2010), 344–351.

[9] Warren, R. M.; Streicher, E. M.; Gey van Pittius, N. C.; Marais, B. J.; van der Spuy, G. D.; Victor, T. C.; Sirgel, F.; Donald, P. R.; van Helden, P. D. The clinical relevance of Mycobacterial pharmacogenetics. *Tuberculosis*. 2009, pp 199–202.

[10] Chen, P.; Bishai, W. R. Novel Selection for Isoniazid (INH) Resistance Genes Supports a Role for NAD -Binding Proteins in Mycobacterial INH Resistance. *Infect. Immun.* **1998**, *66* (11), 5099–5106.

[11] Ando, H.; Miyoshi-Akiyama, T.; Watanabe, S.; Kirikae, T. A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2014**, *91* (3), 538–547 DOI: 10.1111/mmi.12476.

[12] Van Rie, A.; Victor, T. C.; Richardson, M.; Johnson, R.; Van Der Spuy, G. D.; Murray, E. J.; Beyers, N.; Van Pittius, N. C. G.; Van Helden, P. D.; Warren, R. M. Reinfection and mixed infection cause changing *Mycobacterium tuberculosis* drug-resistance patterns. *Am. J. Respir. Crit. Care Med.* **2005**, *172* (5), 636–642 DOI: 10.1164/rccm.200503-449OC.

[13] Streicher, E. M.; Bergval, I.; Dheda, K.; Böttger, E. C.; Pittius, N. C. G. van; Bosman, M.; Coetzee, G.; Anthony, R. M.; Helden, P. D. van; Victor, T. C.; et al. Mycobacterium tuberculosis population structure determines the outcome of genetics-based second-line drug resistance testing. *Antimicrob. Agents Chemother.* **2012**, *56* (5), 2420–2427.

[14] Muller, B.; Chihota, V. N.; Pillay, M.; Klopper, M.; Streicher, E. M.; Coetzee, G.; Trollip, A.; Hayes, C.; Bosman, M. E.; Pittius, N. C. G. Van; et al. Programmatically Selected Multidrug-Resistant Strains Drive the Emergence of Extensively Drug-Resistant Tuberculosis in South Africa. *PLoS One* **2013**, *8* (8), 1–9 DOI: 10.1371/journal.pone.0070919.

[15] Yang, M.; Gao, C. H.; Hu, J.; Zhao, L.; Huang, Q.; He, Z. G. InbR, a TetR family regulator, binds with isoniazid and influences multidrug resistance in *Mycobacterium bovis* BCG. *Sci. Rep.* **2015**, *5* (April), 1–15 DOI: 10.1038/srep13969.

[16] Jeeves, R. E.; Marriott, A. A. N.; Pullan, S. T.; Hatch, K. A.; Allnutt, J. C.; Freire-Martin, I.; Hendon-Dunn, C. L.; Watson, R.; Witney, A. A.; Tyler, R. H.; et al. *Mycobacterium tuberculosis* is resistant to isoniazid at a slow growth rate by single nucleotide polymorphisms in *katG* codon ser315. *PLoS One* **2015**, *10* (9), 1–21 DOI: 10.1371/journal.pone.0138253.

[17] Machado, D.; Perdigão, J.; Ramos, J.; Couto, I.; Portugal, I.; Ritter, C.; Boettger, E. C.; Viveiros, M. High-level resistance to isoniazid and ethionamide in multidrug-resistant *Mycobacterium tuberculosis* of the Lisboa family is associated with *inhA* double mutations. *J. Antimicrob. Chemother.* **2013**, *68* (8), 1728–1732 DOI: 10.1093/jac/dkt090.

[18] Hernando Hazbon, M.; Brimacombe, M.; Del Valle, M. B.; Cavatore, M.; Guerrero, M. I.; Varma-Basil, M.; Billman-Jacobe, H.; Lavender, C.; Fyfe, J.; Garcia-Garcia, L.; et al. Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **2006**, *50* (8), 2640–2649 DOI: 10.1128/AAC.00112-06.

[19] Torres, J. N.; Paul, L. V; Rodwell, T. C.; Victor, T. C.; Amallraja, A. M.; Elghraoui, A.; Goodmanson, A. P.; Ramirez-Busby, S. M.; Chawla, A.; Zadorozhny, V.; et al. Novel katG mutations causing isoniazid resistance in clinical *M. tuberculosis* isolates. *Emerg. Microbes Infect.* **2015**, *4*, 1–8 DOI: 10.1038/emi.2015.42.

[20] Manson, A. L.; Cohen, K. A.; Abeel, T.; Desjardins, C. A.; Armstrong, D. T.; Iii, C. E. B.; Brand, J.; Global, T.; Consortium, G.; Chapman, S. B.; et al. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. *Nat. Publ. Gr.* **2017**, *49* (3), 395–402 DOI: 10.1038/ng.3767.

[21] Loots, D. T. An Altered *Mycobacterium tuberculosis* Metabolome Induced by *katG* Mutations Resulting in Isoniazid Resistance. *Antimicrob. Agents Chemother.* **2014**, *58* (4), 2144–2149 DOI: 10.1128/AAC.02344-13.

[22] Justin L. Kandler; Mercante, A. D.; Dalton, T. L.; Ezewudo, M. N.; Cowan, L. S.; Burns, S. P.; Methchock, B.; Investigators, G. P.; Cegielski, P.; Posey, J. E. Validation of novel *Mycobacterium tuberculosis* isoniazid resistance mutations not detectable by common molecular tests. *CDC* **2018**, 1–43.

[23] Seifert, M.; Catanzaro, D.; Catanzaro, A.; Rodwell, T. C. Genetic mutations associated with isoniazid resistance in *Mycobacterium tuberculosis* : A systematic review. *PLoS One* **2015**, *10* (3), 1–13 DOI: 10.1371/journal.pone.0119628.

[24] Pym, A. S.; Saint-Joanis, B.; Cole, S. T. Effect of *katG*mutations on the virulence of *Mycobacterium tuberculosis* and the implication for transmission in humans. *Infect. Immun.* **2002**, *70* (9), 4955–4960 DOI: 10.1128/IAI.70.9.4955-4960.2002.

[25] Pretorius, G. S.; Van Helden, P. D.; Sirgel, F.; Eisenach, K. D.; Victor, T. C. Mutations in *katG* gene sequences in isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis* are rare. *Antimicrob. Agents Chemother.* **1995**, *39* (10), 2276–2281 DOI: 10.1128/AAC.39.10.2276.

[26] Kit, G.; Siu, H.; Yam, W. C.; Zhang, Y.; Kao, R. Y. T. An Upstream Truncation of the furA-katG Operon Confers High- Level Isoniazid Resistance in a *Mycobacterium tuberculosis* Clinical Isolate with No Known Resistance-Associated Mutations. *Antimicrob. Agents Chemother.* **2014**, *58* (10), 6094–6100 DOI: 10.1128/AAC.03277-14.

[27] Tsolaki, A. G.; Hirsh, A. E.; DeRiemer, K.; Enciso, J. A.; Wong, M. Z.; Hannan, M.; Goguet de la Salmoniere, Y.-O. L.; Aman, K.; Kato-Maeda, M.; Small, P. M. Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (14), 4865–4870 DOI: 10.1073/pnas.0305634101.

[28] Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S. The catalase—peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **1992**, *358* (6387), 591–593 DOI: 10.1038/358591a0.

[29] Coll, F. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* **2015**, *7* (1) DOI: http://dx.doi.org/10.1186/s13073-015-0164-0.

[30] Walker, T. M.; Kohl, T. A.; Omar, S. V; Hedge, J.; Del, C.; Elias, O.; Bradley, P.; Iqbal, Z.; Feuerriegel, S.; Niehaus, K. E.; et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance : a retrospective cohort study. *Lancet Infect. Dis.* **2015**, *15* (10), 1193–1202 DOI: 10.1016/S1473-3099(15)00062-6.

[31] Ando, H.; Kitao, T.; Miyoshi-akiyama, T.; Kato, S.; Mori, T.; Kirikae, T. Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2011**, *79* (February), 1615–1628 DOI: 10.1111/j.1365-2958.2011.07547.x.

[32] Gagneux, S.; Burgos, M. V.; DeRiemer, K.; Enciso, A.; Munoz, S.; Hopewell, P. C.; Small, P. M.; Pym, A. S. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog.* **2006**, *2* (6), 603–610 DOI: 10.1371/journal.ppat.0020061.

[33] Nusrath, A.; Subbian, S.; Elizabeth, L.; Selvakumar, N. Infection , Genetics and Evolution Overview on mechanisms of isoniazid action and resistance in *Mycobacterium tuberculosis*. *MEEGID* **2016**, *45* (1), 474–492 DOI: 10.1016/j.meegid.2016.09.004.

[34] Slayden, R. a; Barry, C. E. The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *Mycobacterium tuberculosis*. *Tuberculosis (Edinb).* **2002**, *82* (4-5), 149–160 DOI: 10.1054/tube.2002.0333.

[35] Vilchèze, C.; Jacobs Jr, W. R. Resistance to Isoniazid and Ethionamide in *Mycobacterium tuberculosis*: Genes, Mutations, and Causalities. *Micrlogy Spectr.* **2014**, *2* (4), 1–21 DOI: 10.1128/microbiolspec.MGM2-0014-2013.

[36] Farr, S. B.; Kogoma, T. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiol. Rev.* **1991**, *55* (4), 561–585.

[37] Takayama, K.; Schnoes, H. K.; Armstrong, E. L.; Boyle, R. W. Site of inhibitory action of isoniazid in the synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Lipid Res.* **1975**, *16*, 308–317.

[38] van Kessel, J. C. Van; Hatfull, G. F. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering : characterization of antimycobacterial drug targets. **2008**, *67* (January), 1094–1107 DOI: 10.1111/j.1365-2958.2008.06109.x.

[39] Cardoso, R. F.; Cardoso, M. A.; Leite, C. Q. F.; Sato, D. N.; Mamizuka, E. M.; Hirata, R. D. C.; De Mello, F. F.; Hirata, M. H. Characterization of *ndh* gene of isoniazid resistant and susceptible *Mycobacterium tuberculosis* isolates from Brazil. *Mem. Inst. Oswaldo Cruz* **2007**, *102* (1), 59–61 DOI: 10.1590/S0074-02762007000100009.

[40] Mycobrowser https://mycobrowser.epfl.ch/genes/MSMEG_0306 (accessed May 11, 2018).

[41] NCBI| Nucleotide BLAST https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (accessed May 2, 2018).

[42] Ruesen, C.; Riza, A. L.; Florescu, A.; Chaidir, L.; Editoiu, C.; Aalders, N.; Nicolosu, D.; Grecu, V.; Ioana, M.; Van Crevel, R.; et al. Linking minimum inhibitory concentrations to whole genome sequence-predicted drug resistance in *Mycobacterium tuberculosis* strains from Romania. *Sci. Rep.* **2018**, *8* (1), 1–8 DOI: 10.1038/s41598-018-27962-5.

Chapter 6: Limitations and Future Studies

6.2 Limitations of study

The data presented in this thesis was investigated to derive the reason for discrepancies between the LPA and DST with the purpose of informing the current drug resistance testing algorithm for INH. However, there were a number of limitations that could have an impact on the results obtained. The first is the lack of clinical data for the isolates investigated, as there may be multiple isolates from a single patient. This may result in over or underestimation of the prevalence of specific SNPs in this cohort. Secondly, whole genome data was only obtained for a subset of 130 of the 397 isolates. As a result, some resistance causing mutations in alternative genes (excluding the *katG* gene and *inhA* promoter region) may have been missed. The homologous recombineering experiment also had various limitations. For instance, the use of *M. smegmatis* as a model allows inference as to the potential effects of the point mutations in *M. tuberculosis* but it is not conclusive that the impact will be the same in the virulent bacteria. Additionally, the vector used could potentially have initiated other undesireable recombination events which may have influenced INH resistance.

6.2 Future Studies

Future studies will minimise the impact of the abovementioned limitations. Applications for ethical approval to access to clinical data are currently underway. Whole genome sequencing will be done on the entire dataset to ensure the correct mechanism of resistance has been elucidated for each isolate. Furthermore, homologous recombineering will be initiated in H37Rv and the entire genome will be investigated via means of WGS. This will allow comparison between the genome of wild-type and recombineered H37Rv in order to confirm there are no alternative mechanisms that are causative of resistance. Alternatively, further possible recombineering events caused by the *Che9c 61* gene can be eliminated via curing of the pJV75amber plasmid by several passages on media lacking selective drugs. Succesful removal of the plasmid with persisting resistance to INH is indicative of the mutation inserted being the cause of resistance.

Chapter 7: Conclusion

As hypothesised, some of the discrepancies between LPA and DST can be attributed to unknown mechanisms of resistance in this cohort. Novel mutations or deletions as well as mutations present elsewhere in the *M. tuberculosis* genome have been identified in a large proportion of cases as the cause for INH resistance. Thus, an alternative or updated drug resistance testing algorithm for INH resistance may be required.

Per the results of this study, it is evident low-level INH resistance is an underrated component of the global TB epidemic worldwide, and may be a significant problem in terms of progression to antibiotic resistance and treatment outcome. Low-level resistance can lead to progression to multi-drug resistance. We need to investigate this carefully and test whether personalised diagnostics and targeted high dose treatment with INH will reduce the incidence of isoniazid mono-resistant and multi-drug resistant (MDR) tuberculosis.

Additionally, a significant amount of isolates have been identified to be resistant to INH but not to have the canonical mutations tested for by molecular test such as the LPA and a small proportion were misdiagnosed due to lab error. Thus, the current drug resistance testing algorithm used is logical. The data suggests that it is not sufficient to employ the LPA as the sole INH resistance testing mechanism in all cases that are diagnosed as rifampicin resistant. Instead, it serves as evidence that the secondary phenotypic DST may have to be employed in all cases, whether the isolates are found to be susceptible or resistant to INH and at more than one drug concentration. Multiple concentrations of drug in DST would uncover the extent of resistance to ensure correct diagnoses are made to inform the antibiotic regimen. Although, it should be noted, that this is time consuming and expensive and should thus not be the sole solution to the current problem of discrepancies in detection of resistance causing mutations for routine laboratories.

Our recommendation is that WGS should be implemented in order to minimize the chances of missing non-canonical INH resistance mutations outside targeted hotspots. Future studies should obtain WGS data for all isolates in this study in order to prevent novel SNPs from being undiagnosed, and to ensure there is no bias. The findings of this study may inform diagnostic algoritms and WGS analysis for improved testing for INH resistance. Additionally, a better understanding of the mechanisms of resistance will aid in devising new strategies to overcome drug-resistance.

Supplementary data

Reagents used for MGIT

MGIT medium

A MGIT 960 tube contains 7.0 ml of modified Middlebrook 7H9 broth base. The approximate formula per 1000 ml purified water is:

- Modified Middlebrook 7H9 broth base -------- 5.9 g
- Casein peptone ---------- 1.2 g
- Add supplements as required.

MGIT Growth supplement (OADC)

MGIT growth supplement is used for the preparation of subcultures and inoculate and contains 15 ml of the following per liter of water.

- Bovine albumin ---------- 50.0 g
- Dextrose---------- 20.0 g
- Catalase ---------- 0.03 g
- Oleic acid ---------- 0.1 g
- Polyoxyethyline state ---------- 1.1 g

Solutions used for spoligotyping

Working stocks

500ml 2xSSPE/0.1%SDS:

- 100ml 10x SSPE
- 5ml 10% SDS
- dH2O to 500ml

1000ml 2xSSPE/0.5%SDS:

- 200ml 10x SSPE
- 50ml 10% SDS
- dH2O to 1000ml

1000ml 1% SDS:

- 100ml 10% SDS
- dH20 to 1000ml

500ml 20mM EDTA:

- 25ml 0.5M EDTA
- dH2O to 500ml

Stock solutions

10x SSPE:

- 13.7g Sodium hydrogen Phosphate – 100mM
- 105.19g Sodium Chloride – 1.8M
- 3.36g EDTA – 10mM
- dH2O to 1000ml
- pH = 7.40

10% SDS:

- 50g SDS
- dH2O to 500ml

0.5M EDTA:

- 93g EDTA
- dH2O to 500ml