Streptomycin and ethambutol resistance associated mutations of multidrug-resistant *Mycobacterium tuberculosis* clinical isolates in Lusaka, Zambia

（ザンビアで分離された多剤耐性結核菌におけるストレプトマイシンおよびエタンブトール耐性関連遺伝子変異の特徴）

Precious Bwalya

Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
DNA	deoxyribonucleic acid
DST	drug susceptibility testing
EMB	ethambutol
embB	gene encoding arabinosylindolylacetylinositol synthase
gidB	gene encoding glucose-inhibited division protein B
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
INH	isoniazid
IS6110	insertion sequence 6110
LF-LAM	lateral flow lipoarabinomannan
LSP	large sequence polymorphism
MDR	multi drug resistant
Mtb	*Mycobacterium tuberculosis*
RFLP	Restriction fragment length polymorphism
RIF	rifampicin
RNA	ribonucleic acid
rpoB	gene encoding RNA polymerase subunit B
rpoC	gene encoding DNA-directed RNA polymerase beta′ subunit
rpsL	gene encoding 30S ribosomal protein S12 RpsL
rrs	gene encoding ribosomal RNA 16S
SNP	single nucleotide polymorphism
STR	streptomycin
TB	tuberculosis
WHO	world health organization

LIST OF DEGREE RELATED PAPERS

Precious Bwalya, Tomoyuki Yamaguchi, Eddie Samuneti Solo, Joseph Yamweka Chizimu, Grace Mbulo, Chie Nakajima, Yasuhiko Suzuki. Characterization of mutations associated with streptomycin resistance in multidrug-resistant *Mycobacterium tuberculosis* in Zambia. Antibiotics 2021;10:1169. doi:10.3390/antibiotics10101169.

PREFACE

Tuberculosis (TB) is an infectious disease caused by the bacillus *Mycobacterium tuberculosis (Mtb)*. It has remained a health concern since its declaration as a public health problem in 1993 [1]. Although global incidences of TB have been reducing in past years, in 2020, nearly 9.9 million people developed TB with most cases being reported in the WHO regions of South-East Asia (43%), Africa (25%), and Western Pacific (18%) (figure 1) [2]. It is among the leading causes of mortality worldwide. An estimated 1.3 million people died from TB in 2020 alone, an increase from 1.2 million reported in 2019 among HIV-negative people. In HIV-positive patients, mortality increased from 209 000 in 2019 to 214 000 in 2020 [2].

Mtb the causative agent of TB, is a gram-positive bacillus belonging to the phylum *Actinobacteria*. Members of this phylum have a high G+C content in their genome, exhibit pleomorphic morphologies, and can be found freely living in wide-ranging ecological environments including soil, water, plants, animals, and human [3]. The *Mycobacterium* genera of *Actinobacteria* constitute important human pathogens. Members of this genus are characterized by a complex cell wall and a waxy cell envelope. The cell wall structure can be divided into 3 major components: a peptidoglycan, arabinogalactan, and mycolic acid layers [4]. The peptidoglycan layer functions to provide shape and rigidity to the cell, protecting against turgor pressure, and anchoring other outer cell components [3]. Arabinogalactan a polysaccharide, links mycolic acids to the peptidoglycan layer and is important for the cell growth [5]. A distinctive feature of the members of the genus *Mycobacterium* is the presence of mycolic acids, unique fatty acids that make up the mycobacteria cell envelope and responsible for the characteristic acid fastness. Mycolic acids are essential for the cell viability and virulence [5]. The mycolic acids make an intercalating complex with other lipids to form a hydrophobic outer membrane highly impermeable to many compounds and enabling it to survive in stressful environmental conditions such as drug presence [6]. *Mtb* and *Mycobacterium leprae* causing tuberculosis and leprosy, respectively, are the most important obligate human pathogens of the genus *Mycobacterium*.

Infection by *Mtb* is by inhalation of droplet nuclei containing live bacteria. It primarily causes pulmonary TB, the most infectious form of TB. Upon inhalation, the bacteria are phagocytosed by alveolar macrophages, forming a phagosome where they persist by escaping the macrophage mechanism to clear intracellular pathogens such as preventing phagosome and lysosome fusion and resistance to reactive nitrogen species [4][7].

Infection by *Mtb* has 3 possible outcomes including 1) clearance by host immune system, 2) containment as latent infection, and 3) progression to active disease. In latent TB infection (LTBI), the host immune response restricts *Mtb* growth and maintains the bacteria in a state of dormancy surrounded by monocytes, dendritic cells, lymphocytes, and fibroblasts [8]. This lesion containing bacteria also known as a granuloma or tubercle, is the hallmark of TB. The granuloma

prevents bacteria from spreading throughout the tissue and other body organs, and live bacteria can persist in the lesion for decades. An estimated 21% to 25% of the global population has LTBI [9]. People with LTBI are asymptomatic but, when the immune system weakens, they can develop symptomatic TB, thus LTBI presents a large reservoir for TB disease. Approximately 10% of people with LTBI have a risk of developing active TB at some point in their life [10]. Active TB is characterized by persistent cough with bloody phlegm, weight loss, night sweat, fatigue, and fever. For clarity, in this thesis, active TB will be referred to as TB.

The immune system suppression predisposes to TB development by compromising the integrity of the granuloma causing it to rupture and release live bacteria. Factors such as diabetes, malnutrition, smoking, alcohol, age, immunosuppressive therapy, cancer, and HIV contribute to immune system suppression and increase the risk of developing TB [11–13]. In addition to increasing the risk of TB reactivation from LTBI, HIV can also increase progression to TB disease after initial infection or reinfection [14].

Recent studies have also confirmed the role of host genetic factors to susceptibility to TB infection as well as the development of disease. Some studies have pointed to a genetic variant on chromosome 11p13 associated with resistance to TB. A single nucleotide polymorphism (SNP) rs2057178 near *WT1* on this chromosome had a protective association against TB in study populations from Ghana (OR 0.77, p value = 2.63×10^{-9}), Gambia (OR 0.80, p value = 4.87×10^{-4}), Russia (OR 0.91, p value = 0.02), South Africa (OR 0.62, p value = 2.7×10^{-6}), and Morocco (OR 0.78, p value = 0.43) [15–17]. A genome-wide association study conducted on the Icelandic population found 2 HLA class II variants that affected the risk of pulmonary TB. The SNP rs557011 [T] appeared to increase the odds of developing the disease (OR 1.25, p value = 5.8×10^{-12}) while rs9271378[G] protected against disease development (OR 1.25, p value = 5.8×10^{-12}), possibly by affecting antigen presentation to T cells [18]. Host genetic susceptibility to TB was also exemplified in the Lübeck disaster in which 251 children were accidentally inoculated with Bacille Calmette-Guérin (BCG) vaccine contaminated with variable concentrations of pathogenic *Mtb*. Out of the 251 inoculated children, 228 developed different forms of TB. Seventy-seven deaths were recorded within a year of exposure, and 72 of these were attributed to TB. Remarkably, 17 (6.8%) children had no clinical signs of TB while 68% of children who had developed symptomatic TB, spontaneously cured stressing the role of host genetics in TB disease resistance and severity [19]. The observed variability in TB outcome could also be attributed to different dosage of *Mtb* administered to the children suggesting that bacteria dosages at exposure also influences the outcome of infection. Undoubtedly, host genetic variability and exposure dosage can influence the risk of developing TB.

Mtb has intrinsic resistance to many antibiotics due to the highly impermeable lipid-rich cell envelope and presence of drug inactivating enzymes such as beta-lactamases. Nonetheless, TB is a treatable disease. The recommended first-line anti-TB drug combination of rifampicin (RIF),

isoniazid (INH), ethambutol (EMB), and pyrazinamide (PZA) for the treatment of drug-susceptible TB has approximately 85% treatment success [20]. The emergence of drug resistance to these drugs is a public health concern globally, particularly rifampicin resistance (RR) and multi-drug resistant TB (MDR-TB) which is resistance to the two most potent anti-TB drugs RIF and INH [21]. For both forms of drug resistance (RR and MDR), second-line drugs are used for treatment. History of previous TB treatment, poor quality drugs, inconsistent supply of TB drugs, delayed diagnosis, accessibility of health care facilities, and patient-related factors such as non-adherence to treatment, social-economic factors increase the likelihood of drug resistance emergence [18]. Human immunodeficiency virus (HIV) co-infection has also been documented to increase the likelihood of developing drug resistance due to drug malabsorption, leading to suboptimal drug plasma concentration [22]. The emergence of MDR/RR-TB is problematic compared to drug-susceptible TB because treatment of this form of TB is highly costly, uses toxic drugs producing adverse side effects; in addition, the patient has to endure a high pill burden for a prolonged treatment period [23]. Moreover, MDR-TB treatment often has poor outcome with only 57% global success rate [20]. MDR/RR-TB emergence is thus a threat to the global efforts to end TB.

Drug resistance in *Mtb* is mediated by the acquisition of chromosomal mutations in genes that encode proteins targeted by the drugs [6]. These mutations can cause drug target structural alteration such as RNA polymerase mutations that cause RIF resistance [24], upregulation of drug target proteins as is the case with NADH-dependent enoyl-acyl carrier protein mutations conferring resistance to INH [25], and overexpression of efflux pumps such as mutations in Rv0678 leading to upregulation of the MmpL5 efflux pump and causing resistance to bedaquiline and clofazimine [26]. Some drugs are prodrugs and require enzymatic activation to exert their antimicrobial properties. Mutations in genes encoding enzymes responsible for activating the drugs abrogate drug activation. An example of prodrugs is PZA and INH [6].

It has been recognized that an interplay of bacterial determinants influences the evolution of drug resistance-conferring mutations and their maintenance. The biological cost of mutations on the organism's fitness is an important predictor of the acquisition of drug resistant confering mutations. The degree of impact on fitness depends on the specific mutation, the environment, and genetic background of the bacteria. Forexample, among RIF resistance confering mutations, *rpoB* S450L is a least fitness cost in *Mtb* [27]. Addittionally, the fitness cost of a mutation also differs in different growth conditions [28]. Clearly, drug resistance mutations may affect the fitness of an organism.

In most experiments, the inference of *in vivo* fitness cost of mutations on *Mtb* is done by *in vitro* comparison of the growth rate of mutant and susceptible strains. This method fails to capture other factors that may have an effect on mutation acquisition such as the host immune system and genetics, nutrient availability, and co-morbidities [28][29]. An alternative way to measure fitness is by quantifying resistant variant frequency in a population which correlates well with *in vivo* fitness.

The most frequently observed mutations in clinical isolates are considered to have reduced *in vivo* fitness cost. For example, a genome-wide analysis found that the *rpoB* mutation S450L previously reported by *in vitro* experiments as the least costly mutation was the most frequent mutation among 6,465 MDR-*Mtb* clinical isolates at 64.2% [27,30]. An allelic exchange experiment revealed that *rpsL* mutation K42R in *Mycobacterium smegmatis* had no fitness cost and the corresponding mutation in *Mtb* K43R accounted for 42.2% of STR resistance in MDR-TB clinical isolates [30,31]. Mutations with higher fitness costs, however, may also be found in reasonable frequency due to the occurrence of compensatory mutations that restores fitness to the bacteria [6]. The mutations with high fitness costs can thus persist in the population and transmit easily causing outbreaks of MDR-TB. For example, an outbreak of MDR-TB in South Africa and Eswatini was caused by strains having a rare *rpoB* mutation I491F whose fitness cost in some strains was alleviated by the acquisition of a *rpoC* mutation E1033A leading to increased transmissibility of the strains [32]. This interaction between a drug resistance-conferring mutation and a compensatory mutation is an example of epistasis which is defined as a non-additive interaction between two genes/alleles that produces an effect on a phenotype such as fitness [6].

The genomic differences of *Mtb* strains is another important bacterial determinant of drug resistance as well as several other observed phenotypes of the organism such as transmissibility and virulence. The strain's genomic SNPs, insertions and deletions (indels), large genomic deletions or duplications, and mobile and repeat elements produce genetic diversity which characterize genetic background of a strain [6]. The genetic background of a strain may influence the impact of a drug resistance mutation on fitness. Variable fitness was observed among 5 clinical isolates having the same S450L amino acid substitution in *rpoB*, underlining the influence strain genetic background may have on an organism's fitness upon acquisition of resistance-conferring mutations [27]. Moreover, the level of drug resistance conferred by a particular mutation can differ depending on the strain's genetic background [33].

Based on genomic differences, *Mtb* can be divided into Lineages 1 to 4, and Lineage 7, whereas Lineages 5 and 6 are *Mycobacterium africanum* [34]. These Lineages show differences in geographic distribution and varied associations to drug resistance. Lineage 1 also known as Indo-Oceanic or EAI, is dominant around the Indian Ocean, while Lineage 3 known as Central Asian Strain (CAS) is predominantly found in East Africa and Central and South-Asia, and Lineage 7 is found exclusively in Ethiopia [35][36]. On the other hand, Lineages 2 (also referred to as East-Asian or Beijing strains) is gradually expanding globally while Lineage 4 (referred to as Euro-American) is globally distributed. The Beijing genotype of Lineage 2 has been shown to have a high propensity to acquire drug resistance due to a higher mutation rate, and based on isolate clustering, has been associated with increased transmission of drug-resistant strains compared to Lineage 1 [37–39]. Genetic background can influence the pathway of drug resistance as shown by Lineage 1 having higher odds of acquiring

inhA mutations -15C/T for INH resistance compared to other lineages (OR 6.4, *p* value = 0.002) [33]. In addition, MDR-TB outbreaks have been attributed to different sublineages in different geographical areas. MDR-TB outbreaks in Vietnam and Thailand were attributed to the Beijing genotype [38,39], in South Africa and Eswatini to S genotype of Lineage 4 [32], in South Africa's KwaZulu Natal region to LAM4 [40], in Argentina to sublineage 4.1.2.1 of Lineage 4 [41], and in India to CAS1_Delhi [42]. Taken together, bacterial determinants such as fitness cost, compensatory mutations, and genomic variability play a key role in driving drug resistance mutation acquisition.

To study *Mtb* genetic diversity, several molecular methods have been developed, including insertion sequence 6110 (IS6110) restriction fragment length polymorphism (RFLP), spacer oligonucleotide typing (spoligotyping), mycobacterial interspersed repetitive-units-variable number tandem-repeat (MIRU-VNTR) typing, large sequence polymorphisms (LSP), and SNP analysis. The IS6110 RFLP method is based on different copy numbers of the IS6110 and the position of insertions in the genome of *Mtb* isolates to generate a strain-specific fingerprint [43]. Spoligotyping detects the presence or absence of 43 variable spacers in the direct repeat (DR) locus and can differentiate *Mtb* to sublineage level such as CAS1_Kili, LAM, T, X, Beijing, etc [44]. Unlike spoligotyping which targets one locus, MIRU-VNTR targets multiple loci containing a variable number of tandem repeats and is thus more discriminatory than spoligotyping [45]. LSP is based on the detection of regions of difference due to deletions in the genome of *Mtb* strains while SNP analysis identifies strain-specific mutations in isolates compared to the reference strains and both can classify *Mtb* into Lineages and sublineages [46,47]. Genotyping of *Mtb* is important in the epidemiological surveillance of TB, outbreak investigation, and the identification of transmission clusters in different geographical settings.

As already discussed, drug resistance mutation frequencies are influenced by bacterial genotypes. The distribution of these genotypes is different depending on geographical area. Therefore, regional setting studies including drug resistance mutations and genotypes are relevant to inform specific control strategies to curb antibiotic resistance and contribute to the development of rapid diagnostic tools.

Zambia has seen a downward trend of TB incidences in the past decade (figure 2), yet still, it is regarded as a high TB burden country [2]. An estimated 59, 000 people became ill with TB in 2020 and only 68% of these were started on treatment [2]. The approximately 30% with delayed treatment initiation can become the source of infection in the population. Factors such as death, untraceable patients due to incorrect or change of address, socioeconomic factors, administrative factors, long diagnostic result turnaround time, and health care provider behavior and attitude; may contribute to delayed treatment initiation [48]. In addition to delayed treatment, approximately 40% of TB cases in Zambia do not have access to microbiological testing [49]. The lack of diagnosis of these TB patients may be due to difficulties to access health care and TB facilities, delay to seek health care,

and unsuspected TB [49][50]. The undiagnosed and untreated TB patients perpetuate TB transmission and could contribute to the high TB burden in Zambia.

Several methods are used for confirmation of a TB diagnosis in Zambia including smear microscopy, lateral flow urine lipoarabinomannan assay (LF-LAM), Xpert MTB/RIF assay, line probe assay (LPA), and culture. Smear microscopy involves demonstrating by sputum smear staining the presence of acid-fast bacilli (AFB). This technique is widely used in TB screening in high TB burden countries, but it has moderate sensitivity [51]. The LF-LAM assay uses urine for diagnosis of TB and is used in patients who cannot produce sputum such as HIV patients [52]. Isolation of *Mtb* by culture is more sensitive than smear microscopy but, in Zambia, it is limited to all retreatment cases, suspected treatment failure, and drug resistance surveillance [53]. The purpose of culturing for these classes of patients is to perform drug susceptibility testing (DST) to guide treatment, however, this method's turnaround time is several weeks to months. In 2018, the molecular tools Xpert MTB/RIF assay and LPA were formally adopted as a first-line tool for the diagnosis of TB in Zambia as well as for DST to first and second-line drugs, respectively. These tools are more rapid and sensitive compared to culture and have improved drug resistance case detection in the country [51][54].

Zambia has seen an increasing trend of MDR/RR-TB in recent years and as a result, was included on the 2021 updated list of high MDR-TB burden countries in the world [55]. Figure 3 shows a gradual increase of laboratory-confirmed MDR-TB over a period of 12 years in Zambia from 2000 to 2011 [56]. It can be said, therefore, that MDR-TB in Zambia has been silently increasing for the past 2 decades. The adoption of molecular-based tools such as Xpert MTB/RIF and second-line DST tool LPA could be one reason for the observed recent steep rise in the number of laboratory-confirmed new MDR/RR-TB cases [54]. The rollout of Xpert MTB/RIF in Zambia started in 2017, and that year saw a significant increase of laboratory-confirmed MDR/RR-TB from 180 in 2016 to 546 (figure 4).

Despite the improved MDR-TB detection, there are still significant gaps in DST and prompt treatment initiation for MDR/RR-TB in Zambia. In 2018 for example, there were 3100 estimated incidences of MDR/RR-TB. In that year, 627 laboratory-confirmed MDR/RR-TB cases were reported, 506 were started on MDR/RR-TB treatment, and yet still, only 150 were tested for second-line drug resistance (figure 4) [21]. That means a significant proportion of MDR/RR-TB patients remain untested against second-line drugs and are treated empirically without DST.

Further complicating the management of MDR/RR-TB in Zambia, is the fact that only a few drugs used in the treatment of MDR/RR-TB are covered by the adopted rapid molecular tools. Currently, the rapid molecular test LPA is used to detect Rif, INH, fluoroquinolones, and an aminoglycoside amikacin (AMK), while phenotypic DST is used for EMB, streptomycin (STR), clofazimine (CFZ), bedaquiline (BDQ), and linezolid (LZD). Because of the limited DST coverage and the number of drugs tested out of those used in MDR/RR-TB treatment, there may be patients on inadequate treatment. Inadequate treatment is a two-edged sword. On one hand, it leads to patients remaining

infectious for long periods of time and thereby becoming the sources of transmission, and on the other hand, predisposes to the emergence and amplification of drug resistance and poor treatment outcome for the patient [57].

Using DNA sequencing and genotyping methods, recent studies have revealed frequent mutations that lead to resistance to INH and RIF in Zambia. Analysis of those mutations revealed clustering of MDR-TB isolates and an association between MDR-TB and CAS1_Kili and LAM1 genotypes [58]. They have thus shown that the rise of MDR-TB cases in Zambia is being driven by the emergence and transmission rather than cross border transmission of MDR-*Mtb* strains [59][60].

Therefore, more effort is needed to hasten the progress that Zambia has been making to control MDR/RR-TB. To reverse the trend of increasing MDR/RR-TB cases, it is important to recognize that early resistance detection and promptly initiating treatment guided by susceptibility profiles is key. Thus, scaling up coverage of testing is critical. Mutation analysis by DNA Sequencing is proving useful in the management of TB patients as well as in surveillance of drug resistance. It has a short results turnaround time and does not require specialized facilities compared to the conventional phenotypic DST methods which requires Biosafety level 3 and is thus limited to a few facilities in Zambia. Additionally, sequencing can be used to determine resistance patterns to a wide range of drugs, explore clustering, and transmission of *Mtb* strains. Information about frequency of dug resistant conferring mutations can also inform the development of easier to use and fast diagnostic test methods.

The aim of this thesis was to characterize mutations conferring resistance to STR and EMB, two drugs used in management of MDR-TB and evaluate the utility of DNA sequencing for DST of MDR-TB isolates in Zambia. Additionally, the correlation between drug resistance-conferring mutations and *Mtb* genotypes was determined using the spoligotyping genotyping method. This method detects the variability in the Direct Repeat Locus. This locus has 36 base pair repeat sequences interrupted by non-repetitive sequences of 34 to 41 base pairs (figure 5) [44].

This thesis has two chapters. Chapter I describes mutations in *rpsL, rrs,* and *gidB* genes known to confer resistance to STR and discusses their correlation with the circulating genotypes in Zambia. In chapter II, mutations in *embB* conferring resistance to EMB are described. Additionally, I discuss the disparity of resistance conferring mutations and phenotypic EMB DST. Using publicly available data on TB profiler website, an interesting phenomenon of possible genotypic background influence on EMB drug resistance trajectory is discussed in chapter II.

Figures

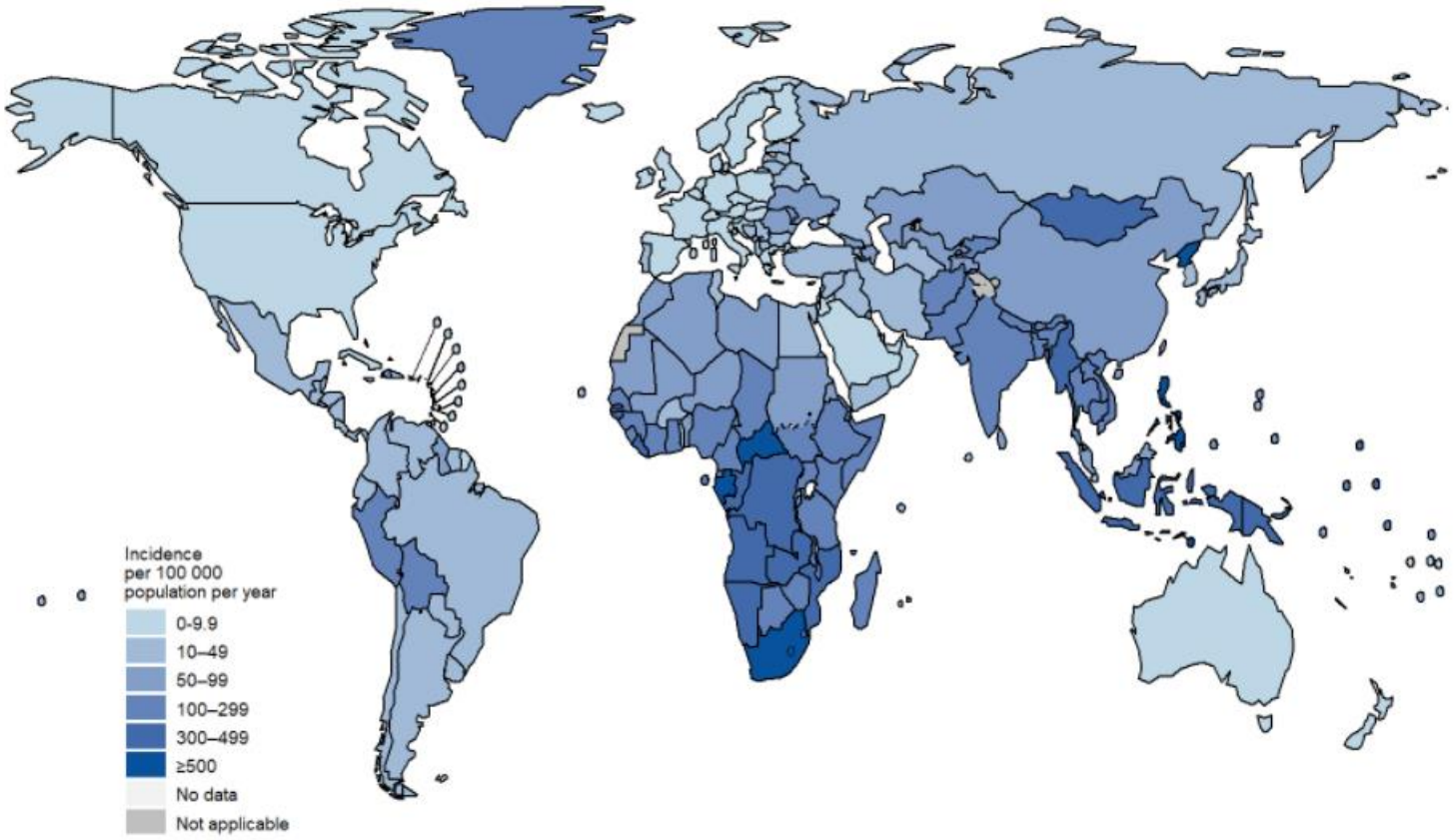


Figure 1: Estimated TB incidence rates, 2020 [2]

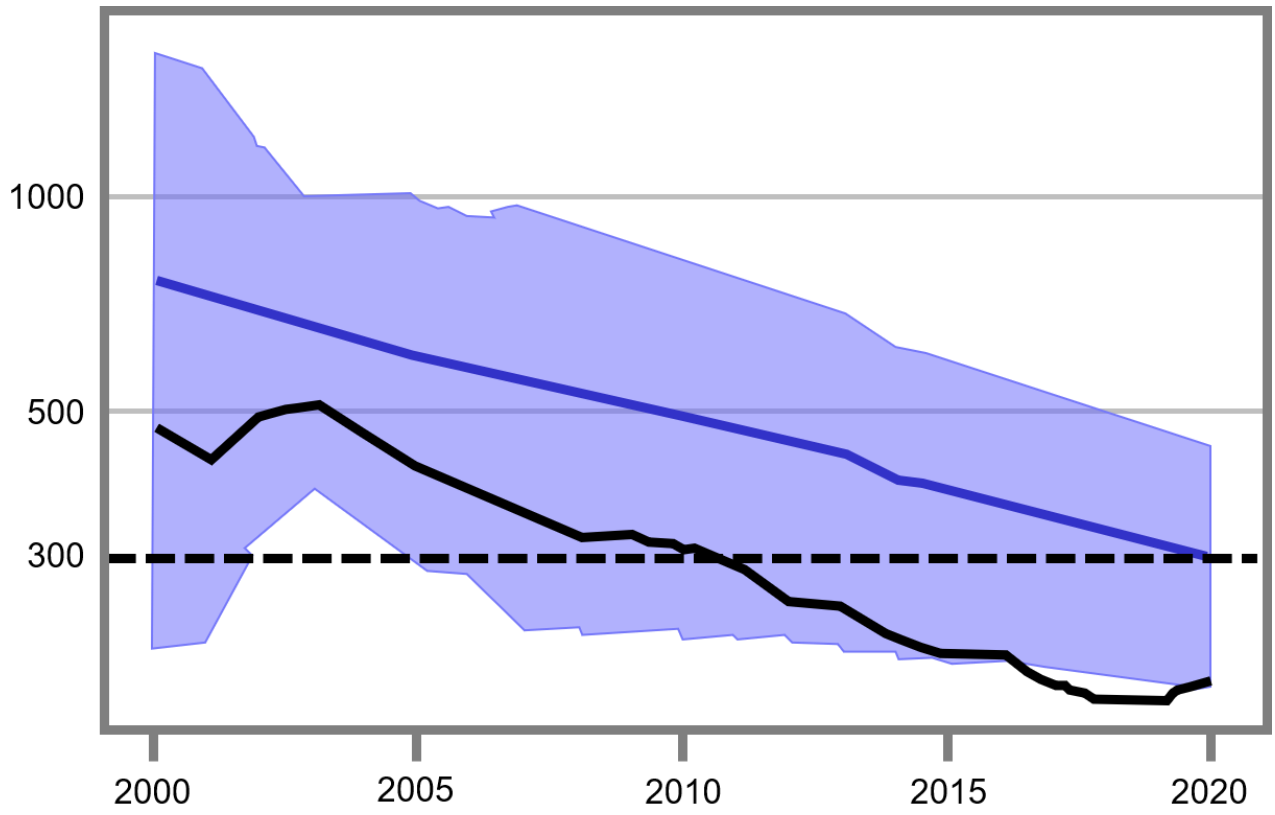


Figure 2: Trend in estimated TB incidence rate in Zambia compared with notifications of new and relapse cases, 2000-2020. The blue line shows incidence rates (shaded area shows uncertainty intervals) while the solid black line indicates notification of new and relapse cases for comparison with estimated total incidence rate. The horizontal dashed line show the 2020 milestone of the end TB strategy [2]. (Accessed 5th November 2021)

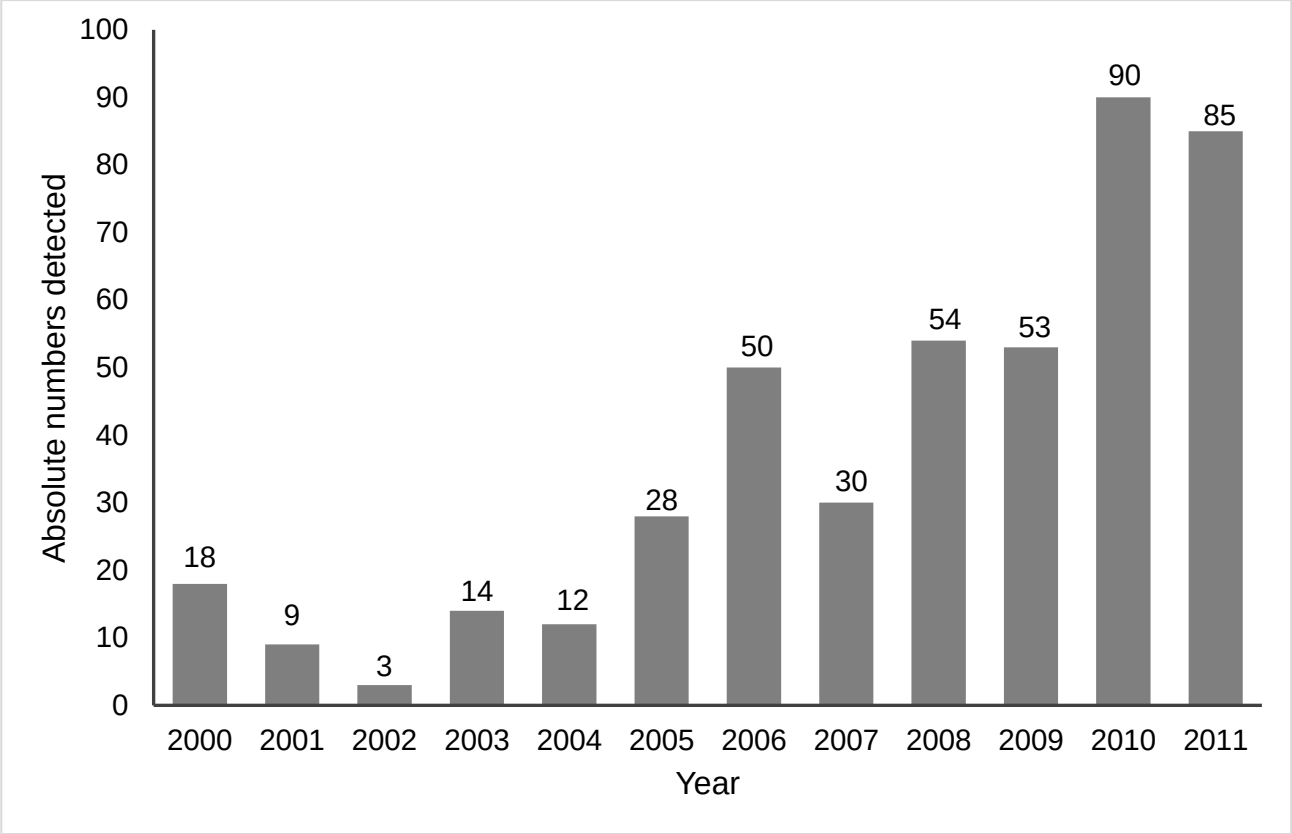


Figure 3: Number of MDR-TB diagnosed by year from 2000 to 2011 in Zambia [56]

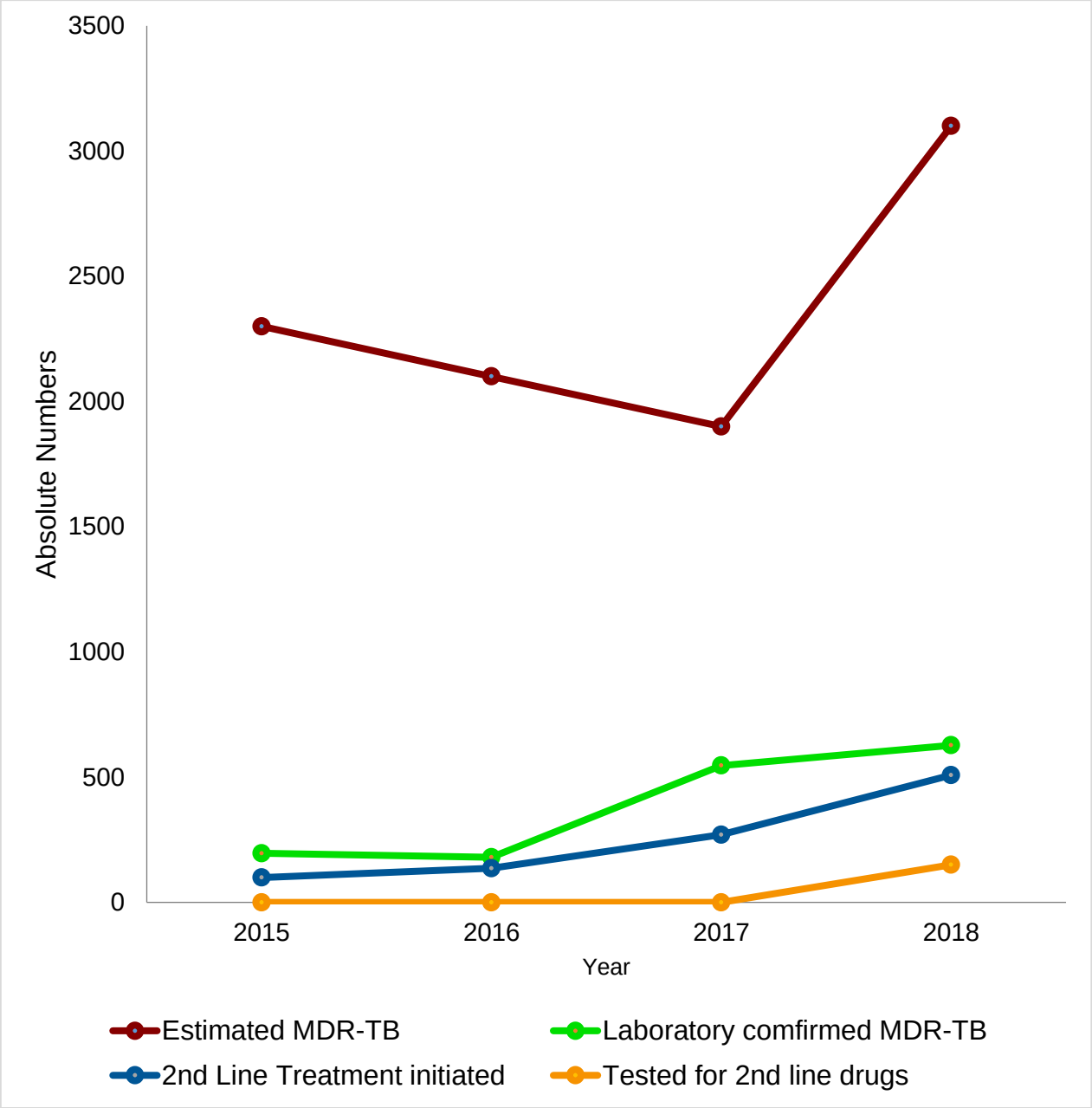


Figure 4: Laboratory confirmed MDR/RR-TB and notified cases in Zambia [21,61–63]

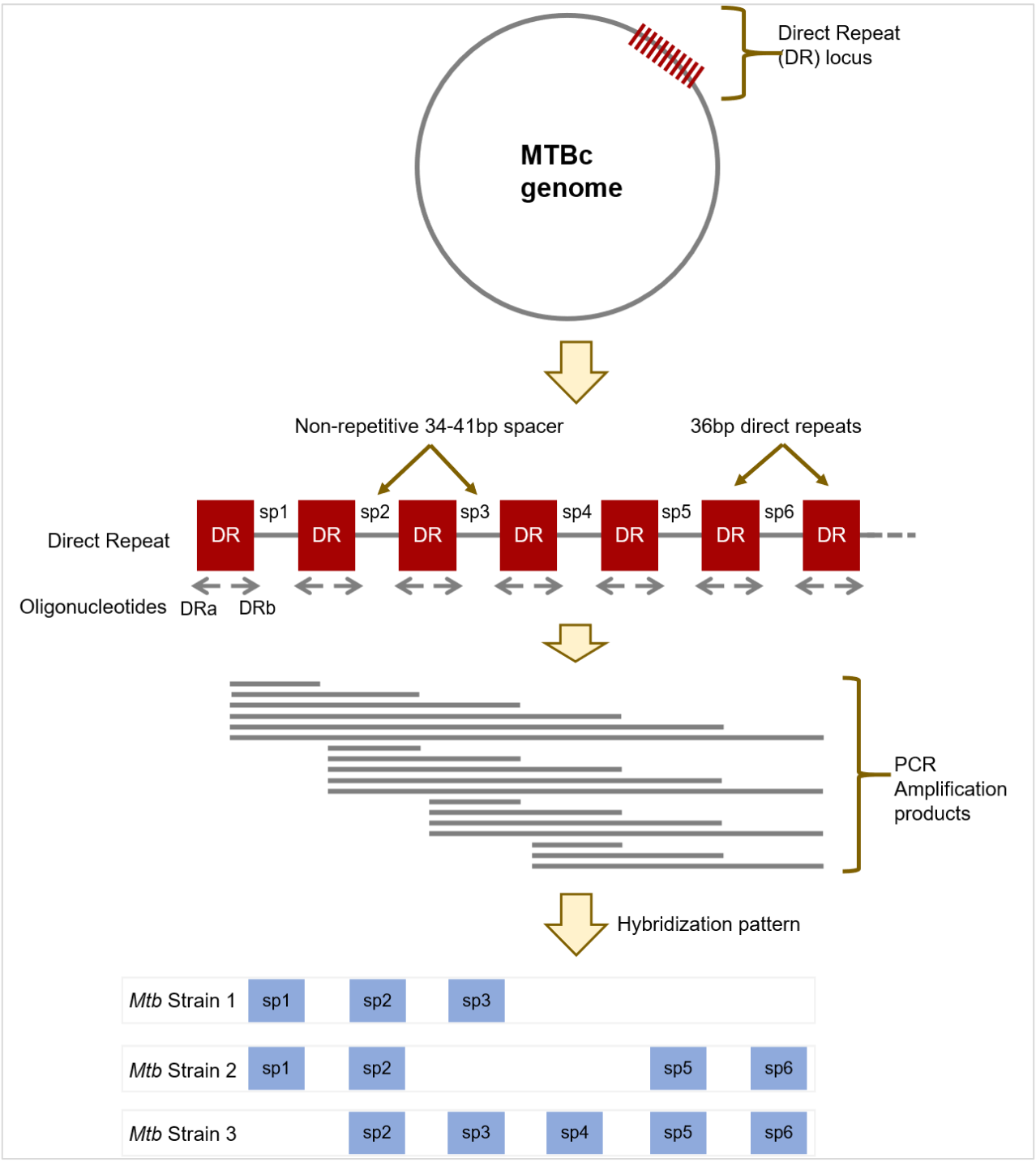


Figure 5: Spoligotyping principle (modified from Kamerbeek et al., 1997) [44]

CHAPTER I

Characterization of mutations associated with streptomycin resistance in multidrug resistant *Mycobacterium tuberculosis* in Zambia

Introduction

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (*Mtb*) is a global public health concern ranked among the top 10 causes of mortality. In 2019, an estimated 1.2 million deaths among human immunodeficiency virus (HIV) negative and 208 000 deaths among HIV positive people were attributed to TB [20]. Zambia is among the 30 countries with the highest TB burden in the world. The emergence of drug resistance raises even more serious concern as it poses a threat to the strides made in controlling the disease. Effort to reduce the incidences of drug resistance, particularly, multidrug resistant tuberculosis (MDR-TB) which is resistance to rifampicin (RIF) and isoniazid (INH), have led to a steady decline of global cases. On the contrary, in Zambia, the incidences of MDR-TB show an upward trend. From 2014 to 2019, MDR-TB has risen from 0.3% to 2.4% in new cases and 8.1% to 18% in previously treated cases [64]. This observed increase of MDR-TB in Zambia can be attributed to an increase in coverage of resistance testing and/or the emergence and spread of MDR-*Mtb* due to inadequate control measures. Earlier guidelines in Zambia recommended that drug susceptibility testing (DST) should be performed on all retreatment cases and suspected treatment failure. However, adherence to that guideline was low [65], leading to usage of drugs without confirmation of susceptibility testing. This practice could have led to undetected emergence and silent spread of drug resistant TB. In a previous study, it was demonstrated that the increase in MDR-TB in Zambia is not the result of transborder transmission but local emergence and possible spread within Zambia [58][60]. The mismanagement of MDR-TB could further the evolution of resistance to other anti-TB drugs. Recently, a first confirmed case of extensively drug resistant TB (XDR-TB) was reported in Zambia [21], stressing the importance of selecting effective drugs for treatment to avoid amplification of resistance to other drugs and prevent transmission of MDR-*Mtb*. The adoption of tools that rapidly detect resistance could guide selection of effective drugs and help control MDR-TB.

Streptomycin (STR) is one of two aminoglycosides recommended in TB treatment. It is currently recommended for limited use as a second-line drug in personalized treatment regimens of MDR-TB patients [23]. With the observed increased drug resistance in Zambia, STR still has a significant role in MDR-TB treatment. It is thus important to assess the level of STR resistance among MDR-TB patients for whom the drug is currently recommended and understand the molecular mechanism of resistance. This will help to assess the usefulness of STR in MDR-TB treatment and of a rapid

molecular based tool for determination of STR resistance in Zambia.

STR induces mistranslation of mRNA to protein by tightly binding to 16S rRNA and ribosomal protein S12 encoded by *rrs* and *rpsL,* respectively (figure 6) [66]. Mutations in *rpsL* are associated with medium to high level STR resistance while mutations in the 530 loop and 912 region of *rrs* are associated with intermediate STR resistance [67]. In addition, mutations in *gidB* encoding a methyltranferase enzyme responsible for methylating a G nucleotide at position 518 of the *rrs* are known to confer low level STR resistance [68,69]. These mutations in *gidB* are at times misclassified as susceptible at the recommended critical concentration of 1 µg/ml and may co-occur with high level resistance conferring mutations to amikacin, the only other aminoglycoside recommended in TB treatment [70]. Moreover, studies have revealed a difference in the types of mutations in different geographical locations and among different genotypes of *M. tuberculosis*. It is therefore important to investigate mutations conferring STR resistance in local settings and this information could assist in the development and adoption of rapid diagnostic tools and consequently in the control of MDR-TB. Thus, I aimed at screening and characterizing mutations in *rpsL, rrs* and *gidB* that confer resistance to STR and examine their association with the circulating genotypes in Zambia.

Materials and Methods

Samples and DST

Hundred and thirty-eight isolates were randomly sampled from among multidrug resistant *Mtb* isolates archived between 2011 and 2017 at University Teaching Hospital Tuberculosis (UTH-TB) laboratory in Lusaka. The laboratory is a reference culture and DST facility covering Eastern, Lusaka and Western provinces of Zambia. Demographic data as well as phenotypic drug susceptibility test results was obtained from the laboratory information system.

Drug susceptibility testing was performed by the laboratory following the manufacturer's manual (BD BACTEC^{TM} MGIT^{TM} 960 SIRE kit) (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) for STR, RIF, INH, and EMB at concentration 1.0 µg/ml, 1.0 µg/ml, 0.1 µg/ml, and 5.0 µg/ml, respectively. The BD BACTEC M960 machine was used for incubation.

DNA extraction, PCR amplification and sequencing of *rpsL, rrs* and *gidB* genes

Extraction of DNA was done by heating at 95ºC for 15 minutes followed by overnight freezing. The two steps were performed twice. Next, equal amount of TE buffer (10 mM Tris-HCL pH 8.0, 1mM EDTA pH 8.0) was added and stored at -20ºC until use.

PCR amplification and DNA sequencing was carried out using primers previously described (Table 1) [71]. The purified amplified DNA was sequenced with Big Dye Terminator v3.1 (Thermo Fisher Scientific, Waltham, MA, USA) using a 3500 Genetic sequencer. Sequence results were aligned to H37Rv (NC_000962.3) using BioEdit sequence alignment editor software [72]. In addition, RIF and INH resistance conferring mutations obtained in a previous study were used to confirm MDR [60].

Spoligotyping

Spoligotyping was conducted as previously described (figure 5) [44,58]. Briefly, the direct repeat region was amplified by PCR using primers DRa GGTTTTGGGTCTGACGAC and DRb CCGAGAGGGGACGGAAAC, and the resulting products were hybridized to 43 spacer-specific oligonucleotides probes on a membrane. The resulting spoligotype pattern was converted to binary and octal formats. The later format was compared with SpolDB4 to determine the spoligo-international type (SIT) and spoligotypes [73].

Statistical analysis

Chi square (X^2) and Fisher's exact test were used to determine association between mutations and genotypes. Odds ratio and simple proportions were used to describe the data. Statistical significance was set at p value <0.05. Statistical software IBM SPSS 26 was used for analysis.

Results

Correlation between Genotypes and STR resistance

The 138 MDR-*Mtb* analyzed included 69 (50%) isolates belonging to the LAM family, 34 (24.6%) to CAS family, 19 (13.8%) to T family, 6 (4.3%) to H family, 7 (5.1%) to X family, 2 (1.4%) to EAI and 1 (0.7%) to S family (Table 2). STR resistance was observed in 65.9% (91/138) of the MDR-TB isolates. Out of all the genotypes analysed, CAS1_Kili (SIT 21) clade had higher odds of having STR resistance compared to non-CAS1_Kilispoligotype clades (OR 3.9, *p* value = 0.009), while LAM1 (SIT 20) had lower odds of having STR resistance compared to non-LAM1 isolates (SIT 20) (OR 0.12, *p* value = 0.0005).

Mutations in *rpsL*

Three types of amino acid substitutions (K43R, K88Q and K88R) associated mutations were identified in *rpsL* in 33% (30/91) of STR resistant MDR-TB isolates (Table 3). K43R was observed in 70% of isolates with mutations in *rpsL* followed by K88R at 26.7% and then K88Q in 3.3%. A silent mutation C117T (T39T) was observed in 2 isolates.

Mutations in *rrs*

Eleven STR resistant isolates harbored 4 different types of mutations in *rrs* including A514C, C517T, A906G and A907C (Table 3). Mutations in *rpsL* and *rrs* were mutually exclusive and observed in resistant isolates only.

Mutations in *gidB*

Fifteen different nucleotide substitutions and 5 deletions were observed in *gidB* in 45 resistant and 12 susceptible isolates (Table 3). Ten mutations were exclusively found in resistant isolates. Three silent mutations at positions C12T (I4I), G211C (V110V) and A615G (A205A) were detected. STR resistant isolates with *gidB* amino acid substitution G71R and A183V were found to harbor additional mutations in *rrs* (A906G and C517T) respectively. A mutation involving two nucleotide substitution (T611A; C612A), producing a stop codon at 204, was observed in 5 isolates.

The nucleotide polymorphism T47G (L16R) was observed exclusively in LAM genotype while A615G (A205A) was observed in CAS and EAI genotypes. Excluding the genotype specific L16R, mutations with amino acid substitutions in *rpsL* and *gidB* were mutually exclusive.

The proportion of STR resistant isolates with mutations in *gidB* was 49.5%. Eight STR resistant isolates did not have mutations in any of the analyzed genes.

Association of genotype and mutations in *rpsL, rrs* and *gidB*

A significant correlation was found between CAS1_Kili genotype and a mutation at codon 43 in *rpsL* (OR 17.6, *p* value <0.0001*)* (Table 4). Mutations in *gidB* were significantly correlated with LAM genotype (*p* <0.05). Clustering of isolates based on SIT number and STR resistance conferring mutations was observed (figure 7).

Molecular determination of STR resistance

The sensitivity of sequencing prediction of STR resistance using *rpsL* was 33% while that of *rrs* was 12.1% and *gidB* was 49.5% (Table 5). All 3 genes had an improved combined sensitivity of 91.2% and 74.5% specificity.

Discussion

STR was used as a first-line drug in Zambia until late-1990s [74,75]. In the early 2000s, it was recommended for the management of retreatment cases [53]. The prolonged use of STR as a first-line drug in new and retreatment cases could have led to the emergence of resistance. This is the first study to describe mutations conferring STR resistance in MDR-*Mtb* and their relationship with the circulating genotypes in Zambia.

I observed a high proportion (65.9%) of STR resistance among the MDR-*Mtb* in this study, which was lower than that observed in Uganda (83.9%) [76]. The high proportion of STR resistance in Uganda was estimated from a national surveillance program and thus reflects the national prevalence. In this study however, isolates were sampled from a limited population. This signifies the importance of rapidly determining STR susceptibility profile before its inclusion in the individualized MDR-TB regimen in Zambia.

Majority of isolates I used belonged to the LAM family (50%) with LAM11_ZWE (SIT 59) clade being the predominant, followed by LAM1 (SIT 20). Thus, LAM11_ZWE remains the major circulating *Mtb* genotype in Zambia [77,78]. The second most dominant genotype was CAS1_Kili (SIT 21) clade (24.6%). Previous studies had reported few isolates belonging to LAM1 (SIT 20) and CAS1_Kili (SIT 21) clades, suggesting a potential recent expansion of these strains. Compared to other clades, CAS1_Kili (SIT 21) had higher odds of STR resistance (OR 3.9, *p* value =0.009) (Table 2). Conversely, LAM1 (SIT 20) clade was less likely to develop STR resistance (OR 0.12, *p* value =0.0005).

I found mutations in *rpsL* were exclusively in 33% of resistant isolates. Mutations in this gene account for more than 65% of resistant isolates in regions with a high prevalence of MDR-*Mtb* Beijing genotype (Lineage 2), such as Myanmar and China [71,79]. In contrast to these regions, Lineage 4 is predominant in Zambia, particularly LAM genotype. The results were consistent with reports from regions where the Beijing genotype circulation in the population is low, and *rpsL* mutations among MDR-*Mtb* isolates account for less than 50% of STR resistance [67,80,81].

Notably in my present study, amino acid substitution K43R of *rpsL* was significantly associated with CAS1_Kili. Analysis of 7,346 of MDR, pre-XDR and XDR isolates publicly available in TB-Profiler web based database, revealed that Lineage 2 had high odds to acquire K43R amino acid substitution for STR resistance than non-Lineage 2genotypes (OR 6.42, *p* value <0.0001) (Table 6) [82]. However, this trend was not observed in the Lineage 3 to which CAS1_Kili belongs. Consequently, the observed high proportion of CAS strain harboring K43R in *rpsL* in this study indicates that, in addition to being a low fitness cost mutation [31], there is possibly clonal expansion of a *Mtb* strain with this mutation.

Mutations C517T and A906G in *rrs* have been reported in isolates with minimum inhibitory concentration (MIC) between 10 µg/ml and 100 µg/ml indicating that they sufficiently confer STR

resistance [83]. In my study results, an isolate with C517T mutation and two isolates with A906G mutation had additional mutations in *gidB.* One of the additional mutations occurred at codon 71 which is close to GidB active site and a mutation at this codon had been reported to have MIC of 20 µg/ml [84,85]. Thus, the additional mutations may lead to increased level of STR resistance. Mutations in *rrs* accounted for 12.1% of STR resistance, which is similar to China and South Korea but lower than Panama and Russia (Table 7) [86,87].

Mutations in *gidB* encoding a methyltransferase result in non-methylation of G518 nucleotide of *rrs* and disruption of STR binding leading to low level resistance [88]. Though *gidB* is not a very long gene, its active site is made up of 16 codons, likely increasing the mutational target size [84] and leading to the observed polymorphic mutations. Among the mutations that I found; 19 mutations had been previously described while 4 were novel. I found the mutation C447G (S149R) previously detected in a resistant isolate with MIC of 10 µg/ml [83], in one susceptible isolate. Among the novel mutations found in this study included 25_88del (64 base pair deletion), C548T (A183V), 575_576delGC and a double mutation T611A/C612A (204stop). Two of these (25_88del and 575_576delGC) were exclusively detected in STR resistant isolates confirming their role in conferring resistance. The double mutation T611A/C612A (204stop) was found in 4 resistant isolates and 1 susceptible isolate. The presence of resistance conferring *gidB* mutations in some susceptible isolates indicate that the MGIT 960 drug susceptibility test method may misclassify low level STR resistance. This can lead to inclusion of STR in a regimen against an otherwise STR resistant *Mtb* strain. These mutations, especially deletions and nonsense mutations, are expected to cause conformational changes in GidB affecting its functionality [84]. Adoption of molecular testing would improve the detection and management of STR resistant *Mtb*.

The non-synonymous mutation at codon 16 in *gidB* (L16R) has been associated with the LAM family in concordance with this study (p value <0.0001) [89]. The CAS and EAI families had a *gidB* synonymous mutation A615G (A205A) observed in all isolates belonging to these families. This synonymous mutation has been reported in other lineages suggesting that nucleotide A at the position 615 in *gidB* in H37Rv reference genome is a polymorphism specific to Lineage 4 [47,89].

A synonymous (V110V) mutation and a non-synonymous (G71R) mutation were detected exclusively in isolates belonging to EAI genotype in this study. The V110V synonymous mutation has been described as an EAI genotype polymorphism [89]. In contrast, G71R may be a resistance conferring genotype-independent mutation [47].

Clusters of 2 or more isolates were observed based on *rpsL*, *rrs*, *gidB*, *katG*, *rpoB* mutations and spoligotype SIT (figure 7). Five clusters were seen in CAS1_Kili (SIT 21) with the largest cluster having 16 isolates and harboring K43R amino acid substitution in *rpsL*. It is postulated that CAS1_Kili (SIT 21) evolved in Tanzania and is dominant in the country particularly in the city of Dar es Salaam [90,91]. This coastal city is of great economic importance as a trade harbor between

Zambia and Indo-Pacific regions and could be the origin of CAS1_Kili found in Zambia. However, CAS1_Kili (SIT 21) clade in Tanzania has not been associated with drug resistance [92]. It possibly was introduced to Zambia as a STR susceptible clone before late-2000 when STR was being used as first-line drug for TB treatment and the acquisition of amino acid substitution K43R may have occurred in Zambia. It is possible that progression to MDR emerged independently in some patients subsequent to infection with a K43R STR resistant clone because these isolates had varied RIF resistance conferring mutations in *rpoB*. Four additional clusters based on mutations in *gidB* and *rpoB* observed in CAS1_Kili (SIT 21) suggest several clones of STR resistant MDR-*Mtb* CAS1_Kili (SIT 21) are spreading in Zambia (figure 7).

The spoligotype clade LAM1 (SIT 20) though not associated with STR resistance could be considered one cluster of 16 isolates having same spoligotype, *rpoB,* and *katG* mutations. Amino acid substitutions S450L in *rpoB* and S315T in *katG* were observed in 100% and 94.1% respectively, in the LAM1 (SIT 20) isolates (figure 7). The frequency of *rpoB* S450L and *katG* S315T mutations in this clade exceeds the global estimates of 65% and 72% respectively, among MDR-*Mtb* isolates [30,93]. These results support clonal expansion of a STR susceptible LAM1(SIT 20) clone. LAM1 is dominant in Angola and Namibia [94]. While there is limited information regarding *Mtb* drug resistance in Namibia, MDR-*Mtb* among retreatment cases was reported to be as high as 71% in Angola [95]. It is likely that an MDR-*Mtb* LAM1 clone susceptible to STR had been introduced in Zambia from Angola by displacement of people due to war. This could have been after STR was no longer used as first-line drug, hence the low STR resistance in this cluster.

Among isolates belonging to LAM11_ZWE (SIT 59) clade, clusters of 5 or more isolates harboring amino acid substitutions Y22H, V124G and F204stop in *gidB* were observed (figure 7). The substitution V124G may have emerged first, followed by acquisition of mutations conferring RIF resistance and are currently spreading in Zambia as STR resistant MDR-*Mtb*, since two types of *rpoB* mutations were observed in this cluster. Isolates with Y22H and F204stop were restricted to Lusaka province only. However, amino acid substitution V124G was found in isolates from Lusaka and Western provinces, with the largest cluster (5 isolates) being from Western province (Mongu district). This observation suggests that *Mtb* strain with V124G is disseminating to other areas from Western Province.

The clustering of isolates in this study was not unique to a specific genotype but was observed in almost all genotypes indicating that the increase of MDR-TB in Zambia is largely being driven by the spread of MDR-*Mtb* clones. I analyzed total number and proportions of different genotypes and clades with time to see if there was any variation in their isolation frequency (Table 8). Although the sample size of isolates from 2011 to 2013 was small, I observed a general increase in the isolation frequency in several clades over the study period. This further supports the multi-clonal expansion of MDR-TB strains in Zambia. Furthermore, one isolate belonging to a cluster harboring a *gidB* G37R

amino acid substitution was isolated from a 9-year-old patient, indicating recent transmission and supporting the hypothesis that MDR-TB is spreading in Zambia. The clustered clones had the same spoligotype SIT number, and *rpoB*, *katG* and *gidB* mutations (figure 7). Though these mutations in *rpoB* and *katG* are frequently observed, mutations in highly polymorphic *gidB* permits for the inference of clones even in the absence of spoligotype results. Outbreaks of MDR-*Mtb* characterized by mutations in *gidB* are not unprecedented [41]. Therefore, *gidB* sequencing maybe useful in checking MDR-*Mtb* transmission.

I calculated the sensitivity of mutations in *rpsL*, *rrs* and *gidB* to predict STR resistance to be 91.2% similar to a report from China (94.6%) but higher than that reported in Myanmar (83.7%), and South Korea (84.3%) [71,79,96]. However, the sensitivity of *rpsL* in this study was lower (32.6%) when compared to these countries (Table 7). This could be attributed to the difference in the genotypes circulating the different countries. Studies have shown that *gidB* mutations are significantly correlated with Lineage 4, while *rpsL* mutations are significantly correlated with Lineage 2 [70,97]. Accordingly, the observed higher sensitivity of *gidB* mutations (49.5%) compared to other genes in predicting STR resistance in this study could be attributed to high number of isolates belonging to Lineage 4 (102/138). In contrast, in China and Myanmar, majority of STR resistant isolates belonged to Lineage 2 (the Beijing family) supporting the observed association between *rpsL* mutations and Lineage 2. This indicates that genotype diversity of *Mtb* may influence the types of STR resistance conferring mutations [98].

A limitation of this study was that I did not include polyresistant and pan-susceptible isolates. I could not therefore confirm whether the observed mutations are also present in those groups of isolates in Zambia.

This study has highlighted the high level of STR resistance among MDR-*Mtb* in Zambia. Mutations in *rpsL*, *rrs* and *gidB* can adequately predict STR resistance in Zambia. Therefore, gene sequencing can be used to determine STR susceptibility before its inclusion on the personalized treatment regimen. This study has also shown that sequencing of drug-resistance conferring mutation can be used to profile drug resistance evolutionary history of isolates and detect transmission. In this study I also highlight that the rise in MDR-TB could be attributed to possible spread of several clones of MDR-*Mtb* with resistance to STR in Zambia. To halt the spread; development and adoption of a rapid molecular tool for STR resistance testing, prompt initiation of effective therapy and implementation of mask wearing by all TB patients in high-risk environments is recommended.

Summary

Streptomycin (STR) is recommended for the management of multidrug resistant tuberculosis (MDR-TB). STR resistance conferring mutation types and frequency are shown to be influenced by genotypes of circulating strains in a population. My aim was to characterize the mutations in MDR-TB isolates and examine their relationship with the genotypes in Zambia. From the analyzed 138 MDR-TB isolates stored at University Teaching Hospital TB Reference Laboratory in Zambia, STR resistance was observed in 65.9%. Mutations in *rpsL*, *rrs* and *gidB* accounted for 33%, 12.1% and 49.5%, respectively. Amino acid substitution K43R in *rpsL* was strongly associated with CAS1_Kili genotype ($p<0.0001$). Sequencing the three genes could predict 91.2% of STR resistance. Thus, DNA sequencing can be useful in rapidly determining STR resistance in Zambia. Sequencing and spoligotyping also revealed clustering of isolates, suggesting that the increase of STR resistant MDR-TB in Zambia is largely due to the spread of resistant strains owing to delayed diagnosis and inadequate treatment. Therefore, I recommend rapid detection of STR resistance using DNA sequencing before its use in MDR-TB treatment in Zambia.

Tables and figures

Table 1: Oligonucleotides used for DNA sequencing

Gene	Primer set	
rrs	Foward	GATGACGGCCTTCGGGTTGT
	Reverse	AGGCCACAAGGGAACGCCTA
rpsL	Foward	GGCCGACAAACAGAACGT
	Reverse	GTTCACCAACTGGGTGAC
gidB	Foward	CGCCGAGTCGTTGTGCT
	Reverse	AGCCTGGCCCGACCTTA

Table 2: Correlation between genotypes and phenotypic STR susceptibility

Spoligotype		Streptomycin Resistant		Streptomycin susceptible		Significance	
SIT	Clade	N=91	Proportion (%)	N=47	Proportion (%)	Odds ratio (95%CI)	*p* value
21	CAS1_Kili	29	31.9	5	10.6	3.9 (1.4 to 11.0)	0.009
59	LAM11_ZWE	21	23.1	8	17.0	1.5 (0.6 - 3.6)	0.4096
20	LAM1	4	4.4	13	27.7	0.1 (0.04 - 0.39)	0.0005
815	LAM11_ZWE	10	11.0	3	6.4	1.8 (0.5 - 6.9)	0.39
53	T1	8	8.8	4	8.5	1.0 (0.3 - 3.6)	0.96
Orphan	LAM11_ZWE	4	4.4	1	2.1		
137	X2	3	3.3	4	8.5		
52	T2	3	3.3	2	4.3		
Orphan	H1	2	2.2	2	4.3		
42	LAM9	1	1.1	2	4.3		
4	H3/T1	0	0.0	1	2.1		
Orphan	EAI	2	2.2	0	0.0		
34	S	1	1.1	0	0.0		
50	H3	1	1.1	0	0.0		
73	T	0	0.0	1	2.1		
317	T2	1	1.1	0	0.0		
811	LAM11_ZWE	1	1.1	0	0.0		
2173	LAM11_ZWE	0	0.0	1	2.1		

Table 3: Distribution of mutations in *rpsL*, *rrs,* and *gidB* among STR resistant and susceptible isolates

rpsL (DNA)	RpsL (protein)	*rrs*	*gidB* (DNA)	GidB (protein)	STR Resistant	STR Susceptible
A128G	K43R				21	0
A263G	K88R				8	0
A262C	K88Q				1	0
		A514C			3	0
		C517T			2	0
		C517T	C548T	A183V	1	0
		A906G	G211C	G71R	2	0
		A907C			3	0
			25_88del	9fs	1	0
			T64C	Y22H	3	2
			98delG	34fs	3	1
			G109A	G37R	3	1
			112delC	39fs	4	1
			C223T	P75S	1	0
			G227A	G76D	1	0
			T242C	I81T	1	0
			T298C	F100L	1	1
			347delG	117fs	5	2
			T371G	V124G	5	1
			C401A	A134E	1	0
			G412C	A138P	3	1
			C447G	S149R	0	1
			T455C	L152S	2	0
			G469C	G157R	2	0
			575_576delGC	193fs	2	0
			T611A, C612A	204 Stop	4	1
	No mutations				8	35

Table 4: Distribution of *rpsL, rrs,* and *gidB* mutations among different genotypes

rpsL (DNA)	RpsL (protein)	*rrs*	*gidB* (DNA)	GidB (protein)	CAS	LAM	H	T	S	X	EAI
A128G	K43R				16	2	1		1	1	
A263G	K88R					4	1	3			
A262C	K88Q					1					
		A514C				3					
		C517T				2					
		C517T	C548T	A183V		1					
		A906G	G330T	G71R							2
		A907C						3			
			25_88del	9fs				1			
			T64C	Y22H		5					
			98delG	34fs	4						
			G109A	G37R				4			
			112delC	39fs	5						
			C223T	P75S		1					
			G227A	G76D		1					
			T242C	I81T		1					
			T298C	F100L		2					
			347delG	117fs	3					4	
			T371G	V124G		6					
			C401A	A134E			1				
			G412C	A138P		4					
			C447G	S149R	1						
			T455C	L152S		2					
			G469C	G157R		2					
			575_576delGC	193fs	2						
			T611A, C612A	204 Stop		5					
	No mutations				3	27	3	8		2	
	Total				34	69	6	19	1	7	2

fs- frame shift, del- deletion

Table 5: STR resistance prediction by mutations in *rpsL, rrs,* and *gidB*

Correlation of drug resistance genotype and phenotype								
	Susceptible		Resistant		Sensitivity	Specificity	PPV	NPV
	Mutation	No mutation	Mutation	No mutation	(%)	(%)	(%)	(%)
rpsL	0	47	30	61	33.0	100	100	43.5
rrs	0	47	11	80	12.1	100	100	37.0
gidB	12	35	45	46	49.5	74.5	78.9	43.2
rrs or rpsL	0	47	41	50	45.1	100	100	48.5
rpsL or rrs or gidB	12	35	83	8	91.2	74.5	87.4	81.4

Table 6: The occurrence of K43R in various lineages among 7346 global MDR, pre-XDR, and XDR-*Mtb* isolates [82]

Lineages	# of isolates	# of isolates with K43R	Proportion	Odds ratio	95% CI	*p* value
L1	272	16	0.06	0.11	0.06-0.18	<0.0001
L2	3154	1827	0.58	6.42	5.78-7.15	<0.0001
L3	559	117	0.21	0.47	0.38-0.57	<0.0001
L4	3308	595	0.18	0.23	0.21-0.25	<0.0001
L5	44	9	0.21	0.48	0.21-0.95	0.048
L6	5	1	0.2	0.46		
L9	1	0	0	-		
M.bovis	3	2	0.67	3.73		

Table 7: Comparison of *rpsL*, *rrs,* and *gidB* mutations among MDR-*Mtb* streptomycin resistant isolates from various countries

Comparison of mutations among MDR-*Mtb* streptomycin resistant isolates in various countries						
Gene	Zambia Current study	Myanmar [71]	China [79]	Russia [87]	Panama [86]	South Korea [96]
	Proportions (%)					
rpsL	33.0	69.5	78.3	51	8.5	64.7
K43R	23.1	64.4	58.7	40.8	8.5	47.1
K88R	8.8	4	19.6	10.2	0	17.6
K88Q	1.1	0	0	0	0	0
rrs	12.1	3.4	9.8	44.3	66	9.8
rpsL or rrs	45.1	73	88	94.3	74.5	74.5
gidB	49.5	9.9	9.8	1.8	19	9.8
rpsl or rrs and/or gidB	91.2	83.7	94.6	94.6	93.6	84.3

This table shows the proportion of isolates with mutations in a specific gene or a combination of genes, as well as the proportion of a specific *rpsL* amino acid substitution among STR resistant in different countries.

Table 8: Frequency of isolation of different genotypes over time for the 138 MDR-*Mtb* isolates

Genotype Clade	2011	2012	2013	2014	2015	2016	2017 ~Apr.	Total
	Number of isolates (%)							
CAS1_KILI (SIT 21)		2 (5.9)	1(2.9)	4 (11.8)	8 (23.5)	14 (41.2)	5 (14.7)	34
LAM11_ZWE (SIT 59)	1 (3.4)	1 (3.4)	2 (6.9)	4 (13.8)	6 (20.7)	12 (41.4)	3 (10.3)	29
LAM11_ZWE (SIT 815)		1 (7.7)		4 (30.8)	3 (23.1)	5 (38.5)		13
LAM11_ZWE (SIT 811)				1 (100)				1
LAM11_ZWE (SIT 2173)							1 (100)	1
LAM11_ZWE (Orphan)				1 (20)		3 (60)	1 (20)	5
LAM9 (SIT 42)			1 (33.3)	1 (33.3)		1 (33.3)		3
LAM1 (SIT 20)		1 (5.9)	2 (11.8)	2 (11.8)	5 (29.4)	3 (17.6)	4 (23.5)	17
T1 (SIT 53)		1 (8.3)		2 (16.7)	3 (25)	5 (41.7)	1 (8.3)	12
T (SIT 73)							1 (100)	1
T2 (SIT 52)				1 (20)	2 (40)	1 (20)	1 (20)	5
T2 (SIT 317)						1 (100)		1
X2 (SIT 137)			1 (14.3)	2 (28.6)	2 (28.6)	2 (28.6)		7
S (SIT 34)				1 (100)				1
H3/T1 (SIT 4)				1 (100)				1
H3 (SIT 50)		1 (100)						1
H1 (Orphan)				1 (25)	1 (25)	2 (50)		4
EAI (Orphan)				2 (100)				2

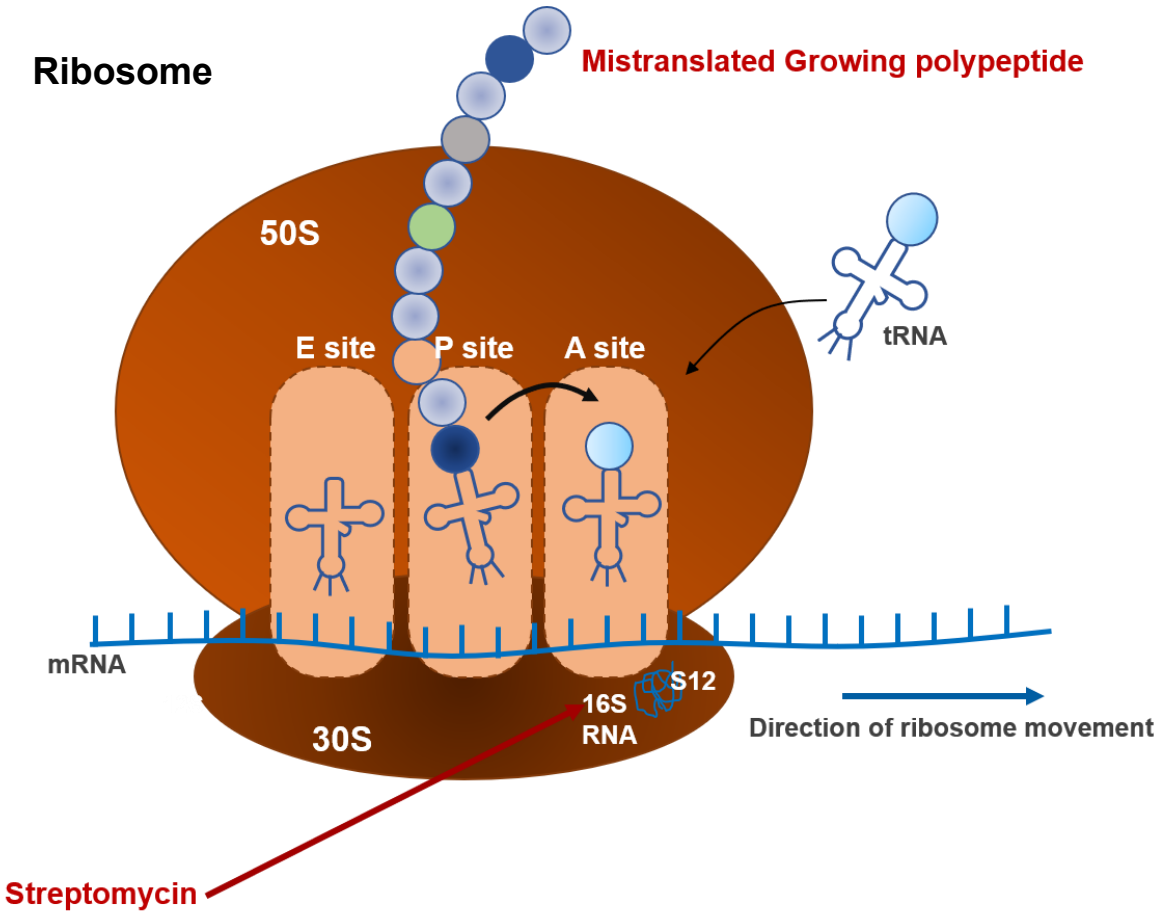
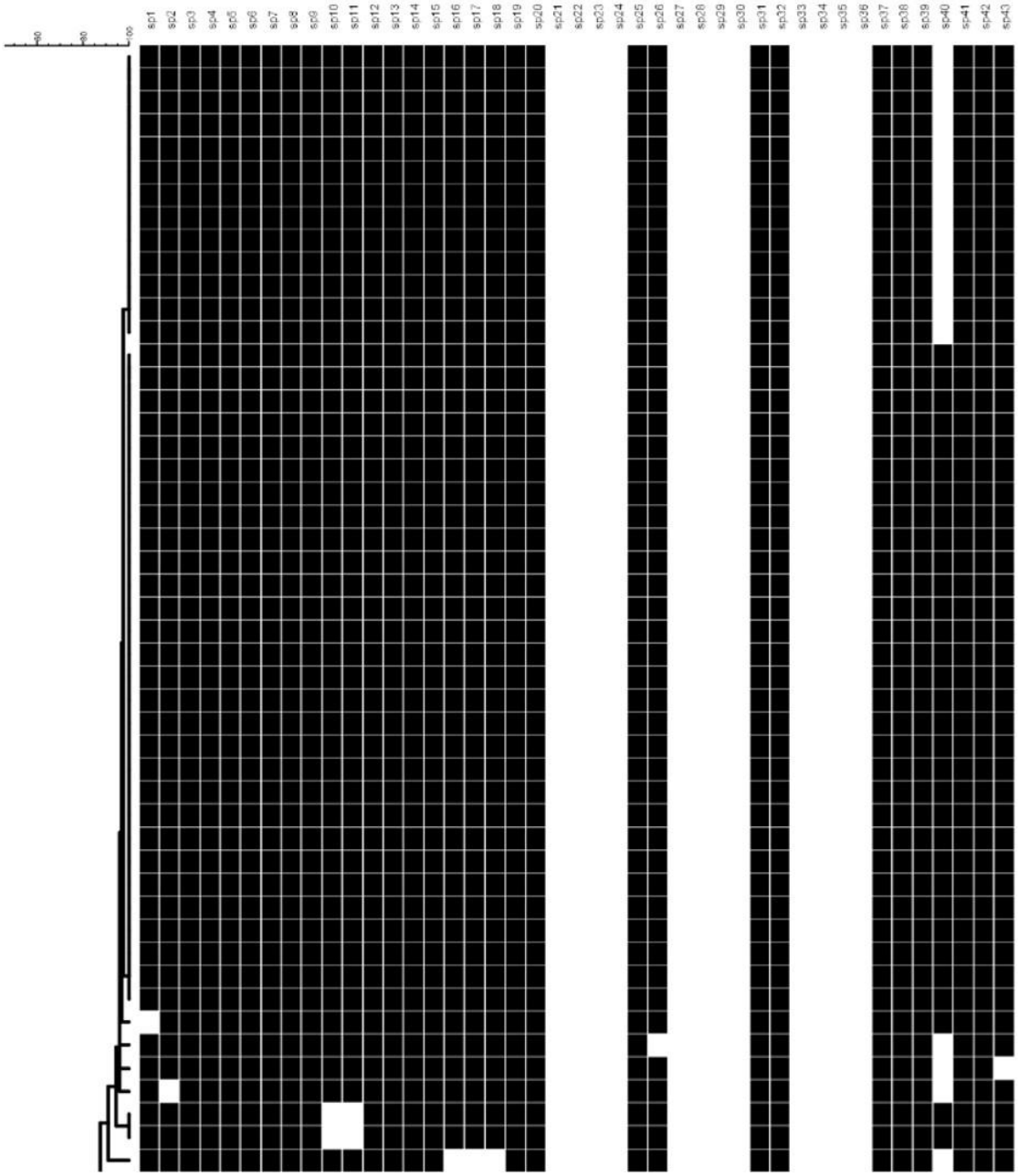
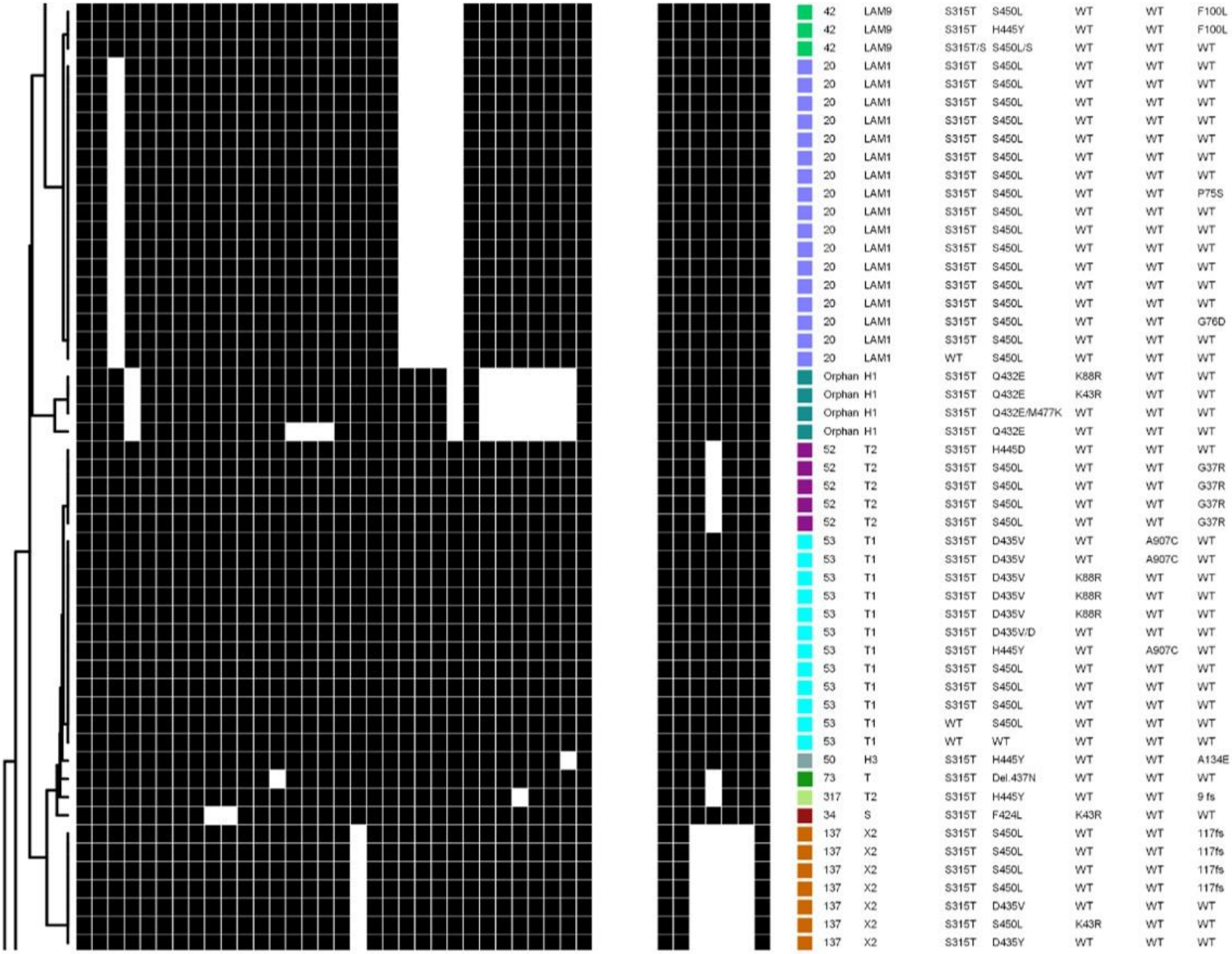


Figure 6: Mechanism of action of streptomycin. Streptomycin binds to 16S rRNA and protein S12 of 30S rRNA leading to mistranslation of mRNA [66]



SIT	Clade	*katG*	*rpoB*	*rpsL*	*rrs*	*gidB*
815	LAM11_ZWE	S315T	S450L	WT	WT	A138P
815	LAM11_ZWE	S315T	S450L	WT	WT	A138P
815	LAM11_ZWE	S315T	S450F	WT	WT	A138P
815	LAM11_ZWE	S315T	S450F	WT	WT	A138P
815	LAM11_ZWE	S315T	S450L	WT	WT	G157R
815	LAM11_ZWE	S315T	S450L	WT	WT	G157R
815	LAM11_ZWE	S315T	H445Y	T39T, K43R	WT	WT
815	LAM11_ZWE	S315T	S450L	WT	WT	WT
815	LAM11_ZWE	S315T	S450L	WT	A514C	WT
815	LAM11_ZWE	S315T	S450L	WT	C517T	WT
815	LAM11_ZWE	S315T	S450L	WT	C517T	WT
815	LAM11_ZWE	S315T	S450L	WT	WT	WT
815	LAM11_ZWE	WT	WT	WT	WT	WT
59	LAM11_ZWE	S315T	S450L	WT	C517T	A183V
59	LAM11_ZWE	S315T	S450W	WT	WT	V124G
59	LAM11_ZWE	S315T	S450L	WT	WT	V124G
59	LAM11_ZWE	S315T	S450L	WT	WT	V124G
59	LAM11_ZWE	S315T	S450W	WT	WT	V124G
59	LAM11_ZWE	S315T	S450W	WT	WT	V124G
59	LAM11_ZWE	S315T	S450W	WT	WT	V124G
59	LAM11_ZWE	S315T	L430R/D435Y	WT	WT	F204Stop
59	LAM11_ZWE	S315T	L430R/D435Y	WT	WT	F204Stop
59	LAM11_ZWE	S315T	L430R/D435Y	WT	WT	F204Stop
59	LAM11_ZWE	S315T	L430R/D435Y	WT	WT	F204Stop
59	LAM11_ZWE	S315T	L430R/D435Y	WT	WT	F204Stop
59	LAM11_ZWE	S315T	S450L	WT	WT	Y22H
59	LAM11_ZWE	S315T	S450L	WT	WT	Y22H
59	LAM11_ZWE	S315T	S450L	WT	WT	Y22H
59	LAM11_ZWE	S315T	S450L	WT	WT	Y22H
59	LAM11_ZWE	S315T	S450L	WT	WT	Y22H
59	LAM11_ZWE	S315T	D435V/D	K43R	WT	WT
59	LAM11_ZWE	S315T	S450L	K88R	WT	WT
59	LAM11_ZWE	S315T	S450L	K88R	WT	WT
59	LAM11_ZWE	S315T	S450L	K88R	WT	WT
59	LAM11_ZWE	S315T	E423A/S450W	K88R	WT	WT
59	LAM11_ZWE	S315T	S450L	WT	WT	WT
59	LAM11_ZWE	D329E	D435F	WT	WT	WT
59	LAM11_ZWE	S315T	D435V	WT	WT	WT
59	LAM11_ZWE	WT	S450L	WT	WT	WT
59	LAM11_ZWE	S315T	S450L	WT	A514C	WT
59	LAM11_ZWE	WT	WT	WT	WT	WT
59	LAM11_ZWE	WT	Del.	WT	WT	WT
2173	LAM11_ZWE	WT	WT	WT	WT	WT
811	LAM11_ZWE	S315T	S450L	K88Q	WT	WT
Orphan	LAM11_ZWE	S315T	S450L	T39T	WT	I81T
Orphan	LAM11_ZWE	S315T	S450L	WT	A514C	WT
Orphan	LAM11_ZWE	S315N	D435V/D	WT	WT	L152S
Orphan	LAM11_ZWE	S315N	D435V	WT	WT	L152S
Orphan	LAM11_ZWE	S315T	S450Q	WT	WT	WT



42	LAM9	S315T	S450L	WT	WT	F100L
42	LAM9	S315T	H445Y	WT	WT	F100L
42	LAM9	S315T/S	S450L/S	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	P75S
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	S315T	S450L	WT	WT	G76D
20	LAM1	S315T	S450L	WT	WT	WT
20	LAM1	WT	S450L	WT	WT	WT
Orphan	H1	S315T	Q432E	K88R	WT	WT
Orphan	H1	S315T	Q432E	K43R	WT	WT
Orphan	H1	S315T	Q432E/M477K	WT	WT	WT
Orphan	H1	S315T	Q432E	WT	WT	WT
52	T2	S315T	H445D	WT	WT	WT
52	T2	S315T	S450L	WT	WT	G37R
52	T2	S315T	S450L	WT	WT	G37R
52	T2	S315T	S450L	WT	WT	G37R
52	T2	S315T	S450L	WT	WT	G37R
53	T1	S315T	D435V	WT	A907C	WT
53	T1	S315T	D435V	WT	A907C	WT
53	T1	S315T	D435V	K88R	WT	WT
53	T1	S315T	D435V	K88R	WT	WT
53	T1	S315T	D435V	K88R	WT	WT
53	T1	S315T	D435V/D	WT	WT	WT
53	T1	S315T	H445Y	WT	A907C	WT
53	T1	S315T	S450L	WT	WT	WT
53	T1	S315T	S450L	WT	WT	WT
53	T1	S315T	S450L	WT	WT	WT
53	T1	WT	S450L	WT	WT	WT
53	T1	WT	WT	WT	WT	WT
50	H3	S315T	H445Y	WT	WT	A134E
73	T	S315T	Del.437N	WT	WT	WT
317	T2	S315T	H445Y	WT	WT	9 fs
34	S	S315T	F424L	K43R	WT	WT
137	X2	S315T	S450L	WT	WT	117fs
137	X2	S315T	S450L	WT	WT	117fs
137	X2	S315T	S450L	WT	WT	117fs
137	X2	S315T	S450L	WT	WT	117fs
137	X2	S315T	D435V	WT	WT	WT
137	X2	S315T	S450L	K43R	WT	WT
137	X2	S315T	D435Y	WT	WT	WT

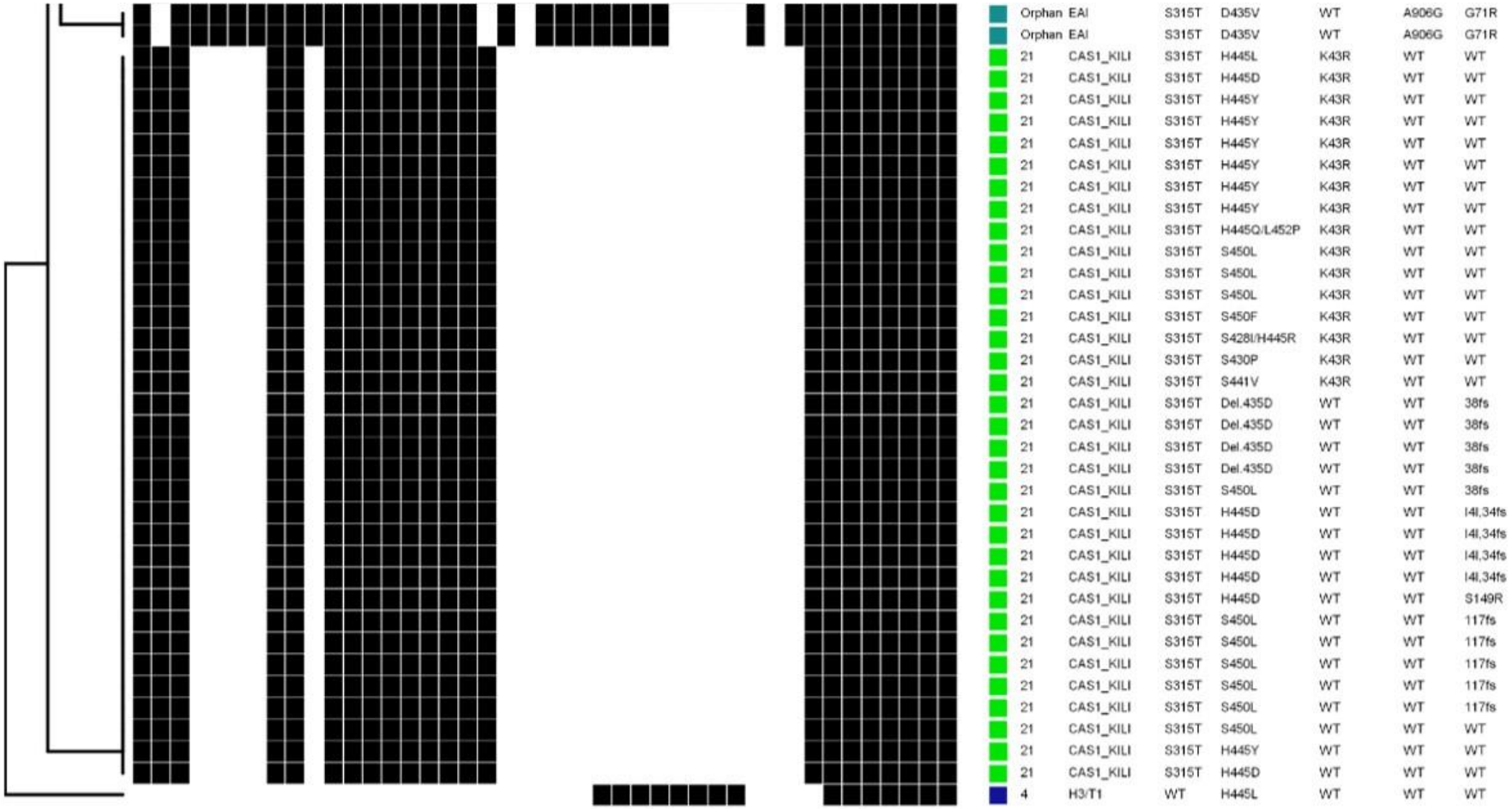


Figure 7: Phylogenetic tree based on spoligotype SITs and *katG, rpoB, rpsL, rrs*, and *gidB* mutations for 138 MDR-TB isolates

CHAPTER II

Characterization of *embB* mutations involved in ethambutol resistance in multi-drug resistant *Mycobacterium tuberculosis* isolates in Zambia

Introduction

The emergence and transmission of drug-resistant tuberculosis (TB) is a major obstacle to the ongoing global efforts of achieving an 80% reduction of TB incidences by 2030 [20]. In recent years, Zambia has seen an increasing trend of multi-drug resistant (MDR) TB and was recently included in the list of high MDR-TB burden countries in the world [55]. In 2019, an estimated proportion of 2.4% of new cases and 18% of previously treated cases had rifampicin resistance (RR)/MDR-TB [99]. An earlier study showed that the increasing cases of MDR-TB were due to local transmission of MDR-TB strains in Zambia [59]. Undiagnosed and unsuspected or diagnosed but inadequately treated MDR-TB patients are the likely source of transmission in Zambia [50]. To control the spread of MDR-TB in Zambia, active case finding, and adequate treatment is imperative. The adoption and implementation of rapid molecular-based diagnostic tools such as GeneXpert (Cepheid, Sunnyvale, CA) and Line Probe Assay (Hain Lifescience GmbH, Nehren, Germany) have improved MDR-TB case detection and subsequent treatment. However, Zambia has not yet adopted the use of molecular tools for resistance testing for some drugs used in MDR-TB treatment.

Ethambutol (EMB) is an integral part of the first-line drug regimen as well as for short course MDR-TB regimen and can be included in individualized longer regimens for MDR-TB treatment depending on drug susceptibility testing (DST) results [23]. EMB inhibits arabinosyltranferases EmbC, EmbA, and EmbB involved in the synthesis of cell wall components and subsequently compromising the cell wall integrity leading to cell death [100]. The EmbA and EmbB are involved in the synthesis of arabinogalactan while EmbC is involved in the synthesis of lipoarabinomannan (figure 2) [5]. Arabinogalactan has a vital role in linking mycolic acids to the peptidoglycan layer and mycolic acids are essential for mycobacterial cell impermeability and viability [5]. Thus, arabinogalactan is essential for *Mycobacterium tuberculosis* (*Mtb)* viability. Lipoarabinomannan is an important *Mtb* virulence factor that prevents phagosome maturation, mycobacterium-induced macrophage apoptosis, interferes with dendritic cell maturation, and induces the release of immunosuppressive cytokine [101].

Resistance to EMB has been attributed to mutations in the *embCAB* locus encompassing 3 contiguous genes *embC*, *embA,* and *embB* [102–104]. The *embB* gene mutations have the predominant role accounting for 54% to 93.7% of EMB resistance, particularly at codons 306, 406, and 497, which are considered hotspot resistance codons [100,103,105,106]. Codon 306 was shown to be directly involved in EMB binding while codons 406 and 497 are not directly involved.

Nevertheless, mutations at codon 497 cause conformational changes that affect codon 327, one of the EMB binding sites. Codon 406 mutations may also affect drug binding by causing protein conformation changes [100].

Despite the documented evidence of *embB* involvement in EMB resistance, there is an apparent discord with conventional phenotypic DST. The high EMB critical concentration (5.0 μg/ml) of MGIT 960 categorizes low-level EMB resistance as susceptible [107]. In addition, phenotypic DST is considered unreliable and unreproducible, thus WHO recommends molecular detection of resistance for EMB [108]. Therefore, it is important to investigate the mutations responsible for EMB resistance in order to develop the strategy to use molecular-based EMB DST in Zambia. In this study, I describe *embB* mutations involved in EMB resistance among MDR-TB isolates and evaluate the concordance with phenotypic DST in Zambia.

Materials and Methods

Samples and phenotype drug susceptibility testing

Mtb isolated from patient samples referred to The University Teaching Hospital TB Reference Laboratory between January 2014 to April 2017 were included in this study. The DST was done as part of the routine testing for rifampicin (RIF), isoniazid (INH), streptomycin (STR), and EMB at a critical concentration of 1.0 µg/ml, 0.1 µg/ml, 1 µg/ml, and 5.0 µg/ml, respectively using the MGIT M960 liquid culture systems following manufacturer's instructions (BD BACTEC™ MGIT™ 960 SIRE kit). A total of 132 MDR-TB isolates were randomly selected. The isolates information was extracted from the Laboratory information system.

DNA extraction

DNA was extracted by the boiling method as previously described [109]. The extracted DNA was stored at -20^{o}C until use.

DNA sequencing

The *embB* gene was amplified using the primers *embB*-F (5'- CGACGCCGTGGTGATATTCG-3') and *embB*-R (5'- CGACGCCGTGGTGATATTCG-3'). The PCR reaction volume of 20 µl contained DDW, 5x Go Tag buffer green (Promega Corp, Madison, WI, USA), 25 mM dNTP (Promega Corp), 25 mM MgCl, 5 M betaine, 10 µM primers, and GoTaq DNA polymerase (Promega Corp). The amplified product was purified using ExoSAP-IT™ Express PCR product cleanup (Thermo Fisher Scientific Inc., Santa Clara, USA) as instructed by the manufacturer. Purified DNA was sequenced using the BigDye Terminator V3.1 (Thermo Fisher Scientific Inc., Waltham, MA, USA) on an ABI 3500 genetic analyzer. Bioedit software was used to align the sequences to the H37Rv reference sequence (NC_000962.3) [72].

Spoligotyping

Spoligotyping was done as previously described using the DRa and DRb primers (figure 5) [44][58] and the resulting hybridized spoligotype pattern was converted to the binary code and compared to SpolD4 database for determination of the Spoligo-International Type (SIT) and spoligotypes [73].

Data analysis

The data was described using proportions and the odds ratio was used for statistical analysis. A two-tailed *p* value was used, and significance was set at <0.05. Sensitivity and specificity for *embB* sequencing method were calculated by comparing to MGIT 960 DST as the reference standard.

Results

Frequency of *embB* mutations in MDR-TB isolates

The analysis of phenotypic DST results showed that 46.2% (61/132) isolates were EMB resistant. Sequencing analysis of *embB* revealed mutations in 75.8% (100/132) of MDR-*Mtb* isolates. Among EMB resistant isolates, 86.9% (53/61) had mutations in *embB* gene. Compared to isolates with only INH and RIF resistance, isolates with phenotypic resistance to 3 or more drugs had higher odds of having *embB* mutations (Table 9).

I observed a total of 14 single nucleotide mutations resulting in 11 amino acid substitutions in *embB*. Codon 306 was the most mutated, accounting for 82% (82/100) of the isolates. Amino acid substitution M306I was the most predominant and found in 42% (42/100) of the isolates. Among isolates with mutations leading to M306I (G918A, G918C, G918T) amino acid change, a transition mutation G918A was found in 38 of the 42 isolates. The second dominant amino acid change was M306V and was found in 35% (35/100) of isolates, followed by Q497R and M306L detected in 6% (6/100) and 5% (5/100) of isolates, respectively. One mutation G982T (D328Y) and a double mutation G1215C and G1225C (E405D and A409P) were exclusively found in EMB resistant isolates. The remaining mutations were observed in either susceptible isolates only or both susceptible and resistant isolates. Mutations at codon 306 (*embB306*) were detected in both resistant and susceptible isolates with no significant difference (Table 12). Codons 497 and 406 were mutated in equal proportion. Table 10 summarizes the mutations detected among MDR-*Mtb* isolates in this study.

Occurrence of *embB* mutations in different genotypes

Spoligotyping revealed 7 major genotypes (Table 13). Among these genotypes, CAS1_Kili had high odds of acquiring *embB* mutations (OR 20.9, *p* value = 0.004) compared to the largest genotype in this study (LAM). Stratification of spoligotype SITs and *embB* gene mutations revealed 12 clusters of isolates (Table 11). The largest cluster had 26 isolates belonging to CAS1_Kili (SIT 21) clade and harboring G918A (M306I) mutation. One isolate of CAS1_Kili (SIT 21) clade had wildtype *embB* gene. I identified 6 clusters having mutation A916G (M306V). The 6 clusters included 13 isolates belonging to LAM11_ZWE (SIT 59); 6 isolates to LAM11_ZWE (SIT 815); 4 isolates each to T1 (SIT 53), T2 (SIT 52), and X2 (SIT 137); and 2 isolates to LAM1 (SIT 20). Mutation A916G (M306V) was found in only one isolate of CAS1_Kili (SIT 21) clade. CAS1_Kili (SIT 21) had significantly high odds of harboring amino acid substitution M306I in *embB* gene compared to the largest genotype of LAM (OR 14.9, *p* value <0.0001).

Discussion

In Zambia, routine phenotypic DST for EMB is performed using the MGIT 960 culture system. However, this method is considered unreliable and unreproducible [108]. Consequently, the external quality assurance for EMB DST often performs poorly. Unreliable results lead to insufficient treatment of patients which can drive emergence, transmission, or amplification of drug resistance. Detection of mutations in *embCAB* locus is used to infer resistance to EMB, but *embB* mutations account for the majority of isolates. The WHO recommends mutation analysis for inference of EMB resistance over phenotypic testing [108].

In this study, I revealed phenotypic EMB resistance in 46.2% of MDR-TB isolates whereas sequencing found *embB* mutations in 75.8% of the total isolates. In Kuwait, EMB resistance in MDR-TB was found in 44.1% of the isolates, while *embB* mutations were detected in 81.7% of the total MDR-*Mtb* isolates [110]. These results show that fewer isolates are determined as EMB resistance by phenotypic DST as compared to *embB* mutation analysis. Safi and colleagues demonstrated that *embB* mutations are involved in EMB resistance and they raise EMB MIC, albeit modestly [111,112]. For some isolates to gain high-level EMB resistance, the acquisition of additional mutations in other genes such as *ubiA* and *embC* is required [113]. Thus, *embB* gene mutations are considered as the initial step to acquiring high-level EMB resistance and should be treated as clinically resistant isolates, although additional studies linking mutations to clinical outcomes would be needed.

Previous studies have revealed that *embB* mutations are significantly associated with resistance to RIF, INH STR, and/or EMB [114,115]. In agreement with these findings, I also found that mutations in *embB* were more likely to occur in isolates with additional resistance to RIF, INH, STR, and EMB when compared to isolates with phenotypic resistance to only RIF and INH. This shows that *embB* mutations predispose to drug resistance amplification [116] and underscores the need for adopting a more reliable and rapid method of EMB resistance testing to receive appropriate treatment.

Among the three amino acid changes at *embB306* (M306I, M306L, and M306V), M306V was associated with EMB resistant isolates (OR 6.4, p value = 0.0002) while M306I was associated with EMB susceptible isolates (OR 2.9, p value = 0.0121). This conformed to the previous results from an allelic exchange experiment that showed that the mutations G918A and G918C producing amino acid change M306I, raises MIC close to the break-point of EMB resistance (5 to 7.5 µg/ml) [111]. Therefore, M306I would more likely appear among susceptible isolates in the MGIT 960 system which has a critical concentration value of 5 µg/ml for resistance determination. Among the three nucleotide substitutions leading to amino acid change M306I observed in this study, a transition mutation G918A was more frequent (90.5%) than the transversions G918C and G918T. This disproportionately high occurrence of transition mutation at this codon can be explained in part by the translation bias previously described in the genome of *Mtb*, wherein, ATG>ATA translation was 1.8 times more frequent than the transversions ATC and ATT [117]. In addition, I hypothesize that the

high frequency of the transition mutation seen in this study compared to the reported transition:transversion ratio, could reflect clonal expansion.

Mutation G1217A leading to amino acid change G406D was seen only in susceptible isolates. Nonetheless, this mutation had been proven to raise EMB MIC by 5 fold in a previous study and thus should be considered significant in eventual evolution to high-level EMB resistance [112,113].

Mutations at *embB*306 account for the majority of mutations in *embB* with an estimated global frequency of 47.5% among MDR-TB isolates, followed by codon 406 at 11.3% and then codon 497 at 7.9%, respectively [82] (Table 12). In Tanzania, Mexico, and South Korea, where the burden of MDR-TB is low, *embB306* mutations were found in 20.8%, 27.8%, and 38.5% of MDR-TB isolates, respectively [20,92,115,118]. In high MDR-TB burden countries of South Africa, Thailand, and China, *embB306* mutations were detected in 60%, 50%, and 30.3% of MDR-TB isolates, respectively [119–121]. In Russia, a high MDR-TB burden country, *embB306* mutations were detected in 30.7% of phenotypically determined MDR-TB isolates [87]. In South Africa where the frequency of *embB306* mutation was high, most isolates were clustered MDR-*Mtb* isolates. In my study, I found mutations at this codon in 62.1% of the MDR-TB isolates, higher than the global estimate and the frequency reported in high MDR-TB countries, but similar to that reported in South Africa suggesting the clonal expansion of EMB resistant MDR-TB isolates in Zambia. Mutations at codons 406 and 497 were both observed at a frequency of 4.5% and were below global frequency.

I found that 86.9% of phenotypically EMB resistant isolates and 66.2% of EMB susceptible isolates had *embB* mutations. Another study using the MGIT M960 method for phenotypic testing found *embB* mutations in 73.1% of EMB susceptible MDR-TB isolates [110]. Using the Lowenstein Jensen (LJ) proportion method, studies from South Korea, Poland, China, and Thailand found *embB* mutations in 30%, 42.5%, 45%, and 45.5%, respectively, of EMB susceptible MDR-TB isolates [105,106,118,120]. The MGIT 960 culture system was previously shown to produce the lowest agreement (77.1%) with sequencing, as compared to the LJ proportion method (81.4%) which has a critical concentration of 2 µg/ml and the microtiter alamarBlue assay (MABA) (84.7%). In my study, the sensitivity of *embB* mutations was 86.9% but the specificity was very low at 33.8% (Table 14). The poor specificity of sequencing in this study, I suppose is caused by the limitation of the phenotypic testing method. The LJ proportion method slightly improves EMB resistance detection as seen from previous studies. However, phenotypic DST is not reproducible and is unreliable, thus not recommended by WHO [108]. Therefore, the reliance on phenotypic testing alone for EMB DST in Zambia would fail to detect resistance in a considerable number of MDR-TB patients and expose these patients to inadequate treatment and promote the spread of MDR-TB. I recommend the adoption of the DNA sequencing method for more accurate results of EMB susceptibility in Zambia.

Mutations in *embB* were significantly associated with CAS1_Kili (SIT 21) in this study. In fact, the largest cluster of 26 isolates identified in this study belonged to CAS1_Kili (SIT 21) clade and

had a G918A (M306I) transition mutation (Figure 8). The size of this cluster suggests clonal expansion and may reflect increased transmissibility of CAS1_Kili (SIT 21) in Zambia. Previous reports have associated CAS1_Kili (SIT 21) with MDR-TB and STR resistance in Zambia [58,109]. The association of CAS1_Kili (SIT 21) with drug resistance and increased transmission in Zambia, makes this genotype a major concern and should be prioritized for tracking and identification of hotspot regions of transmission. The second largest cluster belonged to LAM11_ZWE (SIT 59) followed by LAM11_ZWE (SIT 815) both having M306V amino acid change. In addition, I also observed several smaller clusters, and these may have the potential to expand. This is suggesting multi-clonal transmission events happening in Zambia.

Interestingly, I found that CAS1_Kili (SIT 21) high odds to have the M306I amino acid change (OR 14.9, *p* value < 0.0001), with the odds of having M306V being 0.066 (*p* value = 0.0097) when compared to the major LAM genotype. To check whether this observation was true for global isolates, I analyzed 7,294 MDR-TB isolates from 4 major genotypes from TB profiler (Lineage 1 to Lineage 4). The results revealed that Lineage 3 to which CAS1_Kili belongs had 1.8 times odds of having M306I amino acid change (*p* value < 0.0001) and lower odds of having M306V amino acid change (OR 0.59, *p* value < 0.0001) compared to non-Lineage 3 for EMB resistance (Tables 15 and 17) [82]. Conversely, Lineage 2 showed opposite tendencies by preferring valine (OR 1.7, *p* value = 0.0004) with lower odds of acquiring isoleucine compared to non-Lineage 2 (OR 0.8, *p* value < 0.0001). For Lineage 4, M306I and M306V had equally high odds of being found in this genotype (OR 1.52, *p* value <0.0001) (Tables 15 and 17). Since Lineage 4 has similar odds of harboring M306I and M306V, I use it as a reference standard to compare occurrence of the Isoleucine and Valine in Lineages 1, 2, and 3. I found that compared to Lineage 4, Lineage 3 still had high odds of having M306I (OR 1.617, *p* value <0.0001) and Lineage 2 still had high odds of having M306V (OR 1.619, *p* value <0.0001) (Tables 16 and 18). Although both isoleucine and valine are hydrophobic amino acids with only a methyl group difference, the substitution of methionine with valine at *embB306* produces moderate to high-level of EMB resistance compared to isoleucine which produces low to moderate-level resistance [111]. The observed *embB306* amino acid replacement preferences of Lineage 2 and Lineage 3 may be due to genetic background differences that may favor a specific amino acid. Noteworthy, however, is the fact that these 2 amino acids offer differing degrees of resistance to EMB, with M306V leading to intermediate level resistance while M306I causing low level resistance. It would be interesting to explore any impact this *embB306* site-specific amino acid substitution predilection may have on transmissibility and virulence of Lineage 2 and Lineage 3.

The primary limitation of this study was that I could not perform MIC tests to correlate with detected mutations. In addition, I inferred clustering and transmissibility from spoligotyping and *embB* mutations only. This may overestimate clustering due to low sensitivity. I also did not sequence other genes such as *ubiA*, *embA* and *embC* known to contribute to EMB resistance.

In conclusion, my study highlights the high number of MDR-TB cases with mutations in *embB*, undetected by the MGIT 960 culture system. These mutations were associated with resistance to 3 or more drugs and can predispose progression to high-level EMB resistance. Thus, isolates with *embB* mutations should be considered clinically resistant to EMB. The consequences of inaccurately determining EMB susceptibility for the patient include delayed initiation of effective treatment, which in turn will lead to prolonged treatment, poor adherence to treatment, financial burden, prolonged morbidity, and poor prognosis. The delayed diagnosis and treatment will also affect the National TB Control Program by increasing the number of loss to follow-up cases, amplification of resistance to other drugs, the transmission of drug-resistant strains, poor treatment success rate, and the overall cost of MDR-TB management. I therefore recommend the adoption of genotypic testing to improve EMB resistance detection and management of MDR-TB patients. Genotype CAS1_Kili (SIT 21) was associated with *embB* mutations and had a large cluster of isolates having M306I mutation. This suggests increased transmission and I recommend tracking this genotype, as well as further investigation to determine hot spot areas of transmission for optimized interventions. Finally, an important finding in this study is the *embB306* site-specific amino acid preference for substitution in Lineage 2 and Lineage 3, suggesting possible influence of Lineage-specific genetic background on *Mtb* drug resistance evolution.

Summary

In this study I investigated *embB* mutations among MDR-TB isolates and analyzed their correlations with phenotypic drug susceptibility testing (DST) in Zambia. I used a total of 132 MDR-TB isolates collected from January 2014 to April 2017 and archived at UTH TB laboratory, Zambia. I utilized MGIT 960 systems, *embB* sequencing, and spoligotyping for this study.

Out of 61 phenotypically EMB resistant isolates, 53 had mutations in *embB*. Among EMB susceptible isolates, 47 had *embB* mutations. Sensitivity of *embB* mutations was 86.9% while specificity was 33.8%. CAS1_Kili (SIT 21) had high odds of having *embB* mutations, particularly, G918A (M306I) (OR 16.7, p value < 0.0001). Additionally, Lineage 2 and Lineage 3 showed codon 306 amino acid substitution preference of valine (M306V) and isoleucine (M306I), respectively.

I found that molecular EMB resistance testing can greatly improve the detection of EMB resistance among MDR-TB patients in Zambia compared to the unreliable phenotypic testing being used currently. Clusters of isolates with same mutation and the same spoligotype SIT point to the transmission of EMB resistant strains, I, therefore, recommend a future detailed investigation to track and determine transmission hotspot area for MDR-TB could help optimize control strategies. I also described the unique site-specific amino acid preferences of Lineages 2 and 3 suggest the possible influence of genotype diversity on drug resistance evolutionary trajectory. To the best of my knowledge, I was the first to describe this phenomenon for *embB* locus. A further study should be designed to explore the impact of this amino acid preference on cell wall components and virulence due to this amino acid preference between the two Lineages.

Tables and figure

Table 9: Drug resistance profiles and demographic characteristics of the MDR-TB isolates

Drug resistance profile	*embB* mutations	No *embB* mutations	Odds ratio	95% CI	*p* value
INH, RIF, EMB, STR	38	3	19	4.58-78.76	<0.0001
INH, RIF, EMB	15	5	4.5	1.24-16.35	0.022
INH, RIF, STR	37	9	6.17	2.09-18.20	0.001
INH, RIF	10	15	-	-	-

INH-isoniazid, RIF-rifampicin, EMB-ethambutol, STR-streptomycin

Table 10: Mutations detected in 132 MDR-*Mtb* isolates

embB (DNA)	EmbB (Protein)	Resistant	Susceptible	Proportion N (%)
G918A	M306I	15	23	38 (28.8)
G918C	M306I	0	2	2 (1.5)
G918T	M306I	1	1	2 (1.5)
A916T	M306L	1	4	5 (3.8)
A916G	M306V	28	7	35 (26.5)
A956C	Y319S	1	1	2 (1.5)
G982T	D328Y	1	0	1 (0.8)
C1204G	L402V	1	1	2 (1.5)
G1215C/G1225C	E405D/A409P	1	0	1 (0.8)
G1217C	G406A	1	2	3 (2.3)
G1217A	G406D	0	3	3 (2.3)
A1490G	Q497R	3	3	6 (4.5)
WT	WT	8	24	32 (24.2)
Total		61	71	132

Table 11: Distribution of *embB* mutations in different *Mtb* spoligotype SITs

Clades (SIT)	*embB* (DNA)	EmbB (Protein)	Resistant	Susceptible
CAS1_Kili (21)	G918A	M306I	7	19
CAS1_Kili (21)	A916T	M306L	1	4
CAS1_Kili (21)	A916G	M306V	0	1
CAS1_Kili (21)	WT		0	1
CAS1_Kili (21)	G982T	D328	1	0
EAI (Orphan)	G1217A	G406D	0	2
H3/T1 (4)	A916G	M306V	0	1
H1 (Orphan)	A1490G	Q497R	1	3
H3 (50)	G1215C/G1225C	E405D/A409P	1	0
LAM9 (42)	G918A	M306I	1	1
LAM9 (42)	G918C	M306I	0	1
LAM1 (20)	A916G	M306V	0	2
LAM1 (20)	G918A	M306I	3	1
LAM1 (20)	A956C	Y319S	1	0
LAM1 (20)	WT		2	8
LAM11_ZWE (2173)	WT		1	0
LAM11_ZWE (59)	A916G	M306V	12	1
LAM11_ZWE (59)	A956C	Y319S	0	1
LAM11_ZWE (59)	G1217C	G406A	1	2
LAM11_ZWE (59)	G918A	M306I	0	1
LAM11_ZWE (59)	WT		2	8
LAM11_ZWE (811)	A1490G	Q497R	1	0
LAM11_ZWE (815)	A916G	M306V	6	0
LAM11_ZWE (815)	G918A	M306I	1	0
LAM11_ZWE (815)	C1204G	L402V	1	0
LAM11_ZWE (815)	WT		2	1
LAM11_ZWE (815)	G918T	M306I	1	1
LAM11_ZWE (Orphan)	WT		1	1
LAM11_ZWE (Orphan)	G918A	M306I	1	0
LAM11_ZWE (Orphan)	A1490G	Q497R	1	0
S (34)	G918A	M306I	1	0
T (73)	G918A	M306I	1	0
T1 (53)	A916G	M306V	3	1
T1 (53)	G918A	M306I	0	1
T1 (53)	G918C	M306I	0	1
T1 (53)	WT		0	2
T1 (53)	G1217A	G406D	0	1
T2 (317)	C1204G	L402V	0	1
T2 (52)	A916G	M306V	4	0
T2 (52)	WT		0	1
X2 (137)	A916G	M306V	3	1
X2 (137)	WT		0	2

Table 12: *embB* codon mutation frequency in this study and among global isolates

Codon	This study			Global isolates [82]		
	EMB^R (n=61)	EMB^S (n=71)	Total (n=132)	Proportion	Absolute number	Proportion (n=7294)
Codon 306	45	37	82	62.1	3491	47.9
Codon 497	3	3	6	4.5	583	8.0
Codon 406	1	5	6	4.5	828	11.4
Codon 402	1	1	2	1.5		
Codon 319	1	1	2	1.5		
Codon 328	1	0	1	0.8		
Codon 405	1	0	1	0.8		
Codon 409	1	0	1	0.8		
embB locus	53	47	100	75.8		

R-resistant, S-susceptible

Table 13: Association of *embB* mutations with MDR-TB genotypes

Genotype	Mutations in *embB*	No mutations in *embB*	Odds ratio	95%CI	*p* value
CAS1_Kili	33	1	20.9	2.70-162.43	0.004
T	13	3	2.8	0.71-10.58	0.14
X	4	2	1.3	0.22-7.42	0.79
H	6	0			
EAI	2	0			
S	1	0			
LAM	41	26			

Table 14: Correlation of *embB* mutations and EMB phenotype DST by MGIT M960

Codon/Locus	EMB resistant (n=61)		EMB susceptible (n=71)		Sensitivity	specificity	PPV	NPV
	Mutation	No mutation	Mutation	No mutation				
Codon 306	45	16	37	34	73.8	47.9	54.9	68.0
Codon 497	3	58	3	68	4.9	95.8	50.0	54.0
Codon 406	1	60	5	66	1.6	93.0	16.7	52.4
Codon 402	1	60	1	70	1.6	98.6	50.0	53.8
Codon 319	1	60	1	70	1.6	98.6	50.0	53.8
Codon 328	1	60	0	71	1.6	100.0	100.0	54.2
Codon 405	1	60	0	71	1.6	100.0	100.0	54.2
Codon 409	1	60	0	71	1.6	100.0	100.0	54.2
embB locus	53	8	47	24	86.9	33.8	53.0	75.0

PPV=positive predictive value, NPV=Negative predictive value

Table 15: Occurrence of M306I amino acid change among 4 major *Mtb* genotypes [82]

Lineage	M306I	No M306I mutation	Total	Odds ratio	95%CI	*p* value
Lineage 1	30	242	272	0.4618	0.32 to 0.68	0.0001
Lineage 2	595	2559	3154	0.8127	0.72 to 0.91	0.0004
Lineage 3	173	386	559	1.80	1.49 to 2.17	< 0.0001
Lineage 4	718	2591	3309	1.5223	1.36 to 1.70	< 0.0001

Table 16: Comparison of occurrence of M306I in Lineages 1, 2, and 3 to Lineage 4

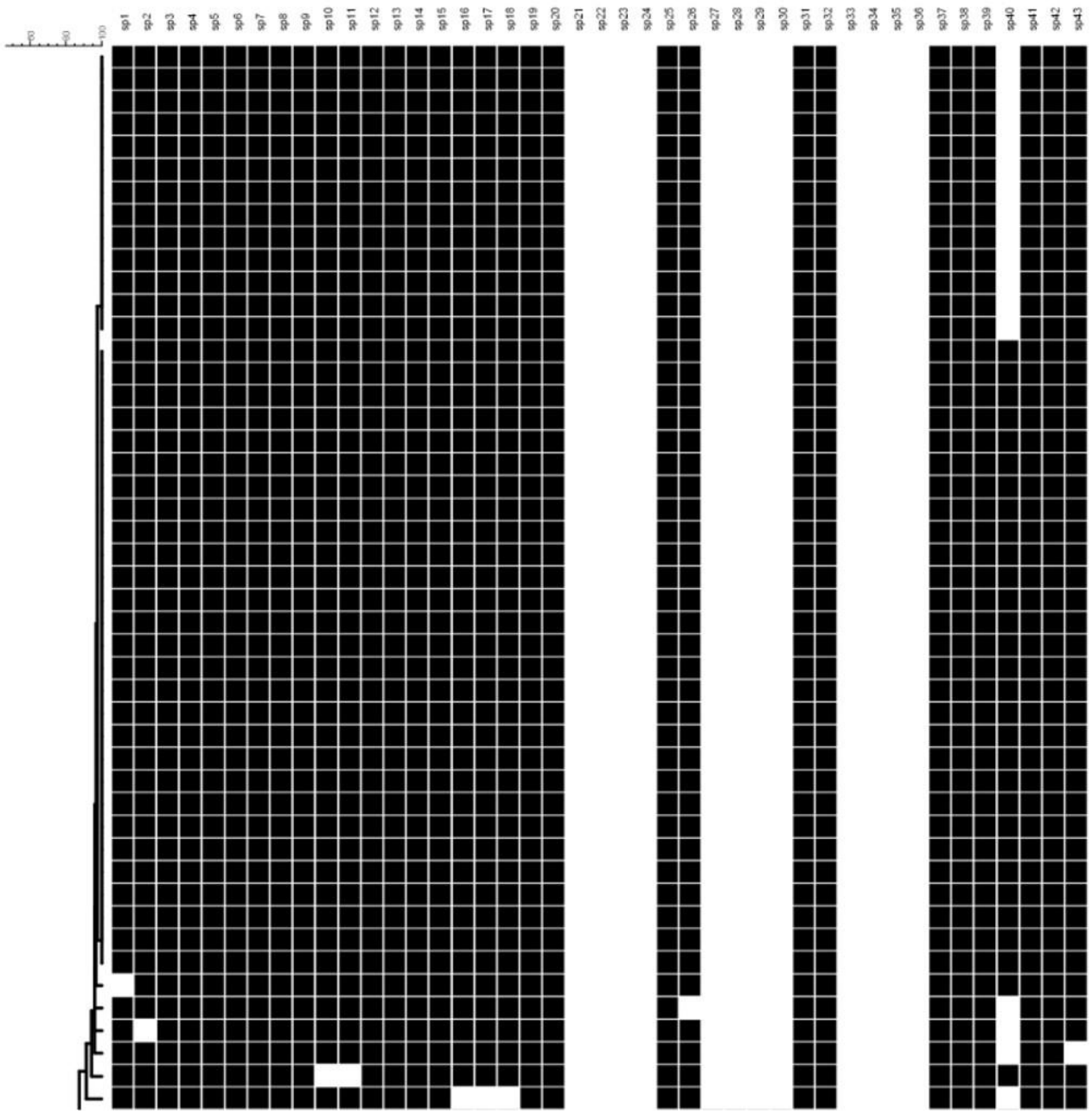
Lineage	M306I	No M306I mutation	Total	Odds ratio	95%CI	*p* value
Lineage 1	30	242	272	0.447	0.30-0.66	<0.0001
Lineage 2	595	2559	3154	0.839	0.74-0.95	0.005
Lineage 3	173	386	559	1.617	1.33-1.97	<0.0001
Lineage 4	718	2591	3309	Reference		

Table 17: Occurrence of M306V amino acid change among 4 major *Mtb* genotypes [82]

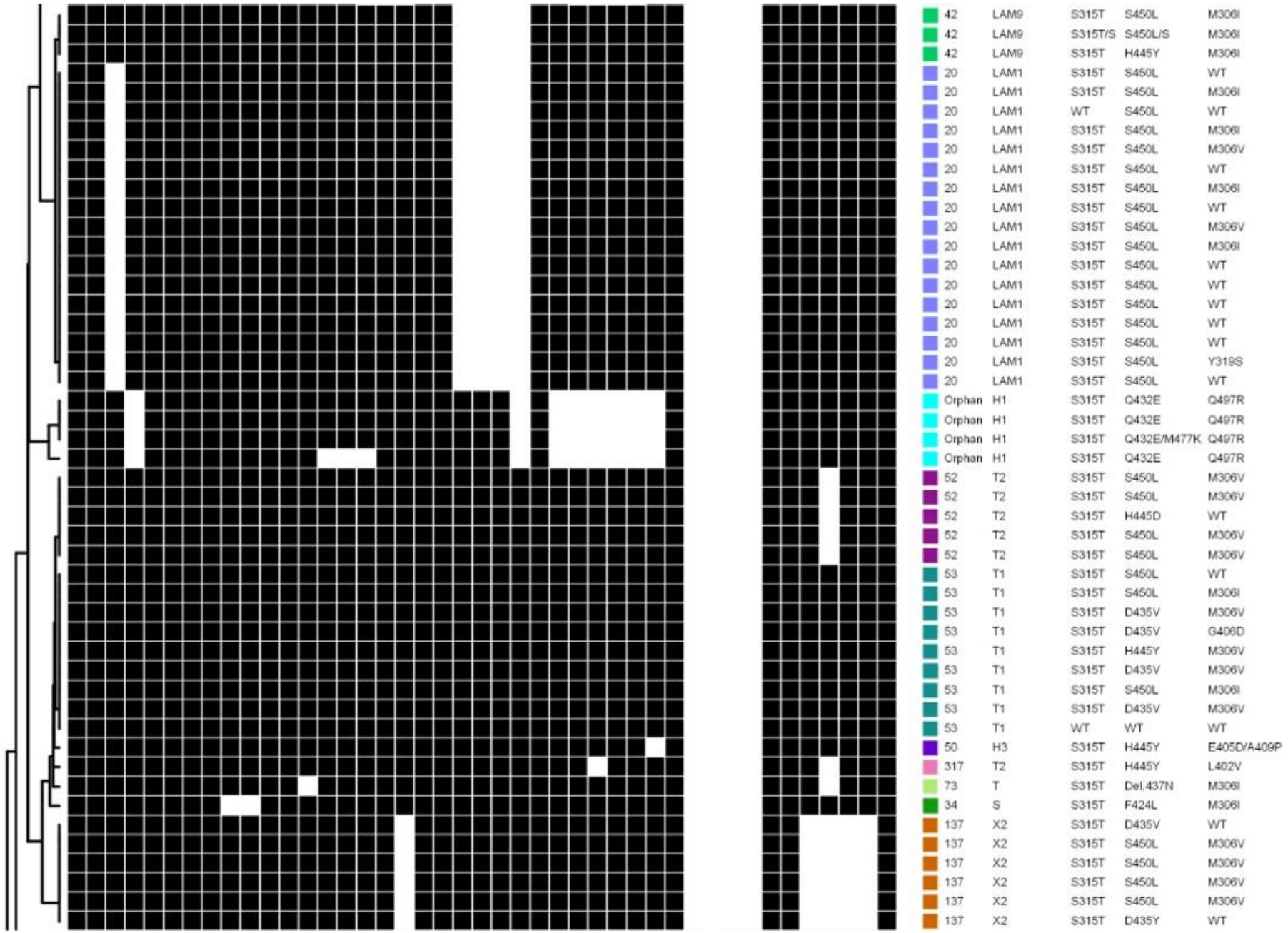
Lineage	M306V	No M306V mutation	Total	Odds ratio	95%CI	*p* value
Lineage 1	63	209	272	0.8797	0.66 to 1.17	0.3812
Lineage 2	977	2177	3154	1.6673	1.50 to 1.85	< 0.0001
Lineage 3	97	462	559	0.59	0.47 to 0.75	< 0.0001
Lineage 4	718	2591	3309	1.5223	1.36 to 1.70	< 0.0001

Table 18: Comparison of occurrence of M306V in Lineages 1, 2, and 3 to Lineage 4

Lineage	M306V	No M306V mutation	Total	Odds ratio	95%CI	*p* value
Lineage 1	63	209	272	1.09	0.81-1.46	0.57
Lineage 2	977	2177	3154	1.619	1.45-1.811	<0.0001
Lineage 3	97	462	559	0.758	0.60-0.96	0.02
Lineage 4	718	2591	3309	Reference		



SIT	Clade	*katG*	*rpoB*	*embB*
815	LAM11_ZWE	S315T	S450L	M306I
815	LAM11_ZWE	S315T	S450L	M306V
815	LAM11_ZWE	S315T	S450L	M306I
815	LAM11_ZWE	S315T	S450L	M306V
815	LAM11_ZWE	S315T	S450L	M306V
815	LAM11_ZWE	S315T	H445Y	M306V
815	LAM11_ZWE	S315T	S450F	M306V
815	LAM11_ZWE	S315T	S450L	L402V
815	LAM11_ZWE	S315T	S450L	WT
815	LAM11_ZWE	S315T	S450L	M306I
815	LAM11_ZWE	WT	WT	WT
815	LAM11_ZWE	S315T	S450L	WT
815	LAM11_ZWE	S315T	S450F	M306V
59	LAM11_ZWE	S315T	L430R/D435Y	M306V
59	LAM11_ZWE	S315T	S450W	M306V
59	LAM11_ZWE	S315T	S450L	G406A
59	LAM11_ZWE	S315T	S450L	Y319S
59	LAM11_ZWE	S315T	L430R/D435Y	WT
59	LAM11_ZWE	S315T	S450L	G406A
59	LAM11_ZWE	S315T	S450L	M306I
59	LAM11_ZWE	S315T	L430R/D435Y	M306V
59	LAM11_ZWE	S315T	S450L	M306V
59	LAM11_ZWE	S315T	D435V/D	M306V
59	LAM11_ZWE	D329E	D435F	WT
59	LAM11_ZWE	S315T	L430R/D435Y	M306V
59	LAM11_ZWE	S315T	S450L	M306V
59	LAM11_ZWE	S315T	S450L	WT
59	LAM11_ZWE	S315T	S450W	M306V
59	LAM11_ZWE	S315T	S450L	WT
59	LAM11_ZWE	S315T	S450L	WT
59	LAM11_ZWE	S315T	S450L	WT
59	LAM11_ZWE	S315T	S450W	M306V
59	LAM11_ZWE	S315T	E423A/S450W	M306V
59	LAM11_ZWE	S315T	S450L	G406A
59	LAM11_ZWE	S315T	D435V	WT
59	LAM11_ZWE	WT	S450L	WT
59	LAM11_ZWE	S315T	S450L	M306V
59	LAM11_ZWE	S315T	L430R/D435Y	M306V
59	LAM11_ZWE	S315T	S450W	M306V
59	LAM11_ZWE	WT	WT	WT
59	LAM11_ZWE	WT	Del.	WT
2173	LAM11_ZWE	WT	WT	WT
811	LAM11_ZWE	S315T	S450L	Q497R
Orphan	LAM11_ZWE	S315T	S450L	WT
Orphan	LAM11_ZWE	S315T	S450L	M306I
Orphan	LAM11_ZWE	S315N	D435V	Q497R
Orphan	LAM11_ZWE	S315T	S450Q	WT



42	LAM9	S315T	S450L	M306I
42	LAM9	S315T/S	S450L/S	M306I
42	LAM9	S315T	H445Y	M306I
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	M306I
20	LAM1	WT	S450L	WT
20	LAM1	S315T	S450L	M306I
20	LAM1	S315T	S450L	M306V
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	M306I
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	M306V
20	LAM1	S315T	S450L	M306I
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	WT
20	LAM1	S315T	S450L	Y319S
20	LAM1	S315T	S450L	WT
Orphan	H1	S315T	Q432E	Q497R
Orphan	H1	S315T	Q432E	Q497R
Orphan	H1	S315T	Q432E/M477K	Q497R
Orphan	H1	S315T	Q432E	Q497R
52	T2	S315T	S450L	M306V
52	T2	S315T	S450L	M306V
52	T2	S315T	H445D	WT
52	T2	S315T	S450L	M306V
52	T2	S315T	S450L	M306V
53	T1	S315T	S450L	WT
53	T1	S315T	S450L	M306I
53	T1	S315T	D435V	M306V
53	T1	S315T	D435V	G406D
53	T1	S315T	H445Y	M306V
53	T1	S315T	D435V	M306V
53	T1	S315T	S450L	M306I
53	T1	S315T	D435V	M306V
53	T1	WT	WT	WT
50	H3	S315T	H445Y	E405D/A409P
317	T2	S315T	H445Y	L402V
73	T	S315T	Del.437N	M306I
34	S	S315T	F424L	M306I
137	X2	S315T	D435V	WT
137	X2	S315T	S450L	M306V
137	X2	S315T	S450L	M306V
137	X2	S315T	S450L	M306V
137	X2	S315T	S450L	M306V
137	X2	S315T	D435Y	WT

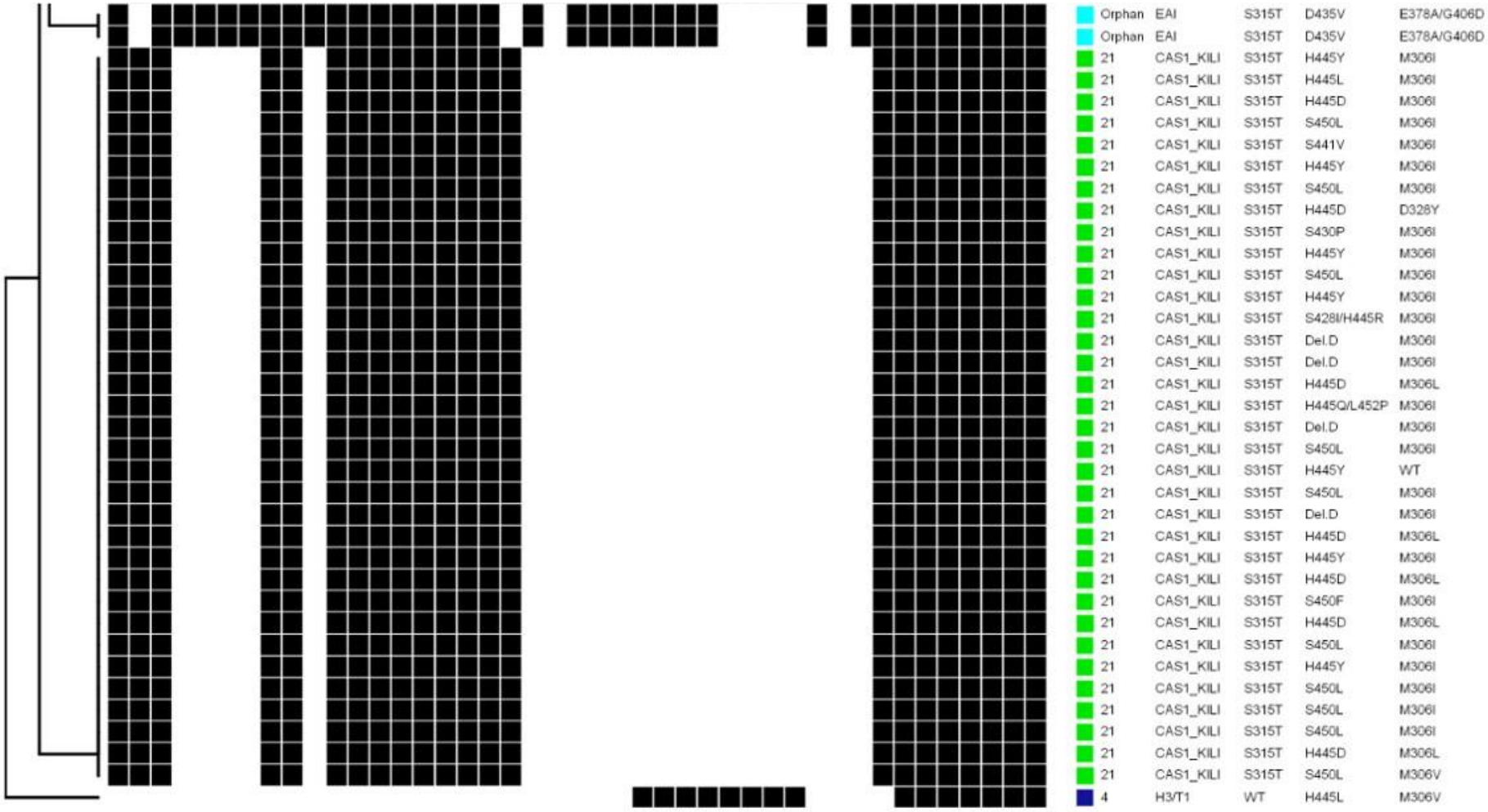


Figure 8: phylogenetic tree based on spoligotype SITs and *katG, rpoB,* and *embB* mutations for 132 MDR-TB isolates

CONCLUSION

In Zambia a resource limited country, MDR-TB is an emerging problem. There are few facilities performing drug susceptibility testing to second-line drugs and the methods used are limited in the scope or by turnaround time. This leads to a delay in initiating effective treatment and consequently, transmission of drug-resistant strains. To control this problem, increased efforts to rapidly diagnose and promptly treat with appropriate drugs are needed. In my Ph.D. studies, I utilized 1) DNA sequencing to characterize resistance conferring mutations and evaluate the utility of mutations for drug resistance prediction and 2) spoligotyping to determine the relationship between drug resistance conferring mutations and *Mtb* genotypes.

In chapter I, I revealed that DNA sequencing would adequately detect STR resistance. In addition, I revealed a correlation between STR resistance mutations and CAS1_Kili genotype in Zambia, particularly K43R mutation. This mutation is associated with Beijing Lineage. I hypothesize that the observed association with CAS1_Kili in my study strongly suggests clonal expansion. The mutation likely emerged at the time STR was used as first-line drug and has since been spreading within Zambia. I have revealed a large cluster of MDR-TB strains belonging to LAM1 genotype that were STR susceptible. This strain was likely introduced in Zambia from neighboring Angola as an MDR-TB strain after STR was no longer used as first-line drug. From this, I showed STR is still a useful drug in the treatment of MDR-TB caused by STR susceptible strains.

In Chapter II, I revealed a widespread EMB resistance using DNA sequencing compared to what could be detected by phenotypic testing. The inability of the MGIT 960 systems to detect low-level EMB resistance means patients with EMB resistant strains will be exposed to the inadequate TB treatment regimen. I recommend the adoption of sequencing for EMB susceptibility testing. In this study also, CAS1_Kili was significantly associated with EMB resistance particularly caused by M306I amino acid change. I revealed an important finding of site-specific amino acid preferences of Lineage 2 and Lineage 3 at *embB306.* The predisposition of Lineage 2 to acquire a mutation that is associated with moderate to high-level resistance, supports the documented association of this genotype with high-level resistance. I hypothesize that this predilection of the two Lineages is due to the different genetic backgrounds, stressing the influence genetic background has on the drug resistance pathways.

In my thesis, I have demonstrated the utility of DNA sequencing for molecular prediction of drug resistance in Zambia. Revealing the gene mutation contribution to STR and EMB resistance is important as it impacts the utilization of rapid molecular tools in the Zambian setting and could potentially affect patient management. I have also demonstrated how the sequencing of drug resistance mutations can be useful in profiling drug resistance evolutionary history of isolates and detect transmission by analyzing clustered isolates with same mutations. Using mutations conferring

resistance to RIF, INH, STR, and EMB, I have revealed multiple transmission events of MDR-TB taking place in Zambia (Figures 7 and 8). CAS1_Kili and LAM1 especially appear to have increased transmissibility due to the large sizes of clusters with similar drug resistance mutations in the genotypes and should carefully be monitored. I recommend a future study to determine hotspot regions for MDR-TB transmission to guide in formulating specific control measures and optimize the use of limited resources. Finally, I recommend using DNA sequencing to dramatically improve resistance detection and early treatment with effective drugs in Zambia and contribute to control of MDR-TB.

ACKNOWLEDGEMENTS

I am profoundly grateful to Prof. Yasuhiko Suzuki of the division of Bioresources, International Institute for Zoonosis Control, Hokkaido University for the guidance, encouragement, stimulating discussions, support, and patience given to me throughout the course of my study. I also express my deepest appreciation to Prof. Chie Nakajima of the division of Bioresources, International Institute for Zoonosis Control, Hokkaido University for the generous support, thought-provoking questions, and insightful advice. I want to express my heartfelt gratitude to my advisor's professor Higashi Hideaki of the division of Infection and Immunity, International Institute for Zoonosis Control, Hokkaido University and associate professor Norikazu Isoda of Laboratory of Microbiology, Faculty of Veterinary Medicine, Hokkaido University for their invaluable advice and suggestions to improve my research and write my thesis.

I also express my gratitude to Mrs. Grace Mbulo and Dr. Eddie S. Solo of The University Teaching Hospital, TB Laboratory, Zambia for recommending me to pursue a Ph.D. course under the guidance of Prof. Yasuhiko Suzuki and for permission to use archived samples for my study. I am grateful to Mr. Timothy Kantenga of The University Teaching Hospital, Department of Pathology and Microbiology, Zambia for encouraging me to progress in my career and for advocating for my study leave from my work so that I was able to pursue this Ph.D.

Many thanks to Ms. Yuko Hidaka and Ms. Kimura Naomi for your warm support that made my school life easy going and comfortable, enabling me to focus on my studies.

To Ms. Fukushima Yukari, Ms. Kunda Kasakwa, and Ms. Mei Koyama, thank you very much for your technical support.

I would like to acknowledge members of Bioresources division members Dr. Jeewan Thapa, Dr. Yayoi Kameda, Dr. Yogendra Shah, Dr. Tomoyuki Yamaguchi, Dr. Hasan Diab, Dr. Nan Aye Thida Oo, Dr. Lai Lai San, Dr. Charitha Mendis, Dr. Kentaro Koide, Dr. Yuki Ouchi, Dr. Park Jong-Hoon, Dr. Ruttana Pachanon, Dr. Thoko Flav Kapalamula, Dr. Dipti Shrestha, Dr. Risa Tsunoda, Dr. Lawrence Belotindos, Dr. Wimonrat Tanomsridachchai, Ms. Janisara Rudeeaneksin, Ms. Mwangala Lonah Akapelwa, Dr. Joseph Yamweka Chizimu, Ms. Nami Ajima, Ms. Jirachaya Toyting, Ms. Pondpan Suwanthada, Mr. Pawarut Narongpun, and Mr. David Barnes, thank you very much for making the research environment warm, friendly, peaceful, and upbuilding.

I express my appreciation to my fellow Students Boniface Lombe Pongombo, Misheck Shawa, Herman Chambaro, Manyando Simbotwe for sharing their experiences.

I would like to thank Japan International Cooperation (JICA) for granting me a scholarship to pursue this Ph.D. course and for the WISE program office for supporting my training.

I want to acknowledge Machiko, Michiyo, Emi, Naoe, Mikako, Satomi, Kana, Tatiana, Lilly, Yuri, Cindy, Nagisa, Abigail, Nobara, and many other friends for being my strengthening aids. Working

alongside all of you in volunteer work gave me balance and grounded me.

Lastly, I want to express my heartfelt gratitude to my dear sisters Glorious, Gracious, and Generous; and brothers Spacious and Dominic and niece for your constant support and for all the laughter that calmed my anxieties and helped me persevere. I would like to dedicate this Ph.D. to my father, for encouraging and supporting me.

REFERENCES

[1] Raviglione MC, Snider Jr DE, Kochi A. Global epidemiology of tuberculosis: Morbidity and mortality of a worldwide epidemic. JAMA 1995;273:220–6. doi:10.1001/jama.1995.03520270054031.

[2] World Health Organization. Global Tuberculosis Report 2021. Geneva, Switzeland: World Health organisation; 2021.

[3] Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk H-P, et al. Taxonomy, physiology, and natural products of *Actinobacteria*. Microbiol Mol Biol Rev 2016;80:1–43. doi:10.1128/mmbr.00044-16.

[4] Dulberger CL, Rubin EJ, Boutte CC. The mycobacterial cell envelope — a moving target. Nat Rev Microbiol 2020;18:47–59. doi:10.1038/s41579-019-0273-7.

[5] Abrahams KA, Besra GS. Mycobacterial cell wall biosynthesis: A multifaceted antibiotic target. Parasitology 2018;145:116–33. doi:10.1017/S0031182016002377.

[6] Gygli SM, Borrell S, Trauner A, Gagneux S. Antimicrobial resistance in *Mycobacterium tuberculosis*: Mechanistic and evolutionary perspectives. FEMS Microbiol Rev 2017;41:354–73. doi:10.1093/femsre/fux011.

[7] Ehrt S, Schnappinger D. Mycobacterial survival strategies in the phagosome: Defence against host stresses. Cell Microbiol 2009;11:1170–8. doi:10.1111/j.1462-5822.2009.01335.x.

[8] Reece ST, Kaufmann SHE. Floating between the poles of pathology and protection: Can we pin down the granuloma in tuberculosis? Curr Opin Microbiol 2012;15:63–70. doi:10.1016/j.mib.2011.10.006.

[9] Cohen A, Mathiasen VD, Schön T, Wejse C. The global prevalence of latent tuberculosis: A systematic review and meta-analysis. Eur Respir J 2019;54:1–14. doi:10.1183/13993003.00655-2019.

[10] Gengenbacher M, Kaufmann SHE. *Mycobacterium tuberculosis*: Success through dormancy. FEMS Microbiol Rev 2012;36:514–32. doi:10.1111/j.1574-6976.2012.00331.x.

[11] Dobler CC, Cheung K, Nguyen J, Martin A. Risk of tuberculosis in patients with solid cancers and haematological malignancies: a systematic review and meta-analysis. Eur Respir J 2017;50:1700157. doi:10.1183/13993003.00157-2017.

[12] Narasimhan P, Wood J, Macintyre CR, Mathai D. Risk factors for tuberculosis. Pulm Med 2013;2013:828939. doi:10.1155/2013/828939.

[13] Zumla A, Malon P, Henderson J, Grange MJ. Impact of HIV infection on tuberculosis. Postgrad Med J 2000;76:259–68. doi:10.1136/pmj.76.895.259.

[14] Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs WRJ, et al. An outbreak of tuberculosis with accelerated progression among persons infected with the human

immunodeficiency virus. N Engl J Med 1992;326:231–5. doi:10.1056/NEJM199201233260404.

[15] Thye T, Owusu-Dabo E, Vannberg FO, Van Crevel R, Curtis J, Sahiratmadja E, et al. Common variants at 11p13 are associated with susceptibility to tuberculosis. Nat Genet 2012;44:257–9. doi:10.1038/ng.1080.

[16] Chimusa ER, Zaitlen N, Daya M, Möller M, Helden PD va., Nicola JM, et al. Genome-wide association study of ancestry-specific TB risk in the South African coloured population. Hum Mol Genet 2014;23:796–809. doi:10.1093/hmg/ddt462.

[17] Grant A V., Sabri A, Abid A, Abderrahmani Rhorfi I, Benkirane M, Souhi H, et al. A genome-wide association study of pulmonary tuberculosis in Morocco. Hum Genet 2016;135:299–307. doi:10.1007/s00439-016-1633-2.

[18] Sveinbjornsson G, Gudbjartsson DF, Halldorsson B V, Kristinsson KG, Gottfredsson M, Barrett JC, et al. HLA class II sequence variants influence tuberculosis risk in populations of European ancestry 2016;48:318–22. doi:10.1038/ng.3498.

[19] Fox GJ, Orlova M, Schurr E. Tuberculosis in Newborns: The Lessons of the "Lübeck Disaster" (1929–1933). PLoS Pathog 2016;12:1–10. doi:10.1371/journal.ppat.1005271.

[20] World Health Organisation. Global Tuberculosis Report 2020. Geneva, Switzeland: World Health Organisation; 2020.

[21] World Health Organization. Global Tuberculosis Report 2019. Geneva, Switzeland: World Health organisation; 2019.

[22] Perlman DC, Segal Y, Rosenkranz S, Rainey PM, Remmel RP, Salomon N, et al. The clinical pharmacokinetics of rifampin and ethambutol in HIV-infected persons with tuberculosis. Clin Infect Dis 2005;41:1638–47. doi:10.1086/498024.

[23] World Health Organization. WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva, Switzeland: World Health organisation; 2020.

[24] Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, et al. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell 2001;104:901–12. doi:10.1016/S0092-8674(01)00286-0.

[25] Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. *ethA*, *inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. Antimicrob Agents Chemother 2003;47:3799–805. doi:10.1128/AAC.47.12.3799-3805.2003.

[26] Bloemberg G V, Keller PM, Stucki D, Trauner A, Borrell S, Latshang T, et al. Acquired resistance to bedaquiline and delamanid in therapy for tuberculosis. N Engl J Med 2015;373:1986–8. doi:10.1056/NEJMc1505196.

[27] Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The competitive cost

of antibiotic resistance in *Mycobacterium tuberculosis*. Science 2006;312:1944–6. doi:10.1126/science.1124410.

[28] Björkman J, Nagaev I, Berg OG, Hughes D, Andersson DI. Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. Science 2000;287:1479–82. doi:10.1126/science.287.5457.1479.

[29] Caminero JA. Multidrug-resistant tuberculosis: Epidemiology, risk factors and case finding. Int J Tuberc Lung Dis 2010;14:382–90.

[30] Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K, Ali S, et al. Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. Nat Genet 2018;50:307–16. doi:10.1038/s41588-017-0029-0.

[31] Sander P, Springer B, Prammananan T, Sturmfels A, Kappler M, Pletschette M, et al. Fitness cost of chromosomal drug resistance-conferring mutations 2002;46:1204–11. doi:10.1128/AAC.46.5.1204.

[32] Makhado NA, Matabane E, Faccin M, Pinçon C, Jouet A, Boutachkourt F, et al. Outbreak of multidrug-resistant tuberculosis in South Africa undetected by WHO-endorsed commercial tests: an observational study. Lancet Infect Dis 2018;18:1350–9. doi:10.1016/S1473-3099(18)30496-1.

[33] Fenner L, Egger M, Bodmer T, Altpeter E, Zwahlen M, Jaton K, et al. Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2012;56:3047–53. doi:10.1128/AAC.06460-11.

[34] Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 2007;7:328–37. doi:10.1016/S1473-3099(07)70108-1.

[35] Firdessa R, Berg S, Hailu E, Schelling E, Gumi B, Erenso G, et al. Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia. Emerg Infect Dis 2013;19:460–3. doi:10.3201/eid1903.120256.

[36] Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLOS Biol 2008;6:e311. doi:10.1371/journal.pbio.0060311.

[37] Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, et al. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat Genet 2013;45:784. doi:10.1038/ng.2656.

[38] Nonghanphithak D, Chaiprasert A, Smithtikarn S, Kamolwat P, Pungrassami P, Chongsuvivatwong V, et al. Clusters of drug-resistant *Mycobacterium tuberculosis* detected by whole-genome sequence analysis of nationwide sample, Thailand, 2014–2017. Emerg Infect Dis 2021;27:813–22. doi:10.3201/eid2703.204364.

[39] Buu TN, van Soolingen D, Huyen MNT, Lan NTN, Quy HT, Tiemersma EW, et al. Increased transmission of *Mycobacterium tuberculosis* Beijing genotype strains associated with resistance to streptomycin: A population-based study. PLoS One 2012;7:e42323. doi:10.1371/journal.pone.0042323.

[40] Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP, et al. Evolution of Extensively Drug-Resistant tuberculosis over four Decades: whole genome sequencing and dating analysis of *Mycobacterium tuberculosis* isolates from KwaZulu-Natal. PLoS Med 2015;12:1–22. doi:10.1371/journal.pmed.1001880.

[41] Eldholm V, Monteserin J, Rieux A, Lopez B, Sobkowiak B, Ritacco V. Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. Nat Commun 2015;6:7119. doi:10.1038/ncomms8119.

[42] Singh AV, Singh S, Yadav A, Kushwah S, Yadav R, Sai DK, et al. Genetic variability in multidrug-resistant *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in North India. BMC Microbiol 2021;21:123. doi:10.1186/s12866-021-02174-6.

[43] Van Soolingen D. Molecular epidemiology of tuberculosis and other mycobacterial infections: Main methodologies and achievements. J Intern Med 2001;249:1–26. doi:10.1046/j.1365-2796.2001.00772.x.

[44] Kamerbeek J, Schouls L, Kolk A, Agterveld M Van, Soolingen D van, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 1997;35:907–14. doi:10.1128/jcm.35.4.907-914.1997.

[45] Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol Microbiol 2000;36:762–71. doi:10.1046/j.1365-2958.2000.01905.x.

[46] David A, W. LD, Hernando HM, S. MA, Weihong Q, D. FR, et al. Role of Large Sequence Polymorphisms (LSPs) in generating genomic diversity among clinical isolates of *Mycobacterium tuberculosis* and the utility of LSPs in phylogenetic analysis. J Clin Microbiol 2007;45:39–46. doi:10.1128/JCM.02483-05.

[47] Feuerriegel S, Köser CU, Niemann S. Phylogenetic polymorphisms in antibiotic resistance genes of the *Mycobacterium tuberculosis* complex. J Antimicrob Chemother 2014;69:1205–10. doi:10.1093/jac/dkt535.

[48] MacPherson P, Houben RMGJ, Glynn JR, Corbett EL, Kranzer K. Loss of follow-up before treatment of TB patients in low- and middle-income countries and in high-burden countries: a systematic review and meta-analysis. Bull World Health Organ 2014;92:126–38. doi:10.2471/BLT.13.124800.

[49] Lungu P, Kerkhoff AD, Kasapo CC, Mzyece J, Nyimbili S, Chimzizi R, et al. Tuberculosis care cascade in Zambia - Identifying the gaps in order to improve outcomes: A population-

based analysis. BMJ Open 2021;11:1–14. doi:10.1136/bmjopen-2020-044867.

[50] Bates M, O'Grady J, Mwaba P, Chilukutu L, Mzyece J, Cheelo B, et al. Evaluation of the burden of unsuspected pulmonary tuberculosis and co-morbidity with non-communicable diseases in sputum producing adult inpatients. PLoS One 2012;7:e40774. doi:10.1371/journal.pone.0040774.

[51] Habeenzu C, Nakajima C, Solo E, Bwalya P, Kajino K, Miller M, et al. Evaluation of in-house loop-mediated isothermal amplification for tuberculosis diagnosis compared with Xpert MTB/RIF. J Infect Dev Ctries 2017;11:440–4. doi:10.3855/jidc.7730.

[52] Bjerrum S, Schiller I, Dendukuri N, Kohli M, Nathavitharana RR, Zwerling AA, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. Cochrane Database Syst Rev 2019;10:CD011420. doi:10.1002/14651858.CD011420.pub3.

[53] Ministry of Health G of the R of Z. The National Tuberculosis and Leprosy Programme: TB Manual 2008:32–3. doi:10.1016/j.pt.2010.06.009.

[54] Kapata N, Mbulo G, Cobelens F, de Haas P, Schaap A, Mwamba P, et al. The Second Zambian National Tuberculosis Drug Resistance survey – a comparison of conventional and molecular methods. Trop Med Int Heal 2015;20:1492–500. doi:10.1111/tmi.12581.

[55] World Health Organization. WHO global lists of high burden countries for tuberculosis (TB), TB / HIV and TB (MDR / RR-TB), 2021-2025. Geneva, Switzeland: World Health organisation; 2021.

[56] Kapata N, Chanda-Kapata P, Bates M, Mwaba P, Cobelens F, Grobusch MP, et al. Multidrug-resistant TB in Zambia: Review of national data from 2000 to 2011. Trop Med Int Heal 2013;18:1386–91. doi:10.1111/tmi.12183.

[57] World Health Organization. WHO guidelines on tuberculosis infection prevention and control: 2019 update. Geneva, Switzeland: World Health organisation; 2019.

[58] Solo ES, Suzuki Y, Kaile T, Bwalya P, Lungu P, Chizimu JY, et al. Characterization of *Mycobacterium tuberculosis* genotypes and their correlation to multidrug resistance in Lusaka, Zambia. Int J Infect Dis 2021;102:489–96. doi:10.1016/j.ijid.2020.10.014.

[59] Chizimu YJ, Solo ES, Bwalya P, Kapalamula FT, Akapelwa LM, Lungu P, et al. Genetic diversity and transmission of multidrug-resistant *Mycobacterium tuberculosis* strains in Lusaka , Zambia. Int J Infect Dis 2022;114:142–50. doi:10.1016/j.ijid.2021.10.044.

[60] Solo ES, Nakajima C, Kaile T, Bwalya P, Mbulo G, Fukushima Y, et al. Mutations in *rpoB* and *katG* genes and the *inhA* operon in multidrug-resistant *Mycobacterium tuberculosis* isolates from Zambia. J Glob Antimicrob Resist 2020;22:302–7. doi:10.1016/j.jgar.2020.02.026.

[61] World Health Organization. Global Tuberculosis Report 2016. Geneva, Switzeland: World Health organisation; 2016.

[62] World Health Organization. Global Tuberculosis Report 2017. Geneva, Switzeland: World

Health organisation; 2017.

[63] World Health Organization. Global Tuberculosis Report 2018. Geneva, Switzeland: World Health Organisation; 2018.

[64] World Health Organization. Tuberculosis Country profile: Zambia 2015. Geneva, Switzeland: World Health organisation; n.d.

[65] Kapata N, Chanda-Kapata P, Bates M, Mwaba P, Cobelens F, Grobusch MP, et al. Multidrug-resistant TB in Zambia: Review of national data from 2000 to 2011. Trop Med Int Heal 2013;18:1386–91. doi:10.1111/tmi.12183.

[66] Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature 2000;407:340–8. doi:https://doi.org/10.1038/35030019.

[67] Tudó G, Rey E, Borrell S, Alcaide F, Codina G, Coll P, et al. Characterization of mutations in streptomycin-resistant *Mycobacterium tuberculosis* clinical isolates in the area of Barcelona. J Antimicrob Chemother 2010;65:2341–6. doi:10.1093/jac/dkq322.

[68] Okamoto S, Tamaru A, Nakajima C, Nishimura K, Tanaka Y, Tokuyama S, et al. Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. Mol Microbiol 2007;63:1096–106. doi:10.1111/j.1365-2958.2006.05585.x.

[69] Wong SY, Javid B, Addepalli B, Piszczek G, Strader MB, Limbach PA, et al. Functional role of methylation of G518 of the 16S rRNA 530 loop by GidB in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2013;57:6311–8. doi:10.1128 /AAC.00905-13.

[70] Cohen KA, Stott KE, Munsamy V, Manson AL, Earl AM, Pym AS. Evidence for expanding the role of streptomycin in the management of drug-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2020;64:e00860-20. doi:10.1128/AAC .00860-20.

[71] Thida Oo NA, San LL, Thapa J, Aye KS, Aung WW, Nakajima C, et al. Characterization of mutations conferring streptomycin resistance to multidrug-resistant *Mycobacterium tuberculosis* isolates from Myanmar. Tuberculosis 2018;111:8–13. doi:10.1016/j.tube.2018.05.003.

[72] Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95–8. doi:citeulike-article-id:691774.

[73] Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, et al. *Mycobacterium tuberculosis* complex genetic diversity: Mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol 2006;6:1–17. doi:10.1186/1471-2180-6-23.

[74] Elliott AM, Halwiindi B, Hayes RJ, Luo N, Mwinga AG, Tembo G, et al. The impact of human

immunodeficiency virus on mortality of patients treated for tuberculosis in a cohort study in Zambia. Trans R Soc Trop Med Hyg 1995;89:78–82. doi:10.1016/0035-9203(95)90668-1.

[75] Hosp M, Elliott AM, Raynes JG, Mwinga AG, Luo N, Zangerle R, et al. Neopterin, β2-microglobulin, and acute phase proteins in HIV-1-seropositive and -seronegative Zambian patients with tuberculosis. Lung 1997;175:265–75. doi:10.1007/PL00007573.

[76] Lukoye D, Adatu F, Musisi K, Kasule GW, Were W, Odeke R, et al. Anti-Tuberculosis Drug Resistance among New and Previously Treated Sputum Smear-Positive Tuberculosis Patients in Uganda: Results of the First National Survey. PLoS One 2013;8:e70763. doi:10.1371/journal.pone.0070763.

[77] Chihota V, Apers L, Mungofa S, Kasongo W, Nyoni IM, Tembwe R, et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in regions of Southern Africa. Int J Tuberc Lung Dis 2007;11:311–8.

[78] Mulenga C, Shamputa IC, Mwakazanga D, Kapata N, Portaels F, Rigouts L. Diversity of *Mycobacterium tuberculosis* genotypes circulating in Ndola, Zambia. BMC Infect Dis 2010;10:177. doi:10.1186/1471-2334-10-177.

[79] Zhao L, Liu H, Sun Q, Xiao T, Zhao X, Li G, et al. Identification of mutations conferring streptomycin resistance in multidrug-resistant tuberculosis of China. Diagn Microbiol Infect Dis 2015;83:150–3. doi:10.1016/j.diagmicrobio.2015.06.020.

[80] Ssengooba W, Meehan CJ, Lukoye D, Kasule GW, Musisi K, Joloba ML, et al. Whole genome sequencing to complement tuberculosis drug resistance surveys in Uganda. Infect Genet Evol 2016;40:8–16. doi:10.1016/j.meegid.2016.02.019.

[81] Feuerriegel S, Oberhauser B, George AG, Dafae F, Richter E, Rüsch-Gerdes S, et al. Sequence analysis for detection of first-line drug resistance in *Mycobacterium tuberculosis* strains from a high-incidence setting. BMC Microbiol 2012;12:90. doi:10.1186/1471-2180-12-90.

[82] Phelan JE, O'Sullivan DM, Machado D, Ramos J, Oppong YEA, Campino S, et al. Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. Genome Med 2019;11:1–7. doi:10.1186/s13073-019-0650-x.

[83] Sun H, Zhang C, Xiang L, Pi R, Guo Z, Zheng C, et al. Characterization of mutations in streptomycin-resistant *Mycobacterium tuberculosis* isolates in Sichuan, China and the association between Beijing-lineage and dual-mutation in *gidB*. Tuberculosis 2016;96:102–6. doi:10.1016/j.tube.2015.09.004.

[84] Verma JS, Gupta Y, Nair D, Manzoor N, Rautela RS, Rai A, et al. Evaluation of *gidB* alterations responsible for streptomycin resistance in *Mycobacterium tuberculosis*. J Antimicrob Chemother 2014;69:2935–41. doi:10.1093/jac/dku273.

[85] Bauskenieks M, Pole I, Skenders G, Jansone I, Broka L, Nodieva A, et al. Genotypic and phenotypic characteristics of aminoglycoside-resistant *Mycobacterium tuberculosis* isolates in

Latvia. Diagn Microbiol Infect Dis 2015;81:177–82. doi:10.1016/j.diagmicrobio.2014.12.004.

[86] Lanzas F, Karakousis PC, Sacchettini JC, Ioerger TR. Multidrug-resistant tuberculosis in panama is driven by clonal expansion of a multidrug-resistant *Mycobacterium tuberculosis* strain related to the KZN extensively drug-resistant *M. tuberculosis* strain from South Africa. J Clin Microbiol 2013;51:3277–85. doi:10.1128/JCM.01122-13.

[87] Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat Genet 2014;46:279–286. doi:10.1038/ng.2878.

[88] Wong SY, Lee JS, Kwak HK, Via LE, Boshoff HIM, Barry CE. Mutations in *gidB* confer low-level streptomycin resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2011;55:2515–22. doi:10.1128/AAC.01814-10.

[89] Coll F, McNerney R, Guerra-Assunção JA, Glynn JR, Perdigão J, Viveiros M, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. Nat Commun 2014;5:4812. doi:10.1038/ncomms5812.

[90] Chihota VN, Niehaus A, Streicher EM, Wang X, Sampson SL, Mason P, et al. Geospatial distribution of *Mycobacterium tuberculosis* genotypes in Africa. PLoS One 2018;13:1–18. doi:10.1371/journal.pone.0200632.

[91] Mbugi E V, Katale BZ, Streicher EM, Keyyu JD, Kendall SL, Dockrell HM, et al. Mapping of *Mycobacterium tuberculosis* complex genetic diversity profiles in Tanzania and other African countries. PLoS One 2016;11:eo154571. doi:10.1371/journal.pone.0154571.

[92] Katale BZ, Mbelele PM, Lema NA, Campino S, Mshana SE, Rweyemamu MM, et al. Whole genome sequencing of *Mycobacterium tuberculosis* isolates and clinical outcomes of patients treated for multidrug-resistant tuberculosis in Tanzania. BMC Genomics 2020;21:1–15. doi:10.1186/s12864-020-6577-1.

[93] Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. Eur Respir J 2017;50:1701354. doi:10.1183/13993003.01354-2017.

[94] Perdigão J, Clemente S, Ramos J, Masakidi P, Machado D, Silva C, et al. Genetic diversity, transmission dynamics and drug resistance of *Mycobacterium tuberculosis* in Angola. Sci Rep 2017;7:42814. doi:10.1038/srep42814.

[95] Rando-Segura A, Aznar ML, Moreno MM, Espasa M, Sulleiro E, Bocanegra C, et al. Drug resistance of *Mycobacterium tuberculosis* complex in a rural setting, Angola. Emerg Infect Dis 2018;24:569–72. doi:10.3201/eid2403.171562.

[96] Jnawali HN, Hwang SC, Park YK, Kim H, Lee YS, Chung GT, et al. Characterization of mutations in multi- and extensive drug resistance among strains of *Mycobacterium*

tuberculosis clinical isolates in Republic of Korea. Diagn Microbiol Infect Dis 2013;76:187–96. doi:10.1016/j.diagmicrobio.2013.02.035.

[97] Oppong YEA, Phelan J, Perdigão J, Machado D, Miranda A, Portugal I, et al. Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance. BMC Genomics 2019;20:252. doi:10.1186/s12864-019-5615-3.

[98] Borrell S, Gagneux S. Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. Clin Microbiol Infect 2011;17:815–20. doi:10.1111/j.1469-0691.2011.03556.x.

[99] World Health Organization. Tuberculosis profile: Zambia 2020. https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&lan=%22EN%22&iso2=%22ZM%22.

[100] Zhang L, Zhao Y, Gao Y, Wu L, Gao R, Zhang Q, et al. Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. Science 2020;368:1211–9. doi:10.1126/science.aba9102.

[101] Koul A, Herget T, Klebl B, Ullrich A. Interplay between mycobacteria and host signalling pathways. Nat Rev Microbiol 2004;2:189–202. doi:10.1038/nrmicro840.

[102] Sun Q, Xiao TY, Liu HC, Zhao XQ, Liu ZG, Li YN, et al. Mutations within *embCAB* are associated with variable level of ethambutol resistance in *Mycobacterium tuberculosis* isolates from China. Antimicrob Agents Chemother 2018;62:1–8. doi:10.1128/AAC.01279-17.

[103] Xu Y, Jia H, Huang H, Sun Z, Zhang Z. Mutations found in *embCAB*, *embR*, and *ubiA* genes of ethambutol-sensitive and -resistant *Mycobacterium tuberculosis* clinical isolates from China. Biomed Res Int 2015;2015:951706. doi:10.1155/2015/951706.

[104] Zhang Z, Wang Y, Pang Y, Kam KM. Ethambutol resistance as determined by broth dilution method correlates better than sequencing results with *embB* mutations in multidrug-resistant *Mycobacterium tuberculosis* isolates. J Clin Microbiol 2014;52:638–41. doi:10.1128/JCM.02713-13.

[105] Bakuła Z, Napiórkowska A, Bielecki J, Augustynowicz-Kopeć E, Zwolska Z, Jagielski T. Mutations in the *embB* gene and their association with ethambutol resistance in multidrug-resistant *Mycobacterium tuberculosis* clinical isolates from Poland. Biomed Res Int 2013;2013:167954. doi:10.1155/2013/167954.

[106] Zhao LL, Sun Q, Liu HC, Wu XC, Xiao TY, Zhao XQ, et al. Analysis of *embCAB* mutations associated with ethambutol resistance in multidrug-resistant *Mycobacterium tuberculosis* isolates from China. Antimicrob Agents Chemother 2015;59:2045–50. doi:10.1128/AAC.04933-14.

[107] Li MC, Chen R, Lin SQ, Lu Y, Liu HC, Li GL, et al. Detecting ethambutol resistance in *Mycobacterium tuberculosis* isolates in China: A comparison between phenotypic drug

susceptibility testing methods and DNA sequencing of *embAB*. Front Microbiol 2020;11:1–7. doi:10.3389/fmicb.2020.00781.

[108] World Health Organization. Technical manual for drug susceptibility testing of medicines used in the treatment of tuberculosis. Geneva, Switzeland: World Health organisation; 2018.

[109] Bwalya P, Yamaguchi T, Solo ES, Chizimu JY, Mbulo G, Nakajima C, et al. Characterization of mutations associated with streptomycin resistance in multidrug-resistant *Mycobacterium tuberculosis* in Zambia. Antibiotics 2021;10:1169. doi:10.3390/antibiotics10101169.

[110] Al-Mutairi NM, Ahmad S, Mokaddas EM. Molecular characterization of multidrug-resistant *Mycobacterium tuberculosis* (MDR-TB) isolates identifies local transmission of infection in Kuwait, a country with a low incidence of TB and MDR-TB. Eur J Med Res 2019;24:1–13. doi:10.1186/s40001-019-0397-2.

[111] Safi H, Sayers B, Hazbón MH, Alland D. Transfer of *embB* codon 306 mutations into clinical *Mycobacterium tuberculosis* strains alters susceptibility to ethambutol, isoniazid, and rifampin. Antimicrob Agents Chemother 2008;52:2027–34. doi:10.1128/AAC.01486-07.

[112] Safi H, Fleischmann RD, Peterson SN, Jones MB, Jarrahi B, Alland D. Allelic exchange and mutant selection demonstrate that common clinical *embCAB* gene mutations only modestly increase resistance to ethambutol in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2010;54. doi:10.1128/AAC.01288-09.

[113] Safi H, Lingaraju S, Amin A, Kim S, Jones M, Holmes M, et al. Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. Nat Genet 2013;45:1190–7. doi:10.1038/ng.2743.

[114] Ahmad S, Jaber AA, Mokaddas E. Frequency of *embB* codon 306 mutations in ethambutol-susceptible and -resistant clinical *Mycobacterium tuberculosis* isolates in Kuwait. Tuberculosis 2007;87:123–9. doi:10.1016/j.tube.2006.05.004.

[115] Cuevas-Córdoba B, María Juárez-Eusebio D, Almaraz-Velasco R, Muñiz-Salazar R, Laniado-Laborin R, Zenteno-Cuevas R. Mutation at *embB* codon 306, a potential marker for the identification of multidrug resistance associated with ethambutol in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2015;59:5455–62. doi:10.1128/AAC.00117-15.

[116] Hazbón MH, Bobadilla Del Valle M, Guerrero MI, Varma-Basil M, Filliol I, Cavatore M, et al. Role of *embB* codon 306 mutations in *Mycobacterium tuberculosis* revisited: A novel association with broad drug resistance and IS6110 clustering rather than ethambutol resistance. Antimicrob Agents Chemother 2005;49:3794–802. doi:10.1128/AAC.49.9.3794-3802.2005.

[117] Payne JL, Menardo F, Trauner A, Borrell S, Gygli SM, Loiseau C, et al. Transition bias influences the evolution of antibiotic resistance in *Mycobacterium tuberculosis*. PLOS Biol 2018;17:e3000265. doi:10.1371/journal.pbio.3000265.

[118] Park YK, Ryoo SW, Lee SH, Jnawali HN, Kim CK, Kim HJ, et al. Correlation of the phenotypic ethambutol susceptibility of *Mycobacterium tuberculosis* with *embB* gene mutations in Korea. J Med Microbiol 2012;61:529–34. doi:10.1099/jmm.0.037614-0.

[119] Dookie N, Sturm AW, Moodley P. Mechanisms of first-line antimicrobial resistance in multi-drug and extensively drug resistant strains of *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa. BMC Infect Dis 2016;16:609. doi:10.1186/s12879-016-1906-3.

[120] Tulyaprawat O, Chaiprasert A, Chongtrakool P, Suwannakarn K, Ngamskulrungroj P. Distribution of *embB* mutations of Thai clinical isolates of ethambutol-resistant *Mycobacterium tuberculosis*. J Glob Antimicrob Resist 2019;18:115–7. doi:10.1016/j.jgar.2019.05.033.

[121] Zhang D, Liu B, Wang Y, Pang Y. Rapid molecular screening for multidrug-resistant tuberculosis in a resource-limited region of China. Trop Med Int Heal 2014;19:1259–66. doi:10.1111/tmi.12359.